



**HAL**  
open science

# Calibrating a hydrological model in stage space to account for rating curve uncertainties: general framework and key challenges

A.E. Sikorska, Benjamin Renard

► **To cite this version:**

A.E. Sikorska, Benjamin Renard. Calibrating a hydrological model in stage space to account for rating curve uncertainties: general framework and key challenges. *Advances in Water Resources*, 2017, 105, pp.51-66. 10.1016/j.advwatres.2017.04.011 . hal-02606271

**HAL Id: hal-02606271**

**<https://hal.inrae.fr/hal-02606271>**

Submitted on 11 Jan 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



**University of  
Zurich**<sup>UZH</sup>

**Zurich Open Repository and  
Archive**

University of Zurich  
University Library  
Strickhofstrasse 39  
CH-8057 Zurich  
[www.zora.uzh.ch](http://www.zora.uzh.ch)

---

Year: 2017

---

## **Calibrating a hydrological model in stage space to account for rating curve uncertainties: general framework and key challenges**

Sikorska, Anna E ; Renard, Benjamin

**Abstract:** Hydrological models are typically calibrated with discharge time series derived from a rating curve, which is subject to parametric and structural uncertainties that are usually neglected. In this work, we develop a Bayesian approach to probabilistically represent parametric and structural rating curve errors in the calibration of hydrological models. To achieve this, we couple the hydrological model with the inverse rating curve yielding the rainfall–stage model that is calibrated in stage space. Acknowledging uncertainties of the hydrological and the rating curve models allows assessing their contribution to total uncertainties of stages and discharges. Our results from a case study in France indicate that (a) ignoring rating curve uncertainty leads to changes in hydrological parameters, and (b) structural uncertainty of hydrological model dominates other uncertainty sources. The paper ends with discussing key challenges that remain to be addressed to achieve a meaningful quantification of various uncertainty sources that affect hydrological model, as including input errors.

DOI: <https://doi.org/10.1016/j.advwatres.2017.04.011>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-137115>

Journal Article

Accepted Version

Originally published at:

Sikorska, Anna E; Renard, Benjamin (2017). Calibrating a hydrological model in stage space to account for rating curve uncertainties: general framework and key challenges. *Advances in Water Resources*, 105:51-66.

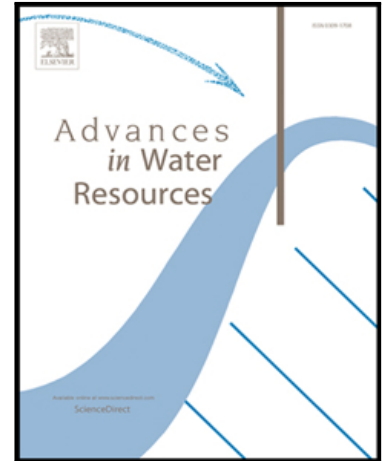
DOI: <https://doi.org/10.1016/j.advwatres.2017.04.011>

## Accepted Manuscript

Calibrating a hydrological model in stage space to account for rating curve uncertainties: general framework and key challenges

Anna E. Sikorska, Benjamin Renard

PII: S0309-1708(16)30391-8  
DOI: [10.1016/j.advwatres.2017.04.011](https://doi.org/10.1016/j.advwatres.2017.04.011)  
Reference: ADWR 2826



To appear in: *Advances in Water Resources*

Received date: 31 August 2016  
Revised date: 12 April 2017  
Accepted date: 13 April 2017

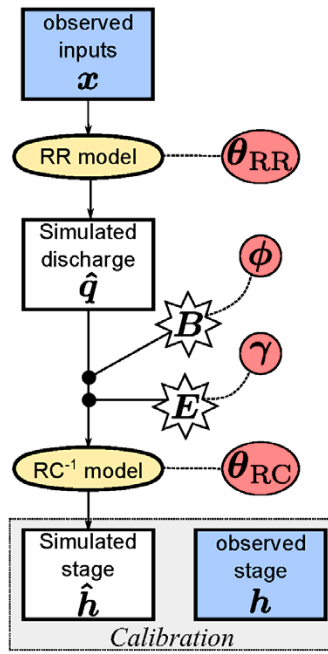
Please cite this article as: Anna E. Sikorska, Benjamin Renard, Calibrating a hydrological model in stage space to account for rating curve uncertainties: general framework and key challenges, *Advances in Water Resources* (2017), doi: [10.1016/j.advwatres.2017.04.011](https://doi.org/10.1016/j.advwatres.2017.04.011)

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

**Highlights**

- Method for quantifying rating curve uncertainties in discharge prediction is proposed
- A rainfall-stage model is developed and calibrated in stage space
- Such a rainfall-stage model couples a hydrological model with an inverse rating curve
- We consider both structural and parametric uncertainties of the rating curve
- Shares of these errors in the total uncertainty of stages and discharges are assessed
- Structural uncertainties of hydrological model dominates other uncertainty sources
- Ignoring rating curve errors affects the estimation of hydrological model parameters

## Graphical Abstract



# Calibrating a hydrological model in stage space to account for rating curve uncertainties: general framework and key challenges

Anna E. Sikorska<sup>a,b,c,\*</sup>, Benjamin Renard<sup>a</sup>

<sup>a</sup>*Irstea, UR HHLY, Hydrology-Hydraulics, 5 rue de la Doua CS70077, 69626  
Villeurbanne Cedex, France*

<sup>b</sup>*University of Zurich, Department of Geography, Winterthurerstr. 190, 8057 Zürich,  
Switzerland*

<sup>c</sup>*Warsaw University of Life Sciences – SGGW, Department of Hydraulic Engineering,  
Nowoursynowska 166, 02-787 Warsaw, Poland*

---

## Abstract

Hydrological models are typically calibrated with discharge time series derived from a rating curve, which is subject to parametric and structural uncertainties that are usually neglected. In this work, we develop a Bayesian approach to probabilistically represent parametric and structural rating curve errors in the calibration of hydrological models. To achieve this, we couple the hydrological model with the inverse rating curve yielding the rainfall-stage model that is calibrated in stage space. Acknowledging uncertainties of the hydrological and the rating curve models allows assessing their contribution to total uncertainties of stages and discharges. Our results from a case study in France indicate that a) ignoring rating curve uncertainty leads to changes in hydrological parameters, and b) structural uncertainty of hydrological model dominates other uncertainty sources. The paper ends with discussing key challenges that remain to be addressed to achieve a meaningful quantification of various uncertainty sources that affect hydrological model, as including input errors.

*Keywords:* rating curve, rainfall-stage model, structural uncertainty, parametric uncertainty, Bayesian inference, hydrological modeling

---

\*Corresponding author at: University of Zurich, Department of Geography, Winterthurerstr. 190, 8057 Zürich, Switzerland

*Email address:* [as@annasikorska.eu](mailto:as@annasikorska.eu) (Anna E. Sikorska)

---

## 1. Introduction

### 1.1. The importance of rating curve uncertainty in hydrological modeling

Flood risk analysis relies on estimates of hydrological models and associated uncertainty [1, 2, 3]. This uncertainty results mainly from four components: (i) parametric uncertainty of the hydrological model, (ii) its limited approximation of the catchment hydrological processes (model structural error), (iii) uncertainty in external model inputs (typically rainfall, temperature or evapotranspiration), and in (iv) output calibration data (typically discharge series) [4, 5, 6, 7, 3, 8, 9].

Among these four uncertainty contributors, input errors are considered to be one of the major uncertainty sources in hydrological models [10, 11, 12, 3] and thus more research has been devoted to investigate their effect on hydrological predictions than the effect of output uncertainty. Hence, different techniques have been proposed to represent input uncertainty which include a rainfall multiplier approach [10, 11, 13, 3], an addition to the bias [14, 8], or a more advanced stochastic description [15]. All of these studies, however, indicated that the inclusion of input errors raises several challenges. First, the computational cost is much higher than with traditional calibration. But even more importantly, substantial difficulties arise from the interaction between input errors and other uncertainty components. For instance, Renard et al. [6] discussed the challenge of identifying both input and structural errors; similarly, Del Giudice et al. [15] reported difficulties in distinguishing between different observational errors (input and output) if they have similar properties, i.e., are systematic. Hence, in this study we do not describe input errors explicitly, to be able to focus entirely on the effect of output uncertainty (due to the rating curve) on calibration and prediction of a hydrological model. Input errors will be implicitly encompassed in the structural error of the hydrological model.

As opposed to input errors, less attention has been given to the output uncertainty which is often assumed to be relatively small in comparison to the other three parts and thus has been evenly neglected in uncertainty analysis frameworks [16, 17]. Such a strong assumption might be justified for a direct measurement of discharge, for which measurement errors of 5% on average could be assumed [18, 19]. For practical applications, however, measuring

35 discharge continuously becomes impossible [20, 21]. Instead, a measure of dis-  
36 charge is obtained from an observed stage using a stage - discharge relation-  
37 ship (called rating curve) [22, 23]. This relationship needs to be established  
38 at a hydrometric station with few direct (discrete) measurements of gaug-  
39 ing pairs (stage and discharge) [18]. Using pre-established rating curves to  
40 compute discharges therefore allows deriving continuous quasi-observed dis-  
41 charge series [22, 23], which next serve for calibration of hydrological models  
42 [24, 25].

43 Awkwardly, these computed discharge series are often communicated to  
44 modellers or practitioners without any uncertainty statement [22, 26]. It is  
45 however clear that such estimated discharge series contain several errors. It  
46 has been reported in literature that although these errors are on average  
47 about 3 – 6% of an estimated value, they may increase to about 20% under  
48 poor measurement conditions [27], and to more than 25% outside the range  
49 of measured stage-discharge pairs [16, 28]. However, the level of these errors  
50 is case-specific [29] and results from many sources: measurement errors of  
51 gauging pairs (instrumental errors, measurement technique), temporal shifts  
52 in the rating curve (unstable stream channel due to vegetation, bank erosion,  
53 sediment deposition, ice jams, etc.), transient hydrological conditions during  
54 measurement of gauging pairs, hysteresis effect, and rating curve parametric  
55 and structural uncertainties [12, 30, 31, 26, 32].

56 All these errors affect calibration of the hydrological model and have  
57 serious implications for discharge simulations [12, 26, 23], flood frequency  
58 analysis [33, 34, 35], and for regionalization of model parameters [36]. As  
59 these errors are often not explicitly considered in uncertainty estimation,  
60 their effect on discharge uncertainty cannot be quantified. Moreover, when  
61 fully neglected, the uncertainty caused by rating curve errors may be wrongly  
62 attributed to other uncertainty source(s), leading to biased estimates that  
63 might be misunderstood by practitioners [35]. Given the above considerations  
64 and the number of studies dealing with calibration of hydrological models  
65 based on such quasi-observed discharge series, an accurate assessment of  
66 the rating curve uncertainties and their impact on the hydrological model  
67 becomes essential for flood risk assessment and management.

### 68 *1.2. Existing approaches to describe rating curve uncertainty*

69 Although a number of recent studies have investigated different aspects  
70 of rating curve uncertainties [20, 22, 37, 24, 38, 30, 39], the contribution  
71 of the rating curve to the uncertainty in hydrological simulations has not



72 been assessed systematically so far. In many uncertainty frameworks, rating  
73 curve errors are either not explicitly represented or are combined with other  
74 error sources. For instance, a common practice in uncertainty analysis is  
75 to pool all uncertainties (apart from parametric uncertainty but including  
76 rating curve uncertainty) into a lumped error term, which properties need  
77 to be mathematically described [40]. We call the latter approach when only  
78 parametric and structural errors of hydrological model are represented and  
79 the model is calibrated against discharges computed from rating curves as  
80 *a standard uncertainty estimation approach*. Another possible solution is  
81 mapping all uncertainty sources (including rating curve errors) to parameter  
82 uncertainty as in the original GLUE (generalized likelihood uncertainty esti-  
83 mation) methodology [41]. Further developments allowed to relate "limits of  
84 acceptability" with the rating curve uncertainty, although the need to extend  
85 these limits to account for other error sources (input errors in particular) was  
86 recognized [42]. Other approaches allow distinguishing input and structural  
87 errors [43, 11]. However, they don't explicitly represent rating curve errors,  
88 which are hence implicitly merged with structural errors. Finally, a recently  
89 introduced bias addition approach [14] gives the possibility to distinguish,  
90 aside from the parametric uncertainty, two different structural error types  
91 of the hydrological model, i.e., systematic and random errors. These er-  
92 rors are interpreted as structural and observational errors respectively. The  
93 bias approach pools however all observational errors (i.e., input and output)  
94 together and thus the uncertainty linked to the rating curve cannot be as-  
95 sessed. Hence, the major drawback of all these different approaches available  
96 to assess uncertainty of hydrological models is their inability to quantify the  
97 uncertainty contribution of the rating curve in total uncertainty estimates of  
98 hydrological models.

99 One possibility to indirectly tackle rating curve uncertainty is to prop-  
100 agate rating curve errors to discharge series which are then represented as  
101 spaghetti lines or uncertainty bands [44]. Such multiple realizations of dis-  
102 charge series yield however a practical question of how to calibrate a hydro-  
103 logical model with hundreds of "observed" discharges.

104 As an alternative, Sikorska et al. [26] and Thyer et al. [45] have recently  
105 proposed to avoid the issue of multiple "observed" discharges by simulating  
106 directly stages instead of discharges. Thus, they proposed to couple the  
107 hydrological model with the inverse rating curve yielding a so-called *rainfall-*  
108 *stage* model, for which uncertainty was evaluated in the stage space. In this  
109 way, rating curve uncertainty could be directly incorporated into simulations

110 of the hydrological model and the contribution of the rating curve uncertainty  
111 could be assessed. Yet, this method was mostly suitable to estimate stages  
112 while it was lacking the possibility to provide discharge predictions along with  
113 their uncertainty estimates (as discharge was only an intermediate step and  
114 was not directly modelled). Moreover, only the assessment of the parametric  
115 rating curve uncertainty was possible, while the structural errors of the rating  
116 curve could not be separated from those of the hydrological model.

117 Finally, other authors proposed specific error models to describe rating  
118 curve errors, based on an analysis of the rating curve itself [24]. Thyer et  
119 al. [37] and Renard et al. [12] proposed a specific error model within the  
120 Bayesian total error analysis methodology (BATEA) of Kavetski et al. [43,  
121 10], to represent structural errors of rating curves in discharge data along  
122 other uncertainty components (input and structural errors of hydrological  
123 model). In this way, contributions of those three main uncertainty compo-  
124 nents could be evaluated. Yet, they did not make an explicit distinction  
125 between parametric and structural uncertainties of rating curves, pooling all  
126 rating curve errors into a lumped structural error.

### 127 *1.3. Objectives*

128 Therefore, within this work, we further advance uncertainty quantifica-  
129 tion of rating curves by developing a Bayesian approach to probabilistically  
130 represent rating curve errors in the estimation of the hydrological model.  
131 In contrast to previous works, for the first time, we explicitly represent the  
132 parametric and the structural uncertainties of both the hydrological and the  
133 rating curve models. To achieve this, we couple the hydrological model with  
134 the inverse rating curve yielding the rainfall-stage model that can be cali-  
135 brated in stage space, as previously proposed by Sikorska et al. [26]. Specifi-  
136 cally, we describe structural errors of the hydrological model as an Ornstein-  
137 Uhlenbeck process [46] in the form implemented by Sikorska et al. [47], and  
138 the structural errors of the rating curve as Gaussian errors with a zero mean  
139 and a standard deviation proportional to the discharge value following the  
140 BaRatin method [23]. Because of such an explicit consideration of different  
141 uncertainty components of the rating curve and the hydrological model, the  
142 coupled total error can be decomposed into its constitutive sources. Hence,  
143 the approach is suitable for providing both stage and discharge simulations  
144 along with their associated uncertainties.

145 Specifically, we formulate the following objectives for this study:

- 146 1. Propose a generic framework for quantifying parametric and structural  
 147 uncertainties of rating curves in hydrological models, and derive the  
 148 corresponding inference equations;
- 149 2. Examine the effects of ignoring a specific source of rating curve uncer-  
 150 tainty (parametric or structural) in the inference of model parameters  
 151 and in model simulations;
- 152 3. Discuss pros and contras of using an advanced calibration approach  
 153 (representing both structural and parametric rating curve errors ex-  
 154 plicitly) over a “standard” uncertainty estimation approach (when un-  
 155 certainty is attributed only to parametric and structural errors of the  
 156 hydrological model and uncertainties of rating curve are neglected).

157 Our approach is developed and tested on a medium-size study catchment in  
 158 France. This study restricts its attention solely to investigate uncertainties  
 159 in output (discharge) of hydrological models, while uncertainty in input data  
 160 (typically rainfall), although non negligible, is not explicitly acknowledged  
 161 and is implicitly represented in structural errors of the hydrological model.  
 162 We debate possible consequences of this assumption in the discussion part.  
 163 Moreover, we recognize that an explicit and reliable treatment of all error  
 164 sources remains a key challenge for hydrologic modeling: while not the ob-  
 165 jective of this paper, we also discuss this long-term objective in section 5.5.

## 166 2. Uncertainty representation

### 167 2.1. Rating curve

#### 168 2.1.1. Rating curve model

169 We describe an instantaneous discharge at time  $t$  predicted with the rating  
 170 curve (RC),  $\check{q}_t$ , as

$$\check{q}_t = f_{\text{RC}}(h_t, \boldsymbol{\theta}_{\text{RC}}) \quad (1)$$

171 where  $f_{\text{RC}}(h_t, \boldsymbol{\theta}_{\text{RC}})$  is the deterministic RC equation,  $h_t$  is the instantaneous  
 172 stage at time  $t$  and  $\boldsymbol{\theta}_{\text{RC}} = (\theta_{\text{RC}_1}, \dots, \theta_{\text{RC}_w})$  are parameters of the RC. Because  
 173 parameters of the RC are unknown, they must be calibrated and thus they  
 174 will introduce parametric uncertainty to the RC (see Sect. 3.1 for description  
 175 of model calibration).

#### 176 2.1.2. Structural error

177 The rating curve equation is a simplified mathematical representation of  
 178 the true stage-discharge relationship prevailing at the gauging station. We

179 therefore introduce a structural error  $E_t$  to describe the difference between  
 180 the RC-predicted discharge  $\check{q}_t$  and the (unknown) true discharge  $q_t$ :

$$\check{q}_t = q_t + E_t(\gamma) \quad (2)$$

181 The structural error  $E_t$  is assumed to be a realization from a Gaussian distri-  
 182 bution with mean zero and standard deviation varying with the RC-predicted  
 183 discharge as parameterized below:

$$E_t \stackrel{indep}{\sim} N(0, g(\check{q}_t, \gamma)^2); \quad g(\check{q}_t, \gamma) = \gamma_1 + \gamma_2 \cdot \check{q}_t \quad (3)$$

184 where  $\gamma = (\gamma_1, \gamma_2)$  are the unknown parameters of the RC structural error  
 185 model. This equation calls for the following comments:

- 186 1. The assumption that the standard deviation of structural errors is an  
 187 affine function of the RC-predicted discharge is made to account for  
 188 heteroscedasticity, which is often observed in practice (see e.g. [22,  
 189 23]). A homoscedastic model can easily be obtained by fixing  $\gamma_2 = 0$ .  
 190 Conversely, more complex heteroscedasticity models can in principle  
 191 be derived by replacing the affine function  $g$  by another function (e.g.  
 192 an higher-order polynomial), at the cost of introducing more unknown  
 193 parameters;
- 194 2. Since the true discharge  $q_t$  is unknown, we assume that the standard  
 195 deviation of structural errors is a function of the RC-predicted discharge  
 196  $\check{q}_t$ ;
- 197 3. Eq. 3 also makes the strong assumption that structural errors are in-  
 198 dependent in time. This will be further discussed in section 5.2.

### 199 2.1.3. Gauging measurement error

200 The RC is typically calibrated using gaugings, i.e., pairs of stage-discharge  
 201 values measured at different stage levels and flow conditions [48, 49, 50].  
 202 The measurement error on stage is assumed to be negligible. Conversely, the  
 203 measurement error on the gauged discharge can be considerable. Hence, we  
 204 represent the gauged discharge observed at time  $t$ ,  $\tilde{q}_t$ , as the sum of the true  
 205 discharge  $q_t$  and a measurement error  $W_t$ :

$$\tilde{q}_t = q_t + W_t \quad (4)$$

206 The measurement error  $W_t$  is further assumed to be a realization from a  
 207 Gaussian distribution with mean zero and known standard deviation  $\delta_t$ :

$$W_t \stackrel{indep}{\sim} N(0, \delta_t^2), \quad (5)$$

208 This equation calls for the following comments:

- 209 1. We assume that  $\delta_t$  is known because the uncertainty of the gauged discharge can be quantified before RC estimation by analyzing the measurement process (see e.g. [51, 52, 23]). Note that each gauging has its  
210 specific uncertainty;  
211
- 212 2. As for structural errors, eq. 5 also makes the assumption that measurement errors are independent in time. However this assumption is  
213 probably much more realistic here.  
214  
215

## 216 2.2. Hydrological model

### 217 2.2.1. Rainfall-runoff model

218 For simplicity sake, we prefer to substitute the hydrological model with a  
219 rainfall-runoff model which abbreviates to RR since  $h$  notation is restricted  
220 for stage and thus could be confused with the abbreviation of a hydrological  
221 model. We represent a RR-predicted discharge at time  $t$ ,  $\hat{q}_t$ , as:

$$\hat{q}_t = f_{\text{RR}}(\mathbf{x}_{1:t}, \boldsymbol{\theta}_{\text{RR}}) \quad (6)$$

222 where  $f_{\text{RR}}(\mathbf{x}_{1:t}, \boldsymbol{\theta}_{\text{RR}})$  represent the deterministic RR equations,  $\mathbf{x}_{1:t}$  are in-  
223 puts time series up to time  $t$  and  $\boldsymbol{\theta}_{\text{RR}} = (\theta_{\text{RR}_1}, \dots, \theta_{\text{RR}_z})$  are the parameters.  
224 Note that for simplicity this notation makes initial conditions implicit. Sim-  
225 ilarly to the parameters of RC, parameters of the RR are unknown and they  
226 must be estimated from observations. Hence they will introduce parametric  
227 uncertainty to the RR model (see further Sect. 3.2 describing model calibra-  
228 tion).

### 229 2.2.2. Structural error

230 To account for the imperfect nature of the RR model, a structural error  $B_t$   
231 is introduced to describe the mismatch between the RR-predicted discharge  
232 and the (unknown) true discharge  $q_t$ :

$$\psi(\hat{q}_t) = \psi(q_t) + B_t(\boldsymbol{\phi}) \quad (7)$$

233 where  $\psi(\cdot)$  is a transformation function applied to the true and the RR-  
234 predicted discharges (typically, a Box-Cox transformation, see appendix sec-  
235 tion Appendix A). The aim of this transformation is to make the proba-  
236 bilistic model used to describe  $B_t$  (described next) more realistic.

237 In order to explicitly describe the autocorrelated nature of structural errors,

238  $B_t$  is represented as an Ornstein-Uhlenbeck (OU) process [46] with param-  
 239 ters  $\phi = (\phi_1, \phi_2)$ .

$$B_t \sim OU(\phi_1, \phi_2) \quad (8)$$

240 The OU process is a continuous-time equivalent of more standard time se-  
 241 ries models such as the autoregressive (AR) error model, which are only  
 242 defined for data sampled at regular discrete times. Such a continuous-time  
 243 model allows dealing with unequally spaced data, which are commonly used  
 244 for routine monitoring of instantaneous water stage or discharge (typically,  
 245 more frequent records during floods than during low flows). We choose the  
 246 correlation structure of  $B_t$  in such a way that it becomes similar to the AR(1)  
 247 model [14, 47] with the variance at time  $t_i$  conditioned on a previous time  
 248 step  $t_j$  being equal to:

$$\text{Var}(B_{t_i|j}) = \phi_1^2 \cdot \left( 1 - \exp\left(-\frac{2 \cdot |t_i - t_j|}{\phi_2}\right) \right) \quad (9)$$

249  $\phi_1$  can be interpreted as the asymptotic standard deviation (for infinitely-  
 250 spaced time points) and  $\phi_2$  is a characteristic correlation time.

### 251 2.3. Rainfall-stage model

252 The basic idea behind the construction of the rainfall-stage (RS) model is  
 253 to apply the inverse of the RC to the discharge simulated by the RR model  
 254 [26]. The advantage of such a RS model is that its parameters encompass  
 255 both the RR and the RC parameters, which allows explicitly accounting for  
 256 RC uncertainty in the calibration of the RR parameters. However, the struc-  
 257 tural errors affecting both the RR and the RC models propagate to the RS  
 258 model and therefore need to be accounted for, as described next.

#### 259 2.3.1. Structural error

260 Let  $h_t$  denote the true stage value at time  $t$ . From the RC model eqs. 1  
 261 and 2 we get:

$$f_{\text{RC}}(h_t, \boldsymbol{\theta}_{\text{RC}}) = q_t + E_t(\boldsymbol{\gamma}) \quad (10)$$

262 Inverting the RC therefore yields the following relation:

$$h_t = f_{\text{RC}}^{-1}(q_t + E_t(\boldsymbol{\gamma}), \boldsymbol{\theta}_{\text{RC}}) \quad (11)$$

263 Moreover from the RR structural error model eq. 7 we get:

$$q_t = \psi^{-1}(\psi(\hat{q}_t) - B_t(\phi)) \quad (12)$$

264 where  $\psi(\cdot)$  and  $\psi^{-1}(\cdot)$  are the forward and the backward transformation.

265 Combining eqs. 11 and 12, the true instantaneous stage at time  $t$  can be  
266 written as:

$$h_t = f_{\text{RC}}^{-1} \left( \psi^{-1} \left[ \underbrace{\psi \left( f_{\text{RR}}(\mathbf{x}_{1:t}, \boldsymbol{\theta}_{\text{RR}}) \right)}_{\text{RR model}} - \underbrace{B_t(\phi)}_{\text{RR structural error}} \right] + \underbrace{E_t(\gamma)}_{\text{RC structural error}}, \boldsymbol{\theta}_{\text{RC}} \right) \quad (13)$$

267 We stress that the structural error model described in eq. 13 is a pure con-  
268 sequence of the individual error models used for the RR and the RC models:  
269 no new assumption has been made to derive eq. 13.

### 270 2.3.2. Input/output measurement errors

271 The RS model needs to be calibrated using observations of its input/output  
272 variables. The input variables typically comprise precipitation and potential  
273 evapotranspiration, while the output variable is stage.

274 In this paper, we make the strong assumption that measurement errors  
275 in all input/output variables are negligible. We acknowledge that this as-  
276 sumption is unrealistic in most studies. For instance, errors in estimating  
277 areal precipitation may be large when the raingauge density is small (see e.g.  
278 [53, 12]). Similarly, continuously-measured stage values may be affected by  
279 non-negligible errors, of both random and systematic nature. Typically, the  
280 inherent uncertainty of the stage sensor corresponds to a random error, while  
281 the periodic recalibration of the stage sensor with respect to the staff gauge  
282 produces an unknown systematic error between two successive recalibrations  
283 (for more details, see e.g. [32]).

284 Making this restrictive assumption allows focusing entirely on the un-  
285 certainty induced by the rating curve while minimizing possible interactions  
286 between input and output errors. In practice, unaccounted input/output er-  
287 rors will be implicitly absorbed by the structural error terms ( $B_t$  and  $E_t$ ).  
288 One should therefore keep in mind that while these terms are intended to  
289 represent structural errors, they may also encompass the effect of ignored  
290 input/output errors.

### 291 3. Calibration

292 In this paper, we apply Bayesian estimation to estimate all unknown  
 293 parameters. The posterior distributions are explored by means of an adaptive  
 294 Markov Chain Monte Carlo sampler described in Haario et al. [54]. The  
 295 convergence of the chains is assessed visually by plotting the simulated chains  
 296 and verifying their stationarity.

297 The general calibration strategy is made of two successive steps. We  
 298 first estimate the RC using available gauging pairs (these gaugings are not  
 299 used afterwards). In a second stage, we estimate the RS model combining  
 300 the RC and the RR submodels (thus the RC model is re-calibrated). Since  
 301 the RS model comprises parameters related to the RC (namely,  $\boldsymbol{\theta}_{\text{RC}}$  and  
 302  $\boldsymbol{\gamma}$ , see section 2.1), the posterior distribution of these parameters obtained  
 303 after stage 1 becomes their prior distribution in stage 2. Note that this  
 304 informative prior for the RC model, based on an analysis of rating curve data,  
 305 strongly constrains the inference. This allows avoiding non-identifiability and  
 306 equifinality problems in the estimation of all parameters during stage 2.

#### 307 3.1. Stage 1: rating curve calibration

308 From the assumptions described in section 2.1, the gauged discharge at  
 309 time  $t$  can be written as follows (combining equations 2 and 4):

$$\tilde{q}_t = f_{\text{RC}}(\tilde{h}_t, \boldsymbol{\theta}_{\text{RC}}) - E_t(\boldsymbol{\gamma}) + W_t \quad (14)$$

310 Conditional on unknown parameters, the gauged discharge  $\tilde{q}_t$  is therefore a  
 311 realization from a Gaussian distribution with mean  $\check{q}_t = f_{\text{RC}}(\tilde{h}_t, \boldsymbol{\theta}_{\text{RC}})$  and  
 312 variance  $(\gamma_1 + \gamma_2 \cdot \check{q}_t)^2 + \delta_t^2$ . The likelihood function can therefore be written:

$$p(\tilde{\boldsymbol{q}} | \boldsymbol{\theta}_{\text{RC}}, \boldsymbol{\gamma}, \tilde{\boldsymbol{h}}) = \prod_{k=1}^{N_{\text{gauging}}} f_G(\tilde{q}_{t_k}; \check{q}_{t_k}, (\gamma_1 + \gamma_2 \cdot \check{q}_{t_k})^2 + \delta_{t_k}^2) \quad (15)$$

313 where  $f_G(u; m, v)$  is the Gaussian pdf with mean  $m$  and variance  $v$ , evalu-  
 314 ated at  $u$ .

315 The posterior distribution is then computed up to a constant of proportion-  
 316 ality using Bayes' theorem:

$$p(\boldsymbol{\theta}_{\text{RC}}, \boldsymbol{\gamma} | \tilde{\boldsymbol{q}}, \tilde{\boldsymbol{h}}) \propto p(\tilde{\boldsymbol{q}} | \boldsymbol{\theta}_{\text{RC}}, \boldsymbol{\gamma}, \tilde{\boldsymbol{h}}) \cdot p(\boldsymbol{\theta}_{\text{RC}}, \boldsymbol{\gamma}) \quad (16)$$



317 The prior distribution for RC parameters  $\boldsymbol{\theta}_{\text{RC}}$  is derived from an analysis of  
 318 the hydraulic configuration of the gauging station, as will be described in  
 319 the case study (for more general considerations, see Le Coz et al. [23]). For  
 320 the parameters  $\gamma$  governing the standard deviation of structural errors, wide  
 321 non-informative priors are used.

### 322 3.2. Stage 2: rainfall-stage model calibration

323 Let  $\mathbf{h} = (h_{t_k})_{k=1:N}$  denote the observed time series of stage values used  
 324 to calibrate the RS model. Computing the likelihood requires deriving the  
 325 distribution of  $\mathbf{h}$  conditional on all inferred quantities. Unfortunately, this  
 326 cannot be done directly on the basis of eq. 13. Indeed, this conditional  
 327 distribution is not Gaussian, because the Gaussian error terms  $E_t$  and  $B_t$   
 328 transit through nonlinear models (the backward transformation  $\psi^{-1}$  and the  
 329 inverse rating curve  $f_{\text{RC}}^{-1}$ ). Moreover, this non-Gaussian pdf cannot be derived  
 330 analytically. Indeed, eq. 13 involves the sum of two independent random  
 331 variables. The pdf of this sum can be obtained by convolution, but this  
 332 convolution has no analytical solution because one of the random variables  
 333 is not Gaussian.

334 In order to circumvent this issue, we partly linearize eq. 13 as described next.  
 335 We introduce the following shorthand notation for this section:

$$\begin{aligned} \hat{q}_t(\boldsymbol{\theta}_{\text{RR}}) &= f_{\text{RR}}(\mathbf{x}_{1:t}, \boldsymbol{\theta}_{\text{RR}}) \\ d_t^{(\psi)}(\boldsymbol{\theta}_{\text{RR}}) &= \psi'(\hat{q}_t) \end{aligned} \quad (17)$$

336 Using this notation and linearizing the backward transformation  $\psi^{-1}$ , eq. 13  
 337 can be approximated as follows (see Appendix B for details):

$$h_t \approx f_{\text{RC}}^{-1} \left( \underbrace{\hat{q}_t(\boldsymbol{\theta}_{\text{RR}}) - \frac{B_t(\phi)}{d_t^{(\psi)}(\boldsymbol{\theta}_{\text{RR}})} + E_t(\gamma)}_{Z_t}, \boldsymbol{\theta}_{\text{RC}} \right) \quad (18)$$

338 The term  $Z_t$  is now the sum of a constant plus two Gaussian terms, and  
 339 is therefore itself Gaussian. More precisely, the vector  $\mathbf{Z} = (Z_{t_1}, \dots, Z_{t_N})$   
 340 follows a multivariate Gaussian distribution, with mean vector  $\boldsymbol{\mu}$  (size  $N$ )  
 341 and covariance matrix  $\boldsymbol{\Sigma}$  (size  $N \times N$ ) defined as follows:

$$\boldsymbol{\mu}(\boldsymbol{\theta}_{\text{RR}}) = (\hat{q}_{t_1}(\boldsymbol{\theta}_{\text{RR}}), \dots, \hat{q}_{t_N}(\boldsymbol{\theta}_{\text{RR}})) \quad (19)$$

$$\Sigma(\boldsymbol{\theta}_{\text{RR}}, \boldsymbol{\phi}, \boldsymbol{\gamma}) = \mathbf{D}^{(\psi)} \Sigma^{(\text{RR})} \mathbf{D}^{(\psi)} + \Sigma^{(\text{RC})} \quad (20)$$

342 In the latter equation,  $\mathbf{D}^{(\psi)}$  denotes the square  $N \times N$  diagonal matrix whose  
 343 diagonal terms are equal to  $1/d_t^{(\psi)}$ , while  $\Sigma^{(\text{RR})}$  and  $\Sigma^{(\text{RC})}$  are the  $N \times N$   
 344 covariance matrices of RR and RC structural errors:

$$\mathbf{D}^{(\psi)}(i, i) = \frac{1}{d_{t_i}^{(\psi)}(\boldsymbol{\theta}_{\text{RR}})}; \quad \mathbf{D}^{(\psi)}(i, j) = 0 \text{ if } i \neq j \quad (21)$$

345

$$\Sigma^{(\text{RR})}(i, j) = \phi_1^2 \cdot \exp\left(-\frac{|t_i - t_j|}{\phi_2}\right) \quad (22)$$

346

$$\Sigma^{(\text{RC})}(i, i) = (\gamma_1 + \gamma_2 \cdot \hat{q}_{t_i})^2; \quad \Sigma^{(\text{RC})}(i, j) = 0 \text{ if } i \neq j \quad (23)$$

347 Having derived the pdf of  $\mathbf{Z}$ , the pdf of  $\mathbf{h} \approx f_{\text{RC}}^{-1}(\mathbf{Z})$  (eq. 18) can be ob-  
 348 tained by applying the change-of-variables formula. After some computation  
 349 (see Appendix B for details), this yields the following likelihood:

$$p(\mathbf{h} | \boldsymbol{\theta}_{\text{RR}}, \boldsymbol{\theta}_{\text{RC}}, \boldsymbol{\phi}, \boldsymbol{\gamma}, \mathbf{x}) = f_{MG}(f_{\text{RC}}(\mathbf{h}, \boldsymbol{\theta}_{\text{RC}}); \boldsymbol{\mu}(\boldsymbol{\theta}_{\text{RR}}), \Sigma(\boldsymbol{\theta}_{\text{RR}}, \boldsymbol{\phi}, \boldsymbol{\gamma})) \times \prod_{k=1}^N |f'_{\text{RC}}(h_{t_k}, \boldsymbol{\theta}_{\text{RC}})| \quad (24)$$

350 where  $f_{MG}(\mathbf{u}; \mathbf{m}, \mathbf{v})$  is the multivariate Gaussian pdf with mean vector  $\mathbf{m}$   
 351 (size  $N$ ) and covariance matrix  $\mathbf{v}$  (size  $N \times N$ ), evaluated at vector  $\mathbf{u}$  (size  
 352  $N$ ).

353 The posterior distribution is then computed up to a constant of proportion-  
 354 ality using Bayes' theorem:

$$p(\boldsymbol{\theta}_{\text{RR}}, \boldsymbol{\theta}_{\text{RC}}, \boldsymbol{\phi}, \boldsymbol{\gamma} | \mathbf{h}, \mathbf{x}) \propto p(\mathbf{h} | \boldsymbol{\theta}_{\text{RR}}, \boldsymbol{\theta}_{\text{RC}}, \boldsymbol{\phi}, \boldsymbol{\gamma}, \mathbf{x}) \cdot p(\boldsymbol{\theta}_{\text{RR}}, \boldsymbol{\theta}_{\text{RC}}, \boldsymbol{\phi}, \boldsymbol{\gamma}) \quad (25)$$

355 The prior distribution for RC-related parameters  $\boldsymbol{\theta}_{\text{RC}}$  and  $\boldsymbol{\gamma}$  is set to the  
 356 posterior distribution obtained after calibration of the RC using gaugings  
 357 at stage 1 (eq. 16). For the parameters of the RR model ( $\boldsymbol{\theta}_{\text{RR}}$ ), priors are  
 358 case-specific and related to the RR model and available information. For the  
 359 parameters  $\boldsymbol{\phi}$  governing the properties of RR structural errors, wide non-  
 360 informative priors are used.

361 Note that the RS model is calibrated against time series with observed  
 362 stages. However, during the evaluation both the output of the RS model,  
 363 stage, and the output of the RR model, discharge, will be examined. This is  
 364 possible thanks to the explicit treatment of RC and RR errors.

365 *3.3. Calibration strategies*

366 The posterior distribution in eq. 25 corresponds to a full calibration strat-  
 367 egy, schematized in Figure 1: the parameters related to both the RR model  
 368 and the RC are estimated together, thus enabling interactions between them  
 369 and hence assessing how RC uncertainties impact the estimation of RR pa-  
 370 rameters. In particular, both parametric ( $\theta_{RC}$ ) and structural ( $\mathbf{E}$ ) uncer-  
 371 tainties of the RC are accounted for. In order to understand in more depth  
 372 the impact of these two types of uncertainty, we also implement incomplete  
 373 calibration strategies, where some uncertainty sources are ignored. As shown  
 374 in Table 1, these strategies are the following:

- 375 1. Strategy NoS ignores RC structural uncertainty. This corresponds to  
 376 assuming that  $\mathbf{E} = 0$ , which is achieved by using  $\Sigma^{(RC)} = 0$  in eq. 20.  
 377 A similar representation of RC uncertainty has been used by Steinbakk  
 378 [55] in the context of flood frequency analysis, and by Sikorska et al. [26]  
 379 in the context of model calibration.
- 380 2. Strategy NoP ignores RC parametric uncertainty. This is achieved by  
 381 removing  $\theta_{RC}$  from the list of inferred parameters. The RC is therefore  
 382 used with a fixed parameter vector  $\hat{\theta}_{RC}$ , taken as the maxpost esti-  
 383 mate (i.e. the vector maximizing the stage-1 posterior of eq. 16). This  
 384 strategy is similar to the representation of RC uncertainty used by e.g.  
 385 Thyer et al. [37] or Renard et al. [6].
- 386 3. Strategy NoPNoS ignores both RC parametric and structural uncer-  
 387 tainty, hence using both a fixed parameter vector  $\hat{\theta}_{RC}$  and setting  
 388  $\Sigma^{(RC)} = 0$ . In this strategy, there is no explicit representation of RC  
 389 uncertainty, which corresponds to the most widely-used approach in  
 390 hydrological modeling (*standard uncertainty estimation approach*).
- 391 4. Strategy FULL\* is similar to the full strategy, except that the prior for  
 392 RC parameters  $\theta_{RC}$  is truncated. More precisely, we set the prior pdf  
 393 to zero outside of 95% probability intervals for each component of  $\theta_{RC}$ .  
 394 This strategy strongly limits the possible interactions between  $\theta_{RC}$  and  
 395 other inferred parameters. It guarantees that after calibration of the  
 396 RS model, the RC parameters will still be within the 95% credibility  
 397 intervals derived by calibrating the RC to gaugings. Note that bluntly  
 398 truncating the prior as done here makes the resulting distribution un-  
 399 normalized; however this is not problematic in the Bayesian-MCMC  
 400 context of this paper since the posterior only needs to be known up to  
 401 a normalizing constant.

402 5. Similarly, strategy NoS\* is a variation of the NoS strategy, with a  
403 truncated prior for  $\theta_{RC}$ .

#### 404 4. Case study: the Ardèche river at Meyras

##### 405 4.1. Ardèche catchment

406 The river Ardèche is a right tributary of the River Rhône and has its  
407 sources in the Massif Central in France (Figure 2). The gauging station  
408 Meyras, located at 318 m a.s.l., controls an area of 98.43 km<sup>2</sup>. The mean  
409 elevation of this catchment is 899 m a.s.l., with the highest point located at  
410 1467 m a.s.l. The catchment is quite steep with an average slope of 23.4 %  
411 and it is in 68% covered by forests [56]. The average annual precipitation,  
412 estimated based on fifty years of observations at the station Péreyres (840  
413 m a.s.l.), is 1774 mm/yr in this region, whereof approximately 40% is lost  
414 to evaporation. With the yearly mean daily temperature equal to 9.25°C  
415 and the snowfall ratio of less than 3% of the annual precipitation, the snow  
416 processes can be neglected to model this catchment.

##### 417 4.2. Rating curve

418 As a RC model (eq. 1), we use a piecewise combination of power functions  
419 of the form  $q = a(h - b)^c$ . This combination is defined by the succession of  
420 hydraulic controls governing the stage-discharge relationship, as explained in  
421 more details by Le Coz [23]. At the Meyras gauging station, three controls  
422 can be identified (Figure 3). Low flows are first governed by a natural gravel  
423 riffle (control 1). When the stage gets above a certain level, this riffle is  
424 drowned and a channel control takes over (control 2). Finally, for very high  
425 stage values, the main channel may be full and some flow may also occur in  
426 the floodplain (control 3). This configuration leads to the following rating  
427 curve equation:

$$f_{RC}(h_t, \theta_{RC}) = \begin{cases} a_1 (h_t - b_1)^{c_1} & \text{if } \kappa_1 < h_t \leq \kappa_2 \text{ (control 1)} \\ a_2 (h_t - b_2)^{c_2} & \text{if } \kappa_2 < h_t \leq \kappa_3 \text{ (control 2)} \\ a_2 (h_t - b_2)^{c_2} + a_3 (h_t - b_3)^{c_3} & \text{if } \kappa_3 < h_t \text{ (control 2 + 3)} \end{cases} \quad (26)$$

428 where  $(\kappa_1, \kappa_2, \kappa_3)$  are the (unknown) activation stages for each control. Note  
429 that parameters  $b_1$ ,  $b_2$  and  $b_3$  do not need to be inferred because they can be  
430 deduced by continuity of the RC as shown in eq. 27 below. Consequently,

431 the parameters of the rating curve are  $\boldsymbol{\theta}_{\text{RC}} = (\kappa_1, a_1, c_1, \kappa_2, a_2, c_2, \kappa_3, a_3, c_3)$   
 432 where the relationships between  $\kappa$  and  $b$  are as follows:

$$b_1 = \kappa_1; b_2 = \kappa_2 - \left( \frac{a_1}{a_2} \cdot (\kappa_2 - b_1)^{c_1} \right)^{\frac{1}{c_2}}; b_3 = \kappa_3 \quad (27)$$

433 The parameters  $\boldsymbol{\theta}_{\text{RC}}$  are related to physical characteristics of the gauging  
 434 section, which opens the possibility to specify informative priors. For in-  
 435 stance, the first control by a natural riffle can be approximated using a rect-  
 436 angular weir formula, as shown in Table 2. This formula indicates that the  
 437 exponent  $c_1$  should be close to 1.5. Moreover, the parameter  $a_1$  is linked  
 438 to the weir width  $B_w$  and to a discharge coefficient  $C_r$ . The width can be  
 439 approximated at <sup>1</sup>  $B_w = (8 \pm 2)$  m, while literature suggests values of the  
 440 coefficient  $C_r = 0.4 \pm 0.1$  (see [48, 23]). These two uncertainties can be com-  
 441 bined by using the uncertainty propagation formula recommended by the  
 442 Guide to the Expression of Uncertainty in Measurement [57]. This yields  
 443 the Gaussian prior distribution for  $a_1$  shown in Table 2. Lastly, the eleva-  
 444 tion of the weir crest, which defines the activation stage  $\kappa_1$ , is estimated at  
 445  $\kappa_1 = (-0.05 \pm 0.05)$  m.

446 A similar approach can be used to specify priors for parameters of controls  
 447 2 and 3, using the Manning-Strickler formula for wide rectangular channels  
 448 (see Table 2). For the main channel, the Strickler coefficient is set to  $K_S =$   
 449  $(25 \pm 2.5) \text{ m}^{1/3} \cdot \text{s}^{-1}$ , the channel width to  $B_w = (15 \pm 2.5)$  m and the slope to  
 450  $S = (3 \pm 1) \text{ m} \cdot \text{km}^{-1}$ . For the floodplain, we use  $K_S = (15 \pm 2.5) \text{ m}^{1/3} \cdot \text{s}^{-1}$ ,  
 451  $B_w = (30 \pm 5)$  m and  $S = (3 \pm 1.25) \text{ m} \cdot \text{km}^{-1}$ . This completes the prior  
 452 specification shown in Table 2.

#### 453 4.3. Rainfall-runoff model (HBV)

454 The rainfall-runoff process within the Ardèche catchment is modelled with  
 455 a HBV model [58, 59, 60]. The HBV consists of four main routines responsible  
 456 for modelling snow dynamics, soil moisture, runoff response, and flow routing  
 457 in the channel. Because snow processes can be neglected in this catchment,  
 458 we use a simplified version of the HBV model, i.e., with an inactive snow  
 459 component. To further simplify the model, we model the catchment as a  
 460 single subcatchment without any elevation-dependent correction factors for  
 461 inputs. This further reduces the number of inferred parameters to 6 (Table 3).

---

<sup>1</sup>in the notation  $x \pm s$ ,  $s$  is the standard deviation.

462 Such a simplified HBV model requires mean areal precipitation and long  
 463 term evaporation estimates as input, while temperature data responsible for  
 464 modeling the snow component are not strictly required. In this study, the  
 465 HBV model is run at hourly time steps. Since the HBV model was not  
 466 applied before in this catchment, no previous knowledge was available for its  
 467 parameters. Thus, we formulate prior for each HBV parameter as a uniform  
 468 distribution restricted to possible ranges that were defined for each parameter  
 469 independently (Table 3).

#### 470 4.4. Calibration data

##### 471 4.4.1. stage 1: rating curve calibration

472 To infer the RC parameters  $\theta_{RC}$  and  $\gamma$ , we use 41 gaugings made between  
 473 2001 and 2008, for a period with no noticeable shift of the RC. For each  
 474 gauged discharge, we assume a constant relative uncertainty of  $\pm 3.5\%$ , i.e.  
 475 for a gauged discharge equal to  $q_t$ , the standard deviation  $\delta_t$  in eq. 5 is set to  
 476  $\delta_t = 0.035 \cdot q_t$ . The gaugings and their uncertainty can be seen in Figure 4b.

##### 477 4.4.2. stage 2: rainfall-stage model calibration

478 The RS model described with the Eq. 13 requires mean areal precipita-  
 479 tion at the hourly time step as input. Yet, the stage observations at the  
 480 gauging station are recorded by the limnigraph with unequal time steps ad-  
 481 equate to the current dynamics of flow processes (i.e., between 1 hour and  
 482 10/15 days). Hence, we chose to use directly these data instead of convert-  
 483 ing them into the hourly estimates, which would yield additional errors due  
 484 to the stage approximation. Note that this involves interpolating the HBV-  
 485 discharge simulations on the temporal (irregular) grid used for stage values.  
 486 Using irregularly spaced data is possible with the correlated error term on  
 487 the hydrological model introduced (Eqs. 8 and 9).

#### 488 4.5. Results: rating curve calibration (stage 1)

489 Figure 4a shows the prior RC resulting from the hydraulic analysis of  
 490 the gauging station (Table 2). Figure 4b shows the posterior RC and illus-  
 491 trates the uncertainty reduction resulting from the information brought by  
 492 the gaugings. The posterior RC is overall quite precise, especially for stages  
 493 smaller than 1 m. For such relatively small stages, parametric uncertainty  
 494 is only a small part of the total uncertainty, which is hence dominated by  
 495 structural uncertainty. For stage values beyond 1 m, total uncertainty in-  
 496 creases, mostly due to an increase of parametric uncertainty which becomes

497 dominant for such high stages. In particular the parameter  $\kappa_3$  representing  
 498 the activation stage of the third control is not precisely estimated (between  
 499 1 m and 1.5 m, see green band in Figure 4b).

500 The posterior distribution of RC parameters  $\theta_{RC}$  and  $\gamma$  obtained after this  
 501 first stage is now being used as a prior distribution for the second stage. Note  
 502 that the posterior on RC is in fact represented with Monte Carlo samples.  
 503 Hence, to specify the prior distribution for the second stage of calibration,  
 504 we fit a multivariate Gaussian distribution to the Monte Carlo samples from  
 505 the first stage. The resulting corresponding marginal distributions can be  
 506 seen as gray boxplots in Figure 5.

#### 507 4.6. Results: parameter estimates (stage 2)

508 Posteriors for the RS model for all six calibration strategies are plotted  
 509 as boxplots against prior information (obtained from stage 1) in Figure 5.  
 510 For parameters of the RR and RC sub-models and of the structural error  
 511 of the RR model, we observe that parameters tend to form three groups  
 512 in terms of their posterior behaviours. These groups are shaped as follows:  
 513 (1) calibration strategies FULL and NoS, (2) NoP and NoPNoS, and (3)  
 514 FULL\* and NoS\*, as seen in the figure. It appears that this grouping is  
 515 driven by the way of accounting for RC parametric uncertainty, i.e.: fully  
 516 accounting (group 1), non-accounting (group 2), and accounting but within  
 517 the constrained truncated prior (group 3). The grouping effect is obviously  
 518 not visible for parameters responsible for the RC structural uncertainty ( $\gamma$ )  
 519 as these parameters are excluded from the inference in the strategies NoS,  
 520 NoPNoS and NoS\*.

##### 521 4.6.1. Hydrological model

522 With respect to the HBV parameters ( $\theta_{RR}$ , i.e., *PERC:MAXBAS*, top  
 523 two rows in Figure 5), which mainly control the response and routing func-  
 524 tion, we specifically observe that posterior parameters vary among three pat-  
 525 tern groups and particularly between FULL and NoPNoS strategy (*standard*  
 526 *uncertainty estimation approach*). We observe that using a simplified de-  
 527 scription of errors as in NoPNoS leads to different values of inferred model  
 528 parameters than when explicitly representing all major contributing sources.  
 529 Such modified parameters of the RR model should mostly transfer to altered  
 530 discharge simulations (being an intermediate step within the RS model) and  
 531 might lead to biased estimates. Confronting posterior ranges of different cal-  
 532 ibration strategies indicates that generally parametric uncertainty of the RR

533 model is very similar for most parameters in all strategies. The interpretation  
 534 of individual parameter uncertainty is however difficult due to their interactions.  
 535 Therefore, not the uncertainty of individual parameters but rather the  
 536 resulting parametric uncertainty in predicted discharge is our major interest  
 537 (see Sect. 4.8.2).

538 Diagnosis of the structural error model of the RR model ( $\phi$ ) show that the  
 539 error standard deviation ( $\phi_1$ ) is the smallest for the FULL and the largest for  
 540 NoPNoS strategy (Fig. 5). This seems logical as in the FULL strategy the to-  
 541 tal residual variance is decomposed into two contributing sources originating  
 542 from the RR model ( $\phi$ ) and from the RC model ( $\gamma$ , see further below), while  
 543 in the strategy NoPNoS all variance is explained with  $\phi$  only. Hence, only  
 544 this error can be increased to capture the mismatch between the observed  
 545 and the simulated stage. Posterior error standard deviations of all other cal-  
 546 ibration strategies lie between these two strategies. This seems reasonable as  
 547 they represent transitional steps between FULL and NoPNoS strategies in  
 548 terms of the level of the variance decomposition from the simplest strategy  
 549 (NoPNoS) towards the most complex strategy (FULL). As it also seems logi-  
 550 cal, excluding RC parameters ( $\theta_{RC}$ ) from the inference results in an increased  
 551 error  $\phi_1$  (NoP and NoPNoS) in comparison to strategies which include  $\theta_{RC}$   
 552 into the inference (NoS) even if restricted prior is used (FULL\* or NoS\*).  
 553 The error autocorrelation length ( $\phi_2$ ) generally follows the behavior of the  
 554 error standard deviation and is the longest for strategy NoP and NoPNoS,  
 555 and the shortest for NoS\* and FULL\*.

#### 556 4.6.2. Rating curve model

557 Posterior RC parameters ( $\theta_{RC}$ ) are presented in three bottom rows in  
 558 Figure 5 (parameters  $k_1$  till  $c_3$ ). Note that RC parameters in strategies NoP  
 559 and NoPNoS are not altered during the inference and are kept at the values  
 560 of maxpost from the calibration stage 1 from section 4.5. For other four  
 561 strategies, again two groups of parameter behaviors can be observed.

562 Specifically, using informative but unbounded prior for RC parameters  
 563  $\theta_{RC}$  during the inference (FULL and NoS) results in a significant shift of  
 564 posteriors often outside of the 95% prior bounds. This effect appears to be  
 565 a result of a possible compensation for other uncertainty sources, specifically  
 566 for the one originating from the RR model parameters. As it appears, nine  
 567 parameters of the RC in addition to six parameters of the RR model gives  
 568 a higher level of freedom for modifying model simulations to match stage  
 569 observations. It is worth recalling that although RC parameters are related



570 to physical characteristics of the gauging station and informative prior is  
 571 used, prior information on the third control is very imprecise as it is con-  
 572 strained with only very few gaugings (see section 4.5). Indeed, posteriors on  
 573 parameters of the third control are strongly modified during the inference.  
 574 As all RS parameters are inferred at the same time, the possible compensa-  
 575 tion between parameters of the RR and the RC sub-model cannot be avoided  
 576 given unbounded priors on all parameters being inferred. This shifting of the  
 577 RC parameters outside of hydraulically reasonable boundaries is expected to  
 578 have a consequence on the shape of an updated RC (see section 4.7).

579 Using a truncated prior on RC parameters indeed prevents from a strong  
 580 modification of RC parameters (FULL\* and NoS\*). As it is visible in the  
 581 Figure 5, posteriors attempt to move towards the values from unbounded  
 582 strategies but remain within the 95% limits set. This also results in a smaller  
 583 RC parametric uncertainty in strategies FULL\* and NoS\* than in unbounded  
 584 strategies FULL and NoS.

585 Finally, the structural error of the rating curve ( $\gamma$ ) varies in different cal-  
 586 ibration strategies. As it is represented with two parameters, the combined  
 587 structural error of the RC cannot be easily quantified from estimated poste-  
 588 riors. We observe, however, an inverse relationship between its behavior and  
 589 the behavior of the RR structural error. This relationships seems also logical  
 590 as both structural and parametric uncertainties of the RC are decomposed  
 591 from the total uncertainty in FULL and FULL\* strategies.

#### 592 4.7. Results: updated rating curve (stage 2)

593 Updated RCs for four strategies accounting for RC parametric uncertainty  
 594 (i.e., FULL, NoS, FULL\* and NoS\*) are plotted in Figure 6 with uncertainty  
 595 bands (blue polygons) against the prior (red polygons). RCs for strategies  
 596 NoP and NoPNoS are not plotted as their parameters are not altered during  
 597 the calibration. As expected, a strong shift in the RC posterior parameters  
 598 observed for strategies which use non-bounded prior (FULL and NoS) leads  
 599 to a strong modification of the RC shape. This effect is especially visible in  
 600 the range of the third control for which the updated RC distinctly transcends  
 601 the prior ranges (red polygons) by pushing the RC towards assigning smaller  
 602 discharge values for the same stages. As previously mentioned, the prior  
 603 for parameter inference on the third RC control is established with very few  
 604 measures and thus is very uncertain (see Figure 4), which allows for freely  
 605 modifying these parameters. It is clear that setting bounded priors on RC

606 parameters in strategies FULL\* and NoS\* prevents from destroying the RC  
607 shape which remains within the 95% prior limits.

608 This issue of using the RC parameters to compensate for limitations of  
609 the RR model has clearly serious implications for using such updated RCs  
610 and will be further discussed in section 5.1.

#### 611 4.8. Results: predictive uncertainty (stage 2)

##### 612 4.8.1. Total uncertainty bands

613 Total uncertainty bands (TUB) for the FULL strategy are plotted in  
614 Figure 7 for stages and in Figure 8 for discharges (top panels). For both  
615 variables TUB appear to be reasonable as they cover most of the data points  
616 and are smaller for low flow and higher for high flow conditions (assessed  
617 visually). The smaller uncertainty during low flows is more apparent for  
618 discharges than for stages.

619 Widths of TUB for all other strategies are plotted in Figure 9 for both  
620 stages (top) and discharges (bottom). The TUB width in the FULL strategy  
621 is used as a reference. Widths of TUB for all other strategies are represented  
622 with respect to the FULL TUB width and thus are plotted as curves (a  
623 value larger than one representing a TUB width larger than that of the  
624 FULL strategy). The top panel of Figure 9 shows that for stage, the TUB  
625 widths of all strategies are larger than that of the FULL strategy for almost  
626 the entire calibration and validation periods. Specifically, during low flow  
627 periods (e.g. around the vertical red line), TUB widths are larger than that  
628 of the FULL strategy by a factor of up to 2, while during high flows this  
629 factor decreases to about 1.25. Similar patterns are observed with respect to  
630 discharges apart for the strategy NoS\*, for which TUB width is similar to  
631 the FULL strategy on average.

632 Although the effect of obtaining the smallest TUB width for the FULL  
633 strategy is visible for both stage and discharge, it has greater implications for  
634 modeling discharge. Much smaller TUB for FULL strategy clearly demon-  
635 strates a benefit compared to the strategy NoPNoS (*standard uncertainty*  
636 *estimation approach*). This finding indicates that accounting for both (struc-  
637 tural and parametric) RC uncertainties allows for removing these uncertainty  
638 parts from the total discharge uncertainty and this results in narrower TUB  
639 in comparison to strategies which do not present such ability (NoPNoS, NoP,  
640 NoS). Using bounded prior on RC parameters results in wider TUB in com-  
641 parison to their respective unbounded strategies, which confirms that the

642 structural error of the RR model is used for compensation of other unrepre-  
643 sented uncertainty components.

#### 644 4.8.2. *Uncertainty contributors*

645 An explicit representation of different uncertainty components within the  
646 TUB (i.e. of rating curve and of hydrological model) allows for their relative  
647 assessment. Depending on the strategy, these are parametric and structural  
648 uncertainty of the RR model and/or the RC model. Clearly, the most inter-  
649 esting is the FULL strategy which makes it possible to assess all four uncer-  
650 tainty components in predictions of stages and two components in prediction  
651 of discharges. Note that by an explicit representation of the parametric and  
652 structural uncertainties of RC in the FULL strategy, these uncertainty com-  
653 ponents can be decomposed from the TUB and thus are not propagated on  
654 the discharge (since the aim is to predict the “true” discharge and not the  
655 RC-estimated one). On the contrary, not accounting for structural or para-  
656 metric uncertainty of the RC does not allow for removing these uncertainty  
657 parts from the total uncertainty and thus they will be implicitly propagated  
658 on the discharge simulations. Uncertainty contributions for the FULL strat-  
659 egy are presented visually in Figure 7 for stages and in Figure 8 for discharges  
660 (bottom panels).

661 With respect to stages, it can be seen that the structural error of the  
662 RR model ( $\phi$ ) represents the majority of the total uncertainty while the  
663 next major contributor is the structural error of the RC model. Parametric  
664 uncertainty of the RR model and of the RC model are both less relevant.  
665 These contributions vary slightly over time and the contribution of the RC  
666 structural error is slightly higher during recession periods, whereas the con-  
667 tribution of the RC parametric error increases during high flows. The con-  
668 tribution of the parametric uncertainty of the RR model is higher during high  
669 flows and successive recession periods, while it is smaller during low flows.  
670 In a similar fashion, the structural error of the RR model accounts for the  
671 majority of the total uncertainty of discharge prediction.

672 The visual assessment of uncertainty contributions is accompanied by  
673 the time-averaged relative contributions of each uncertainty source for all six  
674 strategies and these are presented in Table 4. Generally, we observe quite  
675 stable uncertainty contributions of different error sources in all six calibration  
676 schemes. For all calibration strategies, the structural error of the RR model  
677 explains the majority of the total uncertainty which is ranging from 81% in  
678 the FULL strategy to 94% in NoPNoS for stages, and from 92% in FULL to

679 94% for NoP, NoPNoS, FULL\* and NoS\* for discharges. Both the parametric  
680 uncertainties of the RR and of the RC model vary only insignificantly and  
681 are much less relevant than other two uncertainty components. Hence, the  
682 change in contribution shares is thus mainly due to structural uncertainties  
683 of both the RC but mostly the RR model. The latter component is thus used  
684 to compensate for all unrepresented uncertainty source(s).

## 685 5. Discussion

### 686 5.1. Feasibility of accounting for rating curve uncertainty through a rainfall- 687 stage model

688 The approach proposed in this paper to explicitly account for RC un-  
689 certainty in the calibration of a RR model is to include both RC and RR  
690 parameters within a rainfall-stage (RS) model. However, the results of the  
691 case study show that the initial RC (established using gaugings) is strongly  
692 modified after calibration of the RS model (strategies FULL and NoS), unless  
693 a restrictive truncated prior is used for RC parameters (strategies FULL\* and  
694 NoS\*). We consider that the extent to which the RC is modified is hardly  
695 defensible; we therefore do not consider this modification as a meaningful  
696 improvement of the RC, but rather as a sign that the results produced by  
697 strategies FULL and NoS should be taken with caution.

698 It is of interest to further discuss this issue in terms of the information  
699 content used in the two successive calibration stages. The initial RC is estab-  
700 lished using the information brought by 41 independent gaugings, along with  
701 the prior information derived from the hydraulic analysis (the latter being  
702 informative but still quite imprecise). The posterior distribution of calibra-  
703 tion stage 1 reflects this quantity of information. During calibration stage 2,  
704 this posterior is used as a prior, but the RC can be further modified by the  
705 information brought by more than 1000 stage values used for calibration of  
706 the RS model. At first sight, the information imbalance between 41 gaugings  
707 vs. 1000 stage values may explain why the RC is strongly modified by the  
708 calibration of the RS model (well beyond the prior constraint induced by the  
709 gaugings). However, one should keep in mind the following points:

- 710 1. Since the RR structural error uses an autocorrelation component, the  
711 information content of these 1000 stage values does not correspond to  
712 that of 1000 independent data;

713 2. The information contained in the 41 gaugings is only used to estimate  
714 the RC, while the information contained in the 1000 stage values is  
715 used to infer both the RC and the RR model.

716 These clarifications notwithstanding, the strong modification of the RC is a  
717 sign that the error models we used do not convincingly weight the information  
718 brought by the gaugings and the stage time series. We see at least two  
719 avenues to improve this:

- 720 1. Improve the error models, as discussed in the next section 5.2;
- 721 2. Do not re-estimate the RC during calibration stage 2. This can be  
722 achieved by means of a propagation approach, as discussed in section  
723 5.3

#### 724 5.2. Limitation of the error models

725 The RC structural error model in Eq. 3 assumes independent errors, which  
726 is questionable at least for time steps close to each other. In principle, it is  
727 feasible to avoid this independence assumption e.g. by using an autocorrela-  
728 tion component. However, identifying an autocorrelation structure based on  
729 gaugings is difficult in practice, if not impossible, because gaugings are made  
730 too sporadically. Typically two successive gaugings are separated by weeks  
731 or months, which makes shorter autocorrelation structures non-identifiable.  
732 While implementing dedicated high-frequency gauging strategies might be  
733 feasible, we do not see any obvious solution with existing operational gaug-  
734 ing datasets.

735 Unlike RC structural errors, RR structural errors are not assumed inde-  
736 pendent and instead an explicit autocorrelation component is used (Eq. 8).  
737 This autocorrelation structure is identifiable because the stage time series is  
738 sampled at a high frequency. However, due to the particular dynamics of the  
739 RR model, even this autocorrelation structure is too simplistic. In partic-  
740 ular, autocorrelation properties are likely very different during dry periods  
741 and rainy periods, when quick-flow components are activated. As input er-  
742 ror is implicitly encompassed into the model bias, these different properties  
743 cannot be distinguished with the RR error model used here and the inferred  
744 bias is “averaged” over dry and wet conditions. More flexibility should hence  
745 be added to the autocorrelation component to allow distinguishing these dis-  
746 tinct properties, for instance by making a bias dry/wet period dependent or  
747 input-related (for further discussion on input error see Sect. 5.5).

748 Moreover, a common limitation of both RC and RR structural error mod-  
 749 els is the lack of a systematic component. A structural error is indeed defined  
 750 as the difference between the model prediction (forced with perfect inputs)  
 751 and the unknown truth. For a given set of inputs, this error is likely to  
 752 have a non-zero mean, because it is (at least partly) due to model struc-  
 753 tural deficits that will systematically manifest themselves when the model is  
 754 forced with similar inputs. Such a non-zero mean can also be interpreted as  
 755 a “conditional bias” (conditional to the inputs and initial conditions). The  
 756 fact that the structural error models we used ignore this conditional bias (as  
 757 do the error models we are aware of in the literature) probably explains the  
 758 undesired modification of the RC discussed in section 5.1: the calibration can  
 759 only use parameters  $\theta_{RR}$  and  $\theta_{RC}$  (whose modification induces a systematic  
 760 difference in model prediction) to minimize this conditional bias. Deriving  
 761 an error model that explicitly describes the conditional bias is an important  
 762 perspective in our opinion, but also a challenging one: its formulation and  
 763 identifiability from the data in the absence of prior knowledge are open ques-  
 764 tions. Finally, we note that this discussion has some links with the problem  
 765 of describing epistemic errors with statistical models, which motivated the  
 766 development of “informal” likelihoods for hydrological [61] and rating curve  
 767 [39] models.

### 768 5.3. An alternative: propagating RC uncertainty

769 An alternative to the approach used in this paper is to propagate RC  
 770 uncertainty by performing many calibrations of the RS model, with each  
 771 calibration using a distinct RC. As an illustration, consider the box-plots  
 772 shown in Figure 10. They have been obtained by performing 5 calibrations  
 773 using the strategy NoP, where parameters  $\theta_{RC}$  are fixed to 5 distinct values  
 774 randomly chosen in the MCMC simulations of calibration stage 1. For the  
 775 sake of simplicity, we demonstrate this approach on the basis of only one RR  
 776 and RC parameter ( $\theta_{RR}$  and  $\theta_{RC}$  respectively). One given box-plot represents  
 777 the uncertainty in RR parameter  $\theta_{RR}$ , conditional on one particular RC.  
 778 Merging all 5 box-plots together allows “unconditioning”, i.e. representing  
 779 the total uncertainty in RR parameter  $\theta_{RR}$ , given all plausible RCs. In a  
 780 similar fashion, “unconditional” estimates are derived for all RR parameters  
 781  $\theta_{RR}$ . This propagation approach has been used for instance by Steinbakk [55]  
 782 or Petersen-Overleir [34] in a flood frequency analysis context.

783 Formally, the propagation approach leads to the following pdf represent-  
 784 ing uncertainty in RR parameters  $\theta_{RR}$ :

$$p_{propa}(\boldsymbol{\theta}_{RR}|\mathbf{h}) = \int p(\boldsymbol{\theta}_{RR}|\mathbf{h}, \boldsymbol{\theta}_{RC}) p(\boldsymbol{\theta}_{RC}) d\boldsymbol{\theta}_{RC} \quad (28)$$

785 By contrast, the approach used in this paper represents uncertainty in RR  
786 parameters  $\boldsymbol{\theta}_{RR}$  using its marginal posterior distribution, defined as follows:

$$\begin{aligned} p(\boldsymbol{\theta}_{RR}|\mathbf{h}) &= \int p(\boldsymbol{\theta}_{RR}, \boldsymbol{\theta}_{RC}|\mathbf{h}) d\boldsymbol{\theta}_{RC} \\ &= \int p(\boldsymbol{\theta}_{RR}|\mathbf{h}, \boldsymbol{\theta}_{RC}) p(\boldsymbol{\theta}_{RC}|\mathbf{h}) d\boldsymbol{\theta}_{RC} \end{aligned} \quad (29)$$

787 The difference between the two approaches appears clearly in these equations:  
788 the latter uses the information contained in the stage calibration data to  
789 update the inference of RC parameters (term  $p(\boldsymbol{\theta}_{RC}|\mathbf{h})$  in Eq. 29), while the  
790 former ignores this information and only uses the prior RC estimates (term  
791  $p(\boldsymbol{\theta}_{RC})$  in Eq. 28), i.e. the RC inferred with gaugings only.

792 Future work should investigate the pros and cons of each approach. The  
793 propagation approach is akin to repeating the NoP approach many times,  
794 except that the RC parameters are not fixed at their maxpost estimate, but  
795 are rather sampled from the prior distribution derived from the analysis of  
796 rating curve data. The advantage compared to NoP is that RC parametric  
797 uncertainty is not ignored. However, an obvious drawback of the propagation  
798 approach is its computational cost, since a potentially costly calibration has  
799 to be repeated many times.

#### 800 5.4. Limitation of the proposed approach in terms of time steps

801 The approach proposed in this paper uses the inverse of the RC to de-  
802 rive a RS model. This only makes sense if the RC is invertible, or in other  
803 words, if the stage-discharge relationship can be represented by a bijective  
804 function. This may not be the case under some particular circumstances  
805 (e.g. hydraulic hysteresis or variable backwater effects). But even more gen-  
806 erally, the stage-discharge relationship can only be represented by a bijective  
807 function at a nearly-instantaneous time step. Consider for instance a given  
808 daily-averaged discharge value: for this particular day, an infinity of stage  
809 time series could lead to the same daily discharge. Consequently, it is not  
810 possible to relate this daily discharge to a single daily stage indicator (e.g.  
811 daily mean/median/etc.).

812 Consequently, the approach proposed here is restricted to time steps for  
813 which the within-step variability of stage can be neglected. Whether and how  
814 the approach can be extended for larger time steps remains unclear yet. A  
815 possible strategy would be to define a RC between e.g. daily-averaged stage  
816 and discharge, equipped with a stochastic component in order to account  
817 for the non-uniqueness of the daily discharge associated with a given daily  
818 stage. The variability of this stochastic component would directly depend of  
819 the within-step variability of stage.

#### 820 *5.5. Towards a complete decomposition of input/output/structural errors*

821 Deriving a complete uncertainty framework that allows explicitly repre-  
822 senting all uncertainty sources remains a major challenge of hydrologic mod-  
823 eling. Several methodological frameworks have been proposed for this pur-  
824 pose, e.g., SODA [62], BATEA [63], Kalman and particle filters (e.g. [64, 65])  
825 and many more. These methodological frameworks need to be equipped with  
826 specific error models to describe the various sources of uncertainty (input,  
827 output and structural errors), and realistic error models are a prerequisite  
828 for a meaningful uncertainty analysis. Moreover, specifying precise and ac-  
829 curate prior distributions to characterize input and output errors is another  
830 prerequisite to limit the interactions between the various error sources (e.g.,  
831 [6, 15]). Consequently, studies focusing on a specific uncertainty component  
832 are valuable to derive realistic error models and investigate their properties.  
833 For instance, previous research was devoted to investigate properties of in-  
834 put errors (e.g., [13, 15]) and their impact on model calibration. Following  
835 the same line of thought, we focus in this paper on rating curve output  
836 errors, and their impact on model calibration. The specific error models  
837 we propose could later be included into a more general framework such as  
838 SODA, BATEA or a Kalman/particle filter. Finally, we stress that there is no  
839 unique solution to uncertainty estimation in hydrologic modelling. Instead,  
840 varied and flexible error models are necessary to adapt to the objective of  
841 the study, the available information, etc. As an illustration, we note that the  
842 output error model we propose in this paper requires a significant amount  
843 of information (hydraulic analysis of the gauging station, gaugings and their  
844 uncertainty). While this allows making valuable use of local information, it is  
845 primarily adapted to the detailed analysis of a small number of catchments.  
846 This information may not be available for larger-scale analyses that may in-  
847 volve hundreds or thousands of catchments. In this case, an alternative error



848 model would need to be considered (see, e.g., the nonparametric discharge  
849 uncertainty estimate of Vrugt et al. [62]).

## 850 6. Conclusions

851 In this work, we develop a Bayesian approach to probabilistically repre-  
852 sent parametric and structural uncertainties of the rating curve in the esti-  
853 mation of the hydrological model. To achieve this, we couple the hydrological  
854 model with the inverse rating curve yielding the *rainfall-stage model* that is  
855 calibrated in the stage and not in the discharge space. Such a model de-  
856 scription enables us for explicitly representing and quantifying uncertainties  
857 associated with both the hydrological and the rating curve model in the total  
858 uncertainty of stage and discharge predictions. For a case study in France,  
859 we consider six different calibration strategies with a different representation  
860 level of rating curve uncertainties (parametric and/or structural). Our re-  
861 sults show that a) ignoring rating curve uncertainty leads to visible changes  
862 in hydrological model parameters, and b) structural uncertainty of the hy-  
863 drological model dominates other uncertainty sources. The major limitation  
864 of the current method arises from a strong modification of the rating curve  
865 shape if rating curve parameters are re-estimated during the calibration of  
866 the rainfall-stage model and unbounded prior is used. We see this problem  
867 to be related to the shortcomings of the error models used to describe cor-  
868 related errors of the hydrological model and structural errors of the rating  
869 curve. Thus, the next step should be to test the method with a more ad-  
870 vanced description of errors and/or to explore the proposed alternative of  
871 propagating rating curve parametric uncertainty in more detail.

## 872 Acknowledgments

873 The support of the Ambassade de France en Suisse in the enforcement  
874 of this research is thankfully acknowledged. The authors thank the editor  
875 Harrie-Jan Hendricks-Franssen and three anonymous reviewers for their use-  
876 ful comments, which helped improving the manuscript.

## 877 References

- 878 [1] A. Montanari, What do we mean by uncertainty? the need for a consis-  
879 tent wording about uncertainty assessment in hydrology, *Hydrol. Pro-*  
880 *cess.* 21 (6) (2007) 841–845. doi:10.1002/hyp.6623.

- 881 [2] M.-H. Ramos, T. Mathevet, J. Thielen, F. Pappenberger, Communicat-  
882 ing uncertainty in hydro-meteorological forecasts: mission impossible?,  
883 Meteor. Appl. 17 (2) (2010) 223–235. doi:10.1002/met.202.
- 884 [3] A. E. Sikorska, A. Scheidegger, K. Banasik, J. Rieckermann, Bayesian  
885 uncertainty assessment of flood predictions in ungauged urban basins for  
886 conceptual rainfall-runoff models, Hydrol. Earth Syst. Sci. 16 (4) (2012)  
887 1221–1236. doi:10.5194/hess-16-1221-2012.
- 888 [4] N. K. Ajami, Q. Duan, S. Sorooshian, An integrated hydrologic bayesian  
889 multimodel combination framework: Confronting input, parameter, and  
890 model structural uncertainty in hydrologic prediction, Water Resour.  
891 Res. 43 (1) (2007) W01403. doi:10.1029/2005WR004745.
- 892 [5] G. Kuczera, B. Renard, M. Thyer, D. Kavetski, There are no hydrologi-  
893 cal monsters, just models and observations with large uncertainties!, Hy-  
894 drol. Sci. J. 55 (6) (2010) 980–991. doi:10.1080/02626667.2010.504677.
- 895 [6] B. Renard, D. Kavetski, G. Kuczera, M. Thyer, S. W. Franks, Under-  
896 standing predictive uncertainty in hydrologic modeling: The challenge  
897 of identifying input and structural errors, Water Resour. Res. 46 (5)  
898 (2010) W05521. doi:10.1029/2009WR008328.
- 899 [7] H. McMillan, T. Krueger, J. Freer, Benchmarking observational uncer-  
900 tainties for hydrology: rainfall, river discharge and water quality, Hydrol.  
901 Process. 26 (2012) 4078–4111. doi:10.1002/hyp.9384.
- 902 [8] A. E. Sikorska, J. Seibert, Value of different precipitation data for flood  
903 prediction in an alpine catchment: A bayesian approach, J. Hydrol. - (-)  
904 (2016) in press. doi:10.1016/j.jhydrol.2016.06.031.
- 905 [9] A. Montanari, D. Koutsoyiannis, A blueprint for process-based modeling  
906 of uncertain hydrological systems, Water Resour. Res. 48 (9) (2012)  
907 W09555. doi:10.1029/2011WR011412.
- 908 [10] D. Kavetski, G. Kuczera, S. W. Franks, Bayesian analysis of input un-  
909 certainty in hydrological modeling: 2. application, Water Resour. Res.  
910 42 (3) (2006) W03408. doi:10.1029/2005WR004376.

- 911 [11] J. Vrugt, C. ter Braak, M. Clark, J. Hyman, B. Robinson, Treatment  
912 of input uncertainty in hydrologic modeling: Doing hydrology back-  
913 ward with Markov chain Monte Carlo simulation, *Water Resour. Res.*  
914 44 (2008) W00B09. doi:http://dx.doi.org/10.1029/2007WR006720.
- 915 [12] B. Renard, D. Kavetski, E. Leblois, M. Thyer, G. Kuczera, S. W.  
916 Franks, Toward a reliable decomposition of predictive uncertainty  
917 in hydrological modeling: Characterizing rainfall errors using condi-  
918 tional simulation, *Water Resour. Res.* 47 (11) (2011) n/a–n/a, w11516.  
919 doi:10.1029/2011WR010643.
- 920 [13] H. McMillan, B. Jackson, M. Clark, D. Kavetski, R. Woods,  
921 Rainfall uncertainty in hydrological modelling: An evaluation of  
922 multiplicative error models, *J. Hydrol.* 400 (1-2) (2011) 83–94.  
923 doi:10.1016/j.jhydrol.2011.01.026.
- 924 [14] P. Reichert, N. Schuwirth, Linking statistical bias description to multi-  
925 objective model calibration, *Water Resour. Res.* 48 (9) (2012) W09543.  
926 doi:10.1029/2011WR011391.
- 927 [15] D. Del Giudice, C. Albert, J. Rieckermann, P. Reichert, Describing the  
928 catchment-averaged precipitation as a stochastic process improves pa-  
929 rameter and input estimation, *Water Resour. Res.* 52 (4) (2016) 3162 –  
930 3186. doi:10.1002/2015WR017871.
- 931 [16] G. Di Baldassarre, A. Montanari, Uncertainty in river discharge obser-  
932 vations: A quantitative analysis, *Hydrol. Earth Syst. Sci.* 13 (6) (2009)  
933 913–921.
- 934 [17] G. Di Baldassarre, P. Claps, A hydraulic study on the applicabil-  
935 ity of flood rating curves, *Hydrology Research* 42 (1) (2011) 10–19.  
936 doi:10.2166/nh.2010.098.
- 937 [18] WMO, Guide to hydrological practice, volume i, hydrology – from mea-  
938 surement to hydrological information, 6th edn, World Meteorological  
939 Organisation, Geneva, Switzerland, 2008, p. pp. 296.
- 940 [19] J. Le Coz, A literature review of methods for estimating the uncertainty  
941 associated with stage-discharge relations (2012).

- 942 [20] R. Clarke, Uncertainty in the estimation of mean annual flood due  
943 to rating-curve indefinision, *J. Hydrol.* 222 (1-4) (1999) 185–190.  
944 doi:10.1016/S0022-1694(99)00097-9.
- 945 [21] T. Morlot, C. Perret, A.-C. Favre, J. Jalbert, Dynamic rating curve as-  
946 sessment for hydrometric stations and computation of the associated un-  
947 certainties: Quality and station management indicators, *J. Hydrol.* 517  
948 (2014) 173–186. doi:http://dx.doi.org/10.1016/j.jhydrol.2014.05.007.
- 949 [22] A. Petersen-Øverleir, A. Soot, T. Reitan, Bayesian rating curve inference  
950 as a streamflow data quality assessment tool, *Water Resour. Manag.*  
951 23 (9) (2009) 1835–1842. doi:10.1007/s11269-008-9354-5.
- 952 [23] J. Le Coz, B. Renard, L. Bonnifait, B. F., R. Le Boursicaud, Combining  
953 hydraulic knowledge and uncertain gaugings in the estimation of hydro-  
954 metric rating curves: A bayesian approach, *J. Hydrol.* 509 (2014) 573 –  
955 587. doi:http://dx.doi.org/10.1016/j.jhydrol.2013.11.016.
- 956 [24] H. McMillan, J. Freer, F. Pappenberger, T. Krueger, M. Clark, Im-  
957 pacts of uncertain river flow data on rainfall-runoff model calibration  
958 and discharge predictions, *Hydrol. Process.* 24 (10) (2010) 1270–1284.  
959 doi:10.1002/hyp.7587.
- 960 [25] Q. Shao, J. Lerat, G. Podger, D. Dutta, Uncertainty estimation with  
961 bias-correction for flow series based on rating curve, *J. Hydrol.* 510  
962 (2014) 137 – 152. doi:http://dx.doi.org/10.1016/j.jhydrol.2013.12.025.
- 963 [26] A. E. Sikorska, A. Scheidegger, K. Banasik, J. Rieckermann, Considering  
964 rating curve uncertainty in water level predictions, *Hydrol. Earth Syst.*  
965 *Sci.* 17 (11) (2013) 4415–4427. doi:10.5194/hess-17-4415-2013.
- 966 [27] V. B. Sauer, R. W. Meyer, Determination of error in individual discharge  
967 measurements (1992).
- 968 [28] A. Domeneghetti, A. Castellarin, A. Brath, Assessing rating-curve un-  
969 certainty and its effects on hydraulic model calibration, *Hydrol. Earth*  
970 *Syst. Sci.* 16 (4) (2012) 1191–1202. doi:10.5194/hess-16-1191-2012.
- 971 [29] G. Coxon, J. Freer, I. K. Westerberg, T. Wagener, R. Woods, P. J.  
972 Smith, A novel framework for discharge uncertainty quantification ap-  
973 plied to 500 UK gauging stations, *Water Resour. Res.* (2015) 5531 –  
974 5546doi:10.1002/2014WR016532.

- 975 [30] G. Di Baldassarre, F. Laio, A. Montanari, Effect of observation errors  
976 on the uncertainty of design floods, *Phys. Chem. Earth, Parts A/B/C*  
977 42–44 (2012) 85 – 90.
- 978 [31] J. Guerrero, I. K. Westerberg, S. Halldin, C.-Y. Xu, L.-C. Lundin, Tem-  
979 poral variability in stage–discharge relationships, *J. Hydrol.* 446–447  
980 (2012) 90 – 102. doi:http://dx.doi.org/10.1016/j.jhydrol.2012.04.031.
- 981 [32] I. Horner, J. Le Coz, B. Renard, F. Branger, G. Pierrefeu, et al., Ac-  
982 counting for stage measurement errors in the uncertainty analysis of  
983 streamow records, *J. Hydrol* (2016) in prep.
- 984 [33] G. Kuczera, Correlated rating curve error in flood frequency inference,  
985 *Water Resour. Res.* 32 (7) (1996) 2119–2127. doi:10.1029/96WR00804.
- 986 [34] A. Petersen-Øverleir, T. Reitan, Accounting for rating curve imprecision  
987 in flood frequency analysis using likelihood-based methods, *J. Hydrol.*  
988 366 (1-4) (2009) 89–100, cited By 20. doi:10.1016/j.jhydrol.2008.12.014.
- 989 [35] M. Lang, K. Pobanz, B. Renard, E. Renouf, E. Sauquet, Extrap-  
990 olation of rating curves by hydraulic modelling, with application  
991 to flood frequency analysis, *Hydrol. Sci. J.* 55 (6) (2010) 883–898.  
992 doi:10.1080/02626667.2010.504186.
- 993 [36] I. K. Westerberg, T. Wagener, G. Coxon, H. K. McMillan, A. Castel-  
994 larin, A. Montanari, J. Freer, Uncertainty in hydrological signatures for  
995 gauged and ungauged catchments, *Water Resour. Res.* 52 (3).
- 996 [37] M. Thyer, B. Renard, D. Kavetski, G. Kuczera, S. W. Franks, S. Srikan-  
997 than, Critical evaluation of parameter consistency and predictive un-  
998 certainty in hydrological modeling: A case study using bayesian  
999 total error analysis, *Water Resour. Res.* 45 (12) (2009) W00B14.  
1000 doi:10.1029/2008WR006825.
- 1001 [38] I. Westerberg, J.-L. Guerrero, J. Seibert, K. Beven, S. Halldin, Stage-  
1002 discharge uncertainty derived with a non-stationary rating curve in  
1003 the choluteca river, honduras, *Hydrol. Process.* 25 (4) (2011) 603–613.  
1004 doi:10.1002/hyp.7848.

- 1005 [39] H. K. McMillan, I. K. Westerberg, Rating curve estimation un-  
1006 der epistemic uncertainty, *Hydrol. Process.* 29 (7) (2015) 1873–1882.  
1007 doi:10.1002/hyp.10419.
- 1008 [40] G. Schoups, J. A. Vrugt, A formal likelihood function for parameter and  
1009 predictive inference of hydrologic models with correlated, heteroscedas-  
1010 tic, and non-Gaussian errors, *Water Resour. Res.* 46 (2010) W10531.  
1011 doi:10.1029/2009WR008933.
- 1012 [41] K. Beven, A. Binley, The future of distributed models: Model cali-  
1013 bration and uncertainty prediction, *Hydrol. Process.* 6 (1992) 279–298.  
1014 doi:10.1002/hyp.3360060305.
- 1015 [42] Y. Liu, J. Freer, K. Beven, P. Matgen, Towards a limits of ac-  
1016 ceptability approach to the calibration of hydrological models: Ex-  
1017 tending observation error, *J. Hydrol.* 367 (1–2) (2009) 93 – 103.  
1018 doi:http://dx.doi.org/10.1016/j.jhydrol.2009.01.016.
- 1019 [43] D. Kavetski, G. Kuczera, S. W. Franks, Bayesian analysis of input un-  
1020 certainty in hydrological modeling: 1. Theory, *Water Resour. Res.* 42 (3)  
1021 (2006) W03407. doi:10.1029/2005WR004368.
- 1022 [44] H. Sellami, I. La Jeunesse, S. Benabdallah, M. Vanclooster, Parameter  
1023 and rating curve uncertainty propagation analysis of the swat model  
1024 for two small mediterranean catchments, *Hydrol. Sci. J.* 58 (8) (2013)  
1025 1635–1657. doi:10.1080/02626667.2013.837222.
- 1026 [45] M. Thyer, B. Renard, D. Kavetski, G. Kuczera, Can hydrological model  
1027 predictions be improved by developing streamflow measurement error  
1028 models using rating curve data?, *STAHY workshop - Advances in Sta-*  
1029 *tistical Hydrology*, 2010.
- 1030 [46] G. E. Uhlenbeck, L. S. Ornstein, On the theory of the Brownian Motion,  
1031 *Phys. Rev.* 36 (1930) 823.
- 1032 [47] A. E. Sikorska, D. Del Giudice, K. Banasik, J. Rieckermann,  
1033 The value of streamflow data in improving TSS predictions –  
1034 Bayesian multi-objective calibration, *J. Hydrol.* 530 (-) (2015) 241–254.  
1035 doi:10.1016/j.jhydrol.2015.09.051.

- 1036 [48] R. W. Herschy, *Hydrometry: principles and practice*, John Wiley & Sons  
1037 Ltd, 1998.
- 1038 [49] ISO, *Hydrometry – measurement of liquid flow in open channels using*  
1039 *current-meters or floats* (2007).
- 1040 [50] J. Le Coz, B. Camenen, X. Peyrard, G. Dramais, *Uncertainty in open-*  
1041 *channel discharges measured with the velocity–area method*, *Flow Meas.*  
1042 *Instrum.* 26 (0) (2012) 18–29.
- 1043 [51] P. M. Pelletier, *Uncertainties in the single determination of river dis-*  
1044 *charge: a literature review*, *Can. J. Civil Eng.* 15 (5) (1988) 834–850.
- 1045 [52] T. A. Cohn, J. E. Kiang, R. R. M. Jr., *Estimating discharge measure-*  
1046 *ment uncertainty using the interpolated variance estimator*, *J. Hydrol.*  
1047 *Eng.* 139 (5) (2013) 502–510.
- 1048 [53] R. Linsley, M. Kohler, *Hydrology for Engineers*, McGraw Hill, London,  
1049 1988.
- 1050 [54] H. Haario, E. Saksman, J. Tamminen, *An adaptive metropolis algo-*  
1051 *rithm*, *Bernoulli* 7 (2001) 223–242.
- 1052 [55] G. H. Steinbakk, T. L. Thorarinsdottir, T. Reitan, L. Schlichting,  
1053 S. Hølleland, K. Engeland, *Propagation of rating curve uncertainty*  
1054 *in design flood estimation.*, *Water Resour. Res.* (2016) in press-  
1055 doi:10.1002/2015WR018516.
- 1056 [56] M. Adamovic, I. Braud, F. Branger, J. W. Kirchner, *Assessing the sim-*  
1057 *ple dynamical systems approach in a mediterranean context: application*  
1058 *to the ardèche catchment (france)*, *Hydrol. Earth Syst. Sci.* 19 (5) (2015)  
1059 2427–2449. doi:10.5194/hess-19-2427-2015.
- 1060 [57] J. C. for Guides in Metrology, *Evaluation of measurement data – guide*  
1061 *to the expression of uncertainty in measurement* (2010).
- 1062 [58] S. Bergström, *The HBV model - its structure and applications*, no. RHO  
1063 4, SMHI Hydrology, Norröping, 1992, p. 35 pp.
- 1064 [59] J. Seibert, *Estimation of parameter uncertainty in the HBV model*,  
1065 *Nord. Hydrol.* 28 (4/5) (1997) 247–262.

- 1066 [60] J. Seibert, M. J. P. Vis, Teaching hydrological modeling with a user-  
1067 friendly catchment-runoff-model software package, *Hydrol. Earth Syst.*  
1068 *Sci.* 16 (9) (2012) 3315–3325. doi:10.5194/hess-16-3315-2012.
- 1069 [61] K. Beven, J. Freer, Equifinality, data assimilation, and uncertainty  
1070 estimation in mechanistic modelling of complex environmental sys-  
1071 tems using the GLUE methodology, *J. Hydrol.* 249 (2001) 11–29.  
1072 doi:10.1016/S0022-1694(01)00421-8.
- 1073 [62] J. Vrugt, C. Diks, H. Gupta, W. Bouten, J. Verstraten, Improved treat-  
1074 ment of uncertainty in hydrologic modeling: Combining the strengths of  
1075 global optimization and data assimilation, *Water Resour. Res.* 41 (2005)  
1076 W01017. doi:10.1029/2004WR003059.
- 1077 [63] D. Kavetski, S. W. Franks, G. Kuczera, Confronting input uncertainty  
1078 in environmental modelling, *Calibration of watershed models* (2003) 49  
1079 – 68doi:10.1029/WS006p0049.
- 1080 [64] A. Weerts, G. El Serafy, Particle filtering and ensemble Kalman filtering  
1081 for state updating with hydrological conceptual rainfall-runoff models,  
1082 *Water Resour. Res.* 42 (2006) W09403. doi:10.1029/2005WR004093.
- 1083 [65] P. Salamon, L. Feyen, Disentangling uncertainties in distributed  
1084 hydrological modeling using multiplicative error models and se-  
1085 quential data assimilation, *Water Resour. Res.* 46 (2010) W12501.  
1086 doi:10.1029/2009WR009022.
- 1087 [66] G. E. P. Box, D. R. Cox, An analysis of transformations revisited, re-  
1088 butted, *R. J. Am. Stat. Ass.* 77 (1982) 209–210.



Table 1: Calibration strategies.

Strategy	$\theta_{RC}$		$E_t$		$\pi(\theta_{RC})$	
	infer	fix	active	zero	full	truncated
FULL	✓		✓		✓	
NoS	✓			✓	✓	
NoP		✓	✓			N/A
NoPNoS		✓		✓		N/A
FULL*	✓		✓			✓
NoS*	✓			✓		✓

Table 2: Prior distributions for rating curve parameters. The Manning–Strickler equation is a simplified version valid for wide rectangular channels.

Control	Idealized formula	$\pi(\kappa)$	$\pi(a)$	$\pi(c)$
	Rectangular weir			
Control 1 natural riffle	$q = \underbrace{C_r B_w \sqrt{2g}}_a (h - \underbrace{h_0}_b) \underbrace{1.5}_c$	$N(-0.05, 0.05^2)$	$N(14, 5^2)$	$N(1.5, 0.025^2)$
	Manning–Strickler equation			
Control 2 main channel	$q = \underbrace{K_S B_w \sqrt{S}}_a (h - \underbrace{h_0}_b) \underbrace{5/3}_c$	$N(0.1, 0.05^2)$	$N(20, 5^2)$	$N(1.67, 0.025^2)$
Control 3 floodplain	Manning–Strickler equation	$N(1.2, 2^2)$	$N(25, 7.5^2)$	$N(1.67, 0.025^2)$

Table 3: HBV parameters being inferred during calibration and their prior.

Parameter	Significance [unit]	Prior min	Prior max
PERC	Percolation threshold parameter [ $mm h^{-1}$ ]	0	2
UZL	Groundwater runoff threshold parameter [ $mm$ ]	0	100
K0	Recession coefficient of the 1st storage [ $h^{-1}$ ]	0	0.4
K1	Recession coefficient of the 2nd storage [ $h^{-1}$ ]	0	0.2
K2	Recession coefficient of the 3rd storage [ $h^{-1}$ ]	0	0.1
MAXBAS	Length of the triangular weighing function [ $h$ ]	1	10

Table 4: Time-averaged relative contribution (in %) of each source of uncertainty.

Prediction of Calibration strategy	stage				discharge	
	$\theta_{RR}$	$B_t$	$\theta_{RC}$	$E_t$	$\theta_{RR}$	$B_t$
FULL	6	81	5	8	8	92
NoS	7	89	4	0	7	93
NoP	6	88	0	6	6	94
NoPNoS	6	94	0	0	6	94
FULL*	5	87	2	6	6	94
NoS*	6	92	2	0	6	94

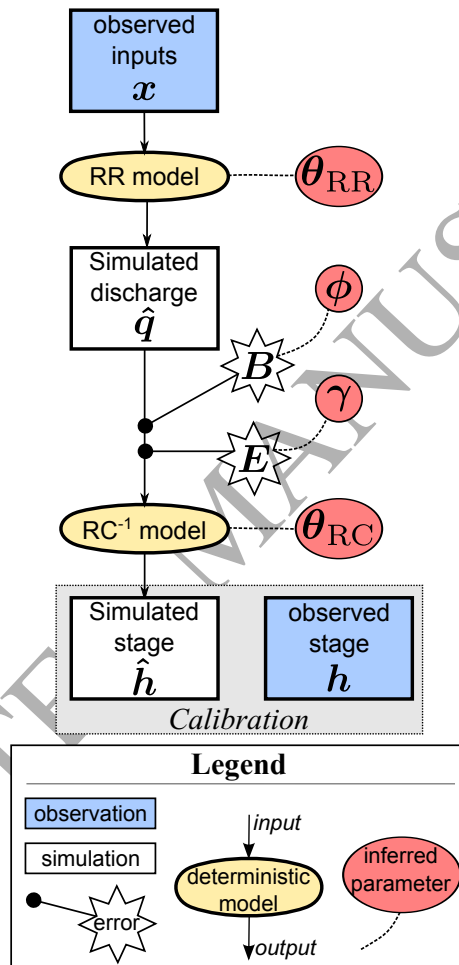


Figure 1: Schematic of the full calibration strategy.

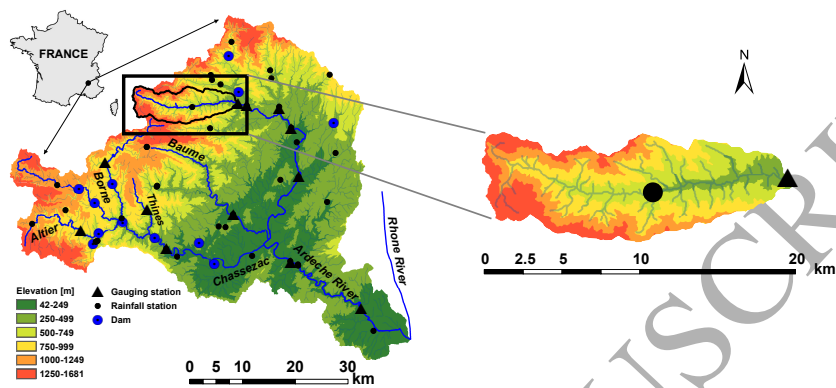


Figure 2: Overview of the catchment Ardèche. Left panel: entire catchment until its confluence with the Rhone River; right panel: Ardèche catchment at Meyras gauging station. Modified from Adamivic et al. [56].

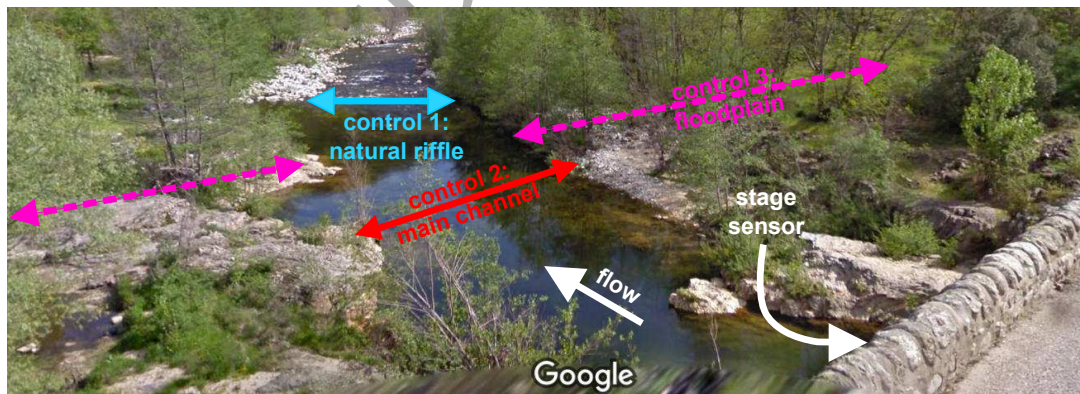


Figure 3: View of the gauging station for the Ardèche at Meyras. Picture from Google Maps, taken in April 2010.

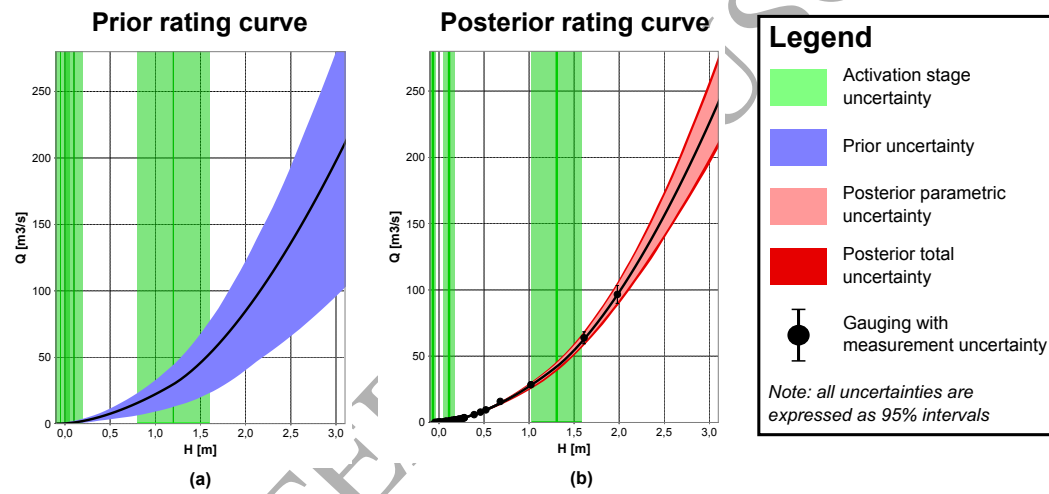


Figure 4: Prior (a) and posterior (b) rating curves in the 1st stage of calibration.

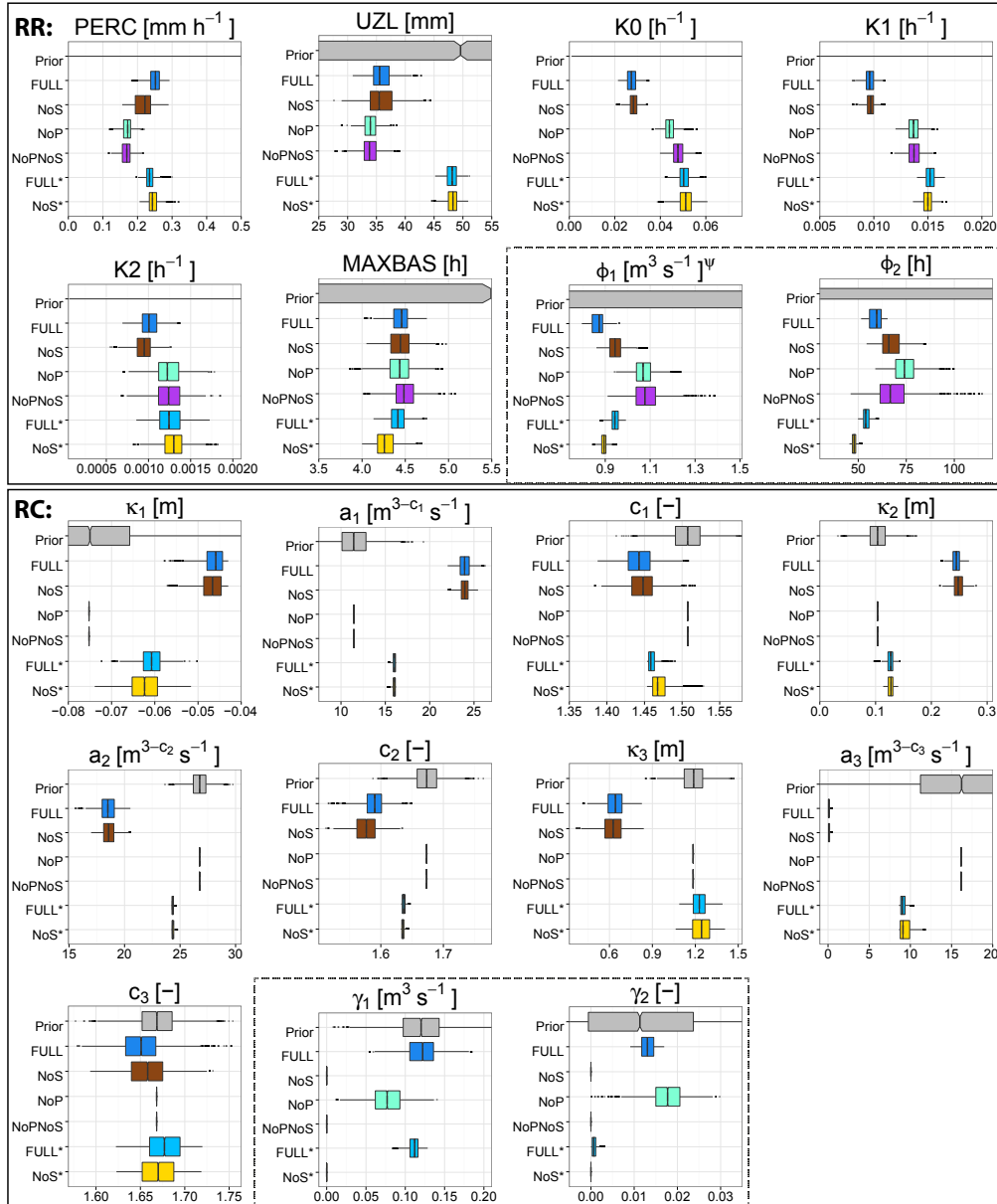


Figure 5: Calibration of the rainfall-stage model (stage 2): boxplots of prior (gray) and posterior (colored) distributions obtained with the six calibration strategies for the RR model (top panel) and RC model (bottom panel). Both error models (of RR and RC) are marked in the dashed boxes.

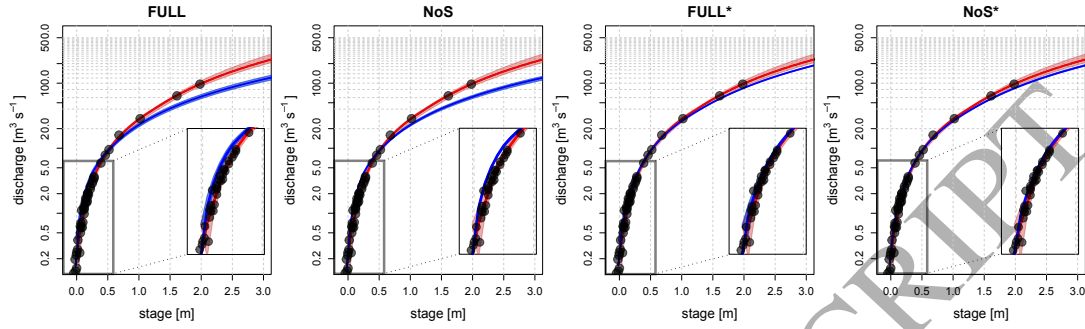


Figure 6: Comparison of rating curves before (red) and after (blue) calibration of the rainfall-stage model (stage 2) for the four calibration strategies accounting for RC parametric uncertainty.

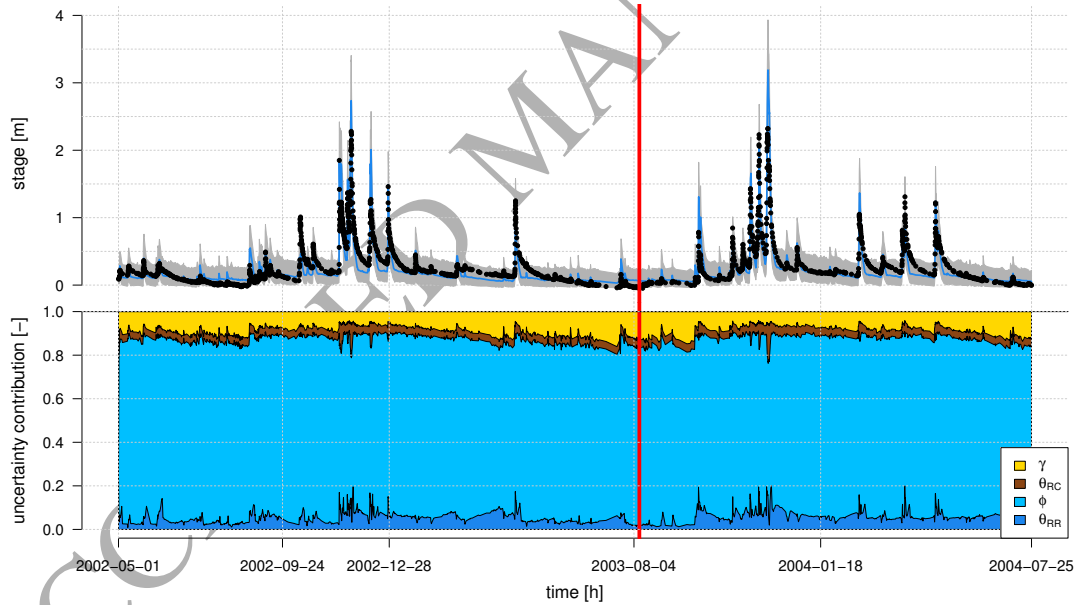


Figure 7: Stage prediction using the FULL calibration strategy. Top panel shows predicted vs. observed stage along with 95% intervals representing the total uncertainty. Bottom panel shows the relative contribution of each source of uncertainty. The calibration period is before the vertical line. Note the irregular time step of observed stages as demonstrated on the x-axis.

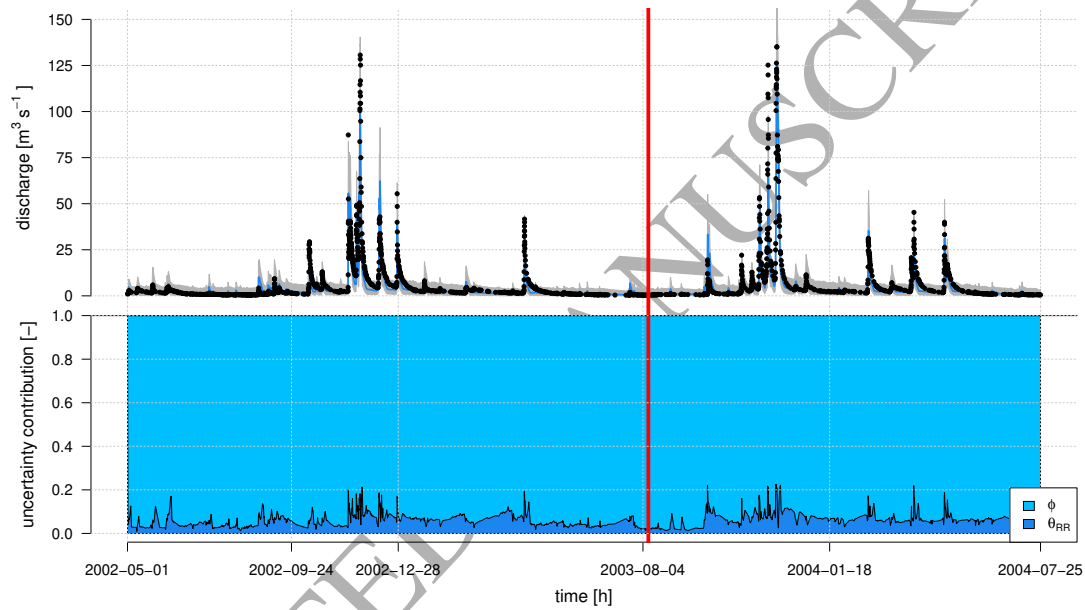


Figure 8: Discharge prediction using the FULL calibration strategy. Top panel shows predicted vs. observed discharge along with 95% intervals representing the total uncertainty. Bottom panel shows the relative contribution of each source of uncertainty. The calibration period is before the vertical line. Note that the irregular time step of discharges results from the irregular time step of observed stages.



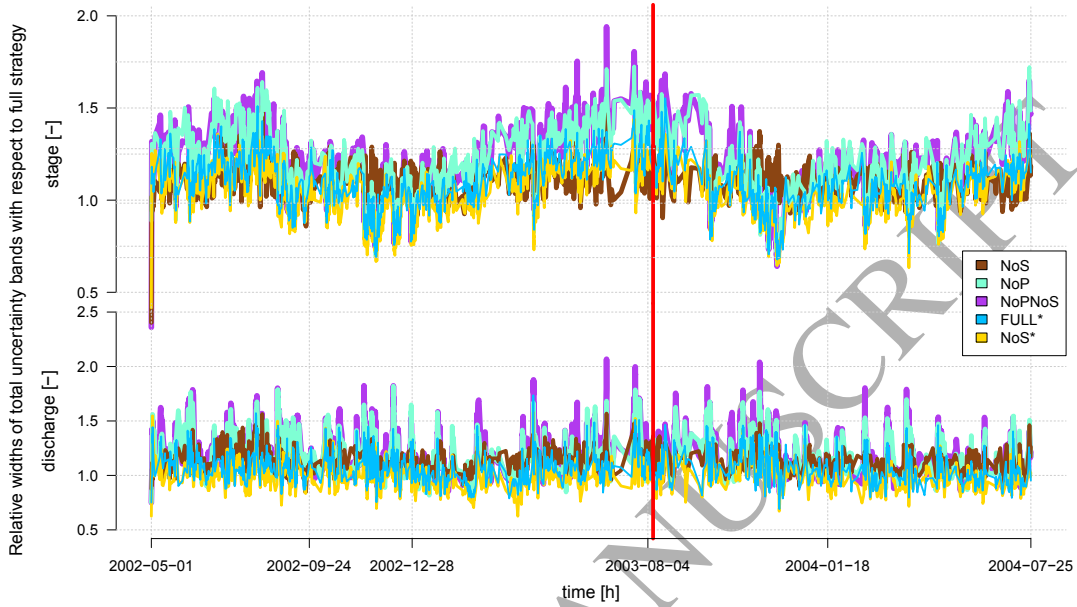


Figure 9: Comparison of total uncertainty for six calibration strategies. Each curve shows the ratio of 95% interval widths between the considered strategy and the reference strategy (FULL) in stage (top) and discharge (bottom) space. The calibration period is before the vertical line. Note the irregular time steps of both stages and discharges.

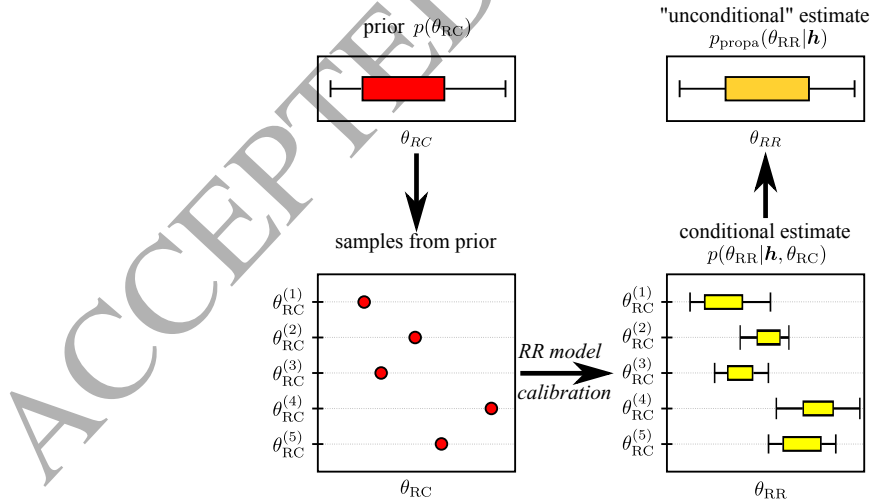


Figure 10: Schematic overview of the RC parametric uncertainty propagation approach.

1089 **Appendix A. Box-Cox transformation  $\psi(\cdot)$**

1090 The Box-Cox transformation [66] with parameters  $\lambda_1$  and  $\lambda_2$  can be writ-  
1091 ten as follows:

$$\psi(y) = \begin{cases} \frac{(y+\lambda_2)^{\lambda_1}-1}{\lambda_1} & \text{if } \lambda_1 \neq 0 \\ \ln(y+\lambda_2) & \text{if } \lambda_1 = 0 \end{cases} \quad (\text{A.1})$$

1092 Parameter  $\lambda_2 \geq 0$  aims at ensuring that the term  $y + \lambda_2$  remains positive.  
1093 Note that for  $(\lambda_1 = 1, \lambda_2 = 1)$ , the Box-Cox transformation is the identity,  
1094 while for  $(\lambda_1 = 0, \lambda_2 = 0)$  it simplifies to a logarithmic transformation.  
1095 Typically parameter  $\lambda_1$  is taken between 0 and 1.

1096 The inverse of the Box-Cox transform and its derivative can be written as  
1097 follows:

$$\psi^{-1}(\hat{y}) = \begin{cases} (\lambda_1 \cdot \hat{y} + 1)^{1/\lambda_1} - \lambda_2 & \text{if } \lambda_1 \neq 0 \\ \exp(\hat{y}) - \lambda_2 & \text{if } \lambda_1 = 0 \end{cases} \quad (\text{A.2})$$

$$\psi'(y) = (y + \lambda_2)^{\lambda_1 - 1} \quad (\text{A.3})$$

1098 **Appendix B. Likelihood computation for the RS model**

1099 The task is to derive the joint pdf of  $(h_{t_1}, \dots, h_{t_N})$ , where  $h_t$  is given by  
1100 eq. 13 (recalled below in a simplified form):

$$h_t = f_{\text{RC}}^{-1}(\psi^{-1}[\psi(\hat{q}_t) - B_t] + E_t) \quad (\text{B.1})$$

1101 The first step is to use a first-order approximation of the backward transform  
1102  $\psi^{-1}$  based on a first-order Taylor expansion, whose general form can be  
1103 written as:

$$f(x + e) \approx f(x) + f'(x) \times e \quad (\text{B.2})$$

1104 Applied to the function  $\psi^{-1}$  in eq. B.1, this yields:

$$\psi^{-1}[\psi(\hat{q}_t) - B_t] \approx \psi^{-1}[\psi(\hat{q}_t)] - (\psi^{-1})'[\psi(\hat{q}_t)] \times B_t \quad (\text{B.3})$$

1105 We then use here the inverse-derivative rule:

$$(\psi^{-1})'(z) = \frac{1}{\psi'(\psi^{-1}(z))} \quad (\text{B.4})$$

1106 Plugging this back into eq. B.3 yields:

$$\psi^{-1} [\psi(\hat{q}_t) - B_t] \approx \hat{q}_t - \frac{B_t}{\psi'(\psi^{-1}[\psi(\hat{q}_t)])} = \hat{q}_t - \frac{B_t}{\psi'(\hat{q}_t)} \quad (\text{B.5})$$

1107 Finally, eq. B.1 becomes:

$$h_t \approx f_{\text{RC}}^{-1} \left( \underbrace{\hat{q}_t - \frac{B_t}{\psi'(\hat{q}_t)}}_{Z_t} + E_t \right) \quad (\text{B.6})$$

1108 The second step is to deduce the joint pdf of  $(h_{t_1}, \dots, h_{t_N})$  from that of  
 1109  $(Z_{t_1}, \dots, Z_{t_N})$ . We use the change-of-variables formula for this purpose, which  
 1110 can be written in general terms as follows. Let  $\mathbf{y} = (y_1, \dots, y_N) = \mathbf{r}(x_1, \dots, x_N)$ ,  
 1111 where  $\mathbf{r}$  is a one-to-one transformation. The pdf of  $\mathbf{y}$  can be deduced from  
 1112 the pdf of  $\mathbf{x}$  using the following formula:

$$p_{\mathbf{y}}(y_1, \dots, y_N) = p_{\mathbf{x}}(\mathbf{r}^{-1}(\mathbf{y})) |\det(J_{\mathbf{r}^{-1}}(\mathbf{y}))| \quad (\text{B.7})$$

1113 where  $J_{\mathbf{r}^{-1}}(\mathbf{y})$  is the  $N \times N$  Jacobian matrix (partial derivatives) of the in-  
 1114 verse transform  $\mathbf{r}^{-1}$ .

1115 Applying the change-of-variables formula above to the transformation  $(h_{t_1}, \dots, h_{t_N}) =$   
 1116  $(f_{\text{RC}}^{-1}(Z_{t_1}), \dots, f_{\text{RC}}^{-1}(Z_{t_N}))$  yields the following formula:

$$\begin{aligned} p_{\mathbf{h}}(h_{t_1}, \dots, h_{t_N}) &= p_{\mathbf{Z}}(f_{\text{RC}}(h_{t_1}), \dots, f_{\text{RC}}(h_{t_N})) \left| \det \begin{pmatrix} f'_{\text{RC}}(h_{t_1}) & & 0 \\ & \dots & \\ 0 & & f'_{\text{RC}}(h_{t_N}) \end{pmatrix} \right| \\ &= p_{\mathbf{Z}}(f_{\text{RC}}(h_{t_1}), \dots, f_{\text{RC}}(h_{t_N})) \prod_{k=1}^N |f'_{\text{RC}}(h_{t_k})| \end{aligned} \quad (\text{B.8})$$

1117 which corresponds to the likelihood function from section 3.2 (eq. 24).