Combining hydraulic knowledge and uncertain gaugings in the estimation of hydrometric rating curves: a Bayesian approach

J. Le Coz, Benjamin Renard, L. Bonnifait, F. Branger, R. Le Boursicaud

Combining hydraulic knowledge and uncertain gaugings in the estimation of hydrometric rating curves: A Bayesian approach

J. Le Coz[*,a], B. Renard[a], L. Bonnifait[a,b], F. Branger[a], R. Le Boursicaud[a]

[a]*Irstea, UR HHLY, Hydrology-Hydraulics*
5 rue de la Doua CS70077 69626 Villeurbanne cedex, France
[b]*CETE Ile-de-France, 12 rue Teisserenc de Bort, 78197 Trappes cedex, France*

Abstract

Discharge time series in rivers and streams are usually based on simple stage-discharge relations calibrated using a set of direct stage-discharge measurements called gaugings. Bayesian inference recently emerged as a most promising framework to build such hydrometric rating curves accurately and to estimate the associated uncertainty. In addition to providing the rigorous statistical framework necessary to uncertainty analysis, the main advantage of the Bayesian analysis of rating curves arises from the quantitative assessment of (i) the hydraulic controls that govern the stage-discharge relation, and of (ii) the individual uncertainties of available gaugings, which often differ according to the discharge measurement procedure and the flow conditions. In this paper, we introduce the BaRatin method for the Bayesian analysis of stationary rating curves and we apply it to three typical cases of hydrometric stations with contrasted flow conditions and variable abundance of hydraulic knowledge and gauging data. The results exemplify that the thorough analysis of hydraulic controls and the quantification of gauging uncertainties are required to obtain reliable and physically sound results.

Key words: stage-discharge relations, rating curves, hydrometry, uncertainty analysis, Bayesian inference, hydraulic controls

1. Introduction

1.1. Physical basis of stage-discharge relations

Most often, the discharges of water streams are monitored by converting continuous water level records using a stage-discharge relation [e.g., 41, 49, 56, 19]. Such a hydrometric rating curve is usually calibrated using a set of direct stage-discharge measurements, which are called gaugings. McMillan et al. [29] provide a comprehensive review of the uncertainty values for gaugings and rating curves that were reported in the literature. Due to technical constraints, the gaugings are often scarce, especially at extremely high or low discharge, and may be affected by large and variable uncertainty, typically 5% to 20% of the measured discharge. Note that in this document, by default uncertainty is expressed at 95% confidence level, which corresponds to the convention most often used in hydrometry, as recommended by the Hydrometry Uncertainty Guide (HUG, 16). Stage-discharge relations often have to be extrapolated beyond the range of available gaugings, which may produce systematic errors as high as 100% or even more, resulting in wide credibility intervals associated with flood quantile estimates [23] and obviously also with drought discharge values. According to the expertise of the hydrometer and to available information, establishing and updating rating curves involve analyzing the hydraulic conditions at the study site. Managing rating curves and assessing their uncertainty hence remain difficult tasks which are not fully standardized yet.

---

[*]Corresponding author. Tel: +33 472208786; fax: +33 478477875.
Email addresses: jerome.lecoz@irstea.fr (J. Le Coz), benjamin.renard@irstea.fr (B. Renard), laurent.bonnifait@developpement-durable.gouv.fr (L. Bonnifait), flora.branger@irstea.fr (F. Branger), raphael.le-boursicaud@irstea.fr (R. Le Boursicaud)

A simple rating curve is a monotonic function relating the discharge, Q, to the water level, h, which is assumed to prevail at a cross-section of the flow in the reference hydraulic conditions. This reference hydraulic regime is seldom explicitly defined. Most often, the reference regime refers to the hydraulic conditions which usually prevail in the considered flow [49], i.e. steady flow (negligible transient effects) and usual hydraulic controls (e.g. no variable backwater effects, no change in channel roughness or in the geometry of the cross-section). Nevertheless, any time the flow deviates from the reference regime, significant errors in the discharge estimate may appear. Such errors must be distinguished from the errors directly related to the reference stage-discharge relation.

When the reference regime is permanently changed, e.g., in case of changed channel geometry after a flood, the rating curve is no longer valid, and a new one must be established corresponding to the new reference regime. Temporary changes of the reference regime may occur, due for instance to seasonal vegetation growth [56], variation of the downstream boundary condition [39], hysteresis due to transient flow effects [26], or dune-flat bed transitions during floods [50]. Non-stationarity in the stage-discharge relation may impose the use of different rating curves according to time periods, or even of rating curves with time-varying parameters.

The physical characteristics of the channel which govern the relation between stage h and discharge Q at a section constitute the hydraulic control. Basically, two kinds of hydraulic controls may be distinguished: section versus channel controls [55, 19]. When section control holds, the flow is mainly regulated by the geometry of a cross-section or a hydraulic work where the flow becomes critical due to a water fall (e.g., riffle, weir, sill) or due to a constriction (e.g., Venturi, Parshall flumes). When channel control holds, usually for medium to high flows, the flow is mainly regulated by the geometry and roughness of a portion of the channel. In non-uniform flow cases, the downstream boundary condition may also influence the stage-discharge relation (backwater effects).
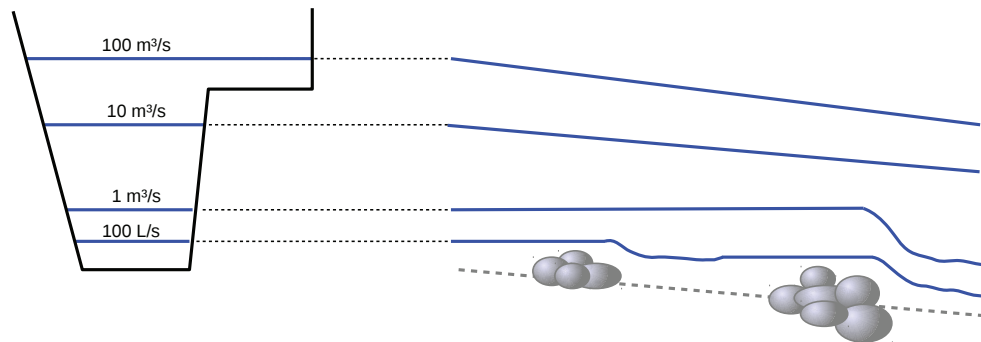

Figure 1: Illustration of the succession of section and channel hydraulic controls for a typical hydrometric station without an artificial control: bottom and water lines for different discharge values (right); the water levels are plotted against the river cross-section at the station (left).

Depending on the discharge, hydraulic controls may change, with some controls disappearing and others appearing. Typically, the section control exerted by a sill will disappear when water stage exceeds a given level of submersion or when the backwater effect is repelled downstream of the station. Fig. 1 shows how the transition between different section and channel controls may occur at a typical hydrometric station without an artificial control. With increasing discharge, the stage at the station is successively controlled by a small natural riffle, then by a higher one, then by the main channel only, and eventually by the main channel and the floodplain. More complex hydraulic controls may be activated successively or simultaneously.

Based on simplifications acceptable for hydrometry purposes, the usual hydraulic formulas for uniform channel controls and for conventional section controls can be expressed as the following power function [19, 55]:

$$Q = a \, (h - b)^c \tag{1}$$

2

where Q is the discharge, h ($h \geq b$) is the water level relative to a given datum (usually at the staff gauge), a is a scaling coefficient related to the characteristics of the control section or channel, b is a cease-to-flow reference level, and c is an exponent related to the type of hydraulic control.

Hydraulic theory provides nominal values for c such as 5/3 for a wide, uniform, rectangular channel control (derived from the Manning-Strickler equation), 3/2 for a rectangular weir control, 5/2 for a triangular weir control (derived from the critical flow equation). The value of the hydraulic exponent may also be determined experimentally for some control structures: $c \approx 1.55 - 1.60$ for commercial Parshall flumes, typically. The value of the exponent, c, may show some variability around the nominal value (say, ± 0.1) due to complex cross-sectional geometry or overbank flow processes. However, it is crucial to keep realistic values for c to allow for the physical derivation of coefficient a values, following the hydraulic formulas. Indeed, different values for the (a, c) couple with no physical meaning may better fit the observations, but yield very poor predictions in extrapolation.

1.2. Uncertainty analysis of stage-discharge relations

The methodology for assessing the uncertainty associated with stage-discharge relations is an important open scientific issue which received some attention in the recent literature. A first approach based mainly on hydraulic analysis of the stage-discharge relation can produce a valuable quantification of errors, since the physical basis of such errors is explicitly defined [49]. Sensitivity analysis of the parameters of a hydraulic model provides a realistic and site-specific estimation of error bounds [4, 23, 32, 5, 6]. However, translating these worst-case errors into probabilistic distributions from which uncertainty may be derived and combined is usually not a straightforward task.

The second family of approaches is based on the statistical analysis of gaugings. The work by Venetis [53] seems to be the first published statistically sound method for computing the uncertainty associated with rating curves, based on nonlinear regression of a single segment power function (cf. Eq. 1). In the same way, Dymond and Christian [8] suggested a new method accounting not only for rating curve error and stage error, but also for errors caused by ignoring all physical parameters other than stage. In works by Herschy [13], Clarke [1], Clarke et al. [2], the rating curve uncertainty analysis is based on the residual variance from regression of a power function like Eq. 1, and possibly on the standard error of the parameter estimates.

Petersen-Overleir [35] proposed a heteroscedastic model to take into account the usually observed heteroscedasticity of stage-discharge relations, which is not captured by classical non-linear least squares methods. The same author extended the non-linear regression approach to more complex stage-discharge relation cases, including multi-segment (or piecewise) power functions [38], hysteresis [36], and overbank flow in rivers with floodplains [37]. While this seminal work constituted a significant advance in stage-discharge analysis, the physical basis of the assumptions seems too loose since unrealistic hydraulic exponents (> 3, > 4) were sometimes obtained.

In recent years, the application of Bayesian inference [10] and Markov chain Monte-Carlo (MCMC) simulations has brought new solutions to the problem of rating curve uncertainty analysis. With computational costs that are now acceptable, MCMC simulation is an efficient way to estimate the posterior probability density function of the parameters of a rating curve, hence the corresponding uncertainty [28]. From the practitioner point of view, the Bayesian framework shows decisive advantages for the estimation and management of hydrometric rating curves. First of all, hydraulic knowledge of the stage-discharge relation can be explicitly translated into prior distributions of the parameters of the assumed rating curve equation. Second, as with any likelihood technique, it is possible to derive a likelihood function that accounts for the uncertainty in individual gaugings, leading to heteroscedastic models similar to that of Petersen-Overleir [35]. Based on the corresponding uncertainties, information from observations (likelihood function) and hydraulic knowledge (priors) are combined into a posterior distribution. Also, as the technique is directly based on probability density functions, uncertainty analysis can be achieved in an easy and clear way using percentiles of the posterior distribution.

To our best knowledge, Moyeed and Clarke [30] authored the first published Bayesian analysis of stage-discharge relations, based on a power function and Box-Cox transformation. However, some assumptions of this pioneering work appear questionable: no formal analysis of the hydraulic relations and of the discharge measurements was performed; the uncertainty on discharge data was quite arbitrarily assumed to be proportional to $\sqrt{h - h_0}$; and again, unrealistic hydraulic exponents were obtained. Other

Bayesian/MCMC studies of stage-discharge relations focused on the uncertainty due to the lack of gaugings for a single-segment power function [43], on the establishment, including extrapolation, of a piecewise power function [44], and on the rating procedures for gauging stations that are subject to variable backwater, using stage-fall-discharge relations for twin gauges [39]. While those published works proved the potential of Bayesian analysis for investigating all acknowledged sources of errors, the interpretation of hydraulic controls and the quantification of gauging uncertainties did not receive the attention corresponding to their practical interest.

Most of the aforementioned uncertainty analysis methods assumed that changes in the reference hydraulic regime are negligible or exceptional within each period of validity of the studied rating curve. Recently, non-stationary stage-discharge relations in unstable channels were addressed by different methods ranging from variographic analysis [17] to Bayesian analysis [45], weighted fuzzy regression within a moving time window [54] and the Generalised Likelihood Uncertainty Analysis (GLUE) methodology [12]. Again, the two main limitations of such studies are that uncertainties in individual gaugings were not sufficiently analyzed (in some cases, the errors in some gaugings may be as large as the detected non-stationarity); and that the analysis of hydraulic controls was too crude to account for the physical processes governing the non-stationarity. This is especially true when only one power-law segment was used while most often, the lower segment is much more prone to morphogenic or vegetation-driven non-stationarity than the upper ones (e.g., 5).

Most of the works cited in this literature review addressed the uncertainty of the stage-discharge relation. To assess the uncertainty of the discharge time series derived from the rating curve, it is further necessary to propagate the uncertainty of the stage time series proportionally to the inverse of the sensitivity of the stage-discharge relation. Only Olivier et al. [33], following the GUM [18] framework, and Freestone [9], in a more empirical way, accounted for the uncertainty in the water stage records. Indeed, in low flow conditions, for a poorly sensitive hydraulic control, this stage-related uncertainty component is often much larger than the uncertainty of the rating curve itself.

1.3. Objectives of this paper

In this paper, we introduce the BaRatin method for the Bayesian analysis of stationary rating curves. This framework is primarily based on the thorough elicitation of hydraulic priors. Methods for quantifying the uncertainty in individual gaugings are also introduced and the Bayesian inference assumptions are described. Then, we present the results obtained using the BaRatin method from three typical cases of hydrometric stations with contrasted site and flow configurations.

The application cases include (1) a densely gauged river with poor prior information on piecewise natural controls; (2) a small stream with precise prior knowledge on compound artificial controls; (3) a river where high-flow extrapolation of the rating curve is sensitive to the adequate interpretation of some tricky hydraulic controls. These application cases exemplify the importance of building physically-based hydraulic priors and checking the consistency of the obtained results. The value of more uncertain flood gaugings or discharge estimates is also highlighted, provided that their uncertainties are correctly assessed.

All the results presented in this paper were produced using the BaRatin software and its graphical user interface BaRatinAGE, which are freely available on request to the authors. Multilingual versions are supported, including French and English languages.

2. The Bayesian rating curve framework (BaRatin)

This section describes the statistical and hydraulic foundations of the BaRatin framework. We stick here to a factual presentation of the models and hypotheses behind BaRatin. Limitations and avenues for improvement will be thoroughly discussed in the discussion Section 4.

2.1. Bayesian inference

BaRatin is the application of the Bayesian paradigm [10] to the problem of estimating rating curve parameters, based on uncertain gaugings and hydraulic knowledge. It is based on the following statistical model.

Measurement error models. Gauging measurements $(\widetilde{H}_i, \widetilde{Q}_i)_{i=1:N}$ are estimations of the real values $(H_i, Q_i)_{i=1:N}$ of water levels and associated discharges. The following error models are assumed:

$$\widetilde{H}_i = H_i \tag{2}$$

$$\widetilde{Q}_i = Q_i + \epsilon_i^Q \qquad\qquad \epsilon_i^Q \sim \mathcal{N}(0, u_{Q_i}) \tag{3}$$

Eq. 2 assumes that the errors in stage measurements are negligible. Eq. 3 assumes that discharge measurements are affected by Gaussian errors with mean zero (i.e. no bias) and known standard deviation u_{Q_i}, i.e. the so-called standard uncertainty.

Rating curve. The rating curve is formalized as a function $f(h|\boldsymbol{\theta})$, where h is the stage and $\boldsymbol{\theta} = (\theta_1, \dots, \theta_{N_{par}})$ are the rating curve parameters. In most cases, the function f is given by the elementary hydraulic control equation 1, or a piecewise combination of several elementary hydraulic controls, as will be introduced in next Section 2.2. In statistical terms, the fact that we are using a piecewise formulation of the rating curve corresponds to a segmented regression context.

The stage-discharge relationship is then formalized as follows:

$$Q_i = f(H_i|\boldsymbol{\theta}) + \epsilon_i^f \qquad\qquad \epsilon_i^f \sim \mathcal{N}(0, \sigma_f) \tag{4}$$

The error term ϵ_i^f is the remnant error, and represents the fact that the rating curve, in its assumed mathematical form, would not perfectly represent the stage-discharge relationship, even if the true stage/discharge values were known. Note that the standard deviation σ_f is unknown and therefore has to be estimated. We also make the assumption that the remnant error ϵ_i^f and the measurement error ϵ_i^Q are independent.

Combining Eqs. 2, Eq. 3 and Eq. 4 yields the following relationship between observed stage-discharge values:

$$\widetilde{Q}_i = f\left(\widetilde{H}_i|\boldsymbol{\theta}\right) + \epsilon_i^f + \epsilon_i^Q \qquad\qquad \epsilon_i^f + \epsilon_i^Q \sim \mathcal{N}\left(0, \sqrt{\sigma_f^2 + u_{Q_i}^2}\right) \tag{5}$$

Bayesian equations. Eq. 5 involves several unknown quantities: the rating curve parameters $\boldsymbol{\theta}$ and the remnant standard deviation σ_f. Inference on these quantities is performed through the posterior distribution, the probability density function (pdf) of which is defined as:

$$\underbrace{p\left(\boldsymbol{\theta}, \sigma_f \mid \widetilde{\mathbf{H}}, \widetilde{\mathbf{Q}}\right)}_{\text{posterior}} \propto \underbrace{p\left(\widetilde{\mathbf{Q}} \mid \boldsymbol{\theta}, \sigma_f, \widetilde{\mathbf{H}}\right)}_{\text{likelihood}} \underbrace{p\left(\boldsymbol{\theta}, \sigma_f\right)}_{\text{prior}} \tag{6}$$

The likelihood of observed discharge values $\widetilde{\mathbf{Q}}$, given rating curve parameters $\boldsymbol{\theta}$, remnant standard deviation σ_f and observed stage values $\widetilde{\mathbf{H}}$ is given by:

$$p\left(\widetilde{\mathbf{Q}} \mid \boldsymbol{\theta}, \sigma_f, \widetilde{\mathbf{H}}\right) = \prod_{i=1}^{N} p_N\left(\widetilde{Q}_i \mid f(\widetilde{H}_i \mid \boldsymbol{\theta}), \sqrt{\sigma_f^2 + u_{Q_i}^2}\right) \tag{7}$$

where $p_N(z \mid m, s)$ is the pdf of a Gaussian distribution with mean m and standard deviation s, evaluated at value z. The prior pdf is derived by assuming independent priors on each inferred quantity:

$$p\left(\boldsymbol{\theta}, \sigma_f\right) = p\left(\sigma_f\right) \prod_{i=1}^{N_{par}} p(\theta_i) \tag{8}$$

Note that hydraulic knowledge can be used to derive informative priors on at least some of the parameters θ_i, depending on the type of hydraulic control. In particular, the exponents in the rating curve equation can be quite precisely elicited for each hydraulic control (see discussion in next section 2.2). Conversely, it is much more difficult to elicit an informative prior distribution for the remnant standard deviation σ_f. We will therefore use a wide uniform distribution (between 0 and 10,000) for this parameter.

5

Inference. The posterior distribution (Eq. 6) is explored by sampling realizations using a Markov Chain Monte Carlo (MCMC) sampler (see [46] for a detailed description). This generates an ensemble of realizations $\left(\boldsymbol{\theta}_j, \sigma_{f,j}\right)_{j=1:N_{sim}}$, with $N_{sim} = 10^4 - 10^5$, typically. For each of these realizations, a rating curve is computed from parameters $\boldsymbol{\theta}_j$. From this large set of rating curves, the statistics computed at each stage value allows assessing the *parametric uncertainty* of the stage-discharge relation. To assess the *total uncertainty*, a Gaussian noise with mean zero and standard-deviation $\sigma_{f,j}$ is added to each rating curve.

Note that there is room to improve the MCMC sampling algorithm by customizing it to the peculiarities of the BaRatin model. However, we never encountered convergence issues over the 50 gauging stations analyzed so far with BaRatin. This is most probably due to the use of informative, hydraulics-based priors, which makes the inference well-behaved. In the absence of such priors, the posterior distribution would probably be much more complex (e.g. multimodality), especially in this context of segmented regression.

As a summary, taking as input the prior knowledge on hydraulic controls and the gauging data with their individual uncertainties, the Bayesian simulator provides the *MaxPost* rating curve that corresponds to the maximum-posterior parameter values, along with its uncertainty interval. Following the common practice in hydrometry [16], uncertainty intervals are defined as 95% credibility intervals. It is important to acknowledge that observations (gaugings) should not be used to achieve the elicitation of priors. If some gaugings are absolutely necessary to do so, they must be removed from the observation dataset provided to the Bayesian simulator.

2.2. Elicitation of hydraulic priors

Identification of elementary hydraulic controls. The main step of the Bayesian Rating curve framework (BaRatin) is to conduct a hydraulic analysis of the study site in order to identify the hydraulic controls at work at different ranges of stage, which will constitute the prior knowledge of the Bayesian inference. Such a hydraulic analysis may be achieved based on the observation of the channel geometry and roughness, as well as the flow patterns over a range of discharge values, within the minimum and maximum water heights for which the hydrometric station is designed to be operated. Field surveys, expert knowledge of the station managers, topography campaigns, photographs and movies, water level records, flood marks and any other kind of documentation are usually very helpful. Numerical hydraulic models may also be used to predict the stage-discharge relation, including the transitions between successive hydraulic controls, especially for situations including overbank flow, flow singularities or artificial structures.

As will be illustrated hereafter with some application cases, most of the hydraulic controls in rivers can be acceptably accounted for using Eq. 1. Each parameter, a, b and c, of each identified control must be assigned a prior distribution. In the following, we will assume that they follow Gaussian distributions and provide the modal value (prior mean) and standard deviation (prior width) of each parameter. Non-Gaussian priors can also be used without additional difficulty.

Control type	ideal assumptions	a	b	c
Channel	wide rectangle, steady and uniform flow	$K_s\, B\, \sqrt{S_e}$	average bed level	5/3
Section (rectangle)	perpendicular to mean flow, no backwater	$C_r\, B_r\, \sqrt{2g}$	crest level	3/2
Section (triangle)	perpendicular to mean flow, no backwater	$C_t\, \tan\beta\, \sqrt{2g}$	lower vertex level	5/2
Section (orifice)	perpendicular to mean flow, no backwater	$C_o\, A_o\, \sqrt{2g}$	orifice centre level	1/2

Table 1: Mean prior values for parameters, a, b and c, given by classical hydraulic formulas for most common types of hydraulic controls. The definitions of all the variables are provided in the text. We always set $g = 9.81$ m/s^2 for gravity acceleration.

Tab. 1 summarizes the mean prior values for parameters a, b and c given by textbook hydraulic formulas for most common hydraulic control types encountered in rivers and canals. For a channel control, the Manning-Strickler equation is usually applied to more or less steady, uniform flows:

$$Q = A\, K_s\, S_e^{1/2}\, R_h^{2/3} \tag{9}$$

with A, the wetted area, K_s, the Strickler flow resistance coefficient, S_e, the flow energy slope, and R_h, the hydraulic radius. For a wide rectangular cross-section, Eq. 9 reduces to Eq. 1 with $a = BK_sS_e^{1/2}$, $b = z_b$, and $c = 1.67$. The flow width, B, and the average bed level, z_b, can be derived from the geometry of the

6

uniform reach, modeled by an equivalent rectangle. Note that upstream as well as downstream elements can govern the stage-discharge relation in case they have a large influence on the average geometry of the channel in which is located the station. S_e can be estimated either from the water surface slope or from the average bed or valley slopes. K_s, or equivalently the Manning coefficient, $n = 1/K_s$, can be estimated from usual default values depending on relevant indicators of head losses, such as grain size, bedforms, sinuosity and land cover of the main channel or the floodplain. Most natural channels and artificial canals present cross-sectional shapes that can be fairly approximated by a wide rectangle, or a mixture of wide rectangles, which is therefore the method we suggest to follow. Considering a realistic concave cross-section, the highest sensitivity would correspond to a triangular channel control (of which we personally know no example in natural streams), which leads to $c = 2.33$. Hydraulic exponents well above this value should therefore be considered as suspicious.

Formulas for section controls include a discharge coefficient with nominal values $C_r = 0.4$ for an ideal rectangle weir, $C_t = 0.31$ for an ideal triangle weir, $C_o = 0.6$ for an ideal orifice. Deviations to nominal values may be ±0.2 or even more in complex, real situations. The cross-section geometry is defined by B_r (rectangle weir width, perpendicular to flow), β (triangle half-angle), A_o (orifice area), and the reference level, b. When necessary, the geometry of the control section must be projected perpendicularly to the mean flow, as a simple approximation. Prior widths on a, b, and c can be easily derived by conducting a sensitivity analysis on all these parameters in the corresponding formulas. When numerical simulation is performed using a 1-D or 2-D (more seldomly 3-D) hydraulic model, sensitivity analysis on salient parameters of the model can also be conducted to document prior widths on a, b, and c. The Strickler flow resistance coefficient, K_s, is usually the main calibration coefficient to be varied over a range of realistic values.

Note that the prior equation 8 assumes independent priors for each parameter a, b, and c. This is a reasonable hypothesis in general, because they are deduced from distinct properties of the section/channel. If deemed too restrictive, the assumption of prior independence can easily be relaxed by directly specifying the joint prior of all parameters.

A versatile rating curve equation. As soon as hydraulic controls are identified and priors are specified, it is possible to model most stage-discharge relations using compound or piecewise combinations of power functions like Eq. 1, which are individually suited for different segments of the stage-discharge relation. This can be written using the following versatile equation for N_{range} stage ranges and $N_{control}$ hydraulic controls:

$$Q = \sum_{r=1}^{N_{range}} \left(\mathbf{1}_{[\kappa_{r-1};\kappa_r]}(h) \times \sum_{j=1}^{N_{control}} M(r,j) \times a_j \left(h - b_j\right)^{c_j} \right) \qquad (10)$$

The function $\mathbf{1}_{[\kappa_{r-1};\kappa_r]}$ is equal to 1 if $\kappa_{r-1} \leq h \leq \kappa_r$, to 0 otherwise. The elevation κ_r is the transition water level between stage ranges r and $r + 1$. Such transition levels also have to be defined as parameters of the hydraulic priors, with provided modal values and standard deviations. For stage ranges higher than the first one, the values of parameters b_j will be actually computed so as to maintain the continuity of the stage-discharge function at each transition level.

The hydraulic control matrix, M, has to be defined too: $M(r,j) = 1$ if hydraulic control j is active in the stage range r; $M(r,j) = 0$ otherwise. The generality of this rating curve equation (Eq. 10) allows for representing most hydraulic situations. Typically, the matrix of the hydraulic controls depicted in the example in Fig. 1 comprises 4 ranges (4 rows) with 4 columns for the hydraulic controls by the first riffle, the second riffle, the main channel and the floodplain, respectively. It simply writes:

		riffle 1	riffle 2	main channel	floodplain	
	range 1	1	0	0	0	
$M(r,j) =$	range 2	0	1	0	0	(11)
	range 3	0	0	1	0	
	range 4	0	0	1	1	

This example illustrates that if one hydraulic control is drowned, then a new one takes over (rows 1-2-3). But a new control can add to an existing one, yielding this control to be active in two or more segments (e.g. rows 3 and 4 for main channel, and main channel+floodplain). In the latter situation, parameters b and κ of the new control are physically equal.

2.3. Observations: uncertainty in individual gaugings

The BaRatin framework also requires the analysis of individual gaugings, i.e., the observations provided to the Bayesian simulator. Indeed, rating curve uncertainty analysis requires the correction of all identified errors in gaugings, and the estimation of the individual uncertainty of each gauging. Gaugings conducted by hydrometry staff consist of direct measurements of stage and discharge over a reduced period of time, in given hydraulic conditions. It should be noted that most authors recognize that uncertainty in stage measurements of individual gaugings are negligible compared with uncertainty in discharge measurements.

At least three kinds of errors should be considered when such measurements are to be used for establishing or evaluating a stage-discharge relation [27]:

1. measurement errors, combining instrumental errors, environmental errors, and spatial integration errors;
2. time integration error due to possible flow variability during the measurement period of time;
3. stage-discharge bias due to non-reference flow conditions, such as the dynamic hysteresis effect (e.g., 31). The latter effect is not a measurement error, but a deviation of the real flow conditions from the ideal conditions for which the rating curve is built.

The literature provides some methods for quantifying discharge uncertainty according to the measurement technique (e.g. for current-meters 14, 34, 15). However, the task remains challenging for many practitioners facing a range of situations and techniques, and some research work is still in progress in order to improve the uncertainty analysis of individual discharge measurements [3, 26]. In order to parameterize such error propagation modeling (cf. 18), interlaboratory field comparison tests [27] provide valuable information.

In the application examples of this study, typical uncertainty values were assigned to gaugings according to the measurement technique and instrumentation, and to quality criteria such as the number of velocity points and verticals in velocity-area gaugings [26], or the repeatability coefficient and the extrapolated discharge ratio for ADCP gaugings [25]. These uncertainty estimation procedures are consistent with common practices in hydrometry. Typically, the uncertainties assigned to good-level current-meter and ADCP gaugings are 7% and 5%, respectively, while uncertainties for surface velocity gaugings (handheld radar velocimeter, LSPIV image sequence analysis) can reach 10% or 20%, according to the data quality and the site conditions.

3. Application cases

3.1. Study sites and available information

Three hydrometric stations were selected to exemplify important aspects of the application of the BaRatin method for contrasted situations. The first application case is typical of a natural river with uncertain prior knowledge on the hydraulic controls. BaRatin yields correct results as far as the hydraulic exponents are given realistic values and the uncertainty of available gaugings is correctly assessed. In the second application case, the section controls exerted by an artificial structure can be accurately predicted. BaRatin proves to be able to account for these compound controls with abrupt transitions. The third application case shows that in more complex river situations wrong assumptions on the hydraulic controls may lead to significantly biased results, especially for the high-flow extension of the rating curve. It is therefore essential to verify the consistency of the results with the structure of the hydraulic controls used to define priors. For such situations, numerical simulation using a simple 1D hydraulic model is often of great help for establishing the structure of the control matrix.

A brief description of the three study sites and available information on their hydraulic controls and gaugings is provided hereafter.

Draining a catchment area of 2240 km^2, the Ardèche river at Sauze flows in a straight channel carved in an inerodible limestone bedrock, with a width of ~ 80 *m* and a mean longitudinal bed slope of 0.3‰ [24]. From past hydraulic studies, the Strickler flow resistance coefficient is estimated to be $K_s \approx 33$ m$^{1/3}$/s. The hydrometric station is located in a deep pool with a low-flow section control which is governed by a natural gravel riffle located ~ 500 *m* downstream (see Fig. 2a). This riffle section is ~ 30 *m* wide. The flat gravel crest may be modified due to floods or recreational activities. Over the considered recent period of time, it was estimated to be stable with a mean elevation roughly 0.50 m below the staff gauge zero level. Gaugings are densely distributed from 4.8 m^3/s to 2700 m^3/s, which corresponds to the 5 year flood [24]. The mean annual discharge is 63 m^3 s^{-1}. Gauging techniques include mechanical currentmeters deployed from a cableway (estimated uncertainty: 7%), vessel-mounted ADCP (estimated uncertainty: 5%), and image-based surface velocity gaugings. The latter remote technique induces uncertainties ranging from 10% to 20% in good image conditions, and from 30% to 80% in adverse image conditions [24]. Only image-based surface velocity gaugings conducted in good image conditions were kept for the present study.
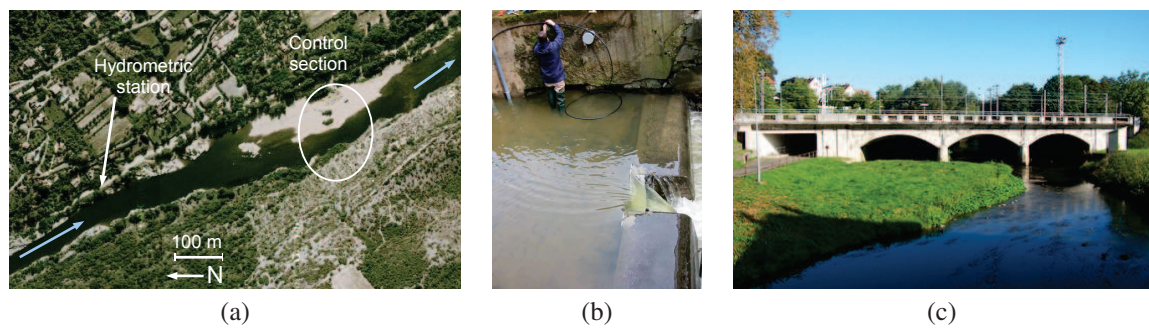

Figure 2: Views of the hydraulic controls of the three application cases: the Ardèche river at Sauze (a); the Charbonnières river at Charbonnières-les-Bains (b); the Sarre river at Sarrebourg, viewed from upstream (c). All three sites are located in France.

The hydrometric station of the Charbonnières River at Charbonnières-les-Bains is located on a small peri-urban river, draining an area of 23 km^2. The hydraulic control is artificial: a V-notch weir was inserted within a flat-crested sill (Fig. 2b). It is therefore a typical case in which simple hydraulic priors can be predicted accurately from the measured geometry of the control section with narrow prior widths. As a consequence, here the hydraulic priors are more informative than the gaugings. Available gaugings span all over the discharge range from 0.02 to 2 m^3/s. Their uncertainties were estimated from conventional values for the optimal application conditions of the variety of streamgauging techniques used: velocity-area with electro-magnetic and acoustic current-meters (7%), acoustic profiler ADCP (5%), saline and fluorescent tracer dilution (5%), handheld radar velocimeter (10%). A larger uncertainty (15%) was attributed to velocity-area gaugings conducted in very shallow flow conditions.

The Sarre River at Sarrebourg is a medium-size, lowland river (catchment area is 311 km^2) with a mild slope ($\approx$ 1 m/km). Since 1953, a pressure gauge is located right upstream a large railway bridge, in the right bank: the cabin of the hydrometric station is visible in Fig. 2c. The bridge has three 10 m-wide arches and there is also a 4 m-wide road tunnel on the left side, the road elevation laying slightly above 3 m at staff gauge. The bed of the main channel is made of sand and fine gravels, with long algae growing during the Spring and Summer periods. A large bar covered with grass formed upstream of the hydrometric station (Fig. 2c). Available gaugings are numerous and quite scattered for low flow conditions, likely due to non-stationary effects (vegetation, changes of the micro-topography) which could not be distinguished clearly from the available information. An interesting set of gaugings are distributed over the medium flow segment. Two recent ADCP gaugings were performed at high flows, showing a sharp increase in flood conveyance. Their validity was confirmed from the comprehensive analysis of the raw ADCP data and the comparison with surface velocities inferred from movies that were recorded during the measurement operations.

3.2. The Ardèche river at Sauze: the importance of gauging uncertainties and hydraulic exponent values

We are facing a typical river case where no comprehensive topography survey and a fortiori no numerical model were used to build precise hydraulic priors. However, the aforementioned information on the site allows to set orders of magnitude for the prior parameters of a rectangle section control, followed by a channel control (see priors in Tab. 2). The transition stage is very roughly estimated to be 1 m at staff gauge, assuming that the natural riffle is submerged for a flow depth around 1.5 m above the crest.

Wide uncertainties were attached to these prior values to define the first set of *priors with uncertain c* (Tab. 2). A second set of priors was set up by considering narrow prior values for *c* exponents (*priors with precise c*), according to the hydraulic formulas corresponding to both types of controls. Each set of priors was tested with BaRatin against either all available gaugings (Fig. 4a) or against a subset including image-based (LSPIV) gaugings only above stage 1 m (Fig. 4b). This subset was designed to test a typical situation where only remote techniques with larger uncertainty could provide flood gaugings. In order to test the importance of accounting for the uncertainties of individual gaugings, all the previous simulations were replicated assuming equal uncertainties for all gaugings. The constant uncertainty was set to 5% as if all gaugings were of good quality. The results of the 8 simulation tests are presented in Tab. 2 and Fig. 4.

Fig. 3 shows the direct output of MCMC sampling (using precise priors for *c* and all gaugings). This representation allows observing prominent features of the posterior distribution. In this case, the most apparent feature is the large correlation between parameters a/b (control 1) and a/c (control 2). All marginal distributions appear to be unimodal and fairly symmetric. Those posterior samples can then be propagated through the rating curve equation, yielding an ensemble of rating curves from which one can extract the MaxPost rating curve and uncertainty envelopes, as shown in Fig. 4.

When all the existing gaugings are used, the 4 options yield MaxPost rating curves with narrow uncertainty envelopes that are in close agreement with each other, even in the high-flow extrapolated section (Fig. 4a). Even when wide uncertainties are used for all κ, a, b, c parameters, the MaxPost results are consistent with the hydraulic interpretation. Thanks to the high quality of the gaugings dataset and to the stability of the hydraulic controls, this validation case confirms the performance of the BaRatin method when many gaugings are available. The values for parameters κ, a and b are consistent with the physical dimensions of the riffle and of the channel. However, the hydraulics-based exponent values for both controls are more accurately reproduced when the gauging uncertainties are correctly assessed ($c_1 = 1.45$ and $c_2 = 1.64$) than when larger uncertainties are ignored ($c_1 = 1.97$ and $c_2 = 1.59$). As expected, setting precise priors for *c* exponents stabilizes the results whatever the assessment of gauging uncertainties.

When only a subset of existing gaugings is used, uncertainty envelopes are wider and the highest parts of the MaxPost rating curves are significantly different for the 4 tested options (Fig. 4b). When gauging uncertainties are correctly assessed, but priors with uncertain *c* are used, the MaxPost values for parameters κ_1 and a_2 are significantly overestimated, while exponents $c_1 = 1.83$ and $c_2 = 1.24$ cannot be related to known hydraulic formulas (Tab. 2). Such unrealistic results arise from the high level of interaction between simulated values for parameters *a* and *c*. This is the reason why it is advised to fix narrow prior uncertainties for the *c* values, which yields realistic values for parameters κ, *a* and *b*, even with a reduced quantity of gaugings with increased uncertainty (see Tab. 2). The results are totally unrealistic when priors with uncertain *c* are used and larger uncertainties in flood gaugings are ignored: meaningless parameters are computed by BaRatin so that the high-flow segment be aligned along the two highest gaugings. In that specific run, some MCMC realizations with $c \approx 0$ lead to a variance greater than the machine overflow for parameter b_2. Here again, forcing the values of the *c* exponents according to the hydraulics theory is useful to stabilize the MaxPost results even when less gaugings are available and/or the uncertainty analysis for each individual gauging is more crude.

Lastly, Tab. 2 also reveals that the remnant standard deviations are much larger when an unduly optimistic 5% uncertainty is assumed for all gaugings. This illustrates that the remnant error compensates for the misspecification of gauging uncertainty, by mistakingly interpreting measurement errors as structural errors. This reiterates the importance of accurately quantifying measurement uncertainty in each individual gauging.
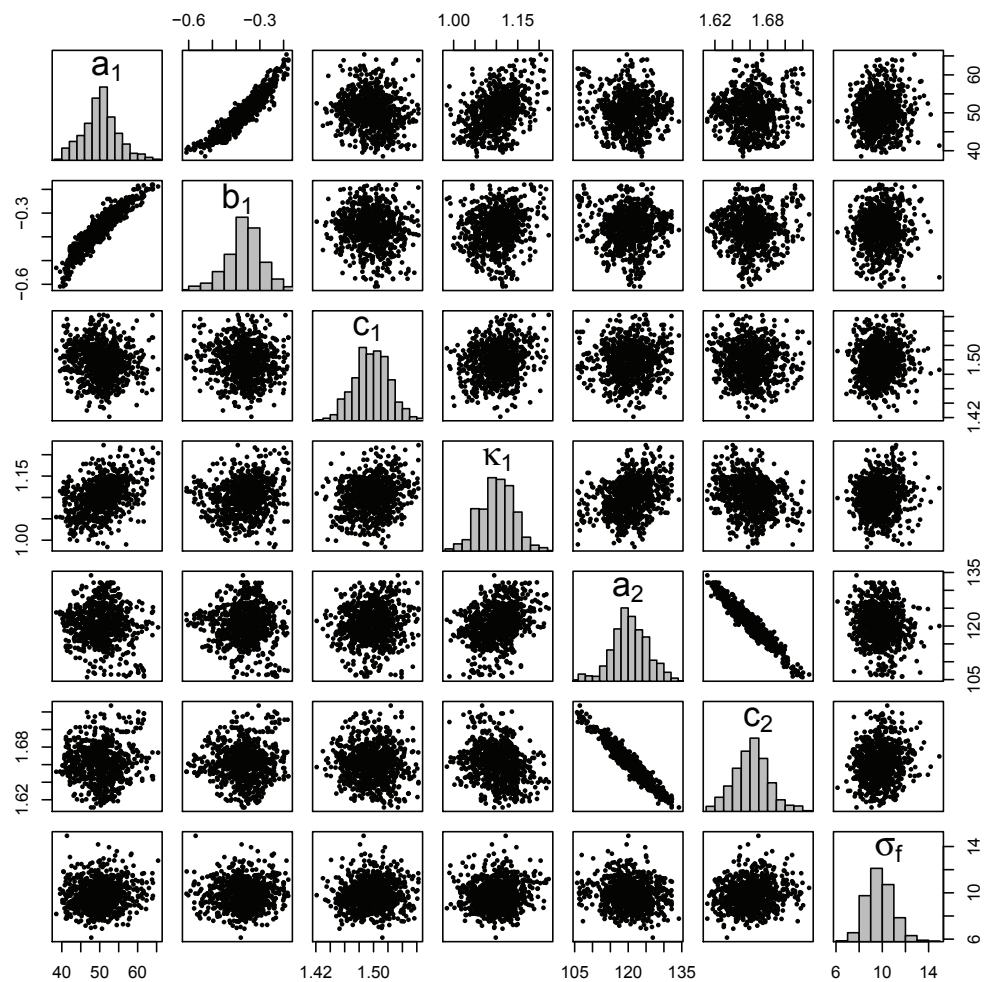
Figure 3: The Ardèche river at Sauze: scatterplot matrix of MCMC samples from the posterior distribution (using precise priors for c and all gaugings.)
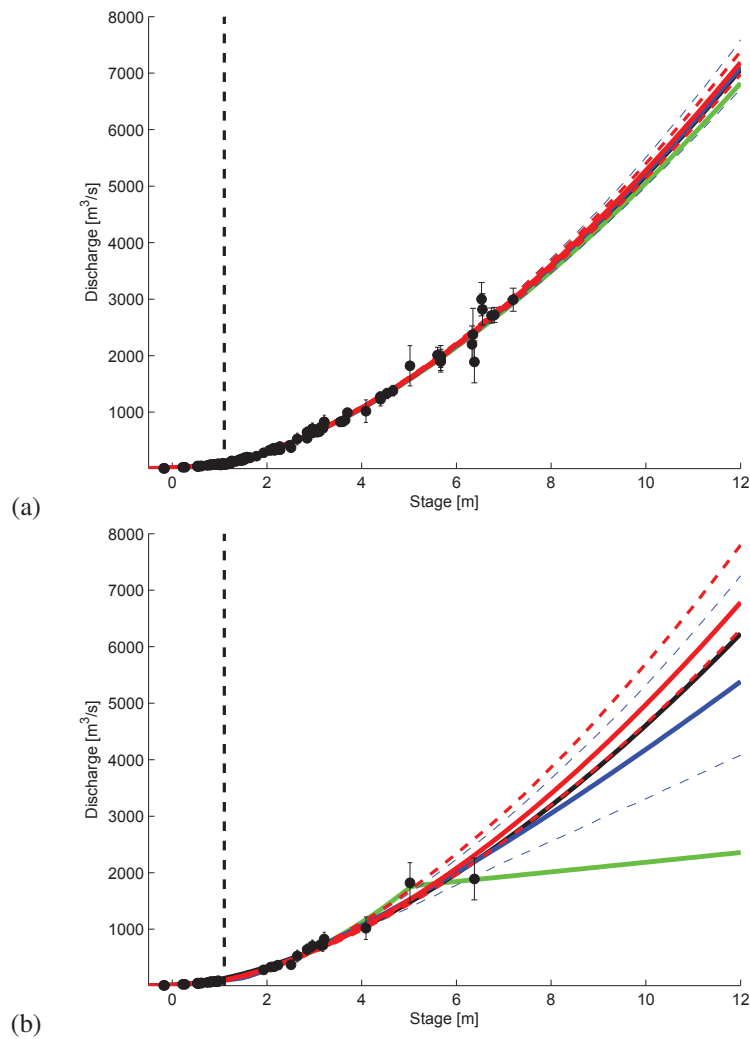
Figure 4: The Ardèche river at Sauze: gaugings with individual uncertainties, MaxPost rating curves and uncertainty intervals, as provided by the BaRatin method using all gaugings (a) or only image-based (LSPIV) gaugings for stage greater than 1 m (b). The MaxPost rating curves are shown as solid lines for precise priors for *c* (red) and wide priors for *c* (blue). Corresponding uncertainty envelopes are shown with dashed lines. Assuming a constant 5% uncertainty for all gaugings, the MaxPost rating curves obtained for precise and uncertain priors for *c* are shown as black and green solid lines, respectively. The computed transition level between section and channel controls is marked with a vertical dashed line (around stage 1.1 m).

	Control 1			κ_1	Control 2			
	a_1	b_1	c_1		a_2	b_2	c_2	σ_f
Priors with uncertain c								
Prior	50 (100)	-0.50 (1)	1.50 (2.5)	1 (1)	100 (200)	-	1.67 (2.5)	*
Posterior 1: all gaugings	49 (5.8)	-0.39 (0.077)	1.45 (0.19)	1.10 (0.052)	124 (12)	0.29 (0.068)	1.64 (0.048)	8.7 (1.2)
Posterior 2: subset	40 (4.0)	-0.51 (0.055)	1.83 (0.11)	1.56 (0.21)	252 (69)	0.89 (0.29)	1.24 (0.17)	1.7 (0.58)
Posterior 3: all gaugings ($U = 5\%$)	36 (16)	-0.50 (0.30)	1.97 (0.69)	1.15 (0.37)	137 (8.0)	0.35 (0.053)	1.59 (0.025)	33 (5.2)
Posterior 4: subset ($U = 5\%$)	59 (7.6)	-0.23 (0.082)	2.04 (0.076)	5.05 (0.11)	58 (128)	-17.8 (∞)	1.09 (0.40)	28 (5.2)
Priors with precise c								
Prior	50 (100)	-0.50 (1)	1.50 (0.25)	1 (1)	100 (200)	-	1.67 (0.25)	*
Posterior 1: all gaugings	48 (4.3)	-0.40 (0.075)	1.50 (0.024)	1.10 (0.037)	123 (4.3)	0.29 (0.026)	1.65 (0.017)	9.1 (1.2)
Posterior 2: subset	55 (3.1)	-0.31 (0.038)	1.50 (0.024)	1.02 (0.12)	107 (9.3)	0.14 (0.091)	1.68 (0.024)	2.4 (0.62)
Posterior 3: all gaugings ($U = 5\%$)	51 (12)	-0.38 (0.23)	1.49 (0.025)	1.05 (0.12)	117 (4.8)	0.23 (0.04)	1.66 (0.018)	33 (5.2)
Posterior 4: subset ($U = 5\%$)	63 (23)	-0.30 (0.29)	1.51 (0.025)	0.63 (0.35)	97 (7.1)	-0.09 (0.11)	1.67 (0.025)	56 (11)

Table 2: The Ardèche at Sauze: prior and posterior modes (and standard deviation) for parameters κ, a, b, and c of the hydraulic controls, as used for the BaRatin runs. The results are presented for 4 simulations: using all gaugings (Posterior 1) or using image-based (LSPIV) gaugings only for stages above 1 m (Posterior 2); the same sets of gaugings but assuming equal uncertainties set to 5% (Posterior 3 and Posterior 4). * a uniform distribution U[0;10,000] is used.

3.3. The Charbonnières river at Charbonnières-les-Bains: accounting for compound artificial controls

In this application case typical of small streams equipped with artificial controls, simple hydraulic formulas provide a much more precise prior knowledge than in the previous example. The problem in using Eq. 10 to describe such compound artificial controls as a triangle weir inserted in a horizontal sill is that when the sill is active, the V-notch remains active but the corresponding flow no longer expands laterally. It is therefore inaccurate to use a simple triangle weir formula for the first control when $h > b_2$, with b_2 the cease-to-flow elevation of the horizontal sill. The following equation has to be used instead of Eq. 10 when $h > b_2$:

$$Q = \underbrace{a_1(h - b_1)^{c_1} - a_1(h - b_2)^{c_1}}_{\text{bounded triangle weir}} + \underbrace{a_2(h - b_2)^{c_2}}_{\text{rectangle weir}} \qquad (12)$$

The first control is defined by a modified version of the triangle weir formula, with an additional negative term accounting for the unnecessary cross-section area of the divergent triangle above the level of the horizontal sill: a_1 is computed with $C_t = 0.31$ and $\beta = 45°$ (cf. Tab. 1); b_1 was accurately measured to be 0.053 m; $c_1 = 2.5$ (triangle weir equation); b_2 is the average measured level of the sill crest (0.403 m). The second control follows the conventional equation for the flow over the horizontal sill (cf. Tab. 1): a_2 is computed with $C_r = 0.4$ and $B = 4.60$ m; $c_2 = 1.5$. Hence, this is a specific situation where the first control formula in Eq. 10 has to be modified for a bounded triangle weir, with an additional parameter, b_2, which is common with the second control formula (cf. Eq 12).

First BaRatin was run using these precise priors on the artificial hydraulic controls. Parameters are listed in Tab. 3 and gaugings, priors and results are plotted in Fig. 5a. As expected, the results are very close to the prior modal values for all parameters. The transition between hydraulic controls by the V-notch and by the horizontal sill is accurately reproduced at the correct stage, and in close agreement with the gaugings. This suggests that the simple continuity condition between segments used in BaRatin allows for control transitions, even the most abrupt ones.

Also noticeable are the apparently too narrow uncertainty envelope (gray area) computed for high flows and the apparently too wide uncertainty envelope computed for low flows. In high flows, the scatter of the uncertain gaugings around the MaxPost rating curve seems too large to be fully attributable to measurement uncertainty. This may be due to the assumption of an additive remnant uncertainty σ_f for the whole rating curve.

As a comparison test, a second run was done using non-informative priors: all parameters were assumed to follow a uniform distribution between 0 and 20 (Tab. 3). BaRatin accurately reproduced the hydraulic formulations by fitting the dense set of available gaugings (Fig. 5b). While slightly different, MaxPost parameter values are still consistent with the accurate priors, regarding the geometry of the weirs (parameters a and b) and the hydraulic exponents: $c_1 = 2.72$ instead of 2.5 (V-notch), and $c_2 = 1.29$ instead of 1.5. The lowest end of the MaxPost curve now deviates from the prior rating curve, and the uncertainty envelope is even larger than before for the triangle weir control. This is mainly due to an increase in the parametric uncertainty (σ_f is lower, cf. Tab. 3), because the priors used in this simulation are non informative.

This application example shows that BaRatin, if necessary with a modified version of Eq. 10, is able to account for complex artificial controls with abrupt transitions. As expected, the influence of the precision in priors is more important where gaugings are less abundant or more uncertain. Also, the assumption of an additive remnant uncertainty, σ_f, possibly leads to questionable uncertainty envelopes.

Note that a reasonable approximation of the hydraulic controls can also be built on Eq. 10 by substituting an equivalent triangular weir to the V-notch when $h > b_2$. This 3-controls approximation leads to equivalent results when precise priors are used, but not when non-informative results are used.
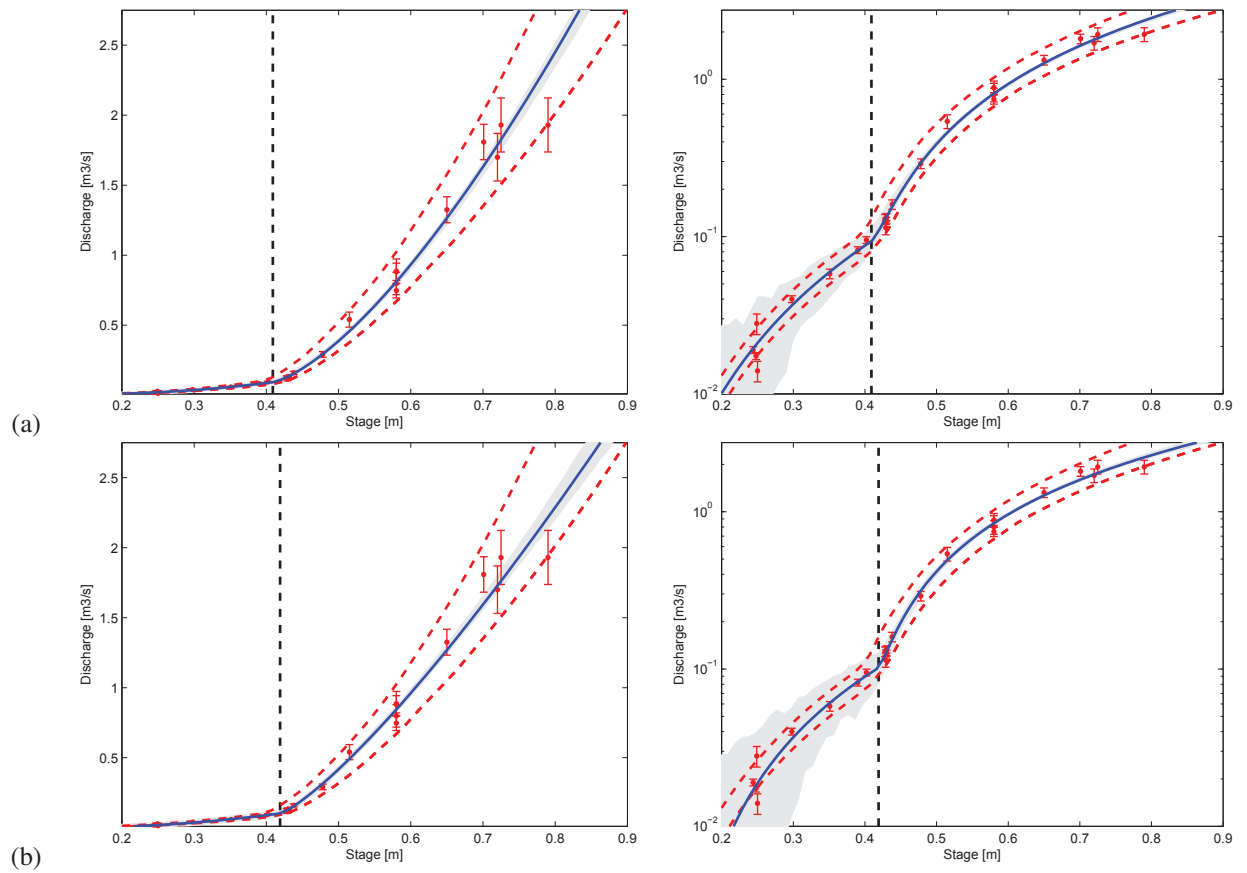
14

Figure 5: The Charbonnières river at Charbonnières-les-Bains: MaxPost rating curve (solid blue line) and its uncertainty interval (gray patch), as provided by the BaRatin method using precise (a) or non informative (b) hydraulic priors. The same gaugings with individual uncertainties (red marks) are used, and the 95% uncertainty envelope of the precise priors (red dashed lines) is plotted, even when non informative priors were used (b). Results are shown with discharge in linear scale (left) and log scale (right). The computed transition level at which the rectangle weir control adds to the bounded triangle weir control is marked with a vertical dashed line (around stage 0.41 m).

| | Control 1 | | | | Control 2 | | | σ_f |
	a_1	b_1	c_1	k_1	a_2	b_2	c_2	
Precise priors								
Prior	1.24 (0.1)	0.053 (0.005)	2.50 (0.025)	0.403 (0.01)	8.0 (1)	0.409 (0.0032)	1.50 (0.025)	*
Posterior	1.25 (0.061)	0.056 (0.0043)	2.49 (0.022)	0.409 (0.0032)	7.92 (0.26)	0.409 (0.0032)	1.48 (0.022)	0.0044 (0.0022)
Non informative priors								
Prior	**	**	**	**	**	-	**	*
Posterior	1.77 (0.31)	0.061 (0.021)	2.72 (0.16)	0.422 (0.0043)	5.88 (0.45)	0.422 (0.0043)	1.29 (0.048)	0.0040 (0.0022)

Table 3: The Charbonnières river at Charbonnières-les-Bains: prior and posterior modes (and standard deviation) for parameters a, b, and c of the hydraulic controls, as used for the BaRatin runs. A modified hydraulic formula for bounded triangle weir is used for the first control (see text). * a uniform distribution U[0;10,000] is used. ** a uniform distribution U[0;20] is used.

3.4. The Sarre river at Sarrebourg: identifying tricky controls using numerical simulation

A 1-D hydraulic model was built to numerically simulate reliable priors for the medium to high flow conditions at the hydrometric station of the Sarre river at Sarrebourg (cf. Fig. 6). The priors for the low flow segment were computed using the rectangle weir formula, since the hydraulic model was not suited to predict low flow conditions accurately. This simulation exercise made us realize that the other hydraulic controls were actually governed by the narrow channel and the bar of the reach located just upstream of the station (cf. Fig. 2c), which was far from obvious. From sensitivity tests on the flow resistance coefficients, prior modal values and widths are determined for both of these controls (cf. Tab. 4). The lowest and highest values for flow resistance coefficients are used to compute the prior uncertainty of parameters a for the medium and high flow controls. Using MCMC simulation to sample possible rating curves from these parameters, the 95% uncertainty envelope of the prior rating curve can be computed (cf. Fig. 6).

To test the influence of an inappropriate control structure for building priors, three sets of hydraulic priors are successively tested, always using narrow prior uncertainties for the values of exponents c as recommended for the BaRatin method (cf. Tab. 4). First, ignoring the modeling results, let us assume that the stage-discharge relation is controlled by a low-flow natural riffle, then by a single channel. The obtained rating curve fails in reproducing the observed flow conveyance increase near 3 m correctly (cf. Fig. 7a). Actually, BaRatin computes the transition level, κ, so low that only the channel control is used to fit the gaugings, which is inconsistent with our prior knowledge of the hydraulic controls.

Therefore, focussing on the cross-section of the bridge located right downstream of the pressure gauge, we now assume that the additional flow through the road tunnel is the cause of that inflexion of the stage-discharge relation. The apparently good agreement between the resulting rating curve and the gaugings (cf. Fig. 7b) is misleading. Indeed, while the results for all other parameters appear to be consistent with the priors (cf. Tab. 4), the far too high value obtained for κ_1 (2.8 instead of 1.2) indicates that BaRatin actually uses the second control (main channel through the bridge) to reproduce the trend of the third identified control (overbank flow through the road tunnel). Again, our hydraulic analysis is erroneous since the road tunnel is actually too high and too narrow to be a credible candidate for the observed high-flow control.

Eventually, the 3 controls predicted after consideration of the hydraulic modeling results are used as priors. Fig. 7c shows that the agreement with the gaugings is as good as previously, while the transitions levels are now realistic as regards the geometry of the river reach. It is worth noting that the extrapolated flood discharges are significantly different than with the previous misleading hydraulic priors.

For the three runs of BaRatin, the 95% uncertainty envelopes of the MaxPost rating curve are far narrower than the prior envelope, within which they are included. The only exception is the highest end of the MaxPost rating curve computed with only two hydraulic controls (first run), which intersects the lowest bound of the prior envelope (cf. Fig. 7a). The inappropriate definition of the control matrix therefore leads to obviously wrong results for the first run. However, the second and third runs of this application case show that a good agreement between the MaxPost curve and the gaugings can be misleading, and that checking the hydraulic meaning of the results is absolutely necessary, especially to avoid hydraulically unrealistic extrapolation.
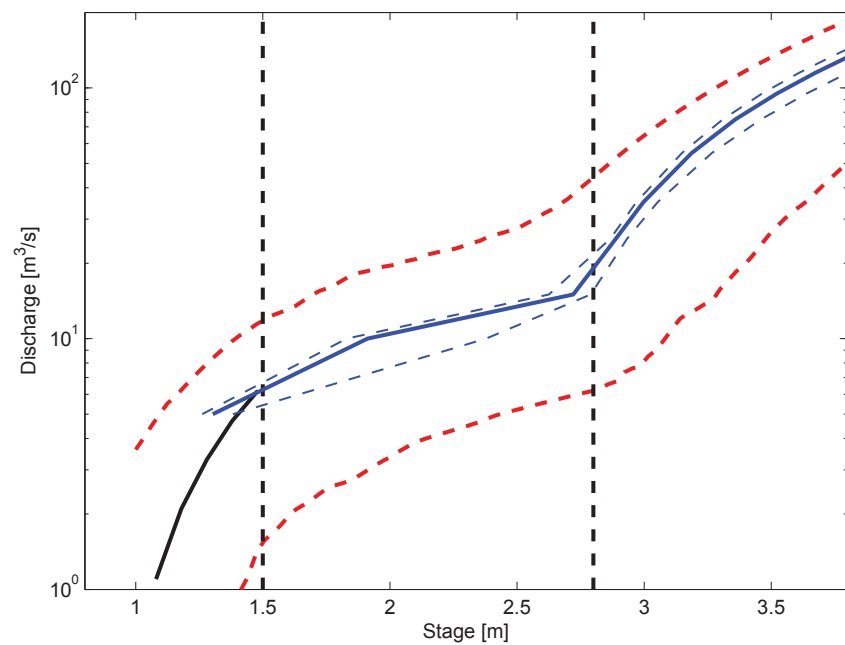
Figure 6: The Sarre river at Sarrebourg: 1-D simulation results with average, rough and smooth flow resistance scenarios (blue solid and dashed lines, respectively) and the low-flow stage-discharge relation computed with a horizontal weir formula (black solid line). The hydraulic priors provided to BaRatin in the third simulation are represented by the 95% uncertainty envelope of the prior rating curve (red dashed lines). The transition stages are shown as vertical black dashed lines.

| | Control 1 | | | | Control 2 | | | | Control 3 | | | |
	a_1	b_1	c_1	κ_1	a_2	b_2	c_2	κ_2	a_3	b_3	c_3	σ_f
2 controls: rectangle then bridge												
Prior	14 (1.5)	0.90 (0.125)	1.50 (0.025)	1.2 (0.25)	35 (5)	-	1.67 (0.025)	-	-	-	-	*
Posterior	14 (1.6)	0.95 (0.013)	1.51 (0.026)	1.0 (0.036)	8.7 (0.26)	0.91 (0.015)	1.65 (0.023)	-	-	-	-	0.68 (0.066)
3 controls: rectangle then bridge and road tunnel												
Prior	14 (1.5)	0.90 (0.125)	1.50 (0.025)	1.2 (0.25)	35 (5)	-	1.67 (0.025)	3 (0.15)	7 (1)	-	1.67 (0.025)	*
Posterior	9.0 (0.15)	0.95 (0.0055)	1.49 (0.019)	2.78 (0.022)	38 (4.0)	2.05 (0.053)	1.67 (0.024)	2.94 (0.12)	6.7 (0.95)	2.94 (0.12)	1.67 (0.025)	0.59 (0.044)
3 controls: rectangle then main channel and floodplain												
Prior	14 (1.5)	0.90 (0.125)	1.50 (0.025)	1.5 (0.25)	3 (2)	-	1.67 (0.025)	2.8 (0.25)	80 (10)	-	1.67 (0.025)	*
Posterior	9.8 (0.27)	0.95 (0.0036)	1.54 (0.022)	1.92 (0.14)	5.3 (0.75)	0.51 (0.18)	1.68 (0.025)	2.73 (0.027)	90 (9.0)	2.73 (0.027)	1.67 (0.025)	0.54 (0.046)

Table 4: The Sarre at Sarrebourg; prior and posterior modes (and standard deviation) for parameters κ, a, b, and c of the hydraulic controls, as used for the BaRatin runs. * a uniform distribution U[0;10,000] is used.
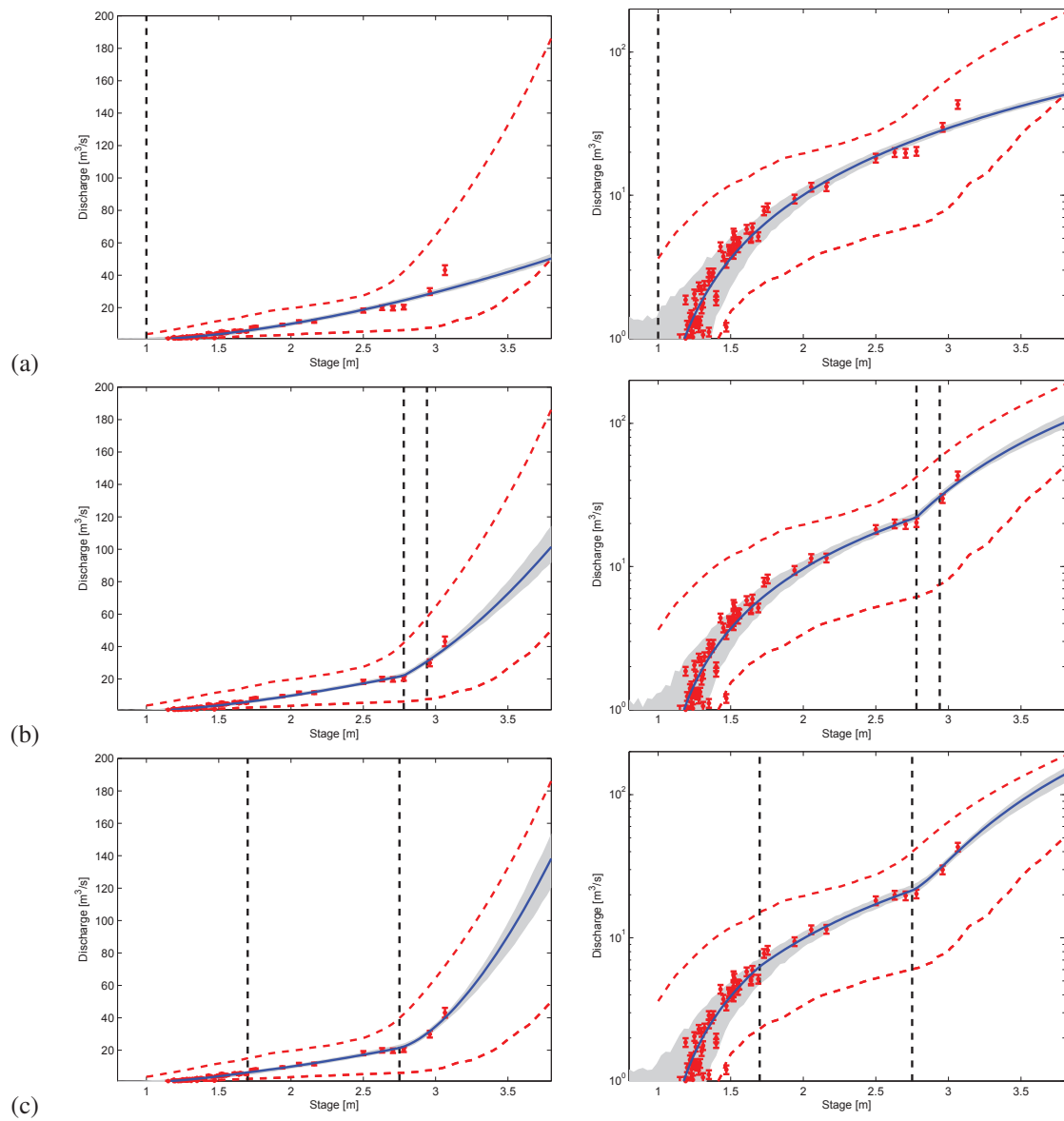
Figure 7: The Sarre at Sarrebourg: gaugings with individual uncertainties (red marks), 95% uncertainty envelope of the priors with appropriate control structure (red dashed lines), MaxPost rating curve (blue solid line) and its 95% uncertainty interval (grey patch), as provided by the BaRatin method for priors with 2 inappropriate controls (a), with 3 inappropriate controls (b), and with 3 appropriate controls (c). The same set of gaugings is used for the three simulations. Vertical dashed lines show the computed transitions between segments. Results are shown with discharge in linear scale (left) and log scale (right).

4. Discussion

Several limitations of the BaRatin framework call for future developments that are discussed in this section.

4.1. Effect of assuming measurement uncertainties are known

The statistical model behind BaRatin assumes that the standard deviation of discharge measurement error u_{Q_i} is known for each individual gauging. This means that this uncertainty has to be assessed by hydrometric services: while this is not yet a generalized practice, this assessment is performed by a growing proportion of hydrometric services. Several operational methods have been proposed (e.g. 26, 3), and some are even implemented in some commercial instruments (e.g. the Sontek ADV FlowTracker). The ability to use those uncertainty estimates in the derivation of the rating curve is a strong motivation for operational services to start or continue quantifying gauging uncertainties.

Notwithstanding these remarks, assuming discharge measurement uncertainties are perfectly known is only an approximation. A possible extension of the model would be to consider them as unknown quantities (with informative priors) to be included in the inference. While this might yield a more realistic quantification of rating curve uncertainties, this approach would also be subject to the following difficulties:

1. It would result in a markedly increased computational complexity, with the inclusion of as many unknowns as there are gaugings.
2. The resulting model would probably be prone to non-identifiability issues, hence still requiring strong priors to ensure the well-posedness of the inference [see 47, for a similar discussion in the context of rainfall-runoff modeling].

Future work will investigate the sensitivity of the BaRatin framework to misspecified discharge measurement uncertainties, and if need be, will further develop the statistical framework. In addition, the interest of considering non-Gaussian measurement errors will also be assessed.

4.2. Effect of assuming error-free stage values

Consider a single-control rating curve equation, $Q(H) = a(H - b)^c$, and let ϵ_H denote a Gaussian stage error with mean 0 and standard deviation σ_H. A first-order approximation of the resulting error in discharge can be obtained as follows:

$$Q(H + \epsilon_H) \approx Q(H) + \frac{\partial Q}{\partial H}(H)\,\epsilon_H \tag{13}$$

$$= Q(H) + ac(H - b)^{c-1}\,\epsilon_H \tag{14}$$

$$= Q(H) + \epsilon_Q \tag{15}$$

where ϵ_Q follows a Gaussian distribution with mean 0 and standard deviation $\sigma_Q = ac(H - b)^{c-1}\sigma_H$. This simple approximation first suggests that the impact of stage errors is larger for insensitive controls (i.e. when $\partial Q/\partial H$ is large). Moreover, it also suggests that the resulting discharge uncertainty depends on the stage.

Ignoring large stage errors is expected to have several types of impact on the inference: (i) it might introduce some bias in rating curve parameters; (ii) it might create compensatory effects through the remnant error term, hence artificially inflating the remnant standard deviation σ_f. A possibility to overcome these drawbacks would be to explicitly recognize stage errors, by considering the stage as an unknown quantity that has to be estimated (with an informative prior representing stage uncertainty). This would be similar to error-in-variables hierarchical approaches introduced in the hydrological modeling literature [e.g. 20]. As for discharge measurement uncertainty, this approach would result in an increased computational complexity, but this increase may be limited: for many gaugings conducted in favorable conditions, stage uncertainty is indeed negligible. This explicit treatment of stage uncertainty could therefore be restricted to gaugings realized in adverse conditions, e.g. during floods or very low flows.

Finally, we reiterate the distinction between errors affecting the stage of gaugings and errors affecting continuously-measured stage values: the latter can easily be taken into account by propagating the corresponding uncertainty to the discharge time series through the estimated rating curve.

21

4.3. Effect of assuming the number of controls is known

The rating curve equation (Eq. 10) involves identifying the number of controls from a purely hydraulic analysis of the gauging station. From our experience in analyzing around 50 stations with BaRatin, this can be achieved in most cases, yielding consistent and acceptable results from a hydrometric point of view. However, there exist cases where this identification is challenging (as illustrated by the Sarre at Sarrebourg case study), with several competing hypotheses being a priori equally sensible.

In such a situation, it might be necessary to work with several rating curve equations in parallel, and evaluate whether gauging data are informative enough to identify the most relevant hypothesis. Alternatively, additional analyses might also be useful (e.g. collecting additional topographic data and/or using hydraulic modeling). An interesting development would be to explicitly recognize the uncertainty related to the number of controls, by implementing a multi-model framework. This could be based on Bayesian Model Averaging [44], or on reversible jump MCMC algorithms [11]. In any case, we believe that letting the number and type of controls free of any hydraulic expertise is the most hazardous option (as illustrated by the last case study, the Sarre at Sarrebourg).

4.4. Improvement of the remnant error model

Experience suggests that the remnant error term leads to uncertainties that are often too large for low flows and too small for high flows (see Section 3.3). This is probably due to its additive, Gaussian and homoscedastic nature (see equation 4). These three hypotheses may be relaxed if necessary. The most straightforward modification is to introduce heteroscedasticity by assuming that the remnant standard deviation is a function of the rating curve discharge, for instance: $\sigma_f(Q) = \gamma_1 + \gamma_2 * Q$, where parameters γ_1 and γ_2 are unknown and are therefore added to the inference list. As an illustration, Fig. 8 shows the results obtained with this heteroscedastic remnant error model for the Charbonnières case study (section 3.3). Comparing homoscedastic and heteroscedastic error models, one can first observe that the MaxPost rating curve remains almost unchanged (blue and red solid lines being hardly discernible in Fig. 8a-b). However, the difference is much more striking in terms of uncertainty envelopes. In high flows (Fig. 8a), the heteroscedastic error model (red) yields a much larger uncertainty than the homoscedastic one (blue, hardly discernible in Fig. 8a), while the opposite holds in low flows (Fig. 8b). The reliability of these two competing error models can be assessed by means of a quantile-quantile plot (QQ-plot) of the standardized residuals. The standardization is performed by dividing the raw residual by its standard deviation $\sqrt{\sigma_f^2 + u_{Q_i}^2}$ (see Eq. 5). In the homoscedastic case, σ_f is constant, while in the heteroscedastic case, it is an affine function of the discharge. The QQ-plot in Fig. 8c confirms that the heteroscedastic treatment is much more realistic, with the red points being closer to the 1:1 line than the blue ones.
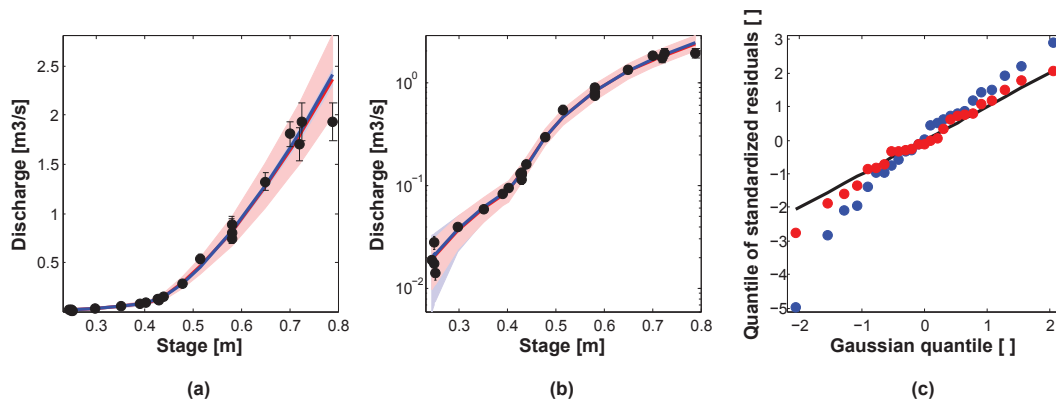

Figure 8: The Charbonnières river at Charbonnières-les-Bains: homoscedastic (blue lines, envelopes and points) vs. heteroscedastic (red) remnant error model. (a) MaxPost rating curves (solid lines), 95% uncertainty intervals (colored patches) and gaugings (points); (b) same as (a) with a logarithm y-axis; (c) Gaussian QQ-plot of the standardized residuals.

22

An additional and more challenging difficulty is related to the nature of remnant errors. Indeed, this error term is *intended* to represent structural errors (e.g. approximation due to a power formulation). However, as illustrated in Section 3.2, it will effectively also compensate for any other misspecification in the model, including ill-specified priors, non-stationarity, ignored stage errors, etc. This may have strong consequences on the handling of the related uncertainty when a stage time series is transformed into a discharge time series using the rating curve. In an ideal case where the remnant error exclusively represents structural uncertainty, then this error should be systematic (and the induced uncertainty will therefore not decrease by temporal averaging of the discharge series). However, if most of the remnant error variance is attributable to underestimated measurement uncertainty, then the remnant error should probably not even be propagated (since the rating curve aims at predicting the true discharge, not the measured one). The practical significance of this challenging difficulty will be evaluated in future work.

4.5. Tackling non-stationary rating curves

A major limitation of BaRatin, in its current version, is related to the assumption of stationarity of the stage-discharge relation. An important perspective is to generalize the method to account for rating shifts and the associated uncertainties. The development of statistical models for sudden, seasonal or episodic rating shifts will require introducing time-varying parameters in the stage-discharge relation, leading to a rating curve that changes over time. As for stationary rating curves, it will be most important to address the non-stationary hydraulic controls on a physical basis. Indeed, detecting and modeling non-stationarity is not a mere statistical problem, but is instead strongly related to the hydraulics governing the gauging sections. As an illustration, if the river bed is eroded after a large flood, the equations of the lower segment(s) should indeed change, but not necessarily the equations of the higher segment(s): having formalized a hydraulics-based framework in the stationary case therefore helps in deciding how the non-stationarity should be modelled (e.g. deciding which parameters of which segments should be made time-dependent, according to which covariates).

5. Conclusions

The BaRatin framework combines Bayesian inference and hydraulic analysis for building steady, multi-segment stage-discharge relations. It is based on three main components:

1. a versatile rating curve equation, based on piecewise compound power functions, that can describe most (though not all) practical stage-discharge relations;
2. the identification of the hydraulic controls and the specification of priors (in particular for hydraulic exponents);
3. the quantification of the individual uncertainty of gaugings.

Such a statistical approach makes the derivation of uncertainty easy and rigorous, unlike with existing purely hydraulic approaches. Meanwhile, the priors and results are always directly related to simple and realistic formulations for hydraulic controls, which is seldom the case in reported Bayesian studies. Hydraulic modeling may be helpful for deriving more accurate priors, which may help in reducing uncertainties for ungauged or poorly gauged high-flow segments of the rating curve. Another important aspect of the method is the attention paid to the variable uncertainty in individual gaugings. This is especially important for making the most of more uncertain flood data based on emerging remote techniques (hand-held radar, image-based techniques, cf. e.g. 7) or on post-flood discharge estimation, using flood marks, hydraulic modeling, etc.

The BaRatin framework was applied to around 50 hydrometric stations, including the 3 contrasted application cases presented in this paper. The application demonstrated the versatility and the reliability of the method once hydraulic controls are correctly identified and gaugings are validated individually with their uncertainty. Already used by some hydrological services in France, BaRatin along with its user-friendly interface BaRatinAGE is released freely on request to the authors.

Beyond the developments reported in the discussion section, the next stage is to make use of the discharge uncertainty provided by BaRatin in hydrological studies. This will require propagating the uncertainty in measured stages, as well as departures from reference hydraulic conditions, to derive the uncertainty in the instantaneous discharge records. Temporal averaging will then yield uncertain discharge time series recorded at any other time step (e.g. daily, monthly).

Taking into account the uncertainties related to the stage-discharge relation in hydrological calculations remains an area of active research. For instance, while most frequency analysis methods are able to take into account sampling uncertainty, taking into account the uncertainties induced by the rating curve is much less common. Solutions have been explored in the literature [see e.g. 21, 22, 42, 40, 32], and could be coupled with the output provided by BaRatin, for instance to estimate flood quantiles taking into account both the sampling uncertainty (extreme paucity of data) and the uncertainty in the data themselves (due to the rating curve). In a similar vein, accounting for rating curve uncertainty in the calibration of hydrologic models is also an important research avenue [28, 52, 48, 51].

Acknowledgements

References

[1] R.T. Clarke. Uncertainty in the estimation of mean annual flood due to rating-curve indefinition. *Journal of Hydrology*, 222 (1-4):185–190, 1999.

[2] R.T. Clarke, E.M. Mendiondo, and L.C. Brusa. Uncertainties in mean discharges from two large South American rivers due to rating curve variability. *Hydrological Sciences Journal/Journal des Sciences Hydrologiques*, 45(2):221–236, 2000.

[3] T.A. Cohn, J.E. Kiang, and R.R. Mason. Estimating discharge measurement uncertainty using the interpolated variance estimator. *Journal of Hydraulic Engineering*, 139(5), 502–510, 2013.

[4] G. Di Baldassarre and A. Montanari. Uncertainty in river discharge observations: a quantitative analysis. *Hydrology and Earth System Sciences*, 13(6):913–921, 2009.

[5] G. Di Baldassarre and P. Claps. A hydraulic study on the applicability of flood rating curves. *Hydrology Research*, 42(1):10–19, 2011.

[6] A. Domeneghetti, A. Castellarin and A. Brath. Assessing rating-curve uncertainty and its effects on hydraulic calibration. *Hydrology and Earth System Sciences*, 16:1191-1202, 2012.

[7] G. Dramais, J. Le Coz, B. Camenen, and A. Hauet. Advantages of a mobile LSPIV method for measuring flood discharges and improving stage-discharge curves. *Journal of Hydro-Environmental Research*, 5(4):301–312, 2011.

[8] J.R. Dymond and R. Christian. Accuracy of discharge determined from a rating curve. *Hydrological Sciences Journal/Journal des Sciences Hydrologiques*, 27(4):493–504, 1982.

[9] H.J. Freestone. Sensitivity of flow measurement to stage errors for New Zealand catchments. *Journal of Hydrology New Zealand*, 22(2):175–181, 1983.

[10] A. Gelman, J.B. Carlin, H.S. Stern, and D.B. Rubin. Bayesian data analysis. Chapman & Hall, 2nd Edition, Texts in Statistical Science, 2004.

[11] P.J. Green. Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika*, 82(4): 711–732, 1995.

[12] J.L. Guerrero, I.K. Westerberg, S. Halldin, C.Y. Xu, and L.C. Lundin. Temporal variability in stage-discharge relationships. *Journal of Hydrology*, 446-447:90–102, 2012.

[13] R.W. Herschy. *Hydrometry, Principles and Practices*. Wiley, 1999.

[14] R.W. Herschy. The uncertainty in a current meter measurement. *Flow Measurement and Instrumentation*, 13:281–284, 2002.

[15] ISO 748 Hydrometry – Measurement of liquid flow in open channels using current-meters or floats. International Organization for Standardization, Geneva, Switzerland, 46 p, 2007.

[16] ISO/TS 25377 (HUG) Hydrometric Uncertainty Guidance (HUG). International Organization for Standardization, Geneva, Switzerland, 2007.

[17] J. Jalbert, T. Mathevet, and A.-C. Favre. Temporal uncertainty estimation of discharges from rating curves using a variographic analysis. *Journal of Hydrology*, 397:83–92, 2011.

[18] JCGM 100:2008 (GUM). Evaluation of measurement data – Guide to the expression of uncertainty in measurement. JCGM member organizations (BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP and OIML), 120 p, 2008.

[19] ISO 1100-2:2010 Hydrometry – Measurement of liquid flow in open channels – Part 2: Determination of the stage-discharge relationship. International Organization for Standardization, Geneva, Switzerland, 28 p., 2010.

[20] D. Kavetski, G. Kuczera and S.W. Franks. Bayesian analysis of input uncertainty in hydrological modeling: 1. Theory *Water Resources Research*, 42, 2006.

[21] G. Kuczera. Correlated rating curve error in flood frequency inference. *Water Resources Research*, 32(7):2119–2127, 1996.

[22] G. Kuczera. Comprehensive at-site flood frequency analysis using Monte Carlo Bayesian inference. *Water Resources Research*, 35(5):1551–1557, 1999.

[23] M. Lang, K. Pobanz, B. Renard, E. Renouf, and E. Sauquet. Extrapolation of rating curves by hydraulic modelling, with application to flood frequency analysis. *Hydrological Sciences Journal/Journal des Sciences Hydrologiques*, 55(6):883–898, 2010. doi: 10.1080/02626667.2010.504186.

[24] J. Le Coz, A. Hauet, G. Dramais, and G. Pierrefeu. Performance of image-based velocimetry (LSPIV) applied to flash-flood discharge measurements in Mediterranean rivers. *Journal of Hydrology*, 394:42–52, 2010. doi: 10.1016/j.jhydrol.2010.05.049.

[25] J. Le Coz, B. Camenen, G. Dramais, M. Ferry, J.-L. Rosique and J. Ribot-Bruno. Contrôle des débits réglementaires *(Inspection of statutory discharges)*. Onema/Cemagref Technical Handbook, 128 p.

[26] J. Le Coz, B. Camenen, X. Peyrard, and G. Dramais. Uncertainty in open-channel discharges measured with the velocity-area method. *Flow Measurement and Instrumentation*, 26:18–29, 2012.

[27] J. Le Coz, K. Pobanz, J.-B. Faure, G. Pierrefeu, B. Blanquart, and Y. Choquette. Stage-discharge hysteresis evidenced by multi-ADCP measurements. *RiverFlow2012 conference*, 1277–1283, 5-7 September 2012, San José, Costa Rica.

[28] H. McMillan, J. Freer, F. Pappenberger, T. Krueger, and M. Clark. Impacts of uncertain river flow data on rainfall-runoff model calibration and discharge predictions. *Hydrological Processes*, 24(10):1270–1284, 2010.

[29] H. McMillan, T. Krueger, and J. Freer. Benchmarking observational uncertainties for hydrology: rainfall, river discharge an water quality. *Hydrological Processes*, 26(26):26, 4078–4111, 2012.

[30] R.A. Moyeed and R. T. Clarke. The use of Bayesian methods for fitting rating curves, with case studies. *Advances in Water Resources*, 28:807–818, 2005.

[31] M. Muste, H.-C. Ho, and D. Kim. Considerations on direct stream flow measurements using video imagery: Outlook and research needs. *Journal of Hydro-environment Research*, 5:289–300, 2011.

[32] L. Neppel, B. Renard, M. Lang, P.A. Ayral, D. Coeur, E. Gaume, N. Jacob, O. Payrastre, K. Pobanz, and F. Vinet. Flood frequency analysis using historical data: accounting for random and systematic errors. *Hydrological Sciences Journal/Journal des Sciences Hydrologiques*, 55(2):192–208, 2010.

[33] A. Olivier, G. Pierrefeu, M. Scotti, and Blanquart B. Incertitudes sur les débits issus des courbes de tarage (uncertainty of discharge measured with relating curve between level and discharge). In *SHF conference, Hydrological measurements and uncertainties*, 2008.

[34] P.M. Pelletier. Uncertainties in the single determination of river discharge: a literature review. *Canadian Journal of Civil Engineering*, 15:834–850, 1988.

[35] A. Petersen-Overleir. Accounting for heteroscedasticity in rating curve estimates. *Journal of Hydrology*, 292(1-4):173–181, 2004.

[36] A. Petersen-Overleir. Modelling stage-discharge relationships affected by hysteresis using the Jones formula and nonlinear regression. *Hydrological Sciences Journal/Journal des Sciences Hydrologiques*, 51(3):365–388, 2006.

[37] A. Petersen-Overleir. Fitting depth-discharge relationships in rivers with floodplains. *Hydrology Research*, 39(5-6):369–384, 2008.

[38] A. Petersen-Overleir and T. Reitan. Objective segmentation in compound rating curves. *Journal of Hydrology*, 311(1-4): 188–201, 2005.

[39] A. Petersen-Overleir and T. Reitan. Bayesian analysis of stage-fall-discharge models for gauging stations affected by variable backwater. *Hydrological Processes*, 23(21):3057–3074, 2009.

[40] A. Petersen-Overleir and T. Reitan. Accounting for rating curve imprecision in flood frequency analysis using likelihood-based methods. *Journal of Hydrology*, 366(1-4):89–100, 2009.

[41] S.E. Rantz. *Measurement and computation of streamflow*, volume 1, Measurement of stage and discharge, Water-Supply Paper 2175. U. S. Geological Survey, Washington, 1982.

[42] D.S. Reis and J. R. Stedinger. Bayesian MCMC flood frequency analysis with historical information. *Journal of Hydrology*, 313(1-2):97–116, 2005.

[43] T. Reitan and A. Petersen-Overleir. Bayesian power-law regression with a location parameter, with applications for construction of discharge rating curves. *Stochastic Environmental Research and Risk Assessment*, 22(3):351–365, 2008.

[44] T. Reitan and A. Petersen-Overleir. Bayesian methods for estimating multi-segment discharge rating curves. *Stochastic Environmental Research and Risk Assessment*, 23(5):627–642, 2009. DOI 10.1007/s00477-008-0248-0.

[45] T. Reitan and A. Petersen-Overleir. Dynamic rating curve assessment in unstable rivers using Ornstein-Uhlenbeck processes. *Water Resources Research*, 47:W02524, 2011. doi: 10.1029/2010WR009504.

[46] B. Renard, V. Garreta, and M. Lang. An application of Bayesian analysis and Markov chain Monte Carlo methods to the estimation of a regional trend in annual maxima. *Water Resources Research*, 42:W12422, 2006. doi: 10.1029/2005WR004591.

[47] B. Renard, D. Kavetski, M. Thyer, G. Kuczera and S.W. Franks. Understanding predictive uncertainty in hydrologic modeling: The challenge of identifying input and structural errors *Water Resources Research*, 46, 2010.

[48] B. Renard, D. Kavetski, E. Leblois, M. Thyer, G. Kuczera and S.W. Franks. Towards a reliable decomposition of predictive uncertainty in hydrological modeling: Characterizing rainfall errors using conditional simulation. *Water Resources Research*, 47, 2011.

[49] A.R. Schmidt. *Analysis of stage-discharge relations for open-channel flows and their associated uncertainties*. PhD thesis, University of Illinois, Urbana-Champaign, 2002.

[50] Y. Shimizu, S. Giri, S. Yamaguchi and J. Nelson. Numerical simulation of dune-flat bed transition and stage-discharge relationship with hysteresis effect. *Water Resources Research*, 45:W04429, 2009. doi: 10.1029/2008WR006830.

[51] A.E. Sikorska, A. Scheidegger, K. Banasik and J. Rieckermann. Considering rating curve uncertainty in water level predictions. *Hydrol. Earth Syst. Sci. Discuss.*, 10:3, 2955–2986, 2013.

25

[52] M. Thyer, B. Renard, D. Kavetski, G. Kuczera, S.W. Franks and S. Srikanthan. Critical evaluation of parameter consistency and predictive uncertainty in hydrological modelling: a case study using Bayesian total error analysis. *Water Resources Research*, 45, 2009.

[53] C. Venetis. A note on the estimation of the parameters in logarithmic stage-discharge relationships with estimate of their error. *Bulletin of the International Association of Scientific Hydrology*, XV(2):105–111, 1970.

[54] I. Westerberg, J.-L. Guerrero, J. Seibert, Beven K.J., and S. Halldin. Stage-discharge uncertainty derived with a non-stationary rating curve in the Choluteca River, Honduras. *Hydrological Processes*, 25:603–613, 2011.

[55] World Meteorological Organization. Manual on Stream Gauging. Volume I – Fieldwork *WMO-No. 1044*, 250 p., 2010.

[56] World Meteorological Organization. Manual on Stream Gauging. Volume II – Computation of discharge *WMO-No. 1044*, 195 p., 2010.