Are historical stage records useful to decrease the uncertainty of flood frequency analysis ? A 200-year long case study

Mathieu LUCAS *[1], Benjamin RENARD[2], Jérôme LE COZ[1], Michel LANG[1], Antoine BARD[3], and Gilles PIERREFEU[4]

[1]INRAE, UR RIVERLY, Villeurbanne, France

[2]INRAE, UR RECOVER, Aix-en-Provence, France

[3]ESDB, Briançon

[4]CNR, Lyon

June 12, 2023

Abstract

Flood frequency analysis (FFA), a widely used method to estimate flood hazard, is affected by several sources of uncertainty. Extending flood samples by reanalyzing historical continuous stage records has the potential to reduce sampling uncertainty, but the historical flood discharges derived from this reanalysis are generally affected by large uncertainties. This paper explores whether historical stage records improve design flood estimates through a chain of uncertainty estimation methods for FFA. Uncertainties are estimated and propagated from stage and rating curves to design flood estimates using Monte Carlo procedures. The role of both streamflow and sampling uncertainties in design flood estimation is examined. This procedure is applied to the 205-year long continuous stage series of the Rhône River at Beaucaire, France (95 590 km^2). The estimated streamflow 95% uncertainty varies from 30% (XIX$^{\text{th}}$ Century) to 5% (1967-2020). The total uncertainty of design flood is significantly reduced when the length of the series increases from 20 to 100 years due to sampling uncertainty reduction. However, the total uncertainty remains stable beyond this sample size: this is because large uncertainties affecting the XIX$^{\text{th}}$ Century flood discharges compensate for the reduction in sampling uncertainty. Enlarging the sample size to two centuries leads to including the two largest known floods in 1840 and 1856. In turn, this induces a 15% increase of the 1000-year flood estimates.

Keywords: Flood frequency analysis, Historical stage records, Uncertainty propagation, Streamflow uncertainty, Sampling uncertainty

*Corresponding author at: INRAE, 5 Rue de la Doua, 69100 VILLEURBANNE, FRANCE.

E-mail adress: mathieu.lucas@inrae.fr (M. Lucas).

1 Introduction

Flood frequency analysis (FFA) is a widely used method to estimate flood hazard. It allows linking the magnitude of a flood to its probability of occurrence (Hamed and Ramachandra Rao (2019); Jain and Singh (2019)). Flood estimates for various exceedance probabilities or, equivalently, return periods, are commonly used for population safety policies, land use planning, as well as industrial safety. The standard FFA approach is to estimate a distribution using a sample of flood peaks, typically defined as annual maximum discharges or discharges over a given threshold. This distribution may be extrapolated to reach the desired flood quantile that typically corresponds to a 100 or 1000-year return period (see Le Delliou, 2014 for dam safety regulations in France).

This FFA approach is affected by various sources of uncertainty. First, the hydrological data used to estimate the FFA distribution is uncertain. Indeed, streamflow time series are generally derived from stage time series through rating curve models (Rantz, 1982). This procedure includes three types of errors on hydrological data: stage measurement; discharge measurement (gaugings); stage-discharge model (rating curve). Moreover, the estimated FFA distribution is also affected by sampling uncertainty resulting from the limited size of available streamflow data (Kjeldsen et al., 2011). Considering the importance of decisions relying on FFA results, a consistent treatment of uncertainty all over the data processing chain (including both streamflow and sampling uncertainties) is essential but is usually not performed.

Streamflow series are affected by several sources of uncertainty as described by McMillan et al. (2012). First, a large number of stage error sources are identified in the literature (Van Der Made (1982); Petersen-Øverleir and Reitan (2005); McMillan et al. (2012); Horner et al. (2018)), such as staff gauge reading, levelling of the staff gauge, or stage sensor calibration. The frequency of measurement may also induce time interpolation errors. With modern automatic gauges, stage is generally measured with a time step small enough (e.g. between 15 min and 1 hour) to get negligible time interpolation errors. However, before the rise of automatic gauges, measurements were made by operators who read the staff gauge less frequently (e.g. once or a few times per day), thus possibly missing the flood peak. This issue is particularly critical when old stage series are used. Hamilton and Moore (2012) and Kuentz et al. (2014) estimated the measurement frequency error by sub-sampling recent, sub-hourly measurements. They calculated the difference between the variable of interest (such as the daily maximum stage) derived from scarce data, and the same variable derived from high-frequency measurements. Kuentz et al. (2014) applied the (monthly-averaged) calculated bias to correct old stage series. This correction aimed at taking into account the error due to the daily variability caused by snow melt. However, this type of correction has never been applied to peak stage correction during floods, especially in the case of long stage series.

Rating curve uncertainty is also a major issue when dealing with streamflow series. Transforming stage into discharge requires calibration data (gaugings) to establish the stage-discharge relationship. Note that the term "gauging" corresponds in this paper (and in the French practice) to the sporadic measurement of the stage-discharge couple, which can correspond to the term "discharge measurement" in the British or American practice. Gaugings uncertainty depends on the measurement method (Le Coz et al. (2014a); Puechberty et al. (2017)). Moreover, the rating curve is also affected by uncertainties coming from the imperfection of the chosen model to represent the actual hydraulic configuration, and from parameter estimation. Many methods have been proposed to quantify these uncertainties (Petersen-Øverleir et al. (2009); Juston et al. (2014); Le Coz et al. (2014b); Morlot et al. (2014); Coxon et al. (2015); McMillan and Westerberg (2015); Mansanarez et al. (2019b)). A comparison of several of these methods has been recently proposed by Kiang et al. (2018). Another important issue affecting streamflow data accuracy is rating changes. The stage-discharge relationship is frequently affected by changes caused by various factors, either natural or anthropic, for instance: bed geometry evolution during floods or river works, aquatic vegetation growth and decay, ice cover... A regular monitoring through gaugings is essential to detect those changes (Ibbitt and Pearson, 1987) that can be transient or sudden. Several methods have been proposed to deal with rating changes: estimating rating curves on moving temporal windows (Westerberg et al. (2011); Guerrero et al. (2012)), computing as many rating curves as there are gaugings (Morlot et al., 2014), exploring changes in the annual minimum stages (Łapuszek and Lenar-Matyas, 2015), selecting the 0.5-year return period discharge as a threshold for rating changes (McMillan et al., 2010). More recently, Darienzo et al. (2021) proposed a method based on a recursive segmentation procedure, accounting for both gaugings and rating curve uncertainties. This method has a particular interest when dealing with old and uncertain gaugings. Following the detection of rating shifts, rating curves should be estimated for each stable period. This task may not be straightforward, as the number of gaugings available within a stable period is not always sufficient to properly estimate the stage-discharge relationship for the whole discharge range. A common way to address the lack of gaugings (in particular flood gaugings) within some of the stable periods is to use gaugings from the other stable periods (McMillan et al., 2012; Puechberty et al., 2017). Mansanarez et al. (2019a) proposed an alternative approach to deal with this issue. They developed a stage-period-discharge (SPD) model where rating curve parameters may vary across periods, while others are supposed constant. This method has the advantage of transferring information between periods to improve the rating curve estimation, even when few gaugings are available.

Estimating sampling uncertainty in FFA is a well-established approach. Whatever the chosen distribution and estimation method, standard statistical procedures are available (Coles, 2001). However, these standard procedures only quantify sampling uncertainty, they do not consider data uncertainty. The literature review proposed in the previous paragraphs shows that methods for quantifying individual sources of uncertainty (stage, rating curve and FFA distribution estimation) are available. However, the way these multiple uncertainties propagate through the FFA analysis chain has been less thoroughly studied. A few solutions have emerged to propagate uncertainties in stage time series through uncertain rating curves (Dymond and Christian (1982); Herschy (1998); Petersen-Øverleir and Reitan (2005)), but they assume independent stage errors and therefore neglect systematic errors. Horner et al. (2018) proposed a method for the propagation of both sources of stage uncertainty through uncertain rating curves. Therefore, it is possible to distinguish the effects of independent and systematic stage errors on streamflow uncertainty. Petersen-Øverleir and Reitan (2009), Steinbakk et al. (2016), and Vieira et al. (2022) performed an integrated analysis in which both rating curve parameters and flood frequency distribution are estimated. These studies highlighted the importance of considering rating curve uncertainty for design flood estimations and concluded that, under some conditions, accounting for rating curve uncertainty may notably widen the uncertainty intervals around flood quantiles. However, these studies did not consider stage measurement and time interpolation errors nor rating changes, which may constitute a major source of uncertainty for streamflow data. This is particularly the case when dealing with long streamflow series for which stage uncertainty is large and variable through time, and rating changes may have been missed. Their consideration in a flood frequency framework is therefore not avoidable.

The following questions will be considered in this paper:

1. How to make the most of historical hydrometric (based on stage records) data in flood frequency analysis while accounting for multiple and variable uncertainties at each step of the procedure ?

2. What is the contribution of each source of uncertainty to the flood quantile uncertainty when historical data are taken into account ?

3. To what extent does enlarging streamflow samples by adding increasingly uncertain historical data improve flood quantiles estimation ? How are the relative contributions of sampling and streamflow uncertainties evolving with sample size ?

Note that the "historical data" term used in this paper refers to the use of old but continuously measured stage series, as opposed to sporadic flood marks, prior to regular stage measurements.

This paper illustrates the chained application of methods to quantify and propagate uncertainty from stage records (and their limited time resolution) and stage-discharge rating curves to the estimation of flood distribution with uncertainties. While most of these methods already exist, a key novelty of this work is their combination in a consistent framework (Figure 1) to provide an end-to-end evaluation of the uncertainty affecting FFA estimates. An original method to quantify the stage uncertainty stemming from infrequent readings is also proposed.

The paper is organized as follows. First, the methodology for establishing uncertain streamflow series in a century-long context is presented. It goes through the detection of rating shifts (section 2.1), the estimation of rating curves (section 2.2), and the estimation (section 2.3) and propagation (section 2.4) of stage errors. Then, an approach to propagate streamflow uncertainty through the estimation of extreme flood quantiles is proposed (section 2.5). This procedure is applied to the Beaucaire gauge on the Rhône River (section 3), which previous official FFA only used a 80-year long discharge series (Rigaudière et al., 2000). In France, the official design flood for flood risk mapping is based on the "largest known flood" or the Q100 flood (AEP=0.01) if the latter is greater. Uncertainty is not taken into account in the official rules. The recent works of Pichard et al. (2017) and Bard and Lang (2018) provided a continuous stage series from 1816 to the present time, which makes it the ideal case study for demonstrating this procedure. The results of this application are presented in section 4, and they are discussed in section 5, where avenues for improvements are proposed.

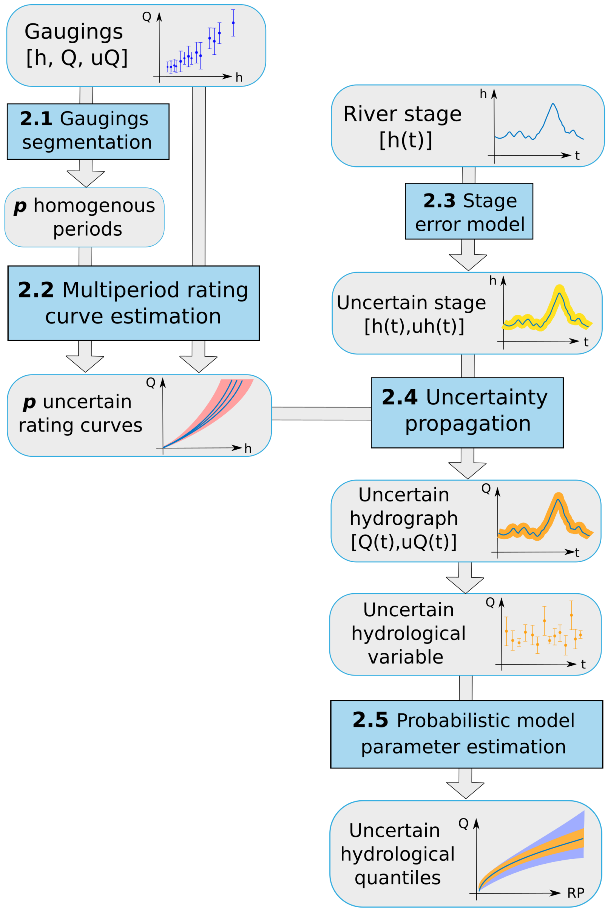


Figure 1: Block diagram of the uncertainty propagation procedure. Grey blocks represent data, blue blocks stand for analysis methods/models that correspond to the sub-sections of this article. h is the water stage, uh the stage uncertainty, Q the discharge, uQ the discharge uncertainty, t is the time and RP the return period of flood quantiles.

2 Uncertainty propagation chain for flood frequency analysis

2.1 Rating shifts detection

The stage-discharge relationship is sensitive to sudden changes caused by morphogenic floods or other causes affecting the flow characteristics. Relying on residuals between the gaugings and the rating curve is the most common approach to monitor the stability of this relationship over time. The method proposed by Darienzo et al. (2021) is used in this work and can be summarized as follows. First, a baseline rating curve is estimated from the whole gaugings dataset. The residuals between gaugings and the rating curve are determined, and a statistical segmentation procedure is applied to them. This procedure accounts for the residuals uncertainty, coming from both the gaugings uncertainty and the rating curve uncertainty. The optimal number of stable sub-periods is determined based on the Bayesian Information criterion (BIC). Then, the same steps are applied recursively to each sub-period. The recursive procedure is stopped when the BIC indicates that a single period is optimal for all sub-periods. The results are not only the dates of the rating shifts but the posterior probability density functions (pdf) of change point times. This allows affecting the shift time to the time of the maximum stage included in the posterior 95% credibility interval. Prior knowledge is provided on the mean of the residuals in each sub-period. The maximum number of segments at each iteration also needs to be specified. All technical details can be found in Darienzo et al. (2021).

2.2 Multi-period rating curves estimation: stage-period-discharge model

Once the stable periods have been identified, the next step is to estimate the rating curves. Mansanarez et al. (2019a) developed a stage-period-discharge (SPD) model "based on the physical interpretation of changes in the stage-discharge relation across a series of stability periods". The SPD model is based on the BaRatin model (Le Coz et al., 2014b).

BaRatin uses the Bayesian paradigm to estimate the parameters of the rating curve equation, a combination of power equations: $Q = a(h-b)^c$, where Q is the discharge, h is the stage, b is an offset (corresponding to the cease-to-flow stage), and a and c are the coefficient and the exponent of the power function. The rating curve equation is deduced from a hydraulic analysis of the gauging station, aimed at identifying the main hydraulic controls governing the stage-discharge relation. The multiple controls can be activated successively or simultaneously. Bayesian inference allows deriving the posterior distribution of rating curve parameters by combining hydraulic information (priors for parameters of each hydraulic control) and information from gaugings with uncertainty (likelihood). Two sources of uncertainty are associated with the estimated rating curve. Parametric uncertainty reflects the uncertainty due to the rating curve parameters estimation because of the limited amount of gaugings and the gaugings uncertainty. Remnant uncertainty comes from the imperfection of the chosen rating curve model to represent the actual hydraulic configuration. The posterior distribution is explored using a Markov Chain Monte Carlo (MCMC) sampler, leading to m realizations of the rating curve parameter vector representing parametric uncertainty. We refer the reader to Le Coz et al. (2014b) for a more thorough description.

The SPD model estimates the rating curves of each stable period based on the same principle by considering that some parameters vary in time, while others remain constant throughout the stable periods. An important step is the identification of those varying parameters based on an hydraulic analysis of the site. Generally, channel depths and/or widths are suspected to change. A distinction is made between "local changes" affecting the lowest control only (for instance the movement of the controlling riffle) and "overall changes" affecting several controls at the same time (for instance, the scouring or filling of the main channel, affecting the offsets of both the low-flow controlling riffle and the main channel itself). Prior specification for varying parameters can be based on the analysis of the yearly lowest stages, which provide information on the evolution of riverbed elevation, as described by Łapuszek and Lenar-Matyas (2015). See Mansanarez et al. (2019a) for a detailed description of prior specification for time-varying rating curves. Specific investigation is needed to correctly account for past flood surveys of the cross section, as complex filling/erosion process may have been encountered during a flood with morphogenic changes.

2.3 Stage uncertainties

Many sources of error having distinct statistical properties can affect stage measurements, as described in Horner et al. (2018). Five different sources of error ($\delta_{1,\dots,5}$) affecting stage measurements are considered. Let $h(t)$ be the measured maximum stage of a day t. The unknown true maximum stage $\hbar(t)$ is assumed to be approximated by the following equation:

$$\hbar(t) = h(t) + \delta_1(t) + \delta_2(t) + \delta_3(t) + \delta_4(t) + \delta_5(t) \tag{1}$$

Staff gauge reading errors $\delta_1 \sim \mathcal{N}(0, \sigma_1)$ originate from operators reading the gauge, where σ_1 depends on the resolution of the graduations (usually 1 cm), and can be increased by waves, especially during floods (McMillan et al., 2012).

Nowadays, most stage measurements are done with automatic sensors of various types such as pressure sensors, floats, radars, and they require a calibration to link the water stage to the measured proxy (respectively the pressure of the water column, the height of a float, or the air draught). Two types of errors arise from this process: sensor errors $\delta_2 \sim \mathcal{N}(0, \sigma_2)$, where σ_2 is usually estimated by the sensor manufacturer, and sensor calibration errors $\delta_3 \sim \mathcal{N}(0, \sigma_3)$ that are related to the corrections made by operators when comparing the stage measured by the sensor to the actual stage at the staff gauge reference. An operator error at this step could affect the stage measurement until the next calibration. Sensor calibration error δ_3 is hence assumed constant between two calibrations and can be represented by drawing a new random value at each operator intervention.

Datum errors $\delta_4 \sim \mathcal{N}(0, \sigma_4)$ are related to changes in the datum reference elevation of the staff gauge zero value and possible discontinuity between successive gauges. Similarly to δ_3, this error is constant between two gauge changes or datum reference measurements.

Measurement frequency errors δ_5 are related to the inadequacy of the frequency of measurement with respect to the rate of stage variations, leading to the true daily maximum occurring in between measurements. Unlike other types of stage errors, this error is hence necessarily positive, which calls for using a positive distribution such as the Exponential distribution. The parameters of this distribution can be estimated with data from the recent period, by analyzing the difference between the daily maximum stage derived from the high-frequency sensor measurement and that from an infrequent fixed-time reading. Note that the frequency errors for hourly (or more frequent) measurements are considered negligible for large rivers with slow variations, such as the Rhône River at Beaucaire.

To sum up, δ_1, δ_2 and δ_5 errors are drawn at each measurement time step, while δ_3 and δ_4 errors are only drawn at specific calibration times. Errors δ_1 to δ_4 are assumed Gaussian with known standard deviations, while δ_5 is assumed Exponential with parameter estimated by subsampling recent measurements. For each error type, 500 realizations are drawn from their respective distribution. Applying eq. 1, the total stage uncertainty is therefore represented by 500 possible realizations of the stage $\hbar(t)$.

2.4 Propagation of stage and rating curve uncertainties to streamflow time series

Stage realizations can be propagated through uncertain rating curves, following the approach described by Horner et al. (2018). Four cases are considered to estimate the contributions of the different sources of streamflow uncertainty:

- **Case 1: Maxpost streamflow.** Stage is taken as the median of the stage time series realizations. This unique stage time series is propagated through the maxpost (Maximum A Posteriori: obtained with parameters maximizing the posterior pdf) rating curve, resulting in a single discharge time series.
- **Case 2: Stage uncertainty.** The n=500 possible stage time series are propagated through the maxpost rating curve. Thus, n discharge time series are obtained.
- **Case 3: Stage and parametric rating curve uncertainty.** The n=500 realizations of stage time series are propagated through m rating curves, corresponding to the m MCMC-simulated parameter vectors described in section 2.2. This leads to $n \times m$ discharge time series.
- **Case 4: Total streamflow uncertainty.** It is obtained by adding remnant rating curve uncertainty (as defined in section 2.2) to case 3. To achieve this, $n \times m$ time series of remnant errors are sampled from their estimated distribution and added to the time series created for case 3.

2.5 Estimation of probabilistic model parameters and flood frequency analysis

The Generalized Extreme Value (GEV) distribution is commonly used to model annual maximum discharges (AMAX) (see Hamed and Ramachandra Rao (2019) or Jain and Singh (2019)). The vector $\boldsymbol{\theta} = (\mu, \sigma, \xi)$ denotes the location, scale and shape parameters of the GEV distribution. The parameters can be estimated based on an independent and identically distributed (iid) sample of j annual maximum discharges $(q_t)_{t=1,\ldots,j}$. Bayesian-MCMC estimation is used in this work, as described in Coles (2001). The posterior distribution quantifies sampling uncertainty and can be represented by r MCMC-generated GEV parameter vectors $\mathbf{\Theta} = (\boldsymbol{\theta}_1, ..., \boldsymbol{\theta}_r)$. The maxpost vector is noted $\hat{\boldsymbol{\theta}}$.

As described in section 2.4, total streamflow uncertainty is represented by $n \times m$ possible realizations of the streamflow, hence of the AMAX series $(q_t^{(i)})_{i=1,...,n\times m;\, t=1,...,j}$, that are subsampled to s=500 realizations to reduce computation time. The estimated flood quantiles should consider both sampling and streamflow uncertainties. Similarly to Steinbakk et al. (2016), the aim is to estimate the contribution of each source to the total uncertainty. For this purpose, three cases can be considered:

- **Case 1: Maxpost quantiles.** The GEV distribution is estimated using the single AMAX series $\hat{\mathbf{q}} = (\hat{q}_t)_{t=1,...,j}$ derived from the maxpost streamflow series (case 1 in section 2.4). Flood quantiles are then computed using the maxpost GEV parameters $\hat{\boldsymbol{\theta}} = (\hat{\mu}, \hat{\sigma}, \hat{\xi})$. In this case, both streamflow and sampling uncertainties are ignored.
- **Case 2: Streamflow uncertainty**. The GEV distribution is estimated for each possible AMAX realization: $\mathbf{q}^{(i)} = (q_t^{(i)})_{t=1,...,j;\, i=1,...,s}$. However, only the maxpost GEV parameters vector is retained for each realization. This results in s vectors of GEV parameters $(\hat{\boldsymbol{\theta}}^{(i)})_{i=1,...,s}$ that represent the effect of streamflow uncertainty of flood quantiles, ignoring sampling uncertainty.
- **Case 3: Total uncertainty.** Similarly to Case 2, the GEV distribution is estimated for each of the s realizations of the AMAX series, but all the r MCMC-simulated GEV parameters are used, leading to $s \times r$ vectors of GEV parameters $(\boldsymbol{\theta}_k^{(i)})_{k=1,...,r;\, i=1,...,s}$. The result thus reflects both sampling and streamflow uncertainties.

3 Case study: The Rhône River at Beaucaire

3.1 Site

The Rhône River at Beaucaire (95 590 km^2) is the lowest gauge of the Rhône River (Figure 2). It captures all the complexity of the Rhône River hydrological regime, from the Alpine area to the oceanic and Mediterranean influences. The annual mean discharge is around 1700 m^3/s (Bard and Lang, 2018), and the maximum known discharge reached 12 500 m^3/s (May 1856, Lang and Coeur (2014)). The station lies in a flood sensitive area, as illustrated by the recent 2003 flood, resulting in 1.1 billion euros worth of damage (Lang and Coeur, 2014). The first stage measurements started in 1816, close to the bridge linking the cities of Beaucaire and Tarascon. This station is named "Pont de Beaucaire" (Kilometric point 267.6 from Lyon). It has been used until the construction of the Vallabrègues hydroelectric scheme in 1967, which led to the derivation of a part of the discharge. Consequently, a new gauging station was installed 2 km downstream from the original one, downstream from the restitution of the derivated discharges. This station, logically named "Beaucaire Restitution" (Kilometric point 269.5), has been used ever since. This resulted in a data gap during the construction process between 1967 and 1970.

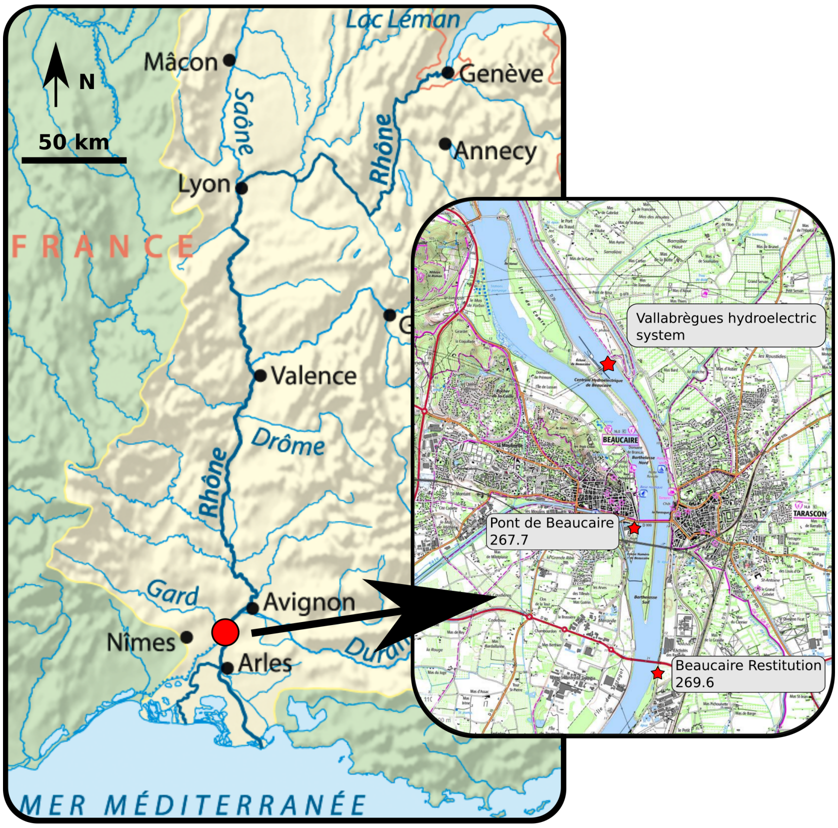


Figure 2: The French Rhône River catchment and Beaucaire gauging stations (from www.geoportail.gouv.fr and www.openstreetmap.org)

3.2 Rating curves

Many gauges of the Rhône River are subject to the effect of variable backwaters caused by the proximity of a dam, and therefore require the use of a stage-fall-discharge (SFD) rating curve model (for example, Valence gauge, 140 km upstream from Beaucaire described by Mansanarez, 2016). Beaucaire is located within a narrowing of the floodplain and there is no dam downstream from the gauge. However, a backwater effect from the sea has been observed at Beaucaire Restitution, but it only affects the very low flows. As this article focuses on floods, we assume that there is no reason to use a SFD model here. Consequently, like the gauge operator (CNR), a stage-discharge (SD) model is used for both Pont de Beaucaire and Beaucaire Restitution gauges.

3.2.1 Pont de Beaucaire

At Pont de Beaucaire, the stage-discharge relationship can be approximated by two additive channel controls: a main channel and a floodway. Thus, the rating curve equation can be written as follows:

$$Q(h) = \begin{cases} a_1(h-b_1)^{c_1}, & \text{if } \kappa_1 < h \leq \kappa_2 \text{ (main channel)} \\ a_1(h-b_1)^{c_1} + a_2(h-b_2)^{c_2}, & \text{if } h > \kappa_2 \text{ (main channel + floodway)} \end{cases} \tag{2}$$

Within the main channel (when water stage is below $\kappa_2 \approx 2$ m), the flow is splitted in two sub-channels (figure 3b) since time immemorial (at least before 1816) as described by Armand (1907). The mobile sandbars separating the flow were progressively fixed by dikes during the XIX$^{\text{th}}$ Century to ease the navigation (figure 3a). These sub-channels are connected upstream and downstream from the gauge location, thus they can be modelled as a single main channel whose average width (≈ 300 m) is the sum of the two sub-channels widths (figure 3b). When stage exceeds κ_2, water starts flowing on the sandbars between the two subchannels. At the gauge location, the total width is limited by unsubmersible levees, but a floodway is also activated a few hundred meters downstream from the station, impacting the stage-discharge relationship at the gauge. The width of this floodway is around 500 m.

The prior distributions of the rating curve parameters are specified using historical material retrieved in regional archives, as described in table 1. Physical parameters that have a direct hydraulic meaning are expressed in the first three lines: channel width (B), slope (S) and Strickler coefficient (K). The resulting prior distribution for the inferred parameter $a = KB\sqrt{S}$ is deduced by Monte Carlo propagation. Log-normal priors are used for positive quantities such as slopes, channel widths and Strickler coefficients. Informative but imprecise priors are assigned to parameters such as channel widths, slopes or offsets which can be difficult to estimate precisely. For c exponents, very precise priors are used because they depend on the control type and shape (here $c = 5/3$ for wide rectangular channel controls based on the simplified Manning-Strickler equation as described by Le Coz et al. (2014b)). Structural uncertainty parameters have uninformative priors.

According to historical profiles and cross-sections, we assume that changes affecting main channel and floodway controls may have occurred due to major floods (in particular 1840, 1856 and 1935 floods) and that channel widths remained constant. Those changes are called "overall changes" and are supposed to affect both main channel and floodway offsets (b_1 and b_2) at the same time. Meanwhile, we assume that local changes due to dike works or sediment depositions from small floods affected the offset (b_1) of the main channel only. As described by Mansanarez et al. (2019a), local and overall changes $\Delta_l^{(k)}$ and $\Delta_g^{(k)}$ affect the offsets of two consecutive periods ($(k-1)$ and k) as follows:

$$\begin{cases} b_1^{(k)} = b_1^{(k-1)} - (\Delta_g^{(k)} + \Delta_l^{(k)}), & \text{(incremental changes in the main channel)} \\ b_2^{(k)} = b_2^{(k-1)} - \Delta_g^{(k)}, & \text{(incremental changes in the floodway)} \end{cases} \tag{3}$$

As the most recent period obtained by gaugings segmentation is assumed to be the most accurately known, it is used as the reference period ($k = 1$) and periods are numbered backward in time. Prior distributions of offset changes are determined in section 3.2.3.

3.2.2 Beaucaire Restitution

Beaucaire Restitution station has a quite stable profile according to 1974-2016 cross-sections (figure 3c left), but the stage-discharge relationship is known to be influenced by the Mediterranean Sea level variations for very low flows (this influence does not apply to the Pont de Beaucaire gauge, located 2 kilometers upstream). This backwater effect can be represented by a channel control with a slope smaller than the slope of the uniform flow, i.e. the mean slope of the channel. The first control (representing low flows influenced by the sea) therefore has the same geometry as the second control (the main channel), but a smaller slope. The main channel control is not influenced by the sea and its slope is close to the longitudinal river slope.

At the gauge location, the 12 meters high banks prevent overbank flows (figure 3c left). However, overbank flows occur further downstream on the left bank, for stages higher than approximately 8 m (figure 3c right). A floodway control (additive to the main channel) is activated above $\approx$ 8 m to model those overbank flows. Therefore, the rating curve equation can be written as follows:

$$Q(h) = \begin{cases} a_1(h - b_1)^{c_1}, & \text{if } \kappa_1 < h \leq \kappa_2 \text{ (main channel, sea-influenced)} \\ a_2(h - b_2)^{c_2}, & \text{if } \kappa_2 < h \leq \kappa_3 \text{ (main channel, non-influenced)} \\ a_2(h - b_2)^{c_2} + a_3(h - b_3)^{c_3} & \text{if } h > \kappa_3 \text{ (main channel + floodway)} \end{cases} \tag{4}$$

Prior distributions of rating curve parameters are specified using recent maps and cross-sections. Priors of the influenced and non-influenced main channels offsets b_1 and b_2 are assumed Gaussian with mean the riverbed elevation that is approximately equal to -5 m. Those offsets b_1 and b_2 are assumed changing in parallel (same local changes as the controls are in the same channel) due to a bed erosion trend described in section 3.2.3, whereas floodway offset b_3 and channel widths are supposed constant because of fixed dikes. These "local changes" $\Delta_l^{(k)}$ are computed backwards in time as follows:

$$\begin{cases} b_1^{(k)} = b_1^{(k-1)} - \Delta_l^{(k)}, & \text{(incremental changes in the main channel)} \\ b_2^{(k)} = b_2^{(k-1)} - \Delta_l^{(k)}, & \text{(incremental changes in the floodway)} \end{cases} \tag{5}$$

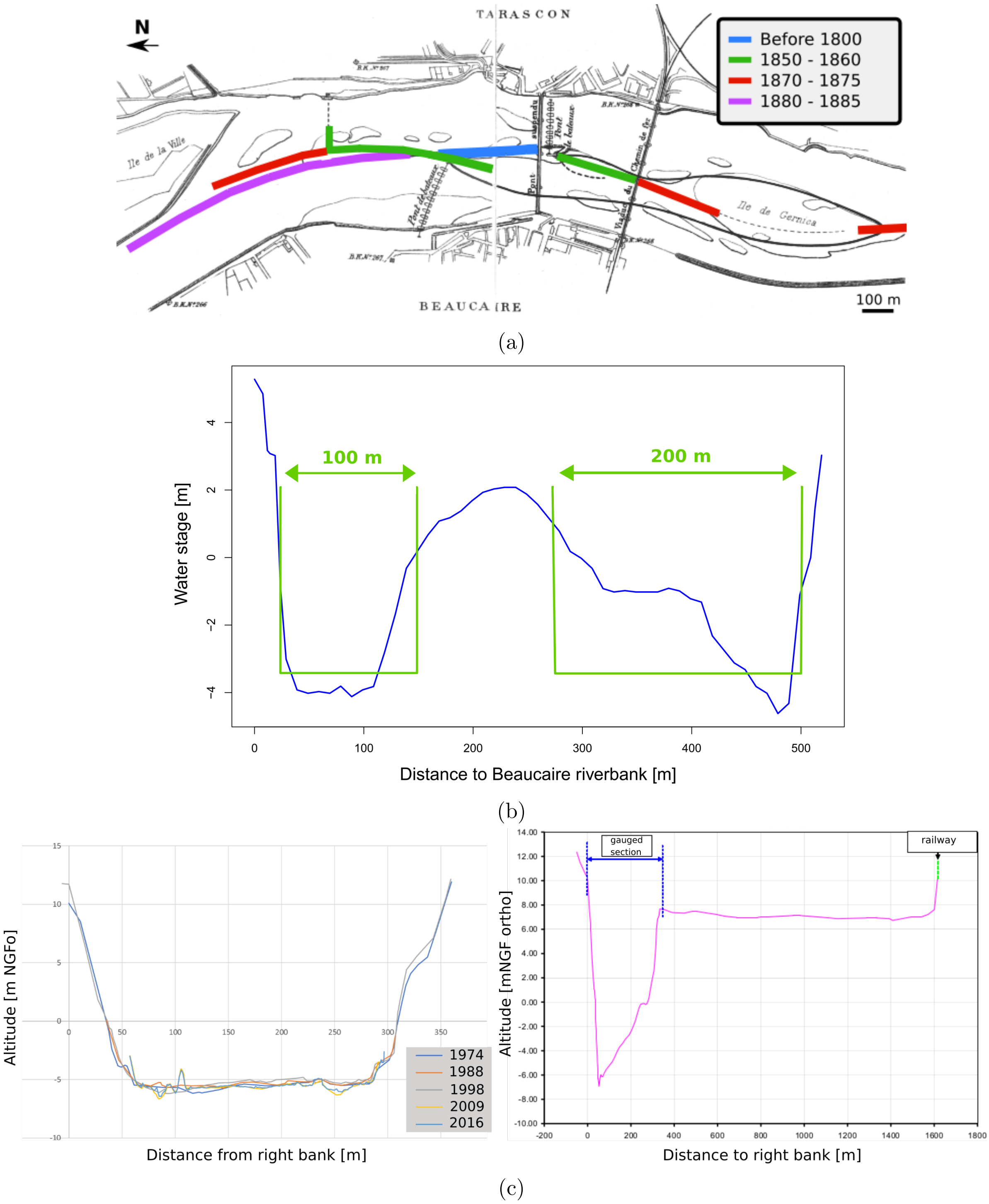


Figure 3: Historical geometry of the Rhône river near Beaucaire: (a) Map of dike evolution between 18th and 20th centuries, adapted from Armand (1907); (b) Approximation of the two subchannels composing the main channel control, based on a 1845 cross-section survey; (c) Profiles from 1974 to 2016 (left) at Beaucaire Restitution station and 2.5 km downstream from the station (right) from CNR data, translated from Bard and Lang (2018) and MEDD (2005)

Priors of incremental bed elevation changes are determined in section 3.2.3.

3.2.3 Prior estimation of bed changes

It is possible to follow the evolution of riverbed elevation through the evolution of yearly lowest stages. Here, the 5% annual stage quantile is considered (figure 4). At Pont de Beaucaire (1816 - 1967), the 5% quantile is oscillating with a 0.3 m standard deviation. Those variations do not seem to be related to the occurrence of major floods. Without more precise information, we assume that prior distributions of local and overall offset changes defined in section 3.2 are Gaussian with mean zero and standard deviation 0.3 (table 1).

At Beaucaire Restitution (1970-2020), the annual 5% quantile shows a large decrease during the first 4 years (more than 1 m). This is a consequence of Vallabrègues hydraulic works between 1967 and 1970 as well as substantial dredgings. A geomorphic adjustment after the works in the channel may have affected the riverbed level as well. After the first years, the channel bottom stabilized, however with a slight scouring trend of about 30 cm in 40 years. The standard deviation of the 5% quantiles reaches 0.5 m. Those bed elevation changes affect both sea-influenced and non-influenced main channel controls offsets. Therefore, the prior distribution of local changes is assumed Gaussian, with mean zero and standard deviation 0.8 m, which is larger than 0.5 m to be more representative of the large changes that occurred during the first years (table 2).

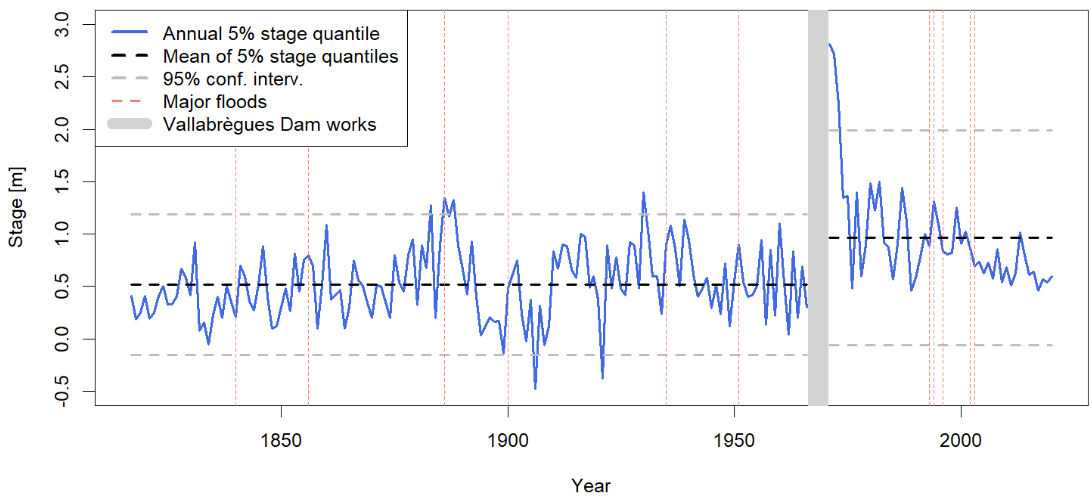


Figure 4: Time series of annual 5% stage quantile at both Pont de Beaucaire (1816-1967) and Beaucaire Restitution (1970-2020) stations.

Physical param.	Meaning	Prior	Inferred param.	Prior
Control 1: main channel				
$b_1[m]$	Offset	$\mathcal{N}(-4, 0.5)$	$b_1[m]$	$\mathcal{N}(-4, 0.5)$
$B_1[m]$	Channel width	$\mathcal{LN}(ln(300), 0.16)$	$a_1[m^{3/2}/s]$	$\mathcal{LN}(ln(128.6), 1.8.10^{-2})$
$K_1[m^{1/3}/s]$	Strickler coeff.	$\mathcal{LN}(ln(35), 0.14)$		
$S_1[m/m]$	Bed slope	$\mathcal{LN}(ln(1.5.10^{-4}), 0.55)$		
$c_1[-]$	Exponent	$\mathcal{N}(5/3, 0.025)$	$c_1[-]$	$\mathcal{N}(5/3, 0.025)$
Control 2: floodway				
$b_2[m]$	Offset	$\mathcal{N}(1.5, 0.5)$	$b_2[m]$	$\mathcal{N}(2, 0.5)$
$B_2[m]$	Channel width	$\mathcal{LN}(ln(500), 0.1)$	$a_2[m^{3/2}/s]$	$\mathcal{LN}(ln(241.9), 1.10^{-2})$
$K_2[m^{1/3}/s]$	Strickler coeff.	$\mathcal{LN}(ln(30), 0.16)$		
$S_2[m/m]$	Bed slope	$\mathcal{LN}(ln(2.6.10^{-4}), 0.34)$		
$c_2[-]$	Exponent	$\mathcal{N}(5/3, 0.025)$	$c_2[-]$	$\mathcal{N}(5/3, 0.025)$
Structural uncertainty parameters				
$\gamma_1[m^3/s]$	Intercept	$\mathcal{U}(0, 1000)$	$\gamma_1[m^3/s]$	$\mathcal{U}(0, 1000)$
$\gamma_2[-]$	Slope	$\mathcal{U}(0, 100)$	$\gamma_2[-]$	$\mathcal{U}(0, 100)$
Multiperiod RC parameters				
$\Delta l[m]$	Local change	$\mathcal{N}(0, 0.3)$	$\Delta l[m]$	$\mathcal{N}(0, 0.3)$
$\Delta g[m]$	overall change	$\mathcal{N}(0, 0.3)$	$\Delta g[m]$	$\mathcal{N}(0, 0.3)$

Table 1: Priors elicitation for Pont de Beaucaire rating curves. $\mathcal{U}(a, b)$ stands for continuous uniform distribution with bounds a and b, $\mathcal{N}(\mu, \sigma)$ for Normal distribution with mean μ and standard deviation σ, and $\mathcal{LN}(\mu, \sigma)$ for Log Normal distribution with mean μ and standard deviation σ.

Physical param.	Meaning	Prior	Inferred param.	Prior
Control 1: Low flows sea-influenced channel				
$b_1[m]$	Offset	$\mathcal{N}(-5, 0.5)$	$b_1[m]$	$\mathcal{N}(-5, 0.5)$
$B_1[m]$	Channel width	$\mathcal{LN}(ln(300), 0.32)$	$a_1[m^{3/2}/s]$	$\mathcal{LN}(ln(49.50), 3.2.10^{-2})$
$K_1[m^{1/3}/s]$	Strickler coeff.	$\mathcal{LN}(ln(35), 0.14)$		
$S_1[m/m]$	Bed slope	$\mathcal{LN}(ln(5.10^{-5}), 0.20)$		
$c_1[-]$	Exponent	$\mathcal{N}(5/3, 0.025)$	$c_1[-]$	$\mathcal{N}(5/3, 0.025)$
Control 2: Main channel				
$b_1[m]$	Offset	$\mathcal{N}(-5, 0.5)$	$b_1[m]$	$\mathcal{N}(0, 0.5)$
$B_1[m]$	Channel width	$\mathcal{LN}(ln(300), 0.32)$	$a_2[m^{3/2}/s]$	$\mathcal{LN}(ln(148.49), 2.4.10^{-2})$
$K_1[m^{1/3}/s]$	Strickler coeff.	$\mathcal{LN}(ln(35), 0.14)$		
$S_1[m/m]$	Bed slope	$\mathcal{LN}(ln(2.10^{-4}), 0.25)$		
$c_1[-]$	Exponent	$\mathcal{N}(5/3, 0.025)$	$c_2[-]$	$\mathcal{N}(5/3, 0.025)$
Control 3: Floodway				
$b_3[m]$	Offset	$\mathcal{N}(8, 0.5)$	$b_3[m]$	$\mathcal{N}(8, 0.5)$
$B_3[m]$	Channel width	$\mathcal{LN}(ln(200), 0.47)$	$a_3[m^{3/2}/s]$	$\mathcal{LN}(ln(241.9), 1.10^{-2})$
$K_3[m^{1/3}/s]$	Strickler coeff.	$\mathcal{LN}(ln(25), 0.20)$		
$S_3[m/m]$	Bed slope	$\mathcal{LN}(ln(2.4.10^{-4}), 0.21)$		
$c_3[-]$	Exponent	$\mathcal{N}(5/3, 0.025)$	$c_3[-]$	$\mathcal{N}(5/3, 0.025)$
Structural uncertainty parameters				
$\gamma_1[m^3/s]$	Intercept	$\mathcal{U}(0, 1000)$	$\gamma_1[m^3/s]$	$\mathcal{U}(0, 1000)$
$\gamma_2[-]$	Slope	$\mathcal{U}(0, 100)$	$\gamma_2[-]$	$\mathcal{U}(0, 100)$
Multiperiod RC parameters				
$\Delta l[m]$	Local change	$\mathcal{N}(0, 0.8)$	$\Delta l[m]$	$\mathcal{N}(0, 0.8)$

Table 2: Priors elicitation for Beaucaire Restitution rating curves. $\mathcal{U}(a, b)$ stands for continuous uniform distribution with bounds a and b, $\mathcal{N}(\mu, \sigma)$ for Normal distribution with mean μ and standard deviation σ, and $\mathcal{LN}(\mu, \sigma)$ for Log Normal distribution with mean μ and standard deviation σ.

3.3 Stage series

3.3.1 Pont de Beaucaire (1816 - 1967)

Thanks to the archival work of Pichard et al. (2017), a continuous stage series at Pont de Beaucaire from 1816 to 1967 is available with daily stage readings from 1816 to 1840, and three stage readings per day from 1841 to 1967. The records were made visually by an operator, at noon during the first years, then at 7am, 12am and 5pm (Figure 5). When three stage readings per day are available, the maximum of the three stages is considered as the daily maximum stage, and before 1840 the unique value at noon is kept as the daily maximum stage. Additionally, after 1840, when the stage was rising above 5 m, the operators made more frequent visual records (supposedly hourly measurements). When these records are available, they are of course used to establish the daily maximum stages.

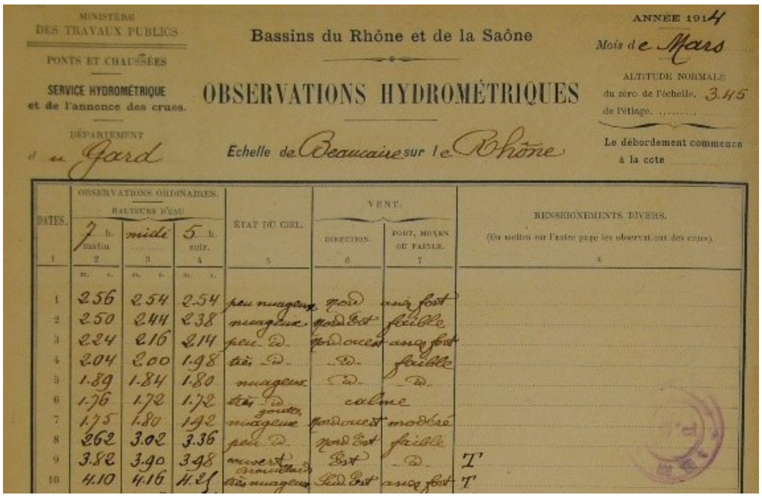


Figure 5: Table of limnimetric surveys at Pont de Beaucaire, March 1914. Operators were supposed to provide the water level at 7am, 12am and 5pm, as well as a few meteorologic details. (Ponts\&Chaussées, 1914)

Stage uncertainties depend on the measurement method, as described in section 2.3. Table 3 summarizes the different sources of stage uncertainties at Beaucaire, here given as standard deviations σ. Staff gauge reading uncertainty σ_1 is taken as 5 cm. Staff gauge precision is centimetric, but as this work is focused on floods, the error is expanded because of the waves that may complicate the reading. Sensor precision σ_2 and sensor calibration uncertainties σ_3 are not considered at Pont de Beaucaire as the stage was read by operators directly on the staff gauge. Bard and Lang (2018) have compiled several elevation measurements of the staff gauge datum between 1912 and 2010. Most of those measurements occurred after the decommissioning of Pont de Beaucaire station. Datum uncertainty σ_4 is assumed to be equal to the standard deviation of those measurements: 6 cm. The datum measurement frequency during the operation of Pont de Beaucaire station is assumed to be 25 years (i.e. the average duration between the retrieved elevation measurements). Hence, datum errors are drawn every 25 years. As described in section 2.3, the distribution of measurement frequency errors δ_5 can be estimated using the frequent stage measurements at Beaucaire Restitution between 1970 and 2020 (50 AMAX values). For the "one stage reading per day" case (mimicking the 1816-1840 period), this error corresponds to the difference between the maximum hourly stage value and the stage at noon of the same day. The "three stage readings per day" error (1840-1967) is the difference between the maximum hourly stage of a day, and the maximum of 7am, noon and 5pm stages of the same day. An exponential distribution is estimated for both errors samples and is used to represent the measurement frequency uncertainty affecting the annual maximum stages at Pont de Beaucaire. According to gauge management instructions, hourly measurements were made by observers after 1840, and for the stages above 5 meters. Hence, measurement frequency error δ_5 can be considered as negligible when stage is above 5 meters after 1840.

SYMADREM (2012), Pichard (2013) and Bard and Lang (2018) suggested that for the floods during which dike breaks happened downstream from Beaucaire, stage measurements should be corrected because the stage measured at the station may lead to underestimating the actual discharge of the flood. The stage corrections for the more thoroughly studied floods of 1840, 1841 and 1856 estimated by SYMADREM (2012) are adopted. For these floods stage uncertainty is represented by a Gaussian distribution, with mean the estimated stage and standard deviation half of the applied correction, chronologically: 0.94, 0.4 and 0.4 m.

3.3.2 Beaucaire Restitution (1970 - 2020)

For Beaucaire Restitution station, most of the stage uncertainty values come from CETIAT (2005) expertise on behalf of Compagnie Nationale du Rhône. They are summarized in table 3. Staff gauge reading uncertainty is considered zero as the measurements are done by automatic sensors. Instrument precision uncertainty: $\sigma_2 = 0.01/\sqrt{3}$ m comes from the sensor manufacturer specifications. The standard deviation of all the re-calibrations made by the operators is equal to 5 cm according to CETIAT (2005). Calibration is also affected by staff gauge reading uncertainties, because the stage read on the staff gauge is the reference used by operators to calibrate the sensor. CETIAT (2005) estimated a 3.35 cm uncertainty for the gauge reading uncertainty. Therefore, gauge reading and calibration uncertainties are combined as follows: $\sigma_3 = \sqrt{0.0335^2 + 0.05^2} = 0.06$ m. As the average time lag between calibrations is 6 months, a new value of the error δ_3 is drawn for each annual maximum stage. Datum reference uncertainty σ_4 is considered negligible because of the precision of modern topographic measurements (< 1 cm). Measurement frequency uncertainty σ_5 is considered negligible, because the sub-hourly measurement frequency is assumed adequate to capture the Rhône River stage variability.

<table>
<tr><th rowspan="2">Date</th><th rowspan="2">δ_1: gauge reading</th><th rowspan="2">δ_2: sensor precision</th><th rowspan="2">δ_3: sensor calibration</th><th rowspan="2">δ_4: datum reference</th><th colspan="2">δ_5: measurement frequency</th></tr>
<tr><th>Stage<5m</th><th>Stage≥5m</th></tr>
<tr><td>Before 1840</td><td>$\mathcal{N}(0, 0.05)$</td><td>-</td><td>-</td><td>$\mathcal{N}(0, 0.06)$</td><td colspan="2">$Exp(2.18)$</td></tr>
<tr><td>1840 - 1967</td><td>$\mathcal{N}(0, 0.05)$</td><td>-</td><td>-</td><td>$\mathcal{N}(0, 0.06)$</td><td>$Exp(8.86)$</td><td>-</td></tr>
<tr><td>1970 - 2020</td><td>-</td><td>$\mathcal{N}(0, 0.01/\sqrt{3})$</td><td>$\mathcal{N}(0, 0.06)$</td><td>-</td><td colspan="2">-</td></tr>
</table>

Table 3: Distributions used for the different sources of stage errors (in meters). $\mathcal{N}$(mean, st. deviation) and Exp(rate) represent Gaussian and Exponential distributions. As a reminder, the periods before 1967 are associated with the Pont de Beaucaire station and the period 1970-2020 with the Beaucaire Restitution station

3.4 Gaugings (discharge measurements)

A set of 244 gaugings from 1840 to 1967 has been compiled at Pont de Beaucaire. After excluding a few gaugings which were considered dubious, 233 measurements remain. The frequency of gaugings is variable in time. No gaugings were retrieved before 1840 and there are several 10- to 20-year gaps without gaugings, which makes the estimation of the stage-discharge relationship over time challenging. The assumed uncertainty of the gaugings at Beaucaire depends on the gauging method according to Bard and Lang (2018) values specified in table 4.

A set of 304 gaugings is available at Beaucaire Restitution. A few of these were out of the period of stage measurements availability and were discarded. Finally, 296 gaugings were selected. As modern hydrometric developments allowed estimating the uncertainty for each individual gauging (particularly for ADCP and current meters), those values are used when available in the CNR archives. If not, values from table 4 are considered.

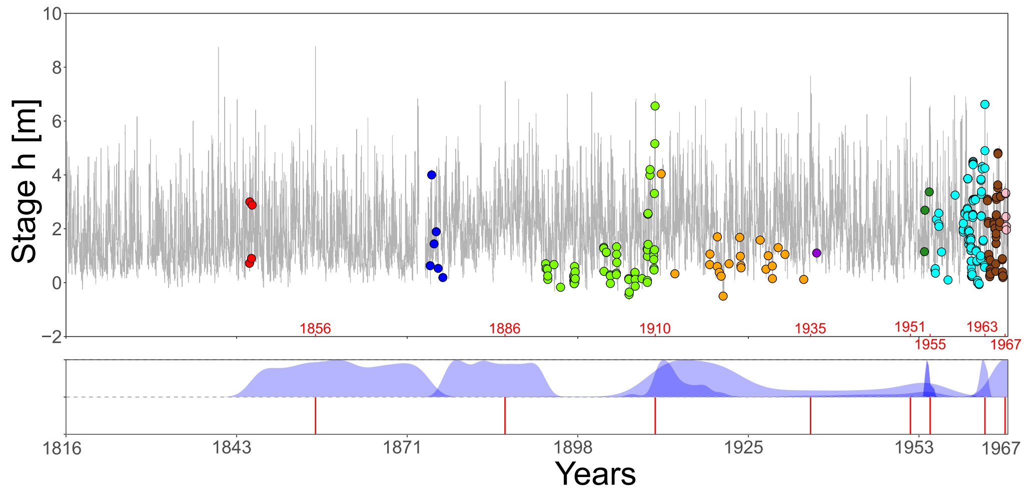
Gauging method	**Standard uncertainty**
Current meter at 0.6 h and surface	5%
Current meter point by point	3.5%
Surface current meter	7.5%
Unknown type	7.5%
ADCP	3.5%
Floats before 1936	10%
Hydrotachymeter before 1936	10%

Table 4: Gaugings uncertainty depending on the method used (hypotheses from Bard and Lang (2018)). Expressed as standard deviations of the measured discharge in %.

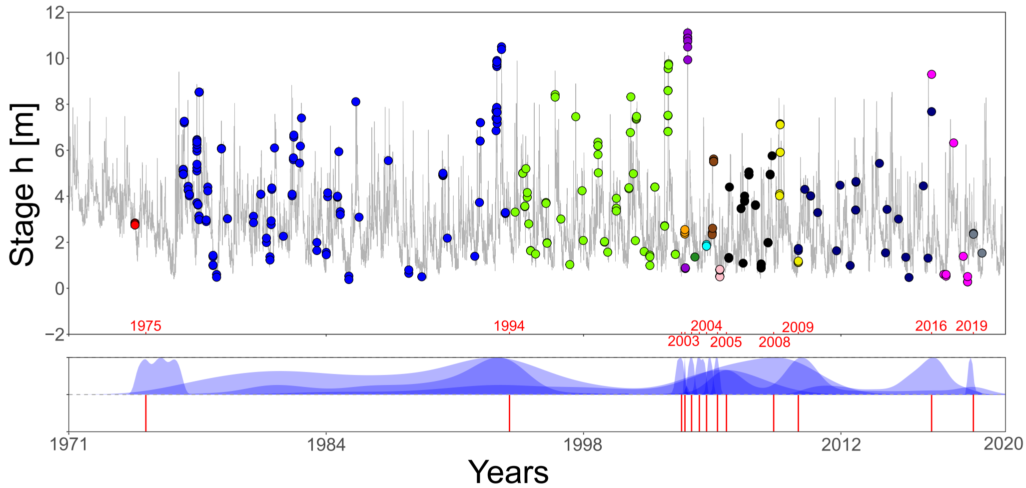
4 Results

4.1 Assessment of rating shifts

Darienzo et al. (2021) segmentation procedure is applied at Beaucaire as described in section 2.1. The prior for the residual mean during each sub-period is taken as a Gaussian distribution with zero mean and a 500 m^3/s standard deviation for both stations. The maximum number of segments at each iteration is fixed at six (see Darienzo et al. (2021) for details on priors and parameters specification).



(a)



(b)

Figure 6: Gaugings segmentation of the Rhône River at Pont de Beaucaire (a) and Beaucaire Restitution (b). Dots represent gaugings with different colors for each stable period. The grey curve is the stage series. Blue ribbons represent posterior pdf of shift times and red segments are the retained shift times taken as the maximum stage included in each posterior pdf interval.

Eight shift times are detected at Pont de Beaucaire (Figure 6 left). The gauging frequency is not constant through the history of the station and some periods include a small number of gaugings. As a consequence, the posterior pdfs of shift times span over many years for the first shifts, and are similar to uniform distributions between sets of gaugings. Without discharge measurement (gaugings) the method is not able to detect any rating shifts. Additional information may be of interest for those first periods. It is no surprise that most of the shifts occur very close to the largest historical floods. 1935 and 1951 shifts respectively correspond to the 3$^{\text{rd}}$ and 4$^{\text{th}}$ largest floods of the history of the station. There is, by construction of the segmentation model, no shift before 1845, year of the first available gaugings. However, we can consider that the 1840 flood (supposedly the largest flood since 1800) is likely to have caused a shift. Therefore, an additional shift time is added at the exact flood date. This brings the total number of stable periods to 10 (table 5).

Thirteen shift times are detected at Beaucaire Restitution. As can be seen in figure 6 (right), gauging frequency is far higher than for Pont de Beaucaire station (except during the first 5 years), resulting in a better determination of the rating shift times. Due to the lack of gaugings at the beginning of the series, only one rating shift is detected but many shifts potentially took place in those first four years as morphological adjustment and dredging occurred (see section 3.2.3). This first shift is assumed to be assigned to the first large flood of the station in 1976, after which the channel stabilized. The next shift occurred during the 1994 flood, one of the largest at the station. The most notable flood at Beaucaire Restitution occurred in 2003 (11 500 m^3/s, with a return period of about 100 years according to MEDD (2005)). Unsurprisingly, the stage-discharge relationship is considerably disrupted by this event, as reflected by the six rating shifts detected from 2003 to 2005. Two out of these six shifts were discarded because the shift amplitude is considered minor based on further analysis of the corresponding rating curve change. The largest flood within posterior intervals of those shifts almost always corresponds to the 2003 flood. This is also the case for 2005, 2008 and 2009 shifts, for which the posterior pdf spans many years including 2003. Therefore, the shift dates are assumed to be located to the maxpost shift times, as several shifts cannot be located at the same date. The last shift of 2019 is also discarded because the shift amplitude is considered minor based on further analysis. Finally, ten rating shifts are retained. This brings the number of stable periods to eleven for Beaucaire Restitution (table 5).

Maxpost shift time	Largest flood within *post. pdf*	Final choice	Period number	Number of gaugings
Pont de Beaucaire (1816-1967)				
No gaugings	No gaugings	1840-11-02	1	0
1860-02-20	1856-06-01	1856-06-01	2	4
1887-05-11	1886-10-29	1886-10-29	3	6
1910-11-21	1910-12-09	1910-12-09	4	58
1921-06-22	1935-11-14	1935-11-14	6	22
1954-08-08	1951-11-23	1951-11-23	6	1
1954-03-30	1955-01-23	1955-01-23	7	3
1963-03-23	1963-11-08	1963-11-08	8	91
1967-01-31	1967-01-31	1967-01-31	9	43
1967-12-31	End of stage series	End of stage series	10	5
Beaucaire restitution (1970-2020)				
1975-02-06	1976-11-11	1976-11-11	1	3
1994-06-10	1994-01-08	1994-01-08	2	122
2003-08-05	2002-11-27	2002-11-27	3	65
2003-10-09	2003-12-04	2003-12-04	4	17
2004-02-16	2003-12-04	No shift	X	X
2004-07-15	2003-12-04	2004-07-15	5	1
2004-12-02	2004-12-02	2004-12-02	6	2
2005-07-03	2004-12-02	No shift	X	X
2005-12-24	2003-12-04	2005-12-24	7	14
2008-06-28	2003-12-04	2008-06-28	8	28
2009-10-20	2003-12-04	2009-10-20	9	7
2016-11-21	2016-11-22	2016-11-22	10	26
2019-02-11	2018-11-24	No shift	X	X
2020-01-01	End of stage series	End of stage series	11	11

Table 5: Beaucaire rating shifts dates

4.2 Multiperiod rating curves estimation

Uncertain rating curves are estimated using Mansanarez et al. (2019a) SPD model, for each stable period detected previously. For Pont de Beaucaire, this leads to ten rating curves that show a good adequacy with gaugings (figure 7a). The evolution of the main channel offset (b_1) gives indications on the evolution of bed elevation (figure 7c). Substantial changes occurred before and after the third stable period with successive increase and decrease of the offset. Those changes may be related to the channel works that occurred during the end of the XIX$^{\text{th}}$ Century. Afterwards, the offset is more stable and only suggests a slight increasing trend which may be a consequence of the filling of the channel noticed in figure 4. The widest uncertainty interval belongs to the first period (1816-1840: dark red) for which no gaugings are available (figure 7a). The expected range of rating curve uncertainties for flood discharges (above 6 m) varies from around 20% for the first period, to less than 10% after 1840. Static parameters are precisely estimated and are presented in figure 7e. The c posterior distributions are as wide as priors because c priors are already very precise as they come from simplified Manning-Strickler formula for which the exponent is exactly 5/3.

Eleven uncertain rating curves were computed at Beaucaire Restitution (figure 7b). The rating curve uncertainty intervals are smaller than at Pont de Beaucaire for usual stages because of a larger number of gaugings and a smaller gaugings uncertainty: around 5% of uncertainty is estimated for floods above 6 m. However, low flows uncertainty is greater than at Pont de Beaucaire, because the sustained flows of the Rhône River limits the exploration of the sea-influenced hydraulic control. Low flows gaugings are unavailable. Thus, the first control offsets b_1 are not precisely estimated (figure 7d), but this has no consequences on the streamflow uncertainty of AMAX floods, for which only controls 2 and 3 are active. The first period rating curve (dark red) is shifted with respect to the other curves due to the channel adjustment and dredging operations after Vallabrègues works (1967-1970). The second control offsets (b_2) globally decreases over time, showing a slight scouring trend of the channel (figure 7d). Static parameters (figure 7f) appear precisely estimated, except for the 3$^{\text{rd}}$ control offset b_3 for which the posterior distribution is as wide as the prior.

There are several reasons for the significant differences between the upper parts of the rating curves for Pont de Beaucaire (figure 7a) and Beaucaire Restitution (figure 7b). First, the gauge datum (the altitude of the stream gauge zero value) correspond to 3.37 m for Pont de Beaucaire and 0.06 m for Beaucaire Restitution. In addition, as the stations are 2 km apart, their cross-sections are very different. Pont de Beaucaire cross-section corresponds to a main channel splitted in two sub-channels (figure 3b), while Beaucaire Restitution cross-section corresponds to a unique channel (figure 3c). A magnified representation of the upper parts of the rating curves is available in supplementary material (figure 2).

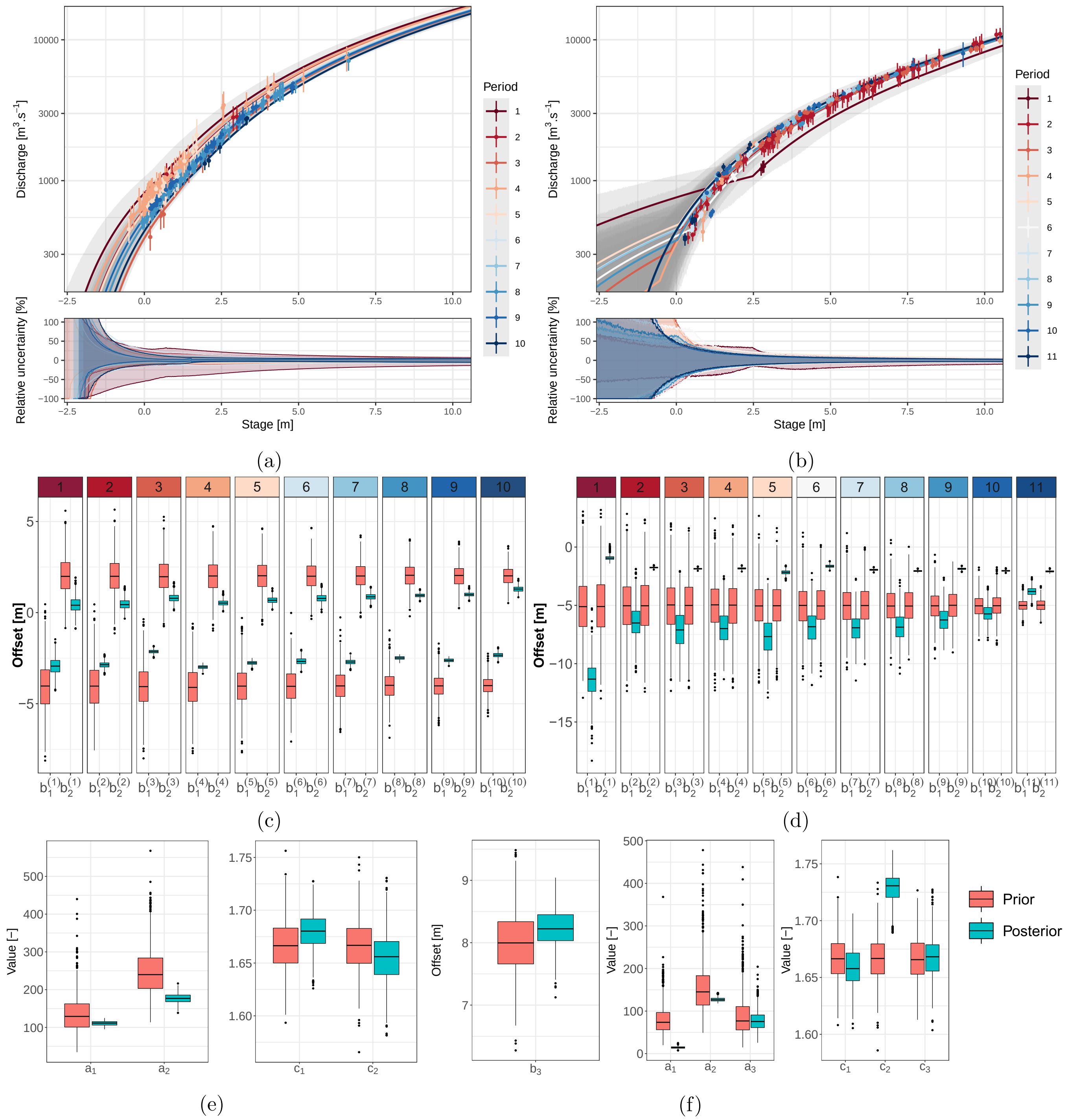


Figure 7: Pont de Beaucaire (a) and Beaucaire Restitution (b) rating curves and relative 95% uncertainty with respect to maxpost, offsets priors and posteriors (c and d) and static parameters priors and posteriors (e and f). Discharges of rating curves are in logarithmic scale, solid lines represent maxpost values, grey transparent envelops represent 95% uncertainty intervals and dots with error bars represent the gaugings with 95% uncertainty. Stable stage-discharge periods are numbered from the oldest to the latest (see table 5).

4.3 Stage uncertainty

The error sources described in table 3 are combined using a Monte Carlo procedure to quantify the uncertainty affecting AMAX stages as described in figure 8. The measured stage is outside and below the stage uncertainty interval before 1840 at Pont de Beaucaire: this is due to the exponential distribution used to model measurement frequency errors which are dominant during this period and are positive by definition. The upper uncertainty bound is sometimes 1.5 meters higher than the measured stage. Therefore, considering this source of uncertainty may have substantial consequences on the final results. The difference between uncertainty bounds and originally measured stages is presented in the bottom part of figure 8. The uncertainty of AMAX stages decreases over time as the measurement frequency and precision improve. The width of the 95% uncertainty interval is 1.7 m before 1840, 0.3 m between 1840 and 1967, and 0.24 m at Beaucaire Restitution (1970-2020). The 5 m threshold above which hourly measurements were done after 1840 explains the large reduction of the uncertainty. After 1840, the uncertainty is controlled by the exceedance of this 5 m threshold, the AMAX below 5 m being penalized by non-negligible measurement frequency errors δ_5.

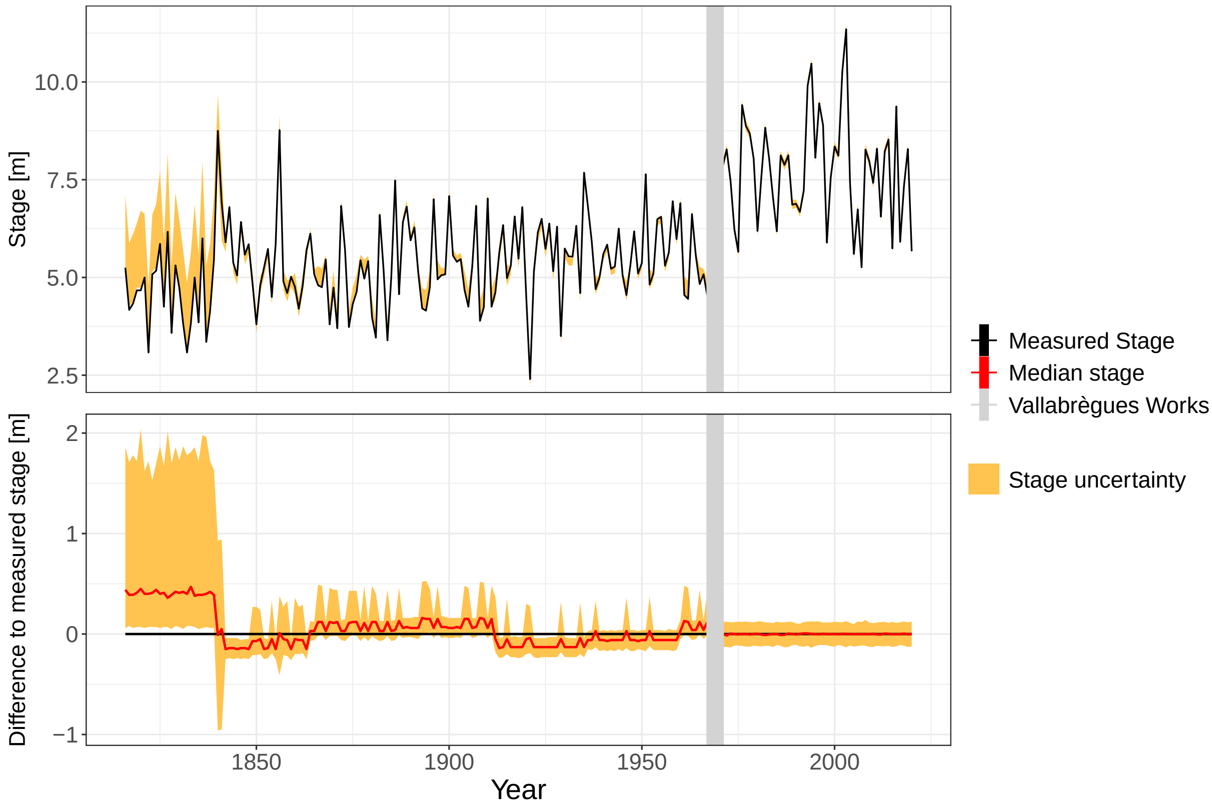


Figure 8: Top: AMAX stages and uncertainty at Beaucaire (1816-2020). Bottom: Difference between 95% stage uncertainty bounds and originally measured stage.

4.4 Total streamflow uncertainties

Stage uncertainties are propagated through uncertain rating curves as described in section 2.4 and the results are shown in figure 9. Streamflow uncertainty, although fluctuating, decreases over time from 30% before 1840, and 10% before 1967, to 5% at Beaucaire Restitution (1967-2020). This 5% uncertainty on recent AMAX discharges is consistent with the results of the international consensus conference on the December

2003 flood : "The most likely estimate of the maximum discharge of the Rhône River at Beaucaire during the December 2003 flood is 11500 m3/s, corresponding to a return period slightly above 100 years. [...] This maximum discharge estimate is subject to an uncertainty of around 5%, resulting from the uncertainty of flow measurements, maximum stage, parameterization and extrapolation of the December 2003 gauging data." (MEDD, 2005).

Stage uncertainty appears dominant at Pont de Beaucaire, as well as rating curve parametric uncertainty, originating from the difficulty to estimate rating curve parameters with only a few gaugings. Thus, parametric uncertainty is reduced for properly gauged periods. During the Vallabrègues hydraulic system construction (1967 - 1970), the waters of the Durance River, one of the major tributaries, were deviated from the Rhône River course. AMAX discharges of these missing years were reconstructed by CNR with upstream gauging stations. The uncertainty around these reconstructed discharges is assumed to be represented by a Gaussian distribution with 10% relative standard deviation. An AMAX flow (with uncertainties) time series plot is available in supplementary material (figure 3).

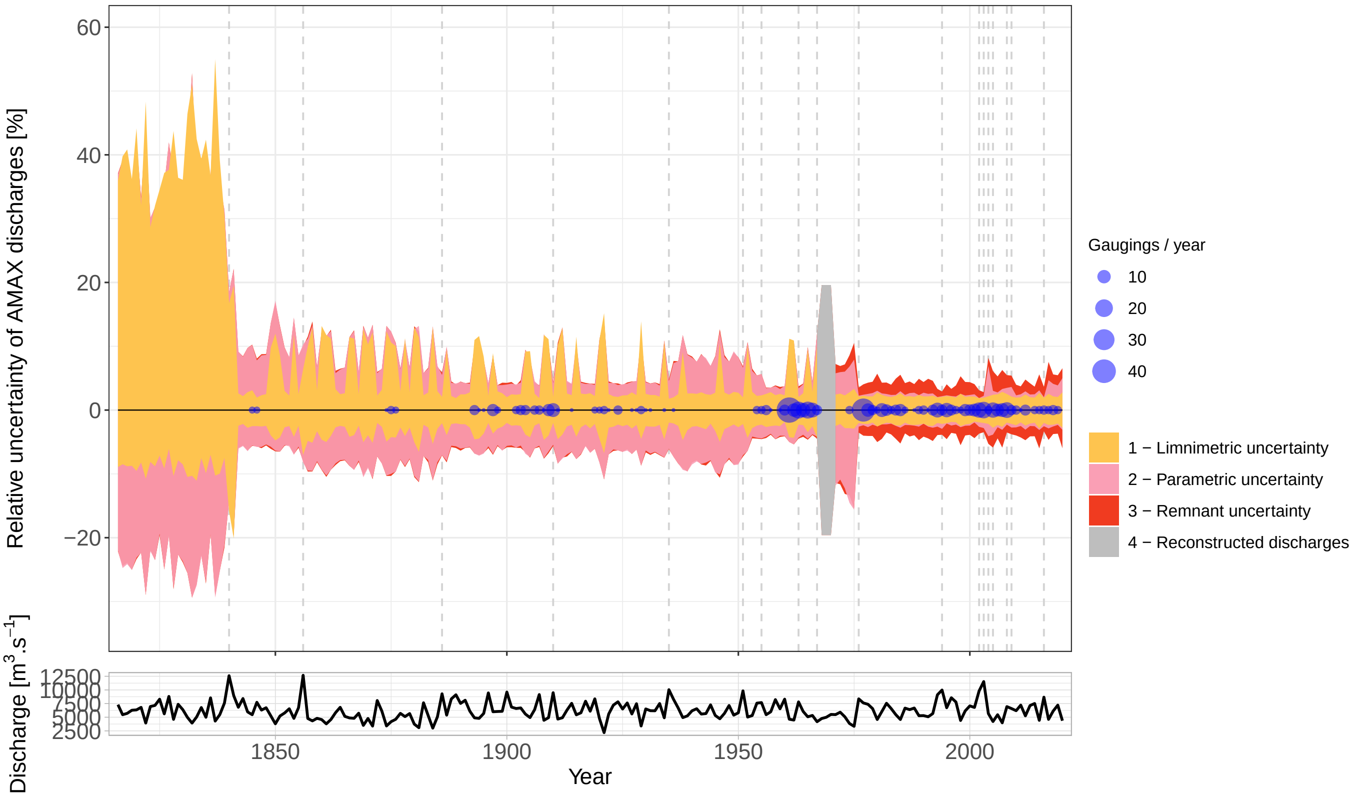


Figure 9: Relative 95% uncertainty of AMAX discharges with respect to the maxpost discharges (black solid line) for the three sources of streamflow uncertainties, at Beaucaire (1816 - 2020). Vertical dotted lines represent rating shifts.

4.5 Flood frequency analysis

4.5.1 Streamflow series homogeneity

The homogeneity of streamflow series is an essential prerequisite as FFA is based on the hypothesis of iid (independent and identically distributed) random variables. In order to check this hypothesis, the Mann-Kendall non-parametric test (Mann (1945); Kendall (1948)) is applied to AMAX series at Beaucaire. As streamflow uncertainty is represented by 500 AMAX discharge realizations (section 2.4), the homogeneity test is applied to each of the 500 AMAX series. Among the Mann-Kendall tests, 81% concluded to the non-rejection of the null hypothesis (there is no trend in the series) with a 0.05 significance level. We assume that this is enough to consider the series as homogeneous and to proceed to FFA.

4.5.2 Flood frequency analysis

The GEV distribution estimation procedure described in section 2.5 is applied to the 205-year long AMAX discharges series at Beaucaire accounting for uncertainties. Vague priors are used for GEV parameters: flat priors for location and scale parameters, and Gaussian with zero mean and 0.2 standard deviation for the shape parameter. This shape parameter prior is consistent with Martins and Stedinger (2000) suggestions. Flood quantiles results (figure 10a) show that streamflow uncertainty dominates for the lowest return periods, but sampling uncertainty becomes dominant when the return period tends toward 1000 years (see figure 10c bottom right for a better understanding of the respective part of each source of uncertainty in this 205 years case). The observed AMAX discharges display a large variability of streamflow uncertainties. The three largest floods of this 205 years sample (1840, 1856 and 2003, by chronological order and from the most uncertain to the most precise) illustrate this point. Thus, not considering 1840 and 1856 floods could have a strong effect on the estimation of the maxpost quantiles values, as well as their uncertainty. This is explored next by varying the sample size.

4.5.3 Sample size influence on quantiles uncertainty

With an exceptionally long sample at Beaucaire, the influence of sample size on flood quantiles estimation in a real case can be quantified, hence assessing the interest of using old hydrometric data when available. Four sample sizes are tested, taken as the last 50 years, 100 years, 150 years, and the largest available sample of 205 years. A figure describing these sub-samples along the AMAX time series is available in supplementary material (figure 3). GEV distributions are estimated and the contribution of both streamflow and sampling uncertainties is computed for each case, following section 2.5 procedure. Total uncertainty is clearly reduced between the 50 years sample and the other samples for the three return periods: 10, 100 and 1000 years (figure 10b). Surprisingly, for the 1000-year flood estimation, the total uncertainty is not reduced between

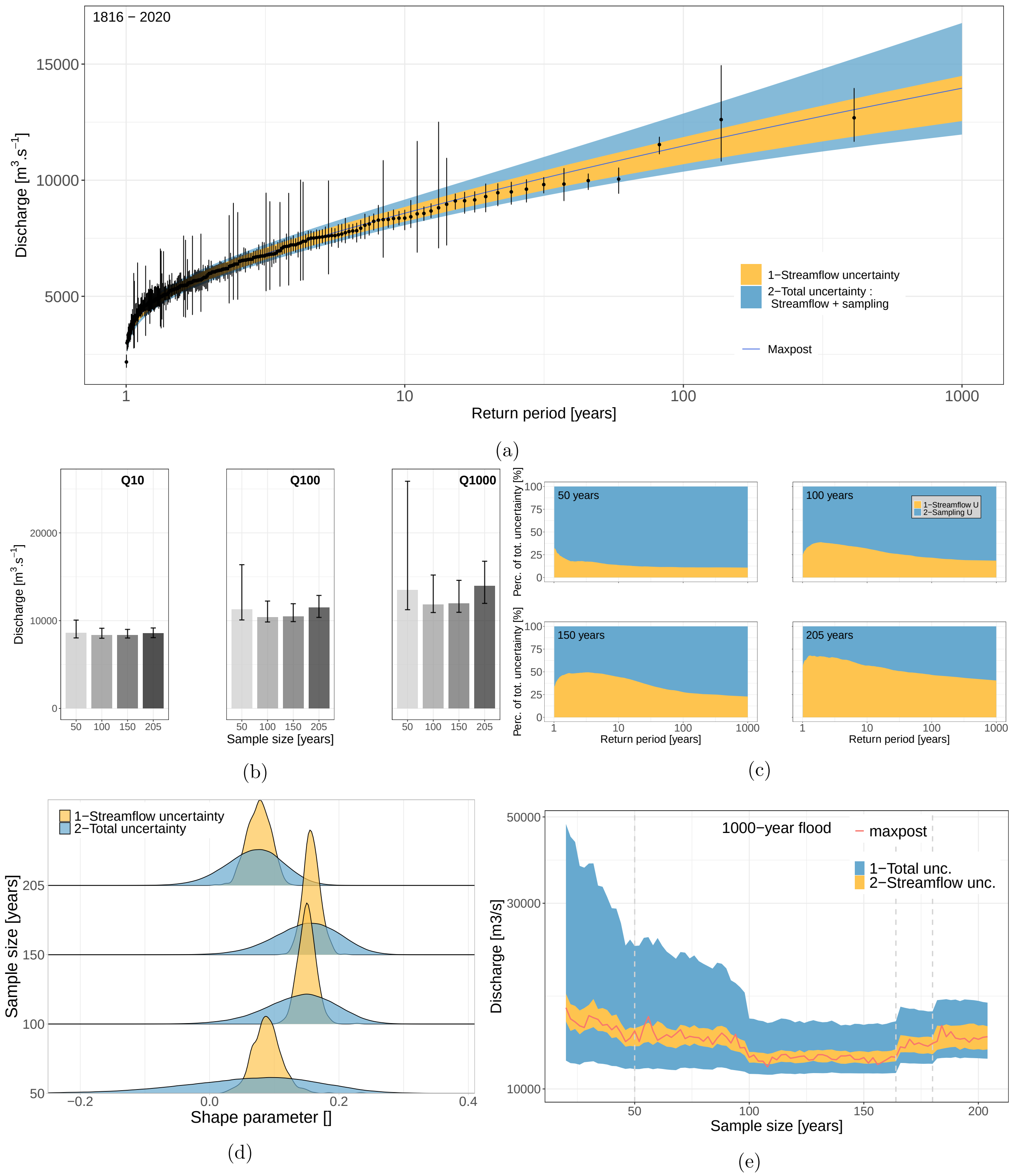


Figure 10: FFA results of the Rhône River at Beaucaire. (a) GEV distribution estimated with the full sample (1816-2020). Error bars represent the observed AMAX discharges with their 95% streamflow uncertainty. (b) Maxpost quantiles estimates for three return periods and four sample sizes. Error bars represent 95% total uncertainty intervals. (c) Contribution of streamflow and sampling uncertainties to the total uncertainty for four sample sizes. (d) GEV shape parameter distributions considering streamflow or total uncertainty for four sample sizes. (e) 1000-year flood estimations for various sample sizes. Grey dotted lines represent large changes in the sample data (the change from Pont de Beaucaire to Beaucaire Restitution gauge en 1967, the inclusion of 1856 flood and the inclusion of 1840 flood to the sample).

the 100 and 205 years samples. This illustrates that the reduction of sampling uncertainty induced by increasing the sample size may be compensated by the increased streamflow uncertainty when going back in time. Figure 10c is a good illustration of this phenomenon, showing the augmentation of the relative part of streamflow uncertainty when increasing the sample size.

The maxpost values of the 205 years sample is higher than the 100 and 150 years samples (figure 10b), probably because of the inclusion of the two largest floods of the history in the 205 years sample (1840 and 1856 floods). Thus, without using those old hydrometric data (1816-1870), the 1000-year flood could have been 15% lower in this specific case. Figure 10d shows that the streamflow uncertainty has a minor impact on shape parameter uncertainty compared with the sampling uncertainty. The sample size impact on flood quantiles estimation is further explored in Figure 10e. The 1000-year flood (maxpost) and the relative contributions of both sources of uncertainty are estimated for several sample sizes, from 20 to 205 years, with a two-year step. A large reduction of sampling uncertainty (and of the total uncertainty as a consequence) appears between 20 and 100 years, along with the reduction of the maxpost value. Then, the maxpost and uncertainty intervals are constant between 100 and 160 years of sample size, until the inclusion of 1856 and 1840 major floods that lead to a slight increase in the estimated quantile. The total uncertainty interval width is not much changed by those flood inclusions but the relative contribution of streamflow uncertainty is larger.

5 Discussion

5.1 Usefulness of disentangling the various sources of uncertainty in FFA

Even though sampling uncertainty is a major concern in FFA, historical continuous stage records are rarely used. The laborious process of gathering and reanalyzing data, and concerns about the reliability of the resulting discharge series may be the potential reasons for this oversight. Moreover, the estimation and the propagation of the various sources of uncertainty throughout the FFA chain is not straightforward to achieve. Several papers performed an integrated analysis in which both rating curve parameters and flood frequency distribution were estimated (Petersen-Øverleir and Reitan (2009); Steinbakk et al. (2016); Vieira et al. (2022)). They concluded that accounting for rating curve uncertainties may notably widen the total uncertainty of flood quantiles. More specifically, Steinbakk et al. (2016) based their case study on several Norwegian rivers, with sample sizes up to 100 years. They found that sampling uncertainty is generally the main contributor (i.e. versus rating curve uncertainty) to design flood estimates and that its contribution decreases when sample size increases.

The approach proposed in this paper does not only consider rating curve uncertainty, but also stage uncertainties (that combine stage measurement and time interpolation errors) and rating changes. Their consideration is crucial when dealing with long streamflow series. The case study investigated in this paper is based on a 205-year long continuous sample, which enables an in-depth evaluation of the contribution of the different sources of uncertainty in design flood estimates from series of several decades to long series exceeding one century of systematic record. The results show that the estimated 95% streamflow uncertainty is not negligible and varies from 30% (XIX$^{\text{th}}$ Century) to 5% (1967-2020). The larger streamflow uncertainty of the XIX$^{\text{th}}$ Century is mostly due to stage uncertainty, which emphasises the importance of paying a particular attention to this uncertainty source when using historical stage measurements. Streamflow uncertainty is then propagated to flood quantiles estimates. For the 205-year long flood sample, streamflow uncertainty dominates for return periods below 100 years but sampling uncertainty becomes dominant when the return period tends toward 1000 years. The 205-year long streamflow series enables exploring the contribution of streamflow and sampling uncertainties for various sample sizes. Sampling uncertainty contribution decreases when sample size increases, unlike streamflow uncertainty contribution which increases when older data are included. Although the relative contribution of each source of uncertainty on extreme quantiles varies, the width of the total uncertainty interval does not change much from 100 to 205 years samples. Nevertheless, in this case study, the maxpost value of extreme quantiles is increased by 15% when considering 205 years rather than 150 or 100 years, as the two largest floods since 1816 occurred during the first 40 years of records. This emphasizes that using longer records does not only affect the uncertainty around flood quantiles, but also the value of the quantiles themselves.

5.2 Potential improvements of the method and further analyses

The streamflow uncertainty analysis procedure proposed in this work is affected by several limitations. The gauging segmentation model proposed by Darienzo et al. (2021) (or more generally any gaugings-based segmentation approach) is limited when gaugings are scarce. Some rating shifts are likely to be missed, and even when one is detected, assigning a precise shift time is difficult when very little information is available about rating changes and channel morphology. Other approaches could be trialed to detect changes in the stage-discharge relationship, such as the analysis of stage or discharge recessions (for instance: Nathan and McMahon (1990); Tallaksen (1995); Vogel and Kroll (1996); Chapman (1999); Lang et al. (2010); Darienzo (2021)) or the occurrences of morphogenic events (Darienzo (2021)).

Another limitation of the method comes from the elicitation of hydraulic priors in a historical context, for which information about hydraulic configurations is scarce. This lack of knowledge is particularly detrimental for stations where changes in bed morphology are frequent. As described by Petersen-Øverleir and Reitan

(2009), the extrapolation of rating curves in this context is more uncertain, thus affecting flood quantile estimates.

The shape parameter is of great importance as it determines the tail behaviour of the GEV distribution. Several regional or local-regional methods have been proposed to reduce the uncertainty of the shape parameter estimation (Burn (1990); Ouarda et al. (2001); Ribatet et al. (2007); Micevski and Kuczera (2009); Haddad and Rahman (2012)). This could be a source of improvement, but a regional approach could be difficult to apply here, because catchments as large as the Rhône River at Beaucaire cover many different climatic influences and are quite unique, which makes the use of other similar catchments challenging. An alternative solution to this approach is to analyse floods in the main subcatchments and their concomitance.

The underlying assumption of stationarity required for FFA is questionnable due to anthropogenic climate change, as described by Milly et al. (2008). However, Madsen et al. (2014) underlined that no particular guideline for climate change adjustment factors on design flood are given in France. Trends in flood magnitudes have been identified in several climatic regions of Europe (Hall et al. (2014); Blöschl et al. (2019)) and France (Giuntoli et al., 2019), but not everywhere and with large regional differences. Many approaches of FFA accounting for non-stationarity have emerged (as reviewed by Salas et al. (2018)) and are sometimes combined with regional approaches (Han et al., 2022). After identifying the regional trends of the Rhône River, a next challenge could be to develop a climate-informed model to account for the effects of climate change on floods.

Finally, a promising development is to use sporadic flood evidences older than systematic stage measurements. This data can come from various origins such as testimonies (Pichard and Roucaute, 2014), flood marks (Renard, 2023), paleoflood evidences based on slack water deposits (Sheffer et al., 2003), lake sediments (Wilhelm et al., 2022) or sediments within the Rhone prodelta (Fanget et al., 2013). Various procedures have been developed in the literature to include such data in FFA through the use of perception threshold and censored data (Brázdil et al., 2006; Kjeldsen et al., 2014; England et al., 2019; Harden et al., 2021). Such approaches have been applied in Europe, including France (Naulet et al. (2005); Lang et al. (2010); Neppel et al. (2010); Payrastre et al. (2011)) and could be interesting for the Lower-Rhône Valley for which many flood evidences (along with information on other climate-related disasters) have been gathered. Pichard and Roucaute (2014) and Pichard et al. (2017) have identified more than 1500 hydro-climatic events in the Lower-Rhône Valley since the XIV$^{\text{th}}$ Century, synthesized in the HISTRHÔNE database (`histrhone.cerege.fr`). The flood events are classified by magnitude of damages, and further investigations are required to identify the perception thresholds corresponding to those different magnitudes.

In addition, as this information is not exhaustive (unlike the stage measurements used in this paper), the extension of the sample before 1816 requires a different statistical treatment. Moreover, the interest of long discharge series and flood evidences is not limited to standard FFA, but is also useful for studying the long-term historical variability of floods (e.g Macdonald and Sangster, 2017).

5.3 Are historical stage records useful for flood frequency analysis?

The interest of including historical stage records to reduce sampling uncertainty in FFA should be balanced against the large streamflow uncertainty induced by those records. For the specific case of Beaucaire, the use of historical stage records is clearly beneficial up to a 100-year sample size, but the added value is not as clear with longer samples (see section 4.5.3 and figure 10). Evaluating the procedure on more stations with historical systematic stage records would be necessary in order to generalize the results. The respective contribution of streamflow and sampling uncertainties could be different for other stations depending on their respective hydraulic configuration and river bed stability.

In this paper, the estimation of flood discharges follows the usual hydrometric process (i.e. converting the measured stages into discharges via the estimation of rating curves from gaugings). However, the stage measurements and gaugings from the XIX$^{\text{th}}$ Century are scarce compared to recent decades, which leads to large uncertainties around flood discharges. Those uncertainties could be reduced in multiple ways, including the use of hydraulic models. This practice is widespread in the literature and is generally applied to floods older than systematic measurements (for instance: Naulet et al. (2005); Neppel et al. (2010); Machado et al. (2015); Ruiz-Bellet et al. (2017); Van der Meulen et al. (2021)). Yet, it requires topographic and bathymetric data that may be even more scarce and uncertain than gaugings. Moreover, the use of flood evidences older than systematic measurements generally leads to the assumption that all the floods greater than an identified perception threshold are know, as well as the magnitude of the threshold itself.

The specific conclusions of the Rhône River at Beaucaire may not be generalized to other climatic regions for which the tail behaviour of flood distribution is influenced by different processes (Merz et al., 2022). More specifically, the Rhône River at Beaucaire has a positive shape parameter, which corresponds to a light-tailed behavior with the GEV parameterization used in this paper. While this might be typical of such large catchments, many smaller catchments display the opposite heavy-tailed behavior. Whether or not the main conclusions drawn in this paper would still hold with such smaller catchments remains to be evaluated.

6 Conclusion

Flood hazard estimation is affected by several sources of uncertainty, including sampling uncertainty that is dominant for usual sample sizes (less than 100 years). It is sometimes possible to gather historical continuous stage measurements in order to enlarge flood samples beyond usual sizes. This process has the potential to reduce the uncertainties of design flood estimates. Nevertheless, the streamflow series derived from these historical stage series are generally affected by much greater uncertainties than modern series. This paper investigates the following questions: to what extent does including historical (and thus uncertain) hydrometric data improves FFA estimates, and what is the contribution of streamflow and sampling uncertainties to the total FFA uncertainty ? Those questions are explored through a general FFA framework accounting for the specific uncertainties affecting long hydrometric series. This uncertainty propagation chain is applied to a 205-year long continuous stage series of the Rhône River at Beaucaire.

The estimated streamflow uncertainty varies from 30% (XIX$^{\mathrm{th}}$ Century) to 5% (1967-2020). This uncertainty is propagated to flood quantiles estimates. When using the full flood sample (205 years), streamflow uncertainty is dominant below the 100-year flood and sampling uncertainty is dominant above. However, this conclusion is sensitive to the available sample size. The sample size impact on design flood estimates is explored by subsampling the full flood sample. The total uncertainty of flood quantiles substantially decreases from 20- to 100-year samples. This decrease is directly induced by the reduction of sampling uncertainty. For sample sizes between 100 and 205 years, the total uncertainty is nearly constant because the sampling uncertainty reduction is offset by the increase of streamflow uncertainty (older flood discharges are more uncertain). However, the central flood quantiles estimates increase by about 15% when increasing sample size between 160 and 205 years, because of the inclusion of the two largest floods that occurred during the first 40 years of measurement. Yet, this 15% increase is slight with respect to the total uncertainty. One should be cautious about generalizing these results beyond the particular case of the Rhône River at Beaucaire, as the contribution of sampling and streamflow uncertainties may strongly depend on the properties of the station and the catchment.

Finally, this article promotes the use of historical stage records to improve design flood estimates and underline the particular importance of estimating and propagating all sources of uncertainty through the estimation process. Discussions are currently underway with the competent authorities in flood risk management concerning the updating of design flood (Q100, Q1000, Q1500) with enlarged samples (from 80 to 205 year-discharge series), and on ways to account for estimation uncertainty. The use of estimates derived from the predictive distribution (also referred to as the expected probability approach, e.g. Kuczera, 1999;

Renard et al., 2013) could be considered since it is naturally suited to the Bayesian approach used here. Interesting improvements may come from the use of sporadic flood evidences older than the systematic stage measurements used in this paper, or from regional or non-stationary FFA approaches. Moreover, beyond standard FFA, such long series also have the potential to shed light on the long-term historical variability of floods.

7 Acknowledgements

The PhD fellowship of Mathieu Lucas is funded by INRAE, the Compagnie Nationale du Rhône (CNR) and EUR H2O'Lyon (ANR-17-EURE-0018) from the University of Lyon. This study was conducted within the Rhône Sediment Observatory (OSR), a multi-partner research program funded through the Plan Rhône by the European Regional Development Fund (ERDF), Agence de l'eau RMC, CNR, EDF, and three regional councils (Auvergne-Rhône-Alpes, PACA, and Occitanie). Data and expert knowledge on the Rhône River at Beaucaire were provided by CNR, the Rhône Sediment Observatory, Pascal Billy and Helene Decourcelle from the DREAL Auvergne-Rhône-Alpes (Ministry of Ecology) and the HISTRHONE database from the CEREGE (Georges Pichard).

The data is available at https://www.plan-rhone.fr/ (up to 2016) and uncertainty propagation code with the whole data set ("QuantilesAmax.txt") is available at `https://github.com/MatLcs/PropagMaxAn`.

References

Armand (1907). "II. Le Rhône à Tarascon (Planche I)". In: *Revue des Études Anciennes* 9.1, pp. 19–21. ISSN: 0035-2004. DOI: 10.3406/rea.1907.1469.

Bard, A. and M. Lang (2018). *Actualisation de l'hydrologie des crues du Rhône. Analyse de la station de Beaucaire.* Irstea pour la DREAL Auvergne-Rhône-Alpes, p. 56.

Blöschl, G., J. Hall, A. Viglione, R. A. P. Perdigão, J. Parajka, B. Merz, D. Lun, B. Arheimer, G. T. Aronica, A. Bilibashi, M. Boháč, O. Bonacci, M. Borga, I. Čanjevac, A. Castellarin, G. B. Chirico, P. Claps, N. Frolova, D. Ganora, L. Gorbachova, A. Gül, J. Hannaford, S. Harrigan, M. Kireeva, A. Kiss, T. R. Kjeldsen, S. Kohnová, J. J. Koskela, O. Ledvinka, N. Macdonald, M. Mavrova-Guirguinova, L. Mediero, R. Merz, P. Molnar, A. Montanari, C. Murphy, M. Osuch, V. Ovcharuk, I. Radevski, J. L. Salinas, E. Sauquet, M. Šraj, J. Szolgay, E. Volpi, D. Wilson, K. Zaimi, and N. Živković (2019). "Changing climate both increases and decreases European river floods". In: *Nature* 573.7772, pp. 108–111. ISSN: 1476-4687. DOI: 10.1038/s41586-019-1495-6.

Brázdil, R., Z. W. Kundzewicz, and G. Benito (2006). "Historical hydrology for studying flood risk in Europe". In: *Hydrological Sciences Journal* 51.5, pp. 739–764. ISSN: 0262-6667, 2150-3435. DOI: 10.1623/hysj.51.5.739.

Burn, D. H. (1990). "Evaluation of regional flood frequency analysis with a region of influence approach". In: *Water Resources Research* 26.10, pp. 2257–2265. ISSN: 1944-7973. DOI: 10.1029/WR026i010p02257.

CETIAT (2005). *Conférence de consensus sur le débit du Rhône à Beaucaire pour la crue de Décembre 2003. Annexe V : Estimation des incertitudes des débits calculés à partir des relations hauteur/débit.* pour le compte de la CNR. Estimation des incertitudes des débits calculés à partir des relations hauteur/débit.

Chapman, T. (1999). "A comparison of algorithms for stream flow recession and baseflow separation". In: *Hydrological Processes* 13.5, pp. 701–714. ISSN: 1099-1085. DOI: 10.1002/(SICI)1099-1085(19990415)13:5<701::AID-HYP774>3.0.CO;2-2.

Coles, S. (2001). "Classical Extreme Value Theory and Models". In: *An Introduction to Statistical Modeling of Extreme Values.* Ed. by S. Coles. Springer Series in Statistics. London: Springer, pp. 45–73. ISBN: 978-1-4471-3675-0. DOI: 10.1007/978-1-4471-3675-0_3.

Coxon, G., J. Freer, I. K. Westerberg, T. Wagener, R. Woods, and P. J. Smith (2015). "A novel framework for discharge uncertainty quantification applied to 500 UK gauging stations". In: *Water Resources Research* 51.7, pp. 5531–5546. ISSN: 0043-1397, 1944-7973. DOI: 10.1002/2014WR016532.

Darienzo, M., B. Renard, J. Le Coz, and M. Lang (2021). "Detection of Stage-Discharge Rating Shifts Using Gaugings: A Recursive Segmentation Procedure Accounting for Observational and Model Uncertainties". In: *Water Resources Research* 57.4. ISSN: 0043-1397, 1944-7973. DOI: 10.1029/2020WR028607.

Darienzo, M. (2021). "Detection and estimation of stage-discharge rating shifts for retrospective and real-time streamflow quantification". These de doctorat. Université Grenoble Alpes.

Dymond, J. R. and R. Christian (1982). "Accuracy of discharge determined from a rating curve". In: *Hydrological Sciences Journal* 27.4, pp. 493–504. ISSN: 0262-6667, 2150-3435. DOI: 10.1080/02626668209491128.

England, J. F. J., T. A. Cohn, B. A. Faber, J. R. Stedinger, W. O. T. Jr, A. G. Veilleux, J. E. Kiang, and J. Robert R. Mason (2019). *Guidelines for determining flood flow frequency — Bulletin 17C*. 4-B5. U.S. Geological Survey.

Fanget, A.-S., M.-A. Bassetti, M. Arnaud, J.-F. Chiffoleau, D. Cossa, A. Goineau, C. Fontanier, R. Buscail, G. Jouet, G. Maillet, A. Negri, B. Dennielou, and S. Berné (2013). "Historical evolution and extreme climate events during the last 400years on the Rhone prodelta (NW Mediterranean)". In: *Marine Geology* 346, pp. 375–391. ISSN: 00253227. DOI: 10.1016/j.margeo.2012.02.007.

Giuntoli, I., B. Renard, and M. Lang (2019). "Floods in France". In: *Changes in Flood Risk in Europe*. 1st ed. CRC Press, p. 13. ISBN: 978-0-203-09809-7.

Guerrero, J.-L., I. K. Westerberg, S. Halldin, C.-Y. Xu, and L.-C. Lundin (2012). "Temporal variability in stage–discharge relationships". In: *Journal of Hydrology* 446-447, pp. 90–102. ISSN: 00221694. DOI: 10.1016/j.jhydrol.2012.04.031.

Haddad, K. and A. Rahman (2012). "Regional flood frequency analysis in eastern Australia: Bayesian GLS regression-based methods within fixed region and ROI framework – Quantile Regression vs. Parameter Regression Technique". In: *Journal of Hydrology* 430-431, pp. 142–161. ISSN: 0022-1694. DOI: 10.1016/j.jhydrol.2012.02.012.

Hall, J., B. Arheimer, M. Borga, R. Brázdil, P. Claps, A. Kiss, T. R. Kjeldsen, J. Kriaučiūnienė, Z. W. Kundzewicz, M. Lang, M. C. Llasat, N. Macdonald, N. McIntyre, L. Mediero, B. Merz, R. Merz, P. Molnar, A. Montanari, C. Neuhold, J. Parajka, R. A. P. Perdigão, L. Plavcová, M. Rogger, J. L. Salinas, E. Sauquet, C. Schär, J. Szolgay, A. Viglione, and G. Blöschl (2014). "Understanding flood regime changes in Europe: a state-of-the-art assessment". In: *Hydrology and Earth System Sciences* 18.7, pp. 2735–2772. ISSN: 1607-7938. DOI: 10.5194/hess-18-2735-2014.

Hamed, K. H. and A. Ramachandra Rao (2019). *Flood Frequency Analysis*. 1st ed. CRC Press. ISBN: 978-0-429-12881-3. DOI: 10.1201/9780429128813.

Hamilton, A. and R. Moore (2012). "Quantifying Uncertainty in Streamflow Records". In: *Canadian Water Resources Journal / Revue canadienne des ressources hydriques* 37.1, pp. 3–21. ISSN: 0701-1784, 1918-1817. DOI: 10.4296/cwrj3701865.

Han, X., R. Mehrotra, A. Sharma, and A. Rahman (2022). "Incorporating nonstationarity in regional flood frequency analysis procedures to account for climate change impact". In: *Journal of Hydrology* 612, p. 128235. ISSN: 0022-1694. DOI: 10.1016/j.jhydrol.2022.128235.

Harden, T. M., K. R. Ryberg, J. E. O'Connor, J. M. Friedman, and J. E. Kiang (2021). *Historical and paleoflood analyses for probabilistic flood-hazard assessments—Approaches and review guidelines*. 4-B6. ISSN: 2328-7055 Publication Title: Techniques and Methods. U.S. Geological Survey. DOI: 10.3133/tm4B6.

Herschy, R. (1998). *Hydrometry Principles and Practices*. 2nd ed. John Wiley. 384 pp.

Horner, I., B. Renard, J. Le Coz, F. Branger, H. K. McMillan, and G. Pierrefeu (2018). "Impact of Stage Measurement Errors on Streamflow Uncertainty". In: *Water Resources Research* 54.3, pp. 1952–1976. ISSN: 0043-1397, 1944-7973. DOI: 10.1002/2017WR022039.

Ibbitt, R. P. and C. P. Pearson (1987). "Gauging frequency and detection of rating changes". In: *Hydrological Sciences Journal* 32.1, pp. 85–103. ISSN: 0262-6667, 2150-3435. DOI: 10.1080/02626668709491164.

Jain, S. K. and V. P. Singh (2019). "Design Flood Estimation". In: *Engineering Hydrology: An Introduction to Processes, Analysis, and Modeling*. First edition. New York: McGraw-Hill Education. ISBN: 978-1-259-64197-8.

Juston, J., P.-E. Jansson, and D. Gustafsson (2014). "Rating curve uncertainty and change detection in discharge time series: case study with 44-year historic data from the Nyangores River, Kenya". In: *Hydrological Processes* 28.4, pp. 2509–2523. ISSN: 08856087. DOI: 10.1002/hyp.9786.

Kendall, M. (1948). *Rank Correlation Methods*. 1. London: Charles Griffin & Company.

Kiang, J. E., C. Gazoorian, H. McMillan, G. Coxon, J. L. Coz, I. K. Westerberg, A. Belleville, D. Sevrez, A. E. Sikorska, A. Petersen-Øverleir, T. Reitan, J. Freer, B. Renard, V. Mansanarez, and R. Mason (2018). "A Comparison of Methods for Streamflow Uncertainty Estimation". In: *Water Resources Research* 54.10, pp. 7149–7176. DOI: https://doi.org/10.1029/2018WR022708.

Kjeldsen, T., N. Macdonald, M. Lang, L. Mediero, T. Albuquerque, E. Bogdanowicz, R. Brázdil, A. Castellarin, V. David, A. Fleig, G. Gül, J. Kriauciuniene, S. Kohnová, B. Merz, O. Nicholson, L. Roald, J. Salinas, D. Sarauskiene, M. Šraj, W. Strupczewski, J. Szolgay, A. Toumazis, W. Vanneuville, N. Veijalainen, and D. Wilson (2014). "Documentary evidence of past floods in Europe and their utility in flood frequency estimation". In: *Journal of Hydrology* 517, pp. 963–973. ISSN: 00221694. DOI: 10.1016/j.jhydrol.2014.06.038.

Kjeldsen, T. R., R. Lamb, and S. D. Blazkova (2011). "Uncertainty in Flood Frequency Analysis". In: *Applied Uncertainty Analysis for Flood Risk Management*. Imperial college press, pp. 153–197. ISBN: 978-1-84816-270-9.

Kuczera, G. (1999). “Comprehensive at-site flood frequency analysis using Monte Carlo Bayesian inference”. In: *Water Resources Research* 35.5, pp. 1551–1557. ISSN: 1944-7973. DOI: 10.1029/1999WR900012.

Kuentz, A., T. Mathevet, D. Cœur, C. Perret, J. Gailhard, L. Guérin, Y. Gash, and V. Andréassian (2014). “Hydrométrie et hydrologie historiques du bassin de la Durance”. In: *La Houille Blanche* 100.4, pp. 57–63. ISSN: 0018-6368, 1958-5551. DOI: 10.1051/lhb/2014039.

Lang, M. and D. Coeur (2014). *Les inondations remarquables en France. Inventaire 2011 pour la directive Inondation.* 1st ed. Hors Collection. Editions Quae. 640 pp. ISBN: 978-2-7592-2260-5.

Lang, M., K. Pobanz, B. Renard, E. Renouf, and E. Sauquet (2010). “Extrapolation of rating curves by hydraulic modelling, with application to flood frequency analysis”. In: *Hydrological Sciences Journal* 55.6, pp. 883–898. ISSN: 0262-6667, 2150-3435. DOI: 10.1080/02626667.2010.504186.

Łapuszek, M. and A. Lenar-Matyas (2015). “Methods of analysis the riverbed evolution. a case study of two tributaries of the upper Vistula river”. In: *Infrastruktura i Ekologia Terenów Wiejskich / Infrastructure and ecology of rural areas* (IV/3/2015), pp. 1313–1327. ISSN: 1732-5587. DOI: 10.14597/infraeco.2015.4.3.095.

Le Coz, J., P.-M. Bechon, B. Camenen, and G. Dramais (2014a). “Quantification des incertitudes sur les jaugeages par exploration du champ des vitesses”. In: *La Houille Blanche* 100.5, pp. 31–39. ISSN: 0018-6368, 1958-5551. DOI: 10.1051/lhb/2014047.

Le Coz, J., B. Renard, L. Bonnifait, F. Branger, and R. Le Boursicaud (2014b). “Combining hydraulic knowledge and uncertain gaugings in the estimation of hydrometric rating curves: A Bayesian approach”. In: *Journal of Hydrology* 509, pp. 573–587. ISSN: 00221694. DOI: 10.1016/j.jhydrol.2013.11.016.

Le Delliou, P. (2014). “Recommandations pour le dimensionnement des évacuateurs de crues de barrages”. In: *La Houille Blanche* 100.5, pp. 54–58. ISSN: 0018-6368, 1958-5551. DOI: 10.1051/lhb/2014050.

Macdonald, N. and H. Sangster (2017). “High-magnitude flooding across Britain since AD 1750”. In: *Hydrology and Earth System Sciences* 21.3. Publisher: Copernicus GmbH, pp. 1631–1650. ISSN: 1027-5606. DOI: 10.5194/hess-21-1631-2017.

Machado, M. J., B. A. Botero, J. López, F. Francés, A. Díez-Herrero, and G. Benito (2015). “Flood frequency analysis of historical flood data under stationary and non-stationary modelling”. In: *Hydrology and Earth System Sciences* 19.6, pp. 2561–2576. ISSN: 1607-7938. DOI: 10.5194/hess-19-2561-2015.

Madsen, H., D. Lawrence, M. Lang, M. Martinkova, and T. R. Kjeldsen (2014). “Review of trend analysis and climate change projections of extreme precipitation and floods in Europe”. In: *Journal of Hydrology* 519, pp. 3634–3650. ISSN: 0022-1694. DOI: 10.1016/j.jhydrol.2014.11.003.

Mann, H. B. (1945). “Nonparametric Tests Against Trend”. In: *Econometrica* 13.3. Publisher: [Wiley, Econometric Society], pp. 245–259. ISSN: 0012-9682. DOI: 10.2307/1907187.

Mansanarez, V., B. Renard, J. L. Coz, M. Lang, and M. Darienzo (2019a). "Shift Happens! Adjusting Stage-Discharge Rating Curves to Morphological Changes at Known Times". In: *Water Resources Research* 55.4, pp. 2876–2899. ISSN: 0043-1397, 1944-7973. DOI: 10.1029/2018WR023389.

Mansanarez, V. (2016). "Non-unique stage-discharge relations : Bayesian analysis of complex rating curves and their uncertainties". PhD thesis. Université Grenoble Alpes.

Mansanarez, V., I. K. Westerberg, N. Lam, and S. W. Lyon (2019b). "Rapid Stage-Discharge Rating Curve Assessment Using Hydraulic Modeling in an Uncertainty Framework". In: *Water Resources Research* 55.11, pp. 9765–9787. ISSN: 0043-1397, 1944-7973. DOI: 10.1029/2018WR024176.

Martins, E. S. and J. R. Stedinger (2000). "Generalized maximum-likelihood generalized extreme-value quantile estimators for hydrologic data". In: *Water Resources Research* 36.3, pp. 737–744. ISSN: 00431397. DOI: 10.1029/1999WR900330.

McMillan, H. K. and I. K. Westerberg (2015). "Rating curve estimation under epistemic uncertainty". In: *Hydrological Processes* 29.7, pp. 1873–1882. ISSN: 08856087. DOI: 10.1002/hyp.10419.

McMillan, H., J. Freer, F. Pappenberger, T. Krueger, and M. Clark (2010). "Impacts of uncertain river flow data on rainfall-runoff model calibration and discharge predictions". In: 24. DOI: https://doi.org/10.1002/hyp.7587.

McMillan, H., T. Krueger, and J. Freer (2012). "Benchmarking observational uncertainties for hydrology: rainfall, river discharge and water quality". In: *Hydrological Processes* 26.26, pp. 4078–4111. ISSN: 08856087. DOI: 10.1002/hyp.9384.

MEDD (2005). *Débit du Rhône à Beaucaire pour la crue de Décembre 2003.* Conférence de consensus. Minitère français de l'écologie et du développement durable.

Merz, B., S. Basso, S. Fischer, D. Lun, G. Blöschl, R. Merz, B. Guse, A. Viglione, S. Vorogushyn, E. Macdonald, L. Wietzke, and A. Schumann (2022). "Understanding Heavy Tails of Flood Peak Distributions". In: *Water Resources Research* 58.6, e2021WR030506. ISSN: 1944-7973. DOI: 10.1029/2021WR030506.

Micevski, T. and G. Kuczera (2009). "Combining site and regional flood information using a Bayesian Monte Carlo approach". In: *Water Resources Research* 45.4. ISSN: 1944-7973. DOI: 10.1029/2008WR007173.

Milly, P. C. D., J. Betancourt, M. Falkenmark, R. M. Hirsch, Z. W. Kundzewicz, D. P. Lettenmaier, and R. J. Stouffer (2008). "Stationarity Is Dead: Whither Water Management?" In: *Science* 319.5863, pp. 573–574. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.1151915.

Morlot, T., C. Perret, A.-C. Favre, and J. Jalbert (2014). "Dynamic rating curve assessment for hydrometric stations and computation of the associated uncertainties: Quality and station management indicators". In: *Journal of Hydrology* 517, pp. 173–186. ISSN: 00221694. DOI: 10.1016/j.jhydrol.2014.05.007.

Nathan, R. J. and T. A. McMahon (1990). "Evaluation of automated techniques for base flow and recession analyses". In: *Water Resources Research* 26.7, pp. 1465–1473. ISSN: 1944-7973. DOI: 10.1029/WR026i007p01465.

Naulet, R., M. Lang, T. B. Ouarda, D. Coeur, B. Bobée, A. Recking, and D. Moussay (2005). "Flood frequency analysis on the Ardèche river using French documentary sources from the last two centuries". In: *Journal of Hydrology* 313.1, pp. 58–78. ISSN: 00221694. DOI: 10.1016/j.jhydrol.2005.02.011.

Neppel, L., B. Renard, M. Lang, P.-A. Ayral, D. Coeur, E. Gaume, N. Jacob, O. Payrastre, K. Pobanz, and F. Vinet (2010). "Flood frequency analysis using historical data: accounting for random and systematic errors". In: *Hydrological Sciences Journal* 55.2, pp. 192–208. ISSN: 0262-6667, 2150-3435. DOI: 10.1080/02626660903546092.

Ouarda, T. B. M. J., C. Girard, G. S. Cavadias, and B. Bobée (2001). "Regional flood frequency estimation with canonical correlation analysis". In: *Journal of Hydrology* 254.1, pp. 157–173. ISSN: 0022-1694. DOI: 10.1016/S0022-1694(01)00488-7.

Payrastre, O., E. Gaume, and H. Andrieu (2011). "Usefulness of historical information for flood frequency analyses: Developments based on a case study". In: *Water Resources Research* 47.8. ISSN: 00431397. DOI: 10.1029/2010WR009812.

Petersen-Øverleir, A. and T. Reitan (2005). "Uncertainty in flood discharges from urban and small rural catchments due to inaccurate head measurement". In: *Hydrology Research* 36.3, pp. 245–257. ISSN: 0029-1277, 2224-7955. DOI: 10.2166/nh.2005.0018.

Petersen-Øverleir, A. and T. Reitan (2009). "Accounting for rating curve imprecision in flood frequency analysis using likelihood-based methods". In: *Journal of Hydrology* 366.1, pp. 89–100. ISSN: 00221694. DOI: 10.1016/j.jhydrol.2008.12.014.

Petersen-Øverleir, A., A. Soot, and T. Reitan (2009). "Bayesian Rating Curve Inference as a Streamflow Data Quality Assessment Tool". In: *Water Resources Management* 23.9, pp. 1835–1842. ISSN: 0920-4741, 1573-1650. DOI: 10.1007/s11269-008-9354-5.

Pichard, G., G. Arnaud-Fassetta, V. Moron, and E. Roucaute (2017). "Hydro-climatology of the Lower Rhône Valley: historical flood reconstruction (AD 1300–2000) based on documentary and instrumental sources". In: *Hydrological Sciences Journal* 62.11, pp. 1772–1795. ISSN: 0262-6667, 2150-3435. DOI: 10.1080/02626667.2017.1349314.

Pichard, G. (2013). *Hauteurs et altitudes aux échelles du bas Rhône*, p. 48.

Pichard, G. and E. Roucaute (2014). "Sept siècles d'histoire hydroclimatique du Rhône d'Orange à la mer (1300-2000). Climat, crues, inondations." In: *Presses Universitaires de Provence* (Hors-série de la revue Méditerranée), p. 194. DOI: https://doi.org/10.4000/geocarrefour.9491.

Ponts\&Chaussées (1914). *Observations hydrométriques du Rhône à Beaucaire, 1914*.

Puechberty, R., C. Perret, S. P. Pitsch, P. Battaglia, A. Belleville, P. Bompart, G. Chauvel, J. Cousseau, G. Dramais, G. Glaziou, A. Hauet, S. Helouin, M. Lang, F. Larrarte, J. L. Coz, P. Marchand, P. Moquet, O. Payrastre, P. Pierrefeu, and G. Rauzy (2017). *Charte qualité de l'hydrométrie. Guide de bonnes pratiques.* Ministère de l'environnement, de l'énergie et de la mer, France. 86 pp.

Rantz, S. E. (1982). *Measurement and computation of streamflow.* USGS Numbered Series 2175. U.S. G.P.O.

Renard, B., K. Kochanek, M. Lang, F. Garavaglia, E. Paquet, L. Neppel, K. Najib, J. Carreau, P. Arnaud, Y. Aubert, F. Borchi, J.-M. Soubeyroux, S. Jourdain, J.-M. Veysseire, E. Sauquet, T. Cipriani, and A. Auffray (2013). "Data-based comparison of frequency analysis methods: A general framework". In: *Water Resources Research* 49.2, pp. 825–843. ISSN: 00431397. DOI: 10.1002/wrcr.20087.

Renard, B. (2023). "Use of a national flood mark database to estimate flood hazard in the distant past". In: *Hydrological Sciences Journal* 0 (ja), null. ISSN: 0262-6667. DOI: 10.1080/02626667.2023.2212165.

Ribatet, M., E. Sauquet, J.-M. Grésillon, and T. B. M. J. Ouarda (2007). "A regional Bayesian POT model for flood frequency analysis". In: *Stochastic Environmental Research and Risk Assessment* 21.4, pp. 327–339. ISSN: 1436-3240, 1436-3259. DOI: 10.1007/s00477-006-0068-z.

Rigaudière, P., A. Dekergariou, C. Laroche, N. Caze, J. Brun, J. Laborde, and W. Berolo (2000). *Etude globale des crues du Rhône.* Tech. rep. SAFEGE Cetiis and Université de Nice.

Ruiz-Bellet, J. L., X. Castelltort, J. C. Balasch, and J. Tuset (2017). "Uncertainty of the peak flow reconstruction of the 1907 flood in the Ebro River in Xerta (NE Iberian Peninsula)". In: *Journal of Hydrology* 545, pp. 339–354. ISSN: 00221694. DOI: 10.1016/j.jhydrol.2016.12.041.

Salas, J. D., J. Obeysekera, and R. M. Vogel (2018). "Techniques for assessing water infrastructure for nonstationary extreme events: a review". In: *Hydrological Sciences Journal* 63.3, pp. 325–352. ISSN: 0262-6667. DOI: 10.1080/02626667.2018.1426858.

Sheffer, N. A., Y. Enzel, G. Benito, T. Grodek, N. Poart, M. Lang, R. Naulet, and D. Cœur (2003). "Paleofloods and historical floods of the Ardèche River, France". In: *Water Resources Research* 39.12. ISSN: 1944-7973. DOI: 10.1029/2003WR002468.

Steinbakk, G. H., T. L. Thorarinsdottir, T. Reitan, L. Schlichting, S. Hølleland, and K. Engeland (2016). "Propagation of rating curve uncertainty in design flood estimation". In: *Water Resources Research* 52.9, pp. 6897–6915. ISSN: 00431397. DOI: 10.1002/2015WR018516.

SYMADREM (2012). *Programme de sécurisation des ouvrages de protection contre les crues du Rhône du barrage de Vallabrègues à la mer*, p. 490.

Tallaksen, L. M. (1995). "A review of baseflow recession analysis". In: *Journal of Hydrology* 165.1, pp. 349–370. ISSN: 0022-1694. DOI: 10.1016/0022-1694(94)02540-R.

Van Der Made, J. (1982). “Determination of the accuracy of water level observations”. In: *IAHS Publications*. Proceedings of the Exeter Symposium 134, pp. 172–184.

Van der Meulen, B., A. Bomers, K. M. Cohen, and H. Middelkoop (2021). “Late Holocene flood magnitudes in the Lower Rhine river valley and upper delta resolved by a two-dimensional hydraulic modelling approach”. In: *Earth Surface Processes and Landforms* 46.4, pp. 853–868. ISSN: 0197-9337, 1096-9837. DOI: 10.1002/esp.5071.

Vieira, L. M. d. S., J. C. L. Sampaio, V. A. F. Costa, and J. C. Eleutério (2022). “Assessing the effects of rating curve uncertainty in flood frequency analysis”. In: *RBRH* 27, e11. ISSN: 2318-0331, 1414-381X. DOI: 10.1590/2318-0331.272220220012.

Vogel, R. M. and C. N. Kroll (1996). “Estimation of baseflow recession constants”. In: *Water Resources Management* 10.4, pp. 303–320. ISSN: 1573-1650. DOI: 10.1007/BF00508898.

Westerberg, I., J.-L. Guerrero, J. Seibert, K. J. Beven, and S. Halldin (2011). “Stage-discharge uncertainty derived with a non-stationary rating curve in the Choluteca River, Honduras”. In: *Hydrological Processes* 25.4, pp. 603–613. ISSN: 08856087. DOI: 10.1002/hyp.7848.

Wilhelm, B., B. Amann, J. P. Corella, W. Rapuc, C. Giguet-Covex, B. Merz, and E. Støren (2022). “Reconstructing Paleoflood Occurrence and Magnitude from Lake Sediments”. In: *Quaternary* 5.1, p. 9. ISSN: 2571-550X. DOI: 10.3390/quat5010009.