Uncertainties of flood frequency estimation approaches based on continuous simulation using data resampling.

P. Arnaud, P. Cantet, J. Odry

1   **Uncertainties of flood frequency estimation approaches based on continuous simulation using data**

2   **resampling.**

3

4   Patrick ARNAUD [a,*], Philippe CANTET [a], Jean ODRY [a]

5   [a] Irstea – Centre d'Aix-en-Provence - 3275 Route Cézanne, 13100 Aix-en-Provence - FRANCE

6   [*] Corresponding author

7   patrick.arnaud@irstea.fr[*], philippe.cantet@irstea.fr, jean.odry@irstea.fr

8

9   **Abstract**: Flood frequency analyses (FFAs) are needed for flood risk management. Many methods exist

10   ranging from classical purely statistical approaches to more complex approaches based on process

11   simulation. The results of these methods are associated with uncertainties that are sometimes difficult to

12   estimate due to the complexity of the approaches or the number of parameters, especially for process

13   simulation. This is the case of the simulation-based FFA approach called SHYREG presented in this paper,

14   in which a rainfall generator is coupled with a simple rainfall-runoff model in an attempt to estimate the

15   uncertainties due to the estimation of the seven parameters needed to estimate flood frequencies. The six

16   parameters of the rainfall generator are mean values, so their theoretical distribution is known and can be

17   used to estimate the generator uncertainties. In contrast, the theoretical distribution of the single hydrological

18   model parameter is unknown; consequently, a bootstrap method is applied to estimate the calibration

19   uncertainties. The propagation of uncertainty from the rainfall generator to the hydrological model is also

20   taken into account. This method is applied to 1112 basins throughout France. Uncertainties coming from the

21   SHYREG method and from purely statistical approaches are compared, and the results are discussed

22   according to the length of the recorded observations, basin size and basin location. Uncertainties of the

23   SHYREG method decrease as the basin size increases or as the length of the recorded flow increases.

24   Moreover, the results show that the confidence intervals of the SHYREG method are relatively small despite

25   the complexity of the method and the number of parameters (seven). This is due to the stability of the

26   parameters and takes into account the dependence of uncertainties due to the rainfall model and the

27   hydrological calibration. Indeed, the uncertainties on the flow quantiles are on the same order of magnitude

28   as those associated with the use of a statistical law with two parameters (here generalised extreme value

1   Type I distribution) and clearly lower than those associated with the use of a three-parameter law (here

2   generalised extreme value Type II distribution). For extreme flood quantiles, the uncertainties are mostly due

3   to the rainfall generator because of the progressive saturation of the hydrological model.

4

5   **Keywords**: flood frequency estimation, uncertainty, flood hazard, stochastic model, bootstrap.

6

7   **1.   Introduction**

8   Flood frequency estimation is essential for flood management. It is used to map floodplain areas, design

9   hydraulic structures (dams, retaining basins, storm water systems) and infrastructures (roads, bridges), and

10   define the frequency of flood events for natural disaster assessments and alert methods (Javelle et al., 2010).

11   The study of hydrological hazard uses flood frequency analysis (FFA) and has led to the development of

12   various methods ranging from purely statistical approaches to simulation approaches. The development of

13   these methods is often influenced by the availability of observation data and by the objectives to be met

14   (Boughton and Droop, 2003; Castellarin et al., 2012; Pathiraja et al., 2012).

15   Purely statistical methods are used to directly fit a probability distribution to the empirical frequency

16   distribution of the hydrological variable studied. The choice of probability distributions used to estimate

17   flood flows is based on the extreme value theory (Coles, 2001). In FFA, common probability distributions

18   are the generalised extreme value (GEV) distribution (Hosking and Wallis, 1993), the generalised Pareto

19   distribution and the three-parameter lognormal distribution. To extend the series length and improve extreme

20   flood frequency estimation, several methods consider historical flood data and sometimes paleo-flood

21   information (Gaume et al., 2010; Payrastre et al., 2011; Stedinger and Cohn, 1986). To study ungauged or

22   poorly gauged sites, a regional flood frequency analysis (RFFA) is used (Darlymple, 1960; Hosking and

23   Wallis, 1993; Hosking and Wallis, 1997; Merz and Blöschl, 2005; Ribatet et al., 2007; Stedinger and Tasker,

24   1985).

25   Due to the nonlinearity of hydrological processes, extrapolating frequency distributions to extreme values

26   remains problematic (Katz et al., 2002). Indeed, calibrating a model based on common observations does not

27   guarantee extrapolation to extreme values. That is why some purely statistical methods rely on estimation of

28   rainfall frequency to extrapolate flow probability distribution (Guillot and Duband, 1967; Margoum et al.,

1   1994). This has resulted in several simulation approaches emerging over the last decade. Such approaches

2   mimic some of the statistical properties of rainfall observations and the rainfall–runoff relationship in order

3   to generate rainfall and runoff series that can be used subsequently as observed series. At the beginning, they

4   were developed especially to fulfil the temporal data requirements associated with design floods (Eagleson,

5   1972). Nowadays, these methods are widespread and tend to be used in the FFA (Li et al., 2014). Models

6   differ in the type of rainfall generator and/or rainfall-runoff model used (Blazkova and Beven, 2004; Onof et

7   al., 2005; Shen et al., 1990), a summary of which is presented in (Boughton and Droop, 2003). In France,

8   there are two different simulation approaches, called the SCHADEX (Simulation Climato-Hydrologique

9   pour l'Appréciation des Débits EXtrêmes) approach developed by Electricité de France (EdF) (Paquet et al.,

10 2013) and the SHYREG (Simulation of Hydrographs for flood frequency estimation – REGionalized)

11 approach developed by Irstea (Arnaud and Lavabre, 2002; Aubert et al., 2014).

12 The SHYREG method and other more classical FFA or RFFA methods were evaluated in a national French

13 research project (national Extraflo project, https://extraflo.cemagref.fr) in which different training/testing

14 decompositions on a set of more than 1,000 French gauging stations are used and reliability and stability

15 indexes are proposed to assess the results (Renard et al., 2013). On one hand this study showed that RFFA

16 approaches are preferable to FFA approaches (Carreau et al., 2013; Kochanek et al., 2014). On the other

17 hand, methods based on the exploitation of rainfall data are preferable for estimating extreme floods

18 (Kochanek et al., 2014). Indeed, during extreme rainfall events, the soil becomes saturated and the behaviour

19 of flood levels mainly depends on the duration and intensity of maximum rainfall. In this comparison project,

20 the SHYREG method was found to be the most stable (ability to yield similar estimates when different data

21 are used for calibration) and one of the most reliable methods tested (Arnaud et al., 2016). Moreover, this

22 method provides realistic flood scenarios directly used for dam design (Carvajal et al., 2009) and takes into

23 account the non-stationarity of phenomena much better (Cantet et al., 2011). These advantages make it a

24 very valuable method, but knowledge of its uncertainties is required.

25

26 FFA or RFFA, assessed by any method, is subject to uncertainties that need to be estimated. These

27 uncertainties stem from the finite sample size of the data set from which the methods were calibrated, which

28 are usually illustrated by confidence intervals. Moreover, the nonlinearity of processes leads to greater

1   uncertainties towards extreme values. Indeed, the observed data, used to calibrate the rainfall model and the

2   hydrological model, do not necessarily correspond to extreme phenomena. Moreover, the observation of

3   extreme peak floods presents greater uncertainties (Ruiz-Bellet et al., 2017) that can increase the

4   uncertainties on flood quantile estimation (Mirzaei et al., 2015). For purely statistical methods, the

5   estimation of confidence intervals can be easily obtained assuming stationarity (which may not be the case in

6   a changing climate). Different approaches can be used, ranging from statistical development based on

7   asymptotic formulae to simulation methods based on resampling processes. For example, confidence

8   intervals of the generalised extreme value distribution can be estimated by formulae depending on the

9   specific parameter estimation techniques (Dupuis and Field, 1998), using the profile-likelihood approach

10   (Coles, 2001), bayesian approaches (Coles et al., 2003) or bootstrap approaches (Burn, 2003; Schendel and

11   Thongwichian, 2015). To reduce uncertainties due to sampling data or shortness of recorded length flow,

12   RFFA methods or partial duration series can be used (Gado and Nguyen, 2016). In the context of non-

13   stationarity, methods can also be employed to take into account this non-stationarity while estimating the

14   associated uncertainties (Sraj et al., 2016).

15   Resampling approaches, also called bootstrap approaches, are nonparametric estimation approaches that

16   present the advantage of avoiding making assumptions on the asymptotic behaviour of distribution and are

17   easy to implement. They involve creating random samples from the original data sample using a

18   bootstrapping process which entails randomly selecting data with replacements (Efron, 1979). This approach

19   was employed for uncertainty analysis of rainfall frequency estimation (Dunn, 2001; Overeem et al., 2008),

20   flood frequency estimation (Burn, 2003; Mirzaei et al., 2015) and flood forecasting (Sehgal et al., 2014;

21   Tiwari and Chatterjee, 2010).

22   When a simulation method is used, estimating uncertainties appears difficult because it involves more

23   complex models, sometimes a combination of models, often with many parameters. Increasing the

24   complexity of the model and the number of parameters can increase the number of degrees of freedom and

25   therefore the sources of uncertainties. In this case, Bayesian approaches are adapted to take into account both

26   observation uncertainties and the imperfect model structure. These methods combine a priori knowledge on

27   parameters with a likelihood function of data to provide an a posteriori distribution of parameter sets. Then

28   using simulation processes with sampled parameters, for example based on the Markov Chain Monte Carlo

4

1   (MCMC) method, uncertainty analysis can be expressed through the posterior results. This is the base of the

2   generalised likelihood uncertainty estimation (GLUE) introduced by Beven and Binley (1992) and widely

3   applied to stochastic rainfall models (Cameron et al., 2000; Cameron et al., 1999) and hydrological models

4   (Beven and Binley, 2014; Beven and Freer, 2001; Freer et al., 1996; Montanari, 2007; Stedinger et al., 2008).

5   These approaches have also been tested on the stochastic rainfall model presented in this study (Muller et al.,

6   2009).

7

8   This paper focuses on the estimation of the uncertainties of the SHYREG simulation-based method with a

9   bootstrap approach. This method was developed by our team and we performed the multiple calibrations

10   implemented by the resampling method on more than 1000 basins in France. The complexity of this method

11   tends to suggest that it would lead to estimated quantiles with substantial uncertainty. For comparison, the

12   uncertainties obtained under the same sampling conditions by fitting classical probability distributions were

13   also investigated. Here distributions from the GEV family were adopted. Since the objective was not to

14   assess the relative merits of different types of distribution, a single family was employed. Instead, the impact

15   of the number of degrees of freedom was studied by fitting both a GEV Type II distribution (three

16   parameters) and a GEV Type I distribution (also called Gumbel distribution, with only two parameters). A

17   resampling approach was chosen to assess the uncertainties of flood quantiles estimated with the SHYREG

18   simulation method. This study focused only on the uncertainties stemming from the sampling of calibration

19   data and their impact on the calibration of the method's parameters. We focus on parameter uncertainty

20   estimation for the rainfall model and on sampling uncertainty for hydrological modelling. These

21   uncertainties created in both the rainfall and the hydrological models separately were investigated, as well as

22   their propagation from one model to the other. Finally, the amplitude of the confidence intervals regarding

23   the time-step, the size of the drainage area, the specific flows and the location of the catchments was also

24   studied.

25

26   **2.   Methodology**

27

28   2.1. *The SHYREG method*

1

2   The SHYREG method is the regionalised version of the simulation-based SHYPRE (Simulation of

3   HYdrograph for flood PREdiction) method. The SHYPRE method was initially developed to simulate flood

4   scenarios on a given catchment. It combines a stochastic hourly rainfall model with a conceptual rainfall-

5   runoff model (Arnaud and Lavabre, 2002; Cernesson et al., 1996). This event-based approach focuses on

6   analysis and modelling of hourly rainy events. A rainy event is defined as a series of daily rainfalls greater

7   than 4 mm among which one exceeds at least 20 mm, but it is analysed at an hourly time step. In France, the

8   number of such events was mapped and varies between three and 25 events per year. Cantet and Arnaud

9   (2014) showed that the rainfall generator succeeds in simulating extreme rainfall of different durations

10   (ranging from 1 h to 72 h). For example, to generate 1,000,000 years of flood events, we generate the

11   number of events per year for each year (using the Poisson distribution law) and the associated independent

12   rainfall events. These are transformed into flood events, which are associated with a simulation of a

13   1,000,000-year period. The method thus generates a set of independent flood hydrographs from which

14   empirical flood quantiles are estimated, for different durations and different return periods, without fitting a

15   theoretical probability law. To ensure the stability of the empirical frequencies, the simulation periods are

16   1000 times longer than the longest return period of interest (Cantet and Arnaud, 2014).

17   The SHYREG method was developed after SHYPRE and is based on the same principle. Its objective is the

18   regional determination of extreme floods. Its implementation requires the regionalisation of both the hourly

19   rainfall generator and the rainfall-runoff model.

20   First the rainfall model is parameterised by 20 parameters estimated for two seasons (June to November and

21   December to May) to take into account a typology of rainfall evens. For each season, these 20 parameters,

22   estimated from hourly data, are regionalised (Arnaud and Lavabre, 2010). During this step, 15 of these

23   parameters are set to a uniform value because they are not very important or not variable in France. The

24   other five are determined directly by linear regressions with three parameters derived from daily data, which

25   are more widespread than hourly data (Arnaud et al., 2008). Consequently, the regionalisation of the hourly

26   rainfall generator parameters is based on the mapping of three daily rainfall parameters. These parameters

27   ($\theta_{i=1\ to\ 6}$) are related to rainfall characteristics of intensity, duration and frequency (see Table 1) and they are

28   determined and mapped using 2,812 rain gauge stations located throughout France, taking into account local

6

1   environmental and topographical characteristics. The regionalised parameters are used to parameterise the

2   hourly rainfall generator and make it possible to simulate at-site hourly rainfall time series. These hourly

3   rainfall time series lead to the empirical estimation of rainfall quantiles for different time durations on a 1-

4   km² square anywhere in France (Arnaud et al., 2008).

5

Parameter	Description	Unit	Range over France or value if imposed.
Rainfall generator parameters			
θ_1 and θ_2	Mean number of rainfall events per year (for two seasons)	N/A	From 1.5 to 15 for each season
θ_3 and θ_4	Mean of the events' maximum daily rainfall (for two seasons)	mm	From 20 to 70
θ_5 and θ_6	Mean of the events' duration (for two seasons)	day	From 1.6 to 3.5
Elementary rainfall-runoff model			
A (imposed)	Maximum capacity of production reservoir: imposed by hydrogeological classes	mm	From 200 to 600
S_0/A	Initial recharge level rate of A reservoir: homogeneous over a basin but optimised for each basin.	N/A	From 0 to 1
B (imposed)	Maximum capacity of routing reservoir: imposed. Its initial level is also imposed at $0,3 \times B$	mm	50 for summer 100 for winter
Areal reduction factor			
$\alpha_1, \beta_1, \gamma_1$ (imposed)	Coefficient of areal reduction function for flows lasting more than 24 h	N/A	$0.01 - 0.24 - 5$
$\alpha_2, \beta_2, \gamma_2$ (imposed)	Coefficient of areal reduction function for flows lasting less than 24 h	N/A	$0.25 - 10 - 0.9$

1   *Table 1: List and description of the method's parameters.*

2

3   Secondly, hourly rainfall events are transformed into flood events by a simplified rainfall-runoff model. The

4   rainfall-runoff model is simplified to allow its regionalisation and to be able to convert hourly rainfall into

5   flood flow at a 1-km² pixel resolution. The transformation to the pixel resolution is necessary because of the

6   at-site nature of the rainfall generator (this is not a rainfall field generator). This elementary hydrological

7   model is the Irstea GR type (www://cemagref.fr/webgr/). It consists of a production reservoir with a capacity

8   A (mm), a routing reservoir with a capacity B (mm) and a 2-h unit hydrograph. This elementary rainfall-

9   runoff model is used in event mode to convert the hourly rainfall scenarios into flood scenarios at the pixel

10   scale (Arnaud et al., 2016; Arnaud et al., 2011). To simplify the model, the A and B parameters are imposed

11   on all French watersheds (citer Yoann) (see Table 1 for further detail). Only the first reservoir's initial

12   recharge level, noted S_0/A, is needed to calibrate the rainfall-runoff transformation. This rate controls the

13   runoff and evolves during the rainy event. A value of (0) corresponds to an initially empty reservoir. A value

14   of (1) means an initially full (or saturated) reservoir and corresponds to a situation in which all the rain water

15   will flow. Between these two values, at each time step, the rain water is shared in two parts: the first one will

16   fill the production tank and the second will contribute to the runoff. The value of S_0/A is the same for all

17   pixels of a given watershed, except those covered by urbanised areas where a value is imposed.

18   Consequently, flood events are simulated for the entire range of possible values for S_0/A, at each pixel. The

19   flood quantiles are extracted empirically from these simulated events for each square kilometre. These

20   'elementary flows' are noted $q_{d,T}$. To estimate river flow quantiles, elementary flow quantiles are aggregated

21   (summed) from all the n pixels in the catchment. Then an areal reduction factor is applied to simultaneously

22   take into account both rainfall areal reduction and flood routing. This factor only depends on the catchment

23   area (S) and the duration of flow (d) and is represented by equations (1), (2) and (3).

24   $Q(d \geq 24\text{h}, T) = \sum_{i=1}^{n} q_{d,T} \cdot f_1(S)$ (1)

25

26   $Q(d < 24h, T) = Q(24\text{h}, T) + \left[\sum_{i=1}^{n} q_{d,T} - \sum_{i=1}^{n} q_{d=24\text{h},T}\right] \cdot f_2(S)$ (2)

27

1   with the term $f_i(S) = \frac{1+\beta_i}{1+\beta_i \cdot S^{0,5 \cdot \gamma_i \cdot (1-S^{\alpha_i})}}$ (3)

2   where S is catchment area in km² and $Q(d, T)$ is the flow quantile (of duration d and return period T)

3   calculated at the catchment outlet. Parameters α_i, β_i and γ_i are assumed constant over metropolitan France

4   and were estimated in a preliminary study with calibration data (Aubert et al., 2014).

5   Then calibrating the SHYREG method consists in determining which value of S_0/A makes it possible to

6   properly simulate flood events statistically equivalent to the observed events at gauged stations. The

7   optimised parameter is the value that minimises the average absolute relative error between six quantiles

8   obtained from observations (peak flows and daily runoff for 2-, 5- and 10-year return periods) and the same

9   six quantiles provided by the SHYREG method. The quantiles from observations are estimated by fitting a

10   GEV probability distribution. A lower bound for the shape coefficient is imposed at 0, which is equivalent to

11   bounding the GEV distribution by a Gumbel distribution. An upper bound is also imposed at 0.4, in order to

12   limit the influence of extreme values when fitting. Note that the choice of probability distribution is

13   relatively insignificant as long as observed frequencies (T<10 years) are used. This probability distribution is

14   here called 'bounded' GEV. Therefore, for each gauged catchment, the SHYREG method can be calibrated

15   by optimising a single parameter, on which the regionalisation process will be based to apply the method

16   over the entire drainage network (including ungauged catchments) (Organde et al., 2013).

17   Overall, only six parameters are needed to simulate rainfall (actually three parameters for two seasons), and

18   an additional one is necessary to perform the rainfall-runoff transformation.

19   Note that currently the areal reduction factor is not parameterised by the slope or the land use. This point is

20   actually the topic of an on-going study but is not within the scope of the present paper. The calibration of the

21   S_0/A parameter partially takes into account a part of the flow variability which may be related to the slope of

22   the basin or the land use.

23   The study presented here concerns only the uncertainties generated at the calibration stage of the SHYREG

24   method, and therefore the uncertainties on the estimation of these seven parameters.

25

26   *2.2. Uncertainty estimation methodology*

27

1   This paper focuses on the uncertainties due to calibration against observations. These observations are

2   rainfall time series used to calibrate the set of parameters of the rainfall generator (noted θ) and flow time

3   series (or by default the annual maximum floods from river gauges with sufficiently long time series to

4   estimate up to 10-year quantiles) to calibrate the hydrological parameter (S_0/A). The uncertainties of the

5   method result partially from uncertainties bound to the estimation of both sets of parameters (θ and S_0/A)

6   associated with both implemented models (the rainfall generator and the hydrological model). Other

7   parameters of the method are assumed to be known throughout the study area (French territory). This is why

8   the uncertainties related to these parameters are not calculated in this study. The uncertainty bound to the

9   calibration of the models' parameters were considered in different ways, as explained below. In the

10   following, we note 'observed' for the parameter estimated with observed data values and 'resampled' for the

11   parameter estimated with resampled data values.

12

13   2.2.1. About the hourly rainfall generator

14   For the rainfall model, we studied the uncertainties due to parameter estimation. Indeed, the set of six

15   parameters ($\theta_{observed_i}$ for i = 1–6) was obtained by calculating six averaged values (three variables for two

16   seasons) over the observed data (see section 3). These average values are calculated over a large number of

17   values (n). Indeed, parameters θ_1 and θ_2 (see Table 1) were estimated by a set of at least 25 values,

18   corresponding to the number of years of observation. The others (θ_3 to θ_6) are estimated on a number of

19   values ranging from 37 to more than 300, corresponding to the product of the number of events per season

20   by the number of years. Thus, each rainfall parameter follows the distribution of an average calculated on a

21   large number of values, which corresponds to a Gaussian distribution if we assume the Laplace central limit

22   theorem can be applied. This Gauss distribution has a mean equal to the average of the variable (μ), and a

23   standard deviation is equal to the standard deviation of the variable (σ) reduced by the square root of the

24   sample size (n). So we can consider that the six-parameter set $\theta_{i=1\,to\,6} \sim \mathcal{N}\left(\mu_i, \sigma_i/\sqrt{n_i}\right)$. According to the

25   Gauss distributions, different sets of parameters ($\theta_{resampled}$) can then be generated. These sets allow one to

26   run the rainfall generator and thus to obtain different frequency distributions on rainfall representing the

27   uncertainty on the calibration stage of the rainfall generator. A previous study was conducted using the data

1   of a single station, based on frequency and Bayesian analysis (Muller et al., 2009). It is assumed that the six

2   parameters are independent and the parameter resampling is processed by independent random selections in

3   the parameter distributions.

4

5   2.2.2. About the rainfall-runoff model

6   For the hydrological model, we studied the uncertainties due to observation sampling. As noted above, the

7   hydrological parameter is calibrated to respect the current flood quantiles (2-, 5- and 10-year return periods

8   of peak flow and daily flow) estimated by fitting a probability distribution on maximum flow values. This

9   parameter is noted $S_0/A_{observed}$. Contrary to the rainfall parameter, the frequency distribution of this

10 parameter is unknown beforehand. To estimate its variability, a bootstrap procedure was employed. The

11 parameter was calibrated for various samples of the observations. Note that the annual maximum values of

12 observed floods were assumed to be independent and identically distributed.

13 The sampled parameter (note $S_0/A_{resampled}$) was then assessed following these steps:

14 -   For a basin on which n is the number of observed annual maximum values, a new sample of n values

15   was generated by a random sampling with replacements.

16 -   A 'bounded' GEV distribution was fitted on this new sample to determine the quantiles of 2-, 5- and 10-

17   year return periods. These quantiles were used to estimate the parameter $S_0/A_{resampled}$ as well as to

18   evaluate the confidence interval of the 'bounded' GEV distribution for higher return periods (up to 1000

19   years). Other distributions, such as the Gumbel distribution, were fitted on this sample to be compared

20   later to GEV and SHYREG values.

21 -   The calibrated parameter $S_0/A_{resampled}$ $S_0/A_{resampled}$ allowed estimating quantiles n from 2 to 1000

22   years using the SHYREG method.

23 -   The above points were repeated NB times (NB is the number of samplings carried out using the

24   bootstrap method).

25 -   From these NB distributions, a confidence interval and the median distribution were calculated.

11

1   Thus, both procedures determined the uncertainties associated with the estimation of the method's

2   parameters. These uncertainties on the parameters were directly expressed in terms of uncertainties in rainfall

3   and flow quantiles provided by the SHYREG method.

4

5   *2.3. Test configuration*

6

7   In this paper, uncertainties were derived from both the rainfall generator and hydrological modelling.

8   Various tests were carried out to study the influence of each cause of uncertainty on the overall uncertainties

9   of the model output. These cases are presented below and illustrated in Figure 1:

10

11   *Figure 1: Configuration of calibration / simulation cases for the SHYRE method.*

12

13   • **Case 1:** This case evaluated only the uncertainty due to the rainfall information by providing

14   distributions of rainfall along with a confidence interval. It also assessed the propagation of this

15   uncertainty on the flow frequency distributions without changing the parameterisation of the

16   hydrological model.

17   • **Case 2:** This case evaluated only the uncertainty due to the flow information, always considering the

18   same source of information for rainfall. We obtained confidence intervals only on the flow frequency

19   distributions.

20   • **Case 3**: This case evaluated the uncertainties due to rainfall and flow observations, considering them

21   independent. Rainfall events were generated from a set of parameters ($\theta_{resampled}$), and then they

22   were converted into flows by a parameter $S_0/A_{resampled\_1}$, which was calibrated from the resampled

23   observed flow with the non-resampled rainfall (generated with $\theta_{observed}$). In this case, independence

24   between the uncertainties of the two models was assumed.

25   • **Case 4:** Contrary to case 3, the parameter $S_0/A_{resampled\_2}$ was determined by taking into account

26   the rainfall generated with the new parameterisation of the rainfall generator ($\theta_{resampled}$). This case

27   assumed that there was a dependency between the rainfall (and its resampling) and the calibration of

1 the rainfall–runoff relationship. This means that the uncertainty caused by the "resampled" rainfall

2 can be balanced by the calibration of the hydrological model. This was already the case when the

3 calibration was performed using the (not resampled) observed rainfall data.

4

5 **3. Data**

6 The SAFRAN data (Quintana-Seguí et al., 2008; Vidal et al., 2010) were used for rainfall to estimate the

7 mean and the standard deviation for the different parameters of the rainfall generator and the sample size on

8 which parameters were calculated. For this study the recorded length of rainfall was set at 25 years, between

9 1977 and 2002. This observation period is related to our availability of data. The use of a longer observation

10 period would have undoubtedly reduced the uncertainty associated with estimating the rainfall model's

11 parameters. For example, the parameter θ_3 is presented in Figure 3. It represents the spatial distribution of

12 the mean and the reduced standard deviation $(\sigma/\sqrt{n})$ of the events' maximum daily rainfall. We can see that

13 the highest values are located in the Mediterranean climate area. To simplify the simulations, the rainfall

14 generator was run with the values of the parameter estimated at the catchment centroids.

15 This study was conducted on data from gauged basins available in the HYDRO French database

16 (http://www.hydro.eaufrance.fr/). These data correspond to 1112 basins located in Figure 2. Their surface

17 areas range from 1 to 2000 km², with a median of 180 km². The monitoring stations were selected in the

18 regionalisation study of the SHYREG method (Aubert et al., 2014; Organde et al., 2013). The selection of

19 the stations was based on different criteria such as the advice of station managers, the presence of a

20 sufficiently long series (over 20 years), the absence of upstream dams and the goodness-of-fit of a GEV

21 distribution.

22 *Figure 2: Location and characteristics of the 1112 basins studied.*

23 These stations have collected hydrometric information for over 20 years (with a median of 40 years). These

24 data allowed us to fit statistical distributions ('bounded' GEV, Gumbel) to estimate flood quantiles and to

25 calibrate the parameter S_0/A.

26

27 *Figure 3: Example illustrating the variability of sampling characteristics for one rainfall generator*

28 *parameter corresponding to the intensity of a rainy event.*

13

1

2 **4. Results and discussion**

3

4 The four cases presented above were tested by carrying out NB=500 simulations using parameter samplings.

5 Therefore, 500 frequency distributions of rainfall and flows provided by the different methods were

6 obtained. From these distributions, the median ($q_{50\%}$) and 5% and 95% (resp. $q_{5\%}$ and $q_{95\%}$) prediction

7 bounds were calculated for each flood (and rainfall) quantile to determine a 90% confidence interval (noted

8 $CI_{90\%}$).

9 For example, Figure 4 shows (in gray) the different distributions obtained from sampled parameters (only

10 100 resampled distributions), and in colour the median distribution (solid line) and 90% confidence intervals

11 (dashed lines) are represented. The example illustrates the difference between confidence intervals associated

12 with each FFA method and their asymptotic behaviour towards extreme frequencies.

13

14 *Figure 4: Confidence intervals of peak flow quantiles associated with the SHYREG method (taking into*

15 *account different sources of uncertainties; cases 1–4), with the Gumbel distribution and with GEV*

16 *distribution: example of the Ill at Didenheim station.*

17

18 The objective of this paper was to analyse the confidence intervals, according to various factors (number of

19 years of observation used to calibrate the models, impact of the hydrological or rainfall patterns, location of

20 stations, return period studied, etc.). To summarise the information on confidence intervals, the ratio between

21 a 90% confidence interval and the median values was calculated on each station and for different quantiles:

$$CIA = \frac{(q_{95\%} - q_{5\%})}{q_{50\%}}$$

22 This criterion provided an idea of the relative 'amplitude' of the confidence interval of a method for a given

23 quantile. The frequency distribution of this amplitude estimated on the 1112 stations is represented with a

24 boxplot that illustrates the 5, 25, 50, 75 and 95% percentile values (*Figure 5*).

25

26 *4.1. Analysis of the different cases*

14

1

2   The confidence interval amplitude (CIA) was calculated for different variables (rainfall or peak flow) of

3   different return periods (2-, 10-, 100-, 1000-year quantiles) using different methods (SHYREG, Bounded

4   GEV, Gumbel). The boxplots in Figure 5 are associated with rainfall estimated by the SHYREG method (in

5   black), peak flow estimated by the SHYREG method (in blue), peak flow estimated by Gumbel distribution

6   (in green) and peak flow estimated by GEV 'bounded' distribution (in red). For peak flow estimated by the

7   SHYREG method, the different boxplots are related to the four cases tested (as described above).

8

9   *Figure 5: Distribution of the confidence interval amplitude for 2-, 10-, 100- and 1000-year quantiles,*

10   *associated with rainfall estimated by the SHYREG method (in black), peak flow estimated by the SHYREG*

11   *method for cases 1–4 (in blue), peak flow estimated by Gumbel distribution (in green) and peak flow*

12   *estimated by 'bounded' GEV distribution (in red).*

13

14   Figure 5a and 5b present the CIA associated with 2- and 10-year quantiles. These quantiles are close to the

15   empirical distribution because they are in the domain of observation. On the one hand, the uncertainties on

16   rainfall were lower than the flow uncertainties. Among the uncertainties on flows, the uncertainty caused by

17   the sole rainfall observations (case 1) was relatively close to that attributed to only the flow observations

18   (case 2). Logically, the uncertainties due to both rainfall and flow considered independently led to larger

19   confidence intervals (case 3). On the other hand, taking into account the dependency between the two

20   sources of uncertainty reduced the amplitude of the confidence intervals of the SHYREG method estimates

21   (case 4).

22   For low frequencies, the uncertainties associated with SHYREG, the Gumbel distribution and the 'bounded'

23   GEV distribution were on the same order of magnitude. This result was understandable because SHYREG is

24   calibrated against low-frequency quantiles estimated with a 'bounded' GEV distribution and is therefore

25   subject to the same source of uncertainty.

26   For extrapolation towards the rare and extreme frequencies, the results were different (Figure 5c and 5d).

27   Globally, the amplitude of the uncertainties on rainfall remained lower than the flow uncertainties. In the

28   domain of extrapolation, very high uncertainties associated with the use of a three-parameter frequency

15

1 distribution were observed. Using this distribution seems to be poorly adapted in an at-site context, notably

2 because of a very heavy tail of the GEV type II (although 'bounded'). To consolidate the estimate of the

3 GEV distribution shape parameter (3rd parameter), a regional approach (Hosking and Wallis, 1993) may be

4 necessary. This can reduce uncertainty over estimates of extreme values. The use of a GEV distribution in a

5 regional approach should lead to confidence intervals close to those resulting from the use of a Gumbel

6 distribution (since the shape parameter is generally fixed on the area studied; see for example Kochanek

7 2014). We can see that uncertainties associated with the use of the Gumbel distribution (two parameters)

8 were close to those associated with the SHYREG method (seven parameters). Even if the uncertainty

9 amplitudes were similar, they were associated with quantiles that can be completely different. The quantile

10 estimates are considered more accurate when using the SHYREG method. Indeed, the conclusions of the

11 ANR Extraflo project do not recommend using a Gumbel distribution to estimate quantiles of extreme floods

12 (Kochanek et al., 2014).

13 The comparison of different cases associated with the SHYREG method is presented in the blue boxplot

14 (*Figure 5*). We can see from the 10-year return period plot that the quantiles estimated by the SHYREG

15 method were marginally influenced by the calibration of the hydrological parameter, in comparison with the

16 uncertainties due to rainfall (comparison of cases 1 and 2). The gradual saturation of the catchment in the

17 hydrological model towards the rare and extreme events makes the method more sensitive to rainfall

18 information than to observed flows. It can also be observed that consideration of the uncertainties stemming

19 from both rainfall and observed flows must be taken into account in a dependent way because the calibration

20 of the hydrological parameter depends on the information available on rainfall. This consideration reduced

21 uncertainties on the flows (comparison of cases 3 and 4).

22

23 *4.2. Detailed analysis*

24

25 Here, we focused on case 4, which took into account the dependence between the two sources of uncertainty

26 (the observations of both rainfall and flows).

27

1   Figure 6 shows the distribution of uncertainty amplitudes for groups of basins classified according to

2   different criteria: the recorded length of observation (graphs a and b), basin size (graphs c and d) and specific

3   flow (to distinguish more or less productive basins) (graphs e and f, where basins are ranked by increasing

4   specific flow). For each criterion, basins were clustered in order to produce groups of equal size. Only

5   uncertainties calculated on 10- and 1000-year quantiles of peak flows are illustrated.

6

7   *Figure 6: Distribution of amplitudes for groups of basins classified according to the recorded length of*

8   *observation (a, b), basin size (c, d) and S0/A value (to distinguish more or less productive basins) (e, f): case*

9   *4.*

10

11   Regarding the impact of length of flow recorded, uncertainties decreased when the length of observation

12   recorded increased. This result, although trivial, made it possible to quantify the variation of the uncertainty

13   with the length of observation recorded. Note that this impact was more visible on current quantiles than

14   extreme quantiles.

15   Regarding the impact of basin size, uncertainties seemed to be greater for small basins, which may be

16   explained by two points: first, observing the extreme values is more difficult on small basins, which are more

17   randomly affected by extreme rainfall events than large basins, and secondly, it is more difficult to reproduce

18   the rainfall-runoff processes accurately on small basins related to available rainfall data (see e.g. Merz and

19   al., 2009).

20   Regarding the impact of the hydrological characteristics (S0/A values), uncertainties tended to be lower on

21   'productive' basins (high S0/A values). This may be explained by the fact that on these basins, saturation is

22   obtained more quickly, so the uncertainty due to the calibration of the hydrological model has a lower

23   impact. This result was obtained for extreme values (1000-year quantiles). Figure 7 illustrates the spatial

24   distribution of the confidence interval amplitudes, which are calculated for 100-year quantiles of rainfall and

25   peak flows.

26

27   *Figure 7: Location of basins with the respective amplitudes of uncertainties, for 100-year quantiles of*

28   *rainfall and peak flow.*

17

We see that rainfall presented greater uncertainties on the Mediterranean region. The flows were characterised by larger uncertainties than rainfall over the entire country. This spatial distribution was difficult to interpret, however, because it mixed basins of different sizes, different lengths of observation recorded and different hydrological parameter values.

5. Conclusion

The work aimed to evaluate the uncertainties related to a flood frequency estimation method based on simulation processes, the SHYREG method, by comparing them with those of purely statistical approaches (GEV type I and type II distributions). This comparison was made on 1112 basins over France. The simulation processes method provides a flood frequency analysis at the basins' outlet from a regionalized hourly rainfall generator and an elementary hydrological model. Calibration consists in searching the hydrological parameter that minimises error between observed and modelled quantiles up to 10-year floods. This paper proposes a methodology to evaluate uncertainties from both rainfall and flow calibration data sampling. We studied parameter estimation uncertainty for the rainfall model and sampling uncertainty for hydrological modelling. In the latter, we used a bootstrap method to resample the flow observations to determine the variability of the hydrological parameter and their impact on the estimation of the quantiles of rainfall and floods.

The main results showed that:

- It is important to take into account the dependence of uncertainties due to rainfall and flow observations and therefore not to overvalue the final uncertainties.

- The uncertainties on the rainfall quantiles were lower than the uncertainties on the flow quantiles.

- The uncertainties of the SHYREG method decreased as the basin size increased or as the recorded length of flow increased.

- The uncertainties on the SHYREG flow quantiles were on the same order of magnitude as those associated with a two-parameter statistical distribution (Gumbel) and clearly lower than those associated with a three-parameter distribution (GEV type II).

1

2　　The paper brings out a number of advantages of using methods based on simulation processes in FFA, such

3　　as the SHYREG method. Using a hydrological model takes into account the non-linearity of the hydrological

4　　processes: i.e. it models the fact that extreme flows are strongly influenced by extreme rainfall, and thus

5　　reduces the uncertainties of extreme flood quantiles compared to methods based only on the use of flows.

6　　Moreover, the uncertainties due to rainfall can be partially reduced during the calibration of the hydrological

7　　model parameters.

8　　More specifically concerning the SHYREG method, its confidence intervals were relatively low compared to

9　　the complexity of the method (seven parameters). These results can be related to the stability of the methods.

10　Compared to Gumbel distribution, which is considered stable but not accurate, the SHYREG method has

11　shown its qualities in terms of accuracy (Arnaud et al., 2016; Kochanek et al., 2014).

12　The final goal of the SHYREG method is to estimate the flow quantiles on ungauged basins in regionalising

13　the hydrological parameter. Here, we studied only the uncertainties associated with the calibration of the

14　method: i.e. we propose a confidence interval of flow quantiles on the gauged catchments. An ongoing

15　project takes into account the uncertainty of regionalisation to propose flow quantiles and their uncertainties

16　on all rivers in France.

17

18

19

20　**6.　References**

21　Arnaud, P., Cantet, P., Aubert, Y., 2016. Relevance of an at-site flood frequency analysis method for
22　　　extreme events based on stochastic simulation of hourly rainfall. Hydrological Sciences Journal, 61(1):
23　　　36-49.

24　Arnaud, P., Lavabre, J., 2002. Coupled rainfall model and discharge model for flood frequency estimation.
25　　　Water Resour. Res., 38(6): 1075-1085.

26　Arnaud, P., Lavabre, J., 2010. Estimation de l'aléa pluvial en France métropolitaine. QUAE, Paris. Update
27　　　Sciences & Technologies, pp. 158 pages.

28　Arnaud, P., Lavabre, J., Fouchier, C., Diss, S., Javelle, P., 2011. Sensitivity of hydrological models to
29　　　uncertainty in rainfall input. Hydrolog Sci J, 56(3): 397-410. DOI:10.1080/02626667.2011.563742

30　Arnaud, P., Lavabre, J., Sol, B., Desouches, C., 2008. Regionalization of an hourly rainfall generating model
31　　　over metropolitan France for flood hazard estimation. Hydrological Sciences Journal, 53(1): 34-47.

1   Aubert, Y., Arnaud, P., Ribstein, P., Fine, J.-A., 2014. The SHYREG flow method—application to 1605 basins
2   in metropolitan France. Hydrological Sciences Journal, 59(5): 993-1005.
3   DOI:10.1080/02626667.2014.902061

4   Beven, K., Binley, A., 1992. The future of distributed models : model calibration and uncertainty prediction.
5   Hydrological Processes, 6(3): 279-298.

6   Beven, K., Binley, A., 2014. GLUE: 20 years on. Hydrological Processes, 28(24): 5897-5918.
7   DOI:10.1002/hyp.10082

8   Beven, K., Freer, J., 2001. Equifinality, data assimilation, and uncertainty estimation in mechanistic
9   modelling of complex environmental systems using the GLUE methodology. Journal of Hydrology, 249(1-
10   4): 11-29.

11 Blazkova, S., Beven, K., 2004. Flood frequency estimation by continuous simulation of subcatchment
12   rainfalls and discharges with the aim of improving dam safety assessment in a large basin in the Czech
13   Republic. Journal of Hydrology, 292(1-4): 153-172.

14 Boughton, W., Droop, O., 2003. Continuous simulation for design flood estimation—a review. Environ
15   Modell Softw, 18(4): 309-318. DOI:http://dx.doi.org/10.1016/S1364-8152(03)00004-5

16 Burn, D.H., 2003. The use of resampling for estimating confidence intervals for single site and pooled
17   frequency analysis / Utilisation d'un rééchantillonnage pour l'estimation des intervalles de confiance lors
18   d'analyses fréquentielles mono et multi-site. Hydrological Sciences Journal, 48(1): 25-38.
19   DOI:10.1623/hysj.48.1.25.43485

20 Cameron, D., Beven, K., Tawn, J., 2000. Modelling extreme rainfalls using a modified random pulse Bartlett-
21   Lewis stochastic rainfall model (with uncertainty). Adv Water Resour, 24(2): 203-211.

22 Cameron, D.S., Beven, K.J., Tawn, J., Blazkova, S., Naden, P., 1999. Flood frequency estimation by
23   continuous simulation for a gauged upland catchment (with uncertainty). Journal of Hydrology, 219(3–
24   4): 169-187. DOI:http://dx.doi.org/10.1016/S0022-1694(99)00057-8

25 Cantet, P., Arnaud, P., 2014. Extreme rainfall analysis by a stochastic model: impact of the copula choice on
26   the sub-daily rainfall generation. Stochastic Environmental Research and Risk Assessment, 28(6): 1479-
27   1492.

28 Cantet, P., Bacro, J.N., Arnaud, P., 2011. Using a rainfall stochastic generator to detect trends in extreme
29   rainfall. Stochastic Environmental Research and Risk Assessment, 2011(25): 429-441.

30 Carreau, J., Neppel, L., Arnaud, P., Cantet, P., 2013. Extreme rainfall analysis at ungauged sites in the South
31   of France: comparison of three approaches. Journal de la Société Française de Statistique, 154(2): 119-
32   138.

33 Carvajal, C., Peyras, L., Arnaud, P., Royet, P., 2009. Probabilistic Modelling of Flood Water Level for Dam
34   Reservoirs. Journal of Hydrologic Engineering, 14(3): 223-232.

35 Castellarin, A. et al., 2012. Review of applied statistical methods for flood frequency analysis in Europe :
36   WG2 of COST Action ES0901.

37 Cernesson, F., Lavabre, J., Masson, J.M., 1996. Stochastic model for generating hourly hyetographs.
38   Atmospheric Research, 42(1-4): 149-161.

1   Coles, S., 2001. An introduction to Statistical Modeling of Extreme Values. Springer Series in Statistics. ed.
2   Springer-Verlag, Heidelberg, Germany.

3   Coles, S., Pericchi, L.R., Sisson, S., 2003. A fully probabilistic approach to extreme rainfall modeling. Journal
4   of Hydrology, 273(1-4): 35-50.

5   Darlymple, T., 1960. Flood-frequency analysis. Water Supply Pap., 1543A - US Geol Surv.

6   Dunn, P.K., 2001. Bootstrap confidence intervals for predicted rainfall quantiles. Int J Climatol, 21(1): 89-94.
7   DOI:10.1002/joc.596

8   Dupuis, D.J., Field, C.A., 1998. A Comparison of confidence intervals for generalized extreme-value
9   distributions. Journal of Statistical Computation and Simulation, 61(4): 341-360.
10   DOI:10.1080/00949659808811918

11 Eagleson, P.S., 1972. Dynamics of flood frequency. Water Resour. Res., 8(4): 878-898.

12 Efron, B., 1979. Bootstrap methods : Another look at the Jacknife. Annals of Statistics, 7(1): 1-16.

13 Freer, J., Beven, K., Ambroise, B., 1996. Bayesian estimation of uncertainty in runoff prediction and the
14   value of data: An application of the GLUE approach. Water Resour. Res., 32(7): 2161-2173.

15 Gado, T.A., Nguyen, V.T.V., 2016. Regional Estimation of Floods for Ungauged Sites Using Partial Duration
16   Series and Scaling Approach. Journal of Hydrologic Engineering, 21(12). DOI:04016044
17 10.1061/(asce)he.1943-5584.0001439

18 Gaume, E. et al., 2010. Bayesian MCMC approach to regional flood frequency analyses involving
19   extraordinary flood events at ungauged sites. Journal of Hydrology, 394(1–2): 101-117.
20   DOI:http://dx.doi.org/10.1016/j.jhydrol.2010.01.008

21 Guillot, P., Duband, D., 1967. La méthode du Gradex pour le calcul de la probabilité des crues à partir des
22   pluies. I.A.S.H., Publication, 84.

23 Hosking, J.R.M., Wallis, J.R., 1993. Some statistics useful in regional frequency analysis. Water Resour. Res.,
24   29(2): 271-281.

25 Hosking, J.R.M., Wallis, J.R., 1997. Regional frequency analysis: an approach based on L-moments.
26   Cambridge University Press.

27 Javelle, P., Fouchier, C., Arnaud, P., Lavabre, J., 2010. Flash flood warning at ungauged locations using radar
28   rainfall and antecedent soil moisture estimations. Journal of Hydrology, 394(1-2): 267-274. DOI:DOI
29   10.1016/j.jhydrol.2010.03.032

30 Katz, R.W., Parlange, M.B., Naveau, P., 2002. Statistics of extremes in hydrology. Adv Water Resour, 25(8-
31   12): 1287-1304.

32 Kochanek, K. et al., 2014. A data-based comparison of flood frequency analysis methods used in France.
33   Natural Hazards and Earth System Sciences, 14: 295-308. DOI:doi:10.5194/nhess-14-295-2014

34 Li, J., Thyer, M., Lambert, M., Kuczera, G., Metcalfe, A., 2014. An efficient causative event-based approach
35   for deriving the annual flood frequency distribution. Journal of Hydrology, 510(0): 412-423.
36   DOI:http://dx.doi.org/10.1016/j.jhydrol.2013.12.035

1   Margoum, M., Oberlin, G., Lang, M., Weingartner, R., 1994. Estimation des crues rares et extrêmes :
2   principes du modèle AGREGEE. Hydrologie continentale, 9(1): 85-100.

3   Merz, R., Blöschl, G., 2005. Flood frequency regionalisation - Spatial proximity vs. catchment attributes.
4   Journal of Hydrology, 302(1-4): 283-306.

5   Merz, R., J. Parajka, and G. Blöschl (2009), Scale effects in conceptual hydrological modeling, Water Resour.
6   Res., 45, W09405, doi:10.1029/2009WR007872.

7   Mirzaei, M. et al., 2015. Uncertainty analysis for extreme flood events in a semi-arid region. Nat Hazards,
8   78(3): 1947-1960. DOI:10.1007/s11069-015-1812-9

9   Montanari, A., 2007. What do we mean by 'uncertainty'? The need for a consistent wording about
10   uncertainty assessment in hydrology. Hydrological Processes, 21(6): 841-845.

11 Muller, A., Arnaud, P., Lang, M., Lavabre, J., 2009. Uncertainties of extreme rainfall quantiles estimated by a
12   stochastic rainfall model and by a generalized Pareto distribution. Hydrological Sciences Journal, 54(3):
13   417-429.

14 Onof, C., Townend, J., Kee, R., 2005. Comparison of two hourly to 5-min rainfall disaggregators.
15   Atmospheric Research, 77(1-4): 176-187.

16 Organde, D. et al., 2013. Régionalisation d'une méthode de prédétermination de crue sur l'ensemble du
17   territoire français : la méthode SHYREG. Revue des sciences de l'eau, 26(1): 65-78.

18 Overeem, A., Buishand, A., Holleman, I., 2008. Rainfall depth-duration-frequency curves and their
19   uncertainties.   Journal   of   Hydrology,   348(1–2):   124-134.
20   DOI:http://dx.doi.org/10.1016/j.jhydrol.2007.09.044

21 Paquet, E., Garavaglia, F., Garçon, R., Gailhard, J., 2013. The SCHADEX method: A semi-continuous rainfall–
22   runoff simulation for extreme flood estimation. Journal of Hydrology, 495(0): 23-37.
23   DOI:http://dx.doi.org/10.1016/j.jhydrol.2013.04.045

24 Pathiraja, S., Westra, S., Sharma, A., 2012. Why continuous simulation ? The role of antecedent moisture in
25   design flood estimation. Water Resour. Res., 48(W06534): 1-15. DOI:doi:10.1029/2011WR010997

26 Payrastre, O., Gaume, E., Andrieu, H., 2011. Usefulness of historical information for flood frequency
27   analyses: developments based on a case study. Water Resour. Res., 47(W08511).
28   DOI:doi:10.1029/2010WR009812

29 Quintana-Seguí, P. et al., 2008. Analysis of Near-Surface Atmospheric Variables: Validation of the SAFRAN
30   Analysis over France. Journal of Applied Meteorology and Climatology, 47(1): 92-107.
31   DOI:10.1175/2007jamc1636.1

32 Renard, B. et al., 2013. Data-based comparison of frequency analysis methods: A general framework. Water
33   Resour. Res., 49(2): 825-843. DOI:10.1002/wrcr.20087

34 Ribatet, M., Sauquet, E., Grésillon, J.-M., Ouarda, T.B.M.J., 2007. A regional Bayesian POT model for flood
35   frequency analysis. Stochastic Environmental Research and Risk Assessment, 21(4): 327-339.

36 Ruiz-Bellet, J.L., Castelltort, X., Balasch, J.C., Tuset, J., 2017. Uncertainty of the peak flow reconstruction of
37   the 1907 flood in the Ebro River in Xerta (NE Iberian Peninsula). Journal of Hydrology, 545: 339-354.
38   DOI:https://doi.org/10.1016/j.jhydrol.2016.12.041

1   Schendel, T., Thongwichian, R., 2015. Flood frequency analysis: Confidence interval estimation by test
2   inversion bootstrapping. Adv Water Resour, 83: 1-9. DOI:10.1016/j.advwatres.2015.05.004

3   Sehgal, V., Tiwari, M.K., Chatterjee, C., 2014. Wavelet Bootstrap Multiple Linear Regression Based Hybrid
4   Modeling for Daily River Discharge Forecasting. Water Resources Management, 28(10): 2793-2811.
5   DOI:10.1007/s11269-014-0638-7

6   Shen, H.W., Koch, G.J., Obeysekera, J.T.B., 1990. Physically based flood features and frequencies. Journal of
7   Hydraulic Engineering, 116(4): 494-514.

8   Sraj, M., Viglione, A., Parajka, J., Bloschl, G., 2016. The influence of non-stationarity in extreme hydrological
9   events on flood frequency estimation. Journal of Hydrology and Hydromechanics, 64(4): 426-437.
10   DOI:10.1515/johh-2016-0032

11 Stedinger, J.R., Cohn, T.A., 1986. Flood Frequency Analysis With Historical and Paleoflood Information.
12   Water Resour. Res., 22(5): 785-793. DOI:10.1029/WR022i005p00785

13 Stedinger, J.R., Tasker, G.D., 1985. Regional hydrologic analysis: 1. Ordinary, weighted and genaralized least
14   squares compared. Water Resour. Res., 21(9): 1421-1432.

15 Stedinger, J.R., Vogel, R.M., Lee, S.U., Batchelder, R., 2008. Appraisal of the generalized likelihood
16   uncertainty   estimation   (GLUE)   method.   Water   Resour.   Res.,   44(W00B06):   1-17.
17   DOI:doi:10.1029/2008WR006822

18 Tiwari, M.K., Chatterjee, C., 2010. Uncertainty assessment and ensemble flood forecasting using bootstrap
19   based   artificial   neural   networks   (BANNs).   Journal   of   Hydrology,   382(1–4):   20-33.
20   DOI:http://dx.doi.org/10.1016/j.jhydrol.2009.12.013

21 Vidal, J.P., Martin, E., Franchisteguy, L., Baillon, M., Soubeyroux, J.M., 2010. A 50-year high-resolution
22   atmospheric reanalysis over France with the Safran system. Int J Climatol, 30(11): 1627-1644.
23   DOI:10.1002/joc.2003
24
25

1   ***Figure captions***

2   *Figure 1: Configuration of calibration / simulation cases for the SHYRE method.*

3   *Figure 2: Location and characteristics of the 1112 studied basins.*

4   *Figure 3: Example illustrating the variability of sampling characteristics for one rainfall generator*

5   *parameter corresponding to the intensity of a rainy event.*

6   *Figure 4: Confidence intervals of peak flow quantiles associated with the SHYREG method (taking into*

7   *account different sources of uncertainties - case 1 to 4), with the Gumbel distribution use and with the GEV*

8   *distribution use: example of the Ill at Didenheim station.*

9   *Figure 5: Distribution of the confidence interval amplitude for 2-, 10-, 100- and 1000-year quantiles,*

10   *associated with rainfall estimated by the SHYREG method (in black), peak flow estimated by the SHYREG*

11   *method for cases 1–4 (in blue), peak flow estimated by Gumbel distribution (in green) and peak flow*

12   *estimated by "bounded" GEV distribution (in red).*

13   *Figure 6: For case 4, distribution of amplitudes for groups of basins classified according to the recorded*

14   *length of observation (a, b), basin size (c, d) and specific flow (to distinguish more or less productive basins)*

15   *(e, f).*

16   *Figure 7: Localisation of station with the respective amplitudes of uncertainties, for 100-year quantiles of*

17   *rainfall and peak flow.*

18
19
20
21   ***Table captions***

22   *Table 1: List and description of the method's parameters.*

23