water

MDPI

Article

Sample Uncertainty Analysis of Daily Flood Quantiles Using a Weather Generator

Carles Beneyto [1,*], Gloria Vignes [1,2], José Ángel Aranda [1] and Félix Francés [1]

[1] Research Institute of Water and Environmental Engineering (IIAMA), Universitat Politècnica de València, 46022 Valencia, Spain
[2] Department of Civil, Chemical, Environmental and Materials Engineering (DICAM), Alma Mater Studiorum—University of Bologna, 40126 Bologna, Italy
* Correspondence: carbeib@alumni.upv.es

Abstract: The combined use of weather generators (WG) and hydrological models (HM) in what is called synthetic continuous simulation (SCS) has become a common practice for carrying out flood studies. However, flood quantile estimations are far from presenting relatively high confidence levels, which mostly relate to the uncertainty of models' input data. The main objective of this paper is to assess how different precipitation regimes, climate extremality, and basin hydrological characteristics impact the uncertainty of daily flood quantile estimates obtained by SCS. A Monte Carlo simulation from 18 synthetic populations encompassing all these scenarios was performed, evaluating the uncertainty of the simulated quantiles. Additionally, the uncertainty propagation of the quantile estimates from the WG to the HM was analyzed. General findings show that integrating the regional precipitation quantile ($X_{T,P}$) in the WG model calibration clearly reduces the uncertainty of flood quantile estimates, especially those near the regional $X_{T,P}$. Basin size, climate extremality, and the hydrological characteristics of the basin have been proven not to affect flood quantiles' uncertainty substantially. Furthermore, it has been found that uncertainty clearly increases with the aridity of the climate and that the HM is not capable of buffering the uncertainty of flood quantiles, but rather increases it.

Keywords: weather generator; hydrological model; uncertainty; Monte Carlo simulation; daily flood quantile

Citation: Beneyto, C.; Vignes, G.; Aranda, J.Á.; Francés, F. Sample Uncertainty Analysis of Daily Flood Quantiles Using a Weather Generator. *Water* **2023**, *15*, 3489. https://doi.org/10.3390/w15193489

Academic Editor: Athanasios Loukas

Received: 9 September 2023
Revised: 28 September 2023
Accepted: 30 September 2023
Published: 6 October 2023

1. Introduction

Accurately designed flood estimation is required to make the best decision possible in various applications, including infrastructure construction, land-use management, and risk assessment [1,2]. Flood frequency analysis (FFA) includes all techniques that aim to decipher the natural random processes that drive the occurrence and magnitude of flood events [3]. Traditional techniques are based on fitting the available annual maxima discharge data to a distribution function for gauged basins or deterministic approaches for ungauged basins. The main problem of these methods is the lack of long systematic observations, mainly concerned with hydrometric measurements, which characterize runoff processes [4].

More recently, hybrid methods based on the use of the two abovementioned approaches, either through their combination or by adding other sources of information or techniques [5], appear to be gaining an important audience. In this vein, a widely adopted approach is the generation of long synthetic flood data series by combining the use of stochastic weather generators (WG) coupled with a hydrological model (HM) in what is called synthetic continuous simulation (SCS). This approach addresses the issue of the short length of the available observations at the same time as it eliminates the problem of the determination of the initial conditions of the basin and the characterization of the

spatiotemporal distribution of the precipitation in the case of deterministic methods such as design storms.

Despite the invaluable contribution that this approach makes to flood estimation, it is true that quantile estimates still present high uncertainty, which becomes magnified when dealing with low-frequency events [6,7]. WGs rely on the amount of information available, meaning they require representative data series of observed extreme precipitation in order to perform adequately [8]. The lack of this has led practitioners to seek potential solutions to reduce the uncertainty of quantile estimations. In this sense, some WG developers have opted for the use of heavy-tailed distribution functions to model precipitation amounts, including Evin et al. [9] and Ahn [10]. Beneyto et al. [11] proposed the incorporation of more robust and reliable studies (e.g., regional precipitation studies) for the parameterization of WGs, obtaining a clear reduction in the uncertainty of the low-frequency precipitation quantile estimates, although not quantifying this uncertainty reduction. The quantification of hydrological prediction uncertainty is crucial in water resource decision-making processes [12].

Major sources of uncertainty in hydrological modeling include input and calibration data, model structures, and parameters [13], with model input data being the primary source of uncertainty [14], especially in regions with limited data availability, such as arid and semi-arid regions [15]. Moges et al. [16] categorized uncertainty analysis methods into six broad classes: (i) Monte Carlo analysis, (ii) Bayesian statistics, (iii) multi-objective analysis, (iv) least squares-based inverse modeling, (v) response surface-based techniques, and (vi) multi-modeling analysis. Based on Monte Carlo analysis, Beneyto et al. [17] evaluated reductions in the uncertainty of the integration of a regional study of annual maximum daily precipitation in the WG parametrization entailed in the precipitation quantiles, concluding that integrating these studies provides more information than longer input data series.

In terms of discharges, many studies have focused on the quantification of input data uncertainty. Sun et al. [18] calculated flood hydrographs for the Finniss River basin in Darwin, Australia, using different approaches to estimate the input rainfalls from the available radar and rain gauge data. Bardossy and Das [19] investigated the influence of the spatial resolution of the rainfall input on the model calibration and application. Moulin et al. [20] built, calibrated, and validated a realistic error model on mean areal precipitation (MAP) estimates and undertook a detailed analysis of the links between MAP estimation uncertainties, basin area, and streamflow simulation uncertainties. This latter study concluded that a large part of the rainfall–runoff modeling errors could be explained by the uncertainties in rainfall estimates, limiting their operational use for flood forecasting.

Therefore, reliable precipitation data series is key to obtaining accurate flood quantile estimates. In the SCS approach, since the sample statistics set the generated precipitation patterns, the precipitation regime may have a direct influence on the flood quantile estimations, i.e., flood quantile estimations in arid or semi-arid climates, with effectively no rain over the year with occasional intense precipitation events, might present different degrees of uncertainty than in the case of humid climates, with less skewed and more homogeneous precipitation.

In this context, and following in the footsteps of the previous work of Beneyto et al. [17], the main objective of this paper is to quantify the uncertainty of flood discharge quantile estimates obtained by SCS in a wide range of hydro-climatic scenarios through the combination of (1) different precipitation regimes, (2) different climate extremality, and (3) different hydrological characteristics of the basin. In other words, we are trying to analyze the influence of hydro-climatic conditions on the uncertainty of the flood quantiles obtained with the SCS approach presented in Beneyto et al. [17]. Additionally, a simple sensitivity analysis for the basin size was conducted, as well as an assessment of the uncertainty transmission from the precipitation to the discharge quantiles. We propose a framework that accounts for all these different scenarios, which are represented by synthetic populations considered

to be "truth" and used as benchmarks to evaluate the sample uncertainty through Monte Carlo simulations.

Since the objective is the uncertainty of the proposed SCS methodology independently of the selected models, and for the sake of simplicity, we will assume that both the meteorological and the hydrological models do not introduce any uncertainty beyond the methodology itself. Specifically, (i) the WG does not present structural uncertainty, and (ii) the HM perfectly reproduces the reality. In any case, this unconsidered uncertainty is much smaller than that corresponding to the precipitation information usually available.

2. Methodology

2.1. Simulation Framework

Different daily precipitation populations were created with different climates and different precipitation extremality, the latter being understood as the number of registered extreme precipitation events per year. These populations were created from existing observations and by handling the adequate WG parameters to increase or reduce the annual precipitation, the annual maximum daily precipitation, and the number of wet days per year. A Monte Carlo simulation was performed with 50 packages of 60-year samples extracted from each of the populations. Following the approach in Beneyto et al. [17] and assuming the population precipitation quantile ($X_{T,P}$) to be perfect (i.e., no uncertainty), the population $X_{100,P}$ was introduced in the WG for its parametrization. A long series of precipitation and both maximum and minimum temperatures were then simulated, which in turn fed a fully distributed HM with two different hydrological regimes (i.e., permanent regime and ephemeral regime) to obtain the discharges and the corresponding flood quantiles. Both basin hydrological regimes were obtained by varying the HM Correction Factors (CF), as will be explained in the following sections.

Precipitation and discharge population quantiles X_{10}, X_{100}, and X_{500} were compared with those of the 50 packages, analyzing the uncertainty through the Relative Root Mean Square Error (RRMSE), the Coefficient of Variation (CV), and the Relative Bias (RB).

Finally, the transmission of uncertainty from the WG to the HM was also evaluated. A workflow diagram of the methodology can be seen in Figure 1.

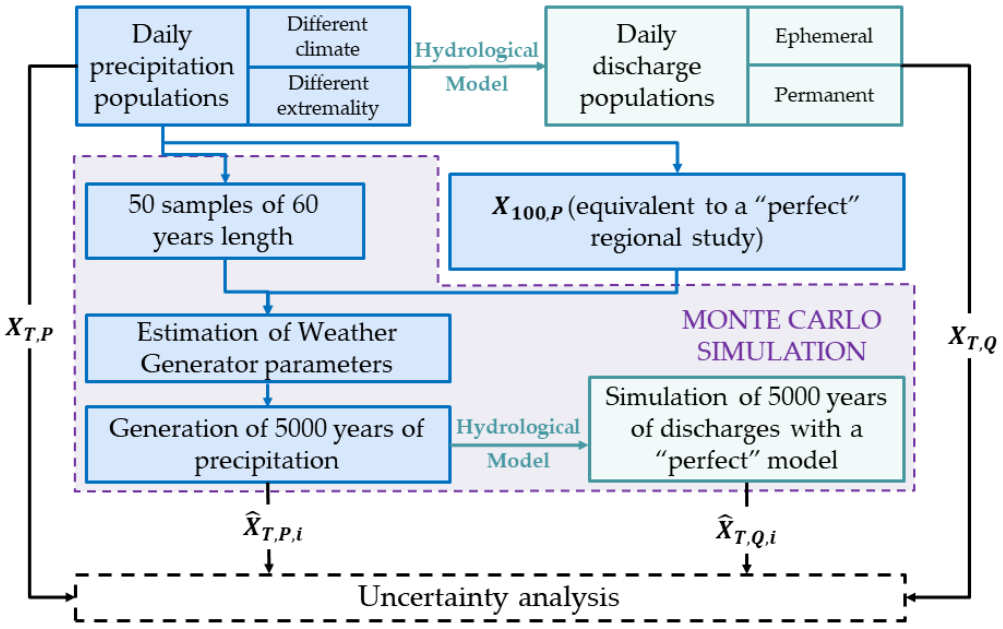

Figure 1. Workflow diagram of the methodology. $X_{T,P}$: Population precipitation quantile. $X_{Q,P}$: Population flood quantile. $X_{T,P/Q,i}$: Precipitation/flood quantile of each realization.

2.2. Stochastic Weather Generator: GWEX

The WG used to perform this study was GWEX [9], a stochastic multi-site WG focused on extreme events. Among the many features this WG presents, and for the present study, it should be highlighted that it incorporates the Extended Generalized Pareto Distribution (E-GPD) function [21] for modeling the precipitation amount, which is basically obtained by raising the Generalized Pareto Distribution to a power of $k > 0$:

$$F(x;\lambda) = \left[1 - \left(1 + \frac{\xi x}{\sigma}\right)_{+}^{-1/\xi}\right]^{\kappa}, \; x > 0 \tag{1}$$

where $\lambda = (k, \sigma, \xi)$ a vector of parameter, where k controls the shape of the lower tail, σ is a scale parameter, and ξ controls the rate of the upper tail decay [22] as demonstrated in Beneyto et al. [17].

This latter parameter was therefore set to three different values {0.05, 0.11, and 0.25} to generate precipitation populations with different extremality, the population with a ξ value of 0.05 being the least extreme, and the one with a ξ value of 0.25 the most extreme. Similarly, this parameter was estimated in the WG calibration process for each realization to obtain the closest value of the sample $X_{100,P}$ to the population one.

2.3. Eco-Hydrological Model: TETIS

The TETIS eco-hydrological model [23] is a conceptual (tank structure) model with physically based parameters and fully distributed in the space. It incorporates an effective split-parameter structure that facilitates the model implementation process, presenting only nine CFs to be calibrated for the hydrological module. Rather than calibrating the value at each cell, the estimated value of each of these CFs is multiplied by the value of each cell in the corresponding input raster map. This considers the spatial and/or temporal effects and the model or input errors, and allows for a quick and simple (manual or automatic) calibration of the different processes represented, taking advantage of the information used in the parameter estimation. The CFs of the TETIS model are as follows: CF1. Static storage; CF2. Evapotranspiration; CF3. Infiltration; CF4. Overland flow; CF5. Percolation; CF6. Interflow; CF7. Deep aquifer flow; CF8. Connected aquifer flow; and CF9. Kinematic wave velocity.

In our case, and starting from an implemented model in previous studies, two different synthetic basins with different hydrological behaviors were built up: ephemeral and permanent. These basins were created by setting the values of CF3, CF5, CF6, CF7, and CF8 based on the authors' expertise. It is worth highlighting that, for our study, we consider our model to be "perfect", i.e., no uncertainty is introduced in the quantile estimations by the model.

3. Synthetic Case Study

3.1. Basin Description

The synthetic basin, obtained from an actual one to guarantee its natural behavior, has a drainage area of approximately 180 km^2. The altitude of the basin ranges from 1.061 m.a.s.l. at the headwaters to 226 m.a.s.l. at the outlet, presenting a main stream of 28.5 km in length and several short tributaries pouring on both sides of the main course (Figure 2).

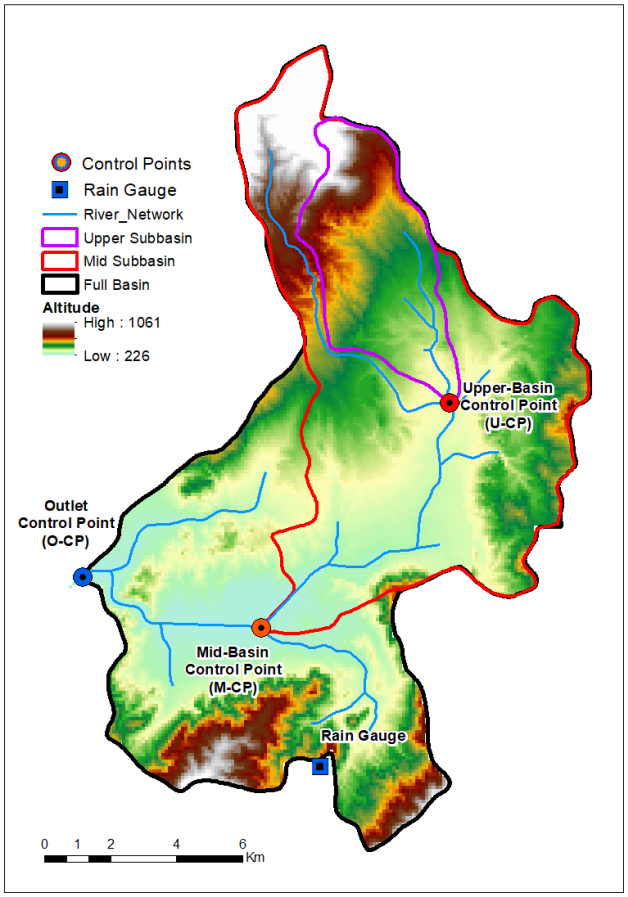

Figure 2. DEM of the synthetic basin (north arrow not shown given the synthetic nature of the basin).

Three synthetic flow gauges (hereafter Control Points (CP)) were defined within the basin as follows: the Upper basin Control Point (U-CP), collecting the waters of an approximately 24 km^2 headwater tributary just before its confluence with the main stream; the Mid-basin Control Point (M-CP), with a basin area of 101.5 km^2; and the Outlet Control Point (O-CP), at the outlet of the full synthetic basin (Figure 2).

Two different hydrological characteristics of the basin were analyzed to test the basin response. The first one, reproducing an ephemeral regime, is characterized by 70% of overland flow and 30% of interflow, being the aquifer disconnected. This regime is typical of arid or semi-arid regions, where rivers only present flow after high precipitation events, remaining dry for most of the year. The second hydrological regime represents a permanent discharge regime, typically found in humid climates where rivers never dry up independently at that time of the year. This latest is characterized by 30% overland flow, 40% of interflow, and the remaining 30% feeds the aquifer or reaches the river channel.

Both hydrological responses were achieved by setting the abovementioned CFs, which basically translates into increasing or decreasing the values of the hydraulic conductivity of the soil maps (obtained from the European Soil Database (ESDB)). Additionally, the infiltration capacity map (obtained from the ESDB) and the percolation capacity map (obtained from the Spanish Geological and Mining Institute) were also slightly amended. Slope, flow directions, and flow accumulation maps, derived from a digital elevation model with a 100 m resolution obtained from the National Geographic Institute, are common for both hydrological regimes.

3.2. Climate Description and Statistics

Two grid points from the Spain02-v5 reanalysis dataset [24,25] were initially used to create the synthetic populations. Specifically, these were "grid3715", representative of a Spanish Mediterranean coast semi-arid climate, and "grid3314", representative of a northern

Spanish humid climate. From this information, six different 15,000 y populations were created by modifying the WG parameters associated with the extremality and the percentage of dry/wet days. Table 1 shows the basic statistics of the nine precipitation populations.

Table 1. Populations' precipitation statistics.

Variable	Statistic	Semi-Arid			Humid			Extremely Humid			Units
		0.05	**0.11**	**0.25**	**0.05**	**0.11**	**0.25**	**0.05**	**0.11**	**0.25**	
Daily Precipitation (P_d)	Mean	1.57	1.57	1.56	2.05	2.05	2.05	3.60	3.60	3.60	mm
	Mean if $P_d > 0.1$ mm	6.32	6.32	6.29	6.59	6.54	6.42	6.20	6.21	6.20	mm
	Standard Deviation (SD)	6.19	6.35	6.90	5.69	5.82	6.35	7.15	7.34	8.05	mm
	SD if $P_d > 0.1$ mm	11.16	11.52	12.73	8.61	8.86	9.91	8.49	8.76	9.78	mm
	% $P_d > 0.1$ mm	24.79	24.79	24.79	31.91	31.91	31.91	57.95	57.95	57.95	%
	Max	249.51	373.15	846.69	173.80	238.80	805.50	208.37	263.20	677.65	mm
Annual Precipitation	Mean	572.46	572.62	569.76	748.94	748.91	748.23	1313.27	1315.27	1313.08	mm
Annual max. Precipitation	Mean	59.56	62.96	70.77	47.61	50.88	60.88	53.51	58.07	72.18	mm
	Coeff. Variation	0.43	0.48	0.67	0.33	0.39	0.60	0.31	0.36	0.57	-
	C. Skewness	1.55	2.02	3.53	1.36	1.75	4.53	1.41	1.81	3.63	-
	C. Kurtosis	7.25	10.68	27.61	6.25	8.62	52.26	6.91	9.54	30.82	-

All populations' annual mean daily temperature and annual mean daily precipitation were checked against De Martonne's aridity index [26], each of them falling within the climate types of semi-arid, humid, and extremely humid, respectively.

In summary, considering the two different basin hydrological regimes, the three different climates, and the three different climate extremality, 18 possible combinations were considered in the present study. Figure 3 below outlines all these considered combinations, which represent 18 synthetic discharge populations with significantly different hydro-climatic characteristics.

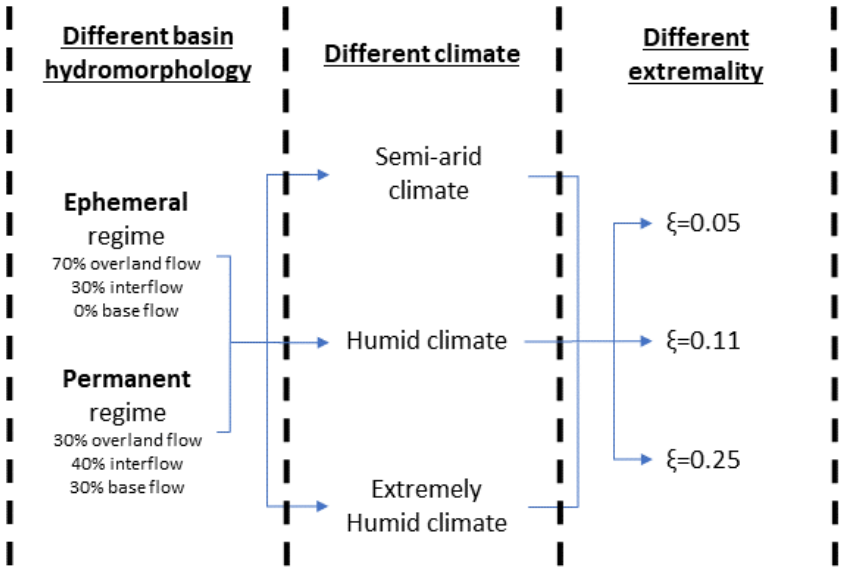

Figure 3. Synthetic populations outline.

4. Results

This section presents the results of the uncertainty analysis of flood quantile estimates in the three different scenarios considered: different basin hydrology, different climate, and different climate extremality. Flood quantiles for return periods of 10, 100, and 500 years were analyzed, representing moderate, low, and very low probability of occurrence.

Additionally, the results of a preliminary analysis assessing the influence of the basin size in the quantile estimates are presented. Finally, the uncertainty transmission through the HM from the precipitation quantile estimates to the flood quantile estimates was also analyzed and quantified.

For better clarity, all RRMSE, RB, and CV values (expressed in percentage) are shown at the end of the section.

4.1. Preliminary Analysis

This analysis was first undertaken to find out whether the size of the drainage area influences the flood quantile uncertainty or not. The analysis was carried out considering a semi-arid climate, medium extremality (ξ = 0.11), and an ephemeral behavior of the basin. Results in the three CPs (24 km^2, 101.5 km^2, and 180 km^2, respectively) are shown in Figure 4, where each boxplot represents the relationship between the population quantile and the simulated quantiles for the 50 realizations. A similar negative value of the RB was obtained for $X_{10,Q}$ and $X_{100,Q}$ for all three CPs, being again similar but positive in the case of $X_{500,Q}$. This latter quantile presented the higher errors, with RRMSE and CV values of up to 8.19% and 6.10%, respectively.

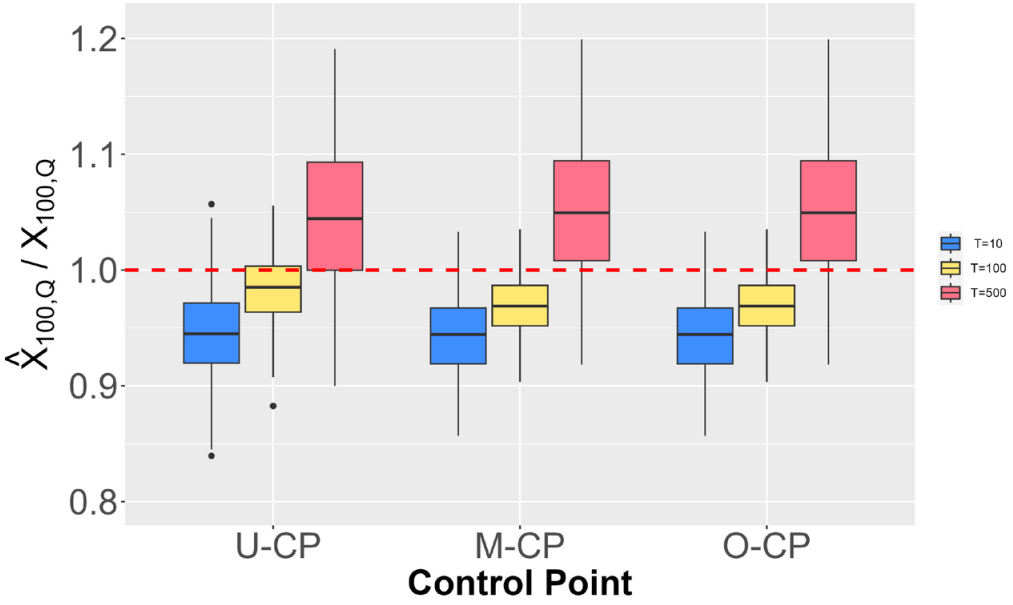

Figure 4. Boxplots of flood quantile estimates in the three CPs for return periods of 10, 100, and 500 years, considering a semi-arid climate, medium extremality (ξ = 0.11), and an ephemeral regime.

Nevertheless, these values tend to vary slightly with the basin size and without a clear pattern. These variations are almost negligible, and they could be attributable to the sample variability or the methodology uncertainty itself. Therefore, only the CP located at the basin outlet (i.e., O-CP) will be considered in further analyses.

4.2. Hydrological Characteristics of the Basin

Two different hydrological characteristics of the basin were analyzed, reproducing an ephemeral (i.e., 70% of overland flow and 30% of interflow) and a permanent discharge regime (i.e., 30% of overland flow, 40% of interflow, and 30% feeding the aquifer or reaching the river channel).

The aim of this analysis was to evaluate if the basin behavior could have a significant influence on the uncertainty of quantile estimations. As with the previous analysis, Figure 5 represents the boxplots for the quantile estimates in a semi-arid climate for medium extremality (ξ = 0.11), both for an ephemeral and a permanent regime. Although flood quantiles in a permanent river present slightly lower values of RRMSE and CV for all

three return periods, again, no significant differences in terms of uncertainty can be found between either discharge regime. Additionally, quantiles for return periods of 10 and 100 years tend to be underestimated (negative RB), whereas quantiles for a return period of 500 years are overestimated (positive RB).

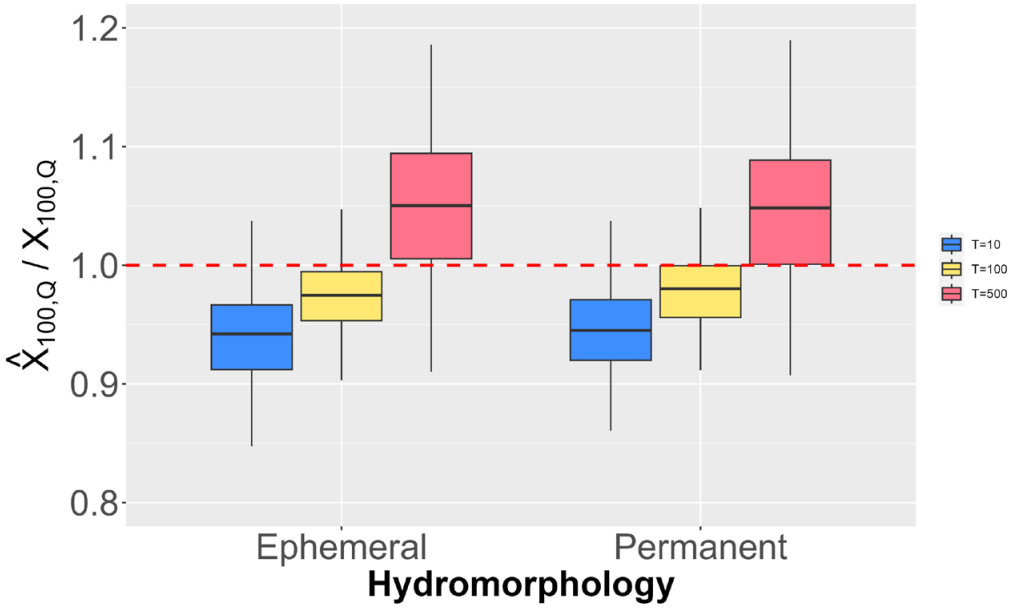

Figure 5. Boxplots of flood quantile estimates in a semi-arid climate and medium extremality ($\xi = 0.11$) for return periods of 10, 100, and 500 years in two different hydrological basin characteristics.

It can be concluded that the hydrological characteristics of the basin are not a significant factor in terms of quantile estimates' uncertainty. Thus, only results for the ephemeral regime will be presented in successive analyses.

4.3. Precipitation Regime

Having set the river regime as ephemeral, three rainfall regimes according to De Martonne's aridity index were analyzed: semi-arid, humid, and extremely humid. As shown in Table 1, the main differences between all three climates relate to the percentage of wet days (ca. 25%, 32%, and 58%, respectively) and the mean annual precipitation (ca. 570 mm, 750 mm, and 1300 mm, respectively). Figure 6 shows the boxplots of the quantile estimates in an ephemeral river with medium climate extremality ($\xi = 0.11$) for the three climates. In this case, significant differences can be observed. Flood quantiles for return periods of 10 years and 100 years present less uncertainty as the climate turns more humid, reducing the values of RRMSE from 7.31% in a semi-arid climate to 4.59% in a very humid climate and from 4.17% to 1.42%, respectively. A similar reduction can be observed for the value of the CV, this being more evident in the case of a very humid climate. Conversely, for return periods of 500 years, both values of RRMSE and CV increase rapidly as the climate turns less arid, reaching values of up to 24.75% and 4.60%, respectively. Although $X_{100,Q}$ is fairly well represented, especially in the case of a very humid climate, $X_{10,Q}$ is systematically underestimated and $X_{500,Q}$ presents a positive RBs for all three climates.

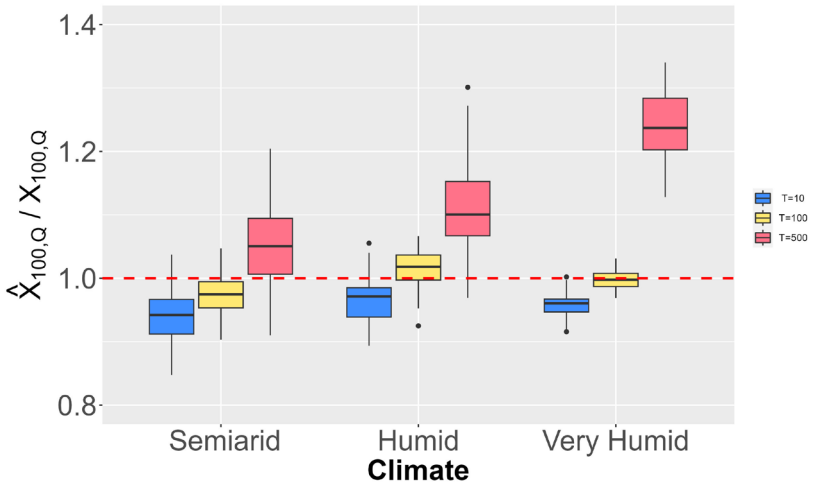

Figure 6. Boxplots of flood quantile estimates for an ephemeral river with medium climate extremality ($\xi = 0.11$) for return periods of 10, 100, and 500 years in the three analyzed climates.

4.4. Climate Extremality

The aim of this analysis is to assess whether the extremality of the population has an impact on the flood quantile estimations. This extremality, understood as the number of registered extreme precipitation events per year, has been synthetically introduced in the populations by means of modifying the shape parameter ξ of the E-GPD, obtaining three climate extremality: low extremality ($\xi = 0.05$); medium extremality ($\xi = 0.11$); and high extremality ($\xi = 0.25$). Figure 7 shows the boxplots of the quantile estimates for an ephemeral river for return periods of 10, 100, and 500 years in the three analyzed climates before, and the three climate extremalities. Results show that $X_{100,Q}$ are generally well estimated independently of the climate extremality, albeit RRMSE and CV values slightly increase as the climate is more extreme. Still, RRMSE values range from 1.42% to 4.17%, which indicates a very good estimation. $X_{10,Q}$ are underestimated for all three climate extremalities, presenting RRMSE values quite similar for all climates and climate extremality, except for the semi-arid climate, where the value increases considerably. In the case of $X_{500,Q}$, and with the only exception of the semi-arid climate with low extremality where the RRMSE value is satisfactory (i.e., 4.66%), quantile estimates systematically overestimate the population quantiles. This is more evident as the population is more extreme, and especially as the climate becomes more humid, reaching RRMSE values of up to 37.43% and CV values of up to 10.50%.

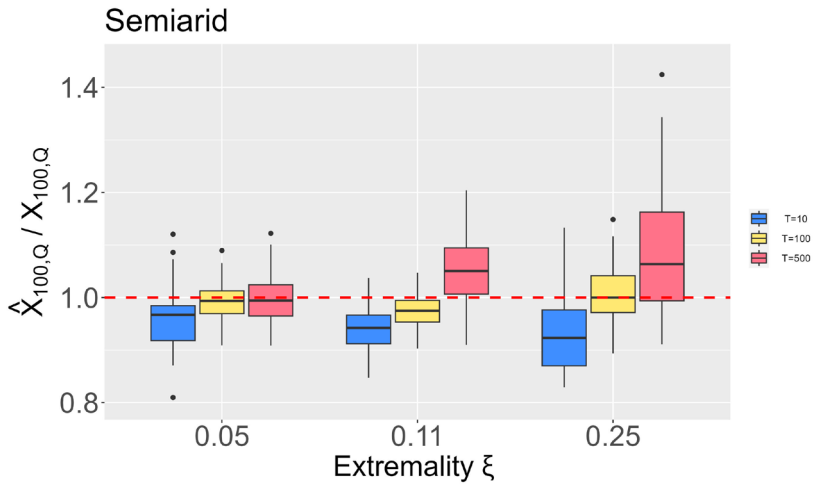

Figure 7. *Cont.*

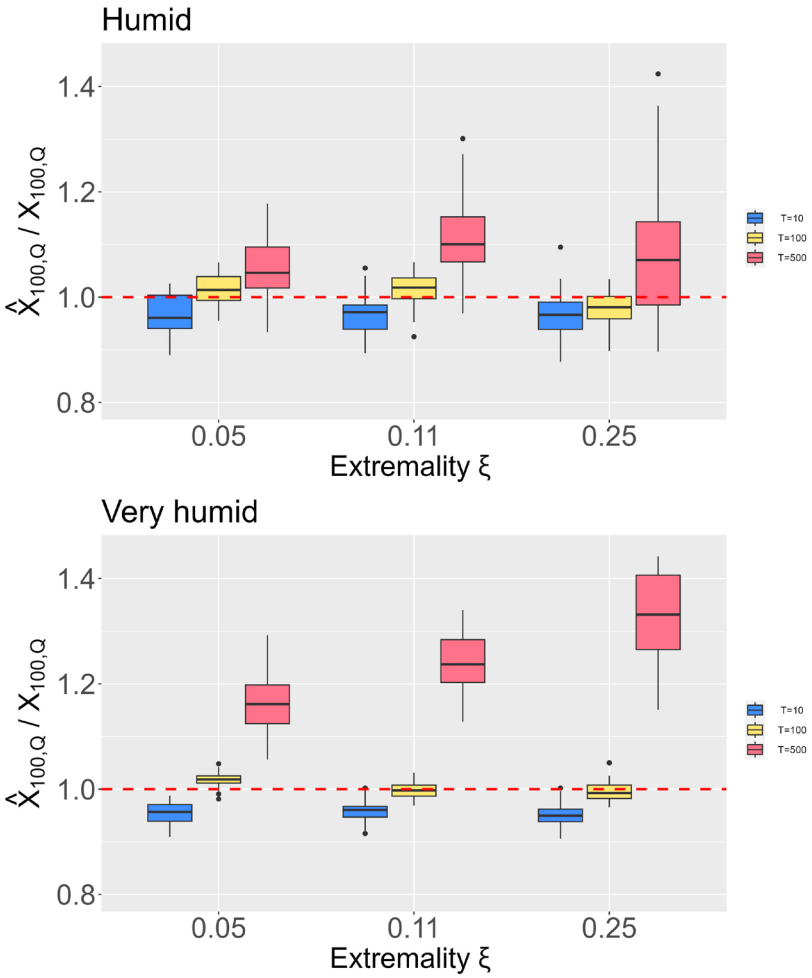

Figure 7. Boxplots of flood quantile estimates for an ephemeral river for return periods of 10, 100, and 500 years in the three analyzed climates and climate extremality.

4.5. Uncertainty Propagation

This section analyzes the uncertainty propagation of the quantile estimates from the WG to the HM. Respective precipitation quantiles to those flood quantiles shown in Figure 7 are now represented in Figure 8.

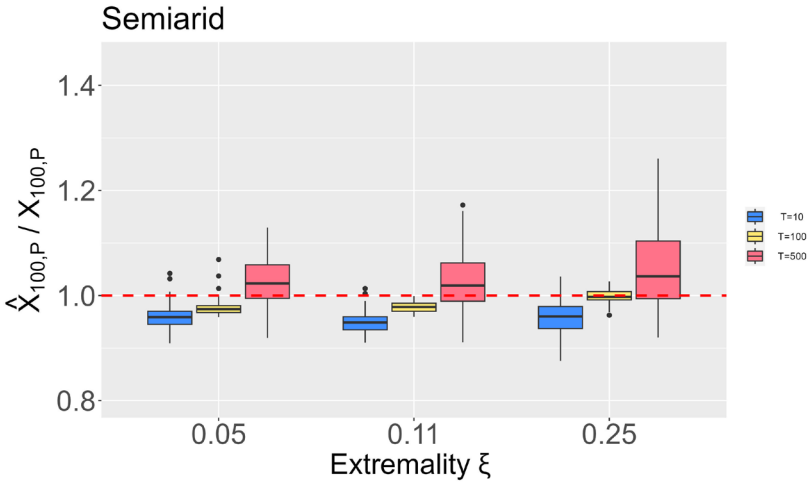

Figure 8. *Cont.*

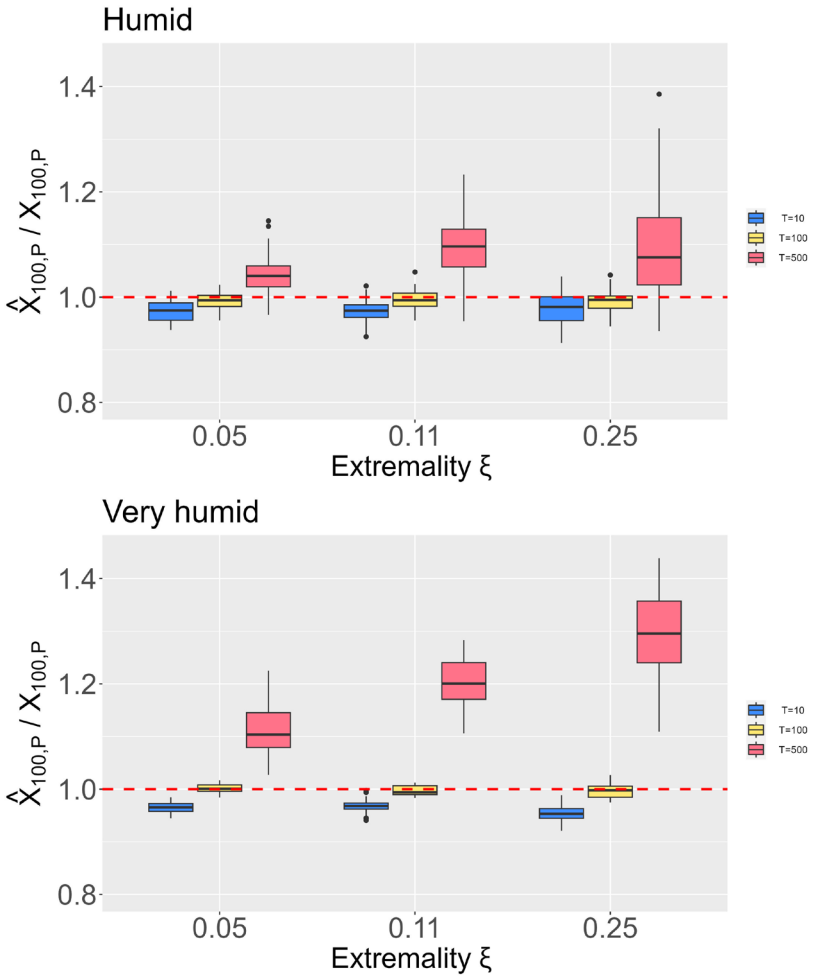

Figure 8. Boxplots of precipitation quantile estimates for an ephemeral river for return periods of 10, 100, and 500 years in the three analyzed climates and climate extremality.

At first glance, it can be appreciated how the size of the boxplots and the distances to the red dashed line are lower, which means that, in general, $X_{T,P}$ present lower uncertainties than their respective $X_{T,Q}$. Table 2 shows the RRMSE, CV, and RB values for precipitation and flood quantiles for all analyses undertaken in this research. As can be observed, there is a systematic increase (or decrease in the case of negative RB) for all scenarios, which means that uncertainty clearly propagates through the HM. Therefore, it can be concluded that the rainfall–runoff transformation, far from acting as a buffer, increases the uncertainty of the flood quantile estimates.

In a more illustrative way, Figure 9 shows the mean (considering all scenarios) value of RRMSE, CV, and RB for both precipitation and flood quantile estimates for the three considered return periods. RRMSE, CV, and RB values are considerably lower for return periods of 100 years, as expected, since $X_{100,T}$ was used for the WG calibration. However, it can be seen an increase in RRMSE and CV for all return periods. In the case of the RB, this presents an increase for return periods of 500 years and a decrease for return periods of 100 years, which indicates an increase in uncertainty. Lastly, the almost-negligible mean RB values are explained by the WG calibration procedure.

Table 2. RRMSE, CV, and RB values for the estimated precipitation and flood quantiles for return periods of 10, 100, and 500 years, the three climate extremality (ξ = 0.05, 0.11, and 0.25), and the three analyzed climates.

			RRMSE		CV		RB	
Climate	ξ	**Return Period (Years)**	**Precipitation (%)**	**Discharge (%)**	**Precipitation (%)**	**Discharge (%)**	**Precipitation (%)**	**Discharge (%)**
Semi-arid	0.05	10	4.8	7.1	2.8	6.0	−4.0	−4.3
		100	2.9	3.8	1.9	3.7	−2.2	−0.9
		500	5.2	4.7	4.5	4.7	2.5	−0.1
	0.11	10	5.5	7.3	2.4	4.8	−5.0	−5.8
		100	2.4	4.2	1.0	3.4	−2.2	−2.6
		500	6.9	8.2	6.0	6.1	3.2	5.2
	0.25	10	5.4	12.4	3.7	9.5	−4.1	−9.1
		100	1.4	5.3	1.4	5.3	−0.1	0.8
		500	9.7	13.9	8.0	10.8	5.0	7.8
Humid	0.05	10	3.4	5.0	2.0	3.8	−2.7	−3.4
		100	1.7	3.3	1.5	2.9	−0.7	1.5
		500	5.8	7.5	3.7	5.1	4.3	5.3
	0.11	10	3.3	5.0	2.2	3.8	−2.6	−3.4
		100	1.9	3.3	1.8	2.9	−0.6	1.5
		500	11.1	13.7	5.4	6.9	9.5	11.4
	0.25	10	3.5	5.5	2.9	4.4	−2.1	−3.5
		100	2.2	3.8	2.1	3.2	−0.7	−2.1
		500	13.5	13.5	9.3	10.5	9.0	7.5
Extremely humid	0.05	10	3.6	4.8	1.0	1.9	−3.4	−4.4
		100	0.8	2.2	0.8	1.3	0.1	1.8
		500	12.2	17.1	4.0	4.6	11.3	16.2
	0.11	10	3.5	4.6	1.3	2.0	−3.3	−4.2
		100	1.0	1.4	0.9	1.4	−0.3	−0.2
		500	20.9	24.8	4.1	4.6	20.3	24.1
	0.25	10	4.9	5.3	1.6	2.0	−4.7	−5.0
		100	1.3	1.8	1.3	1.8	−0.4	−0.5
		500	31.5	37.4	6.8	6.9	30.3	36.2

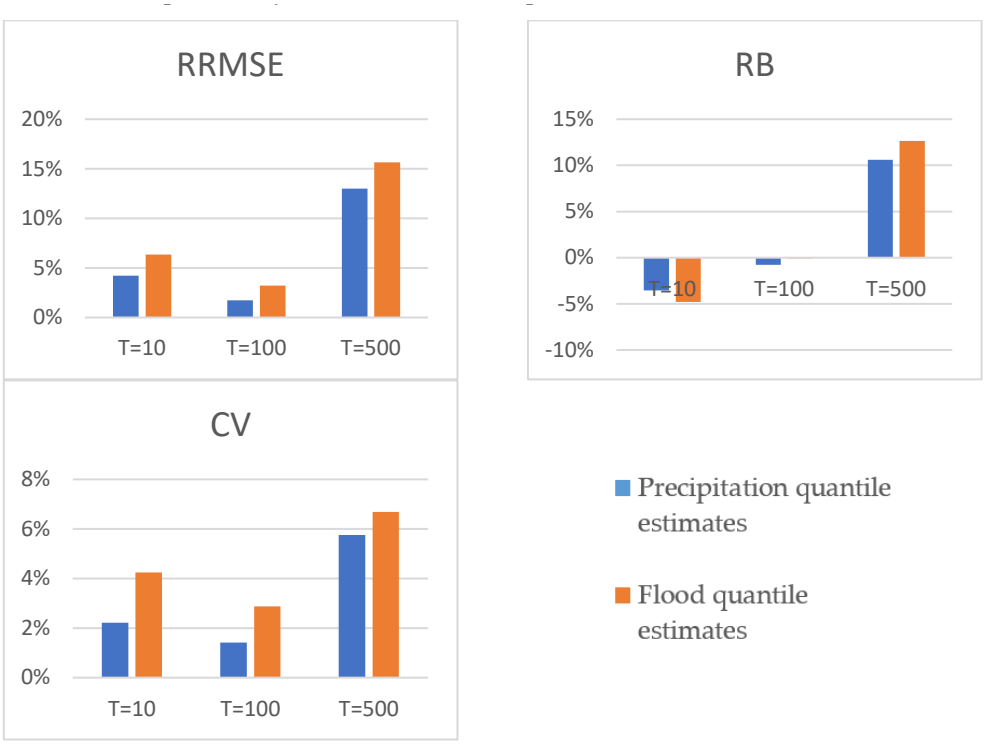

Figure 9. Difference between mean (considering all scenarios) RRMSE, CV, and RB values for the estimated precipitation and flood quantiles.

5. Discussion

SCS is a widely adopted hybrid approach to determine flood quantile estimates, which resolves the issues of former purely statistical or deterministic methodologies (i.e., characterizing the initial conditions of the basin and adequately representing the spatiotemporal distribution of the precipitation) [27]. Nevertheless, low-frequency flood quantile estimates still present high uncertainty. The length of the input data series and the low density of monitoring stations constitute the main source of uncertainty and one of the main challenges to be faced in FFA, especially in arid and semi-arid regions [15]. Longer available input data series would ideally contribute to reducing estimates' uncertainty, but, unfortunately, this is something that only the passage of time can mitigate. Instead, incorporating additional sources of information or improving the model set-up can lead to considerable improvements in the reliability of flood estimates. Moreover, quantifying the uncertainty is required for decision-makers to understand the implications of limited data, model uncertainties, changes in the flooding system over the long term, incommensurate scales of appraisal, and potentially conflicting decision objectives [28]. Recognizing this importance, no studies could be found in the literature quantifying the uncertainty of quantile estimates by SCS associated with the amount of available input information, which is precisely the objective of this paper. Being aware of the different precipitation patterns and basin characteristics around the world and drawing from the previous research by Beneyto et al. [17], this study intends to elucidate how different precipitation regimes, climate extremality, or basin hydrological characteristics impact the uncertainty of the flood quantile estimates.

Results initially obtained in three CPs of the basin through Monte Carlo simulations with samples of 60 years, which is double the 30 years set as a standard reference by the World Meteorological Organization [29], showed no major differences in terms of quantile estimates' uncertainty. This follows the line of the results obtained by Moulin et al. [20], and it can be concluded that the basin size has no major influence on the accuracy of the estimations.

An ephemeral and perennial river were also compared. Far from these authors' initial thoughts, results show again that no significant differences could be found with regard to the quantile estimates accuracy. Ephemeral rivers, usually located in less populated arid or semi-arid areas, pose unique challenges to researchers and practitioners due to the generally limited data records, these being poorly gauged and usually sporadically active, which result in the most hazardous types of floods [15,30]. In view of the results obtained, the difficulties in accurately modeling ephemeral rivers stem from the gauge data availability rather than the few no-zero observations since it has been proven that, under the same conditions (i.e., same sample size), flood quantile estimates in ephemeral rivers present the same degree of uncertainty as in the case of permanent rivers.

In arid and semi-arid areas, with long periods of drought conditions with no flow followed by short, intense precipitation events leading to flash floods, few models are considered adequate for modeling hydrologic processes due to the difficulty in effectively modeling infiltration-excess runoff processes as the dominant generation mechanism [31]. Compared to humid regions, where the information on the internal state of the basin is obtained from streamflow records, most models perform well, mainly because the dominant runoff generation mechanism is saturation excess runoff [32,33]. For this reason, hydrological prediction is more challenging for arid or semi-arid regions than for humid regions [34]. The results obtained in the present study analyze three different climates according to De Martonne's index: semi-arid, humid, and very humid point in the same direction. As shown in Figure 6, both values of RRMSE and CV decrease as the climate turns more humid. This is more evident in the case of the CV, which is explained by the higher precipitation homogeneity of humid and very humid climates, where the internal basin fluxes remain more stable than in semi-arid climates. An exception is found here for $X_{500,Q}$, which is systematically overestimated for all return periods, especially as climate becomes wetter. In fact, this problem originates from the meteorological modeling (Figure 8) and it

is dragged into discharges. After different analysis, it was determined that the problem lay with the high number of non-zero precipitation, and especially with the shape of the E-GPD. Humid and very humid climates, with much more days of low precipitation events than in semi-arid climates, added more information to the left tail of the distribution functions, which resulted in these data governing the fit in detriment of the right tail observations, thus and given the shape of the E-GPD, misrepresenting the higher quantiles.

In terms of climate extremality, three different populations with different extremality were analyzed. This analysis was repeated for the three abovementioned climate types (Figure 7). Apart from the already commented overestimations of $X_{500,Q}$, in general, RRMSE and CV values remain similar for the three climates. There are not many differences between the uncertainty obtained in $X_{10,Q}$ and $X_{100,Q}$ for low extremality and medium extremality (i.e., $\xi = 0.05$ and $\xi = 0.11$, respectively); however, for high extremality ($\xi = 0.25$), RRMSE and CV values increase considerably, which means that the more extreme the climate is the more uncertainty the flood quantile estimates present.

It is worth noting that, following the same methodology as in Beneyto et al. [17], WG calibration was made using the regional $X_{100,T}$. This is the reason the best results in terms of flood quantile uncertainty are presented for $X_{100,Q}$. Practitioners willing to better capture higher flood quantiles should use a different $X_{P,T}$ for the WG calibration (e.g., $X_{500,T}$).

It is expected that for a well-calibrated HM that adequately represents the important runoff processes within the basin, the major factor contributing to the uncertainty in the predicted flows is the uncertainty in rainfall [35]. In our case, having a "perfect" HM, the intention of the analysis was to assess if the uncertainty of the estimated precipitation quantiles propagates through the HM, evaluating if the HM could buffer the uncertainty of flood quantiles. In similar studies, Butts et al. [35] determined the propagation of uncertainty due to uncertainties in the measured rainfall using a Monte Carlo approach, using a total of 200 samples. They concluded that a 50% relative standard deviation in the precipitation estimate (R = 0.5) has only a limited impact on the accuracy of the hydrological simulation when compared to the flow measurement uncertainty and the other sources of uncertainty. Gabellani et al. [36] explored the impact of uncertainties in the spatiotemporal distribution of rainfall on the prediction of peak discharge in a typical mountain basin, concluding that uncertainties in the small-scale statistical properties of forecasted rain fields propagate along the rainfall–runoff chain and affect the prediction of peak discharge. In our case, and as observed in Figure 8, similar results were obtained: RRMSE, CV, and RB values systematically increased for all nine precipitation populations, indicating that the HM, far from acting as a balance component in the SCS approach, magnified the uncertainty of the estimates.

6. Conclusions

SCS has increasingly gained popularity as a means to extend the existing limited hydro-meteorological records. However, this approach heavily relies on the available observations, which in practice are rarely sufficiently long. This may lead to systematic under- or overestimation of flood quantiles, particularly when trying to adequately model extreme events in basins with a lack of hydro-meteorological data.

In a previous study carried out by Beneyto et al. [17], it was proven the necessity to incorporate additional information in the WG calibration process, especially when estimating low-frequency precipitation quantiles. This former work presented the reduction in uncertainty in the precipitation quantile estimates based on the available information used for the model calibration. The present paper presents an extension of this work aimed at analyzing the uncertainty of flood quantiles estimated by SCS in different scenarios: (1) different precipitation regimes, (2) different climate extremality, and (3) different hydrological characteristics of the basin. Thus, 18 "base" populations were used as benchmarks to analyze the flood quantile uncertainty using Monte Carlo simulations.

The findings of this study highlight the significant influence of the precipitation regime on the estimated flood quantile uncertainty. Although it appears that the basin size and

the hydrological characteristics of the basin do not substantially impact flood quantiles' uncertainty, it has been found that uncertainty clearly increases with the aridity of the climate, which should be considered by practitioners when dealing with flood studies in arid and semi-arid climates. Climate extremality has been proven not to be as significant as expected; however, very low-frequency flood quantiles presented a higher degree of uncertainty when the climate was more extreme. Finally, flood quantiles presented higher uncertainty than their precipitation quantile counterparts, indicating that the HM does not act as a balance component in the SCS approach as expected. Additionally, despite the aim of this study not being to evaluate the performance of either the WG or the HM, some doubts have emerged as to the use of the E-GPD distribution function, which will be studied in further research. The general findings of this research reveal that low-frequency flood quantile estimates by SCS are still far from presenting adequate levels of uncertainty for flood studies if additional information is not integrated into the WG implementation. Integrating a regional $X_{P,T}$ (with a relatively high reliability) in the WG model calibration clearly reduces the uncertainty of flood quantile estimates independently of the climate, the extremality, the drainage area, and the hydrological characteristics of the basin (i.e., in a broad range of hydro-climatic conditions). This uncertainty reduction is greater for flood quantiles with return periods near the regional precipitation quantile used $X_{P,T}$.

These findings carry significant implications for advancing the efficiency of flood risk management, particularly in areas with a lack of hydro-meteorological data and in arid and semi-arid climates characterized by substantial variability in their flood patterns. Results from this research will help practitioners using SCS to obtain more accurate flood quantiles should they be working in an extremely arid climate or a mild continental humid region.

Author Contributions: Conceptualization, C.B., J.Á.A., G.V. and F.F.; Data curation, C.B., J.Á.A. and G.V.; Formal analysis, C.B., J.Á.A., G.V. and F.F.; Investigation, C.B., J.Á.A. and F.F.; Methodology, C.B., J.Á.A. and F.F.; Project Administration, F.F.; Resources, C.B. and J.Á.A.; Software, C.B. and J.Á.A.; Supervision, C.B., J.Á.A., G.V. and F.F.; Visualization, C.B. and J.Á.A.; Writing—Original draft, C.B., J.Á.A., G.V. and F.F.; Writing—review and editing, C.B., J.Á.A., G.V. and F.F. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Spanish Ministry of Science and Innovation through the research projects TETISCHANGE (RTI2018-093717-B-100) and TETISPREDICT (PID2022-141631OB-I00). Funding for the open-access charge has been provided by Universitat Politècnica de València.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Acknowledgments: The authors thank AEMET and the UC for the data provided to carry out this work (Spain02 dataset).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Kidson, R.; Richards, K.S. Flood frequency analysis: Assumptions and alternatives. *Prog. Phys. Geogr.* **2005**, *29*, 392–410. [CrossRef]
2. Kim, D.; Cho, H.; Onof, C.; Choi, M. Let-It-Rain: A web application for stochastic point rainfall generation at ungaged basins and its applicability in runoff and flood modeling. *Stoch. Environ. Res. Risk Assess.* **2017**, *31*, 1023–1043. [CrossRef]
3. Gaume, E. Flood frequency analysis: The Bayesian choice. *Wiley Interdiscip. Rev. Water* **2018**, *5*, e1290. [CrossRef]
4. Grimaldi, S.; Nardi, F.; Piscopia, R.; Petroselli, A.; Apollonio, C. Continuous hydrologic modelling for design simulation in small and ungauged basins: A step forward and some tests for its practical use. *J. Hydrol.* **2021**, *595*, 125664. [CrossRef]
5. Salazar-Galán, S.; García-Bartual, R.; Salinas, J.L.; Francés, F. A process-based flood frequency analysis within a trivariate statistical framework. Application to a semi-arid Mediterranean case study. *J. Hydrol.* **2021**, *603*, 127081. [CrossRef]
6. Verdin, A.; Rajagopalan, B.; Kleiber, W.; Katz, R.W. Coupled stochastic weather generation using spatial and generalized linear models. *Stoch. Environ. Res. Risk Assess.* **2015**, *29*, 347–356. [CrossRef]
7. Cavanaugh, N.R.; Gershunov, A.; Panorska, A.K.; Kozubowski, T.J. On the Probability Distribution of Daily Precipitation Extremes. *Geophys. Res. Lett.* **2015**, *42*, 1560–1567. [CrossRef]
8. Soltani, A.; Hoogenboom, G. Minimum data requirements for parameter estimation of stochastic weather generators. *Clim. Res.* **2003**, *25*, 109–119. [CrossRef]

9. Evin, G.; Favre, A.C.; Hingray, B. Stochastic generation of multi-site daily precipitation focusing on extreme events. *Hydrol. Earth Syst. Sci.* **2018**, *22*, 655–672. [CrossRef]
10. Ahn, K.H. Coupled annual and daily multivariate and multisite stochastic weather generator to preserve low- and high-frequency variability to assess climate vulnerability. *J. Hydrol.* **2020**, *581*, 124443. [CrossRef]
11. Beneyto, C.; Aranda, J.Á.; Benito, G.; Francés, F. New approach to estimate extreme flooding using continuous synthetic simulation supported by regional precipitation and non-systematic flood data. *Water* **2020**, *12*, 3174. [CrossRef]
12. Tegegne, G.; Kim, Y.O.; Seo, S.B.; Kim, Y. Hydrological modelling uncertainty analysis for different flow quantiles: A case study in two hydro-geographically different watersheds. *Hydrol. Sci. J.* **2019**, *64*, 473–489. [CrossRef]
13. Pluntke, T.; Pavlik, D.; Bernhofer, C. Reducing uncertainty in hydrological modelling in a data sparse region. *Environ. Earth Sci.* **2014**, *72*, 4801–4816. [CrossRef]
14. Faramarzi, M.; Abbaspour, K.C.; Ashraf Vaghefi, S.; Farzaneh, M.R.; Zehnder, A.J.B.; Srinivasan, R.; Yang, H. Modeling impacts of climate change on freshwater availability in Africa. *J. Hydrol.* **2013**, *480*, 85–101. [CrossRef]
15. Metzger, A.; Marra, F.; Smith, J.A.; Morin, E. Flood frequency estimation and uncertainty in arid/semi-arid regions. *J. Hydrol.* **2020**, *590*, 125254. [CrossRef]
16. Moges, E.; Demissie, Y.; Larsen, L.; Yassin, F. Review: Sources of hydrological model uncertainties and advances in their analysis. *Water* **2021**, *13*, 1–23. [CrossRef]
17. Beneyto, C.; Ángel, J.; Francés, F. Exploring the uncertainty of Weather Generators' extreme estimates in different practical available information scenarios. *Hydrol. Sci. J.* **2023**, 1203–1212. [CrossRef]
18. Sun, X.; Mein, R.G.; Keenan, T.D.; Elliott, J.F. Flood estimation using radar and raingauge data. *J. Hydrol.* **2000**, *239*, 4–18. [CrossRef]
19. Bárdossy, A.; Das, T. Influence of rainfall observation network on model calibration and application. *Hydrol. Earth Syst. Sci.* **2008**, *12*, 77–89. [CrossRef]
20. Moulin, L.; Gaume, E.; Obled, C. Uncertainties on mean areal precipitation: Assessment and impact on streamflow simulations. *Hydrol. Earth Syst. Sci.* **2009**, *13*, 99–114. [CrossRef]
21. Papastathopoulos, I.; Tawn, J.A. Extended generalised Pareto models for tail estimation. *J. Stat. Plan. Inference* **2013**, *143*, 131–143. [CrossRef]
22. Naveau, P. Modeling jointly low, moderate, and heavy rainfall intensities without a threshold selection. *Water Resour. Res.* **2016**, *52*, 2753–2769. [CrossRef]
23. Francés, F.; Vélez, J.I.; Vélez, J.J. Split-parameter structure for the automatic calibration of distributed hydrological models. *J. Hydrol.* **2007**, *332*, 226–240. [CrossRef]
24. Herrera, S.; Fernández, J.; Gutiérrez, J.M. Update of the Spain02 gridded observational dataset for EURO-CORDEX evaluation: Assessing the effect of the interpolation methodology. *Int. J. Climatol.* **2016**, *36*, 900–908. [CrossRef]
25. Kotlarski, S.; Szabó, P.; Herrera, S.; Räty, O.; Keuler, K.; Soares, P.M.; Cardoso, R.M.; Bosshard, T.; Pagé, C.; Boberg, F.; et al. Observational uncertainty and regional climate model evaluation: A pan-European perspective. *Int. J. Climatol.* **2017**, 3730–3749. [CrossRef]
26. De Martonne, E. L'indice d'aridité. *Bull. L'association Géographes* **1926**, *9*, 3–5. [CrossRef]
27. Cameron, D.S.; Beven, K.J.; Tawn, J.; Blazkova, S.; Naden, P. Flood frequency estimation by continuous simulation for a gauged upland catchment (with uncertainty). *J. Hydrol.* **1999**, *219*, 169–187. [CrossRef]
28. Hall, J.; Solomatine, D. A framework for uncertainty analysis in flood risk management decisions. *Int. J. River Basin Manag.* **2008**, *6*, 85–98. [CrossRef]
29. WMO. *Guide to Climatological Practices*; WMO: Geneva, Switzerland, 2011; ISBN 9789263101006.
30. Ortega, J.A.; Razola, L.; Garzón, G. Recent human impacts and change in dynamics and morphology of ephemeral rivers. *Nat. Hazards Earth Syst. Sci.* **2014**, *14*, 713–730. [CrossRef]
31. Pilgrim, D.H.; Chapman, T.G.; Doran, D.G. Problèmes de la mise au point de modèles pluie-écoulement dans les régions arides et semi-arides. *Hydrol. Sci. J.* **1988**, *33*, 379–400. [CrossRef]
32. Dunne, T.; Black, R.D. Partial Area Contributions to Storm Runoff in a Small New, England Watershed. *Water Resour. Res.* **1970**, *6*, 1296–1311. [CrossRef]
33. Hongwei, Z.; Xuehua, Z.; Bao'an, Z. Application of Developed Grid-GA Distributed Hydrologic Model in Semi-Humid and Semi-Arid Basin. *Trans. Tianjin Univ.* **2009**, *15*, 70–74.
34. Bafitlhile, T.M.; Li, Z. Applicability of ε-Support Vector Machine and artificial neural network for flood forecasting in humid, semi-humid and semi-arid basins in China. *Water* **2019**, *11*, 85. [CrossRef]
35. Butts, M.B.; Payne, J.T.; Kristensen, M.; Madsen, H. An evaluation of the impact of model structure on hydrological modelling uncertainty for streamflow simulation. *J. Hydrol.* **2004**, *298*, 242–266. [CrossRef]
36. Gabellani, S.; Boni, G.; Ferraris, L.; von Hardenberg, J.; Provenzale, A. Propagation of uncertainty from rainfall to runoff: A case study with a stochastic rainfall generator. *Adv. Water Resour.* **2007**, *30*, 2061–2071. [CrossRef]