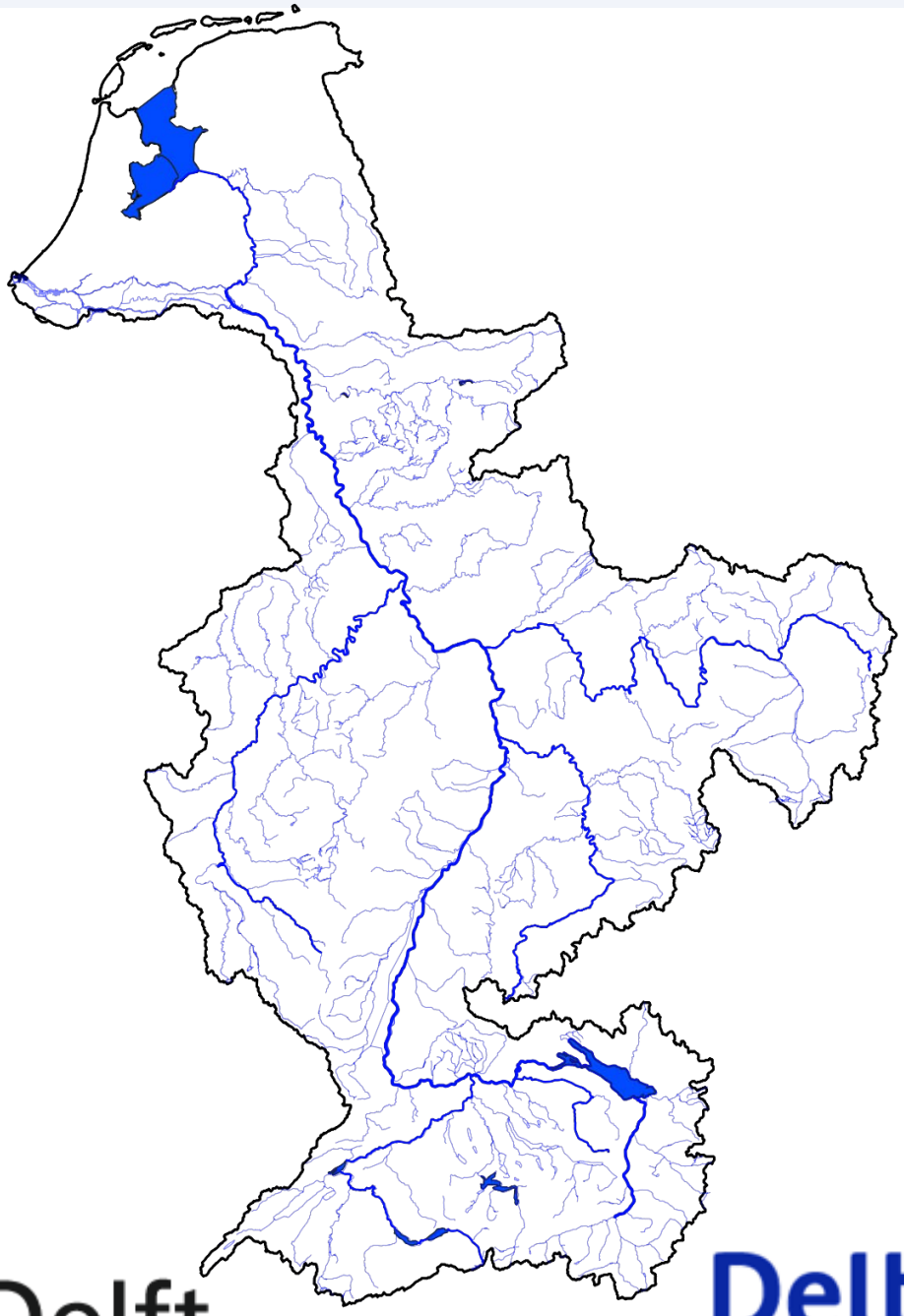
M.d.S. Fonseca Cerda

Flood risk analysis for river confluences

Evaluation of the use of long synthetic time series for the Rhine River



TU Delft

Deltares

Flood risk analysis for river confluences

Evaluation of the use of long synthetic time series for the Rhine River

By

M.d.S. Fonseca Cerda

in partial fulfilment of the requirements for the degree of

Master of Science
in Civil Engineering

at the Delft University of Technology,
to be defended publicly on September 30th 2022.

Supervisor:	Dr. ir. O. Morales Napoles,	TU Delft
Thesis committee:	Dr. ir. W. Luxemburg,	TU Delft
	Dr. ir. H. Winsemius,	Deltates
	Dr. ir. F. Diermanse,	Deltares

Preface

This thesis is the culmination of my MSc programme in Water Management at TUDelf. My dream started long ago, whit the idea of pursuing this personal goal in order to be able to help improve society by addressing the current and future flood problems. It's been 2 years of being away from home, with ups and downs along the way, but mostly full of joy, 'de moeite waard'. This achievement represents the culmination of one of my greatest goals, and the start of a new phase in my career.

I would like to express my deepest gratitude to all my committee members: Oswaldo, Wim, Hessel and Ferdinand. Special thanks to my daily supervisors Hessel and Ferdinand, for their expertise and willingness to explain and help. They always guided me and showed support during our weekly meetings, following my enthusiasm. They were always available and checking on me, especially during the corona lockdowns.

Words cannot express my gratitude to the board of the Justus en Louise van Effen Foundation, who awarded me with the Justus en Luis van Effen Excellence Scholarship 2020-2022 to study my MSc Programme at Delft University of Technology.

I had the pleasure of collaborating with Deltares, which provided the computational resources, the data to be able to do this thesis, and a great working environment. Special thanks to Anaïs, who has given guidance and helped with the setup of the hydraulic model, and who always showed enthusiasm for my research. Thanks to the colleagues from the Catchment & Urban Hydrology and Flood Risk Management departments for the many talks, lunches and borrels we had all along the way.

Last but not least, I would like to extend my sincere thanks firstly to my family (gezin en familie), who have always supported me even though I am far from them, and secondly to my friends, the ones I have met on this journey and have been closely supporting me, and the ones from Mexico who always sent the motivation to continue.

M.d.S. Fonseca Cerda
Delft, September 2022

Abstract

Floods are among the most common and devastating natural hazards worldwide. A key challenge is to select and implement efficient measures to reduce the flood risk, within limitations of budget, time, space, and societal acceptance. Interactions between joining rivers in flood plains and cities lead to several challenges in flood risk studies. Typical 'critical' events of such river systems may differ as small (large) rivers are affected by short duration-high intensity (long duration large volume) rainfall events. The challenges are 1) to define design flood conditions (design events) for systems on which the flood resistance measures can be based, e.g. dimensions of measures like levees or reservoirs, and 2) to understand and determine the risk and consequences that a flood can cause. A multivariate[1] analysis is needed to account for the interactions and statistical dependencies (Bender et al., 2016). Copula functions have been applied in different fields as a multivariate method. Nevertheless, its application in risk analysis for river confluences still needs further evaluation.

In this thesis, the objective is to develop a methodology to determine design flood events that account for the different statistical dependencies and interactions between joining rivers, and that balances the required simulation time with the required accuracy of the results. For this objective, we evaluated the different statistical dependencies and interactions of joining rivers according to the catchment characteristics for each model domain: meteorological, hydrological, and hydraulic. First, identifying how the extreme precipitation events (meteorological) of neighbouring catchments are correlated, and evaluating the differences whit correlations of extreme river discharge events (hydrological). Second, we performed hydraulic simulations at a confluence to evaluate the hydraulic interactions of the joining rivers and the flood impacts, from which a response function was obtained. Subsequently, we evaluated the flood risk by implementing an approach to sample combinations of discharges of joining rivers, and determining the flood impacts from the response function. The developed approach consists of three main steps: 1) the selection of the extremes sets at the confluence, 2) the estimation of the copula parameters and 3) Monte Carlo simulations where the discharges of joining rivers are sampled from their respective marginal probability distributions, and the flooded area is calculated by using the response function. The dependence between the two joining rivers is taken into consideration by using a copula (Gaussian, Gumbel, or Clayton) to construct the joint distribution of the confluence from the marginal distributions.

The meteorological and hydrological evaluation of statistical dependencies was done for the Rhine River basin, where 74 confluences were selected and evaluated by using the synthetic time series of 50,000 years of precipitation and discharge, developed by Deltares and KNMI. Annual maxima sets were identified, and the Spearman rank correlation was calculated between the mainstream peaks and the tributary stream peaks for each confluence (annual maxima and concurrent flow, depending on the selection method), which was used to evaluate the statistical dependencies between neighbouring catchments. The results showed that confluences of two smaller rivers are higher correlated than confluences in which at least one larger river is involved. From this, we conclude that in small, upstream confluences, the extreme events at neighbouring catchments are concurrent with precipitation events which cover in time and space both catchments, whereas in large, downstream confluences, where fewer dependencies were observed, the extreme events may be driven by a combination of different events from upstream areas.

The hydraulic simulations and the flood risk analysis were done for the confluence where the Main River joins the Rhine River, located at the downstream end of the Upper Rhine region. 25 hydraulic simulations of possible discharge combinations were carried out and the response function was calculated: flooded area as a function of both river discharges. The Monte Carlo analysis was carried out in a bootstrapping experiment in such a way that time series of *N* years were sampled from the full set of 50,000 years. For each set of *N* years, after *M* Monte Carlo iterations, we constructed an empirical distribution of the flooded areas to estimate the flooded area that corresponds to a certain probability of exceedance (or return period). The results showed that the Rhine River has more influence on the flood levels. Moreover, from the evaluated copulas, the Gaussian copula led to the closest estimation to the benchmark (which consisted of all 50,000 annual maxima), regardless of the size of the sample *N*. Additionally, a sensitivity analysis indicated that variations

[1] The multivariate analysis of this report is limited to two dimensional distributions (bivariate)

in the results are introduced by the selection of the marginal distribution, for which the size of the sample *N* does influence the variability: the larger the sample, the smaller the variation of the results, which is according to expectation.

The developed approach can be applied for tributary flood risk analysis, which allows selecting the copula and marginals that fit the best each specific confluence.

The available 50,000 years of synthetic data were generated by Deltares and KNMI from a limited set of observations. Therefore, we suggest performing the analysis using the actual records to evaluate the differences/similarities with the results of the current report. Additionally, the results can also be compared to results that are obtained when using a deterministic approach that does not consider statistical dependencies and results when using other multivariate methods that do not involve the use of copulas. Moreover, we only evaluated one of the confluences in detail, hence, we suggest evaluating more confluences to check if the conclusions hold for a confluence with similar characteristics, and to evaluate the different hydraulic interactions and sensitivity of the flood risk approach for confluences with different characteristics.

Contents

List of Figures

List of Tables

List of Acronyms and Symbols

AIC	Akaike Information Criterion
AM	Annual Maxima
A_T	Flooded area that corresponds to a certain return period
BIC	Bayesian Information Criterion
C	Confluence
CDF	Cumulative distribution function
C_{max}	Maxima of the sum of both streams (C=MS+TS)
CM_n	Cramér-von-Mises
CV	Coefficient of variation
DEM	Digital Elevation Model
EVT	Extreme value theory
GEV	Generalized extreme value
GLUE	Generalized Likelihood Uncertainty Estimation
GPD	Generalized Pareto distribution
GRADE	Generator of RAinfall and Discharge Extremes
GUM	Gumbel distribution
h_{max}	Maximum water depth
HVB	Hydrologiska Byråns Vattenbalansavdelning
HYD	Hydrological model domain
k	Duration of the time series in years
MET	Meteorological model domain
MS	Mainstream
MS_{conc}	Concurrent mainstream
MS_{max}	Mainstream maxima
N	Size of the random sample
N_{MC}	Number of Monte Carlo samples
NOR	Normal distribution
N_t	Number of extreme events
$\hat{P}_{exc}$	Exceedance probability
POT	Peaks-over-Threshold
RMSE	Root mean squared error
r_s	Spearman rank correlation
SFINCS	Super-Fast Inundation of CoastS hydraulic model
T	Return period
TS	Tributary stream
TS_{conc}	Concurrent tributary stream
TS_{max}	Tributary stream maxima
Z_i, Z_j	Standard normal variables
λ	Average number of threshold exceedances that occur per year
ρ	Pearson's correlation
τ	Kendall rank correlation

1 Introduction

1.1. Context

Floods are among the most common and devastating natural hazards worldwide. The frequency, magnitude, and impacts of floods are projected to be exacerbated in the (near) future as a result of climate change, sea-level rise, and population growth. Flood risk reduction is therefore high on the agenda of decision-makers. A key challenge is to select and implement efficient measures to reduce the flood risk, within limitations of budget, time, space, and societal acceptance. Such decision-making needs to be supported with detailed knowledge of the existing flood risk as well as the flood risk for various scenarios of climate change, socio-economic developments, and potential flood reducing measures.

Interactions between joining rivers in floodplains and cities lead to several challenges in flood risk studies. Typical 'critical' events of such river systems may differ as small (large) rivers are affected by short duration-high intensity (long duration large volume) rainfall events. The challenges are 1) to define design flood conditions (design events) for systems on which the flood resistance measures can be based, e.g. dimensions of measures like levees or reservoirs, and 2) to understand and determine the risk and consequences that a flood can cause in order to provide flood resilience measures for recovery such as insurances and/or governmental compensations.

This research will contribute to the emerging field of compound flood analysis. First, the research aims not only to have a better understanding of compound floods but also to evaluate how catchment characteristics influence the interactions between rivers. Hence, we aim to determine design events in a more effective way that can be translated to time and cost reduction when performing the risk assessment. Furthermore, the design events have an impact on society in terms of safety and economy. Such design events, and associated designs of risk-reducing measures, may improve people's safety, but also have a cost impact on flood resistance and resilience measures. Moreover, they can also play an important role in the design and implementation of public policies related to flood protection, helping policymakers to make rational choices.

1.2. Problem statement

Univariate analyses have been widely used to assess flood risk, which is constructed based on a selected characteristic (e.g. peak, volume, water level) with target flood frequency (Candela et al., 2014; Grimaldi & Serinaldi, 2006; as cited in Zhou et al., 2019). However, for joining rivers, a multivariate[2] analysis is needed to account for the interactions between the joining rivers and their statistical dependencies (Bender et al., 2016). Moreover, the individual rivers may have different characteristics. Therefore, it is important to be able to model those differences to provide more accurate results.

Even though multivariate analysis is more common nowadays, when implementing it employing traditional bivariate or trivariate distributions (such as normal, lognormal, exponential, gamma and extreme value) all flood components are represented using the same distribution function, therefore the results may still not be accurate as each flood component may have a different distribution. As a step forward, copula functions have been applied in different fields as a multivariate method, overcoming the difficulties of the traditional multivariate functions, allowing the use of different distributions for each flood component, and showing successful results in other studies as mentioned in Tosunoglu et al., (2020). Nevertheless, its application in risk analysis for river confluences still needs further evaluation.

Moreover, the interactions and statistical dependencies of joining rivers may vary according to the scale and meteorological, hydrological, and hydraulic model domain. Therefore, it is important to evaluate the model

[2] The multivariate analysis of this report is limited to two dimensional distributions (bivariate)

domain variations and to identify at what scale they are relevant for the flood risk assessment in order to overcome possible simulation time constraints while balancing the obtained accuracy in the results.

The use of long synthetic time series allows us to perform a significantly more robust risk assessment. Hence, if applied correctly, the results become more accurate. As shown in Deltares (2021), when considering statistical dependencies between joining rivers in combination with long synthetic time series, not only the uncertainties of the results are significantly reduced, but also the design discharge can be lower than when using the design event approach. This can be of importance when designing flood resistance measures as it can be translated into a reduction of construction costs. Unfortunately, it is not always possible to generate such long time series due to limitations in time and budget. Therefore, it is of relevance to evaluate when, where, and how they should be implemented.

1.3. Objective and research questions

The research objective is to develop a methodology to determine design flood events that account for the different statistical dependencies and interactions between joining rivers, and that balances the required simulation time with the required accuracy of the results.

In order to reach the research objective, the following research question needs to be answered:

What are the interactions between joining rivers according to the meteorological, hydrological, and hydraulic model domains, and how can we account for such interactions in order to determine design flood events that balance the required simulation time and the accuracy of the results?

To be able to answer the main research question, we propose three sub-questions:

1 How do the different catchment characteristics and the meteorological and hydrological model domains influence the correlations between joining rivers?

The interactions between joining rivers and their spatial correlations vary according to not only the meteorological but also the hydrological processes that occur in the catchment at different temporal and spatial scales. Moreover, the catchment characteristics also influence such interactions. The following questions will be addressed:

- Which are the meteorological processes involved in the study area and how are they related to the catchment characteristics?
- How are meteorological (precipitation) extreme events spatially correlated and how do they influence the correlations between joining rivers?
- Which are the relevant hydrological processes involved in the study area and how are they related to the catchment characteristics?
- How are hydrological (discharge) extreme events spatially correlated and how do they influence the correlations between joining rivers?
- What are the differences between meteorological and hydrological model domain interactions according to the catchment characteristics?

2 What are the hydraulic interactions between joining rivers?

Due to the time it takes to perform hydraulic simulations, it is necessary to identify the potential locations where hydraulics play a major role. The results of the first experiment (first sub-question) could be a first indication for the selection. Hence, locations, where the hydrological model domain reveals low statistical dependencies in the results, should be analysed in more detail. Moreover, the topography of the area also influences the selection of the locations: in the steepest areas, we would not expect relevant hydraulic interactions between adjoining rivers because of limited backwater effects, as opposed to flat areas.

Moreover, the hydrological processes in the catchments of the joining rivers also play a role in the hydraulic interactions, therefore they need to be evaluated and considered in the simulations. The following questions will be addressed:

- What characteristics should we consider when identifying confluences where interactions might be relevant?
- How can the hydrological processes in different joining tributaries be considered in setting up probabilistic hydraulic modelling simulations?
- Which hydraulic events need to be simulated for carrying out a flood risk analysis?

3 and how should hydrological events be sampled for performing simulations in a flood risk analysis and/or determining design flood events?

The choices we make when sampling the hydrological events will have an impact on the results. Therefore, it is important to evaluate their sensitivity to the sampling approach. First, different methods of extreme event selection can be applied. Second, we need to select the copula that 'best' accounts for the statistical dependencies between the joining rivers. Third, for joining rivers there is more than one possibility of extreme situations that can lead to a flooding event at the confluence, therefore different ways of selecting the extreme events should be analysed and compared. Finally, long time series lead to a more accurate result (less uncertainty). Unfortunately, long time series are generally not available, hence, it is important to evaluate the impact of the time series length on the accuracy of the results. The following questions will be addressed:

- Which extreme event selection method should be used?
- Which copulas should be applied as a bivariate method?
- How should the sets of simultaneous and non-simultaneous events at the confluence be selected and applied?
- What is the impact of the length of available time series on the uncertainty of the derived flood risk?
- How sensitive are these choices for the sampling approach?

Summarizing, throughout this research, we aim to evaluate the different statistical dependencies and interactions of joining rivers according to the catchment characteristics for each model domain: meteorological, hydrological, and hydraulic. First, identifying how the extreme meteorological events of neighbouring catchments are related, and evaluating the differences with respect to the extreme hydrological events. Second, performing hydraulic simulations at a confluence to evaluate the hydraulic interactions of the joining rivers. Subsequently, a response function from the hydraulic evaluation can also be obtained in order to analyse flood risk for hydrological extreme events. Subsequently, we need to determine how to sample hydrological events to perform a flood risk analysis. Finally, determining when, where, and how the use of synthetic data and the different model domains (meteorological, hydrological, and hydraulic) should be applied in order to model compound flooding of joining rivers for analysing flood risk of joining rivers.

1.5. Materials and methods

The applied research is experimental since it aims to evaluate the statistical dependence of the different model domains (meteorological, hydrological, hydraulic) of joining rivers. This is established through the use and evaluation of several experiments. The evaluation will be quantitative and the objectives are exploratory since it is intended to not only explore the field of compound floods but also the use of long synthetic time series. Moreover, it is aimed to account for the statistical dependencies between relevant variables.

The study area of this research is the Rhine River Basin, for which we use 50,000 years of meteorological and hydrological data generated by a combined stochastic-physically based model GRADE (Generator of RAinfall and Discharge Extremes). More information of both the study area and the data is provided in Section 3.

The framework of analysis is shown in Figure 1 and described below.

First, during the literature review not only the state of the art on flood risk analysis is investigated but also the theories to be used and the methodology requirements. Subsequently, the characteristics of the case study area are determined and described in detail. Moreover, the data is explored to verify that it is a reliable representation of the temporal and spatial characteristics. Next, the data is explored according to the different model domains, performing different experiments to answer the research questions. This is done as follows:

1 How do the different catchment characteristics and the meteorological and hydrological model domains influence the correlations between joining rivers?

The meteorological processes in the area of study are identified and evaluated, paying attention to the catchment location and characteristics. An analysis is carried out of statistical dependencies of precipitation over joining rivers at different catchment scales, focusing on spatial correlations, and spatial patterns.

Similar to the meteorological evaluation, the hydrological processes in the area of study are identified and evaluated, and attention is paid to the catchment location and characteristics. An analysis is carried out of statistical dependencies between (peak) discharges of joining rivers in a similar manner as the catchment precipitation.

Subsequently, the results of both model domains (meteorological and hydrological) are compared to identify which catchment characteristics and scales influence the correlation between the joining rivers at the different model domains (meteorological and hydrological).

2 What are the hydraulic interactions between joining rivers?

The results of the first experiment are evaluated. A confluence is selected for which [1] the statistical dependence is weak and [2] the area near the floodplain is relatively flat, as it is expected that such a location is most sensitive to choices made in the approach. The second criterion makes that hydraulic interactions between the two joining rivers are likely to be substantial, and the first criterion makes that a proper selection of data and methods is very relevant.

The hydraulic model is set up for the selected location. This is done using SFINCS which is a 2D hydraulic model engine based on simplified Sain-Venant equations, designed for modelling inundation from pluvial, coastal and riverine sources, developed by Deltares. The input of the model is a grid of hydrological extreme events from the synthetic data, that accounts for different possible combinations of discharges of the joining rivers and their hydrological characteristics. Different hydraulic interactions for each possible discharge combination are found, evaluated and compared to identify the hydraulic processes involved and the relationship with the location characteristics.

3 How should hydrological events be sampled for performing simulations in a flood risk analysis and/or determining design flood events?

In order to answer this question, the interactions between joining rivers are evaluated using different approaches in terms of 1) the extreme event selection method, 2) the copula that is implemented, 3) the sets of simultaneous and non-simultaneous events and 4) the length of the data that is used. This is carried out in a similar experiment as the one done in (Deltares, 2021), but this time investigating in detail the sensitivity of every choice.

Selection method approaches are benchmarked against an “ideal case”, i.e. the full 50,000 years of synthetic meteorological and hydrological data for the Rhine River. Later, intermediate steps will be evaluated to determine the potential scenarios where the simulation time can be significantly reduced without significantly affecting the accuracy of the results.

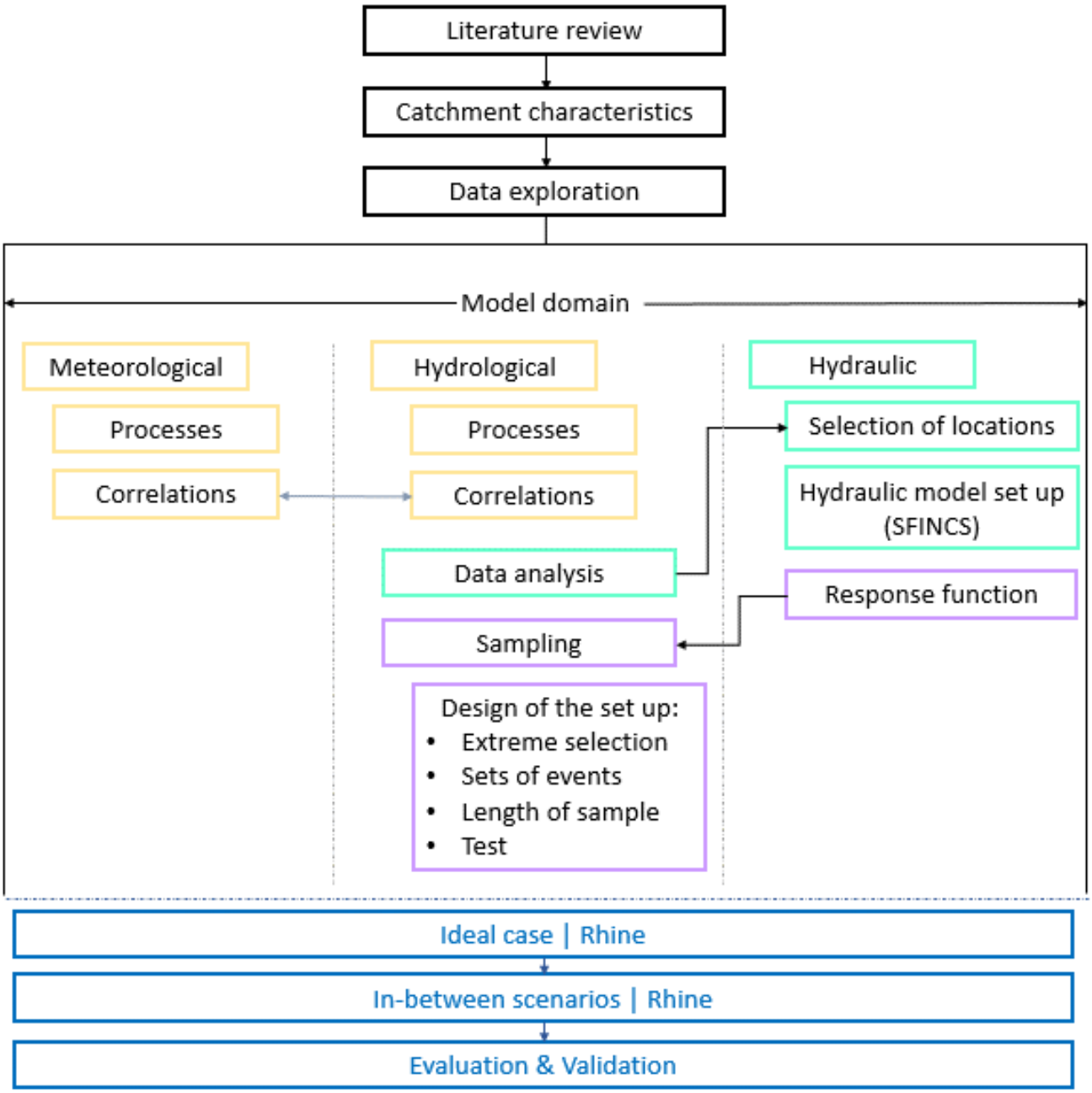


Figure 1 Research process that is followed to answer Research Question 1 (yellow), Research Question 2 (green), and Research Question 3 (purple).

2 Literature review

Most flood control measures require information on return periods (or frequencies) of flood events. (Zhou et al., 2019). A typical flood risk assessment uses a univariate analysis that is constructed based on a selected characteristic (e.g. peak, volume, water level) with target flood frequency (Candela et al., 2014; Grimaldi & Serinaldi, 2006; as cited in Zhou et al., 2019). As explained in Section 2.1, the extreme value theory is often used to select candidate distribution functions to determine the target flood frequencies.

However, when using univariate analysis for joining rivers, only one of the joining rivers is taken into consideration and the dependence between the peak flows of both rivers is not considered. Therefore, multivariate analysis is needed to account for the different interactions between the rivers and their statistical dependencies (Bender et al., 2016). In Section 2.2, we explain the theory for multivariate analysis of compound floods.

2.1 Extreme value theory

The Extreme value theory (EVT) to derive marginal distribution functions has been a fundamental tool in the study of geophysical processes including floods (Zorzetto et al., 2016). Extreme events are selected out of the full data set. The selected events need to be independent, and their magnitude is assumed to follow a specific cumulative distribution function (CDF) that does not change over the period of measurements (stationarity assumption). Two methods of selecting the extremes for the different conditions that can occur at a confluence can be evaluated, Peaks-over-Threshold (POT) and Annual Maxima (AM) which are explained in Sections 2.1.1 and 2.2.2 respectively. Once the parameters of the CDF are determined, the discharge that corresponds to a certain probability/frequency of exceedance (or return period) can be estimated by applying the inverse of the CDF.

2.1.1 Peaks Over Threshold

One of the methods to select the extreme events that can be used is Peaks Over Threshold (POT). According to Pickands' theorem (1975, as cited in Bernardara et al., 2014), the distribution of the data exceeding a given threshold converges as the data set increases towards a generalized Pareto distribution (GPD). Selecting the threshold is crucial to meet the hypothesis of convergence on the GPD and to limit the variability of parameter calibration. Bernardara et al. (2014) describe the two different approaches that can be implemented to select the threshold: 1) Physical declustering by identifying and characterizing independent events and 2) Statistical optimization by setting a threshold for the convergence of the GPD.

The threshold is usually defined by expert prior knowledge, based on the practical consequences of the phenomenon to exceed a certain threshold value (Bernardara et al., 2014). It can be tuned to obtain a certain value of λ, which represents the average number of threshold exceedances that occur per year:

$$\lambda = \frac{N_t}{k} \qquad \text{(Eq. 1)}$$

Where N_t is the number of events, and k the duration of the time series in years. The lower the value of λ the more extreme the selected events are.

Another parameter to be selected when using POT is the window of the event or the time lag. It is used to ensure the independence of the selected events. The most common techniques consist in setting temporal parameters, most of the time based on the minimal time lag between two events (Bernardara et al., 2014).

Once the POT data is derived, the parameters of the GPD need to be determined by fitting the GPD on the selected data. There are several methods available for this task, such as the method of moments, maximum likelihood and probability-weighted moments (Hosking & Wallis, 1997).

2.1.2 Annual Maxima

The second method that can be used to select the extreme events is Annual Maxima (AM), where the event with the maximum magnitude that occurred over 1 year is selected. According to Zorzetto et al, (2016) when selecting the maxima of a certain period length, the distribution converges to a generalized extreme value (GEV) with increasing series length.

Unlike the POT, the extreme selection using AM does not involve crucial decisions such as the selection of the threshold and window. However, it is important to define the start of the hydrological year during the dry season, to avoid duplicating an extreme event for two consecutive years.

Similar to POT, the parameters of the GEV need to be determined by fitting the GEV on the selected data. There are several methods available for this task, such as the method of moments, maximum likelihood and probability-weighted moments (Hosking & Wallis, 1997).

2.2 Compound flood risk analysis

As mentioned before, when using a univariate analysis for tributary compound flood risk analysis, only one of the joining rivers is taken into consideration. Hence, the resulting risk of compound flooding may be under- or overestimated, depending on the assumptions made for the second tributary. Therefore, a bivariate analysis is needed to account for statistical dependencies between joining rivers. First, in Section 2.2.1 we explain how the correlation can be used as a metric to evaluate the statistical dependencies between two variables. Second, in Section 2.2.2 we introduce the copula theory, which is an effective and widely used approach to describe joint probability distributions of correlated variables (Zhou et al., 2019). Subsequently, Monte Carlo simulations can be implemented to construct empirical distributions for estimating the design discharge (Section 2.2.3). Copula functions can be integrated with the Monte Carlo sampling procedure to correctly account for statistical dependencies between variables.

The use of long synthetic time series is not a common practice for compound flood risk analysis. However, as shown in (Deltares, 2021) it allows to perform a significantly more robust risk assessment, hence, the results become more accurate. When considering statistical dependencies between joining rivers in combination with long synthetic time series, not only the uncertainties of the results are significantly reduced, but also the design discharge could be lower (or higher, depending on the choices made) than when using the deterministic approach used in Deltares (2021). Section 2.2.5 describes the possible options we have to select the sets of data to be analysed.

2.2.1 Statistical dependencies | metrics

Correlation coefficients can be used to measure the degree of association between two variables and the direction of the relationship. The coefficient ranges from -1 to 1. The absolute values indicate the strength of the relationship between the two variables. In the case of positive correlation, the two variables tend to increase/decrease simultaneously, resulting in a positive value. Whereas for a negative correlation, one variable tends to increase when the other decrease (Chok, 2010). If the correlation coefficient value is close to 0, the relationship between the variables is weak (Moran, 2022).

Three well-known coefficients are described below: Pearson product-moment correlation, Spearman's rank correlation, and Kendall's tau correlation.

Pearson's correlation

The Pearson product-moment correlation coefficient (Pearson's correlation) is a measure of the strength of a linear relationship between two variables that are bivariate normally distributed. The following formula is used to calculate the Pearson's correlation:

$$\rho = \frac{Cov(X,Y)}{\sigma_X \sigma_Y} \qquad \text{(Eq. 2)}$$

However, it has a disadvantage for bivariate distribution other than the bivariate normal distribution, as Pearson's correlation can be zero for dependent variables (Chok, 2010), neglecting the dependence of the variables.

Spearman correlation

The Spearman rank correlation coefficient measures the strength of a monotonic relationship between two variables. This correlation does not carry any assumptions about the distribution of the data (Moran, 2022) but the variables should have a monotonic relationship, meaning that the size of one of the variables increases/decreases as the other variable increases. The following formula is used to calculate the Spearman rank correlation:

$$r_s = \frac{\sum_{i=1}^{n}((rank(x_i) - \overline{rank(x)})(rank(y_i) - \overline{rank(y)}))}{\sqrt{\sum_{i=1}^{n}(rank(x_i) - \overline{rank(x)}^2 \sum_{i=1}^{n}(rank(y_i) - \overline{rank(y)})^2}} \qquad \text{(Eq. 3)}$$

Kendal correlation

The Kendall rank correlation captures the association between two ordinal variables that are monotonically related. It quantifies the discrepancy between the number of concordant and discordant pairs (Chok, 2010). The following formula is used to calculate the value of the Kendall rank correlation (Moran, 2022):

$$\tau = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} sgn(x_i - x_j) sgn(y_i - y_j)}{n(n-1)} \qquad \text{(Eq. 4)}$$

Where:

$$sgn(x_i - x_j) = \begin{cases} 1 \ if \ (x_i - x_j) > 0 \\ 0 \ if \ (x_i - x_j) = 0 \\ -1 \ if \ (x_i - x_j) < 0 \end{cases} ; sgn(y_i - y_j) = \begin{cases} 1 \ if \ (y_i - y_j) > 0 \\ 0 \ if \ (y_i - y_j) = 0 \\ -1 \ if \ (y_i - y_j) < 0 \end{cases}$$

2.2.2 Copula theory

Copula theory was introduced by Sklar (1959) to model the dependence structure of two or more random variables where a multivariate distribution function can be constructed using the univariate marginal distribution functions in combination with the copula (Dodangeh et al., 2019). Considering *X* and *Y* as a random variables and their respective marginal cumulative distribution functions (CDF) *Fx(x)* and *Fy(y)*, a copula can be constructed to describe the joint behaviour such as *Fx,y(x,y)=C(Fx(x),Fy(y))*, where *Fx,y(x,y)* is the joint cumulative distribution function and *C(Fx(x),Fy(y))* is the copula function (Bender et al., 2016).

The statistical dependence between the joining rivers can be described by a copula function in combination with the (inverse) cumulative distribution function of each river, the mainstream and the tributary stream. Other studies (Bender et al., 2016; Mitková, 2020; Wang et al., 2007; Wang et al., 2009) have demonstrated that copulas can be applied to study compound floods in river confluences. In this study, we analyze three copulas (Gaussian, Gumbel-Hougard, and Clayton) due to their simplicity in the application and to develop the basis for a compound flood risk assessment methodology.

Gaussian Copula

The Gaussian copula is a member of the parametric family and it uses the Spearman rank correlation ρ as a parameter to control the strength of dependence between the variables (Chang, 2019):

$$C(x, y, \rho) = \Phi_2(\Phi^{-1}(x), \Phi^{-1}(y); \rho) \qquad \text{(Eq. 5)}$$

Where Φ is the CDF of a standard normal distribution and $\Phi_2(\cdot; \rho)$ is the joint CDF of (x, y).

Archimedean Copulas

This section is based on Diermanse (2005)

A well-known and widely used type of copula functions are the so-called *Archimedean copulas*. These copulas are related to a generator function $\varphi:[0,1] \rightarrow [0,\infty]$ that has the following properties: For two standard uniformly distributed random variables u and v the following function is a copula function:

$$C_\varphi(u, v) = \varphi^{-1}[\varphi(u) + \varphi(v)] \qquad \text{(Eq. 6)}$$

In other words, C_φ is a bivariate distribution function with marginals that are standard uniformly distributed. All copulas that are generated in this way are called Archimedean copulas. Table 1 contains some Archimedean copulas and their respective generator functions.

Table 1 Some Archimedean copulas and their generator functions (Diermanse, 2005)

name	φ(t)	φ'(t)
Frank:	$-\ln\left[\frac{e^{-\alpha t}-1}{e^{-\alpha}-1}\right]$	$\frac{\alpha}{\left(1-e^{\alpha t}\right)}$
Clayton:	$\frac{1}{\alpha}\left(t^{-\alpha}-1\right)$	$-t^{(-1-\alpha)}$
Gumbel-Hougard:	$\left(-\ln t\right)^{\alpha}$	$\alpha\frac{\left[-\ln(t)\right]^{\alpha}}{t\ln(t)}$

Archimedean copulas have a single parameter α that determines the correlation. However, it is not the same as the Spearman rank correlation coefficient ρ. For starters, α is larger than 1, as opposed to ρ. Parameter α can be derived from Kendall's parameter τ as shown in equations 7 and 8 for the Gumbel and Clayton copulas respectively.

$$\alpha = \frac{1}{1-\tau} \qquad \text{(Eq. 7)}$$

$$\alpha = \frac{2\tau}{1-\tau} \qquad \text{(Eq. 8)}$$

2.2.3 Monte Carlo sampling

Monte Carlo simulations can be used to sample the discharges of joining rivers from their respective marginal probability distributions, taking into account joint probabilities, to generate n samples of the sum of the

discharges of the joining rivers (Q_C=Q_{MS}+Q_{TS}) as shown in Deltares (2021). Having the *n* samples of Q_C, an empirical cumulative distribution can be constructed to estimate the discharge that corresponds to the desired probability (or the desired return period).

Moreover, the discharges of the individual rivers Q_{MS} *and* Q_{TS} can also be sampled individually to account for the interactions of the different possible combinations of discharges while using a response function that accounts for, e.g., the backwater effects, flooding volume, and/or flooding area.

The probability of exceedance (P_{exc}) of a certain threshold (b) can be calculated using Monte Carlo Sampling where for each basic variable X_i *(i=1, …, n)* one simulates *N* realizations $x_{i1}, x_{i2}, \ldots, x_{iN}$. For each set *j (j=1, …, N)* one calculates $g(x_{1j}, x_{2j}, \ldots, x_{Nj})$, in which *g(.)* represents the response of the system to variable *X* . In case *g(.) > b* a counter N_e is increased by one. After *N* simulations it is possible to calculate the exceedance probability:

$$\hat{P}_{exc} = \frac{N_e}{N} \qquad \text{(Eq. 9)}$$

In case $N \rightarrow \infty$ we can obtain the probability of exceedance (P_{exc}). Equation 9 can also be rewritten as:

$$\hat{P}_{exc} = \frac{1}{N}\sum_{j=1}^{N} I[g(x) > b] \qquad \text{(Eq. 10)}$$

Where *I[.]* is the indicator function which is equal to 1 in case the argument of the operator is true and otherwise 0 in case the argument is false (Jonkman et al., 2015)

The required number of samples *N* can be determined according to a target coefficient of variation *CV* (relative error) and probability of exceedance P_{exc}, as shown in the following formula (Jonkman et al., 2015):

$$CV_p = \frac{1}{\sqrt{N\ P_{exc}}} \qquad \text{(Eq. 11)}$$

For example, if the design discharge corresponds to a probability of exceedance P_{exc} *=0.01*, and if the target *CV* is set to 0.05, *N=40,000* Monte Carlo samples are needed.

2.2.4 Goodness of fit test

A goodness of fit test measures the precision of a model to fit a particular sample of observed data, determining the statistical significance of the model (Pham, 2019). Usually, the model is compared with the observations and the difference between the two can be measured in different ways (Pitt et al., 2002), such as the Root mean squared error (RMSE) and the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC), which are described below.

Moreover, when using copulas to model bivariate data, we also want to know which provides the best fit for the observations. There are formal and informal tests which can measure the goodness of fit (Genest & Favre, 2007). The formal methodology is emerging, one of the common statistics that is used is the Cramér-von-Mises (CM_n). Besides, the informal test includes different graphical diagnostics where the scatter plots of the observed data can be compared with the model data (Genest & Favre, 2007). Below, several goodness-of-fit measures that are used throughout this thesis are briefly explained.

Root mean squared error (RMSE)

The RMSE is the square root of the variance of the residuals. This measure quantifies the error, which is the difference between the observed value and the modelled value (Pham, 2019), as follows:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2} \qquad \text{(Eq. 12)}$$

where y_i is the observed value, $\hat{y}_i$ is the fitted value, and N is the number of observations.

A lower value of the RMSE indicates a better fit of the theoretical distribution to the data (Pobočíková et al., 2017).

Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC)

The goodness of fit measures can also take the complexity of a model into account. For example, the AIC was introduced by Akaike (1998; as cited in Pham, 2019) and is calculated as follows:

$$AIC = N \ln\left(\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 / N\right) + 2k \qquad \text{(Eq. 13)}$$

The BIC was introduced by Schwars (1978; as cited in Pham, 2019) and is calculated as follows:

$$BIC = N \ln\left(\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 / N\right) + k \ln(N) \qquad \text{(Eq. 14)}$$

where y_i is the measured (observed) value, $\hat{y}_i$ is the fitted value, N is the number of observations, and k is the number of parameters of a model.

Both AIC and BIC penalize models with many parameters: the penalty is larger when the number of parameters k increases. The difference between the two criteria lies in the way of calculating the penalty term. Whereas in AIC the sample size is not considered, in BIC, the penalty term depends on the sample size N shows how strongly they impact the penalty of the number of parameters in the model (Pham, 2019).

The AIC and the BIC do not test the model in terms of hypothesis testing, but they are useful tools to compare and select models. The smaller the AIC or BIC value, the better the model (Pobočíková et al., 2017).

Cramér-von-Mises (CM_n)

The Cramér-von-Misses statistic can be applied to measure the goodness of fit of copulas. Different attempts have been done to accurately implement it. In this report, we refer to the statistic considered by Genest et al. (2009), for more details the reader is referred to (Couasnon et al., 2018; Genest et al., 2009; Genest & Favre, 2007).

The Cramér-von-Misses statistic compares the squared differences between the empirical copula and the selected parametric copula for a sample of length n (Couasnon et al., 2018):

$$CM_n(u) = \sum_{|u|}\left\{C_{\hat{\theta}_n}(u) - B(u)\right\}^2, \qquad u \in [0,1]^2 \qquad \text{(Eq. 15)}$$

Where $B(u) = \sum 1(U_i \leq u)$ is the empirical copula, and $C_{\hat{\theta}_n}(u)$ is a parametric copula with parameter $\hat{\theta}_n$

The smaller the CM_n, the closer the parametric copula is to the empirical copula.

Semi-correlation test

An approach suggested by Joe (2014, as cited in Couasnon et al., 2018) is based on semi-correlations. First, the pairs of variables *(X, Y)* are transformed to standard normal variables (Z_i, Z_j) as follows:

$$Z_i = \frac{X - \mu}{\sigma};\ Z_j = \frac{Y - \mu}{\sigma} \qquad \text{(Eq. 16)}$$

The transformed variables are split into the four quadrants in the standard normal space, and for each of the quadrants the Pearson's correlation is calculated as follows:

$$\rho_{NE} = \rho\left(Z_i, Z_j | Z_i > 0, Z_j > 0\right) \qquad \text{(Eq. 17)}$$

$$\rho_{SW} = \rho\left(Z_i, Z_j | Z_i < 0, Z_j < 0\right) \qquad \text{(Eq. 18)}$$

$$\rho_{NW} = \rho\left(Z_i, Z_j | Z_i < 0, Z_j > 0\right) \qquad \text{(Eq. 19)}$$

$$\rho_{SE} = \rho\left(Z_i, Z_j | Z_i > 0, Z_j < 0\right) \qquad \text{(Eq. 20)}$$

The Pearson's correlation value for each of the quadrants can give a preliminary indication for the distribution of the data, indicating if the data deviates from a certain parametric copula model (Couasnon et al., 2018).

2.2.5 Event selection

The events to be analyzed are for joining rivers, meaning that for each location where rivers compound there is more than one possibility of extreme situations that can lead to a flooding event. We can differentiate between the mainstream (*MS*), the tributary stream (*TS*), and the confluence (*C*) as shown in Figure 2.

The worst-case scenario would be a simultaneous event in which the maximum of the mainstream (MS_{max}) coincides in time with the maximum of the tributary stream (TS_{max}) (see Figure 3). However, this situation is not likely to always occur (see Figure 4), leading to the need to consider other scenarios as well: mainstream maxima (MS_{max}) with a concurrent discharge of the tributary stream (TS_{conc}), and tributary stream maxima (TS_{max}) with a concurrent mainstream discharge (MS_{conc}) (Bender et al., 2016). Moreover, a third scenario can also be considered at the confluence where the sum of the streams (*C=MS+TS*) represents a maximum of the confluence, even though the individual river discharges of the joining rivers may not be maxima themselves. Hence, the events resulting in peak discharges are not necessarily the same for the three rivers. While the selection of the maxima is a straightforward process, the selection of the concurrent flow can be carried out in different ways: 1) by selecting the concurrent flow at the exact time of the maxima or 2) by selecting the maximum value of the concurrent flow within the same window of the maxima. In this report, the following three data sets are considered:

- *Set 1 | MSmax-TSconc: mainstream maxima (MS_{max}) with a concurrent tributary stream (TS_{conc}),*
- *Set 2 | TSmax-MSconc: tributary stream maxima (TS_{max}) with a concurrent mainstream (MS_{conc}),*
- *Set 3 | Cmax (C=MS+TS): the maxima of the sum of both streams.*

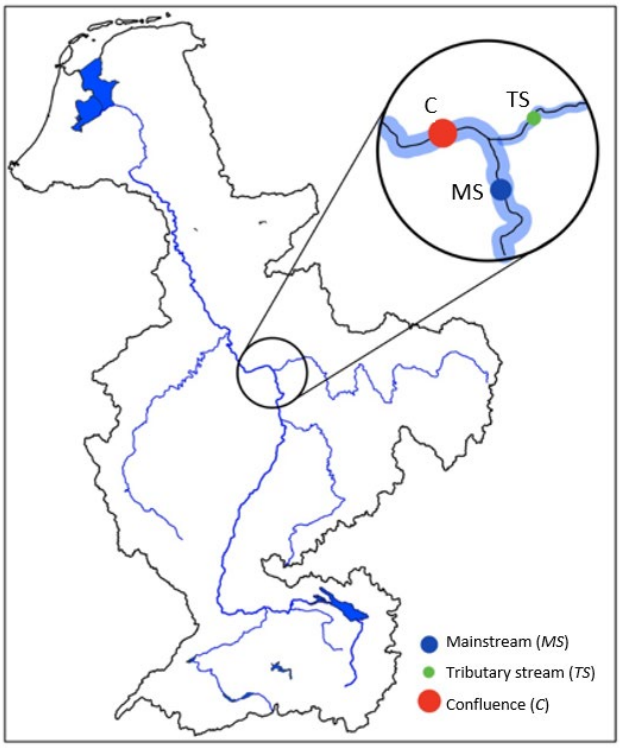


Figure 2 Example of a confluence, when two rivers join is possible to differentiate between the Mainstream (*MS*), the Tributary stream (*TS*) and the Confluence (*C*) (right after the rivers join).

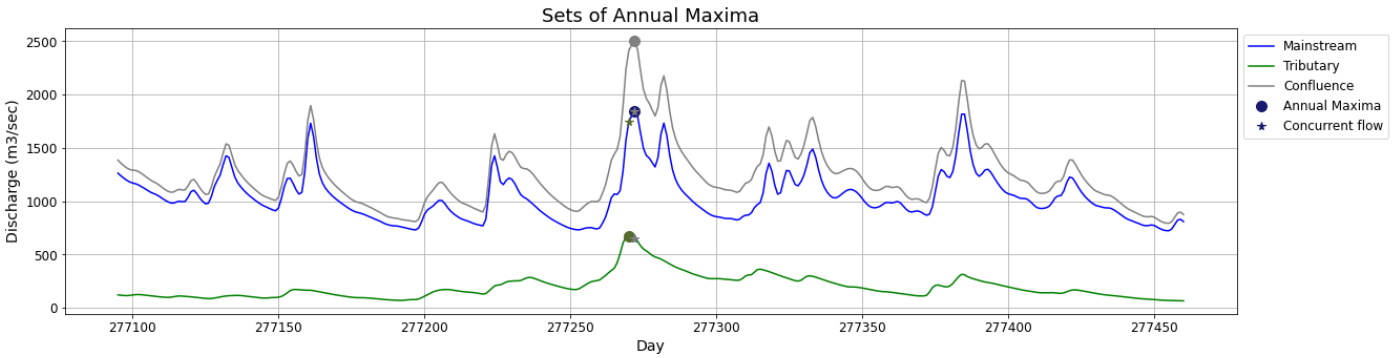


Figure 3 Simultaneous event in which the annual maxima of the mainstream (MS_{max}) coincides in time with the annual maxima of the tributary stream (TS_{max}) and with the annual maxima of the confluence (C_{max})

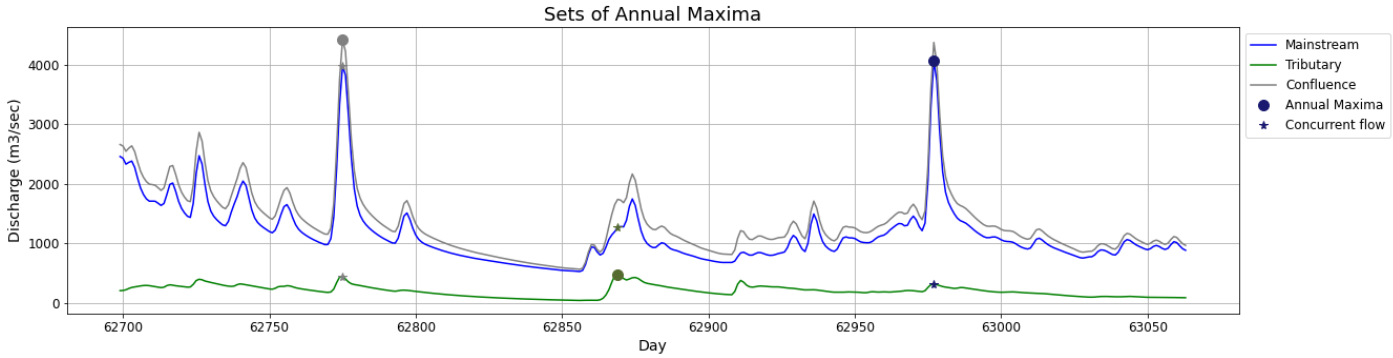


Figure 4 Non-simultaneous event in which the annual maxima of the mainstream (MS_{max}) does not coincide in time with the annual maxima of the tributary stream (TS_{max}), nor with the annual maxima of the confluence (C_{max})

As shown in (Deltares, 2021) the three methods lead to different sets of peaks, demonstrating that the peaks of both rivers do not necessarily occur simultaneously and each set of peaks' combinations can occur and lead to a flood.

3 Case study

3.1. The Rhine River Catchment

3.1.1 Rhine River

This section is based on Uehlinger et al., (2009) and Hegnauer et al., (2014).

The Rhine basin is located in Central and Western Europe, covering 185,260 km^2 in 9 countries: Austria, Belgium, France, Germany, Italy, Liechtenstein, Luxemburg, the Netherlands, and Switzerland. The river is approximately 1,250 km long with an average river flow of 2,300 m^3/s. The main tributaries are Aare, Neckar, Main, and Moselle.

The Rhine River is situated in one of the most important economic regions of Europe. Around 58 million people live along the river and transportation has been historically one of the most important uses of the river. The river is also used for power generation, industrial production, drinking water, and irrigation. These human interventions also affect the hydrology of the river.

The river can be divided into 6 distinct sections: Alpine, High, Upper, Middle, Lower, and Delta Rhine (see Figure 5), which are distributed in three main physiographic regions from south to north: the European highlands (alps), the central upland and plateau regions, and the northern lowland with the coastal plain.

Rhine floods show regional patterns due to the different meteorological conditions and the different hydrological responses of each sub-basin (IKHR 1999, as cited in Uehlinger et al., 2009). This means flood events do not occur simultaneously all over the basin. The flood characteristics are determined by different factors such as the precipitation, topography, soil, vegetation and the hydrological state of the catchment. For example, in the Lower Rhine, flood events are typically caused by multi-day rainfall events in combination with saturated soils during the winter and early spring. In the Alpine Rhine, in Switzerland, floods typically occur in late spring or early summer. From a different genesis, occasionally, snow melt in combination with frozen soils leads to more extreme runoff. The volume and the peak discharge strongly depend on the spatial extent and the sequence of rainfall events due to the size and shape of the Rhine basin. Moreover, the coincidence of the peak flows from different tributaries is important.

3.1.2 Climatic and catchment characteristics

This section is based on Uehlinger et al., (2009) and Hegnauer et al., (2014).

The climate is determined by the location, changing gradually from continental to maritime climate from South to North. General weather patterns during winter are primarily influenced by weather dynamics in the northern and eastern Atlantic and the North Sea.

Temperature and precipitation vary considerably with altitude and local topography. The mean annual temperature of the basin is 8.3°C, 11.2°C in the valley of the Upper Rhine, and <0°C at higher elevations (>3000 m+msl). Precipitation can occur at any time of the year and it averages 945 mm/year. The orographic effect of mountain ranges or uplands results in heterogeneous precipitation patterns at different spatial scales (more details in Table 1). Whereas for the Alpine region ~30% of the annual precipitation falls during summer, for the lower regions seasonal differences are slightly less pronounced.

The elevations of the basin range from 3614 m+msl in the Alpine Rhine to 0 m+msl in the Delta Rhine. The longitudinal profile of the Rhine River is characterized by three base levels of erosion: 1) Lake Constance, 2) the quartzite reef at the beginning of the Middle Rhine, and 3) the sea (see Figure 6 and Figure 7).

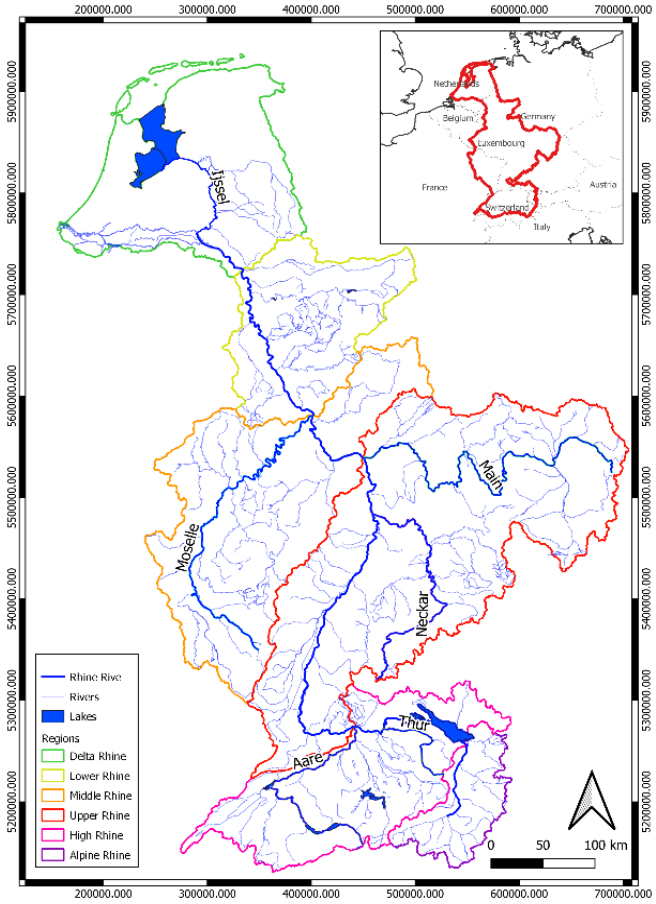


Figure 5 Regions and drainage network of the Rhine River Basin. Data retrieved from Eionet and FAO-UN Land and Water Division

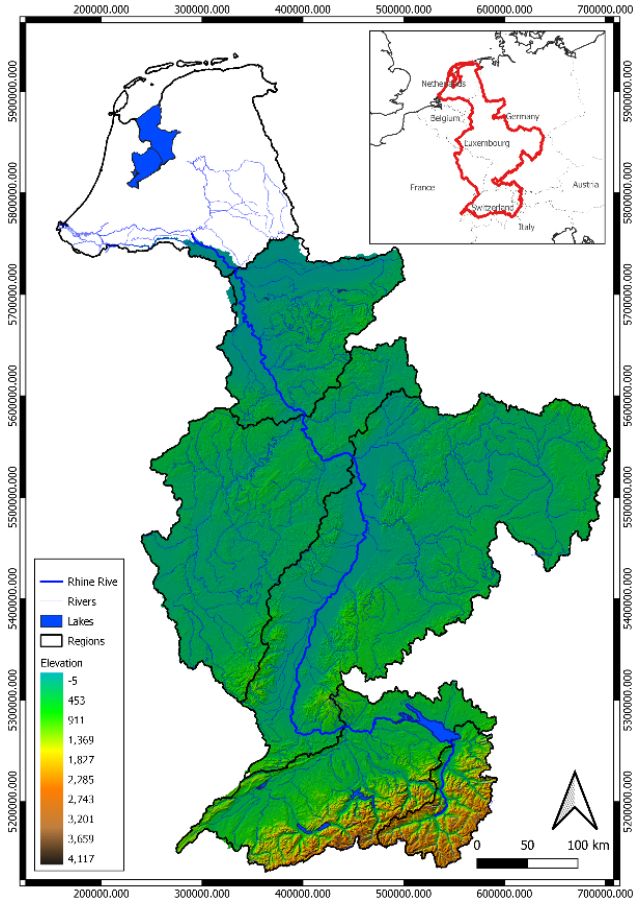


Figure 6 Digital Elevation Model (DEM) of the Rhine River Basin. Data retrieved from NASA's EarthData

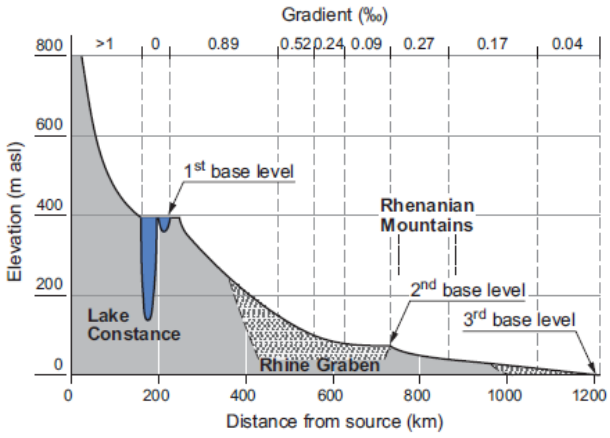


Figure 7 Longitudinal profile of the Rhine River (Uehlinger et al., 2009)

Different hydrological processes are involved, and it is reflected in different discharge patterns along the river due to changes in the contribution of runoff. This shift in the annual cycle of the discharge from upstream to downstream is a typical characteristic of the Rhine River (see Figure 8). During the winter months, snow is stored in the Alps and runoff is limited during this period. During spring, the accumulated snow in the Alps starts to melt, causing the shift in the discharge in time. The downstream part of the basin is under the influence of an oceanic climate and less predictable rain-dominated runoff.

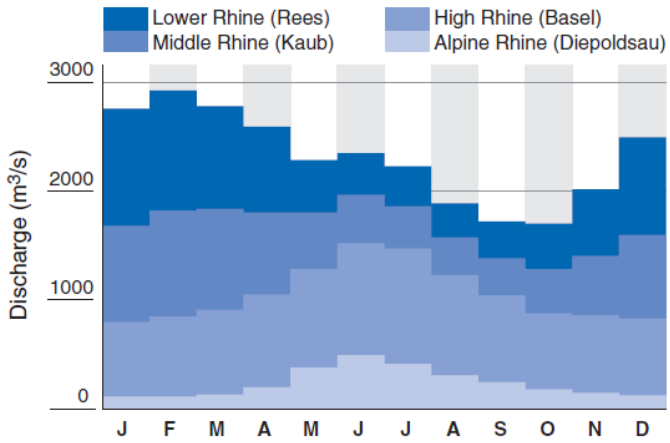


Figure 8 The average monthly discharge (1931-2003) of the Rhine River in four regions: Alpine, High, Middle, and Lower (Uehlinger et al., 2009)

A summary of the main characteristics per region is presented in Table 2, and described below:

Alpine Rhine

The first region from upstream to downstream is the Alpine Rhine. This region is the smallest, covering only 3.3% of the entire catchment with an area of 6155 km^2, and a river length of 90 km. The mean elevation of this subbasin is 1764 m+msl, ranging from 395 m+msl to 3614 m+msl (see Figure 6). The mean air temperature is 2.7 °C. About 1.4% of the Alpine Rhine catchment is covered by glaciers and snowmelt is the primary water source in this region. The mean annual discharge is 7.3km^3 (~231m^3/s) and the mean annual precipitation is 1900mm.

Lake Constance

Between the Alpine Rhine and the High Rhine, the river is interrupted by Lake Constance for approximately 60 km. This lake consists of 2 basins, the upper and lower which have a volume of 47.6 km^3 and 0.8 km^3 respectively and an area of 472 km^2 and 62 km^2. Moreover, it receives ~231 m^3/s and ~140 m^3/s from the Alpine Rhine and the other tributaries respectively, evaporation losses are on average 12 m^3/s (Baumgartner et al. 1983 as cited in Uehlinger et al., 2009) and ~5.5 m^3/s are withdrawn for domestic use.

High Rhine

The High Rhine covers 16.3% of the entire catchment with an area of 30,100 km^2, including the Aare catchment which is one of the main tributary rivers. The river in this region is 142 km long. The mean elevation of this subbasin is 902 m+msl and it ranges from 246 m+msl to 4274 m+msl. Its mean air temperature is 6.8 °C. The mean annual discharge and precipitation of this region are 33.4 km^3 (~1058 m^3/s) and 1349 mm respectively.

Upper Rhine

The Upper Rhine is the largest region, covering 34.0% of the entire catchment with an area of 62,967 km^2, including two of the main tributary rivers: Neckar and Main. The river in this region is around 333 km long. The mean elevation of this subbasin is 348 m+msl and it ranges from 88 m+msl to 1493 m+msl. Its mean air temperature is 8.6 °C. The mean annual discharge and precipitation of this region are 50.1 km^3 (~1587 m^3/s) and 735 mm respectively.

Middle Rhine

The Middle Rhine is the second largest region, which covers 22.6% of the entire catchment with an area of 41,810 km^2, including the Moselle, one of the main tributary rivers. The river in this region is around 159 km long. The mean elevation of this subbasin is 336 m+msl and it ranges from 43 m+msl to 880 m+msl. Its mean

air temperature is 9 °C. The mean annual discharge and precipitation of this region are 64.4 km^3 (~2040 m^3/s) and 811 mm respectively.

Lower Rhine

The Lower Rhine has an area of 18,836 km^2 covering the 10.2% of the entire catchment. The river in this region is around 177 km long. The mean elevation of this subbasin is 202 m+msl. Its mean air temperature is 9 °C. The mean annual discharge and precipitation of this region are 72.4 km^3 (~2040 m^3/s) and 797 mm respectively.

Delta Rhine

The last region is the Delta Rhine. This region covers 13.7% of the entire catchment with an area of 25,347 km^2. The mean elevation of this subbasin is 12 m+msl and the main air temperature is 9.2 °C. The discharge is diverted in three major branches in this region (Nederrijn-Lek, Waal, and Ijssel). Historically, there have been human interventions resulting in a closed dike system and decreased dynamics in the floodplain. This region is out of the scope of this research, as the model data that is central to this research is only available for the other regions.

Table 2 Main hydrological characteristics of the Rine River Regions (Uehlinger et al., 2009)

		Alpine Rhine	High Rhine	Upper Rhine	Middle Rhine	Lower Rhine	Delta Rhine
Mean catchment elevation	[m]	1764	902	348	336	202	12
Catchment area	[km2]	6155	30148	62967	41810	18836	25347
Mean annual discharge	[km3]	7.3	33.4	50.1	64.4	72.4	>72.4
Mean annual precipitation	[mm]	1926	1349	735	811	797	764
Mean air temperature	[°C]	2.7	6.8	8.6	9	9	9.2

3.1.3 Land Use

This section is based on Uehlinger et al., (2009).

Figure 9 and Table 3 show the land use of the Rhine Basin. Out of the total area (185,260 km^2) the major uses are Arable and Forest, 34.6% and 31.7% respectively, which are predominant in the Upper Rhine and Middle Rhine. Pasture covers 15.7% of the total area and it is predominant in the Middle Rhine, followed by High Rhine and Upper Rhine. The biggest urbanized region is the Lower Rhine, followed by the Delta Rhine, Upper Rhine, and Middle Rhine. Urban areas cover in total 8.8% of the whole catchment. Moreover, a big part of the Alpine Rhine and High Rhine is covered by Natural grassland and Spare vegetation, representing ~6.1% of the total area of the catchment. Finally, Wetland and Freshwater bodies occupied a minor part of the catchment (~2.8%), which are located mostly in the Delta Rhine and High Rhine.

Table 3 Land use of the Rine River Regions (Uehlinger et al., 2009)

		Alpine Rhine	High Rhine	Upper Rhine	Middle Rhine	Lower Rhine	Delta Rhine
Urban	[%]	1.9	4.4	8.9	6.8	18.3	11.5
Arable	[%]	2.5	20.7	43.7	36	38.4	31.5
Pasture	[%]	8.6	19.9	7.1	16.4	12.7	34.8
Forest	[%]	22.6	31.8	38.4	38.5	27.9	8.6
Natural grassland	[%]	37.4	11.5	1.2	1.3	0.5	1.8
Spare vegetation	[%]	26.3	6.7	0	0	0	0.2
Wetland	[%]	0.1	0.3	0	0	0	1
Freshwater bodies	[%]	0.5	4.7	0.5	0.6	1	10.22

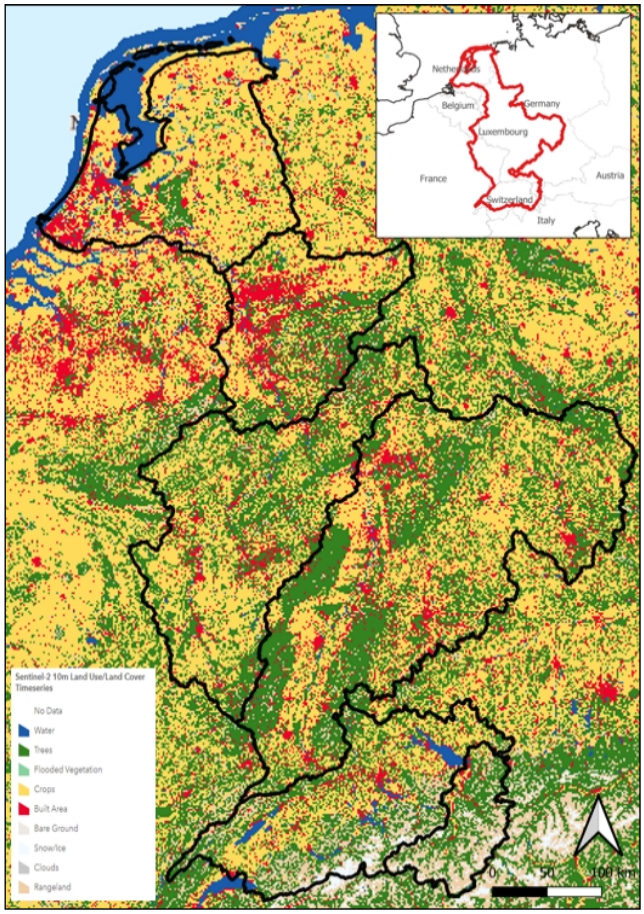


Figure 9 Land use of the Rhine River Basin. Data retrieved from Sentinel2 (ESRI)

3.2. Data | GRADE

In this study, we use 50,000 years of meteorological and hydrological data generated by a combined stochastic-physically based model GRADE (Generator of RAinfall and Discharge Extremes), which provides an alternative of a purely statistical method for deriving design discharges and associated flood hydrographs for the Rhine River in The Netherlands. The advantage when using long synthetic time series is that there is no need for extrapolation of data and upscaling hydrographs in flood risk assessments (Hegnauer et al., 2014) and that relevant flood processes during extreme events can be taken into account.

The GRADE method includes three components, namely 1) stochastic weather generator, 2) Hydrological model, and 3) Hydrologic and hydrodynamic routing. Component 1) provides precipitation and temperature series from a weather generator, the weather series are fed into 2) the hydrological model that provides discharge series that is routed through the river by a hydrodynamic model (Hegnauer et al., 2014). For this study we use the outputs of components 1 and 2, more details are provided in sections 3.2.1 and 3.2.2 respectively.

3.2.1 Meteorological Data

This section is based on Hegnauer et al., (2014).

The first component of GRADE, the stochastic weather generator, produces daily rainfall and temperature series (50,000 years) by resampling historical records with replacement, preserving the statistical properties of the original series. The nearest-neighbour resampling algorithm cannot generate daily rainfall amounts outside the range of the historical data, however, different multi-day temporal patterns are generated, allowing

different values for multi-day rainfall amounts than those observed in the historical data and providing a more accurate estimation of the statistical properties of extreme multi-day extreme events.

The weather generator generates rainfall and temperature at multiple locations simultaneously, preserving both the spatial correlation of daily rainfall over the drainage basin and the dependence of daily rainfall and temperature. Moreover, the autocorrelation of the time series is also incorporated by using the nearest neighbour method, which is restricted by a moving window to also account for the effect of seasonal variations in the dependencies between variables.

The weather generator for the Rhine basin generates daily precipitation and temperature simultaneously over 134 sub-basins of the Rhine upstream of Lobith (Dutch-German border), using two gridded daily precipitation data sets (the HYDRAS 2.0 and the E-OBS (Haylock et al., 2008; Rauthe et al., 2013; as cited in Hegnauer et al., 2014)) of observed daily rainfall and temperature data for the 56 years (1951-2006).

Limitations of the weather generator are important to mention. First, the limited length of the observations leads to considerable uncertainty in return values of multi-day rainfall extremes and simulated discharges. Second, the autocorrelation of the Rhine basin reasonably describes short-range temporal dependence, but there are indications that this is not the case for long-range temporal dependence (Buishand 2004 as cited in Hegnauer et al., 2014). Finally, both for precipitation and temperature the dependence between the spatial patterns on successive days is underestimated by the model (Beersma, 2011 as cited in Hegnauer et al., 2014). The implications of this bias are not clear. Nevertheless, the spatial dependence of the multi-day winter maximum precipitation amounts was adequately reproduced for the Rhine basin (Buishand and Brandsma, 2001, as cited in Hegnauer et al., 2014).

3.2.2 Hydrological Model

This section is based on Hegnauer et al., (2014).

The second component of GRADE, the HVB hydrological model, transforms the synthetic precipitation and temperatures series into discharges. The HVB (Hydrologiska Byråns Vattenbalansavdelning) model is a hydrological conceptual model developed in the early 1970s that has been applied to many river basins all around the world (Lindström et al., 1997 as cited in Hegnauer et al., 2014). The model mimics the behaviour of the runoff generation by using storage compartments and the flows between those compartments, accounting for the hydrological processes of the basin.

The model input consists of daily average precipitation and temperature for each sub-basin. Temperature is needed to account for temporary snow storage as well as evapotranspiration losses, which are calculated by the HBV model. An important assumption is that precipitation within the sub-basin is considered uniform. The HBV model for the Rhine runs with a daily time step.

Like the subdivision of the basins in the weather generator component, the Rhine basin is subdivided into sub-basins in the HVB setup. Four of the initial sub-basins were further subdivided to include four large lakes in Switzerland which have a considerable effect on the discharges, leading to 148 sub-basins in total for the hydrological model. Such a division allows achieving that single values of the model parameters can represent the physical characteristics of the sub-basin and having a more detailed fine-tuning of the calibration of the model for the total basin. Discharge data from the gauging station at the downstream end of the subbasins are used to carry out the calibration. Moreover, within each sub-basin, different zones are identified for which altitude and land use can be differentiated to account for the effect of altitude on the temperature.

A GLUE (Generalized Likelihood Uncertainty Estimation) analysis was used to calibrate the HBV model for each sub-basin, which combines the calibration of the model with an uncertainty estimation of the model parameters. GLUE allows to have multiple parameter sets that give approximately the same results, the performance of those parameter sets are evaluated to select the “behavioural sets” (sets that meet the constraints of the performance criteria). Whereas the GLUE analysis of the HBV model for the Rhine basin downstream of Basel showed good results for most of the HBV sub-basins, reproducing observed discharges

upstream of Basel turned out to be difficult. The overall performance of the HBV model for the whole area upstream of Basel is reasonable, but for some sub-basins the results of the GLUE analysis are poor and often no behavioural parameter sets could be found without loosening the criteria thresholds. The poor performance on sub-basin level is reflected in the relatively large uncertainty in the modelled results.

3.2.2 Configuration of the area of study according to the data

In the meteorological model, the area of study is divided into 134 catchments whereas in the hydrological model the area of study is divided into 148 catchments. To be able to compare the statistical dependencies between neighbouring catchments of both models, we identify 74 confluences that have two rivers that join in a confluence. (see Figure 10)

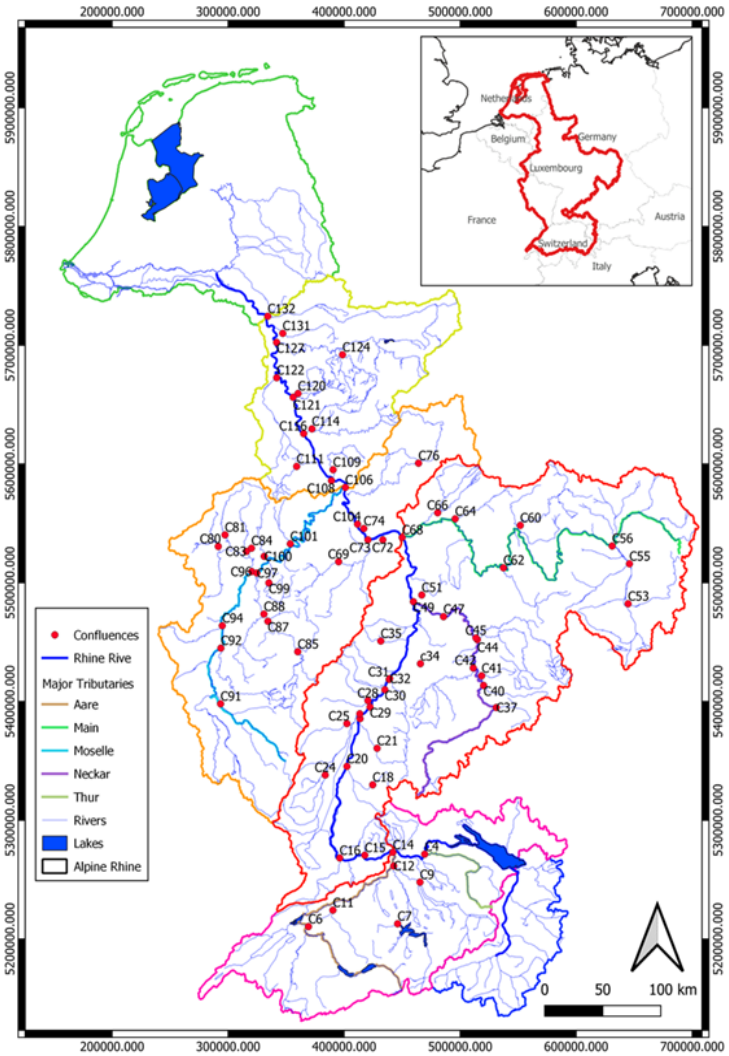


Figure 10 The 74 confluences that are evaluated in the Rhine River Basin. Data retrieved from Eionet, FAO-UN Land and Water Division, and Deltares.

4. Statistical dependencies of Meteorological and Hydrological parameters

In this chapter we answer research question 1 and its sub-questions, which are repeated below:

- *How do the different catchment characteristics and the meteorological and hydrological model domains influence the correlations between joining rivers?*

  The interactions between joining rivers and their spatial correlations vary according to not only the meteorological but also the hydrological processes that occur in the catchment at different temporal and spatial scales. Moreover, the catchment characteristics influence such interactions. The following questions will be addressed:

  1.1. Which are the meteorological processes involved in the study area and how are they related to the catchment characteristics?
  1.2. How are meteorological parameters spatially correlated and how do they influence the correlations between joining rivers?
  1.3. Which are the relevant hydrological processes involved in the study area and how are they related to the catchment characteristics?
  1.4. How are hydrological parameters spatially correlated and how do they influence the correlations between joining rivers?
  1.5. What are the differences between meteorological and hydrological model domain interactions according to the catchment characteristics?

Answers to sub-questions 1.1 and 1.3 are given in section 4.1 based on the background information of the Rhine Basin. In section 4.2 we explain the methodology and the experiment setup that is required to answer sub-questions 1.2 and 1.4. Then in section 4.3, we present the results for both, the meteorological and hydrological domains to be able to evaluate them and answer sub-question 1.5 in section 4.4. Finally, in section 4.5 we conclude this chapter by answering research question 1.

4.1. Meteorological and hydrological processes

We refer the reader to Section 3.1 to find the description of the meteorological and hydrological processes of the Rhine River basin and its flood patterns.

4.2. Methodology and Experiment setup

In Figure 11 we present the experiment setup that was performed in order to evaluate the statistical dependencies of the meteorological and hydrological extreme events between neighbouring catchments, which are also explained below.

Data Processing

The meteorological data consists of 50,000 years of simulated daily precipitation rate and temperature per catchment. For larger basins, consisting of two or more sub-catchments, the weighted mean precipitation over all sub-catchments was calculated according to the topology of the catchment. The weights were taken proportional to the size of the sub-catchment.

Three alternatives are considered for the meteorological model domain. First, by considering 1) daily and 2) multi-day events. For the latter, the precipitation was accumulated over 5 days to account for the travel time of the flood wave between Basel and Lobith (Diermanse, 2000 as cited in Hegnauer et al., 2014). In some instances, 3) precipitation falls as snow during the winter period. Therefore, a third alternative is analysed to

account for this process. This is done according to the temperature data: when the temperature is below 0 °C the precipitation is assumed to be 0. (see Appendix A)

Whereas for the meteorological domain three alternatives are analysed, for the hydrological domain only 1 alternative is considered. The hydrological data consists of the daily mean discharge per catchment. For this domain, no multi-day events are analysed, but the daily discharge is evaluated as the HBV model considers already the hydrological processes and routes the runoff, thus the travel time is already considered.

Extreme event selection

As indicated in Section 2.2.5, we differentiate between the main stream (*MS*), the tributary stream (*TS*), and the confluence (*C*) as shown in Figure 2. The sets of extremes considered for the analysis are non-simultaneous and named as follows:

- *Set 1 | MSmax-TSconc: mainstream maxima (MS_{max}) with a concurrent tributary stream (TS_{conc}),*
- *Set 2 | TSmax-MSconc: tributary stream maxima (TS_{max}) with a concurrent mainstream (MS_{conc}),*
- *Set 3 | Cmax (C=MS+TS): the maxima of the sum of both streams.*

From the 50,000 years of daily data, 50,000 annual maxima are identified for each set.

Bootstrapping

In order to calculate the variability of the results when using different sample sizes by random selection, the experiment is repeated 50 or 100 times (see Figure 11) for each sample size, being able to calculate the standard deviation of the results.

Parameter estimation

To evaluate the statistical dependencies of rainfall and discharge of neighbouring catchments during extreme events, the Spearman rank correlation (see section 2.2.1) is estimated for each of the extreme sets and samples of different sizes (5000, 1000, 500, 100, 50, and 20). When the sample size is other than the full data set of 50,000 years, the bootstrapping experiment is carried out and we calculated the mean and standard deviation of the Spearman rank correlation.

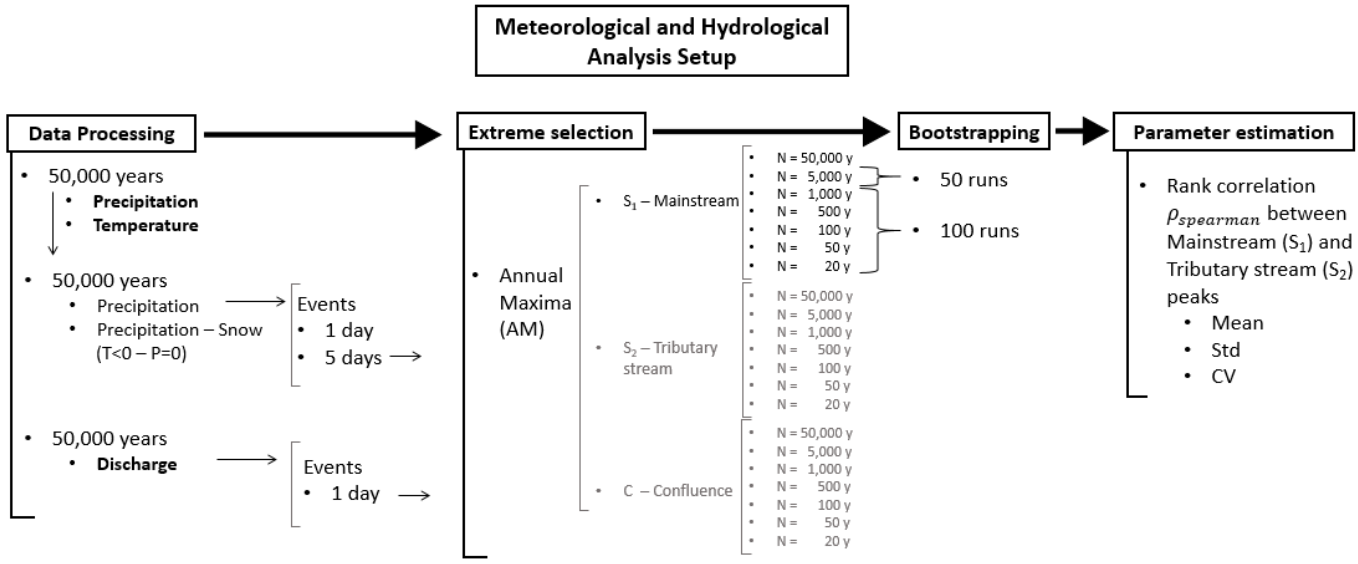


Figure 11 Process to evaluate the rank correlation between extreme events (meteorological: precipitation or hydrological: discharge) of neighboring catchments at each confluence.

4.3. Results and Evaluation

In this section, first, we explore the seasons where most of the peaks occur for both meteorological and hydrological events. Subsequently, we evaluate the statistical dependencies (Spearman rank correlation) between neighbouring catchments for each evaluated alternative and how they are spatially distributed in the catchment. The results represent the different alternatives that are evaluated according to the experiment setup explained in Section 4.2:

- MET | Precipitation 5 days
- MET | Precipitation-snow 5 days
- MET | Precipitation 1 day
- HYD | Discharge 1 day

Finally, we compare 1) the results of the four alternatives, and 2) in more detail the results of the Meteorological (precipitation 5 days) and the Hydrological (Discharge 1 day) alternatives.

Furthermore, in this section only the Annual Maxima Set 1 | MS_{max}-TS_{conc} is considered, using the full data (50,000 peaks). The sensitivity analysis of both extreme sets and the size of the samples are evaluated in sections 7.1.1 and 7.1.2.

Seasonality

The sets of extreme peaks are explored in terms of seasonality. The seasonality of the extreme events is relevant to identify the meteorological and hydrological processes that play a major role in the generation of the peaks. Figures 12, 13, 14, and 15 show the season (indicated by the colour) where most of the peaks occur per confluence (x-axis), moreover, the percentage of peaks is indicated on the y-axis.

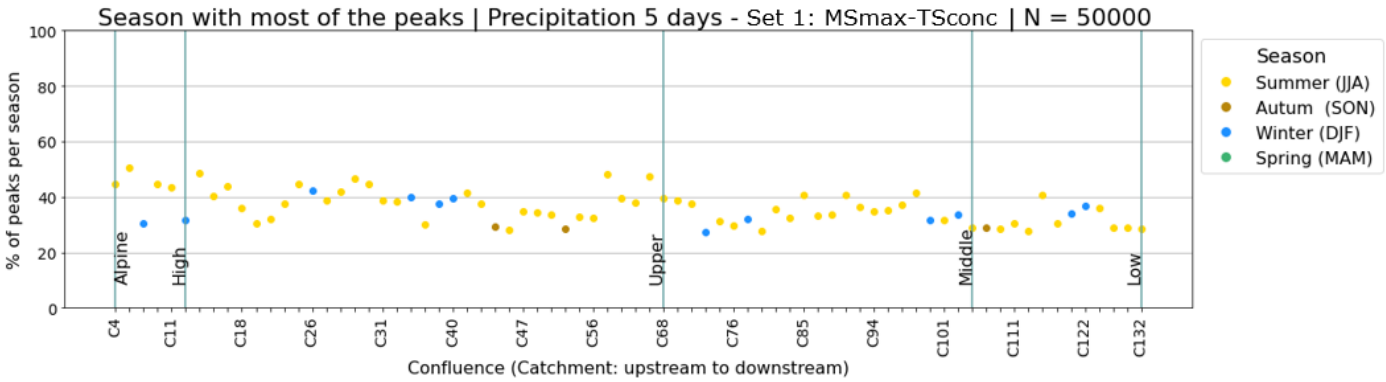


Figure 12 Season with most annual maxima peaks per confluence (x-axis) for the meteorological alternative Precipitation 5 days. The position on the y-axis indicates the percentage of peaks that occur during the season: Summer (yellow), Autumn (brown), Winter (blue), and Spring (green).

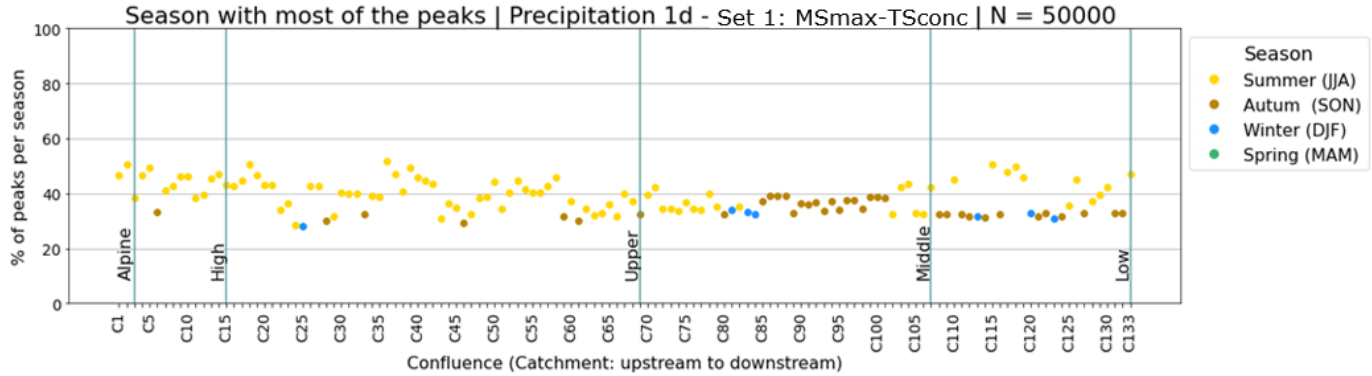


Figure 13 Season with most annual maxima peaks per confluence (x-axis) for the meteorological alternative Precipitation 1 day. The position on the y-axis indicates the percentage of peaks that occur during the season: Summer (yellow), Autumn (brown), Winter (blue), and Spring (green).

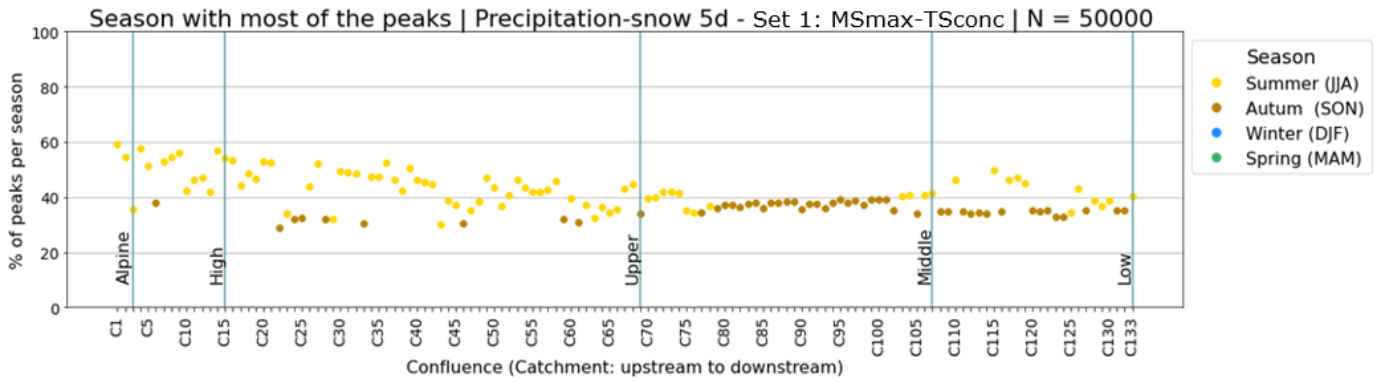


Figure 14 Season with most annual maxima peaks per confluence (x-axis) for the meteorological alternative Precipitation-snow 5 days. The position on the y-axis indicates the percentage of peaks that occur during the season: Summer (yellow), Autumn (brown), Winter (blue), and Spring (green).

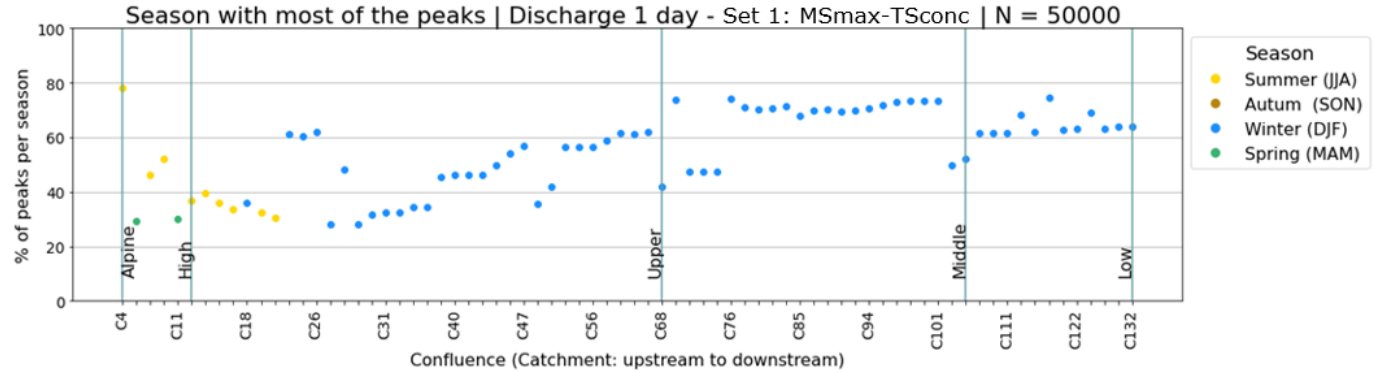


Figure 15 Season with most annual maxima peaks per confluence (x-axis) for the hydrological alternative Discharge 1 day. The position on the y-axis indicates the percentage of peaks that occur during the season: Summer (yellow), Autumn (brown), Winter (blue), and Spring (green).

For the three alternatives of meteorological events we can observe that in the Alpine, High and Upper regions, most of the peaks (~40%) occur during summer. In the Middle and Low regions, when considering precipitation of 5 days, most of the peaks occur during summer (Figure 12) as well, whereas for Precipitation 1 day (Figure 13) and Precipitation-snow 5days (Figure 14) the peaks occur mostly during autumn. Having most of the peaks during summer is consistent with the precipitation patterns, ~30% of the annual precipitation falls during that season as described in Section 3.1.2.

In Figure 15, we can observe that most of the discharge peaks occur during winter, except in the Alpine and High regions where they occur mostly during summer. This seasonality resembles the shift in the annual cycle of the discharge described in Section 3.1.2, where higher discharges are observed during summer for the Alpine and High regions where snowmelt takes place, and during winter for the Upper, Middle and Lower regions where less predictable rain-dominated runoff takes place (see Figure 8).

Statistical dependencies between neighbouring catchments

In this subsection, we present the Spearman rank correlation of extreme peaks between neighbouring catchments that result from the experiment described in Section 4.2. In Figure 16, Figure 17, Figure 18, and Figure 19 we can observe how the correlations are spatially distributed in the catchment for each of the analysed alternatives.

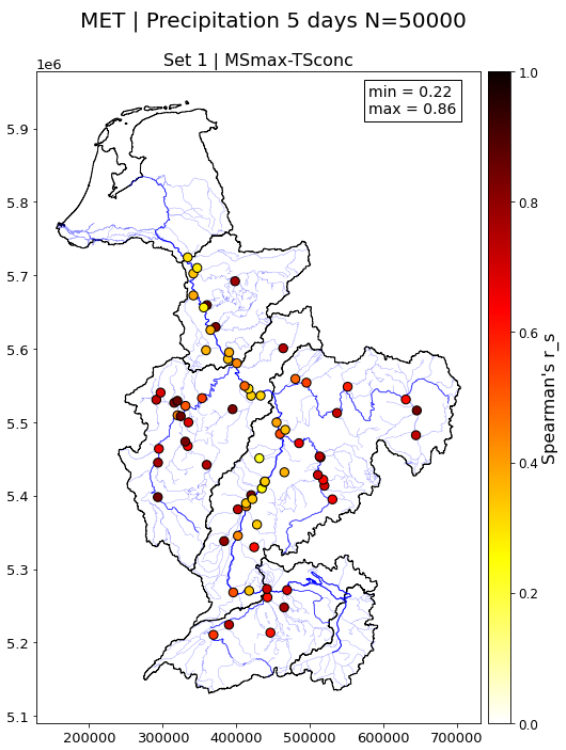


Figure 16 Spearman's rank correlation between extreme events (meteorological: precipitation 5 days) of neighbouring catchments at each confluence

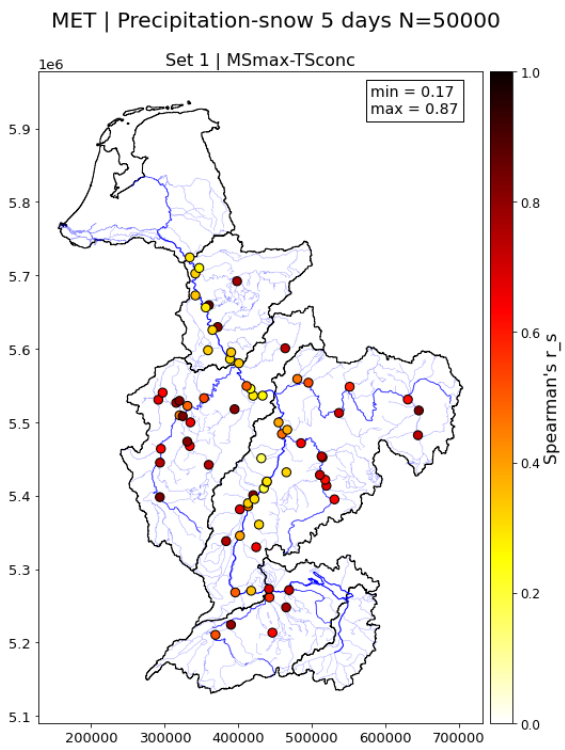


Figure 17 Spearman's rank correlation between extreme events (meteorological: precipitation-snow 5 days) of neighbouring catchments at each confluence

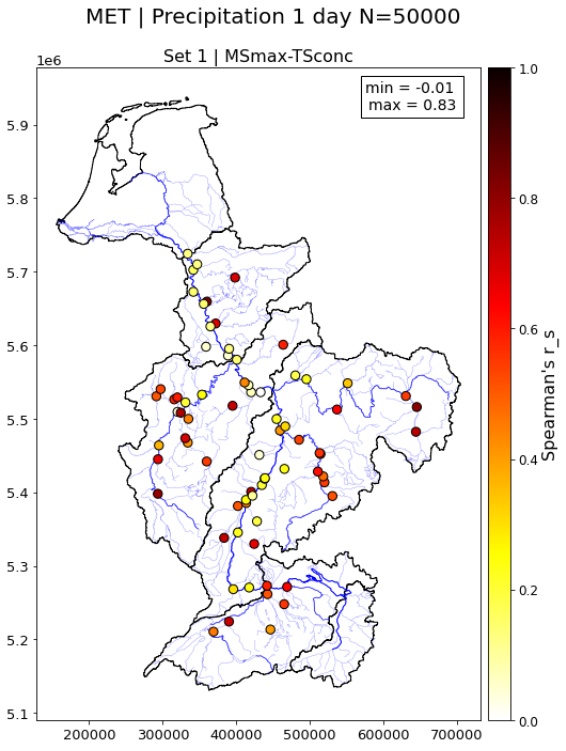


Figure 18 Spearman's rank correlation between extreme events (meteorological: precipitation 1 day) of neighbouring catchments at each confluence

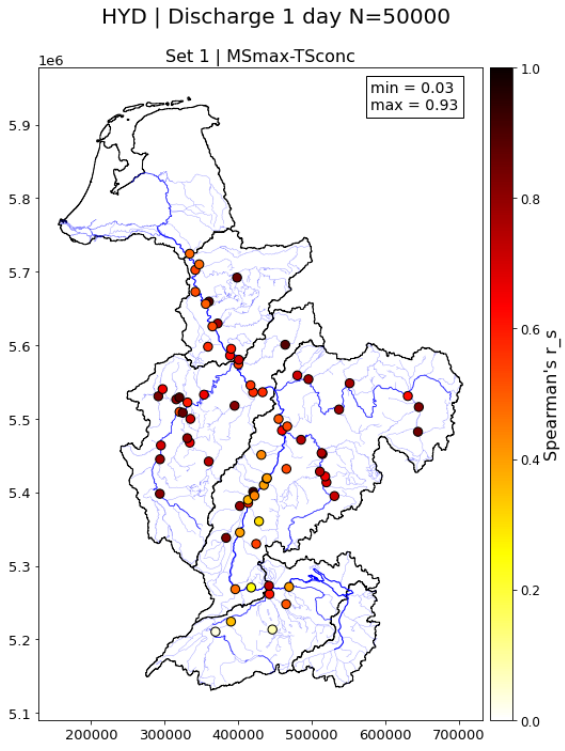


Figure 19 Spearman's rank correlation between extreme events (hydrological: discharge 1 day) of neighbouring catchments at each confluence

First, for the three meteorological alternatives (Figure 16, Figure 17, and Figure 18), in general, high correlations are observed in the upstream confluences of each region, whereas for the confluences in which

at least one large river is involved the correlations are lower. The catchments of the upstream confluences are small; hence, the precipitation events are more likely to be the same event for both of the neighbouring catchments. However, when moving downstream the catchment area of the main river is larger and therefore the precipitation events are less likely to cover the whole area in time and space. Therefore, the low correlations in larger rivers show that the extreme events are driven by a combination of different events from upstream areas.

Second, for the hydrological event (Figure 16) high correlations can be observed all around the catchment. Lower correlations around the larger rivers are also observed, but this pattern is less distinct than it is for the meteorological figures. Similar to the meteorological events, the low correlations at the larger rivers show the combination of extreme (or not extreme) events from different upstream areas.

Differences between all event alternatives

To compare the meteorological and the hydrological event alternatives, in **Error! Reference source not found.**, **Error! Reference source not found.**, Figure 22, and Figure 23 we show the correlation differences, using as a benchmark the MET | Precipitation 5 days event.

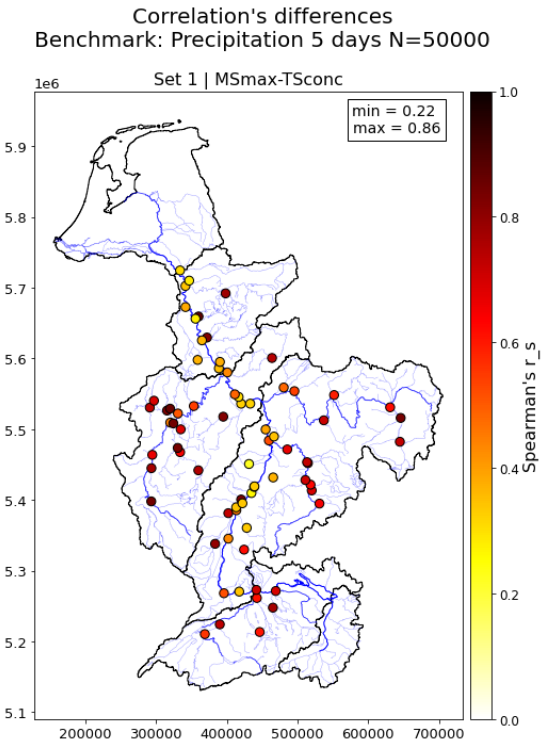


Figure 20 Spearman's rank correlation between extreme events (meteorological: precipitation 5 days) of neighbouring catchments at each confluence

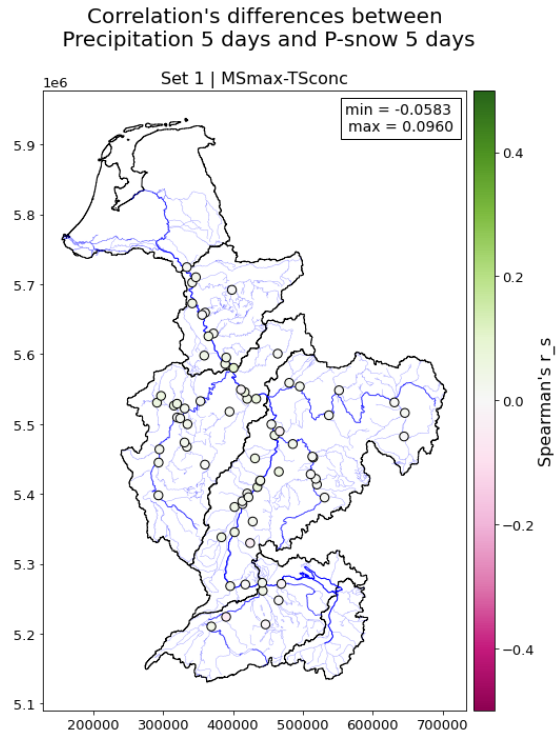


Figure 21 Differences of the Spearman's rank correlation between extreme events (precipitation-snow 5 days) of neighbouring catchments at each confluence, with respect to the precipitation 5 days correlation

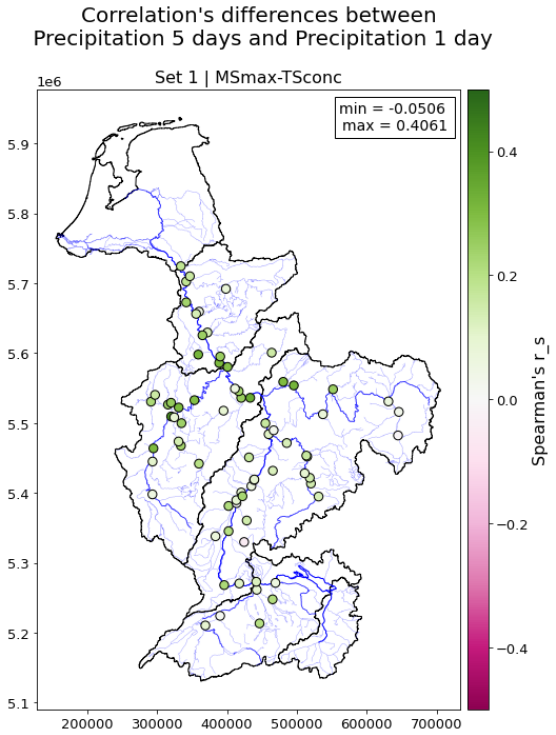


Figure 22 Differences of the Spearman's rank correlation between extreme events (precipitation 1 day) of neighbouring catchments at each confluence, with respect to the precipitation 5 days correlation

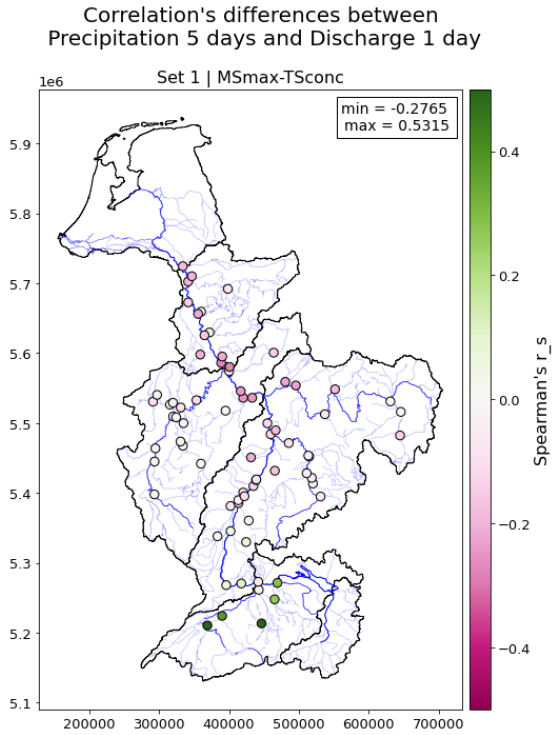


Figure 23 Differences of the Spearman's rank correlation between extreme events (discharge 1 day) of neighbouring catchments at each confluence, with respect to the precipitation 5 days correlation

Differences between the three meteorological alternatives can be observed. First, in **Error! Reference source not found.** we can observe that the sets of Precipitation 5 days and Precipitation-snow 5 days led to similar magnitudes of correlation in all the confluences. As mentioned in Section 4.2, the second alternative assumes that the precipitation is 0 when the temperature is below 0°C, and snow melt is not considered in this alternative. Additionally, as mentioned before, most of the peaks for both alternatives take place during the summer. Certainly, temperatures are below 0°C during winter, therefore the precipitation that falls as snow does not play a major role in the meteorological domain as considered in this study.

In contrast, in Figure 22 we can observe differences between the sets of Precipitation 5 days and Precipitation 1 day, especially in the Upper, Middle and Lower regions. In general, the multi-day events (5 days) are higher correlated than the daily events. It is noticeable that for the confluences located in upstream areas the difference is lower than for those that are in downstream areas. This indicates that for those upstream (and therefore small) catchments, the temporal scale does not play a major role in the statistical dependencies as daily and multi-day (5 days) meteorological events show similar correlations. Whereas for downstream (large) catchments, the temporal scale of meteorological events plays a major role. Downstream confluences are less likely to have a meteorological event in 1 day that covers both neighbouring catchments leading to low correlations for daily events. Five-day events may come from the same weather system, therefore, causing higher correlations between neighbouring catchments.

We also observe some similarities and differences between the meteorological (Precipitation 5 days) and the hydrological (Discharge 1 day) events in Figure 23. Similar correlations are observed in the upstream confluences of the Upper, Middle and Low regions, meaning that for those confluences the extreme events are concurrent with extreme precipitation. Different correlations are observed in the Alpine and High regions, and in confluences that are close to the main river in the other regions. When looking at the confluences involving larger rivers, we can observe higher correlations for the hydrological model domain. This means that the extreme events are not necessarily fully concurrent with the meteorological domain, and other hydrological processes of the catchment take place. When looking at the Alpine and High regions we can

observe lower correlations for the hydrological model domain, which can be related to the snow processes that take place in those regions of the catchment and/or to the lakes that are located in the area.

Differences between Meteorological (Precipitation 5 days) and Hydrological (Discharge 1 day)

In the following figures, we only use one of the meteorological alternatives (Precipitation 5 days) to compare it in detail with the hydrological alternative (Discharge 1 day). Figure 24 depicts the correlations per confluence on the Y-axis, and the confluences on the X-axis, which are ordered according to the size of the total upstream area. Moreover, Figure 25 shows the spatial distribution of the ratio (size ratio) of the two contributing areas per confluence (Area of the mainstream divided by the area of the tributary stream). Finally, Figure 26 shows the relationship between the confluence correlation (y-axis), its total upstream area (x-axis), and the size ratio (colour) for both, meteorological (a) and hydrological (b) domains.

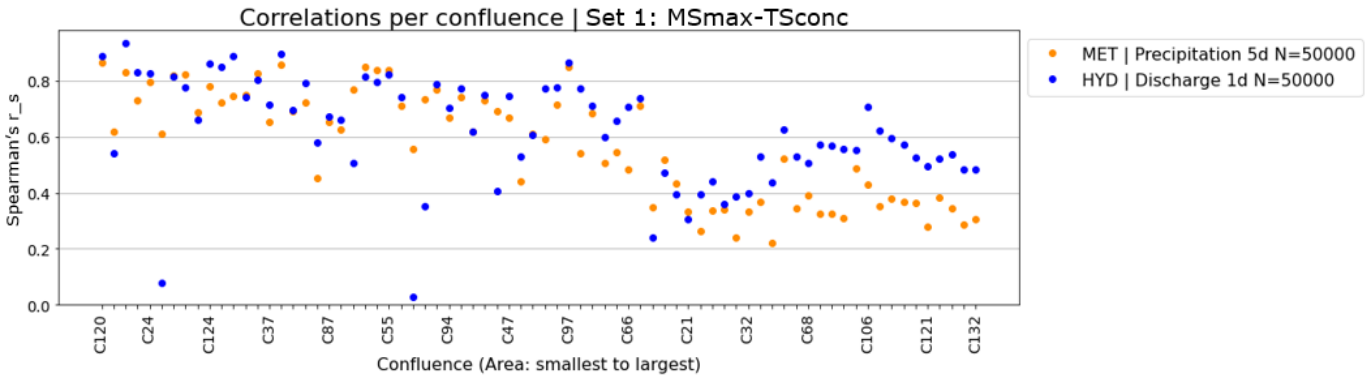


Figure 24 Spearman's rank correlation between extreme events of neighbouring catchments at each confluence. The extreme events represented are precipitation 5 days (orange) and discharge 1 day (blue). The confluences are ordered in the x-axis according to the total contribution area of the two neighbouring catchments, from smallest to largest.

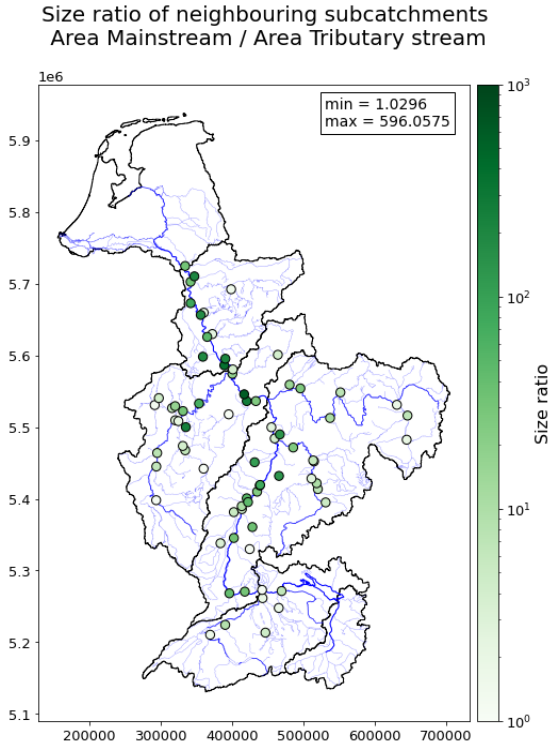


Figure 25 Size ratio of the contributing catchments of each confluence. (Mainstream area (A_{MS}) divided by the tributary stream area (A_{TS}))

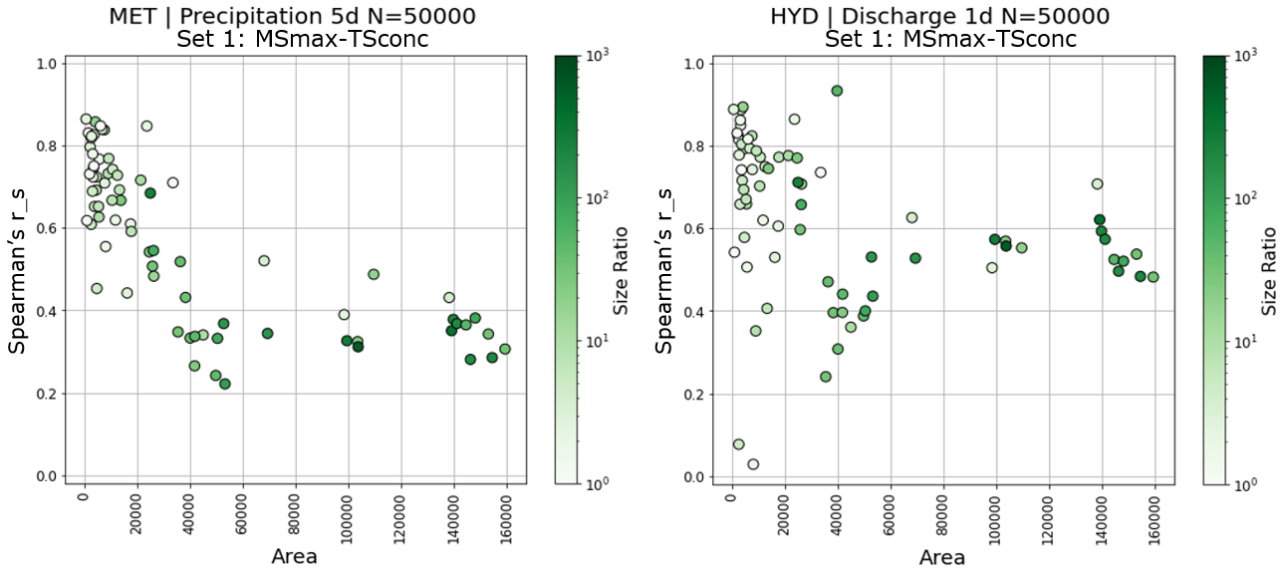


Figure 26 Spearman's rank correlation between extreme events of neighbouring catchments at each confluence according to the total upstream area (x-axis) and the size ratio of the contributing catchments (colour rank). Left: Extreme precipitation 5 days. Right: Extreme discharge 1-day

In Figure 24 we can observe that, in general, the meteorological and hydrological correlations are lower when the size of the catchment is larger. Moreover, a pattern related to the size ratio can also be observed in Figure 26-left, the closer the ratio to 1, the higher the correlation of meteorological events. In other words: the bigger the difference in area between the neighbouring catchments, the less correlated the meteorological events. Whereas for meteorological events the patterns are clear, for the hydrological events they are less distinct. In Figure 26-right, we can observe 2 confluences that have size ratios close to 1 (white) but that are low correlated. Those confluences are in the Alpine and High regions, therefore, the low correlation can be related to the snow processes that take place and/or to the lakes that are located in the area.

4.5. Conclusion and answer to the Research Question 1

In this section, we aim to answer Research Question 1: *How do the different catchment characteristics and the meteorological and hydrological model domains influence the correlations between joining rivers?*

For this objective, the catchment characteristics were explored to determine their influence on the statistical dependencies between neighbouring catchments of extreme events in both meteorological and hydrological domains. 74 confluences were identified and evaluated.

Different alternatives of events were evaluated. On the one hand, for the meteorological domain, three alternatives are evaluated, 1) daily and 2) multiday events, and 3) accounting for the precipitation that falls as snow. On the other hand, only one alternative for the hydrological domain is evaluated: annual maximum of daily discharges. For each confluence, the annual maximum peaks were identified out of the 50,000 years of meteorological (precipitation) and hydrological (discharge) synthetic data from GRADE (Hegnauer et al., 2014).-Subsequently, the sets were explored in terms of seasonality, identifying the season where most of the peaks occur per confluence. Moreover, the Spearman rank correlation was calculated for the sets of extremes, between the mainstream peaks and the tributary stream peaks for each confluence, which is used to evaluate the statistical dependencies between neighbouring catchments.

Based on the catchment characteristics exploration and the seasonality of the peaks, we can conclude that even though precipitation occurs at any time of the year, a seasonal variation is observed, having ~30% of the annual maximum precipitation during summer in the Alpine, High and Upper regions. Hence, most of the extreme precipitation events occur during that season. Moreover, there is a shift in the annual cycle of the

discharge, where higher discharges are observed during summer for the Alpine and High regions, and during winter for the Upper, Middle and Lower regions. This shift resembles the seasonality observed in the extreme discharge events, where snowmelt processes take place during spring/summer in the Alpine and High regions, and less predictable rain-dominated runoff takes place in the Upper, Middle and Lower regions.

Based on the spatial distribution of the calculated Spearman rank correlation for all the alternatives, we observed high correlations in the upstream confluences in all the Rhine regions. The catchments of the upstream confluences are small; hence, the precipitation events are more likely to be the same event for both neighbouring catchments. Moreover, similar correlations for the meteorological and hydrological evaluation are observed in the upstream confluences of the Upper, Middle and Low regions, meaning that for those confluences the extreme hydrological events are concurrent with extreme precipitation (meteorological domain).

When moving downstream and reaching the main river the correlations are lower. Furthermore, we observed higher correlations for the hydrological domain than the meteorological domain in those confluences. The catchment area of the main river is larger and therefore the precipitation events are less likely to cover the whole area in time and space, and hydrological processes play a major role in peak flow generation. The low correlations closer to the main river show that the extreme events are driven by a combination of different events from upstream areas.

The results for the Alpine and High regions showed lower correlations for the hydrological model domain, which can be related to the snow processes that take place in those regions of the catchment and/or to the lakes that are located in the area.

In addition, no differences were observed between the analysis that includes, or leaves out the precipitation that falls as snow Even though the snow process plays a major role in the catchment, the considerations taken in this analysis did not show its importance, which can be attributed to the fact that most of the extreme events occur during summer as shown in the seasonality evaluation and we are not accounting for the snow melt in this alternative. Differences in correlations between the daily and multiday meteorological events are more substantial. In upstream (and therefore small) catchments the temporal scale does not play a major role in the estimation of the statistical dependencies as daily and multiday (5 days) meteorological events show similar correlations. But for downstream (large) catchments, the temporal scale of meteorological events becomes relevant regarding correlations between rainfall in catchments of joining rivers. Downstream confluences are less likely to have a meteorological event in 1 day that covers both neighbouring catchments, leading to low correlations for daily events.

The correlation of meteorological and hydrological events is lower for larger catchments. Along with the total area, the size ratio of the mainstream catchment and the tributary stream catchment also showed a pattern: the bigger the difference in area between the neighbouring catchments, the less correlated the extreme events. This pattern is more distinct for the meteorological domain (catchment rainfall) than the hydrological domain (river discharges).

Throughout this first evaluation, we have identified the confluences where the extreme events of neighboring sub-catchments are strongly/weakly correlated, which allows us to give a first indication of where we shall perform a deeper analysis. Hence, for the confluences where the obtained correlations are weak, it is not likely to have extreme events at both of the neighboring sub-catchments, so we suggest having a closer evaluation as it may be translated into cost reduction when implementing flood measures.

5. Hydraulic evaluation

In this chapter, we aim to answer research question 2 and its sub-questions, which are repeated below.

- *What are the hydraulic interactions between joining rivers?*
  Due to the time it takes to perform hydraulic simulations, it is necessary to identify the potential locations where hydraulics play a major role. The results of the first experiment (first sub-question) could be a first indication for the selection. Hence, locations, where the hydrological model domain reveals low statistical dependencies in the results, should be analysed in more detail. Moreover, the topography of the area also influences the selection of the locations: in the steepest areas, we would not expect relevant hydraulic interactions between adjoining rivers because of limited backwater effects, as opposed to flat areas.
  Moreover, the hydrological processes in the catchments of the joining rivers also play a role in the hydraulic interactions, therefore they need to be evaluated and considered in the simulations. The following questions will be addressed:

  2.1. What characteristics should we consider when identifying confluences where interactions might be relevant?
  2.2. How can the hydrological processes in different joining tributaries be considered in setting up probabilistic hydraulic modelling simulations?
  2.3. Which hydraulic events need to be simulated for carrying out a flood risk analysis?

In Section 5.1 we explain the methodology and introduce the hydraulic model. Additionally, we define the response function that is used to evaluate the flood risk, partially answering sub-question 2.3. Next, in section 5.2, first, sub-question 2.1 is answered, then the location to perform the hydraulic simulation is selected, second, we evaluate the hydrological characteristics of the selected location to answer sub-question 2.2 and, finally, the hydraulic model setup is explained. Subsequently, the results and evaluation of the hydraulic simulations are presented in Section 5.3, giving an answer to sub-question 2.3. Finally, in section 5.5 we conclude this chapter by answering research question 2.

5.1. Methodology and Experiment setup

In this section, we present the experiment setup (Figure 27) that we use to answer Research Question 2 and its sub-questions. Subsequently, we introduce SFINCS, the hydraulic model that is used. Moreover, we define the response function that is obtained from the hydraulic simulation and used to evaluate the risk of flooding.

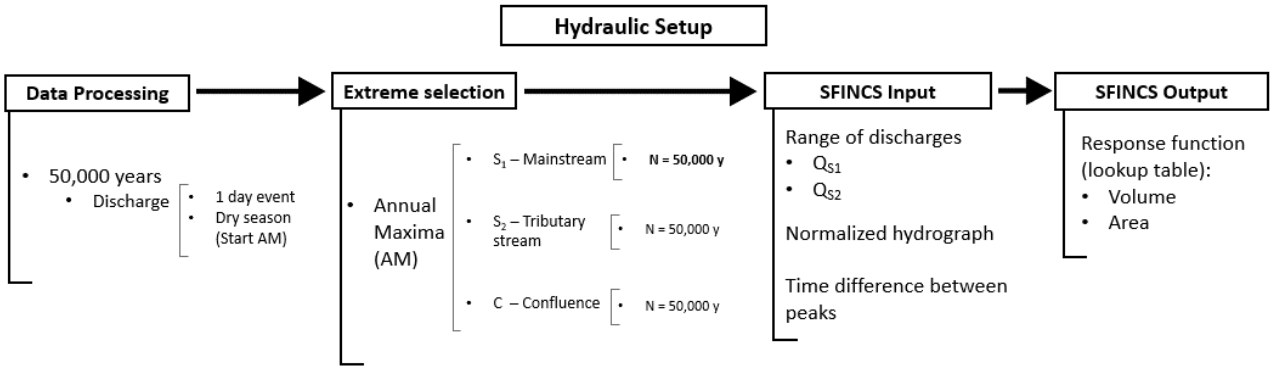


Figure 27 Process to obtain the input, and perform the hydraulic simulations, to obtain the response function

Data Processing

The hydrological data consists of 50,000 years of simulated daily mean discharge per catchment.

Similar to the evaluation of the meteorological and hydrological parameters in Section 4, we did not evaluate multiday events as the HBV model considers already the hydrological processes and routes the runoff, thus the travel time is already considered.

Extreme event selection

As indicated in Section 2.2.5, we differentiate between the mainstream (*MS*), the tributary stream (*TS*), and the confluence (*C*) as shown in Figure 2. The sets of extremes considered for the analysis are non-simultaneous and named as follows:

- *Set 1 | MSmax-TSconc: mainstream maxima (MS_{max}) with a concurrent tributary stream (TS_{conc}),*
- *Set 2 | TSmax-MSconc: tributary stream maxima (TS_{max}) with a concurrent mainstream (MS_{conc}),*
- *Set 3 | Cmax (C=MS+TS): the maxima of the sum of both streams.*

From the 50,000 years of daily data, 50,000 annual maxima peaks are identified for each set.

In this analysis, attention is paid to the start of the block maxima (annual maxima). Therefore, the dry season (the season with the fewest extreme peaks) needs to be identified. The end of the dry season is used as the start of each year to select an annual maximum. This avoids having twice the same event for consecutive years as mentioned in Section 2.1.2.

SFINCS

SFINCS (Super-Fast Inundation of CoastS) is a reduced-complexity model developed at Deltares. It is capable of simulating compound flooding, at the same time it balances high computational efficiency with adequate accuracy. SFINCS solves a set of momentum and continuity equations with a first-order explicit scheme based on Bates et al. (2010, as cited in Leijnse, 2018).

Moreover, SFINCS includes fluvial, pluvial, tidal, wind- and wave-driven processes, which can be modelled in different environments such as coastal, coral reef, tsunami, storm surge, riverine, urban, and flash flood. Finally, in the compound environment, all relevant types of forces can be combined into 1 domain, being able to simulate compound flooding events (Leijnse, 2018).

In this study, the Riverine model is used to simulate a fluvial flood due to two rivers joining at a confluence. For inland riverine types of environments, generally, at the upstream end of rivers, discharge points with discharge time series are provided. Moreover, at the downstream end of rivers, water level time-series need to be specified, which in case of sub-critical flow conditions will influence the flow upstream (Leijnse, 2018).

The input of the model

Due to limitations in time, we cannot simulate all the annual maxima peaks we have identified (50,000 annual maxima peaks of each set). Therefore, we need to evaluate them, identifying the minimum and maximum peak magnitudes for each of the rivers. Moreover, different combinations of the river discharges will lead to different flood responses, hence, a grid of 5 discharges for each river, ranging from the minimum to the maximum peak magnitude observed in the data, leads to 25 possible combinations to perform the hydraulic simulations.

As mentioned in the previous subsection, discharge time series are required as input for the hydraulic model (SFINCS) for each of the rivers that join. Thus, a hydrograph is needed to provide the time series of each peak. To account for the differences in the response of the flood behaviour in terms of time according to the peak magnitude, we classify the discharge's magnitude into 5 ranges of discharge. For each range, an 'average' normalized hydrograph shape is derived. We briefly describe the procedure below.

We normalized all the peaks for each class by first identifying the time series of each peak, considering a window of 15 days previous to and after the peak and second dividing the discharges by the peak magnitude. As a result, we had N time series of 30 days with a peak magnitude equal to 1 for each class of discharges. To derive a representative normalized hydrograph per class, we calculated the vertical mean of all the peak normalized hydrographs. (See Appendix C). The resulting hydrograph was applied to derive the input time series of the SFINCS simulations.

When simulating high discharges of both rivers, it becomes relevant to consider the time difference between the peaks, as it is likely to be the annual maxima for both rivers. Hence, the distribution of the day of the year when the annual maxima peak occurs was identified for both, the mainstream and the tributary stream. Then the difference in time of the peak occurrence is calculated for each year, and a histogram is constructed to evaluate them. So, the time difference between the 2 river peaks is also included in the input time series of the model.

The output of the model: the response function

As we wish to assess how tributary compound flood events may impact flood risk in surrounding areas, the hydraulic simulation results need to be converted into a measure of the consequences of a flood event. The consequences can be evaluated in different terms such as material, ecological damages, injuries, and fatalities, or also be expressed in monetary damage (Jonkman et al., 2021).

To evaluate the damage a flood can cause, a damage function can be constructed as a function of a certain characteristic of the flood (e.g. depth, area, volume, duration, time of occurrence, the velocity of the flooding water). Therefore, the damage functions relate the level of damage to the flood conditions (Jonkman et al., 2021).

As economic analyses to construct a damage function are out of the scope of this study, two response functions are used instead. The responses are calculated in terms of the flooded area and the flooded volume as a surrogate for socio-economic impacts, which with further economic data and evaluation may be translated into damage.

SFINCS allows us to calculate the maximum water depth (h_{max}) per grid cell of the simulation area. The total flooded area is estimated by counting all the cells with $h_{max} > 0$ and multiplying them by the size of the cells. Moreover, the flooded volume is calculated by multiplying each h_{max} by the cell's size and accumulating them.

From the 5x5 grid of possible combinations of discharges (25 simulations) explained before, a grid of responses can be calculated, having a response function in terms of the river discharges. The response function can be used in the next step when evaluating the hydrological samples of Research Question 3 (Section 6).

5.2. Hydraulic model

In this section, the selected location to perform the hydraulic evaluation is introduced. Additionally, a hydrological evaluation to define the boundary conditions for the hydraulic model is performed. Finally, the hydraulic model setup is explained.

5.2.1 Selection of the location

In order to select a location to evaluate the hydraulic interactions, 7 confluence locations were preselected. The preselection was done based on the region of the catchment, the main tributaries, the size of the total contributing catchment and the ratio of the two sub-catchment sizes, the differences/similarities of the 3 sets of extremes, and the statistical dependencies between the neighbouring sub-catchments obtained in Section 4. (see Appendix B)

We performed the hydraulic simulation in only one location as a demonstration case: Confluence 68, located in the outlet of the Upper Rhine region. At the confluence, the Main River, one of the main tributaries, joins the Rhine River. In the right part of Figure 28, we can see that the flood plain is relatively flat, which means we can expect hydraulic interactions to be substantial.

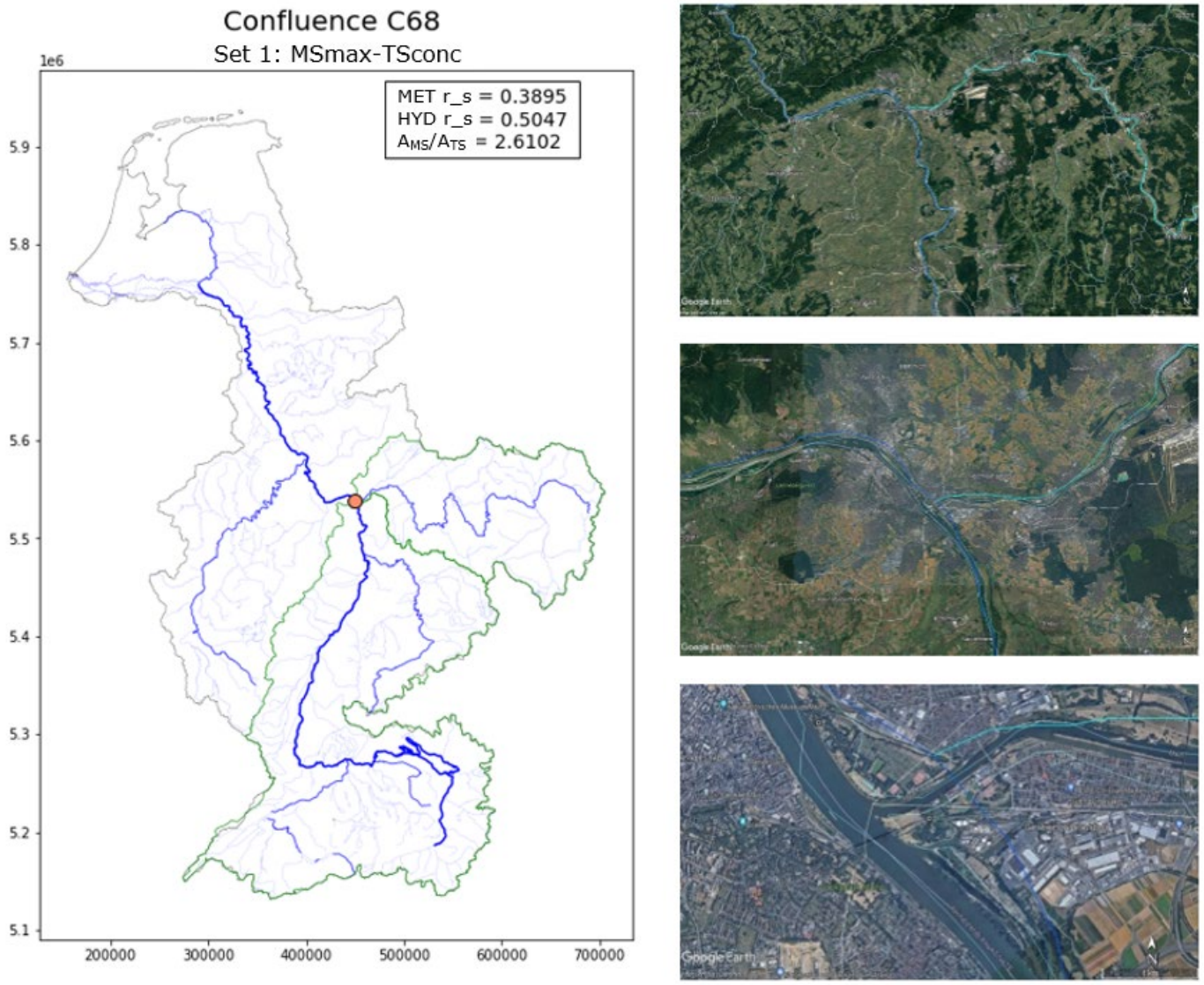


Figure 28 Selected location to perform the hydraulic evaluation: Confluence 68, where the Main River joins the Rhine River. Left: location of the confluence in the catchment. Right: Satellite views retrieved from Google Earth, zoom to the location.

The total contributing catchment is about 99000 km^2, from which 71300 km^2 corresponds to the Rhine sub-catchment and 27300 km^2 to the Main catchment, having a size ratio of 2.6. Due to the large size of the total upstream catchment, we can analyse the hydraulic interactions on a daily time scale.

In Figure 29, we can see that the three selection methods resulted in different meteorological events, with a Spearman rank correlation of 0.39 for Set1. Whereas the hydrological events led to more similar events and a higher correlation, 0.54 (see Table 4).

Table 4 Spearman's (r_s) and Kendall (τ) rank correlations between extreme events (discharge 1 day) of neighbouring catchments at Confluence 68

	r_s	τ
Set 1 | MS_{max}-TS_{conc}	0.54	0.37
Set 2 | TS_{max}-MS_{conc}	0.69	0.50
Set 3 | C_{max} (C=MS+TS)	0.45	0.31

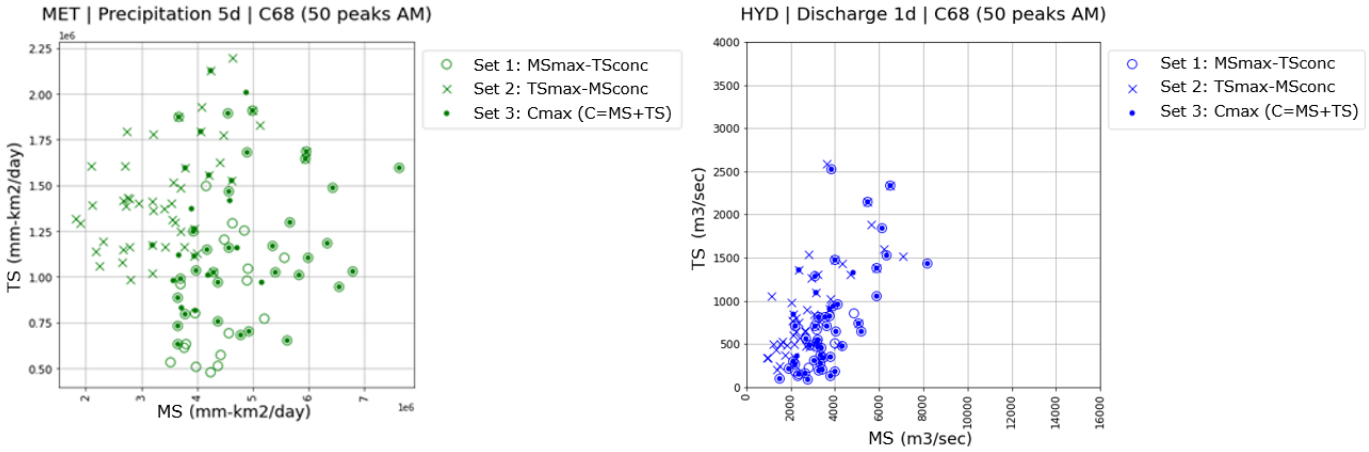


Figure 29 Pairs of extreme events at Confluence 68: Mainstream (*MS*) on the x-axis and Tributary stream (*TS*) on the y-axis. Left: Meteorological events (Precipitation 5 days). Right: Hydrological events (Discharge 1 day)

5.2.2 Hydrological evaluation of the selected location (hydraulic model input)

As explained in Section 5.1, we need to do a hydrological evaluation to determine the hydraulic input and boundary conditions. First, the dry season (the season with the fewest extreme peaks) was identified to define the start of the block maxima. Then, the 5x5 grid of discharges are defined, ranging from the minimum to the maximum peak magnitude observed in the data. Moreover, we classified the discharges into 5 magnitude ranges and calculated the normalized hydrographs that represent each of the classes. Finally, the time difference between the 2 river peaks is evaluated.

Start of the Block Maxima

From the extreme sets identified in the hydrological evaluation of Section 4, we identified the month where each of the peaks occur. As observed in Figure 30, September is the month with the fewest extreme peaks for the three of the extreme sets. Therefore, for the hydraulic evaluation, we started the hydrological year of the block maxima in September.

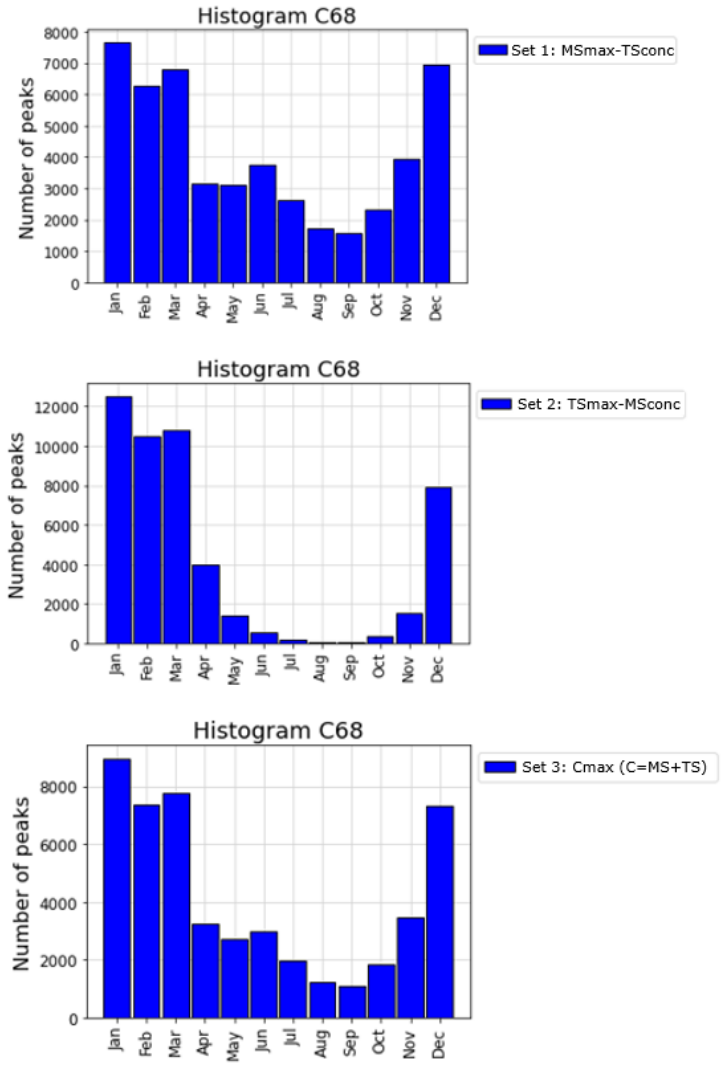


Figure 30 Peak discharge (1 day) histogram for each extreme set (Set 1: upper, Set 2: middle, Set 3: bottom)

Grid of discharges

From all the possible combinations of extreme discharges obtained from the 3 sets of extremes, we selected a grid of 5x5 discharges for each river that ranges from the minimum to the maximum peak magnitude observed in the data, having 25 possible combinations to perform the hydraulic simulations (see Figure 31 and Table 5).

Table 5 Magnitude of discharges for the hydraulic simulations

		Q0 minimum	Q1	Q2	Q3	Q4 maximum
MS-Rhine	m^3/s	724	3404	6083	8763	11443
TS-Main	m^3/s	27	1120	2212	3305	4397

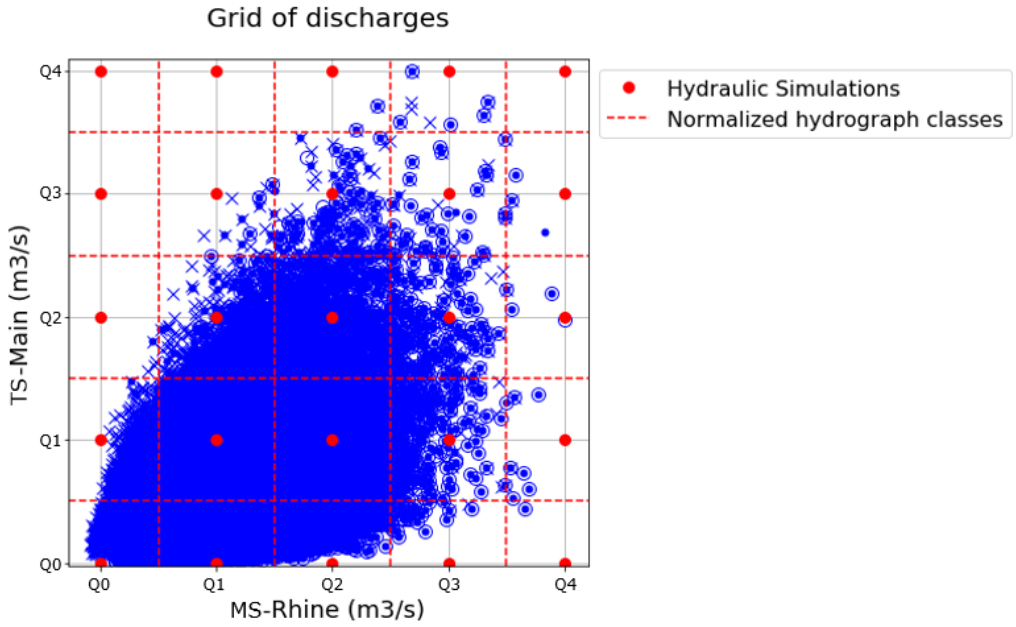


Figure 31 Pairs of extreme events (blue) at Confluence 68: Mainstream (*MS-Rhine*) on the x-axis and Tributary stream (*TS-Main*) on the y-axis, and the grid of discharges (red dot) and the normalized hydrograph classes (red dashed line) for the hydraulic simulations.

Classification of discharges and Normalized hydrographs

A hydrograph is needed to provide the time series of each peak. To account for the differences in the response of the flood behaviour in terms of time according to the peak magnitude, we classify the discharge's magnitude into 5 ranges of discharges as shown in the peaks magnitude histograms (left part) of Figure 32 and Figure 33 (see also Figure 31).

We normalized all the peaks for each class by first identifying the time series of each peak, considering a window of 15 days previous to and after the peak and second dividing the discharges by the peak magnitude. As a result, we had N time series of 30 days with a peak magnitude equal to 1 for each class of discharges. (See Appendix C). Furthermore, in order to have a representative normalized hydrograph per class, we calculated the mean discharge of all the normalized hydrographs, which can be seen in the right part of Figure 32 and Figure 33 for the Rhine and Main Rivers respectively.

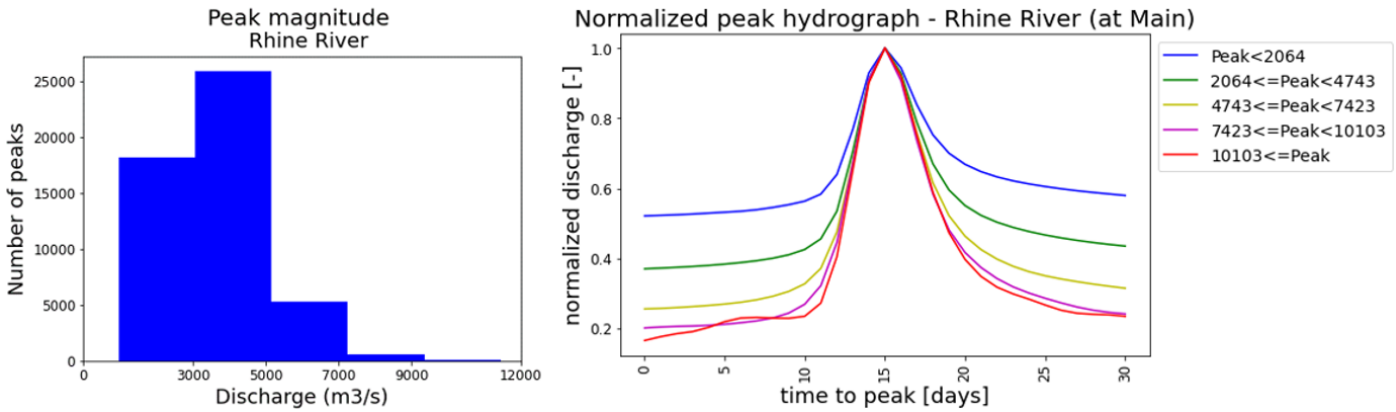


Figure 32 Rhine River discharge at confluence 68. Left: Peak discharge histogram. Right: Normalized peak hydrograph per range of discharge

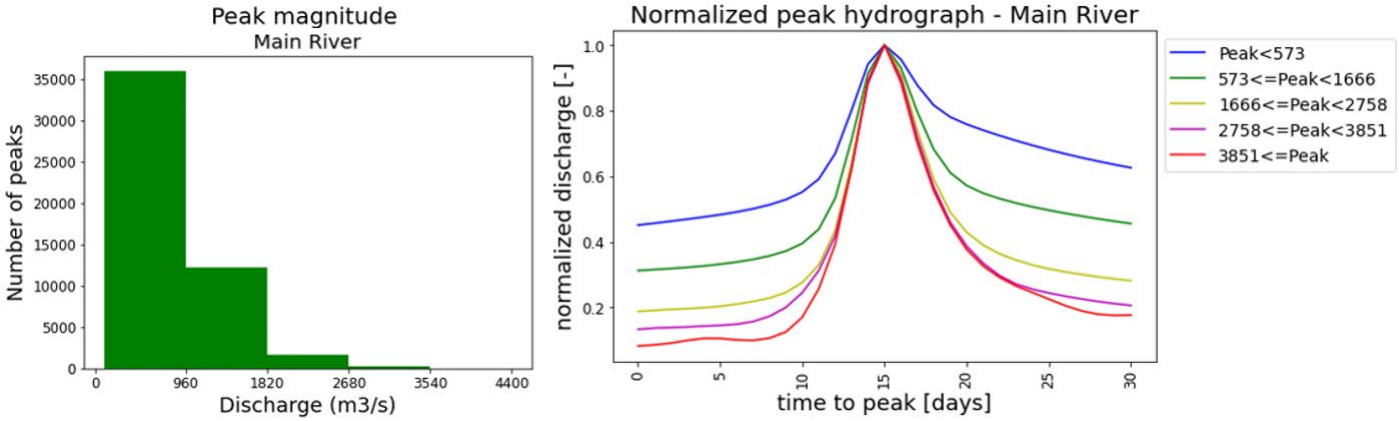


Figure 33 Main River discharge at confluence 68. Left: Peak discharge histogram. Right: Normalized peak hydrograph per range of discharge

Time difference between the 2 river annual maxima peaks

The day of the year when the annual maxima peak occurs was identified for both, the Rhine River (Stream 1) and the Main River (Stream 2). Then the difference in time of the peak occurrence is calculated for each year, and a histogram is constructed to evaluate them as shown in Figure 34.

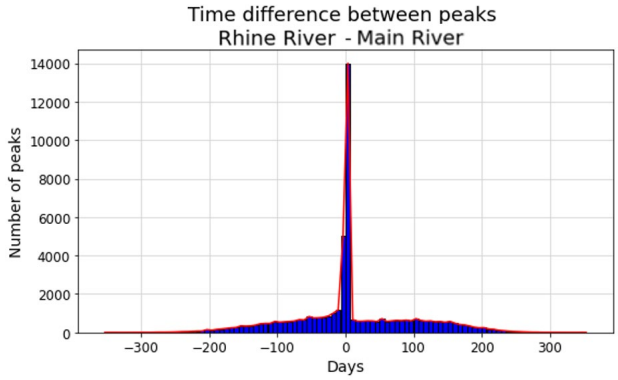


Figure 34 Time difference between the Rhine peaks and the Main peaks at Confluence 68

The most common range of the time difference between the 2 river peaks is from 0 to 1 day (31% of the events). Therefore, the time considered in the hydraulic simulation is 1 day for all the simulations (see Figure 35).

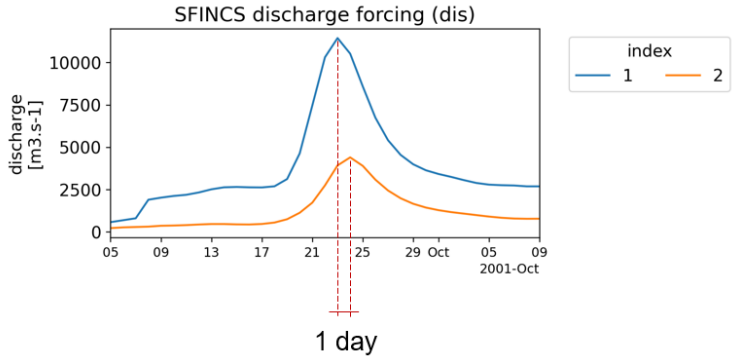


Figure 35 Example of the forcing discharge: time-series of the Mainstream (blue) and Tributary stream (orange), considering 1 day between the peaks

5.2.3 Model setup

The model was set up for the selected location, Confluence 68, where the Main River joins the Rhine River. Figure 36 shows the modelled area and its elevation map, which is complemented by Table 6. The grid resolution of the model was selected to be 200m x 200m to limit the computational time of the simulations.

Table 6 Distance and elevation of the river sections in the hydraulic model (SFINCS)

	Distance	Elevation	
	km	m+msl	m+msl
Rhine - From Upstream to confluence	69.89	88.23	83.32
Rhine - From confluence to downstream	39.21	83.32	77.1
Main - From Upstream to confluence	51.22	98.95	83.32

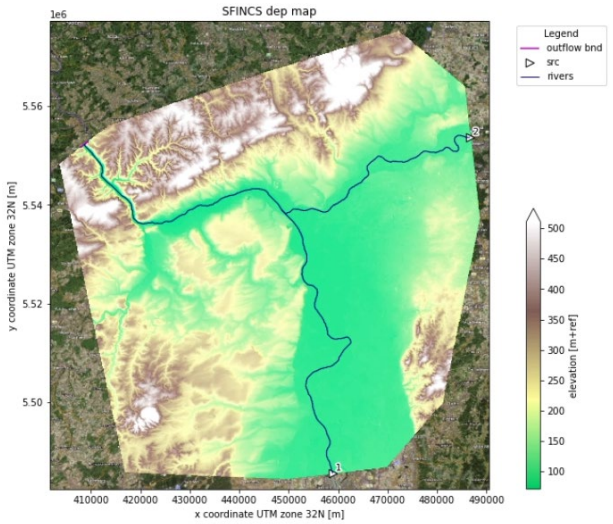


Figure 36 Delimitation of the hydraulic model (SFINCS) and the Digital Elevation Model (DEM)

The model was calibrated by adjusting the river depth to ensure that there is no flooded area when having a discharge that corresponds to a return period of 2 years for each river (Rhine $Q_{T=2years}$=3,454 m^3/s, Main $Q_{T=2years}$=701 m^3/s). The calibrated river depths are -26.6m for the Rhine River and -14.1m for the Main River.

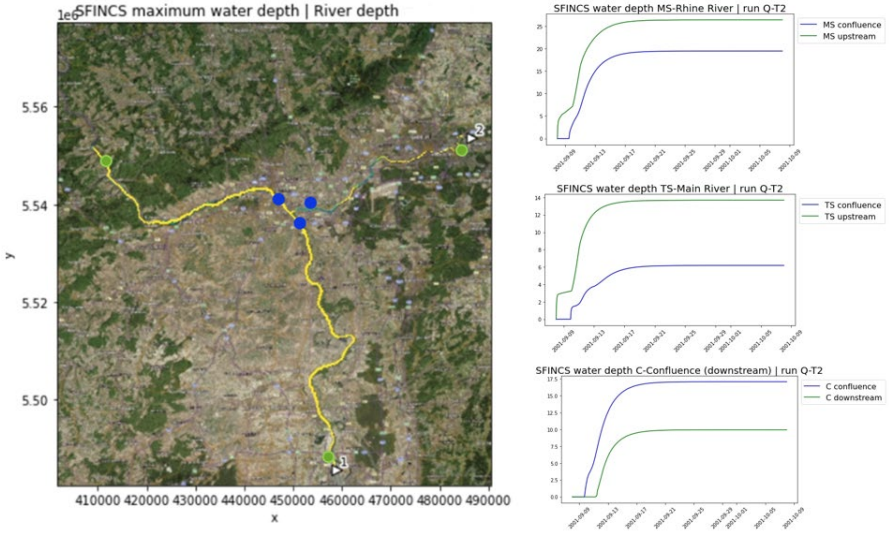


Figure 37 Calibration of the hydraulic model using a 2 years return period discharge ($Q_{T=2years}$). Left: Flooded map. Right: water depth at the Rhine River (upper, *MS*), the Main River (middle, *TS*), and downstream of the confluence (bottom, *C*)

5.3. Results and Evaluation

In this section, first, we present the results of the 25 hydraulic simulations. Subsequently, we evaluate the hydraulic interactions. Finally, we present the response functions.

Simulations

We performed the 25 simulations according to the 5x5 grid of possible discharge combinations (see Section 5.2.2, Figure 31 and Table 5). The flooded maps, according to the maximum water depth (h_{max}) of the simulations, are presented as a grid according to each river discharge in Figure 38, the x-axis represents the 5 Rhine discharges, and the y-axis the 5 Main discharges.

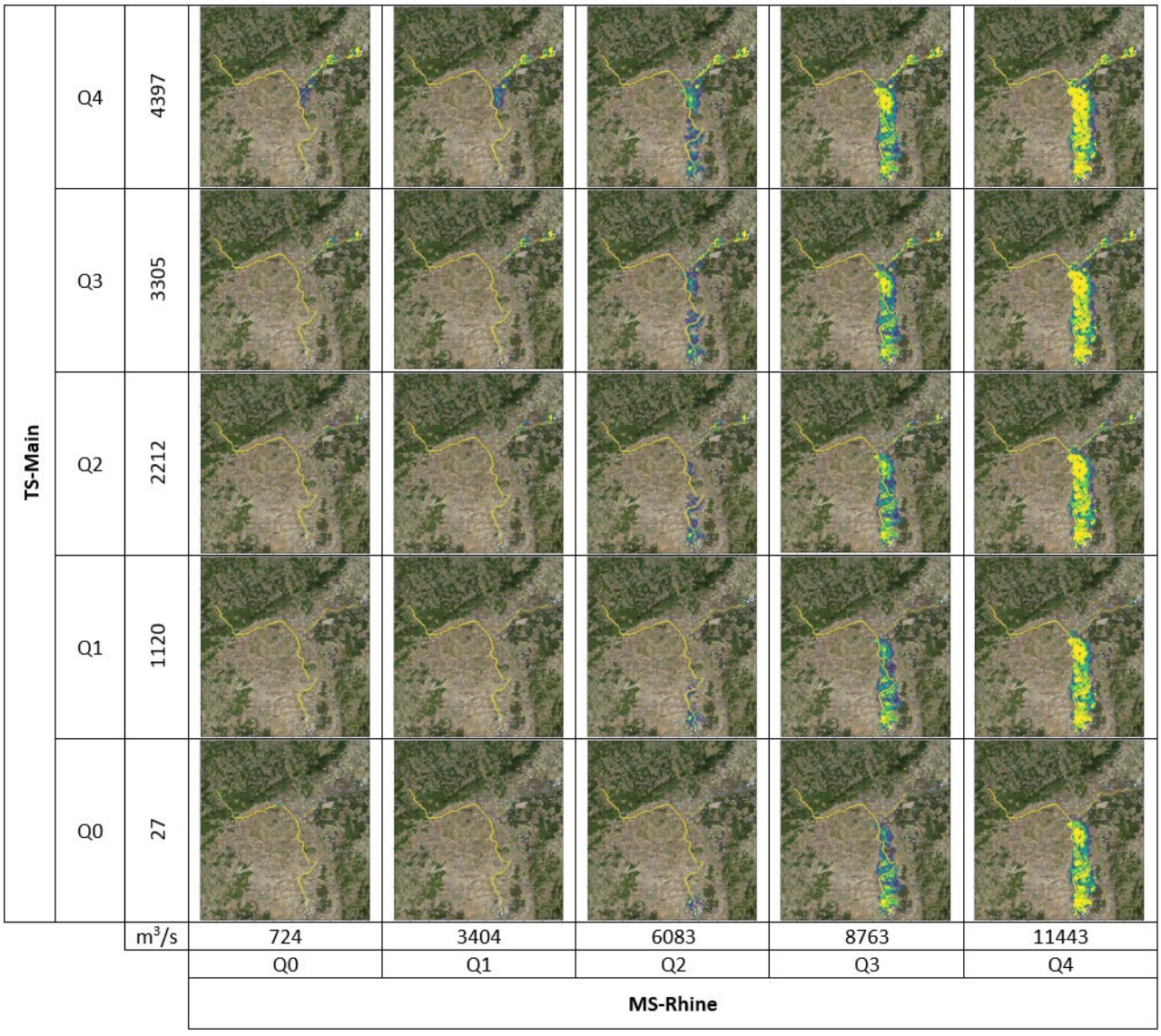


Figure 38 Hydraulic simulation results. Inundation maps (h_{max}) according to the grid of discharges: Mainstream (*MS-Rhine*) on the x-axis and Tributary stream (*TS-Main*) on the y-axis

In Figure 38 we can observe that the flood occurs upstream of the confluence, even in the worst scenario, we do not observe downstream areas with flooding. This is attributed to the topography of the confluence. Figure 36 shows the Digital Elevation Model (DEM) of the hydraulic model, where upstream near the confluence is relatively flat and wide, but downstream (right after the confluence) the flat terrain becomes narrow.

Additionally, in Figure 38 we observe that higher discharges of the Rhine River cause more flooding than higher discharges of the Main River.

Hydraulic interactions

In order to evaluate the hydraulic interactions and flood patterns in more detail, Figure 39 shows the difference in the maximum water depth (h_{max}) between the maximum (*Q4*) and minimum (*Q0*) discharges of the Rhine River, in combination with the 5 discharges of the Main River (from Q0 to Q4).

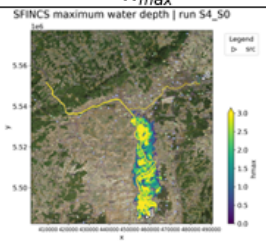
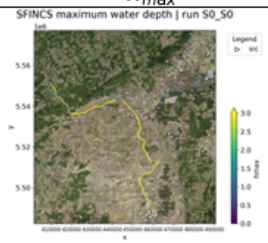
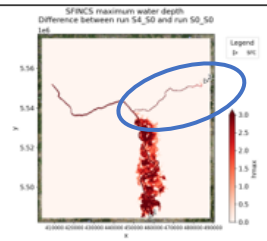
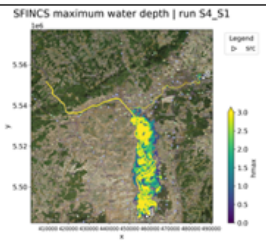
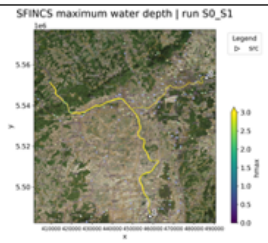
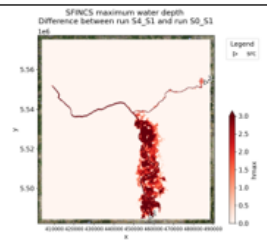
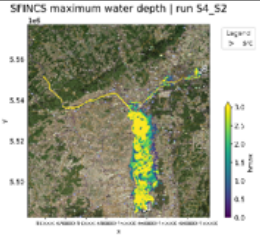
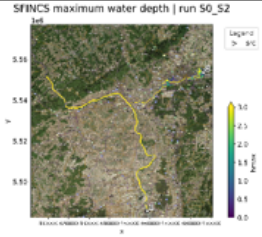
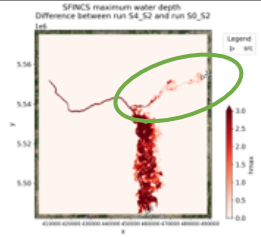
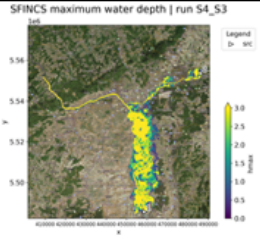
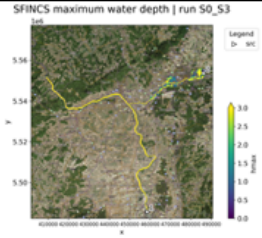
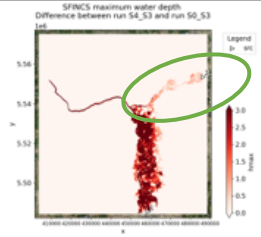
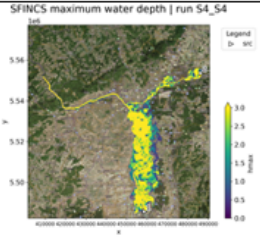
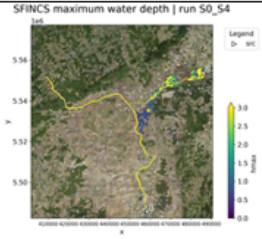
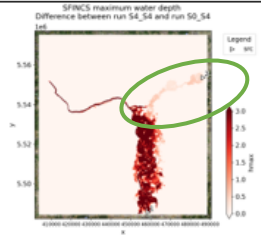
Simulation 1			Simulation 2			Difference
Rhine	Main	h_{max}	Rhine	Main	h_{max}	
Q4	Q0		Q0	Q0		
Q4	Q1		Q0	Q1		
Q4	Q2		Q0	Q2		
Q4	Q3		Q0	Q3		
Q4	Q4		Q0	Q4		

Figure 39 Influence of the Rhine River on the Main River flooding. Difference between the highest and lowest discharges of the Rhine River.

Similarly, Figure 40 shows the difference in the maximum water depth (h_{max}) between the maximum (Q4) and minimum (Q0) discharges of the Main River, in combination with the 5 discharges of the Rhine River (from Q0 to Q4).

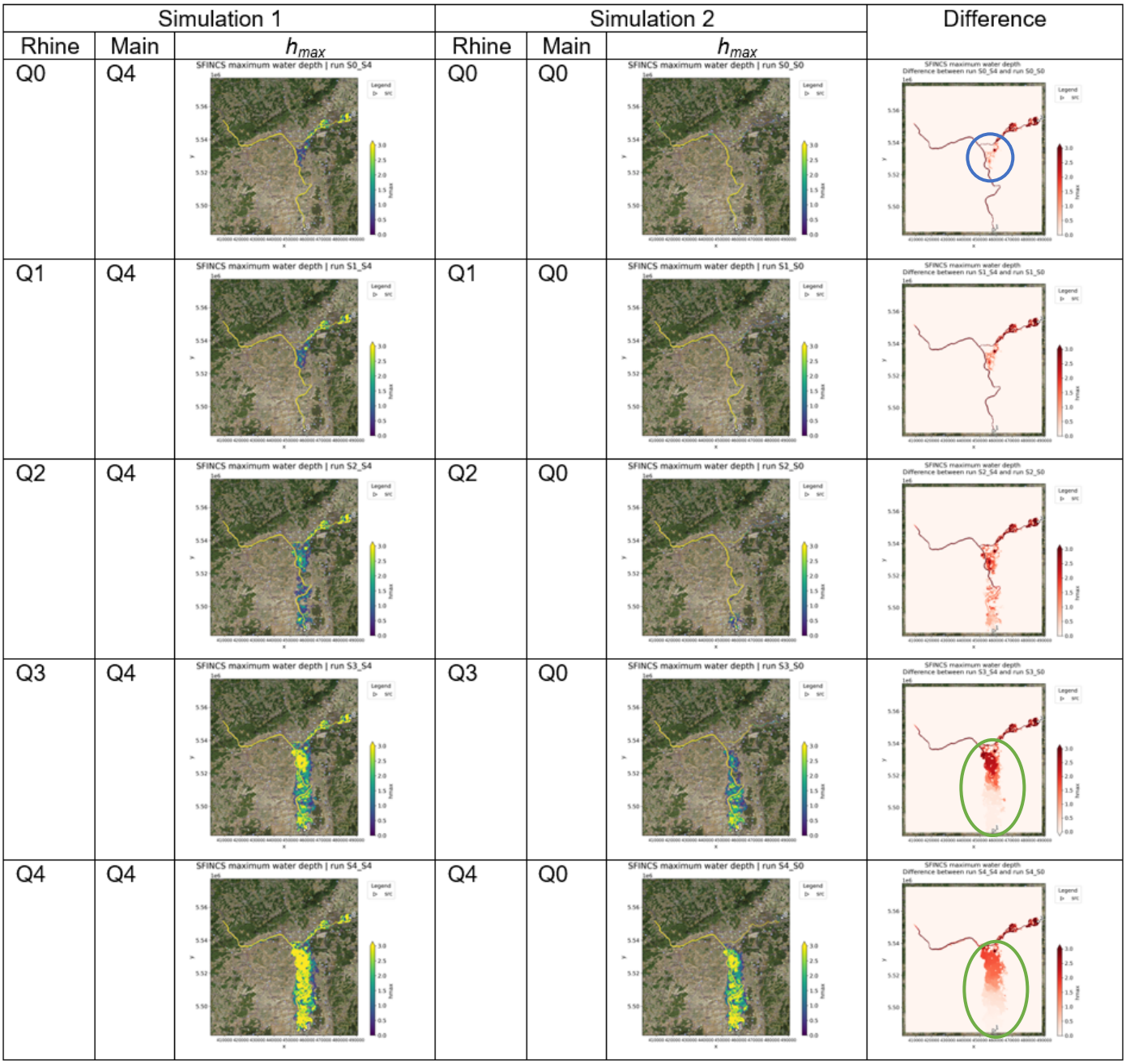


Figure 40 Influence of the Main River on the Rhine River flooding. Difference between the highest and lowest discharges of the Main River.

From Figure 39, we can observe that the Rhine River influences the flood area/pattern of the Main River. By looking at the blue circle, we can observe that there is no flood in the Main when it does not have high discharges ($Main_{Q0}$), even if the Rhine River presents the maximum peak discharge ($Rhine_{Q4}$). However, when looking at the green circles, we can observe that when the Main River has higher discharges (already observed at $Main_{Q2}$), the high discharge at the Rhine ($Rhine_{Q4}$) has an influence on the flood area in the Main, all along the Main River, we observe flooded area. Hence hydraulic interactions play a role in the inundation.

In Figure 40, we can observe that the Main River also influences the flood area/pattern of the Rhine River. When there is a minimum discharge at the Rhine River ($Rhine_{Q0}$), the high discharge of the Main River ($Main_{Q4}$) already causes inundation on the Rhine River side, close to the confluence (see the blue circle). This reflects the effects of the Main River on the flood area along the Rhine. Moreover, by looking at the green circles, when both rivers have high discharges (Q3 and Q4), the influence of the Main River on the flood area along the Rhine reaches ~33 km upstream of the confluence.

Response functions

As explained in Section 5.2, from the 5x5 grid of discharge combinations, two response functions are obtained, area and volume of inundation, both responses are a function of the Rhine and Main River discharges as shown in Figure 41.

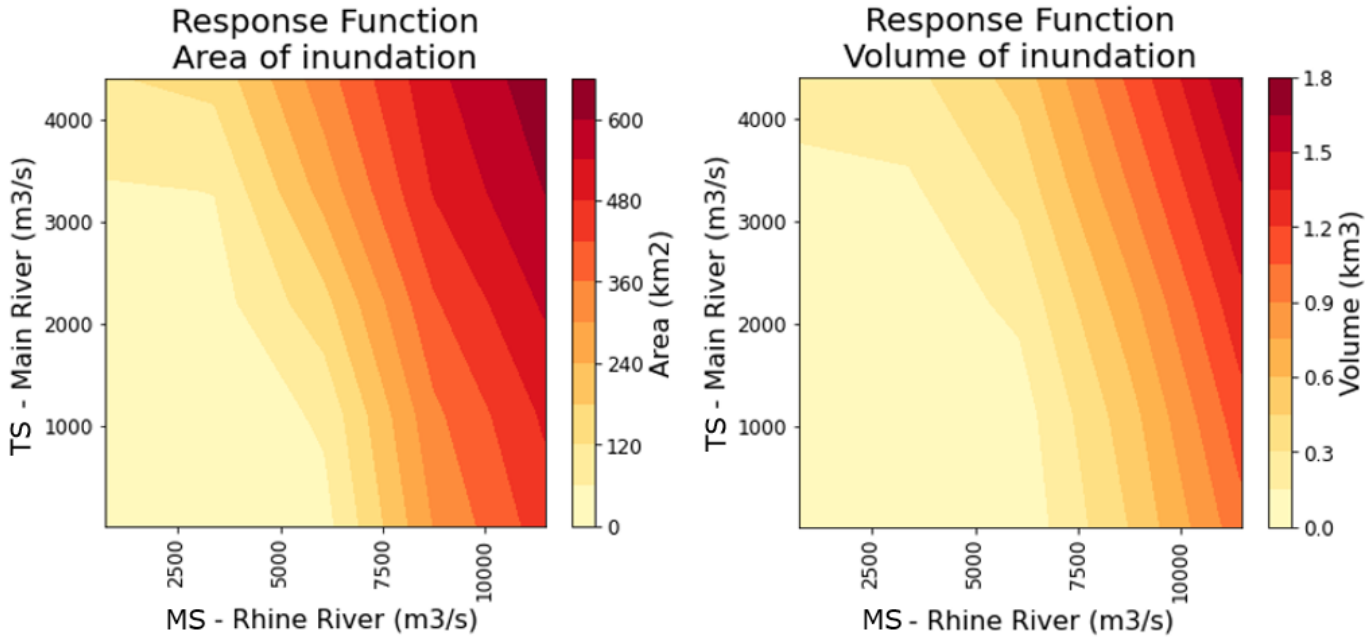


Figure 41 Response function according to the river discharges (Mainstream (MS-Rhine) on the x-axis and Tributary stream (TS-Main) on the y-axis). Left: Area of inundation. Right: Volume of inundation

In general, the increasing pattern of both responses is relatively similar. It is possible to see that the responses do not depend equally on both rivers as they do not increase in a 45° pattern. The Rhine River has more influence on both responses. Moreover, we can observe a difference in the increasing pattern of the flooded area, when having low discharges at the Rhine River and high discharges at the Main River the flooded area and volume already increases. This reflects the influence of the Main River on the Rhine flood area, explained earlier in Figure 40.

5.4. Conclusion and answer to Research Question 2

In this section, we aim to answer Research Question 2: *What are the hydraulic interactions between joining rivers?*

Confluence 68, located in the outlet of the Upper Rhine region where the Main River joins the Rhine River, was selected to perform the hydraulic evaluation. A more detailed hydrological evaluation of the location was performed and the hydraulic model was set up.

From the hydrological evaluation, we defined the boundary conditions for the hydraulic model. First, September was identified as the month with the fewest extreme peaks and therefore selected as the start of the hydrological year of the block maxima. Subsequently, the minimum and maximum peak magnitudes of both rivers were identified from the data and selected as the boundaries of a 5x5 grid of discharges, so in total 25 possible combinations were used to perform the hydraulic simulations. Finally, representative normalized hydrographs (according to the magnitude of the peak discharge) were used to provide the input time series, which at the same time considered 1 day as a time difference between the 2 river peaks.

Overall, 25 hydraulic simulations of possible discharge combinations were carried out. The results were evaluated in terms of the maximum water depth (h_{max}), identifying the differences between each simulation and the hydraulic interactions that take place at the evaluated location. Moreover, two responses function

were obtained, which are calculated in terms of the flooded area and the flooded volume as a surrogate for socio-economic impacts.

We observed that the flood occurs upstream of the confluence, even in the worst scenario, we do not observe downstream areas with flooding. This is attributed to the topography of the confluence: upstream near the confluence is relatively flat and wide, but downstream (right after the confluence) the flat terrain becomes narrow. Additionally, the hydrograph of the water depth directly downstream of the confluence is predominated by the Rhine River and is always lower than directly upstream of the confluence. This difference in depth shows that the water is contained in the flood plain, reducing depths and flooded areas at the downstream end of the confluence.

Based on the evaluation of the hydraulic interaction, we observed that the influence of the Rhine River (mainstream) on the Main River (tributary stream) water depth is substantial due to the backwater effects. The backwater effects on the Main are mainly observed close to the confluence, where the magnitude of the water depth peak increases, in time and magnitude, as the magnitude of the Rhine discharge increases. Additionally, when a high discharge at the Rhine coincides with a high discharge at the Main, we observed not only backwater effects close to the confluence, but also an influence on the flooded area all along the Main River.

The Main River also influences the flood area/patterns of the Rhine River. When there is a low discharge at the Rhine, we observed an increase in the magnitude of the water depth as the Main discharge increases. Nevertheless, when the Rhine has a high discharge, its water depth does not increase with increasing Main discharge, but we still observed changes in the duration of the peak water level, and the influence on the flooded area reaches half of the simulated distance of the Rhine River.

Based on the response functions, the increasing pattern of both the area and volume responses is relatively similar. The worst scenario, as expected, occurs when both rivers have the maximum peak discharges simultaneously. The responses (flooded area and flooded volume) do not depend equally on both rivers. The Rhine River has more influence: higher discharges on the Rhine River cause more flooding than higher discharges on the Main River.

Throughout this second evaluation, we have found the different hydraulic interactions of the Rhine and Main confluence, which vary according to the combination of river discharges. If we assume fully statistical dependence (spearman's correlation equal to 1) of the mainstream and the tributary stream during extreme events, we would expect to have a narrow cloud distribution of the extreme discharge sets. Hence, the range of possible combinations of peak discharges would be reduced and so the possible hydraulic interactions, allowing to have a straighter forward process to evaluate the expected damage scenarios.

However, the obtained spearman's correlation is 0.54 (for Set 1) with a wider distribution of possible discharge combinations, leading to different hydraulic interactions and a challenge to evaluate the expected damage when a hydraulic evaluation is not performed. For the study confluence, if measures to reduce the flood damages want to be taken, they can be focused on the Rhine River as we did observe that it has more influence on the flooding.

6. Design flood event

In this chapter, we answer research question 3 and its sub-questions, which are repeated below.

- *How should hydrological events be sampled for performing simulations in a flood risk analysis and/or determining design flood events?*

  The choices we make when sampling the hydrological events will have an impact on the results. Therefore, it is important to evaluate their sensitivity to the sampling approach. First, different methods of extreme selection can be applied. Second, we need to select the copula that best accounts for the statistical dependencies between the joining rivers. Third, for joining rivers multiple extreme situations can lead to a flooding event at the confluence, therefore different sets of events should be analyzed. Finally, long synthetic time series leads to a more accurate result (less uncertainty). Unfortunately, it is not always possible to generate such long synthetic data, hence, it is important to evaluate whether there is an optimal length that leads to the required accuracy of the results. The following questions will be addressed:

  3.1. Which **extreme selection method** should be used?
  3.2. Which **copulas** should be applied as a bivariate method?
  3.3. How should the **sets of simultaneous and non-simultaneous events** at the confluence be selected and applied?
  3.4. How and where should the **long synthetic time series** be implemented? Is it possible to find an optimal length?
  3.5. How sensitive are these choices for the sampling approach?

In Section 6.1 we explain the methodology that is used to evaluate the different choices we can use for hydrological sampling. Subsequently, in Section 6.2 the results and evaluation are presented. Finally, in Section 6.3 we conclude this chapter by answering research question 3.

This section is carried out in the selected confluence in the previous section: Confluence 68, where the Main River joins the Rhine River.

6.1. Methodology and Experiment setup

In this section, we explain the experiment setup (Figure 42) that is used to answer Research Question 3 and its sub-questions.

Data Processing

The hydrological data consists of 50,000 years of simulated daily mean discharge per catchment.

Similar to the previous Sections, we evaluate daily events as the HBV model considers already the hydrological processes and routes the runoff, thus the travel time is already considered.

Extreme selection

As explained in Section 2.1, there are different methods to evaluate extremes such as Peaks Over Threshold (POT) and Annual Maxima (AM). Due to time constraints, only Annual Maxima is used in this section. However, as shown in Deltares (2021), for its specific case study, AM in that particular case led to higher values of the design discharge than POT. Hence, it is important to corroborate whether that conclusion holds for other cases of study and to evaluate the sensitivity of the extreme selection method.

Moreover, as indicated in Section 2.2.5, we differentiate between the mainstream (*MS*), the tributary stream (*TS*), and the confluence (*C*) as shown in Figure 2. The sets of extremes considered for the analysis are non-simultaneous and repeated below:

- *Set 1 | MSmax-TSconc: mainstream maxima (MS_{max}) with a concurrent tributary stream (TS_{conc}),*
- *Set 2 | TSmax-MSconc: tributary stream maxima (TS_{max}) with a concurrent mainstream (MS_{conc}),*
- *Set 3 | Cmax (C=MS+TS): the maxima of the sum of both streams.*

From the 50,000 years of daily data, 50,000 annual maxima peaks are identified for each set.

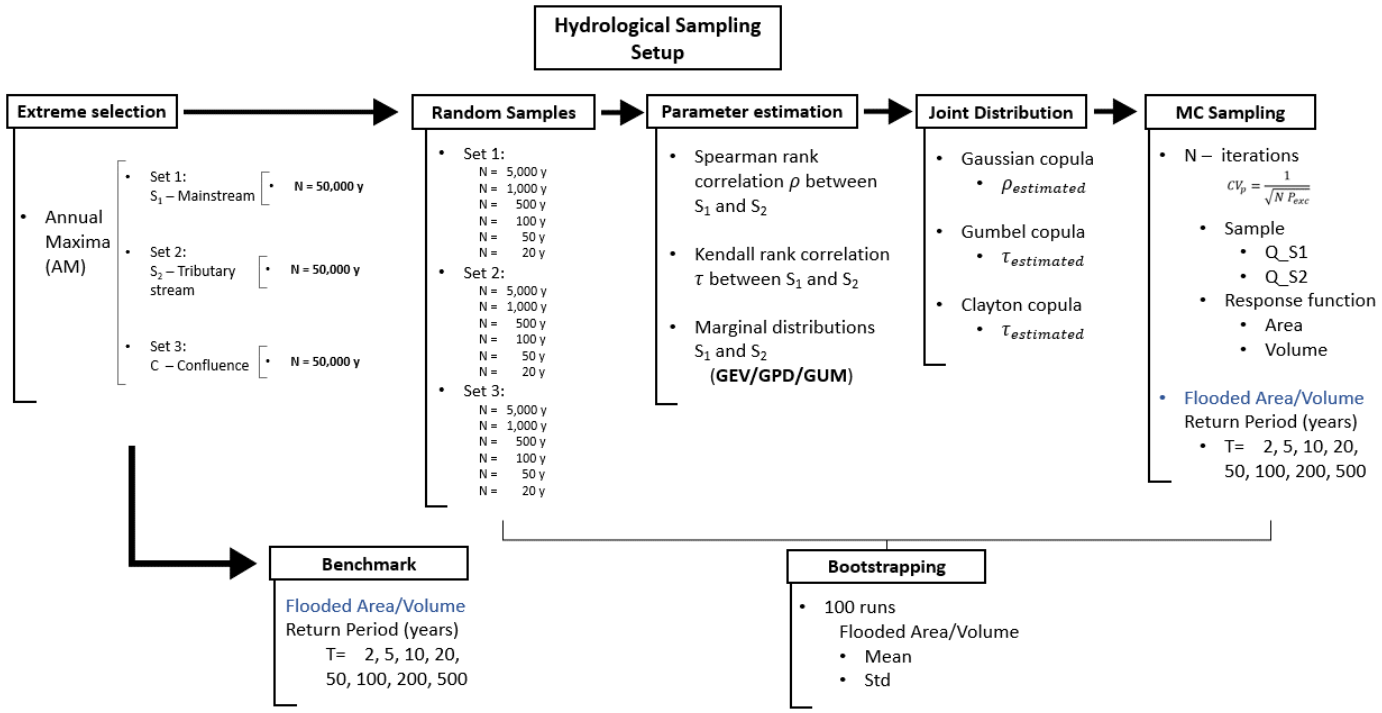


Figure 42 Process to perform the hydrological evaluation and sampling to obtain the flooded area that corresponds to certain return periods

Random Samples

Having 50,000 years of synthetic data is not always possible. If no synthetic data is available, observed data is used. Common hydrological records contain 20 to 100 years of data. Therefore, it is important to evaluate the results and variability of samples with common record sizes.

Even though generating synthetic data out of the records allows us to perform a more robust risk assessment, it is not always possible to generate it due to limitations in time, computational capacity and budget. Therefore it is relevant to evaluate the size of such synthetic series to determine whether there is an optimal length to overcome possible constraints.

The sample sizes that are evaluated in this section are different in length (20, 50, 100, 500, 1000, and 5000 years), and the selection is random as the selected extremes out of the 50,000 years of synthetic data are independent and identically distributed values.

Parameter estimation and Joint distribution

The joint probability distributions of the joining rivers are described/constructed by using three different copulas, Gaussian, Gumbel and Clayton. To derive the parameters of these copulas, rank correlations are needed.

While the spearman rank correlation is used for the Gaussian copula, the Kendall rank correlation is used for both, the Gumbel and Clayton copula. The rank correlations are estimated for the annual maxima peaks and concurrent flows of each set (Set 1 | MS_{max}-TS_{conc}, Set 2 | TS_{max}-MS_{conc}, and Set 3 | C_{max} (C=MS+TS)). Moreover, when using samples of different sizes, the rank correlations are calculated for each set of samples.

Moreover, the marginal distributions of both rivers (peak or concurrent) discharges are needed to construct the bivariate distribution function. As mentioned in Section 2.1.2. when selecting the annual maxima, the distribution converges to a Generalized Extreme Value (GEV). Therefore, for the rivers where the annual maxima are selected, we can use the GEV distribution.

However, the concurrent flow is not necessarily an extreme situation, therefore it is important to evaluate the fit of the distribution. The full data (50,000 years) is used to evaluate the distribution fit. First, constructing the empirical distribution of the variables (river peak/concurrent discharges) of each set (Set 1 | MS_{max}-TS_{conc}, Set 2 | TS_{max}-MS_{conc}, and Set 3 | C_{max} (C=MS+TS)), and second, comparing it to four different probability distributions: Gumbel (GUM), Generalized Extreme Value (GEV), General Pareto Distribution (GPD), and Normal (Nor). Three different scores were evaluated (RMSE, AIC, BIC), and the distributions with the best results were selected (see Table 7). The full evaluation of the distribution fit can be found in Appendix E.

Table 7 Marginal distribution fit of the Rhine (MS) and Main (TS) rivers according to the extreme set

	Rhine (MS)	Main (TS)
Set 1 | MS_{max}-TS_{conc}	GEV	GPD
Set 2 | TS_{max}-MS_{conc}	GEV	GEV
Set 3 | C_{max} (C=MS+TS)	GEV	GPD

The parameters of the selected distributions are determined by fitting the distribution to the selected data. When using samples of different sizes, the parameters are determined for each sample. There are several methods available for this task, such as the method of moments, maximum likelihood and probability-weighted moments. we chose to use the L-moments method (Hosking & Wallis, 1997).

Monte Carlo sampling

Monte Carlo simulations are implemented to sample the river discharges by using the joint probability constructed with the copulas and the marginal distributions and to evaluate the response function obtained from the hydraulic simulations. The procedure for this is as follows:

1. Sample N_{MC} combinations of the 2 rivers (peak and concurrent) discharges
2. Use the response function to estimate the flooded area (obtained in Section 5)
3. Construct the empirical distribution of the N flooded areas
4. Use the empirical distribution to estimate the flooded area (A_T) that corresponds to a certain annual probability of exceedance (or return period)

Different return periods are evaluated, T= 2, 5, 10, 20, 50, 100, 200, and 500 years.

The number of Monte Carlo samples (N_{MC}) is determined using Equation 11. The largest return period that is evaluated is T=500 years, corresponding to an annual probability of exceedance P_{exc} =*0.002*. The target *CV* is set to 0.05, therefore *200,000* Monte Carlo samples are needed.

Bootstrapping

It is relevant to calculate the variability of the results when using different sample sizes by random selection. Therefore, a bootstrapping experiment is carried out, the procedure is repeated 100 times for each sample size, being able to calculate the standard deviation of the results as follows:

1. From the full data set, select a random sample of size N
2. Estimate the parameters, rank correlations and marginal distributions, out of the random sample set
3. Perform the N_{MC} realizations of Monte Carlo simulations, and store the estimated flooded area that corresponds to the desired return periods (A_T)
4. Repeat steps 1-3 100 times
5. Calculate the mean and standard deviation of the flooded area (A_T)

Benchmark

To compare the results, we determined a benchmark from the 50,000 years of synthetic data. First, for each of the extreme sets, we determined the flooded area of all the discharge combinations by using the response function obtained in Section 5, resulting in 3 sets of 50,000 flooded areas. Subsequently, we constructed the empirical distribution of the flooded area for each set and estimated the flooded area (A_T) that corresponds to different return periods (T= 2, 5, 10, 20, 50, 100, 200, and 500 years).

6.2. Results and Evaluation

In this section, we present the benchmark and the results for the estimated flooded area (A_T) when performing the hydrological sampling that is explained in Section 6.1.

The 100-year return period flooded area (A_{T100}) is estimated for the Annual Maxima Set 1 | MS_{max}-TS_{conc} , using different sample sizes (20, 50, 100, 500, 1000, and 5000 years). The sensitivity analysis of both, extreme sets and the different return periods, are evaluated in section 7.2.

The Benchmark

In Figure 43 and Table 8 the benchmarks for each extreme set (Set 1 | MS_{max}-TS_{conc}, Set 2 | TS_{max}-MS_{conc}, and Set 3 | C_{max} (C=MS+TS)) are presented.

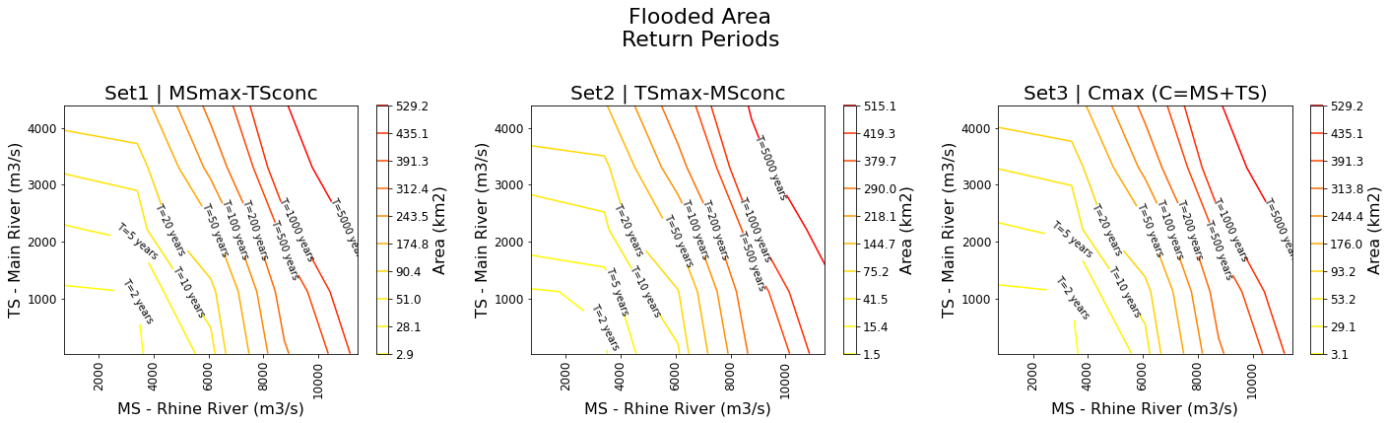


Figure 43 Flooded area for different return periods (T), obtained from the response function using the different sets of extremes: Set 1 (left), Set 2 (middle), and Set 3 (right)

Table 8 Benchmarks: Flooded area for different return periods (T), and each extreme set

	Set 1 | MS_{max}-TS_{conc}	Set 2 | TS_{max}-MS_{conc}	Set 3 | C_{max} (C=MS+TS)
	Flooded Area (km^2)	Flooded Area (km^2)	Flooded Area (km^2)
A_{T2}	2.9	1.5	3.1
A_{T5}	28.1	15.4	29.1
A_{T10}	51.0	41.5	53.2
A_{T20}	90.4	75.2	93.2
A_{T50}	174.8	144.7	176.0
A_{T100}	243.5	218.1	244.4
A_{T200}	312.4	290.0	313.8
A_{T500}	391.3	379.7	391.3
A_{T1000}	435.1	419.3	435.1
A_{T5000}	529.2	515.1	529.2

The joint probability pattern is similar for the three sets of extremes, except for T=5 years. We can observe that for small return periods (2, 5, and 10 years) both rivers have similar importance. However, when the return period is higher, the increasing pattern changes and the responses do not depend equally on both rivers as they do not increase in a 45° pattern, then the Rhine River has more influence on the response.

While the magnitudes of the flooded areas for the different return periods are similar for Set 1 and Set 3, when looking at Set 2, we can observe that they differ, resulting in a smaller flooded area. The difference is lower for larger return periods.

Flooded Area of a 100 years Return period estimation (Different N-samples)

In Figure 44 we can observe the estimation of the 100 years flooded area according to:

- The size of the sample N (x-axis)
- The used copulas: Gaussian (blue), Gumbel (green), and Clayton (yellow)
- The mean flooded area (bar) and its standard deviation (red solid line) (bootstrap experiment)
- The benchmark (red dashed line)

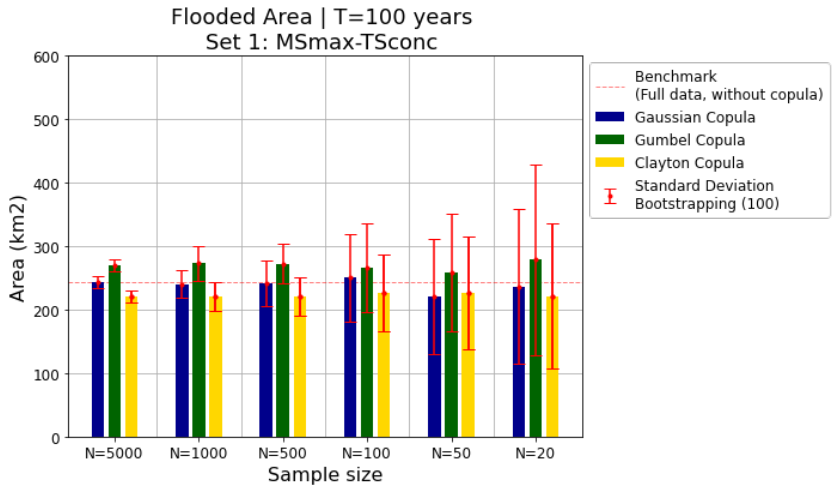


Figure 44 Flooded area that corresponds to a 100 years return period (A_{T100}) according to the size of the sample N (x-axis) and the copulas: Gaussian (blue), Gumbel (green), and Clayton (yellow). The benchmark (red dashed line) and the standard deviation (red solid line)

In Figure 45 we can observe the coefficient of variation (CV, y-axis) for each of the copulas (Gaussian: blue, Gumbel: green, and Clayton: yellow) and sample sizes (x-axis).

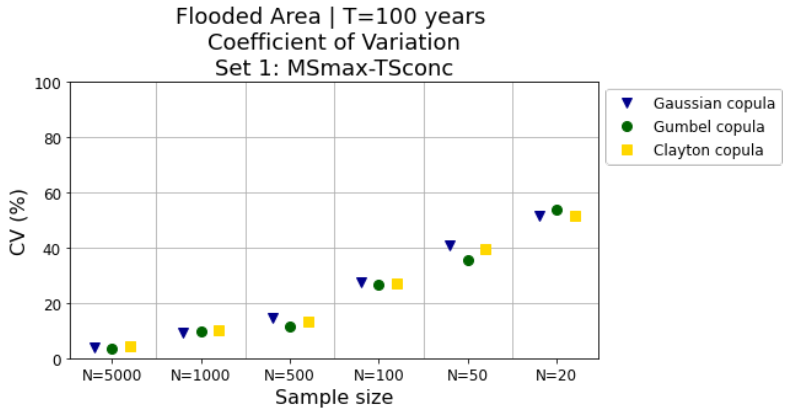


Figure 45 Coefficient of variation of the flooded area that corresponds to a 100 years return period (A_{T100}) according to the size of the sample N (x-axis) and the copulas: Gaussian (blue triangle), Gumbel (green circle), and Clayton (yellow square).

From Figure 44 we can evaluate first the mean flooded area estimated when using different copulas. The Gaussian copula led to the closest estimation to the benchmark, independent of the sample size N. While the Gumbel copula overestimates the results, the Clayton copula underestimates them. This holds for all sample sizes, except for N=50 where the Gaussian copula led to an underestimated result.

When looking at the standard deviation in Figure 44 and the coefficient of variation in Figure 45, we can observe that there are no significant differences between the three copulas. However, the difference in CV between the different sample sizes is significant. The larger the sample, the smaller the standard deviation and the coefficient of variation.

6.3. Conclusion and answer to the Research Question 3

In this section, we aim to answer Research Question 3: *How should hydrological events be sampled for performing simulations in a flood risk analysis and/or determining design flood events?*

For this purpose, we continued evaluating the previously selected confluence 68, located in the outlet of the Upper Rhine region where the Main River joins the Rhine River, and for which we have the response functions, flooded area and flooded volume, which are a function of both river discharges (Section 5).

The annual maxima peaks were identified out of the 50,000 years of hydrological (discharge) synthetic data from GRADE. Three different ways of selecting the maxima were evaluated as mentioned in Section 2.2.5, leading to three extreme sets (Set 1 | MS_{max}-TS_{conc}, Set 2 | TS_{max}-MS_{conc}, and Set 3 | C_{max} (C=MS+TS)). From the sets, using the response function, we calculated the flooded area that can be expected for certain return periods, which were used as benchmarks.

Subsequently, random samples were selected from the full data set to resemble commonly recorded data (20, 50, and 100) and possible lengths for generating synthetic data (500, 1000, and 5000). The parameters of the copula functions, when using copulas to construct the joint probability of the joining river discharges, were calculated for each set. Then, using Monte Carlo simulations we sample the two river discharges and calculate the flooded area from the response function. After *N* Monte Carlo iterations, we constructed an empirical distribution of the flooded areas to later estimate the flooded area that corresponds to a certain probability of exceedance (or return period).

The process was repeated 100 times for each sample in a bootstrapping experiment to be able to quantify the variations in the results.

Based on the benchmarks we can conclude that for small return periods (2, 5, and 10 years) both rivers, the Rhine and the Main, have similar importance. However, when the return period is higher than 10 years, the Rhine River has more influence on the response in terms of the flooded area. Moreover, while the magnitudes of the flooded areas for the different return periods are similar for the extreme sets 1 and 3, Set 2 generally results in smaller flooded areas.

Based on the estimated 100 years return period flooded area (A_{T100}) when using different sizes of the sample and different copulas, we can conclude that the Gaussian copula led to the closest estimation to the benchmark independent of the sample *N*.

Based on the sensitivity analysis, the variation of the results is attributed not only to the size of the sample but also to the marginal distribution choice. We observed lower variation when we used larger samples, meaning that there is more data to fit the distribution, hence, less uncertainty.

Throughout this third evaluation, we have evaluated the different choices we make when sampling hydrological events to determine the design flood events. First, for the evaluated confluence, where the Main River joins the Rhine River, the Gaussian copula was the best fit for the three extreme sets. Nevertheless, for other confluences, we suggest always evaluating the selection of the copula. Second, we suggest considering non-simultaneous events for confluences with weak statistical dependencies of the mainstream and the tributary stream during extreme events. Set 1 | MS_{max}-TS_{conc} (mainstream maxima with a concurrent tributary stream) can be chosen as it is the worst scenario of the three sets of non-simultaneous events. When it is desired to reduce the marginal distribution selection uncertainty (especially with short time series), we recommend generating synthetic data.

7 Sensitivity analysis

In this chapter, we evaluate the sensitivity of the annual maxima selection and the size of the data (In-between scenarios). First, for the Meteorological and Hydrological evaluation (Section 4), and second, for the Hydrological sampling (Section 6).

7.1. Sensitivity analysis of the Meteorological and Hydrological evaluation

In this section, first, we present the results and evaluation for the 50,000 years of synthetic data (namely the benchmark), for which we evaluate the differences in results between the three sets of extreme events. Following, we present the results and evaluation for scenarios where not only the different extreme sets are considered but also different lengths of the size of data, assessing the variability of the results for other lengths than the full data.

7.1.1. Evaluation of the annual maxima of the full data set

As indicated in Section 2.2.5, we differentiate between the mainstream (*MS*), the tributary stream (*TS*), and the confluence (*C*). The sets of extremes considered for the analysis are non-simultaneous and named as follows:

- *Set 1 | MSmax-TSconc: mainstream maxima (MS_{max}) with a concurrent tributary stream (TS_{conc}),*
- *Set 2 | TSmax-MSconc: tributary stream maxima (TS_{max}) with a concurrent mainstream (MS_{conc}),*
- *Set 3 | Cmax (C=MS+TS): the maxima of the sum of both streams.*

We evaluate the differences between the 3 sets of meteorological events (Precipitation 5 days) and hydrological events (Discharge 1 day). In Figure 46 we can observe the meteorological correlations between the annual maxima peaks of the mainstream and the tributary stream per confluence. The map on the left shows the actual values of the Spearman rank correlation of Set 1. The middle and right maps show the difference in the correlation of Set 2 and Set 3 with respect to Set 1.

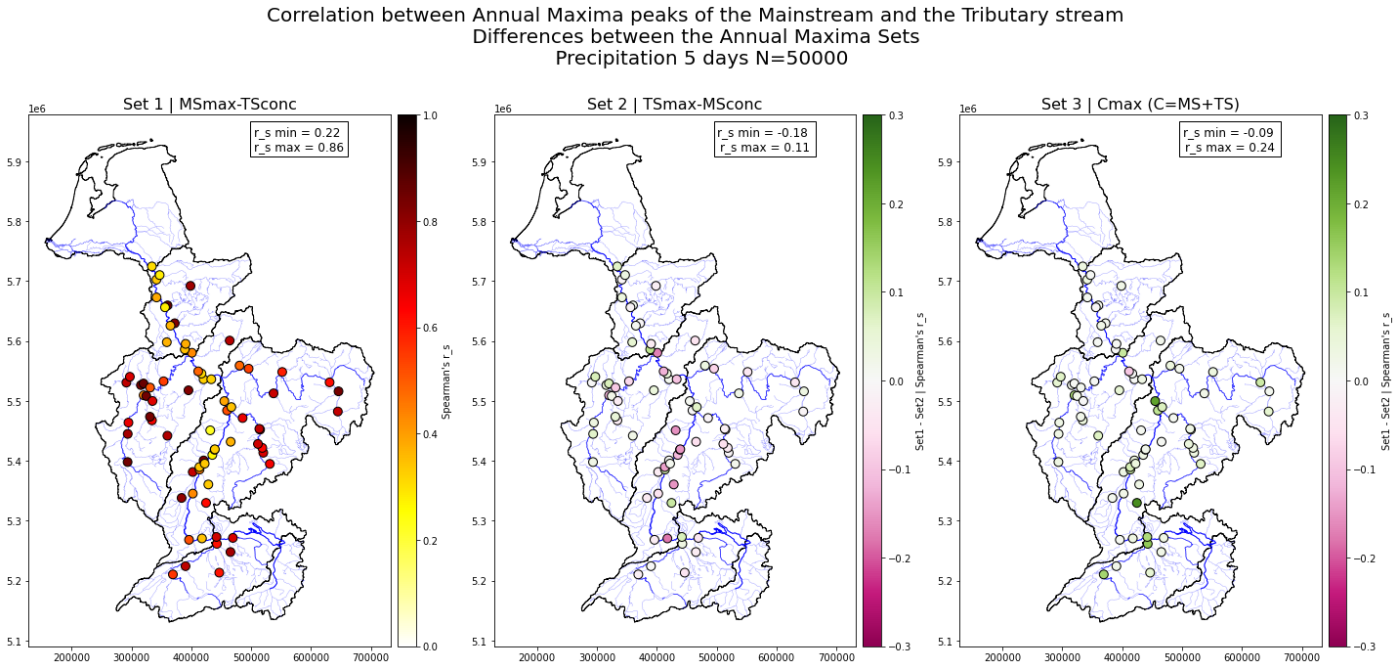


Figure 46 Differences between extreme sets of the Spearman's rank correlation between extreme events (meteorological: precipitation 5 days) of neighbouring catchments at each confluence. Left: Set 1 value of the Spearman's rank correlation. Middle: Difference between Se1 and Set 2. Right: the difference between Set 1 and Set 3.

Similarly, we can see the differences between the extreme sets of the hydrological event (discharge 1 day) in Figure 47.

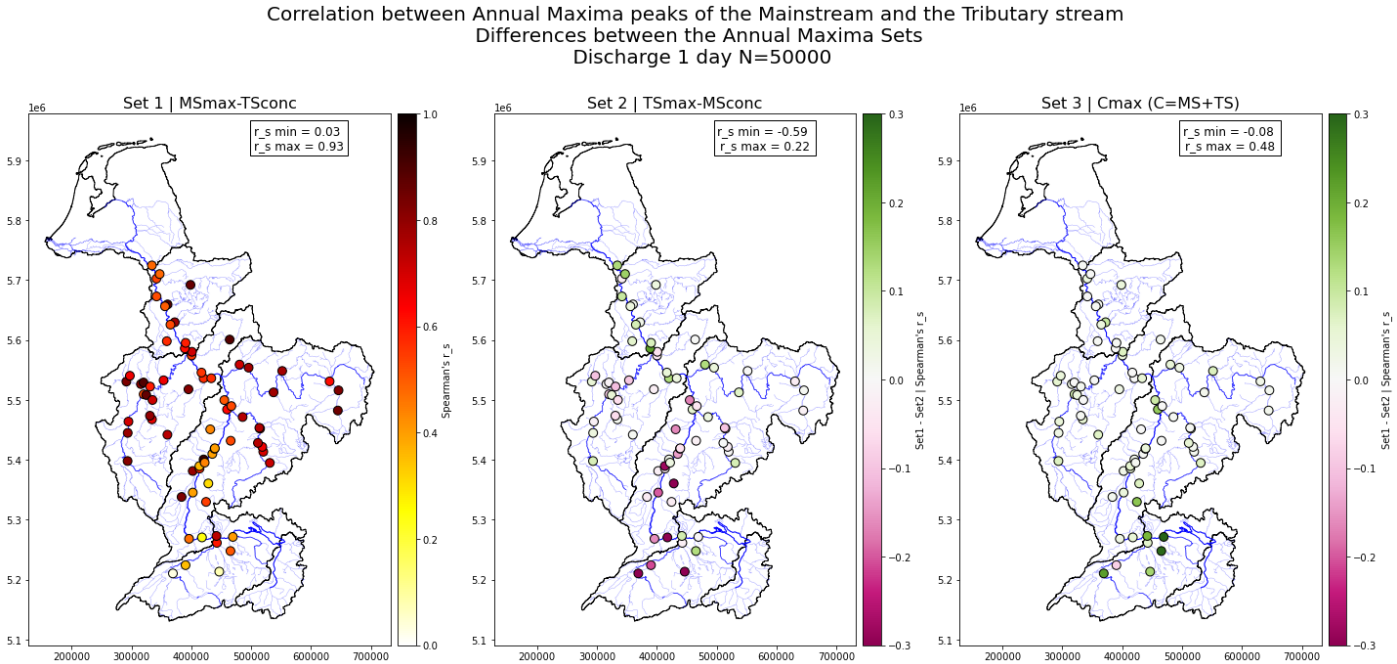


Figure 47 Differences between extreme sets of the Spearman's rank correlation between extreme events (hydrological: discharge 1 day) of neighbouring catchments at each confluence. Left: Set 1 value of the Spearman's rank correlation. Middle: Difference between Se1 and Set 2. Right: the difference between Set 1 and Set 3.

First, when looking at the meteorological results (Figure 46), the biggest differences are observed between Set 1 and Set 2. In general, Set 2 leads to lower correlations (pink dots). Set 3 leads to higher correlations (green dots). The confluences where differences are largest for both sets (Set 2 and Set 3) are located close to or in the main river, especially in the Upper Rhine and the Middle Rhine.

Likewise, when looking at the hydrological results in Figure 47, the largest differences are observed between Set 1 and Set 2. However, the patterns for the hydrological events are different from those of meteorological events. Set 2 leads to lower correlations in the upper regions (Alpine Rhine, High Rhine, and Upper Rhine) and higher correlations in the lower regions (Middle Rhine, and Low Rhine). Correlations of Set 3 are more similar to Set 1, except in the Alpine Rhine and High Rhine, where higher correlations are observed.

The differences between the 3 extreme sets confirm that a simultaneous event in which the maximum of the mainstream (MS_{max}) coincides in time with the maximum of the tributary stream (TS_{max}), which would be the worst-case scenario, is not likely to always occur.

7.1.2. Evaluation of the impact of the sample size

Different sets of peaks were selected randomly out of the long synthetic time series. These sets are different in length (5000, 1000, 500, 100, 50, and 20 years), the selection of years is random as the selected extremes are independent and identically distributed values.

As explained in Section 4.2, a bootstrapping experiment was done to calculate the variability of the results when using different sample sizes by random selection, the experiment is repeated 100 times for each sample size (except for N=5000 which was repeated 50 times due to computational time constrains, see also Figure 11), which makes it possible to calculate the standard deviation of the results.

In Figure 48, we can observe the variability of the calculated Spearman rank correlation per confluence and size of the sample for Meteorological events (precipitation 5 days). The three graphs show, in the x-axis the correlation per confluence, in the y-axis its standard deviation when performing the bootstrapping experiment,

and the colour represents the size of the sample as indicated in the legend. Each graph shows the results for one of the extreme sets, Set 1, Set 2, and Set 3 from left to right.

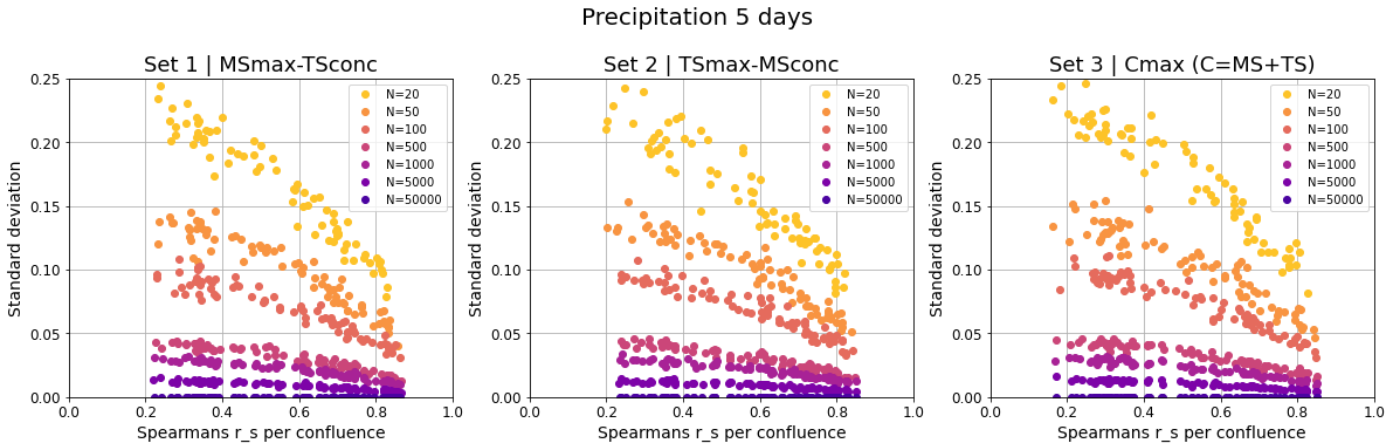


Figure 48 Standard deviation (y-axis) of the Spearman's rank correlation (x-axis) between extreme events (meteorological: precipitation 5 days) of neighbouring catchments per confluence, according to the size of the sample (colour rank). Left: Set 1. Middle: Set 2. Right: Set 3

Similarly, we can see the variability of the sample size for the hydrological events (discharge 1 day) in Figure 49.

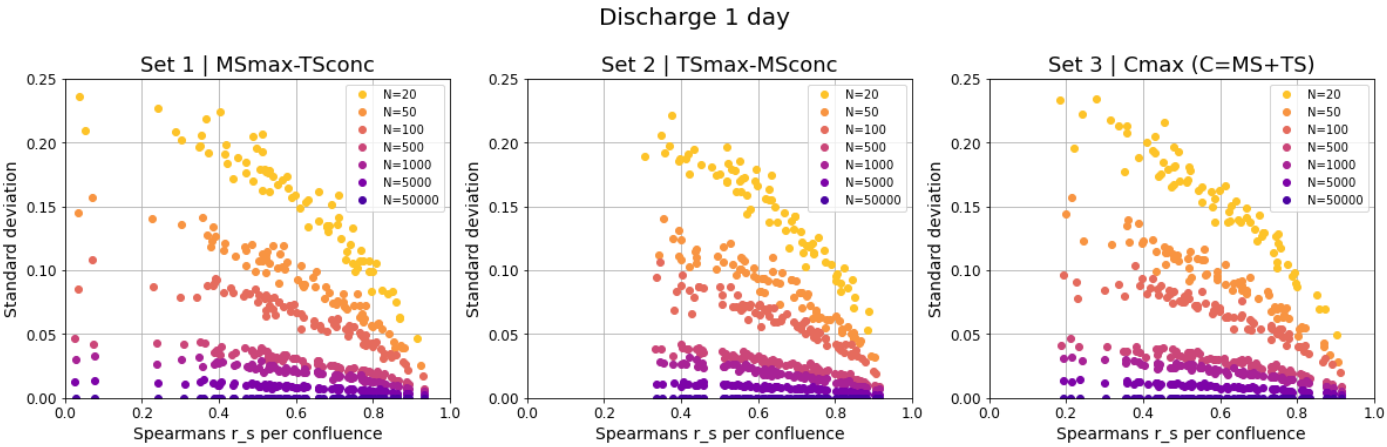


Figure 49 Standard deviation (y-axis) of the Spearman's rank correlation (x-axis) between extreme events (hydrological: discharge 1 day) of neighbouring catchments per confluence, according to the size of the sample (colour rank). Left: Set 1. Middle: Set 2. Right: Set 3

Figure 48 and Figure 49 show similar patterns not only for the two model domains (meteorological and hydrological) but also for the 3 different sets of extremes.

First, we can see a relationship between the correlation and the standard deviation, the higher the correlation, the lower the standard deviation. Recalling one of our conclusions in Section 4.5, a high (low) correlation can be associated with a small (large) total contributing catchment and/or with a small (big) size ratio between the two neighbouring catchments.

We can see the gradual change related to the size of the sample. the smaller the sample, the higher the standard deviation. Moreover, in all the graphs we can see that with a sample of 500 (100) years the standard deviation is reduced to 5% (10%) in low correlated confluences and 2.5% (5%) in high correlated confluences.

7.2. Sensitivity analysis of the hydrological sampling

In this section, we present the sensitivity of the results to the hydrological sampling choices. First, we present the results and evaluation for the different return periods T (2, 5, 10, 20, 50, 100, 200, and 500 years) in section 7.2.1. Following, we evaluate the differences between the three extreme sets. Moreover, we also evaluate the influence of the marginal distribution, in section 7.2.3. Finally, the copulas are evaluated in section 7.2.4.

7.2.1. Evaluation of different Return periods

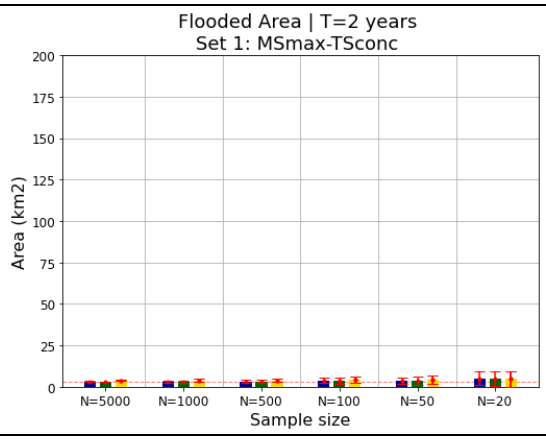
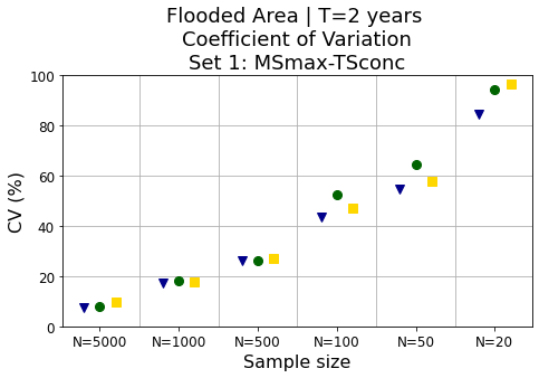
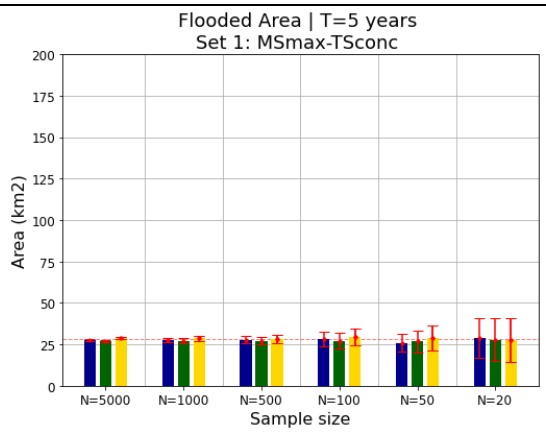
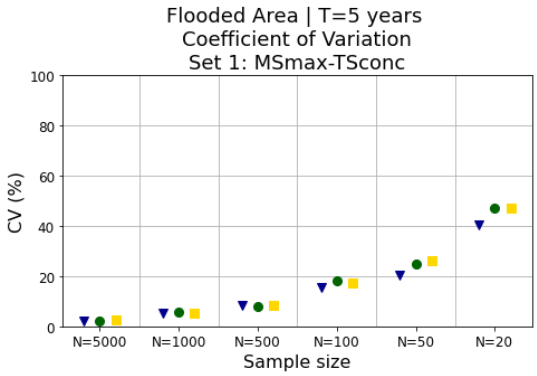
To evaluate the results for the different return periods T (2, 5, 10, 20, 50, 100, 200, and 500 years) we used the annual maxima Set 1 (MS_{max}-TS_{conc}) and all sample sizes N (20, 50, 100, 500, 1000, and 5000).

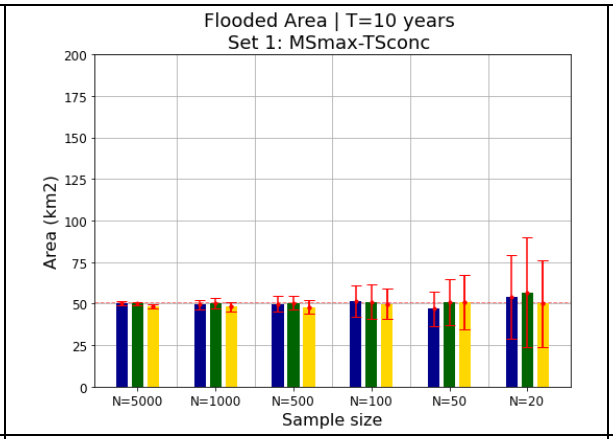
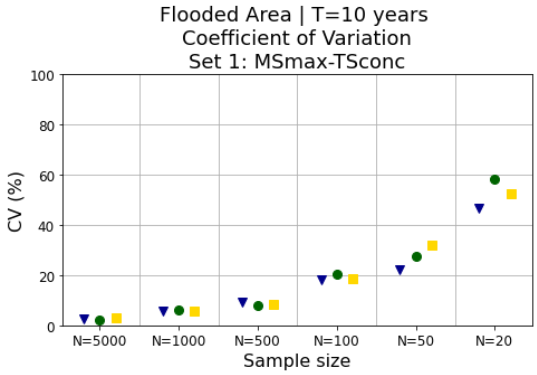
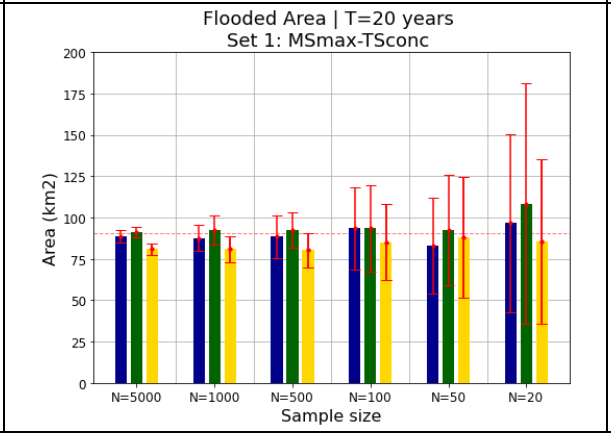
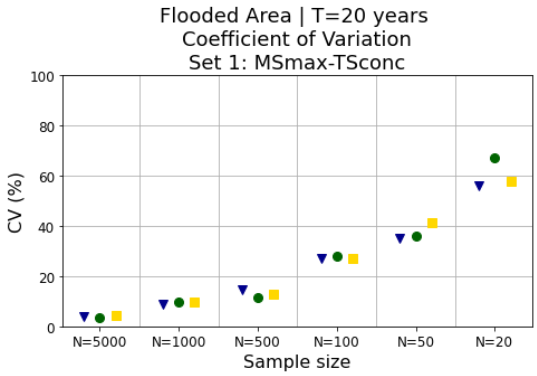
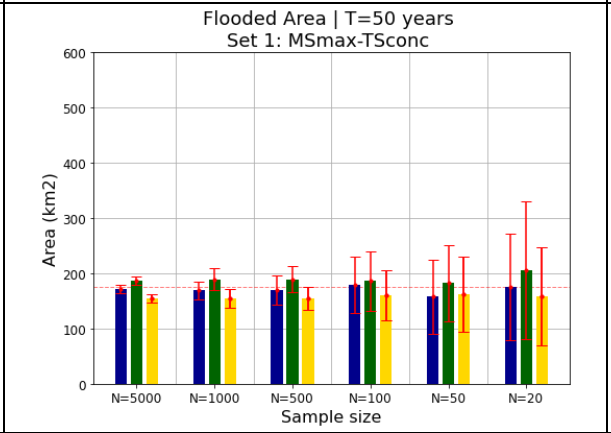
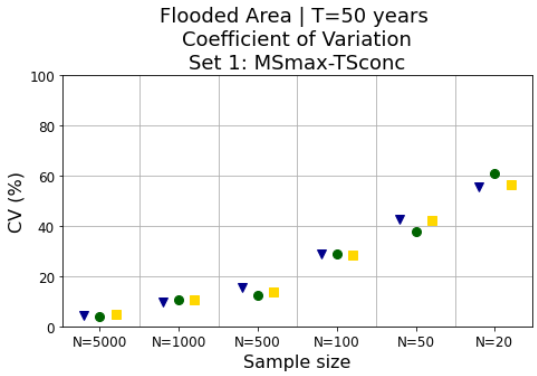
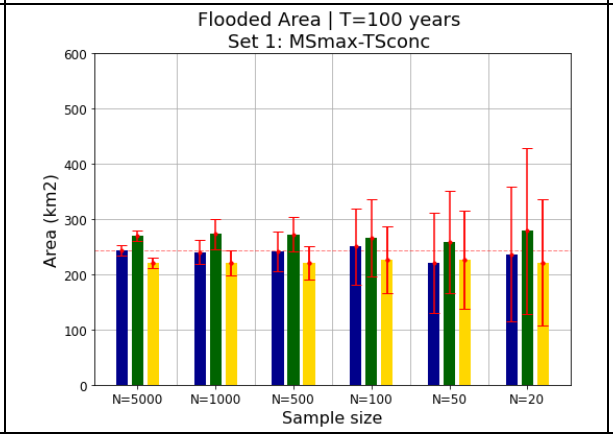
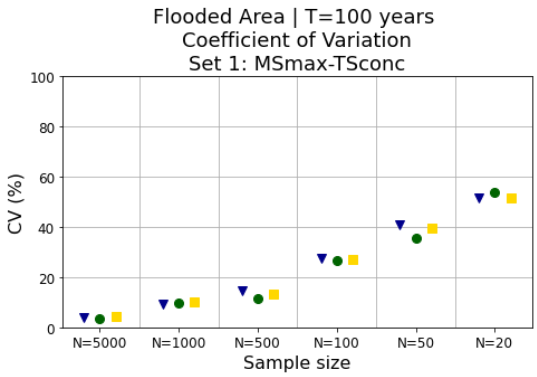
In Table 9, column 'Estimated flooded area', we can observe the estimation of the flooded area for different return periods according to (see also the symbology in Figure 50):

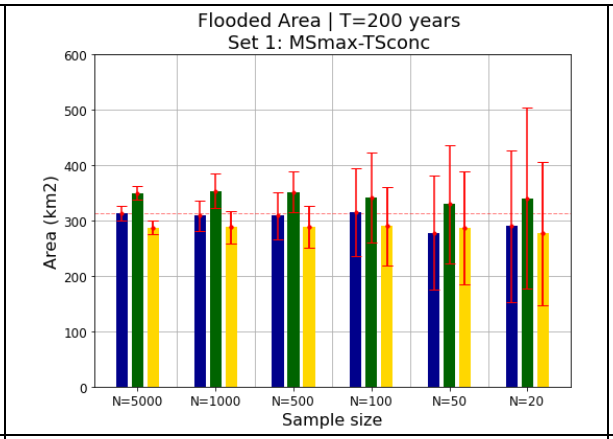
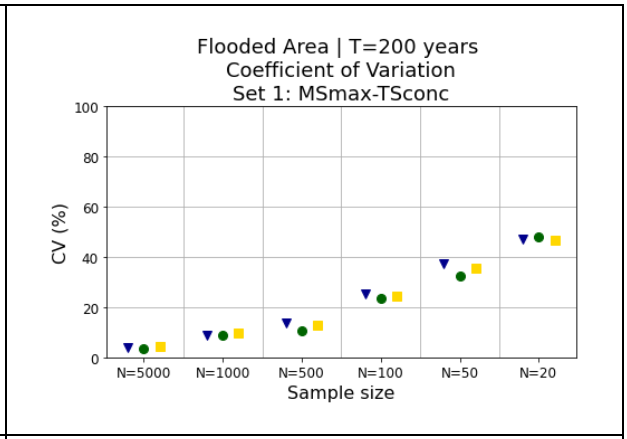
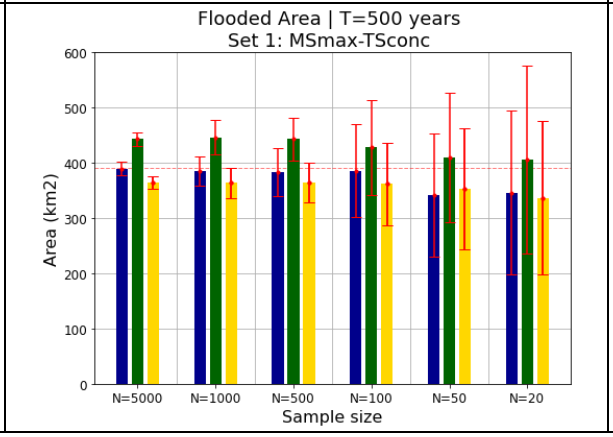
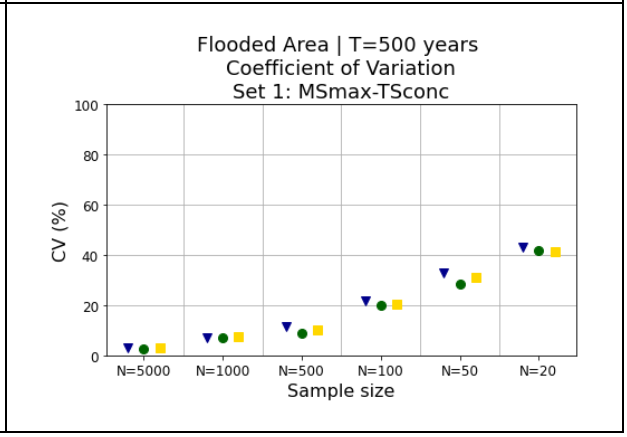
- The size of the sample N (x-axis)
- The adopted copulas: Gaussian (blue), Gumbel (green), and Clayton (yellow)
- The mean flooded area (bar) and its standard deviation (red solid line) (bootstrap experiment)
- The benchmark (red dashed line)

Moreover, in column 'Coefficient of variation', we can observe the coefficient of variation (CV, y-axis) for each of the copulas (Gaussian: blue, Gumbel: green, and Clayton: yellow) and sample sizes (x-axis).

Table 9 Evaluation of the results for the different return periods (T). The flooded area that corresponds to a T return period (A_T) and its Coefficient of Variation according to the size of the sample N (x-axis) and the copulas: Gaussian (blue), Gumbel (green), and Clayton (yellow). The benchmark (red dashed line) and the standard deviation (red solid line)

T	A_T	Flooded area	Coefficient of Variation
years	km^2		
2	3		
5	28		

10	51		
20	90		
50	175		
100	244		

200	312		
500	391		

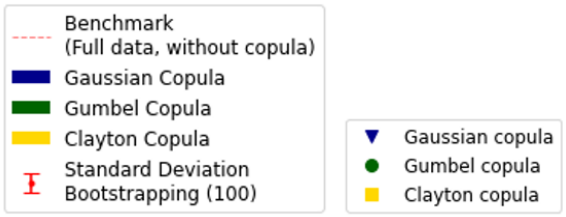


Figure 50 Symbology of the return periods graphs depicted in Table 9

From Table 9, column 'Estimated flooded area', we can evaluate first the mean flooded area estimated when using copulas. For return periods T<20 years, we do not observe a significant difference in the mean estimation of the flooded area between the various combinations of the copula that are used and the size of the sample N. However, for return periods T>20 years we observe larger differences. The Gaussian copula led to the closest estimation to the benchmark. Whereas the Gumbel copula overestimates the results, the Clayton copula underestimates them. This holds for all sample sizes.

There are no significant differences between the standard deviation and coefficient of variation between the three copulas. However, differences are significant for the different sample sizes. The larger the sample, the smaller the standard deviation and the coefficient of variation.

Additionally, we also observe some differences between the estimated return periods. Low return periods such as 2 and 5 years correspond with a large standard deviation and coefficient of variation when using small sample sizes. The standard deviation (and CV) is reduced when estimating a higher return period.

The large variation of the results when estimating a 2 or 5 years return period flooded area can be related to the physical threshold of the river system, the bankfull capacity is usually estimated to be a 2-year return period discharge.

7.2.2. Evaluation of the extreme sets

As indicated in Section 2.2.5, we differentiate between the mainstream (*MS*), the tributary stream (*TS*), and the confluence (*C*). The sets of extremes considered for the analysis are non-simultaneous and named as follows:

- *Set 1 | MSmax-TSconc: mainstream maxima (MS_{max}) with a concurrent tributary stream (TS_{conc}),*
- *Set 2 | TSmax-MSconc: tributary stream maxima (TS_{max}) with a concurrent mainstream (MS_{conc}),*
- *Set 3 | Cmax (C=MS+TS): the maxima of the sum of both streams.*

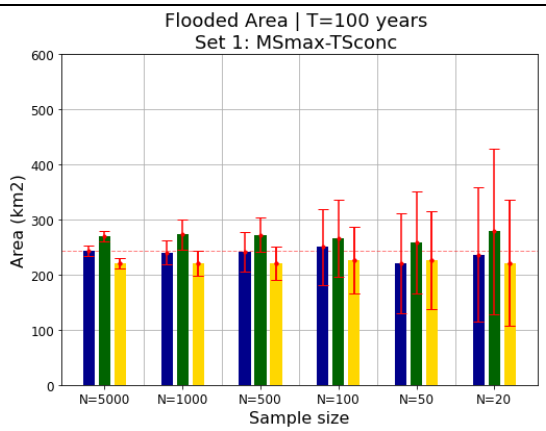
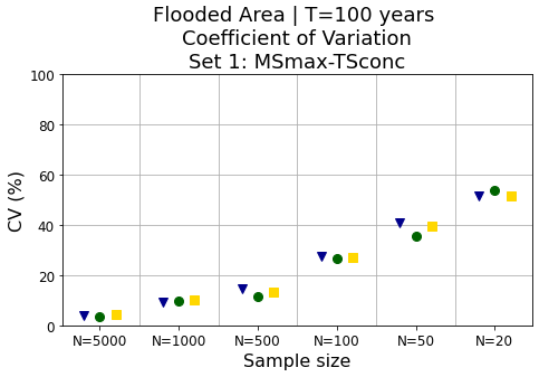
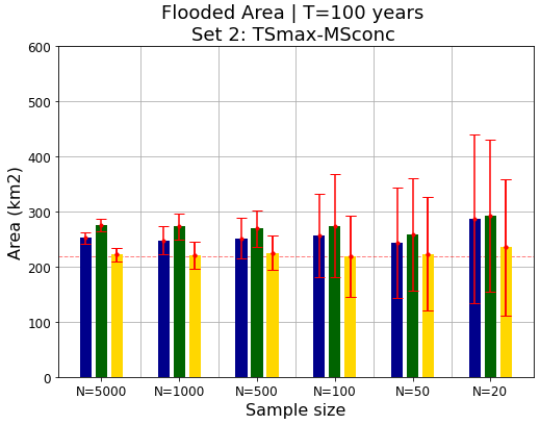
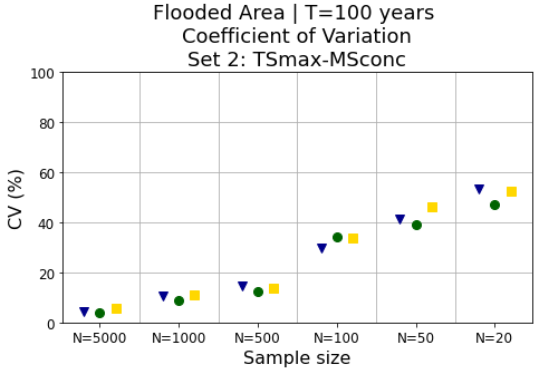
Therefore, it is important to evaluate the differences between the 3 sets when performing the hydrological sampling approach. The evaluation is done with a 100-year return period flooded area, and using all sample sizes N (20, 50, 100, 500, 1000, 5000).

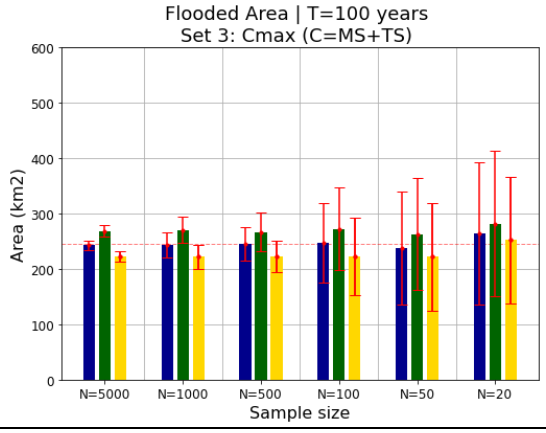
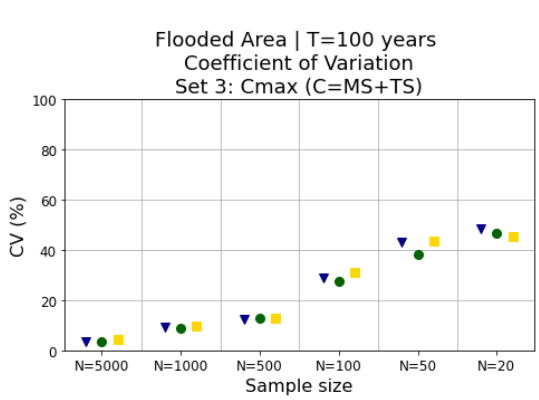
In Table 10, column 'Estimated flooded area', we can observe the estimation of the different 100 year return period flooded area for each of the sets according to (see also the symbology in Figure 50):

- The size of the sample N (x-axis)
- The used copulas: Gaussian (blue), Gumbel (green), and Clayton (yellow)
- The mean flooded area (bar) and its standard deviation (red solid line) (bootstrap experiment)
- The benchmark (red dashed line)

In column 'Coefficient of variation', we can observe the coefficient of variation (CV, y-axis) for each of the copulas (Gaussian: blue, Gumbel: green, and Clayton: yellow) and sample sizes (x-axis).

Table 10 Evaluation of the results for the different extreme sets. The flooded area that corresponds to a 100 years return period (A_{T100}) and its Coefficient of Variation according to the size of the sample N (x-axis) and the copulas: Gaussian (blue), Gumbel (green), and Clayton (yellow). The benchmark (red dashed line) and the standard deviation (red solid line)

Set	A_{T100} km^2	Flooded area	Coefficient of Variation
Set 1	243.5		
Set 2	218.1		

Set 3	244.4		

No significant differences are observed between the three sets of extremes other than the magnitude of the benchmark. The variation of the copula results and the variation due to the sample size patterns are the same as mentioned before in Section 6.2. Even though the benchmark results of Set 2 are different and lower than the benchmark results of Sets 1 and 3, the mean estimation of the flooded area is similar for the 3 sets.

7.2.3. Evaluation of the marginal distribution

When using copulas to describe the joint probability, the marginal distributions of both river (peak or concurrent) discharges are needed. As mentioned in Section 6.1, 'Parameter estimation and Joint distribution', four probability distributions were evaluated (GUM, GEV, GPD, and NOR), selecting the distributions with the best scores (RMSE, AIC, and BIC) for each set.

The log plots of the 4 distributions (see Appendix E) revealed that the end tails are not always well represented by the selected distributions according to the best score. Therefore, a sensitivity analysis is done to evaluate the influence of the selection of the marginal distribution function.

The evaluation is done for the three sets of extreme events, with a 100-year return period flooded area (A_{T100}), and using a sample size N=100. The evaluated marginal distributions are the ones that showed the best fit according to the scores (RMSE, AIC, BIC), the best fit observed in the end tail (log plots), and additionally, the empirical distribution (based on the full data set). (see Table 11).

Table 11 Marginal distribution fit of the Rhine (MS) and Main (TS) rivers for each extreme set, according to the score and the plot

	Best score		Best plot	
	Rhine (MS)	Main (TS)	Rhine (MS)	Main (TS)
Set 1 | MS_{max}-TS_{conc}	GEV	GPD	GEV	GPD
Set 2 | TS_{max}-MS_{conc}	GEV	GEV	GUM	GUM
Set 3 | C_{max} (C=MS+TS)	GEV	GPD	GEV	GUM

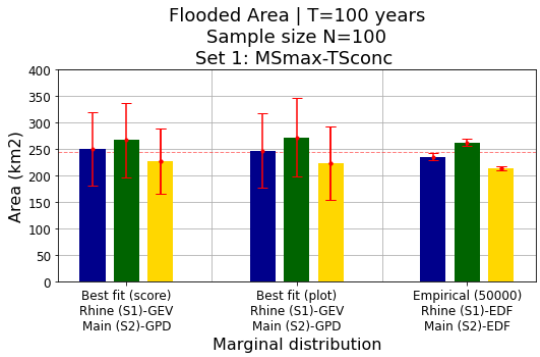
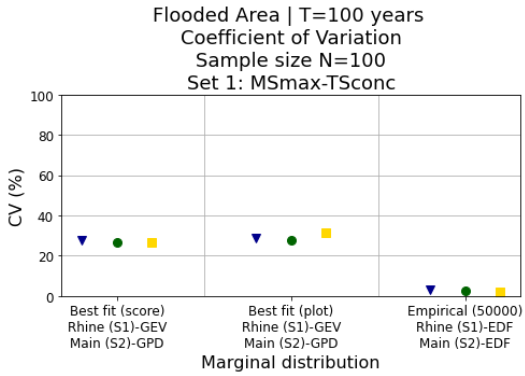
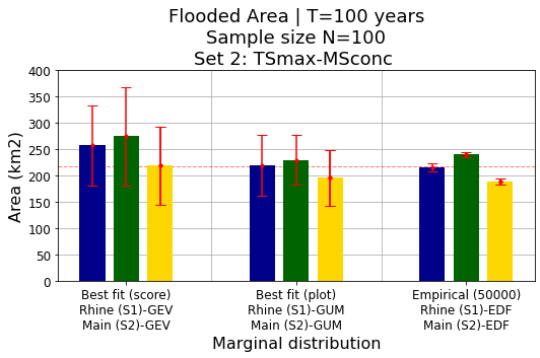
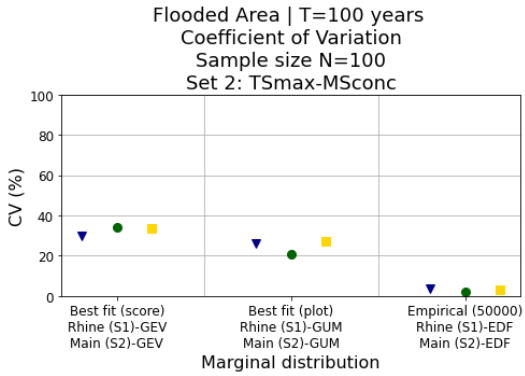
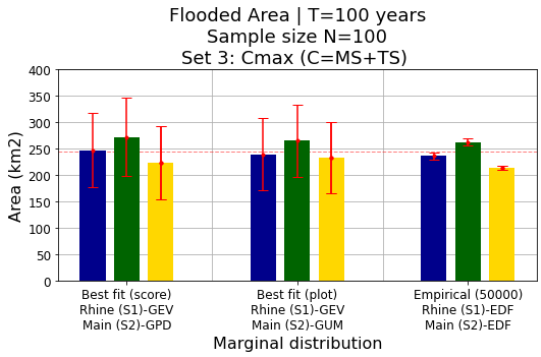
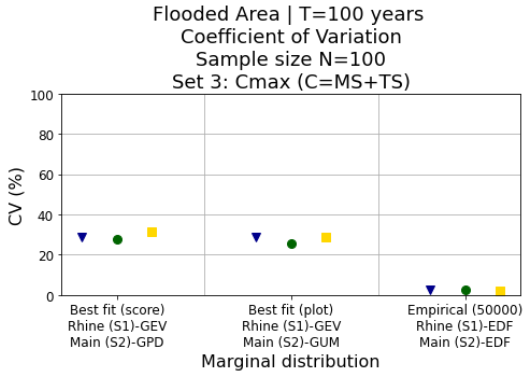
As mentioned in Section 6.1, The parameters of the selected distribution (GEV, GUM, GPD) are determined by fitting the distribution to the selected data. Again, we chose to use the L-moments method for this task out of the several available methods.

In Table 12, column 'Estimated flooded area', we can observe the estimation of the A_{T100} for each of the sets according to (see also the symbology in Figure 50):

- The marginal distributions (x-axis)
- The used copulas: Gaussian (blue), Gumbel (green), and Clayton (yellow)
- The mean flooded area (bar) and its standard deviation (red solid line) (bootstrap experiment)
- The benchmark (red dashed line)

Column ‘Coefficient of variation’ shows the coefficient of variation (CV, y-axis) for each of the copulas (Gaussian: blue, Gumbel: green, and Clayton: yellow) and marginal distributions (x-axis).

Table 12 Evaluation of the results for the different marginal distributions. The flooded area that corresponds to a 100 years return period (A_{T100}) and its Coefficient of Variation according to the marginal distributions (x-axis) and the copulas: Gaussian (blue), Gumbel (green), and Clayton (yellow). The benchmark (red dashed line) and the standard deviation (red solid line)

Set	A_{T100} km^2	Flooded area	Coefficient of Variation
Set 1	243.5		
Set 2	218.1		
Set 3	244.4		

In general, the variation of the three copula results patterns is the same as the one observed before in Section 6.2. The Gaussian copula led to the closest estimation to the benchmark. Whereas the Gumbel copula overestimates the results, the Clayton copula underestimates them.

However, some differences exist in the mean magnitudes when using different marginal distributions. In Set 1 we observe a slightly lower mean estimation when using the empirical distributions. In Set 2 the ‘best score’ marginals led to an overestimation of the mean flooded area, whereas there is no difference observed between the ‘best plot’ and the empirical distribution. The differences in magnitudes when using the GEV distribution (‘best score’) can be expected as the fit in the distribution showed an overestimation in the end tail (See Appendix E). Whereas for Set 1 and Set 2 some differences were observed between the mean results when using different marginal distributions, for Set 3 no significant differences were observed.

For the three sets of extreme events, the empirical distribution led to lower variation in the results. For Set 2, the GUM distribution led to lower variation than when using GEV. Therefore, the variation of the results can be attributed to the marginal distribution choice. This is also corroborated by the analysis presented in Section 6.2, where we observed lower variation when we used larger samples, meaning that there is more data to fit the distribution, hence, less uncertainty.

7.2.4. Evaluation of the copulas

As mentioned in Section 2.2.4, formal and informal tests can be applied for determining the goodness of fit of the copulas. We only performed an informal test, the semi-correlation test, where we not only calculate the Pearson's correlation value for the observed data, but also the modelled data when using the three proposed copulas (Gaussian, Gumbel, and Clayton) (See Table 13).

The evaluation is done for the three sets of extreme events (Set 1, Set2, and Set3). When using the copulas, the empirical distributions (from the full data) were used as the marginal distributions. The size is the same for all samples (N=50000).

Table 13 Semi-correlation test. Pearson's correlation of the quadrants for each set of extremes and each copula

	r_s	Copula	ρ_{NW}	ρ_{NE}	ρ_{SE}	ρ_{SW}	$\Sigma \lvert\rho_{Data} - \rho_{Copula}\rvert$
Set 1 | MS_{max}-TS_{conc}	0.5612	***Full data***	***0.22***	***0.36***	***0.17***	***0.24***	
		Gaussian	0.13	0.27	0.14	0.28	**0.25**
		Gumbel	0.11	0.60	0.07	0.24	0.45
		Clayton	0.11	0.04	0.08	0.51	0.79
Set 2 | TS_{max}-MS_{conc}	0.6786	***Full data***	***0.18***	***0.45***	***0.09***	***0.47***	
		Gaussian	0.16	0.43	0.17	0.44	**0.15**
		Gumbel	0.13	0.72	0.08	0.40	0.40
		Clayton	0.14	0.08	0.08	0.70	0.65
Set 3 | C_{max} (C=MS+TS)	0.4923	***Full data***	***0.18***	***0.32***	***0.15***	***0.15***	
		Gaussian	0.13	0.22	0.13	0.23	**0.25**
		Gumbel	0.10	0.52	0.06	0.18	0.40
		Clayton	0.10	0.05	0.07	0.45	0.73

Additionally, Figure 51 contains scatter plots of the observed data and the model data (in the standard normal space), for Set 1. (See Set 2, and Set 3 in Appendix F)

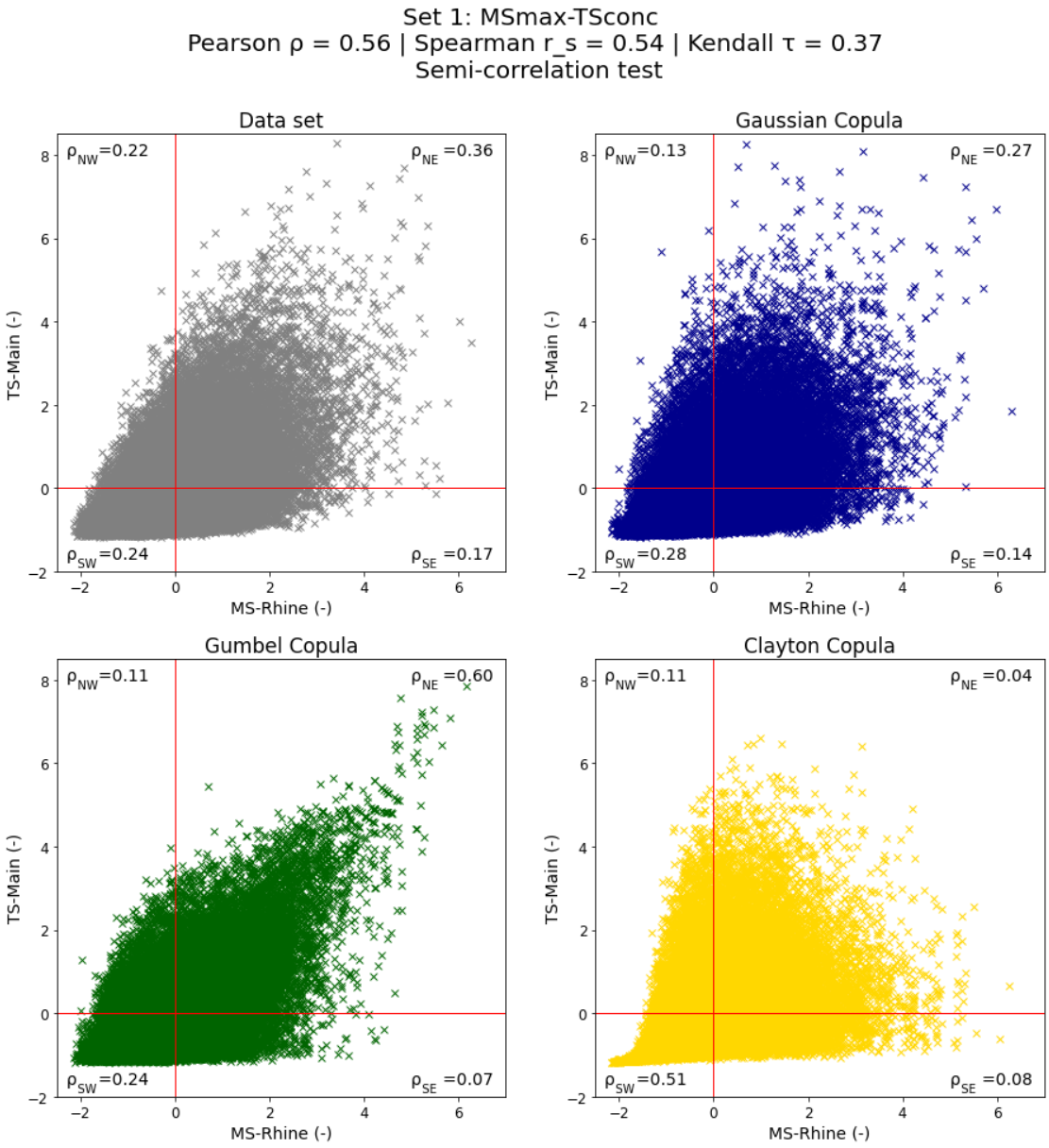


Figure 51 Semi-correlation test of Set 1. Scatter plots of the pairs of data: Mainstream (MS-Rhine) on the x-axis and Tributary stream (TS-Main) on the y-axis. Synthetic data (grey), and the simulated data using the copulas: Gaussian (blue), Gumbel (green), and Clayton (yellow). All of them are in the standard normal space

From Table 13 we can observe that, overall, the Gaussian copula led to less difference in the quadrant Pearson's correlation compare to the full data set. Moreover, from Figure 51, we can also observe that the distribution of the modelled variables that most closely resembles the full data set is the Gaussian copula.

8 Discussion

In this section, we first discuss the data that is used, followed by a discussion on the selection of the confluence and its hydraulic model that was used for answering research questions 2 and 3, and finally, the use of copulas on the sampling approach proposed for the research question 3.

8.1. Data: the 50,000 years of synthetic time series

As described in Section 3.2, in this study we used 50,000 years of meteorological and hydrological data generated by a combined stochastic-physically based model GRADE (Generator of RAinfall and Discharge Extremes). Two of the GRADE components' outputs were used in Section 4 to answer research question 1. First, the precipitation and temperature time series of the 134 sub-basins from the stochastic weather generator for the meteorological model domain evaluation, and second, the discharge time series of the 148 sub-basins from the hydrological HBV model for the hydrological model domain evaluation. The second was also used at a specific location (the Rhine and Main River confluence) in Sections 5 and 6 to answer research questions 2 and 3 respectively.

The limitations of the data and its application are important to mention:

First, the weather generator uses observed daily rainfall and temperature data for the 56 years (1951-2006) to generate the 50,000 years of synthetic data. Due to the limited length of observations, the multi-day rainfall extreme and simulated discharge values contain considerable uncertainty. Moreover, there are indications for long-range temporal dependence in the Rhine basin (Buishand 2004 as cited in Hegnauer et al., 2014), whereas the derived autocorrelation from the GRADE data indicated short-range temporal dependence. Finally, for both precipitation and temperature, the dependence between the spatial patterns on successive days is underestimated by the model (Beersma, 2011 as cited in Hegnauer et al., 2014).

Second, the GLUE analysis of the HBV model uses discharge data from the gauging station at the downstream end of the subbasins to carry out the calibration. Whereas the GLUE analysis of the HBV model for the Rhine basin downstream of Basel showed good results for most of the HBV sub-basins, reproducing observed discharges upstream of Basel turned out to be difficult, for some sub-basins the results of the GLUE analysis are poor, leading to relatively large uncertainty in the modelled results (Hegnauer et al., 2014).

The multi-day rainfall extreme and simulated discharge values uncertainties and the temporal and spatial dependencies uncertainties may affect the analysis. Whereas the multi-day rainfall extreme and simulated discharge values uncertainties and the temporal uncertainty can lead to an under- or overestimation of the extreme event, the spatial uncertainty can lead to changes in the statistical dependencies calculated between extreme events of neighbouring catchments evaluated at each confluence. Additionally, the poor results of the GLUE analysis of the hydrological data cause uncertainty in the rank correlations calculated for the Alpine and High Rhine regions.

An analysis of the statistical dependencies between extreme events of neighbouring catchments at each confluence, using the actual records instead of the synthetic data, was not done to corroborate the results.

Additionally, when evaluating daily events there is no time lag considered, and when evaluating multi-day events, we considered 5 days to account for the travel time of the flood wave, however, the 5 days criterion holds for the travel time between Basel and Lobith (Hegnauer et al., 2014).

Finally, for the hydrological and meteorological analysis carried out in Section 4, we did not carry out an analysis to identify the dry season of the 74 evaluated confluences due to time limitations. Therefore, the 1st of January is considered as the start of the block maxima.

8.2. Selection and analysis of the Confluence 68

Due to time limitations, we selected only one of the 7 proposed locations: Confluence 68, located in the outlet of the Upper Rhine region, where the Main River joins the Rhine River. In that confluence, we performed the hydraulic evaluation and the hydrological sampling evaluation of sections 5 and 6 respectively. Ideally, more confluences would have been closely evaluated, following the same procedure so we could evaluate 1) if the conclusions hold for a confluence with similar characteristics, and 2) different hydraulic interactions and the sensitivity of the hydrological sampling approach when evaluating a confluence with different characteristics.

The hydraulic evaluation of confluence 68 considers some simplifications (hence, some limitations) such as the resolution of the model, the number of simulations, the considerations for the input time series, and the calibration of the model. Which are described below.

First, the resolution of the model was set to be coarse (200m x200 m) to reduce the computational time of the simulations. This means that the results of each simulation are coarse, and therefore the response function that was obtained from all of the simulations may be under- or overestimated. Besides, to obtain the response function, only 25 combinations of the Rhine and Main River discharges (5x5 grid) were simulated. Increasing the number of simulations would allow obtaining a more precise response function.

Additionally, the 25 simulations were done considering a time difference between the peaks of 1 day. For combinations of high-and-low discharges (e.g. S4-S0 see Figure 31), there is no time difference between the simulated discharges because the extreme events that we are considering correspond to the annual maxima of one of the rivers and the concurrent flow of the other. When simulating high discharges of both rivers, it becomes more relevant to consider the time difference between the peaks, as it is likely to be the annual maxima for both rivers. Moreover, the analysis performed in Section 5.2.2 showed that the most common range of time difference is between 0 and 3 days (35% of the events). We are considering 1 day as the time lag between the two peaks and we did not perform a sensitivity analysis of the response function when using different lag times between the peaks.

Finally, the calibration of the model was done by adjusting the rivers' depth to ensure that there is no flooded area when having a discharge that corresponds to a return period of 2 years for each river. The Rhine River close to the confluence is ~350m wide, and the Main ~150m wide, however, the width of both rivers is fixed to be 100m in the model. Therefore the resulting depth of the calibration is overestimated. This leads to uncertainties of the backwater level effects, and may also lead to uncertainties about the flooding patterns.

8.3. The use of copulas in the hydrological sampling approach

As mentioned in Section 7.2.4 we only performed an informal test to measure the goodness of fit of the implemented copulas, the semi-correlation test. Performing a formal test would allow having a metric to determine the statistical significance of the modelled data, complementing the informal test.

In the semi-correlation test, we not only calculated Pearson's correlation value for the synthetic data but also for the modelled data when using the copulas. The correlations calculated for the modelled data were done by samples of the same size as the synthetic data (50,000). As mentioned before, common records of data contain 20, 50, or even 100 years, so it is of interest to also evaluate the copulas using a shorter sample.

Only three copulas were implemented out of the wide range of existing copulas. Even though the Gaussian copula resulted in the 'best' fit for the evaluated confluence, it is of interest to evaluate other copulas. Additionally, the sensitivity analysis of the marginal distributions, performed in Section 7.2.3, indicated that the variations in the results are introduced by the selection of the probability distribution.

The benchmark was obtained from the full synthetic data, the same data that was used to take the random samples when evaluating the different copulas and sample sizes. For larger samples, the variation of the results is lower, and the mean discharge is close to the benchmark. This can be expected as the larger the sample, the more similar it is to the full data.

It would be of interest to compare the results of the proposed approach not only with the benchmark, but also to the results that are obtained when 1) using a deterministic approach without considering the statistical dependencies of the variables, and 2) using another multivariate analysis that does not involve the use of copulas.

9 Conclusions and recommendations

In this section, we first repeat the conclusions of the three research questions given in Sections 4.5, 5.4, and 6.3, and second, we give some recommendations for future research.

9.1. Conclusions and answers to the research questions

In this section, we give answers to the research questions of this research, repeated below.

Research Question 1: ***How do the different catchment characteristics and the meteorological and hydrological model domains influence the correlations between joining rivers?***

In order to answer this research question, the catchment characteristics were explored to determine their influence on the statistical dependencies between neighbouring catchments of extreme events in both meteorological and hydrological domains. 74 confluences were identified and evaluated.

Different alternatives of events were evaluated. On the one hand, for the meteorological domain, three alternatives are evaluated, 1) daily and 2) multiday events, and 3) accounting for the precipitation that falls as snow. On the other hand, only one alternative for the hydrological domain is evaluated: annual maximum of daily discharges. For each confluence, the annual maximum peaks were identified out of the 50,000 years of meteorological (precipitation) and hydrological (discharge) synthetic data from GRADE (Hegnauer et al., 2014). Subsequently, the sets were explored in terms of seasonality, identifying the season where most of the peaks occur per confluence. Moreover, the Spearman rank correlation was calculated for the sets of extremes, between the mainstream peaks and the tributary stream peaks for each confluence, which is used to evaluate the statistical dependencies between neighbouring catchments.

Based on the catchment characteristics exploration and the seasonality of the peaks, we can conclude that even though precipitation occurs at any time of the year, a seasonal variation is observed, having ~30% of the annual maximum precipitation during summer in the Alpine, High and Upper regions. Hence, most of the extreme precipitation events occur during that season. Moreover, there is a shift in the annual cycle of the discharge, where higher discharges are observed during summer for the Alpine and High regions, and during winter for the Upper, Middle and Lower regions. This shift resembles the seasonality observed in the extreme discharge events, where snowmelt processes take place during spring/summer in the Alpine and High regions, and less predictable rain-dominated runoff takes place in the Upper, Middle and Lower regions.

Based on the spatial distribution of the calculated Spearman rank correlation for all the alternatives, we observed high correlations in the upstream confluences in all the Rhine regions. The catchments of the upstream confluences are small; hence, the precipitation events are more likely to be the same event for both neighbouring catchments. Moreover, similar correlations for the meteorological and hydrological evaluation are observed in the upstream confluences of the Upper, Middle and Low regions, meaning that for those confluences the extreme hydrological events are concurrent with extreme precipitation (meteorological domain).

When moving downstream and reaching the main river the correlations are lower. Furthermore, we observed higher correlations for the hydrological domain than the meteorological domain in those confluences. The catchment area of the main river is larger and therefore the precipitation events are less likely to cover the whole area in time and space, and hydrological processes play a major role in peak flow generation. The low correlations closer to the main river show that the extreme events are driven by a combination of different events from upstream areas.

The results for the Alpine and High regions showed lower correlations for the hydrological model domain, which can be related to the snow processes that take place in those regions of the catchment and/or to the lakes that are located in the area.

In addition, no differences were observed between the analysis that includes, or leaves out the precipitation that falls as snow Even though the snow process plays a major role in the catchment, the considerations taken in this analysis did not show its importance, which can be attributed to the fact that most of the extreme events occur during summer as shown in the seasonality evaluation and we are not accounting for the snow melt in this alternative. Differences in correlations between the daily and multiday meteorological events are more substantial. In upstream (and therefore small) catchments the temporal scale does not play a major role in the estimation of the statistical dependencies as daily and multiday (5 days) meteorological events show similar correlations. But for downstream (large) catchments, the temporal scale of meteorological events becomes relevant regarding correlations between rainfall in catchments of joining rivers. Downstream confluences are less likely to have a meteorological event in 1 day that covers both neighbouring catchments, leading to low correlations for daily events.

The correlation of meteorological and hydrological events is lower for larger catchments. Along with the total area, the size ratio of the mainstream catchment and the tributary stream catchment also showed a pattern: the bigger the difference in area between the neighbouring catchments, the less correlated the extreme events. This pattern is more distinct for the meteorological domain (catchment rainfall) than the hydrological domain (river discharges).

Throughout this first evaluation, we have identified the confluences where the extreme events of neighboring sub-catchments are strongly/weakly correlated, which allows us to give a first indication of where we shall perform a deeper analysis. Hence, for the confluences where the obtained correlations are weak, it is not likely to have extreme events at both of the neighboring sub-catchments, so we suggest having a closer evaluation as it may be translated into cost reduction when implementing flood measures.

Research Question 2: ***What are the hydraulic interactions between joining rivers?***

In order to answer this research question, Confluence 68, located in the outlet of the Upper Rhine region where the Main River joins the Rhine River, was selected to perform the hydraulic evaluation. A more detailed hydrological evaluation of the location was performed and the hydraulic model was set up.

From the hydrological evaluation, we defined the boundary conditions for the hydraulic model. First, September was identified as the month with the fewest extreme peaks and therefore selected as the start of the hydrological year of the block maxima. Subsequently, the minimum and maximum peak magnitudes of both rivers were identified from the data and selected as the boundaries of a 5x5 grid of discharges, so in total 25 possible combinations of Rhine and Main peak discharges were used to perform the hydraulic simulations. Finally, representative normalized hydrographs were derived to provide the input time series of the model simulations. In each simulation, 1 day time difference between the 2 river peaks was assumed.

Overall, 25 hydraulic simulations of possible discharge combinations were carried out. The results were evaluated in terms of the maximum water depth (h_{max}), identifying the differences between each simulation and the hydraulic interactions that take place at the evaluated location. Moreover, two responses function were obtained, which are calculated in terms of the flooded area and the flooded volume as a surrogate for socio-economic impacts.

We observed that the flood occurs upstream of the confluence, even in the worst scenario, we do not observe downstream areas with flooding. This is attributed to the topography of the confluence: upstream near the confluence is relatively flat and wide, but downstream (right after the confluence) the flat terrain becomes narrow. Additionally, the hydrograph of the water depth directly downstream of the confluence is predominated by the Rhine River and is always lower than directly upstream of the confluence. This difference in depth shows that the water is contained in the flood plain, reducing depths and flooded areas at the downstream end of the confluence.

Based on the evaluation of the hydraulic interaction, we observed that the influence of the Rhine River (mainstream) on the Main River (tributary stream) water depth is substantial due to the backwater effects. The backwater effects on the Main are mainly observed close to the confluence, where the magnitude of the water depth peak increases, in time and magnitude, as the magnitude of the Rhine discharge increases. Additionally, when a high discharge at the Rhine coincides with a high discharge at the Main, we observed not only backwater effects close to the confluence, but also an influence on the flooded area all along the Main River.

The Main River also influences the flood area/patterns of the Rhine River. When there is a low discharge at the Rhine, we observed an increase in the magnitude of the water depth as the Main discharge increases. Nevertheless, when the Rhine has a high discharge, its water depth does not increase with increasing Main discharge, but we still observed changes in the duration of the peak water level, and the influence on the flooded area reaches half of the simulated distance of the Rhine River.

Based on the response functions, the increasing pattern of both the area and volume responses is relatively similar. The worst scenario, as expected, occurs when both rivers have the maximum peak discharges simultaneously. The responses (flooded area and flooded volume) do not depend equally on both rivers. The Rhine River has more influence: higher discharges on the Rhine River cause more flooding than higher discharges on the Main River.

Throughout this second evaluation, we have found the different hydraulic interactions of the Rhine and Main confluence, which vary according to the combination of river discharges. If we assume fully statistical dependence (spearman's correlation equal to 1) of the mainstream and the tributary stream during extreme events, we would expect to have a narrow cloud distribution of the extreme discharge sets. Hence, the range of possible combinations of peak discharges would be reduced and so the possible hydraulic interactions, allowing to have a straighter forward process to evaluate the expected damage scenarios.

However, the obtained spearman's correlation is 0.54 (for Set 1) with a wider distribution of possible discharge combinations, leading to different hydraulic interactions and a challenge to evaluate the expected damage when a hydraulic evaluation is not performed. For the study confluence, if measures to reduce the flood damages want to be taken, they can be focused on the Rhine River as we did observe that it has more influence on the flooding.

Research Question 3: ***How should hydrological events be sampled for performing simulations in a flood risk analysis and/or determining design flood events?***

In order to answer this research question, we continued evaluating the previous selected confluence 68, located in the outlet of the Upper Rhine region where the Main River joins the Rhine River, and for which we have the response functions, flooded area and flooded volume, which are a function of both river discharges.

The annual maxima peaks were identified out of the 50,000 years of hydrological (discharge) synthetic data from GRADE. Three different ways of selecting the maxima were evaluated as mentioned in Section 2.2.5, leading to three extreme sets (Set 1 | MS_{max}-TS_{conc}, Set 2 | TS_{max}-MS_{conc}, and Set 3 | C_{max} (C=MS+TS)). From the sets, using the response function, we calculated the flooded area that can be expected for certain return periods, which were used as benchmarks.

Subsequently, random samples were selected from the full data set to resemble commonly recorded data (20, 50, and 100) and possible lengths for generating synthetic data (500, 1000, and 5000). The parameters of the copula functions, when using copulas to construct the joint probability of the joining river discharges, were calculated for each set. Then, using Monte Carlo simulations we sample the two river discharges and calculate the flooded area from the response function. After *N* Monte Carlo iterations, we constructed an empirical distribution of the flooded areas to later estimate the flooded area that corresponds to a certain probability of exceedance (or return period).

The process was repeated 100 times for each sample in a bootstrapping experiment to be able to quantify the variations in the results.

Based on the benchmarks we can conclude that for small return periods (2, 5, and 10 years) both rivers, the Rhine and the Main, have similar importance. However, when the return period is higher than 10 years, the Rhine River has more influence on the response in terms of the flooded area. Moreover, while the magnitudes of the flooded areas for the different return periods are similar for the extreme sets 1 and 3, Set 2 generally results in smaller flooded areas.

Based on the estimated 100 years return period flooded area (A_{T100}) when using different sizes of the sample and different copulas, we can conclude that the Gaussian copula led to the closest estimation to the benchmark independent of the sample *N*.

Based on the sensitivity analysis, the variation of the results is attributed not only to the size of the sample but also to the marginal distribution choice. We observed lower variation when we used larger samples, meaning that there is more data to fit the distribution, hence, less uncertainty.

Throughout this third evaluation, we have evaluated the different choices we make when sampling hydrological events to determine the design flood events. First, for the evaluated confluence, where the Main River joins the Rhine River, the Gaussian copula was the best fit for the three extreme sets. Nevertheless, for other confluences, we suggest always evaluating the selection of the copula. Second, we suggest considering non-simultaneous events for confluences with weak statistical dependencies of the mainstream and the tributary stream during extreme events. Set 1 | MS_{max}-TS_{conc} (mainstream maxima with a concurrent tributary stream) can be chosen as it is the worst scenario of the three sets of non-simultaneous events. When it is desired to reduce the marginal distribution selection uncertainty (especially with short time series), we recommend generating synthetic data.

9.2. Recommendations

Based on the discussion elaborated in Section 8, we have some recommendations for future research.

For future research on the statistical dependencies of the evaluation of the meteorological and hydrological parameters (performed in Section 4) we recommend:

- To carry out an analysis to identify the dry season of the 74 evaluated confluences, define the start of the block maxima, and evaluate the sensitivity of the statistical dependencies.
- When evaluating multi-day events, evaluate the travel time of the flood wave according to the topology of the catchment, changing the 5 days considered in this study for the actual travel time.
- To perform an analysis of the statistical dependencies between extreme events of neighbouring catchments at each confluence, but using the actual records instead of the synthetic data.

For future research on the evaluation of the hydraulic interaction (performed in Section 5) we recommend:

- To evaluate more confluences, following the same procedure so it is possible to evaluate 1) if the conclusions hold for a confluence with similar characteristics as the one evaluated in this study, and 2) the different hydraulic interactions for confluences with different characteristics.
- To build the hydraulic model using a finer resolution, and to simulate more discharges to obtain a more precise response function.
- When considering high discharges at both joining rivers, the time difference between the peaks needs to be better evaluated, performing a sensitivity analysis of the response function when using different times.
- To calibrate the model by not employing the rivers' depth but the rivers' width, having a more realistic representation of the river systems and better results for the backwater effects evaluation.

Finally, for future research on the use of copulas and the risk evaluation (performed in Section 6) we recommend:

- To evaluate more confluence following the same procedure so it is possible to evaluate 1) if the conclusions hold for a confluence with similar characteristics as the one evaluated in this study, and 2) the sensitivity of the hydrological sampling approach when evaluating confluences with different characteristics.
- To evaluate the use of more copulas. Performing a formal goodness-of-fit test, and, when evaluating the modelled data, not only evaluate a sample as large as the synthetic data but shorter samples that resemble the common length of observations.
- To compare the results of the proposed approach not only with the benchmark, but also to the results that are obtained when 1) using a deterministic approach without considering the statistical dependencies of the variables, and 2) using another multivariate analysis that does not involve the use of copulas.

Bibliography

Akaike, H. (1998). Information Theory and an Extension of the Maximum Likelihood Principle. In E. Parzen, K. Tanabe, & G. Kitagawa (Eds.), Selected Papers of Hirotugu Akaike (pp. 199–213). Springer. https://doi.org/10.1007/978-1-4612-1694-0_15

Bačová Mitková, V. (2020). Bivariate joint probability analysis of flood hazard at river confluence. Acta Hydrologica Slovaca, 21(1), 29–38. https://doi.org/10.31577/ahs-2020-0021.01.0004

Baumgartner, A., Reichel, E., and Weber, G. (1983). Der Wasserhaushalt der Alpen. Oldenbourg Verlag, Munchen.

Beersma, J.J., (2011). Rainfall generator for the Rhine basin: Sensitivity to the composition of the feature vector and passive simulations. KNMI publication 186-VI, KNMI, De Bilt, The Netherlands.

Bender, J., Wahl, T., Müller, A., & Jensen, J. (2016). A multivariate design framework for river confluences. Hydrological Sciences Journal, 61(3), 471–482. https://doi.org/10.1080/02626667.2015.1052816

Bernardara, P., Mazas, F., Kergadallan, X., & Hamm, L. (2014). A two-step framework for over-threshold modelling of environmental extremes. Natural Hazards and Earth System Sciences, 14(3), 635–647. https://doi.org/10.5194/nhess-14-635-2014

Buishand, T.A. and Brandsma, T., (2001). Multi-site simulation of daily precipitation and temperature in the Rhine basin by nearest-neighbour resampling. Water Resources Research, 37, 2761 – 2776.

Candela, A., Brigandì, G., & Aronica, G. T. (2014). Estimation of synthetic flood design hydrographs using a distributed rainfall–runoff model coupled with a copula-based single storm rainfall generator. Natural Hazards and Earth System Sciences, 14(7), 1819–1833. https://doi.org/10.5194/nhess-14-1819-2014

Chang, B. (2019, May 8). Copula: A Very Short Introduction. Retrieved August 01, 2022, from: https://bochang.me/blog/posts/copula/

Chok, N. S. (2010). Pearson's versus Spearman's and Kendall's correlation coefficients for continuous data (Doctoral dissertation, University of Pittsburgh).

Couasnon, A., Sebastian, A., & Morales-Nápoles, O. (2018). A Copula-Based Bayesian Network for Modeling Compound Flood Hazard from Riverine and Coastal Interactions at the Catchment Scale: An Application to the Houston Ship Channel, Texas. Water, 10(9), 1190. https://doi.org/10.3390/w10091190

Diermanse, F.L.M., van Kappel, R.R., Ogink, H.J.M. and Heynert, K.V., (2000). Analysis of the instrumentation for computing design discharges. WL | delft hydraulics, Delft, The Netherlands.

Diermanse, F. L. M. (2005). Modelling statistical dependence using copulas. 8.

Dodangeh, E., Shahedi, K., Solaimani, K., Shiau, J.-T., & Abraham, J. (2019). Data-based bivariate uncertainty assessment of extreme rainfall-runoff using copulas: Comparison between annual maximum series (AMS) and peaks over threshold (POT). Environmental Monitoring and Assessment, 191(2), 67. https://doi.org/10.1007/s10661-019-7202-0

Deltares. (2021). Risk analysis with tributary compound flooding (p. 29).

Genest, C., & Favre, A.-C. (2007). Everything You Always Wanted to Know about Copula Modeling but Were Afraid to Ask. Journal of Hydrologic Engineering, 12(4), 347–368. https://doi.org/10.1061/(ASCE)1084-0699(2007)12:4(347)

Genest, C., Rémillard, B., & Beaudoin, D. (2009). Goodness-of-fit tests for copulas: A review and a power study. Insurance: Mathematics and Economics, 44(2), 199–213. https://doi.org/10.1016/j.insmatheco.2007.10.005

Gilja, G., Ocvirk, E., & Kuspilić, N. (2018). Joint probability analysis of flood hazard at river confluences using bivariate copulas. Gradevinar, 70(4).

Grimaldi, S., & Serinaldi, F. (2006). Asymmetric copula in multivariate flood frequency analysis. Advances in Water Resources, 29(8), 1155–1167. https://doi.org/10.1016/j.advwatres.2005.09.005

Haylock, M.R., Hofstra, N., Klein Tank, A.M.G., Klok, E.J., Jones, P.D. and New, M. (2008). A European daily high-resolution gridded data set of surface temperature and precipitation or 1950-2006. Journal of Geophysical Research, 113, D20119, doi: 10.1029/2008JD010201.

Hegnauer, M., Beersma, J. J., Van den Boogaard, H. F. P., Buishand, T. A., & Passchier, R. H. (2014). Generator of Rainfall and Discharge Extremes (GRADE) for the Rhine and Meuse basins. Final report of GRADE 2.0. Project Under Commission of the Ministry of Infrastructure and the Environment: Rijkswaterstaat Water, Transport and Environment (RWS-WVL). Deltares, Delft, The Netherlands.

Hosking, J. R. M., & Wallis, J. R. (1997). Regional Frequency Analysis: An Approach Based on L-Moments. Cambridge University Press. https://doi.org/10.1017/CBO9780511529443

IKHR (1999). Eine Hochwasserperiode in Rheingebiet. Extremereignisse zwischen Dezember 1993 and Februar 1995, International Kommission für die Hydrologie des Rheingebietes, Koblenz.

Joe, H. (2014). Dependence Modeling with Copulas. CRC Press.

Jonkman, S. N., Steenbergen, R. D. J. M., Morales-Nápoles, O., Vrouwenvelder, A. C. W. M., & Vrijling, J. K. (2015). Probabilistic design: risk and reliability analysis in civil engineering. Collegedictaat CIE4130.

Jonkman, S. N., Jorissen, R. E., Schweckendiek, T., & van de Bos, J. P. (2021). Flood Defences. Lecture notes CIE5314 4th edition.

Leander, R., and Buishand, T.A., (2004). Rainfall generator for the Meuse basin: Development of a multi-site extension for the entire drainage area. KNMI publication 196-III, KNMI, De Bilt, The Netherlands.

Leijnse, T. (2018) Introduction — SFINCS 1.0 documentation Retrieved July 17, 2022, from https://sfincs.readthedocs.io/en/latest/overview.html

Lindström, G., Johansson, B., Persson, M., Gardelin, M. and Bergström, S., (1997). Development and test of the distributed HBV-96 hydrological model. Journal of Hydrology, 201, 272-288.

Moran, M. (2022). Correlation (Pearson, Kendall, Spearman). Statistics Solutions. Retrieved August 05, 2022, from https://www.statisticssolutions.com/free-resources/directory-of-statistical-analyses/correlation-pearson-kendall-spearman/

Pham, H. (2019). A New Criterion for Model Selection. Mathematics, 7(12), 1215. https://doi.org/10.3390/math7121215

Pickands III, J. (1975). Statistical inference using extreme order statistics. the Annals of Statistics, 119-131.

Pitt, M. A., & Myung, I. J. (2002). When a good fit can be bad. Trends in cognitive sciences, 6(10), 421-425.

Pobočíková, I., Sedliačková, Z., & Michalková, M. (2017). Application of Four Probability Distributions for Wind Speed Modeling. Procedia Engineering, 192, 713–718. https://doi.org/10.1016/j.proeng.2017.06.123

Rauthe, M., Steiner, H., Riediger, U., Mazurkiewicz, A. and Gratzki, A., (2013). A central European precipitation climatology – Part I: Generation and validation of a high resolution gridded daily data set (HYRAS). Meteorologische Zeitschrift, 22, 235-256.

Schwarz, G. (1978). Estimating the Dimension of a Model. The Annals of Statistics, 6(2), 461–464.

Tosunoglu, F., Gürbüz, F., & İspirli, M. N. (2020). Multivariate modeling of flood characteristics using Vine copulas. Environmental Earth Sciences, 79(19), 459. https://doi.org/10.1007/s12665-020-09199-6

Uehlinger, U., Arndt, H., Wantzen, K. M., & Leuven, R. S. E. W. (2009). The Rhine River Basin. In Rivers of Europe (pp. 199–245). Elsevier. https://doi.org/10.1016/B978-0-12-369449-2.00006-0

Wang, C. (2007). A joint probability approach for the confluence flood frequency analysis. Iowa State University.

Wang, C., Chang, N.-B., & Yeh, G.-T. (2009). Copula-based flood frequency (COFF) analysis at the confluences of river systems. Hydrological Processes, 23(10), 1471–1486. https://doi.org/10.1002/hyp.7273

Zhou, T., Liu, Z., Jin, J., & Hu, H. (2019). Assessing the Impacts of Univariate and Bivariate Flood Frequency Approaches to Flood Risk Accounting for Reservoir Operation. Water, 11(3), 475. https://doi.org/10.3390/w11030475

Zorzetto, E., Botter, G., & Marani, M. (2016). On the emergence of rainfall extremes from ordinary events: EXTREMES EMERGE FROM ORDINARY EVENTS. Geophysical Research Letters, 43(15), 8076–8082. https://doi.org/10.1002/2016GL069445

Appendix A – Meteorological Event Alternatives

As mentioned in Section 4.2, the meteorological data consists of 50,000 years of simulated daily precipitation rate and temperature per catchment. For larger basins, consisting of two or more sub-catchments, the weighted mean precipitation over all sub-catchments was calculated according to the topology of the catchment. The weights were taken proportional to the size of the sub-catchment.

This was done for daily and multiday events. For the latter, the precipitation was accumulated over 5 days to account for the travel time of the flood wave between Basel and Lobith (Hegnauer et al., 2014).

In some instances, precipitation falls as snow during the winter period. Therefore, a third alternative is analysed to account for this process. In this case, when the temperature is below 0 °C (see Figure A.1) the precipitation is assumed to be 0 (see Figure A.2). Therefore, as a third alternative, we analyzed the multiday event (5-days) considering the precipitation that falls as snow (see Figure A.3)

Some limitations should be noted. Snowmelt is not considered for the meteorological domain analysis. We only subtracted the precipitation that falls as snow, but it is not added later when the temperature increases and the snow melts. The snowmelt processes are included, though, in the hydrological analysis.

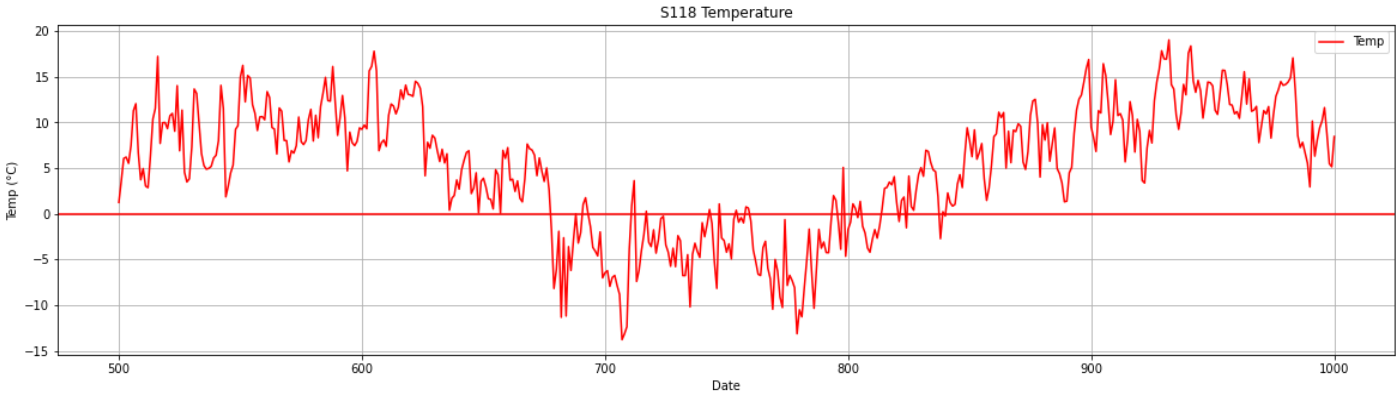


Figure A1. Temperature time series

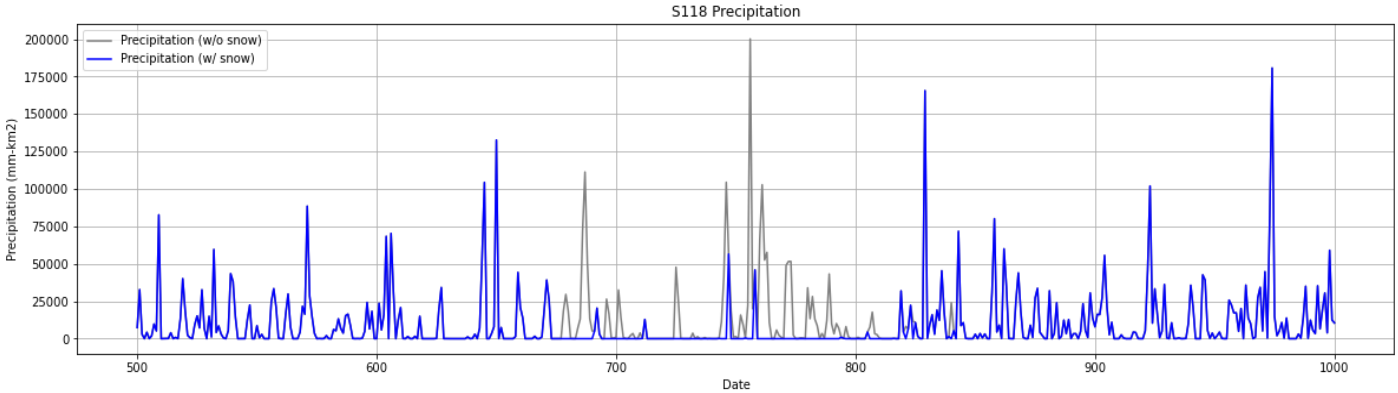


Figure A2. Precipitation time series. Original data (grey) and considering precipitation that falls as snow (blue)

In Figure A2, the grey colour corresponds to the original precipitation data. The blue colour represents the original data for days where the temperature is higher than 0, but when the temperature is lower, the precipitation is considered to be 0.

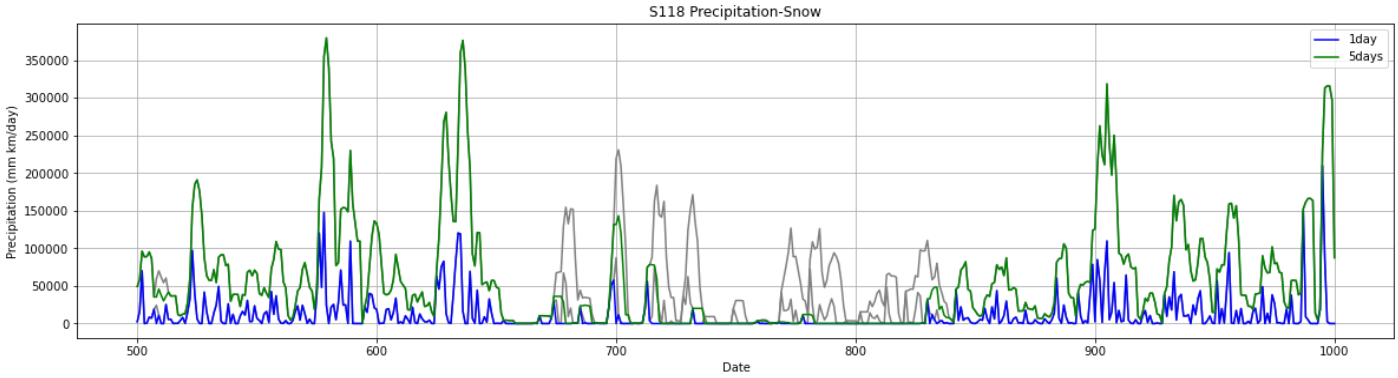


Figure A3. Precipitation time series, daily (blue) and multiday (green) alternatives

In Figure A3, the grey line corresponds to the precipitation data, the blue line represents the daily rainfall (so snow excluded), and the green line, the 5-days accumulated rainfall.

Appendix B – Potential locations to perform the hydraulic simulations

In order to select a location to evaluate the hydraulic interactions, 7 locations were pre-selected. The preselection was done as follows:

For the 7 preselected locations, the area near the flood plain is relatively flat (see Figure B2 and Figures B5 to B12). Moreover, they are located in 5 different regions: Alpine (1), High (1), Upper (3), Middle (1), and Low (1). If we can simulate all of them, we could evaluate the differences concerning the location and the different hydrological processes involved in each region.

Moreover, the 4 main tributaries of the Rhine River are considered: Aare in the High Rhine, Neckar and Main in the Upper Rhine, and Moselle in Middle Rhine. Additionally, a confluence in the upper area of the Main River was also preselected, to evaluate the differences between the upstream confluence and the very downstream confluence (when the Main joins the Rhine).

Catchment characteristics such as the size of the whole contributing area and the size ratio of the sub-catchments were also taken into consideration (see Figure B1). Preselecting from small to large contributing catchments, and locations with different size ratios.

As indicated in Section 2.2.5, we differentiate between the mainstream (*MS*), the tributary stream (*TS*), and the confluence (C) as shown in Figure 2. The sets of extremes considered for the analysis were non-simultaneous (Set 1 | MS_{max}-TS_{conc}, Set 2 | TS_{max}-MS_{conc}, and Set 3 | C_{max} (C=MS+TS)). Therefore, we also looked into the 3 sets of extremes for the 7 preselected locations, observing differences/similitudes between the sets for each location. In that way, the influence of the way of selecting the maxima could also be evaluated.

We also looked at locations where the meteorological and hydrological evaluation revealed not that high Spearman rank correlation (see Figure B3 and Figure B4). Thus, they would be relevant when performing the copula evaluation in Section 7.

From the 7 potential locations, the final selection was done according to the topography. The selected confluence was Confluence 68, located in the outlet of the Upper Rhine region. At the confluence, the Main River, one of the main tributaries, joins the Rhine River. In Figure B2 we can see that the flood plain is relatively flat and wide for both joining rivers in this confluence, which means we can expect hydraulic interactions to be substantial.

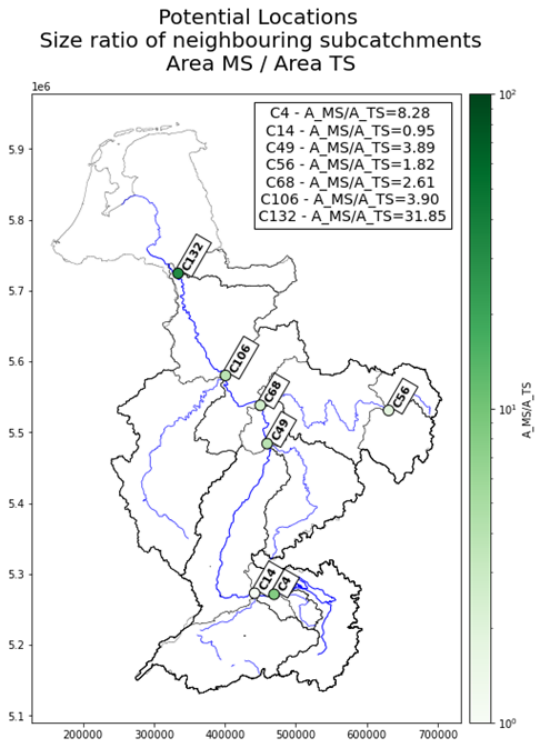


Figure B1. Potential locations to perform the hydraulic model: Size ratio of the contributing catchments of each confluence. (Mainstream area (A_{MS}) divided by the tributary stream area (A_{TS}))

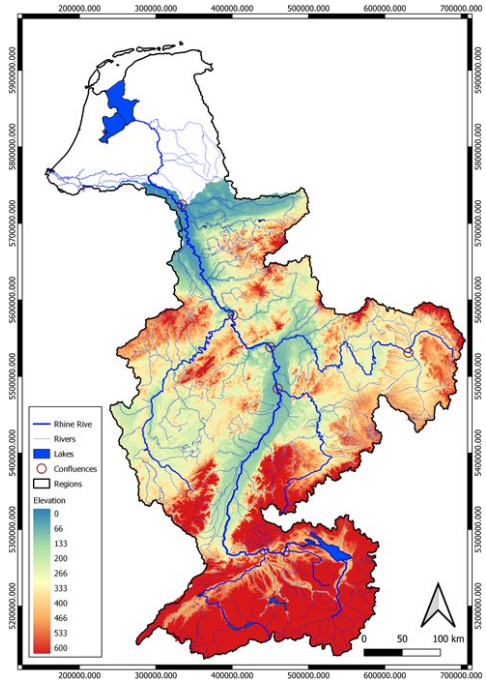


Figure B2. Digital Elevation Model (DEM) of the Rhine River Basin. Data retrieved from NASA's EarthData

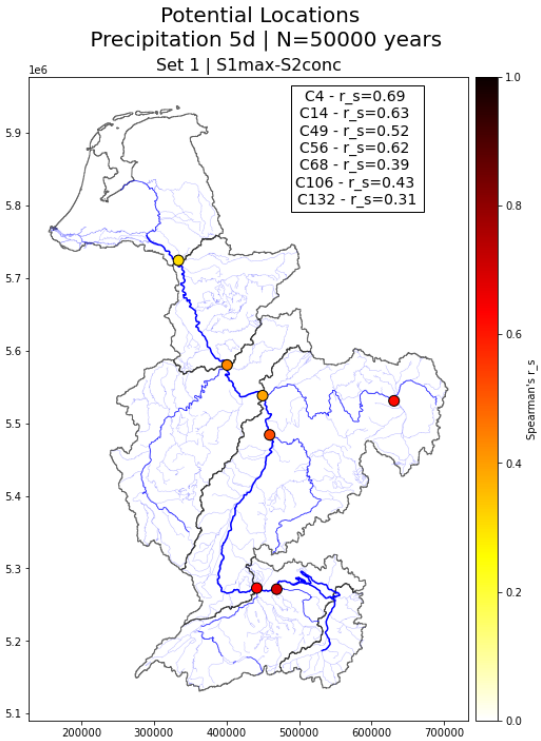


Figure B3. Potential locations to perform the hydraulic model: Spearman's rank correlation between extreme events (meteorological: precipitation 5 days) of neighbouring catchments at each confluence

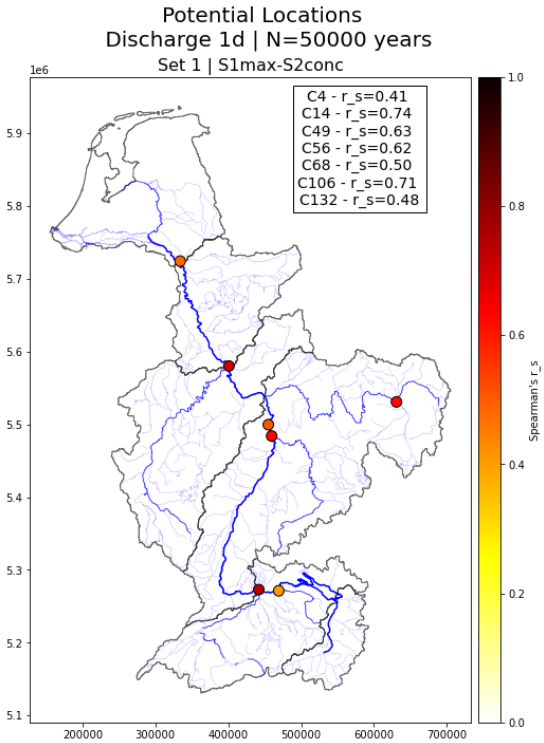


Figure B4. Potential locations to perform the hydraulic model: Spearman's rank correlation between extreme events (hydrological: discharge 1 day) of neighbouring catchments at each confluence

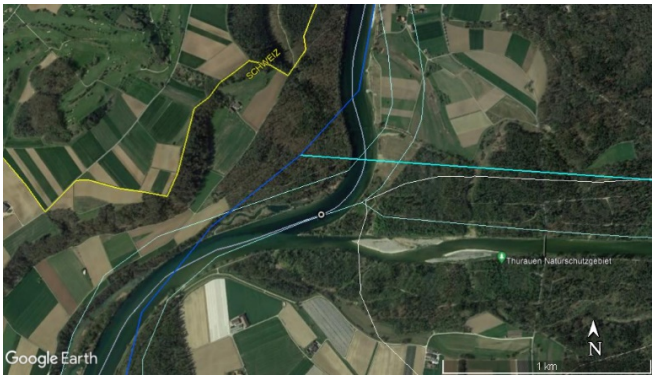


Figure B5. Confluence 4 – Thur and Lake Constance.
Satellite view retrieved from Google Earth

Figure B6. Confluence 14 – Aare.
Satellite view retrieved from Google Earth

Figure B7. Confluence 49 – Neckar.
Satellite view retrieved from Google Earth

Figure B8. Confluence 56 – Upstream Main.
Satellite view retrieved from Google Earth

Figure B9. Confluence 68 – Main.
Satellite view retrieved from Google Earth

Figure B10. Confluence 106 – Moselle.
Satellite view retrieved from Google Earth

Figure B11. Confluence 132 – The most downstream confluence.
Satellite view retrieved from Google Earth

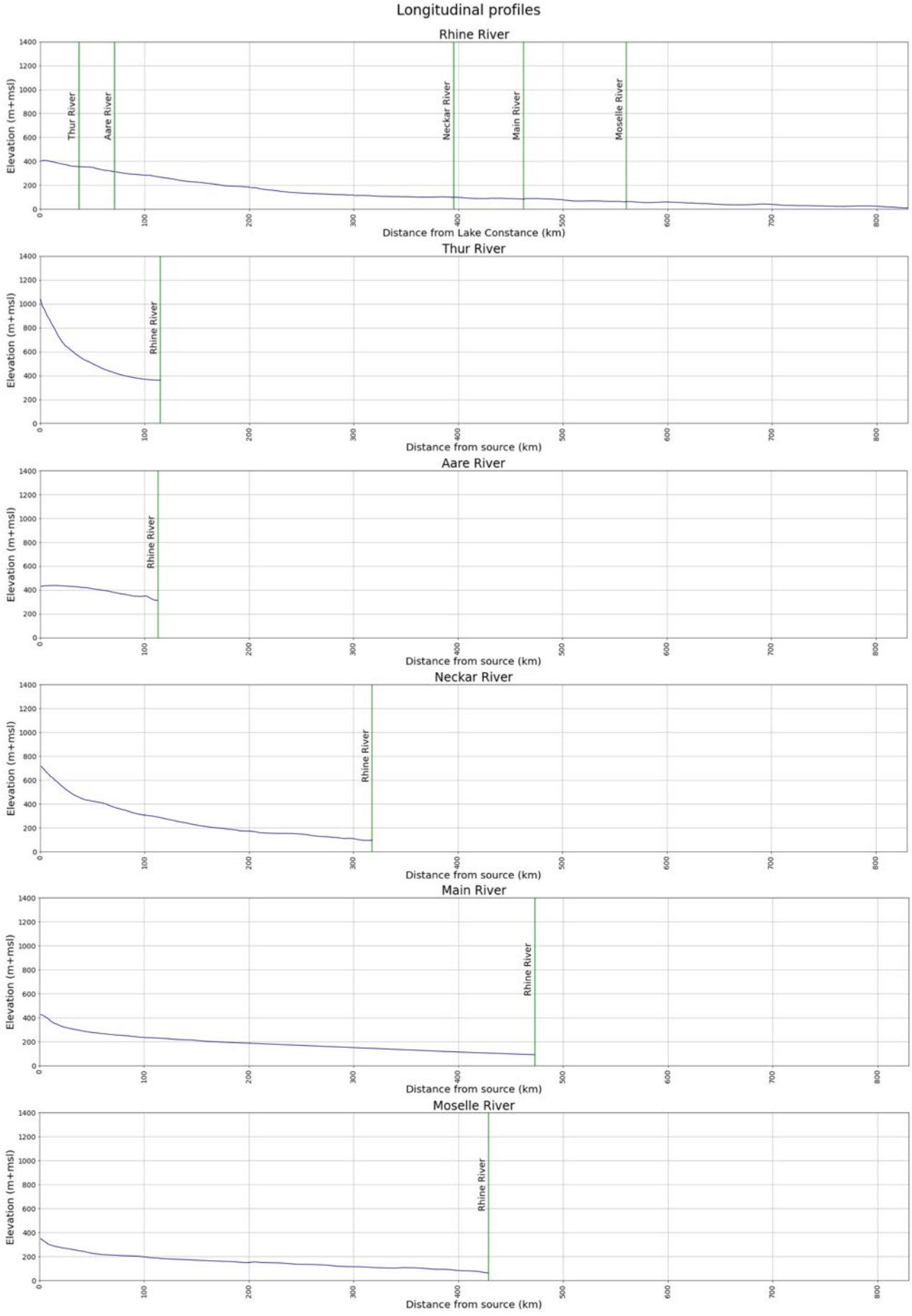


Figure B12. Longitudinal profiles of the potential locations.

Appendix C – Normalized hydrographs (C68)

As mentioned in Section 5.1, discharge time series are required as an input for the hydraulic model (SFINCS) for each of the rivers that join. Thus, a hydrograph is needed to provide the time series of each peak. To account for the differences in the response of the flood behaviour in terms of time according to the peak magnitude, we classify the discharge's magnitude into 5 ranges of discharge (see Figure C1).

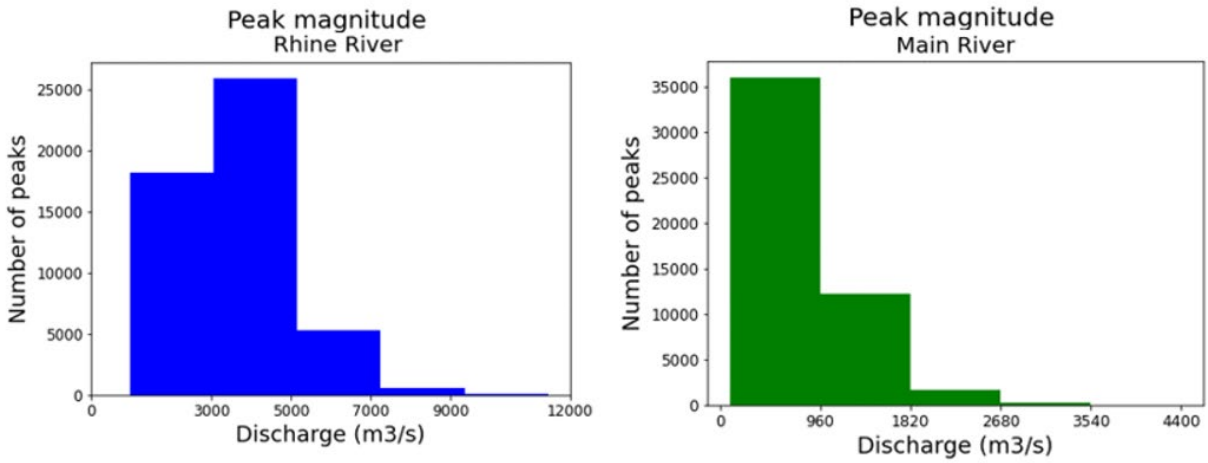


Figure C1. Extreme peak histogram of the Rhine (left) and Main (right) rivers

We normalized all the peaks for each class by first identifying the time series of each peak, considering a window of 15 days previous to and after the peak and second dividing the discharges by the peak magnitude. As a result, we had N time series of 30 days with a peak magnitude equal to 1 for each class of discharges. (see Figure C2 and Figure C3)

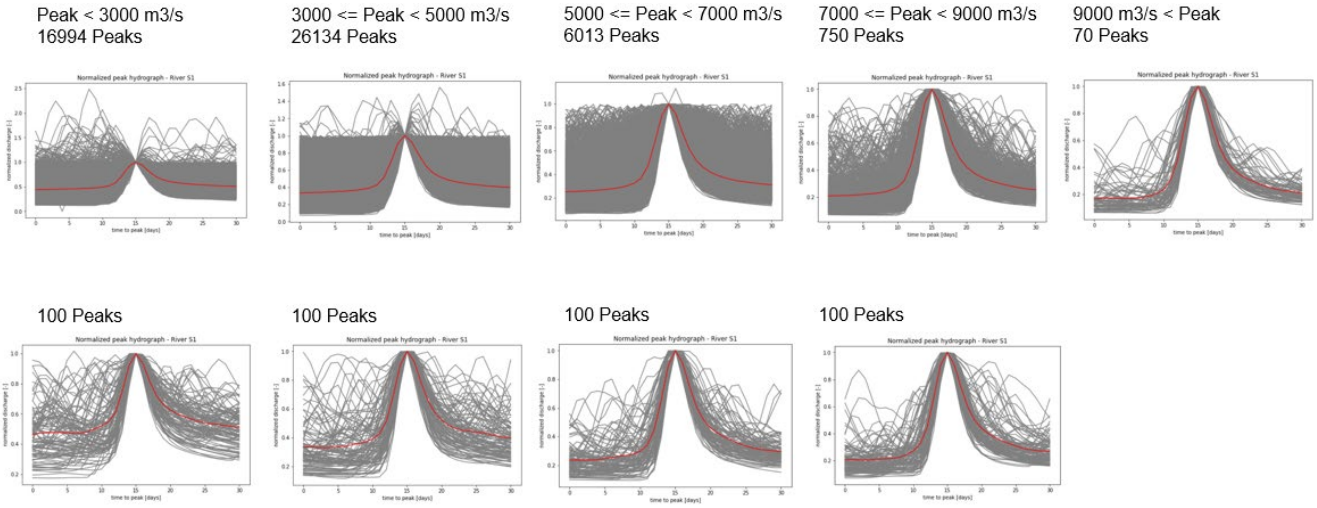


Figure C2. Rhine River normalized hydrographs (grey lines) and mean normalized hydrograph (red line) per class (or discharge ranges). Upper: Use all the peaks. Bottom: using 100 random peaks

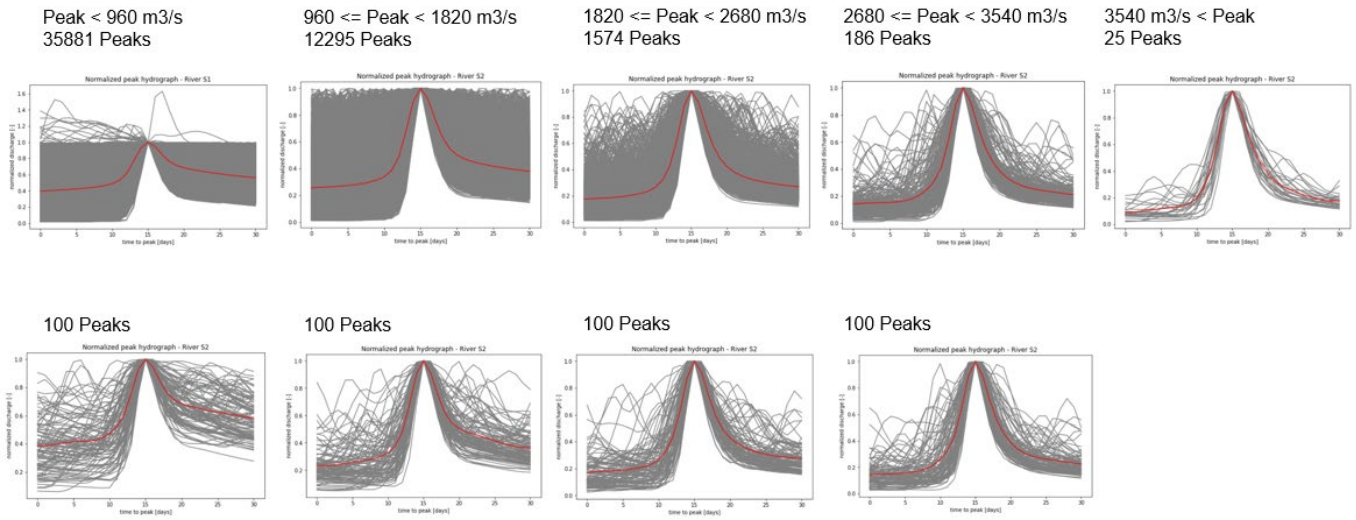


Figure C3. Main River normalized hydrographs (grey lines) and mean normalized hydrograph (red line) per class (or discharge ranges). Upper: Use all the peaks. Bottom: using 100 random peaks

Furthermore, in order to have a representative normalized hydrograph per class, we calculated the mean discharge of all the peak normalized hydrographs (see Figure C4 and Figure C5).

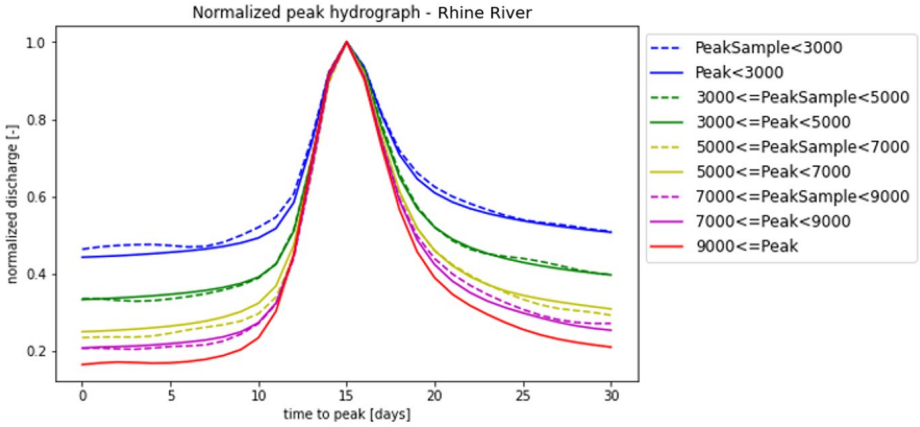


Figure C4. Rhine River mean normalized hydrograph per class. Solid line: using all peaks. Dashed line: using 100 random peaks.

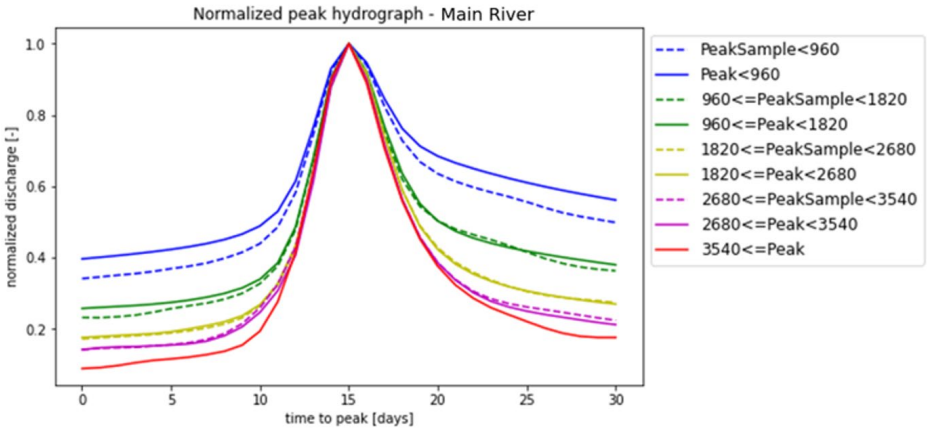


Figure C5. Main River mean normalized hydrograph per class. Solid line: using all peaks. Dashed line: using 100 random peaks.

Appendix D – Hydraulic simulations results (extended)

Analysis of the river depths

This is a preliminary analysis of the backwater effects of the hydraulic evaluation. One should keep in mind that the calibration of the model was done by adjusting the rivers' depth to ensure that there is no flooded area when having a discharge that corresponds to a return period of 2 years for each river. The Rhine River close to the confluence is ~350m wide, and the Main ~150m wide, however, the width of both rivers is fixed to be 100m in the model. Therefore the resulted depth of the calibration is overestimated. This leads to uncertainties about the backwater level effects

In Figures D2, D2, D4 and D5 we present the water levels at different locations: Rhine (upstream and at the confluence), Main (upstream and at the confluence), and at the confluence and downstream area (see Figure D1).

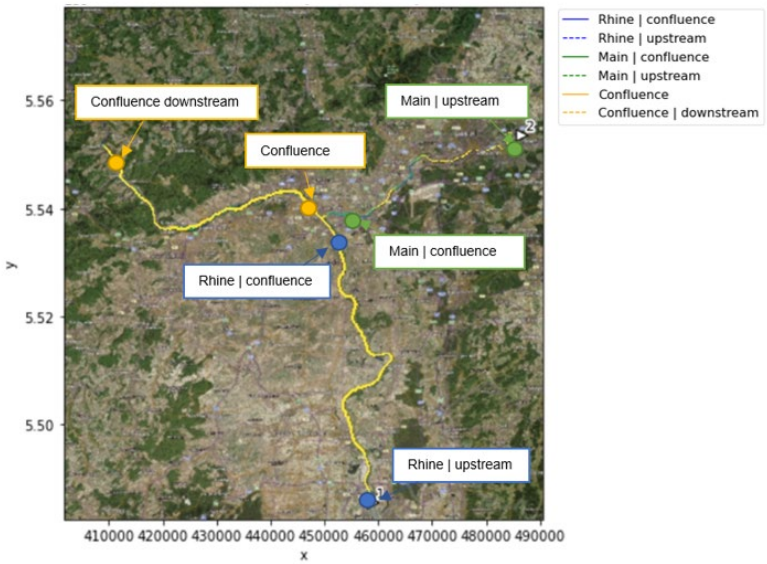


Figure D1 River water levels locations

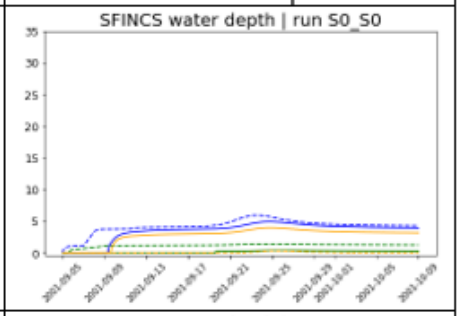
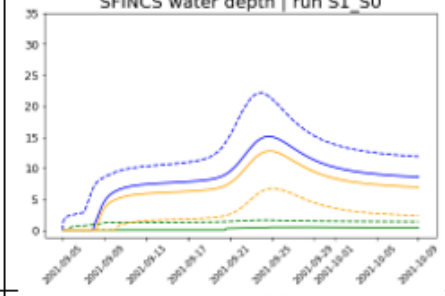
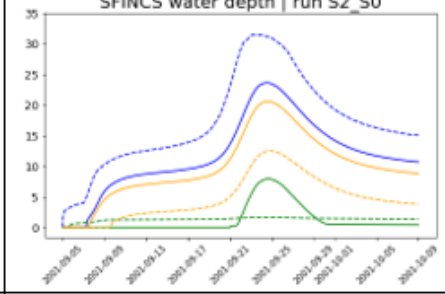
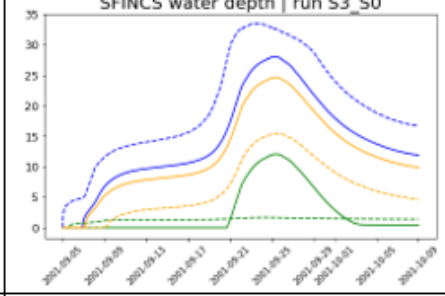
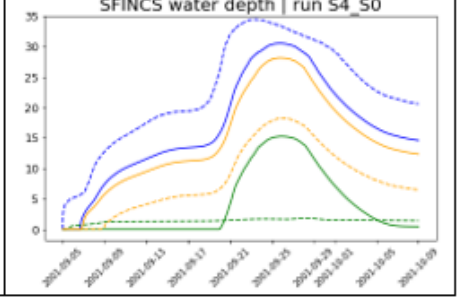
Rhine	Main	River water depth
Q0	Q0	
Q1	Q0	
Q2	Q0	
Q3	Q0	
Q4	Q0	

Figure D2 River water levels (at different locations see Figure D1) according to the hydraulic simulation of the discharges indicated on the left

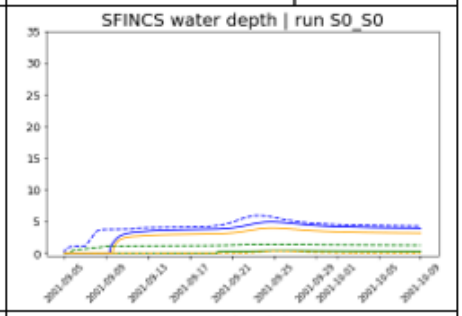
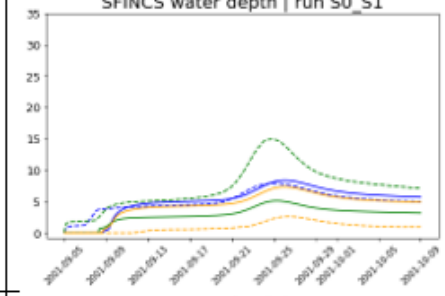
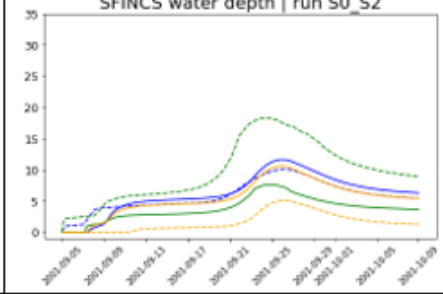
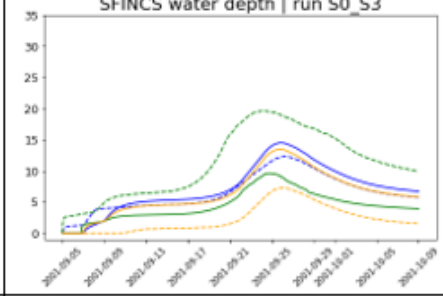
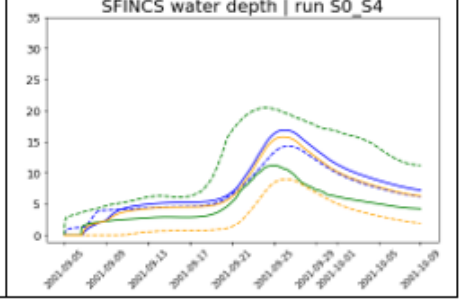
Rhine	Main	River water depth
Q0	Q0	
Q0	Q1	
Q0	Q2	
Q0	Q3	
Q0	Q4	

Figure D3 River water levels (at different locations see Figure D1) according to the hydraulic simulation of the discharges indicated on the left

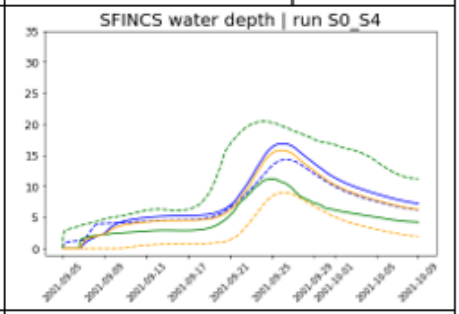
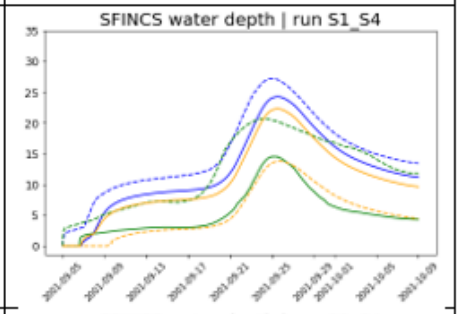
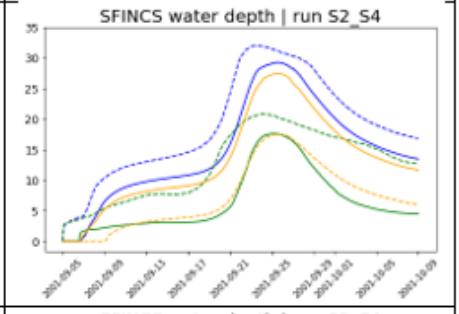
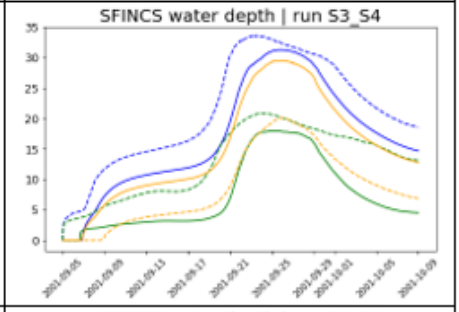
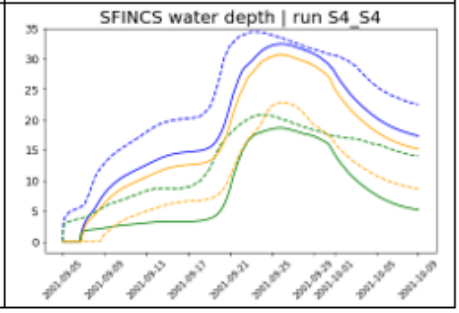
Rhine	Main	River water depth
Q0	Q4	
Q1	Q4	
Q2	Q4	
Q3	Q4	
Q4	Q4	

Figure D4 River water levels (at different locations see Figure D1) according to the hydraulic simulation of the discharges indicated on the left

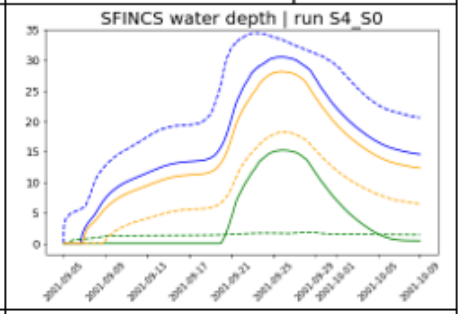
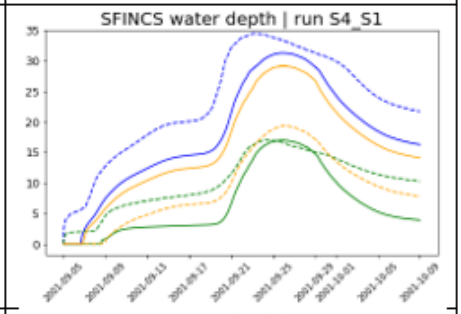
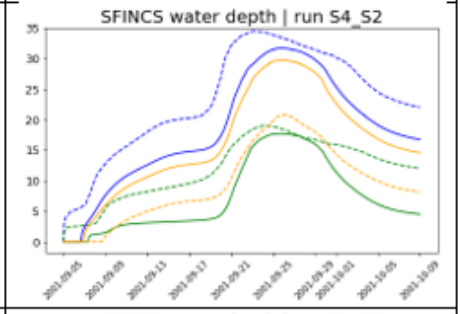
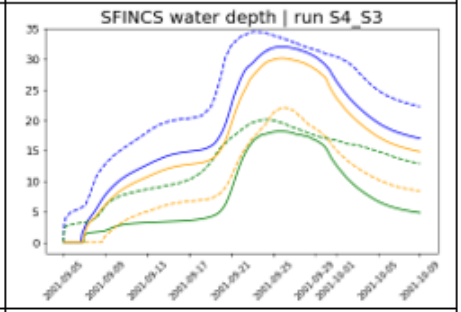
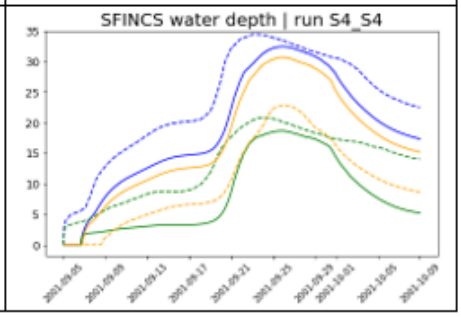
Rhine	Main	River water depth
Q4	Q0	
Q4	Q1	
Q4	Q2	
Q4	Q3	
Q4	Q4	

Figure D5 River water levels (at different locations see Figure D1) according to the hydraulic simulation of the discharges indicated on the left

In general, the water depth directly downstream of the confluence (solid orange line '– confluence') follows the 'shape' of the water depth of the Rhine River at the confluence (solid blue line '– Rhine confluence'), meaning that the Rhine River upstream of the confluence dominates the water depth downstream of the confluence. This could mean that the water is contained in the flood plain, reducing water depths downstream of the confluence.

In Figure D2, we show the water depths for the minimum discharge of the Main ($Main_{S0}$), when having low discharges at the Rhine River we do not observe a peak in the water depth of the Main at the confluence (solid green line '– Main confluence'), but when the Rhine River has higher discharge river we do observe a peak (already observed for simulation $Rhine_{S2}$). The magnitude of the water depth peak on the Main River increases as the magnitude of the Rhine discharge increases. This reflects the influence of the Rhine River on the Main River backwater effect.

In Figure D3, we show the water depths for the minimum discharge of the Rhine ($Rhine_{S0}$), we can observe the influence that the Main River has on the water level of the Rhine River (due to backwater effects). When the Main discharge is low ($Main_{S0}$), the water depth of the Rhine at the confluence (solid blue line '– Rhine confluence') follows the shape of its upstream depth (dashed blue line '-- Rhine upstream'). However, when the discharge at the Main is high (e.g. $Main_{S4}$), even when there is no high discharge at the Rhine, we observe a peak in the water depth at both locations of the Rhine, at the confluence and upstream. As shown in Figure 41, the difference between $Rhine_{S0}$-$Main_{S4}$ and $Rhine_{S0}$-$Main_{S0}$ does not lead to a flooded area in the upstream part, but close to the confluence. Still, the backwater effect does reach the upstream part of the river.

In Figure D4, we show the water depths for the maximum peak discharge at the Main ($Main_{S4}$) we observe differences in its water depth for the different discharges at the Rhine River. As the discharge increases at the Rhine, not only does the water depth of the Main increase but also the duration of high water levels increases.

In Figure D5, we show the water depths for the maximum peak discharge at the Rhine ($Rhine_{S4}$), the water depth does not increase as the Main discharge increases, but we still observe changes in the duration of the peak water levels.

Flooded maps enlargement

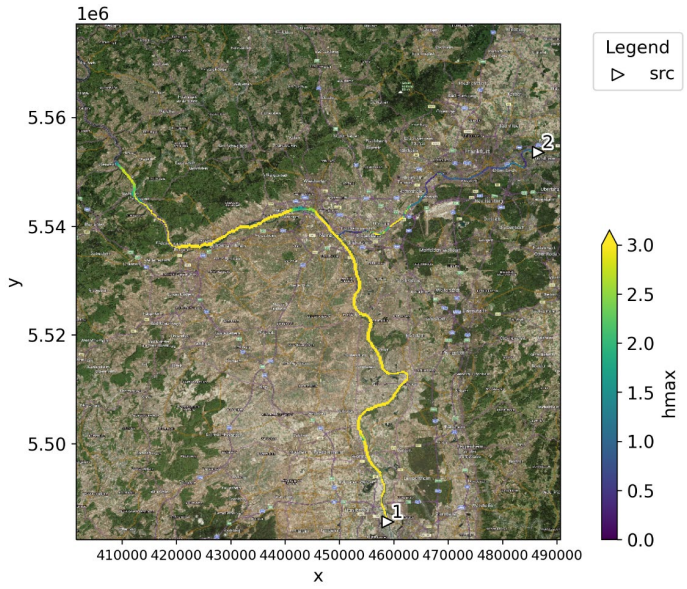
The following figures are an enlargement of the flooded maps presented as a grid according to each river discharge in Figure 38 (Section 5.3). The maximum water depth (h_{max}) of the hydraulic simulations is depicted.

The grid of discharges that were simulated is presented in Figure D6, indicating the name of the simulation, which is also indicated in each of the figures with maximum water depth (e.g. run S0_S4).

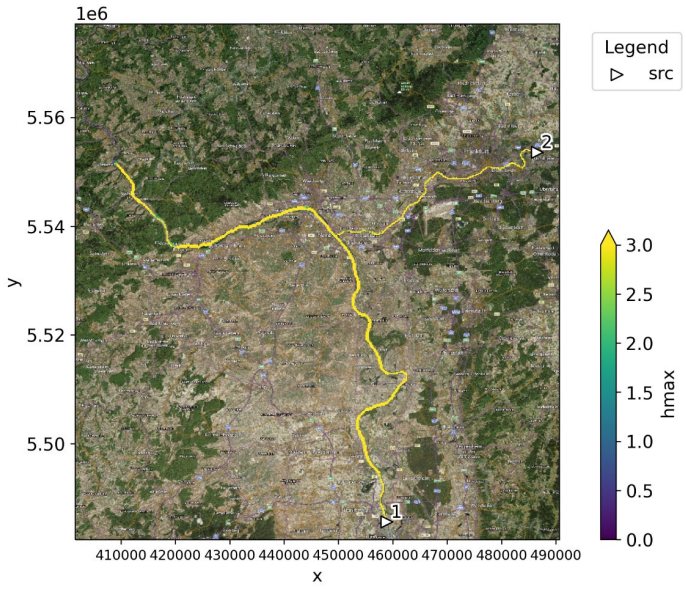
Main	Q4	4397	S0-S4	S1-S4	S2-S4	S3-S4	S4-S4
	Q3	3305	S0-S3	S1-S3	S2-S3	S3-S3	S4-S3
	Q2	2212	S0-S2	S1-S2	S2-S2	S3-S2	S4-S2
	Q1	1120	S0-S1	S1-S1	S2-S1	S3-S1	S4-S1
	Q0	27	S0-S0	S1-S0	S2-S0	S3-S0	S4-S0
		m^3/s	724	3404	6083	8763	11443
			Q0	Q1	Q2	Q3	Q4
			Rhine				

Figure D6 Grid of discharges and key names of the hydraulic simulations. Mainstream (MS-Rhine) in the x-axis and Tributary stream (TS-Main) in the y-axis

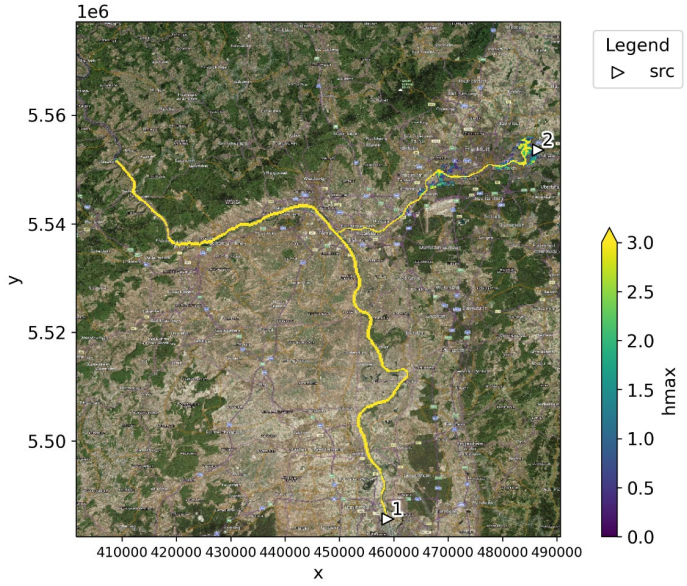
SFINCS maximum water depth | run S0_S0



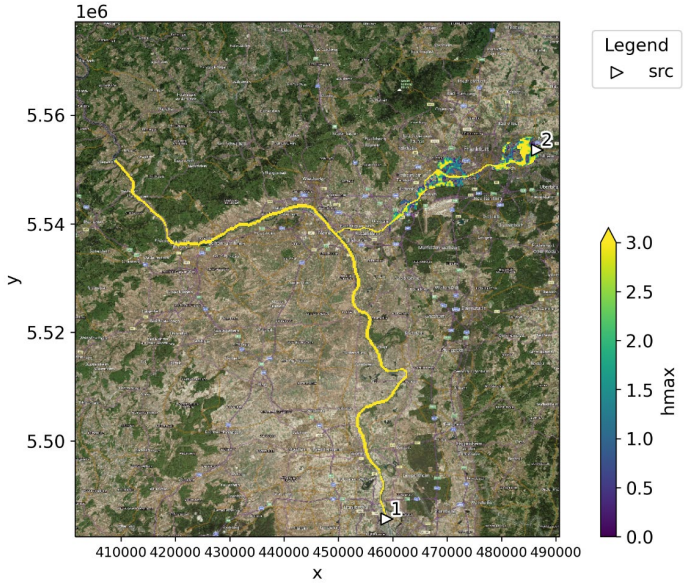
SFINCS maximum water depth | run S0_S1



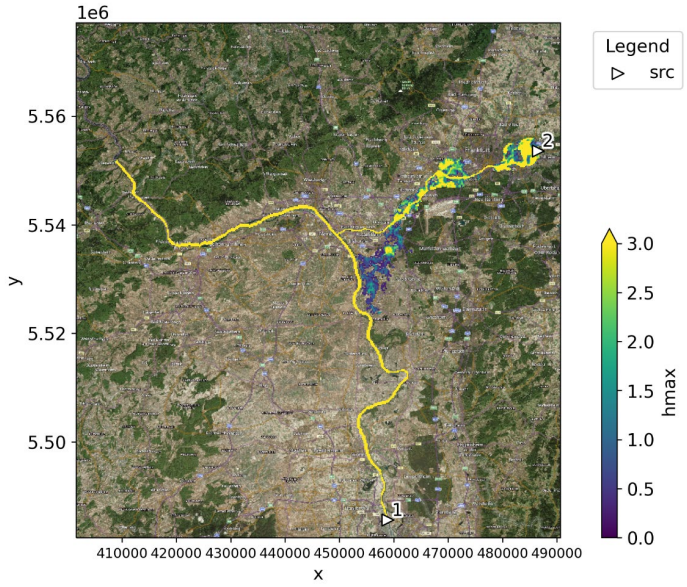
SFINCS maximum water depth | run S0_S2



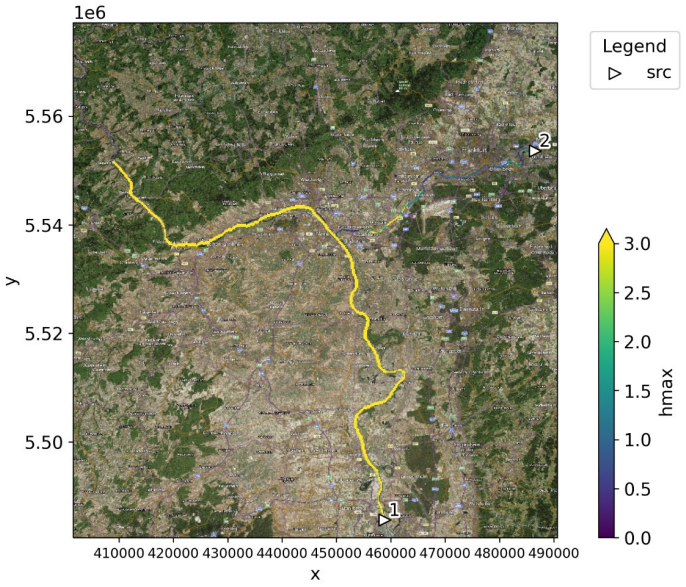
SFINCS maximum water depth | run S0_S3



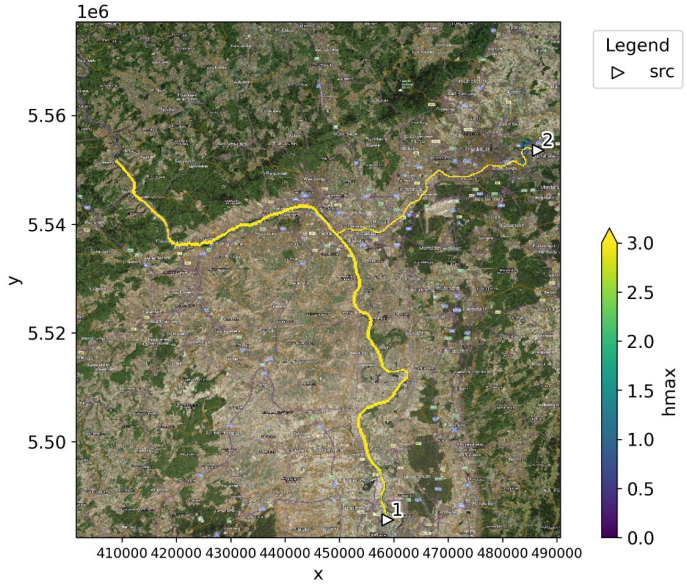
SFINCS maximum water depth | run S0_S4



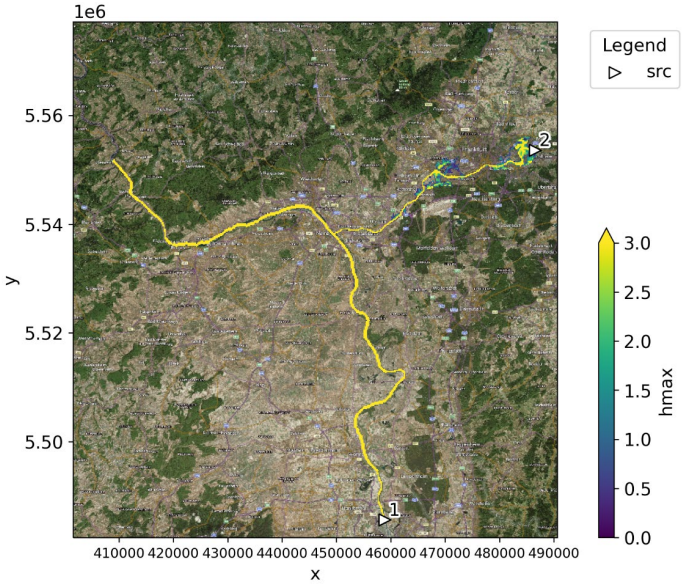
SFINCS maximum water depth | run S1_S0



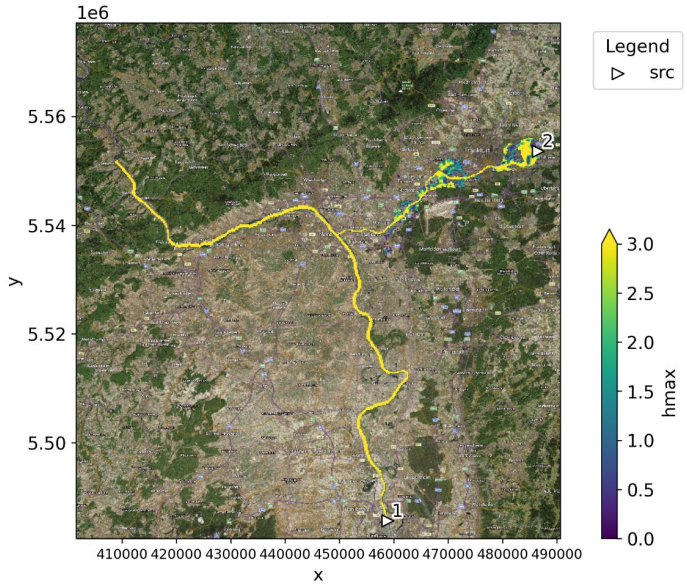
SFINCS maximum water depth | run S1_S1



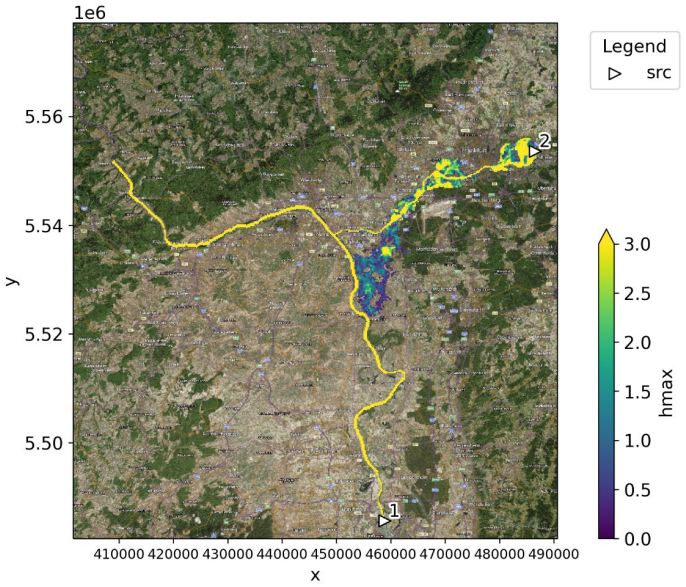
SFINCS maximum water depth | run S1_S2



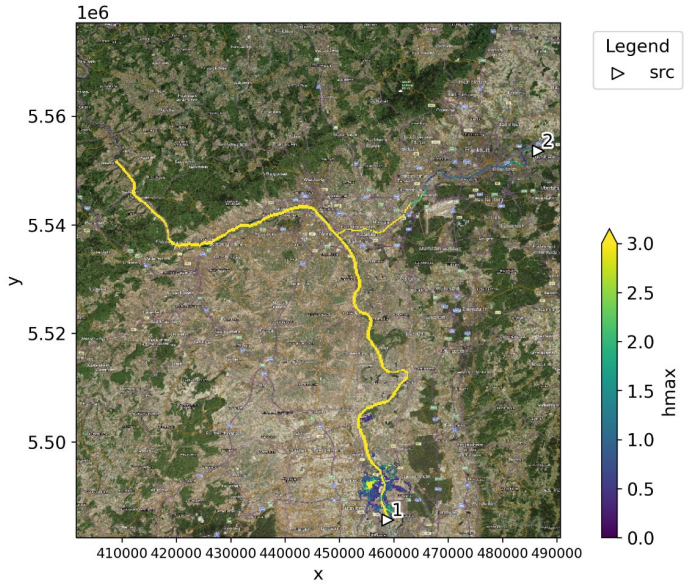
SFINCS maximum water depth | run S1_S3



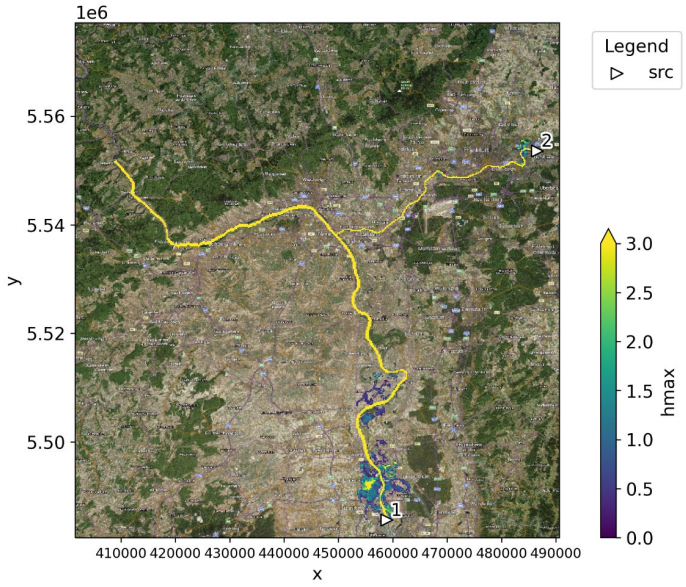
SFINCS maximum water depth | run S1_S4



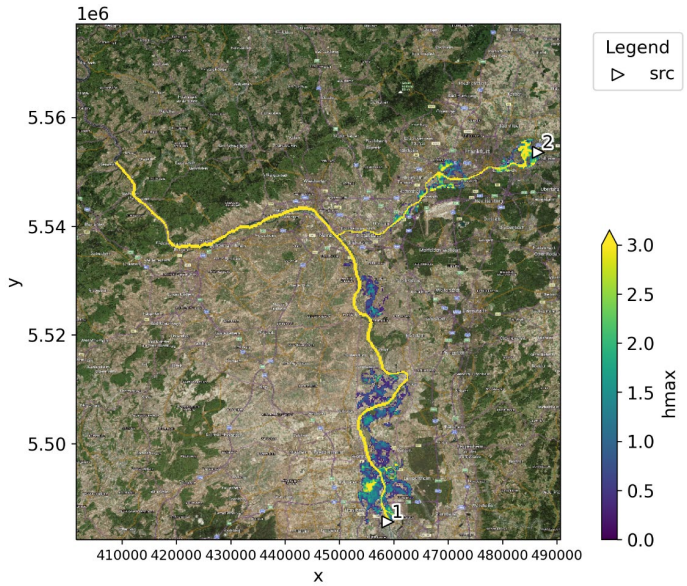
SFINCS maximum water depth | run S2_S0



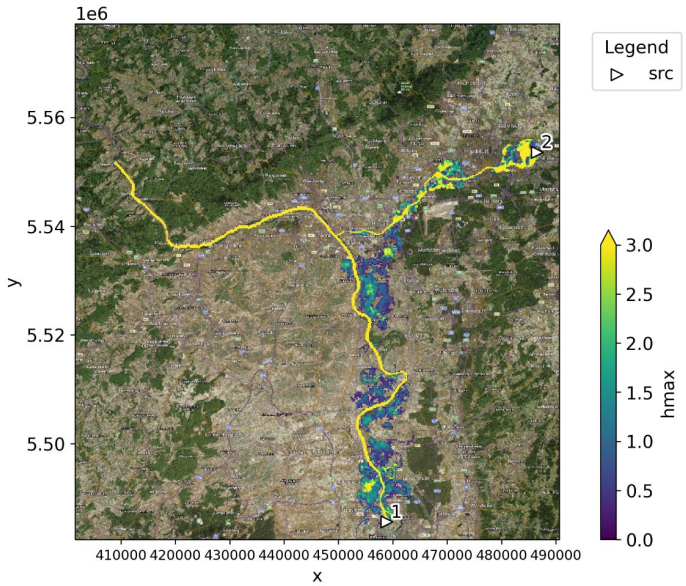
SFINCS maximum water depth | run S2_S1



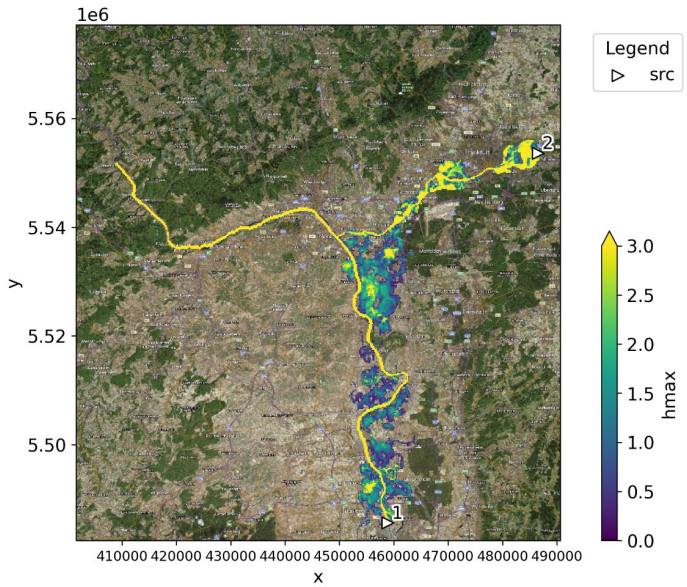
SFINCS maximum water depth | run S2_S2



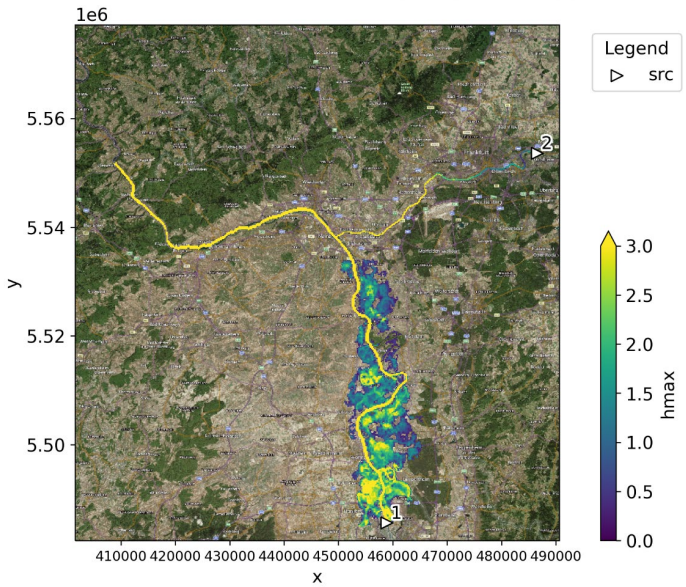
SFINCS maximum water depth | run S2_S3



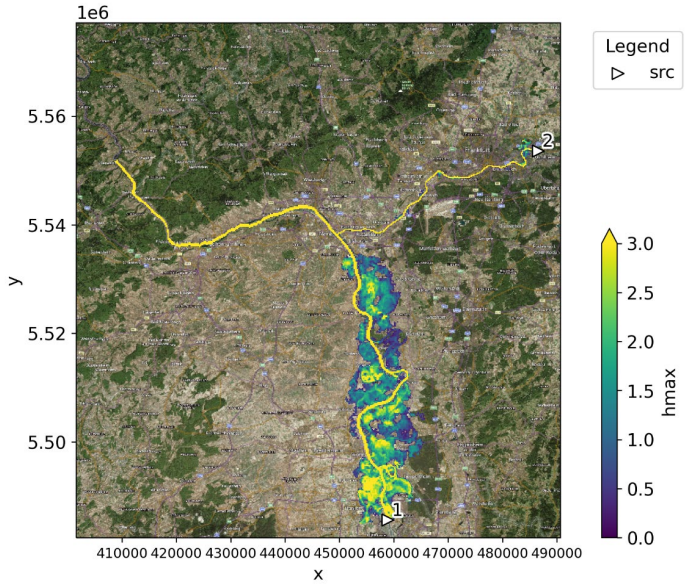
SFINCS maximum water depth | run S2_S4



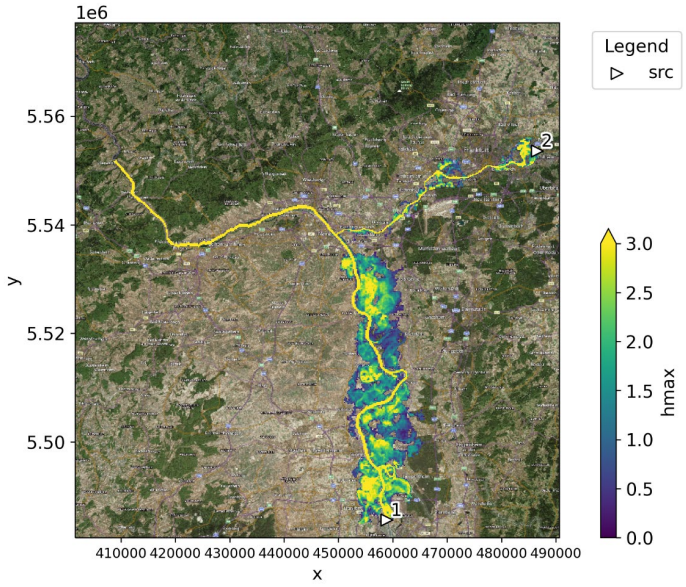
SFINCS maximum water depth | run S3_S0



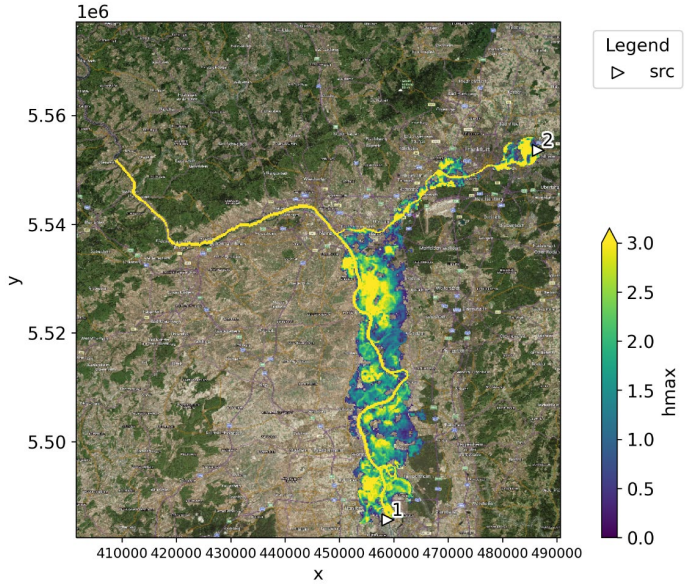
SFINCS maximum water depth | run S3_S1



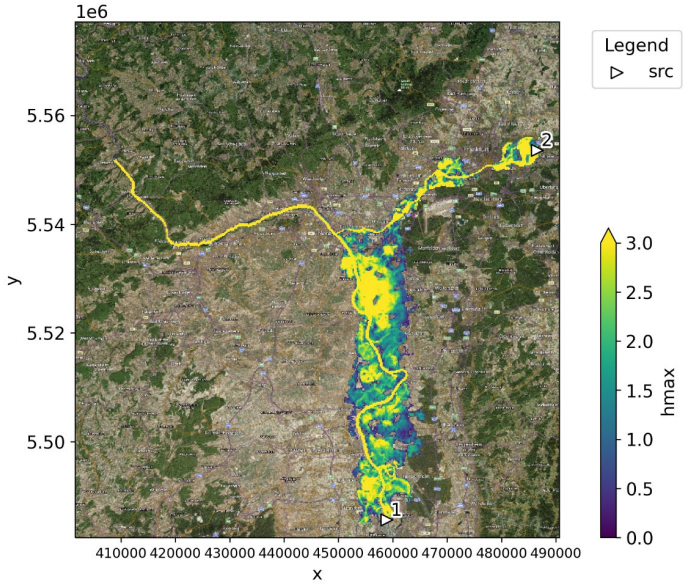
SFINCS maximum water depth | run S3_S2



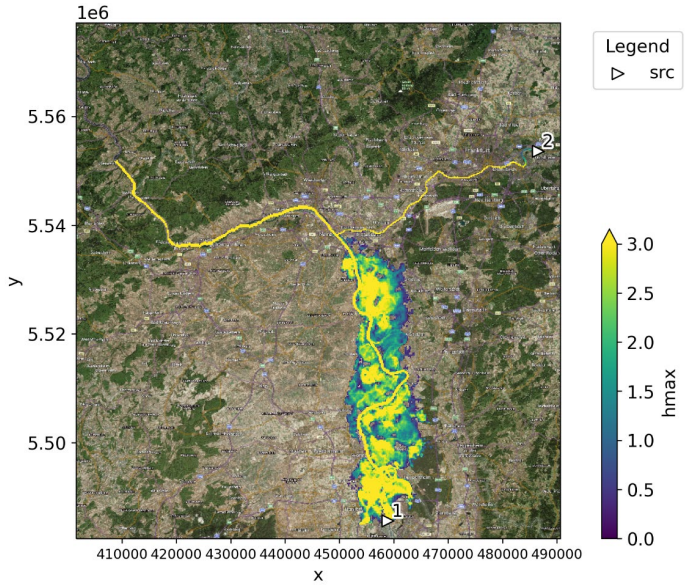
SFINCS maximum water depth | run S3_S3



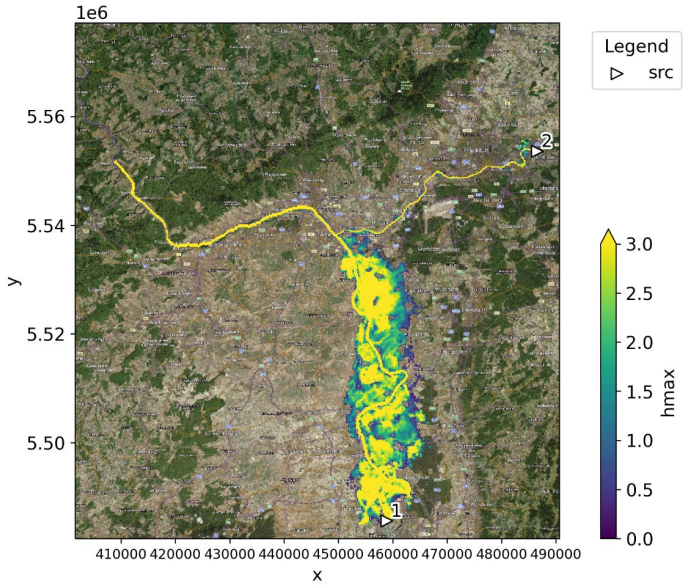
SFINCS maximum water depth | run S3_S4



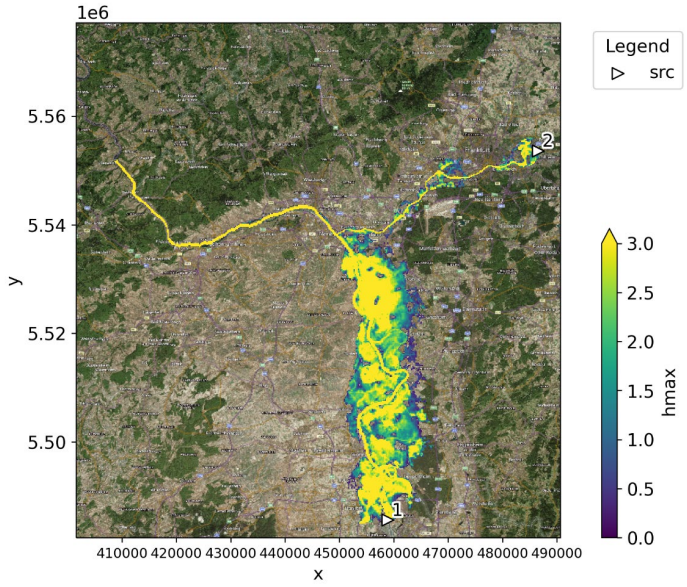
SFINCS maximum water depth | run S4_S0



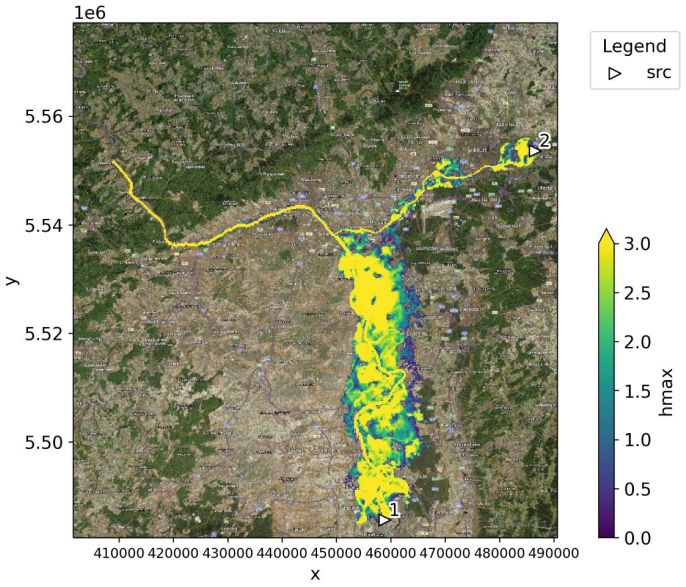
SFINCS maximum water depth | run S4_S1



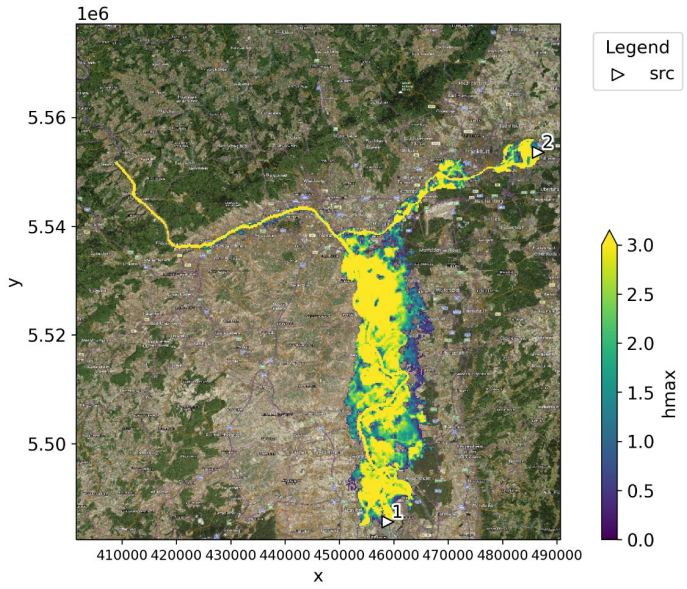
SFINCS maximum water depth | run S4_S2



SFINCS maximum water depth | run S4_S3



SFINCS maximum water depth | run S4_S4



Appendix E – Marginal distribution fits

The evaluation of the marginal distributions of the variables is presented in this Appendix.

The sets of extremes considered for the analysis are non-simultaneous and repeated below:

- *Set 1 | MSmax-TSconc: mainstream maxima (MS_{max}) with a concurrent tributary stream (TS_{conc}),*
- *Set 2 | TSmax-MSconc: tributary stream maxima (TS_{max}) with a concurrent mainstream (MS_{conc}),*
- *Set 3 | Cmax (C=MS+TS): the maxima of the sum of both streams.*

Therefore, the evaluation is done for each of the variables (peak and concurrent discharge) of the three sets of extremes.

As mentioned in Section 2.1.2. when selecting the annual maxima, the distribution converges to a Generalized Extreme Value (GEV). Therefore, for the rivers where the annual maxima are selected, we can use the GEV distribution.

However, the concurrent flow is not necessarily an extreme situation, therefore it is important to evaluate the fit of the distribution. The full data (50,000 years) is used to evaluate the distribution fit.

First, constructing the empirical distribution of the variables (river peak/concurrent discharges) of each extreme set, and second, comparing it to four different probability distributions: Gumbel (GUM), Generalized Extreme Value (GEV), General Pareto Distribution (GPD), and Normal (Nor). The parameters of the evaluated distributions are determined by fitting the distribution to the full data. There are several methods available for this task, such as the method of moments, maximum likelihood and probability-weighted moments. we chose to use the L-moments method (Hosking & Wallis, 1997)

The cumulative distribution function of the four probability functions and the empirical distribution for Set 1, Set 2, and Set 3 can be seen in Figures E1, E2, and E3 respectively

As explained in Section 6.1, three different scores, RMSE, AIC, and BIC, were evaluated (see Section 2.2.4), and the distributions with the best results were selected and used for the hydrological sampling (see Table 7).

Table E1. Marginal distribution fit of the Rhine (MS) and Main (TS) rivers for each extreme set, according to the best scores

	Rhine (MS)	Main (TS)
Set 1 | MS_{max}-TS_{conc}	GEV	GPD
Set 2 | TS_{max}-MS_{conc}	GEV	GEV
Set 3 | C_{max} (C=MS+TS)	GEV	GPD

However, when looking at the plots of the return periods, using a logarithmic scale (Figures E4, E5, and E6), the end tails are not always well represented by the selected distributions according to the best score. Therefore, a second selection of the marginal distribution is done for each set of extremes (see Table E2)

Table E2. Marginal distribution fit of the Rhine (MS) and Main (TS) rivers for each extreme set, according to the score and the plot

	Best score		Best plot	
	Rhine (MS)	Main (TS)	Rhine (MS)	Main (TS)
Set 1 | MS_{max}-TS_{conc}	GEV	GPD	GEV	GPD
Set 2 | TS_{max}-MS_{conc}	GEV	GEV	GUM	GUM
Set 3 | C_{max} (C=MS+TS)	GEV	GPD	GEV	GUM

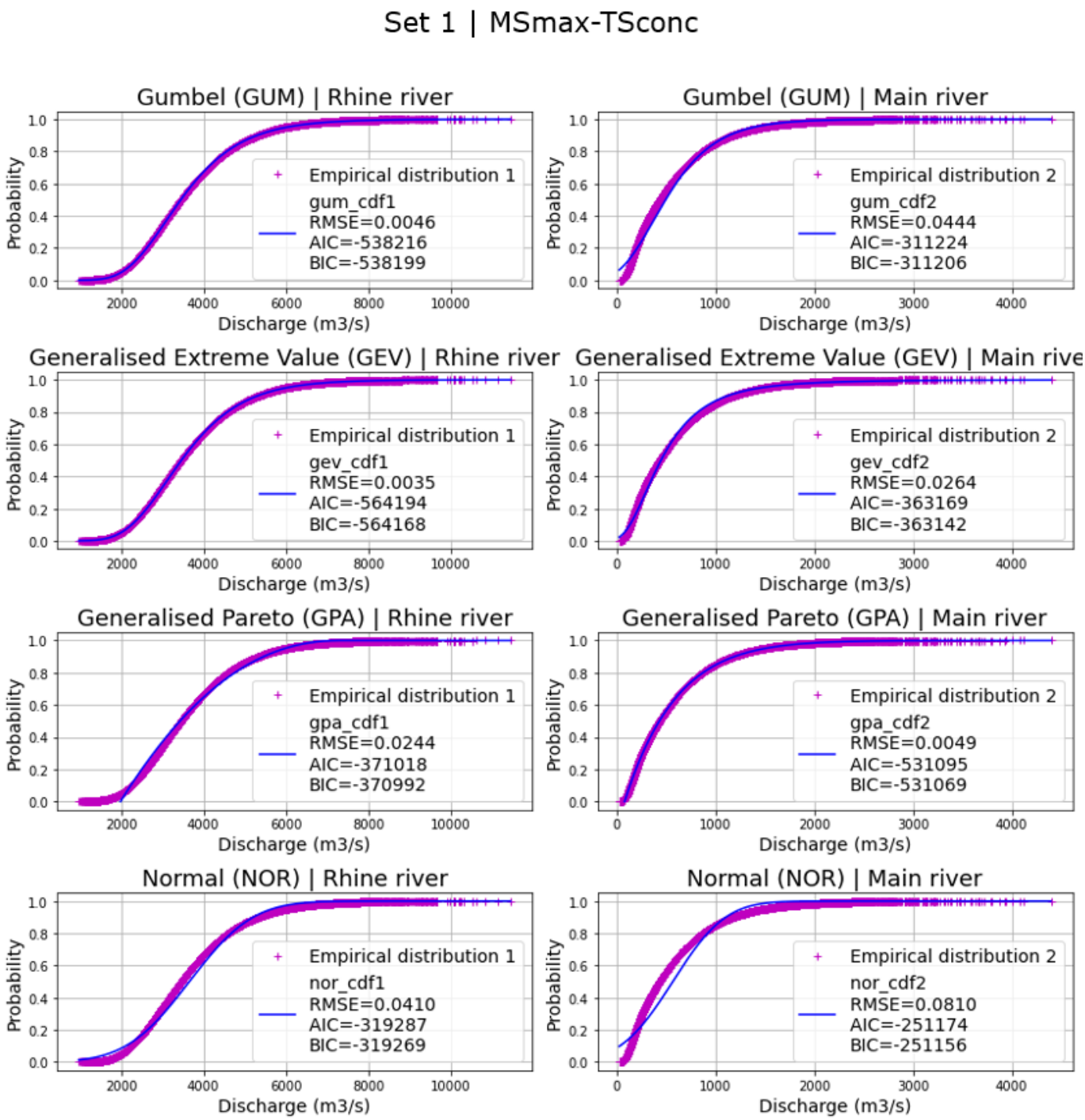


Figure E1. Set 1 – Marginal distribution evaluation: Cumulative Distribution Functions (CDF) of the empirical distribution (pink) and the distribution fits (blue)

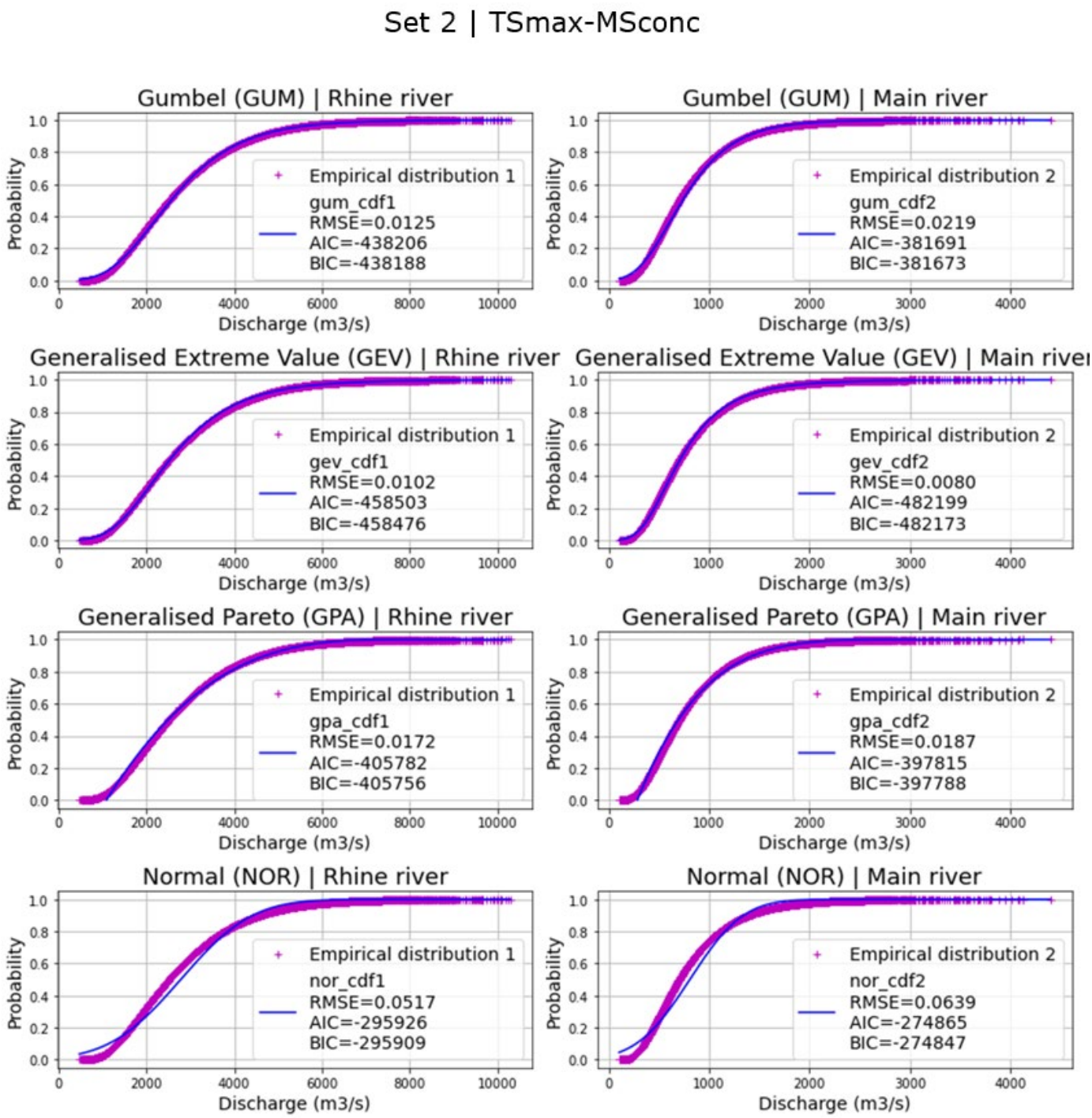


Figure E2. Set 2 – Marginal distribution evaluation: Cumulative Distribution Functions (CDF) of the empirical distribution (pink) and the distribution fits (blue)

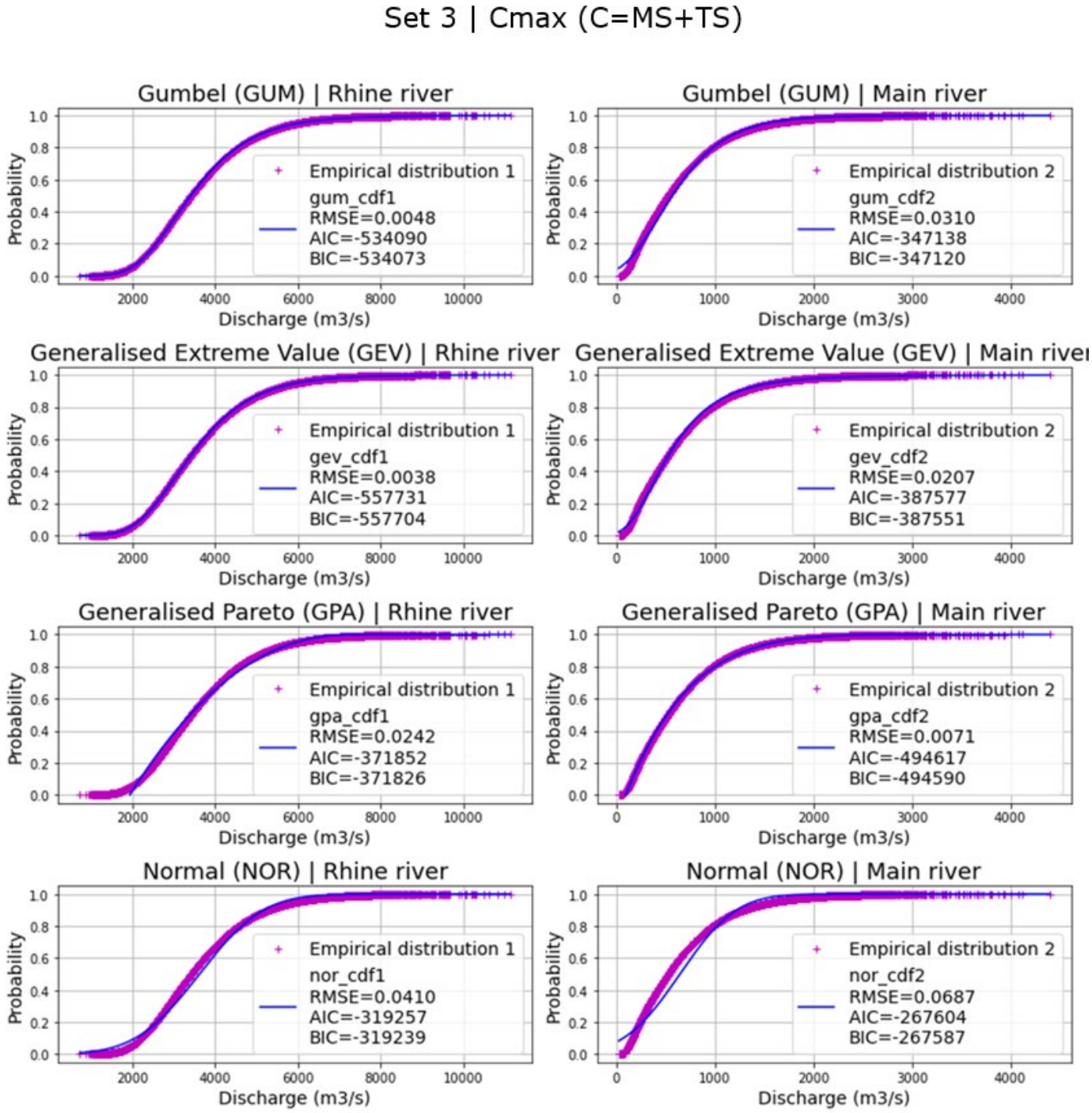


Figure E3. Set 3 – Marginal distribution evaluation: Cumulative Distribution Functions (CDF) of the empirical distribution (pink) and the distribution fits (blue)

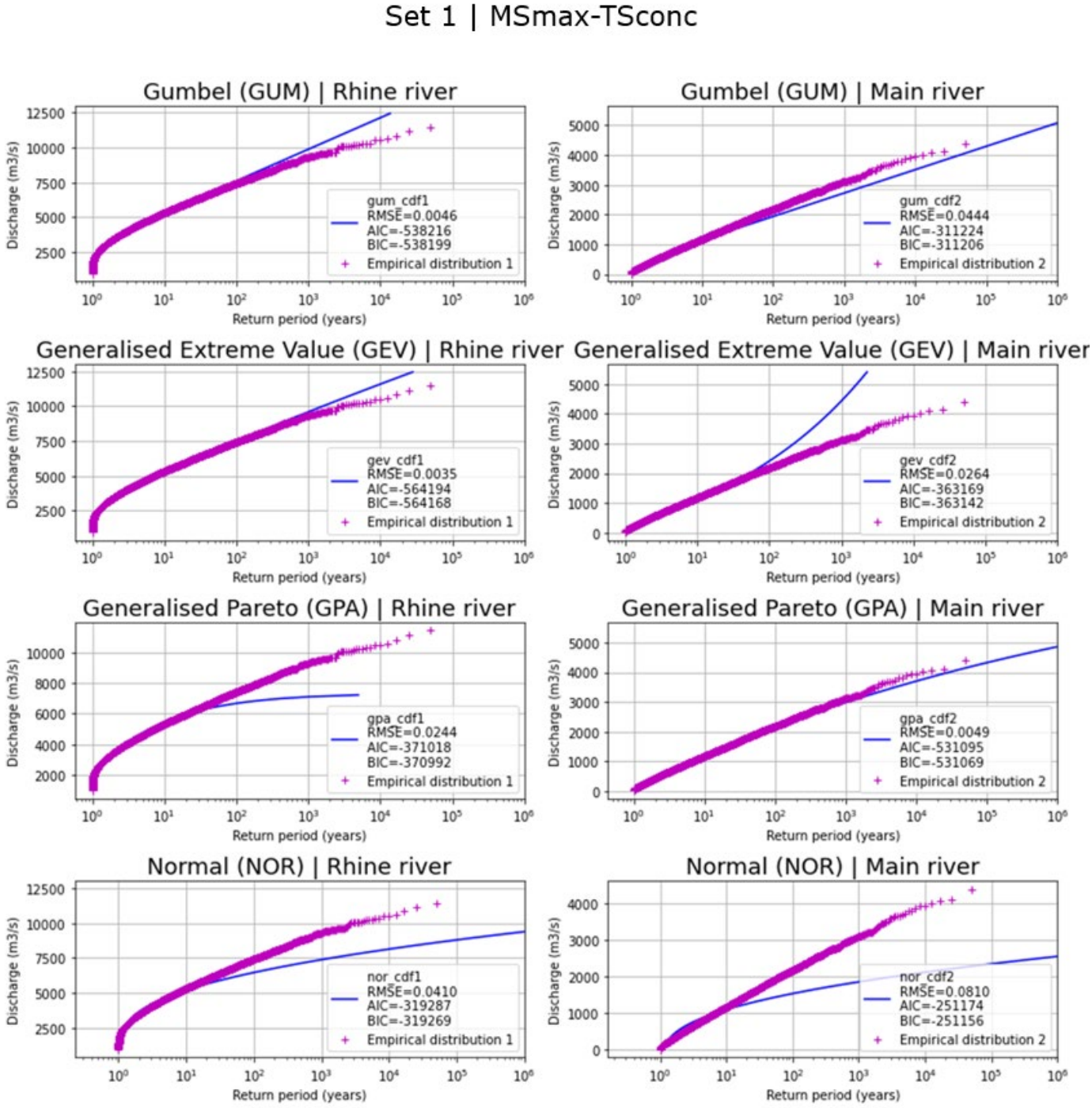


Figure E4. Set 1 – Marginal distribution evaluation: Return periods (x-axis log scale) of the empirical distribution (pink) and the distribution fits (blue)

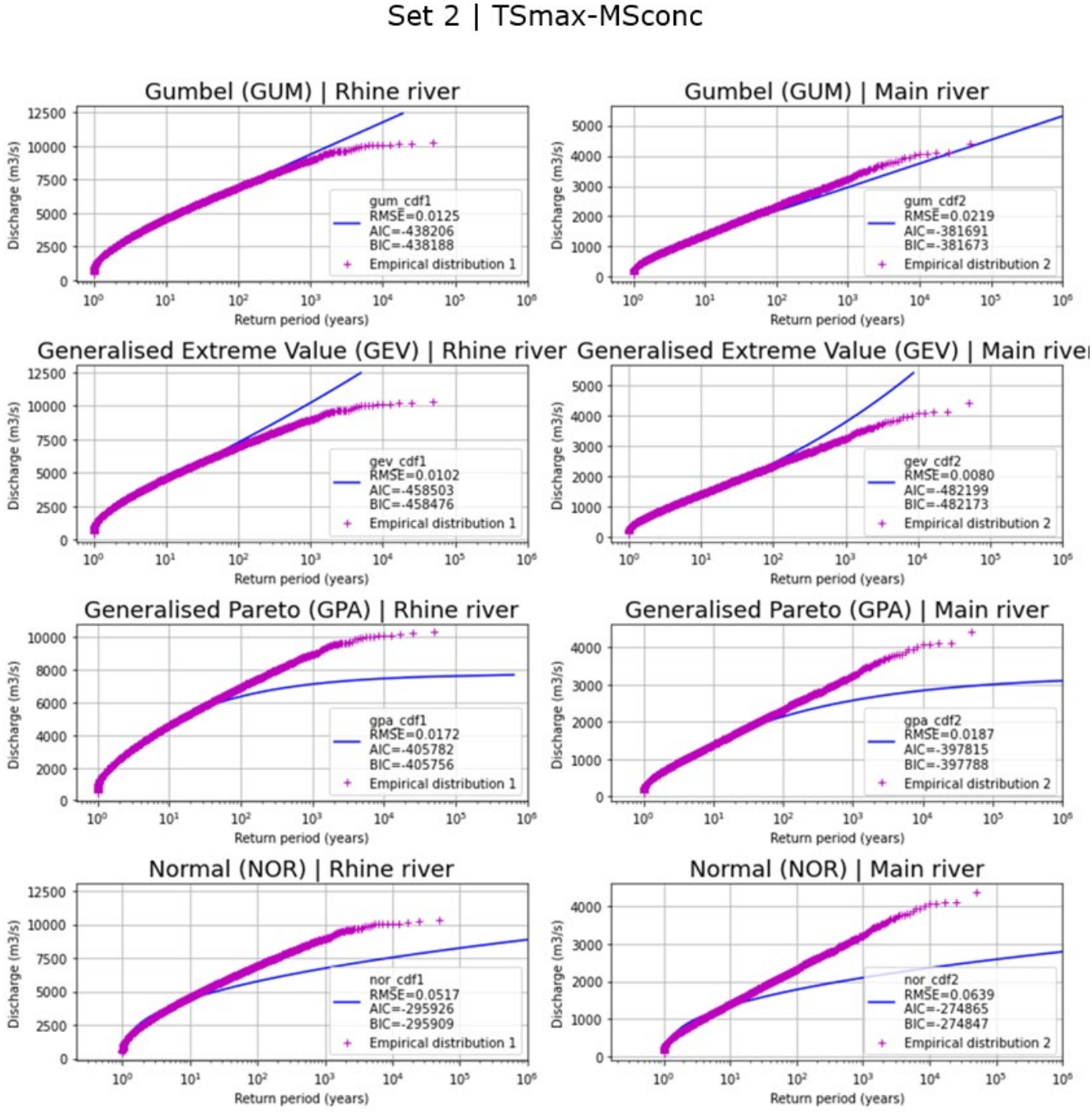


Figure E5. Set 2 – Marginal distribution evaluation: Return periods (x-axis log scale) of the empirical distribution (pink) and the distribution fits (blue)

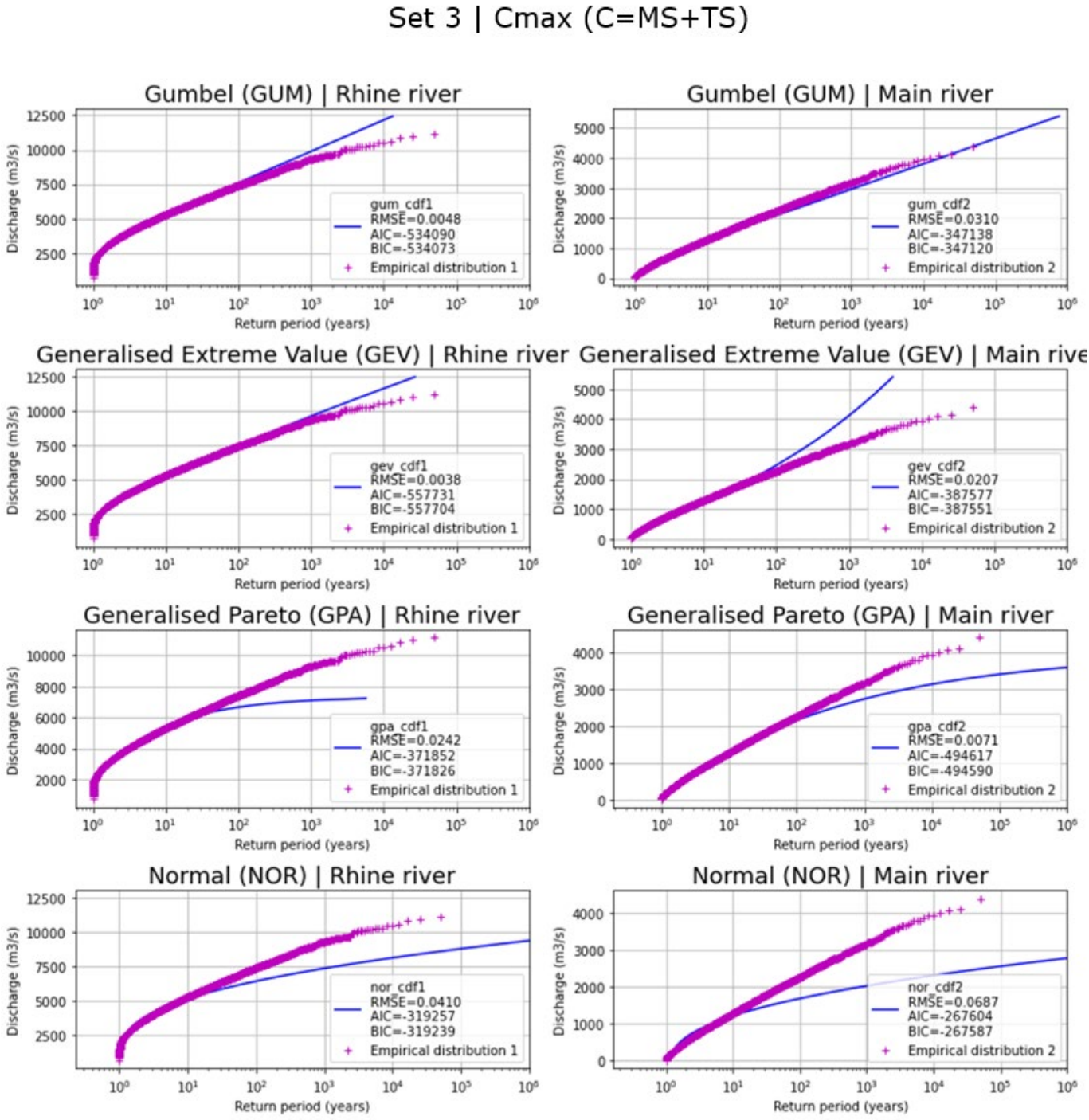


Figure E6. Set 3 – Marginal distribution evaluation: Return periods (x-axis log scale) of the empirical distribution (pink) and the distribution fits (blue)

Appendix F – Semi-correlation tests

As mentioned in Section 7.2.4. an evaluation of the copulas was done using the semi-correlated test. In Figures F1 and F2, we can see the scatter plots of the observed data and the model data (in the standard normal space), for Set 2 and Set 3 respectively.

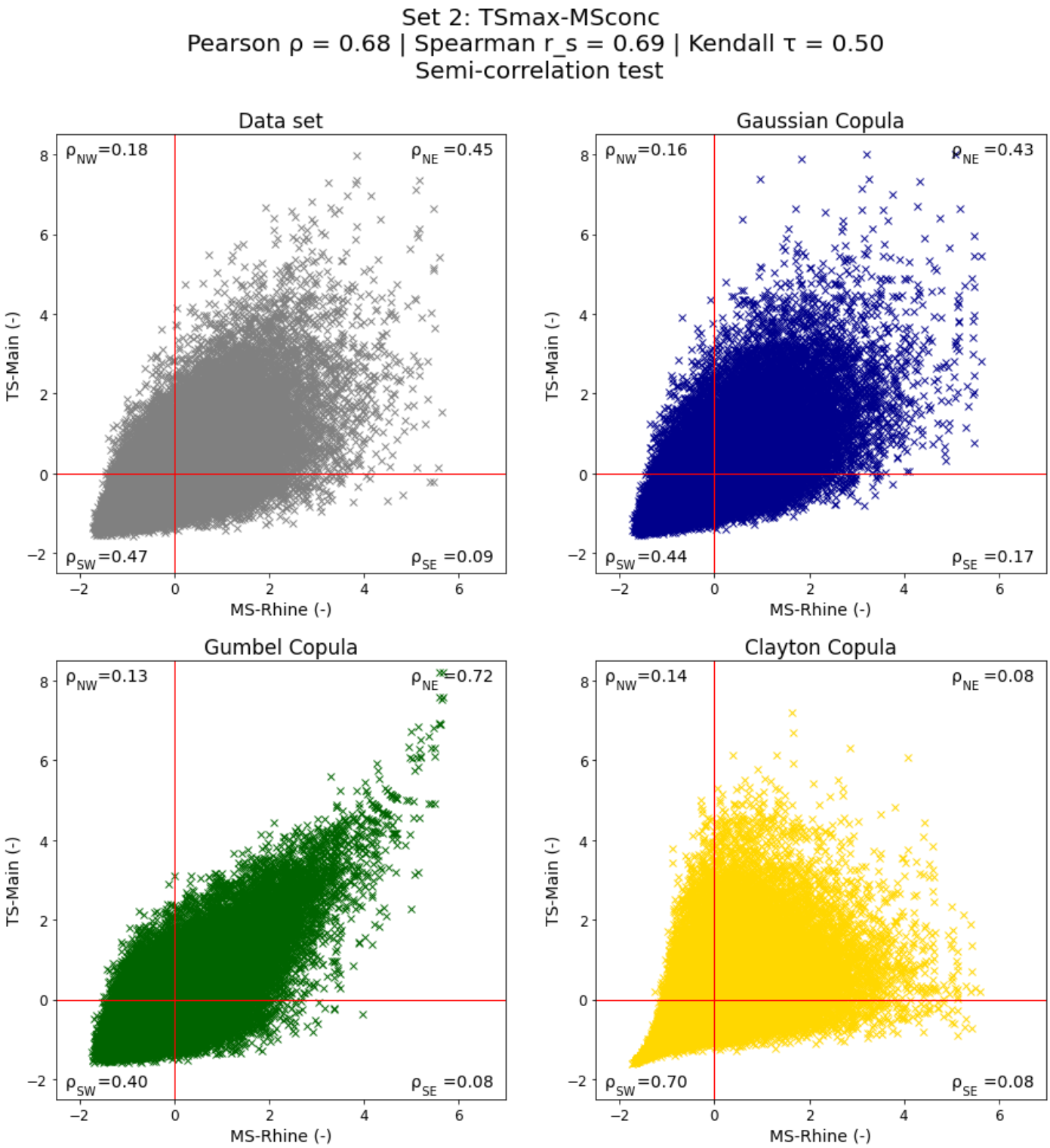


Figure F1. Semi-correlation test of Set 2. Scatter plots of the pairs of data: Mainstream (MS-Rhine) on the x-axis and Tributary stream (TS-Main) on the y-axis. Synthetic data (grey), and the simulated data using the copulas: Gaussian (blue), Gumbel (green), and Clayton (yellow). All of them are in the standard normal space

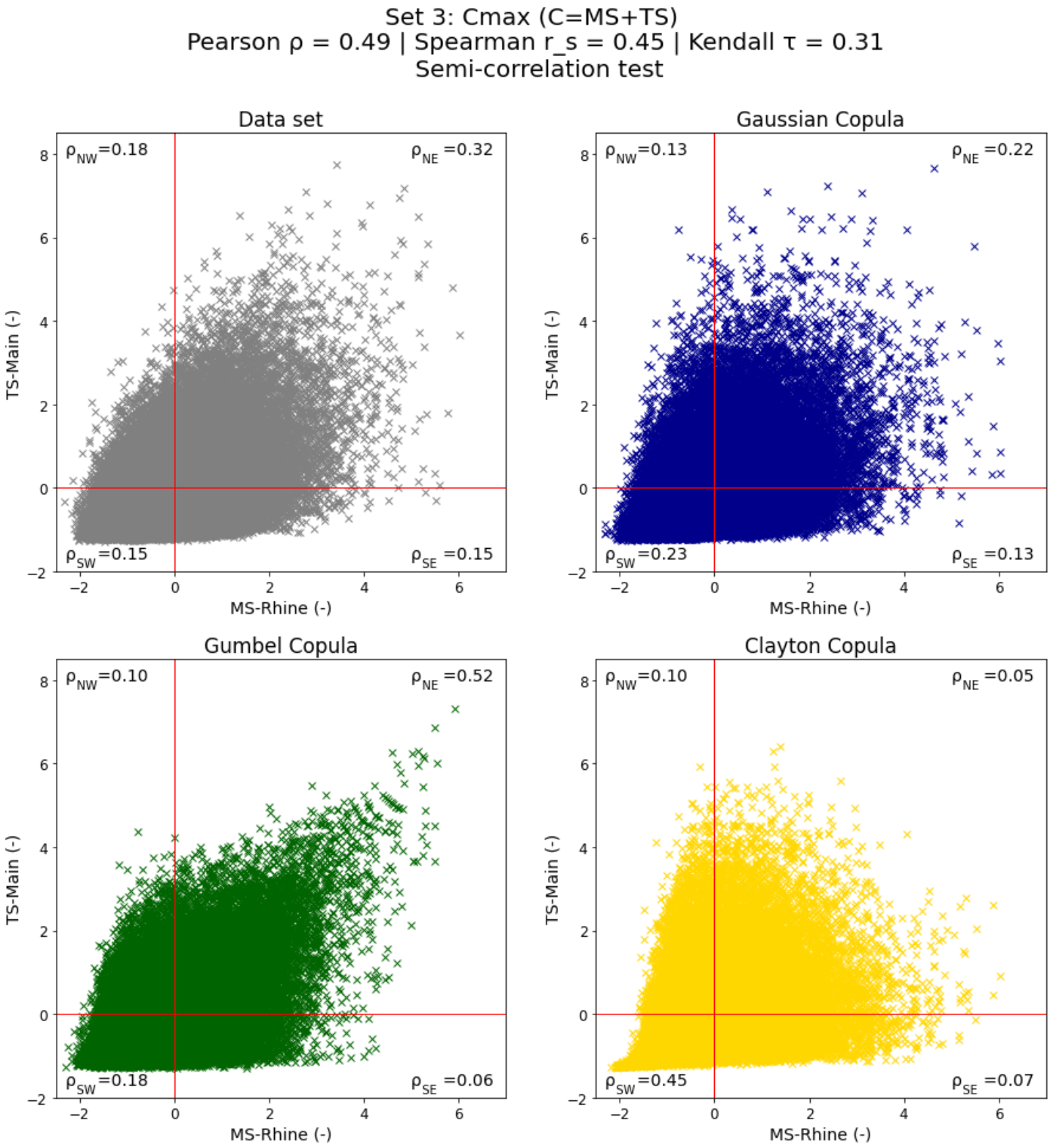


Figure F2. Semi-correlation test Set 3. Scatter plots of the pairs of data: Mainstream (MS-Rhine) on the x-axis and Tributary stream (TS-Main) on the y-axis. Synthetic data (grey), and the simulated data using the copulas: Gaussian (blue), Gumbel (green), and Clayton (yellow). All of them are in the standard normal space