TU Dortmund University
Department of Statistics

DISSERTATION
in partial fulfilment of the requirements for the degree of
Doktor der Naturwissenschaften

Robust estimation methods with application to flood statistics

Svenja Fischer
born on 02.07.1989 in Recklinghausen

August 2017

Advisors: Prof. Dr. A. Schumann, Prof. Dr. M. Wendler, Prof. Dr. R. Fried
Reviewers: Prof. Dr. R. Fried, Prof. Dr. W. Krämer, Prof. Dr. M. Wendler, Prof. Dr. A. Schumann

Abstract

Robust statistics and the use of robust estimators have come more and more into focus during the last couple of years. In the context of flood statistics, robust estimation methods are used to obtain stable estimations of e.g. design floods. These are estimations that do not change from one year to another just because one large flood occurred.

A problem which is often ignored in flood statistics is the underlying dependence structure of the data. When considering discharge data with high time-resolution, short range dependent behaviour can be detected within the time series. To take this into account, in this thesis a limit theorem for the class of GL-statistics is developed under the very general assumption of near epoch dependent processes on absolutely regular random variables, which is a well known concept of short range dependence. GL-statistics form a very general class of statistics and can be used to represent many robust and non-robust estimators, such as Gini's mean difference, the Q_n-estimator or the generalized Hodges-Lehmann estimator. In a direct application the limit distribution of L-moments and their robust extension, the trimmed L-moments, is derived. Moreover, a long-run variance estimator is developed. For all these results, the use of U-statistics and U-processes proves to be the key tool, such that a Central Limit Theorem for multivariate U-statistics as well as an invariance principle for U-processes and the convergence of the remaining term of the Bahadur-representation for U-quantiles is shown. A challenge for proving these results pose the multivariate kernels that are considered to be able to represent very general estimators and statistics.

A concrete application in the context of flood statistics, in particular in the estimation of design floods, the classification of homogeneous groups and the modelling of short range dependent discharge series, is given. Here, well known models (peak-over-thresholds) as well as newly developed ones, for example mixing models using the distinction of floods according to their timescales, are combined with robust estimators and the advantages and disadvantages under consideration of stability and efficiency are investigated. The results show that the use of the new models, that take more information into account by enlarging the data basis, in combination with robust estimators leads to a very stable estimation of design floods, even in high quantiles. Whereas a lot of the classical estimators, like Maximum-Likelihood estimators or L-moments, are affected by single extraordinary extreme events and need a long time to stabilise, the robust methods approach the same level of stabilisation rather fast. Moreover, the newly developed mixing model cannot only be used for flood estimation but also for regionalisation, that is the modelling of ungauged basins. Here, especially when needing a classification of flood events and homogeneous groups of gauges, the use of robust estimators proves to result in stable estimations, too.

[0] Title: flood in Dresden, 2002; source: `http://www.tiesel.de/schwere%20katastrophen%20.html`, last visited: 22.03.2017

Contents

List of Figures

List of Tables

1. Introduction

A flood is defined as time-limited exceedance of discharge-thresholds in a cross-section of a river with a related drainage area caused by meteorological events (DIN 4049-1). This threshold is not defined precisely and depends on the local circumstances. For example, a flood can be defined as a discharge event that covers land that is typically not covered by water. In general, it is assumed that under humid conditions at least one flood occurs in a year (DWA (2012)).

Besides deterministic models (e.g. rainfall-runoff models) statistical models and statistical evaluation are the most commonly used methods to describe the coherences of climatic, meteorological and discharge phenomena and to obtain predictions, for example for design events for flood protection systems.
Since the processes leading to discharges can be explained by physical phenomena, it might seem somewhat artificial to use stochastic models instead of deterministic ones. In fact, this is an often discussed and misunderstood point. The most important difference between deterministic and stochastic models is the handling of errors (see e.g. Bierkens and van Geer (2008)). Stochastic models are developed to predict values at unknown time or at unknown location, where at the same time an assumption on the error can be made. That is, stochastic models give us the information how uncertain the estimation is. Errors in hydrological models can have many sources. First of all, not the real discharge is measured. Instead, often the water level at the gauge is measured and used to calculate the discharge by taking into account the flow velocity, that is a quantity based on empirical experiences. This of course causes uncertainty. Often even the assumptions on the stage-discharge relation change during the years and the discharge series of a gauge have to be calculated anew. Additionally, many of the main hydrological processes of runoff formation leading to discharge cannot be observed since they proceed under the land surface. Groundwater or the permeability of the soil can just approximately be estimated based on information about the time of infiltration or the soil conditions of the point scale. Most importantly, all models can only be seen as an approximation of the complex processes that lead to discharge. Even in deterministic hydrology these errors are taken into account. During the calibration the difference to the residuals is minimized and the parameters are calculated. After this step, though, the errors are no longer taken into account and, therefore, do not appear in the outcomes. In stochastic hydrology, the errors of the outcomes are handled as well. The treatment of discharge as random variable handles these errors indirectly. To make the difference between the two approaches clearer we want to give an example following Bierkens and van Geer (2008). Consider the discharge x. Using a deterministic model this value is represented by the model outcome x'. The made error can then be denoted by $\epsilon = x' - x$, the so called residuum. In stochastic hydrology we now have additional information on the unknown error by considering it as random variable with known probability distribution. Often, not the error is treated as random variable but we assume the whole measurement x to be a random variable, the exact value of which we do not know, such that the error is treated indirectly.
Especially when considering extreme events, the use of stochastic models is advantageous compared to deterministic models, which typically underestimate extreme values. The first explicit connections between statistical methods and hydrological data have been made in the 1940s

(Gumbel (1958)) when among others the mathematician Emil Julius Gumbel has investigated return periods of floods (Gumbel (1941)). Nevertheless, only descriptive statistic has been applied and no stochastic models have been widely used in hydrology until the 1970s, when the books of Vujica Yevjevich have been published (Yevjevich (1972a), Yevjevich (1972b)). Here, Yevjevich tries to make the advantages of stochastic models and the complex information represented by them easier understandable and therefore better applicable by hydrologists by reducing the complexity to the cases relevant for hydrology and avoiding proofs. He explains the meaning of confidence ranges and the importance of statistical tests and estimators besides multivariate consideration of hydrological data. Moreover, he is one of the first hydrologists considering discharge series as time-variant process of stochastic nature, that can be autocorrelated or non-stationary. This topic has been extended by Salas et al. (1980), where different kinds of dependence have been taken into account. This has then been the starting point of a wide use of stochastic models and methods in hydrology and still today many newly developed concepts in statistics are applied in the hydrological context shortly after publication. A good example here are copulas. Nevertheless, hydrological data also exhibit aspects that make statistical procedures more difficult to apply. The main problem when considering discharge series, especially maxima, is the limited period of observation. The construction of gauges and therefore the systematic conception of discharge data is in general not older than 100 years. Many gauges in Germany are in fact not observed for more than 30 years. Since the mostly considered flood series are annual maximum series, this results in sample lengths of about $n = 30$. Many statistical procedures based on asymptotic results fail here and many results are not valid for such small samples.
Additionally, we can find many different types of dependence in the data series. Depending on the time-resolution, discharge series can exhibit long-range dependence (e.g. some daily discharge series), short-range dependence (e.g. daily and monthly discharge series) as well as independence (annual maximum series). Moreover, in many discharge series extraordinarily large events occur that have a large influence on the estimation, especially in small samples. Also the different nature of the discharge series has to be taken into account. Whereas the annual as well as the monthly maximum series consist of flood peak measurements with the highest possible resolution, daily discharge series mostly consist of discharge means of a day. Additionally, anthropological and climatic changes (building of dams, heat periods) can lead to non-stationary behaviour, e.g. heteroscedasticity. All these aspects have to be taken into account when using flood statistics.

In this thesis we place our focus on robust estimation methods in the context of flood statistics. Robust estimation is expected to lead to more stable results that do not vary much over time. That is, single extraordinary discharge events shall not have much influence on the single estimations. Otherwise, the influence of these events in short time series would be much too large. Whereas the asymptotic distribution of robust estimators under independence is well known, few results exist under short-range dependence.
The first part of this thesis therefore develops asymptotic normality as well as a long-run variance estimator under a form of short-range dependence for the class of GL-statistics. These statistics cover many robust estimators, especially the linear as well as the trimmed linear moments, which are among the most commonly used estimators in hydrology. Hence, a concrete form of the limit distributions for these estimators is given. Since the theory is mainly based on the concept of U-statistics and U-processes, very general and in a broad statistical context applicable results are obtained. To take into account the heteroscedasticity of discharge data, we prove that the heteroscedastic model EGARCH exhibits the assumed concept of short-range dependence.
The results concerning robust estimation are then applied in the context of hydrology. Here, different concepts to estimate design floods are developed and their applicability is validated.

Concrete advantages of robust estimators in the hydrological context are shown and the combination of the new models with robust estimators is compared to classical hydrological models. Moreover, a method to detect coherences between gauges in a river basin based on the classification into alert steps is developed and applied to regionalisation where again the robustness of the estimators plays a crucial role.

2. Robust Estimation

The concept of robustness is used in many (statistical) disciplines, often with very different, sometimes even inconsistent, meaning. In general, there is no one-and-only valid definition for robustness and the use of this term often depends on the field of application or even the author. Even in this work robustness has a mathematical as well as an application-oriented meaning. Whereas in statistics robustness can be clearly measured in certain ways, in hydrology (respectively almost all engineering disciplines) robustness is understood as a kind of stability, which is a very vague explanation. Of course, also statisticians want to obtain stability by using robust methods, nevertheless, the terms are defined much more specific. In the following, we will show that still both definitions are compatible and that statistically robust estimators lead to more stable estimations in the hydrological context. Some of the considerations made here can also be found in Fischer et al. (2015) and Fischer and Schumann (2016).

2.1. Definition

Following Huber (1981) we define robustness as "insensitivity to small deviations from the assumptions". These deviations can be model misspecification or the wrong assumption on the convergence rate. In our context of hydrology and in general the most widely considered deviation is that of the shape of the underlying distribution respectively sample. Huber (1981) calls this "distributional robustness". Of course one could argue that this problem can be solved by simply removing outliers from the data, but in practice it becomes evident that it is not always clear, how to detect such outliers. Or it is not sensible to remove extreme values since they really occurred and are not erroneous data. This means they belong to the right tail of the underlying distribution but get too much weight in the estimation. Therefore, robustness is an important aspect when comparing estimators.
Now, even after giving the definition of robustness, there still exist many different aspects. For example Huber (1981) distinguishes between quantitative, qualitative or optimal robustness. For more details on this topic and a more explicit description we refer to Huber (1981). The robustness measures described later on in this chapter focus on different aspects of robustness respectively.

2.2. Robustness in Hydrology

The estimation of flood quantiles with very low exceedance probabilities is a key problem of engineering hydrology. Since the number of recorded series of floods is very limited and seldom longer than 100 years, the needed probabilities of extreme flood events are derived from a fitting of a suitable distribution function and its extrapolation into the realm of very low exceedance probabilities. The selection of the underlying statistical model is crucial in this context. Therefore, robustness is an important practical goal in flood statistics. It becomes evident when an extreme event with an exceedance probability significantly smaller than $1/n$ occurs within a time series of n years. Robustness in this context not only means robustness against extraordinary

extreme events but also against model misspecification or errors in the data. Whereas extraordinary extreme events are an important part of the flood series and contain important information this is not the case for data errors. This difference has to be considered. Robust estimators therefore shall not be used to cut off these extreme events, but they shall be used to reduce the influence of them especially in the presence of small sample lengths to gain stable estimates and to estimate the influence of single extraordinary extreme events.

One problem consists of the temporal variability of statistical characteristics of these series, which results from exceptional extreme events happening occasionally. Those events have a large influence on the estimation of the parameters of the distribution functions and their quantiles temporarily. The impact of this temporal variability is aggravated if the demand for design floods is increasing after disastrous flood events, which often results in step changes of the estimated parameters of distribution functions and of quantiles used as design criteria. In many cases, such changes are smoothed again by subsequent periods of "normal" floods. Figure 2.1 shows an example of step-wise changes of the 99 percent quantile derived from the generalized extreme value (GEV) distribution with probability-weighted moments and a year-by-year extended flood series. If these quantiles are applied for the design of long-lasting hydraulic structures, their temporal variability becomes a problem. There exist several approaches to handle such extraordinary events, causing step changes in hydrological parameters and quantiles. Often, these extraordinary events are treated as outliers and several statistical tests exist to detect such data points that deviate markedly from the remaining data points. By this, those values are treated like they are not drawn from the same population as the remainder data. For example, the Bulletin 17B, which is a benchmark in US-American hydrology, recommends such tests in a case of a skew larger than 0.4 (or smaller than -0.4 if we expect extraordinary low values) (Subcommittee (1981)). A very simple tool to define outliers is the 3-σ-rule, which defines outliers on the basis of the standard deviation (Jeong et al. (2017)). A well known test for these outliers is the maximum-value test of Grubbs (Grubbs (1969)). Here, it is assumed that the data follow approximately a normal distribution, but tests for the detection of outliers in data with other underlying distributions also exist (e.g. Spencer and McCuen (1996)). Also the tests by Dixon-Thompson, Rosner and Chauvenet are often recommended in the context of hydrology (see e.g. McCuen (2003)).

If an outlier is detected, still the question arises how to handle it. The outlier could be an erroneous value, which should be corrected or removed. Outliers in flood statistics may be the result of a mixed population occurrence (Klemeš (1986)). If we exclude these two possibilities, we can conclude that it is just an event from the tail of the distribution. Here we have two options: It could be censored to avoid distorting the analyses, or it could be weighted to reduce the resulting distortion.

The removal of conspicuous events has been the general handling of these values for a long time period (McCuen (2003)). Nevertheless, Gumbel already has remarked that "the rejection of outliers on a purely statistical basis is and remains a dangerous procedure. Its very existence may be a proof that the underlying population is, in reality, not what it was assumed to be" (Kruskal et al. (1960)). During the last 30 years a change of the handling of extraordinary events has begun and several authors and even pamphlets of the federate states in Germany now recommend the consideration of these in the statistics (e.g. Ashkar (1993) and DWA (2012)). The option of weighting of such events became very popular by using L-moments, which are rather robust to the effects of outliers (Hosking (1990)), or LH-moments (Wang (1997)), a generalization of L-moments, for characterizing the upper part of distributions and larger events in data.

There are other hydrological problems where a demand for robustness also exists, e.g. parameter

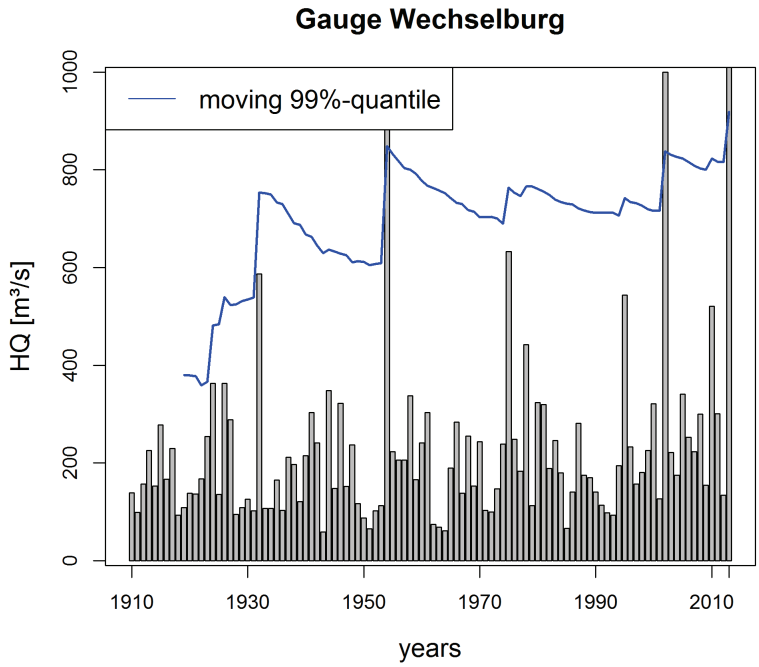

Figure 2.1.: Annual maximum discharges (HQ) for the Wechselburg/Zwickauer Mulde gauge in Saxony (1910-2013) and the estimated 99%-quantile for increasing sample length. A jump in the estimated high quantile can be seen every time an extraordinary large event occurs, leading to an unstable estimation over the years.

calibration procedures for deterministic hydrological models (Guerrero et al. (2013)). Bárdossy and Singh (2008) have specified four criteria for an estimation of parameter vectors of such models in the framework of a "data depth" of observation periods. The parameter vectors should:

- lead to good model performance over the selected time period
- lead to a hydrologically reasonable representation
- not be sensitive against the choice of the calibration period
- be transferable to other time periods.

The third and fourth criteria are especially suitable for the interpretation of robustness used in this research. Since the estimated quantiles for certain annual return periods like $T = 1000$ are used for the design of long-lasting hydraulic structures, it is not desirable that these parameters change much with any extension of the observed time series. From the hydrological point of view a robust estimation is preferable that can mirror the asymptotic behaviour (limit) of the estimated quantile of the AMS to an early point of observation without having these step-changes. In this context we want to focus on the interpretation of robustness as stability. That is, the estimation of extreme quantiles should not change significantly when adding or removing only a few values. This intention shall be emphasised by an introducing example. The estimation methods presented here are not of interest in the moment and will be explained later on. Instead, we want to give an outline of the idea of robustness needed in hydrology. As shown in Figure 2.1, where the 99%-quantile for a year-by-year prolonged series of maximum discharges at the Wechselburg gauge is estimated, the influence of single values on the estimation of high quantiles can be very large, especially when extraordinary large floods occur in very short time series. This instability leads to large problems if these quantiles are used as design floods. To emphasize the vulnerability of estimates of extreme quantiles in the presence of only a short period of observation we use a form of sensitivity curve to outline the influence of single (extraordinary) floods.
For this we take the whole series of annual maxima (AMS) at the Wechselburg gauge $(X_1, \dots, X_n)$, remove the respective annual maximum for every time step and replace it by the median of the whole sample to gain a new sample $\mathbf{X}_i' = (X_1, \dots, X_{i-1}, \text{med}(X_1, \dots, X_n), X_{i+1}, \dots, X_n)$ for $i = 1, \dots, n$. For these new samples a GEV or the Peak-over-threshold approach (POT) is fitted using linear moments (L-moments) or trimmed L-moments (TL-moments) and the quantiles for the annualities $T = 200, 500, 1000$ are calculated:

$$q_{T;i} = G^{-1}_{\mathbf{X}_i'}(T),$$

where $G^{-1}_{\mathbf{X}_i'}$ is the quantile function of the GEV fitted to the sample $\mathbf{X}_i'$. More details on the Peak-over-threshold approach as well as L-moments and TL-moments can be found later on in Sections 6.1.2 and 7.2.1. In this context they should only serve as examples of robust methods. We finally take the difference of $q_{T;i}$ and q_T (the quantile based on the whole original sample) and multiply it with $n + 1$. This is analogous to the classical sensitivity curve introduced in the next section, where instead of replacing one value by the median it is replaced by "one-wild" (Tukey (1960)) to test the sensitivity of an estimator. Here, the "one-wild" observation is a real observation, whereas in statistical simulations often a function of a real-valued variable is used and the deviation for increasing values of this variable is measured. The use of the observations here should emphasise the problem of estimating a design flood, instead.
The results can be found in Figure 2.2, where the year displayed on the x-axis marks the year of the replaced observation.

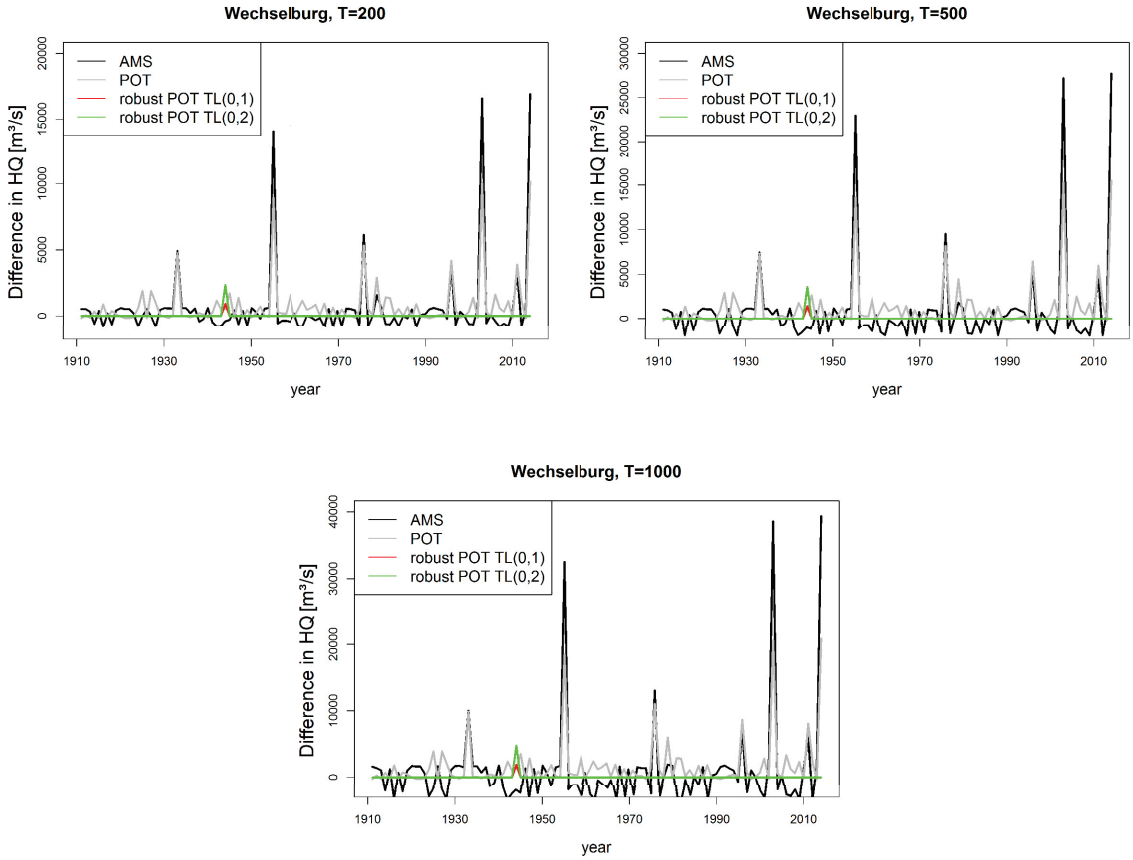

Figure 2.2.: Influence of single annual maxima on the estimation of the quantile with annuality T for different estimators and models by calculating the difference to the quantiles based on the series without this annual maximum (sensitivity). For the non-robust estimators in AMS and POT-model the estimation is influenced a lot by single events. When using robust estimators like the TL-moments the influence of this single events is reduced to almost zero leading to a stable estimation.

It becomes obvious that the extraordinary floods in the years 1954, 2002, and 2013 have a very high influence on the estimation of the quantiles when using classical estimators like the L-moments. This influence increases with increasing annuality. Therefore, the use of such quantile estimators in the design of dams can lead to serious problems since it is highly unstable. Robust estimation approaches like the robust POT reduce this instability and are a noteworthy alternative. Additionally, they can be used to estimate the influence of single events.
Although especially extreme events are of interest in hydrology, since they are the ones causing highest damage, the considerations made above show that also the use of robust estimators in hydrology is of considerable interest.
Additionally, it is not always clear, which kind of distribution function one should use, two- or three-parametric. Whereas the three-parametric distribution function allows a greater flexibility in modelling the tails it is also more uncertain, especially when estimating the shape parameter. Recommendations are for example given by the DWA (DWA (2012)), recommending a two-parametric distribution function for samples with less than 30 years and a three-parametric one for samples with more than 50 years. A distinct recommendation for samples of 30-50 years is not given. A robust estimator used in the hydrological context should therefore be also insensitive against small deviations from the model. Moreover, besides the GEV distribution there exist a lot of other distributions used in hydrology to model floods, for example the PearsonIII-distribution or the Gamma-distribution. Nevertheless, in the context of our considered data samples the GEV distribution is the most commonly used. More details on this context can be found in Section 7.2.1.
All these aspects play a crucial role in flood statistics and should have influence on the used methods and estimators. Robustness could lead to an improvement of the consideration of uncertainty in this point.

2.3. Measures of Robustness

There exist several possibilities to measure robustness. All of them focus on different aspects of the definition made above. Additionally, some of the measures have been developed because of the special challenges in their field of application. We want to define two of the most common statistical measures of robustness as well as one measure that has its origin in hydrology, focussing especially on the right tail. All of them will be used later on to emphasise the robustness of certain models or estimators.

2.3.1. Influence Curve

As mentioned above, one important aspect of robustness is the insensitivity of the estimation against single (extreme) values. This aspect somehow coincides with the hydrological point of view concerning stability: we do not want to obtain large deviations in the estimation if one extraordinary event occurs. The limit of the influence of a single observation x on an estimate $T(F_n)$ of F (just think about random variable X with distribution function F and a sample $x_1, \ldots, x_n$ with empirical distribution function F_n) can be expressed by (see Hampel (1971))

$$IC(x; F, T) = \lim_{\lambda \to 0} \frac{T((1-\lambda)F + \lambda\delta_x) - T(F)}{\lambda},$$

where δ_x denotes the point mass 1 at x. This is the so called **influence curve**, which can be shown to have several interesting properties (see Section 5). In fact it is the first derivative of

the estimator T evaluated at the perturbation of F by δ_x. For a robust estimator we of course want to have a bounded influence curve indicating that also the influence of single observations on the estimate is bounded.

Example 2.1. Assume we have i.i.d random variables $X_1, \ldots, X_n$. The sample mean

$$\hat{X} = \mu(F_n) = \int_{-\infty}^{\infty} x dF_n(x) = \frac{1}{n}\sum_{i=1}^{n} X_i$$

is an estimator for the expected value $\mu(f) = \int_{-\infty}^{\infty} x dF(x)$. It has influence function

$$\begin{aligned} IC(x; F, \mu) &= \lim_{\lambda \to 0} \frac{\mu((1-\lambda)F + \lambda\delta_x) - \mu(F)}{\lambda} \\ &= \frac{d}{d\lambda} \int_{-\infty}^{\infty} z d[(1-\lambda)F + \lambda\delta_x]|_{\lambda=0} = \frac{d}{d\lambda}\left((1-\lambda)\mu(F) + \lambda x\right)|_{\lambda=0} \\ &= x - \mu(F), \end{aligned}$$

which increases unboundedly with increasing x. Thus, the sample mean is not robust.

There also exist several sample versions of the influence curve, where we want to focus on the one proposed by Tukey (1960), the so-called **sensitivity curve**

$$SC_{n-1}(x) = \frac{T(\frac{n-1}{n}F_{n-1} + \frac{1}{n}\delta_x) - T(F_{n-1})}{\frac{1}{n}},$$

where we simply replaced F by F_{n-1} and λ by $\frac{1}{n}$.

2.3.2. Breakdown Point

Often, not only one extreme value occurs and in this case the knowledge of the influence of a single value is not helpful. Here, we are interested in the behaviour of the estimate under the occurrence of many extreme values. The asymptotic **breakdown point** ϵ^* of an estimator $T(F_n)$ of the functional $T(F)$ is defined as

$$\epsilon^*(T, F, d) = \sup_{\epsilon<1}\{\epsilon : \sup_{F: d(F,F_0)<\epsilon} |T(F) - T(F_0)| < \infty\}.$$

It characterises the maximum deviation from the true F_0 for a chosen metric d. For a finite sample $\Omega = X_1, \ldots, X_n$ the sample breakdown point is then defined as

$$\epsilon_n^*(T) = \frac{1}{n}\max\left\{m : \sup_{\Omega_m} \|T(\Omega_m)\| < \infty\right\},$$

where Ω_m is a sample derived from Ω by replacing any m values of Ω with arbitrary ones;
It gives us information on how many outliers can occur until the estimator collapses. A breakdown point of 0 indicates a totally non-robust estimator, whereas equivariant robust estimators can reach a breakdown point of 50%. For example, the sample mean has breakdown point 0, whereas the median has breakdown point 0.5.

2.3.3. Stability Index

The breakdown point or the influence or sensitivity curve are the most frequently used measures for robustness. However, these measures do not consider the special properties of hydrological data. When using flood series, the quantity of available data is very limited, and the asymptotic behaviour of the mathematical procedures are not effective. The special form of the applied models, in which the estimated parameters have an exponential influence on the resulting quantile, leads to large deviations in the results, even if the differences between parameter estimations are small. Therefore, it is not sufficient to check only the parameter estimators for their robustness, but the applied statistical model as a whole plays a crucial role. Hence, we use hydrological measures of stability of quantiles. Typical of most hydrological assessments of stability is the comparison of different calibration (in our case: modelling) and validation subperiods (cf. Brigode et al. (2013)). For stability of quantiles, the criterion $SPAN_T$ measuring the variability (span) of the estimation is used, which has been proposed by Garavaglia et al. (2011) and applied to compare the robustness of fitting methods (Kochanek et al. (2014), Renard et al. (2013)).
The value of $SPAN_T$ for a quantile of the annual return period T at a given site l is calculated by

$$\mathrm{SPAN}_T(l) = \frac{\max\limits_{1\le s\le b}\{\hat{q}_{T;l}(s)\} - \min\limits_{1\le s\le b}\{\hat{q}_{T;l}(s)\}}{\frac{1}{b}\sum\limits_{s=1}^{b}\hat{q}_{T;l}(s)},$$

where $\hat{q}_{T;l}(s)$ is the estimated quantile related to the return period T for one of b non-overlapping subperiods $s = 1, \ldots, b$ at the gauge l. The optimal value of $SPAN_T$, indicating a robust, stationary behaviour of the statistical model, is 0. To compare the $SPAN_T$ for several gauges at the same time, the empirical distribution can be considered for all l. Since in our case the sample length is very limited and the robust estimators need a certain quantity of data, we have to reduce the quantity of subperiods to two, choosing one with a length of 50 years. The $SPAN_T$ criterion can also be applied to compare quantiles of two parts of a time series s_1 and s_2 as follows (Renard et al. (2013))

$$SPAN_T(l) = \frac{|\hat{q}_{T;l}(s_1) - \hat{q}_{T;l}(s_2)|}{\frac{1}{2}\left(\hat{q}_{T;l}(s_1) + \hat{q}_{T;l}(s_2)\right)},$$

where $\hat{q}_T(s_i)$, $i = 1, 2$, is the estimated quantiles related to the annual return period T for subperiods s_1 and s_2 respectively at the gauge l.

In contrast to the two above-mentioned measures, which are well-known in statistical theory, the $SPAN$-criterion is mainly used in the hydrological context. This is not only due to the fact that it measures the stability instead of the influence of one value, but also because we can lay the focus on high quantiles by choosing appropriate T. Because of the comparison of two subsamples it is also possible to compare the influence of two or more values on the estimation. Having in mind the often frequent appearing extraordinary events in hydrology, this is a desirable property. By using the representation by the distribution of $SPAN$ of several gauges it is also possible to detect salience of single gauges. Hence, this measure will be used here especially in the context of hydrological flood series to take into account their special nature. Nevertheless, it is comparable to other statistical measures for robustness.

3. Concepts of Short-Range Dependence

For several years, the concept of independent and identically distributed data has been the common assumption in hydrological statistics. And not only in hydrology, but also in many other applications independence has been assumed. Nevertheless, it is questionable whether discharge series, especially of high time-resolution, are really independent. As an example we take the monthly maximum discharges at the Wechselburg gauge in Germany, see Figure 3.1.

The Wechselburg gauge at the river Zwickauer Mulde belongs to the Mulde river basin located in Saxony in South-East Germany. The time series may look independent, but the autocorrelation function shows a different picture (Figure 3.2). We can see a significant deviation from the confidence bands based on White Noise and therefore from independence. Thus, one can assume a certain dependence in the data. Please note that both discharge series are not related directly, since the maximum values are peak measurements.
If one accepts the presence of dependence in the data, the question arises, which kind of dependence is present.
On the left hand side of Figure 3.2 one can see a fast decay of the autocorrelation function, whereas on the right hand side there is only a slow decay. Nevertheless, the same gauge is considered, only the type of discharges (monthly maxima and daily means) is different.

In general, most of the considered flood series in hydrology can be assumed to be independent or short-range dependent. Moreover, to detect long-range dependent behaviour, the time series considered here are not long enough. In the following we will therefore concentrate on the concept of short-range dependence.
There exist several definitions of different forms of short-range dependence. One of the most common ways to define short-range dependence is by mixing processes.
Bradley (2007) gives an overview over the different forms of mixing. We will consider the case of absolutely regular sequences of random variables.

Definition 3.1. *Let $\mathcal{A}, \mathcal{B} \subset \mathcal{F}$ be two σ-fields on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The absolute regularity coefficient of $\mathcal{A}$ and $\mathcal{B}$ is given by*

$$\beta(\mathcal{A}, \mathcal{B}) = \mathbb{E} \sup_{A \in \mathcal{A}} |\mathbb{P}(A|\mathcal{B}) - \mathbb{P}(A)| .$$

If $(X_n)_{n\in\mathbb{N}}$ is a stationary process, then the absolute regularity coefficients of $(X_n)_{n\in\mathbb{N}}$ are given by

$$\beta(l) = \sup_{n\in\mathbb{N}} \beta(\mathcal{F}_1^n, \mathcal{F}_{n+l}^\infty).$$

$(X_n)_{n\in\mathbb{N}}$ is called absolutely regular, if $\beta(l) \to 0$ as $l \to \infty$.

Absolutely regular random variables are sometimes also called β-mixing and have been introduced by Volkonskii and Rozanov (1959). β-mixing is a stronger assumption than for example α-mixing, since for the α-mixing coefficients α given by

$$\alpha(l) = \sup_{n\in\mathbb{N}} \sup \left\{ |\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)| : A \in \mathcal{F}_1^n, B \in \mathcal{F}_{n+l}^\infty \right\},$$

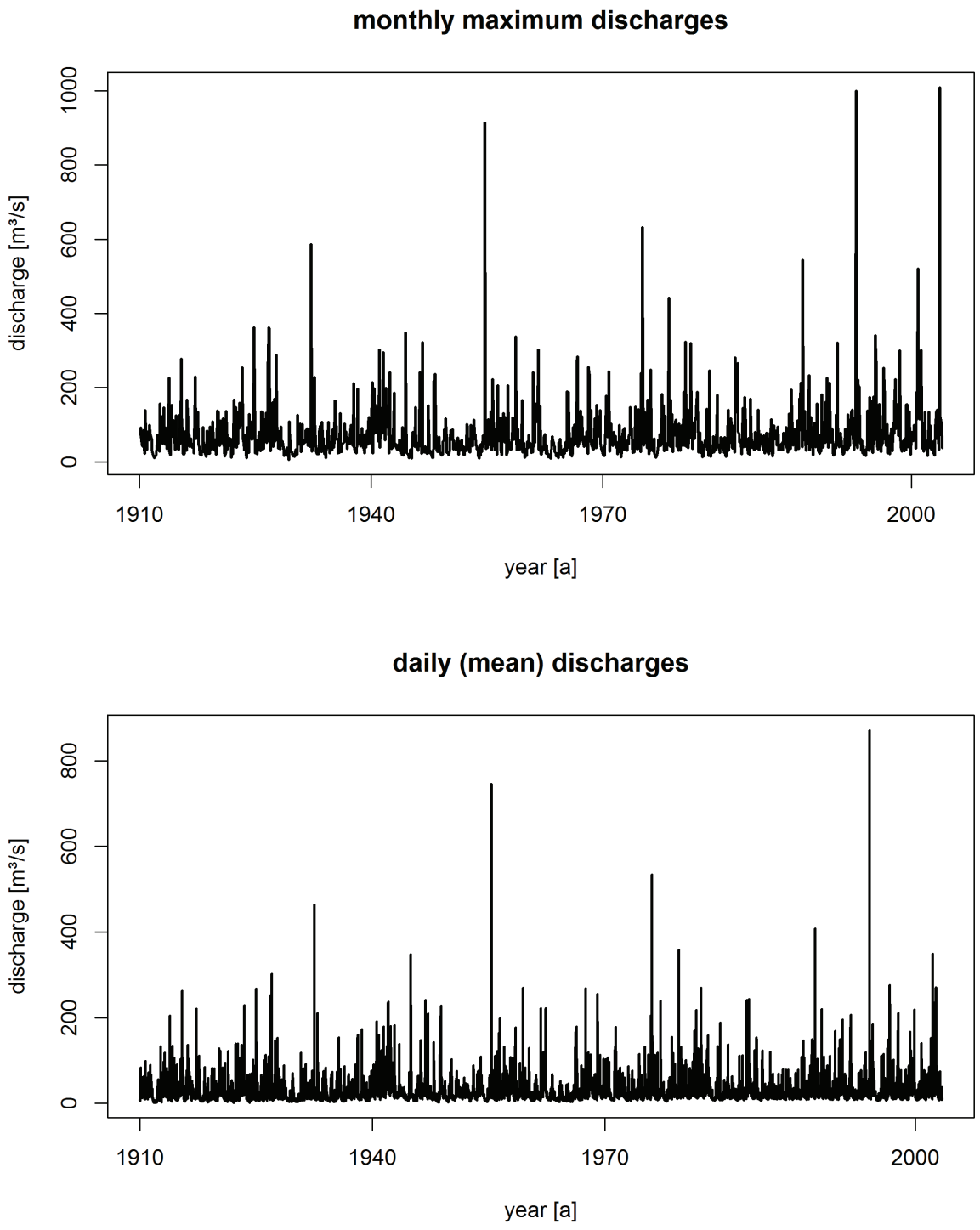

Figure 3.1.: Monthly maximum discharges (top) and daily means (bottom) of the Wechselburg/Zwickauer Mulde gauge. The difference between the peak values and means becomes evident.

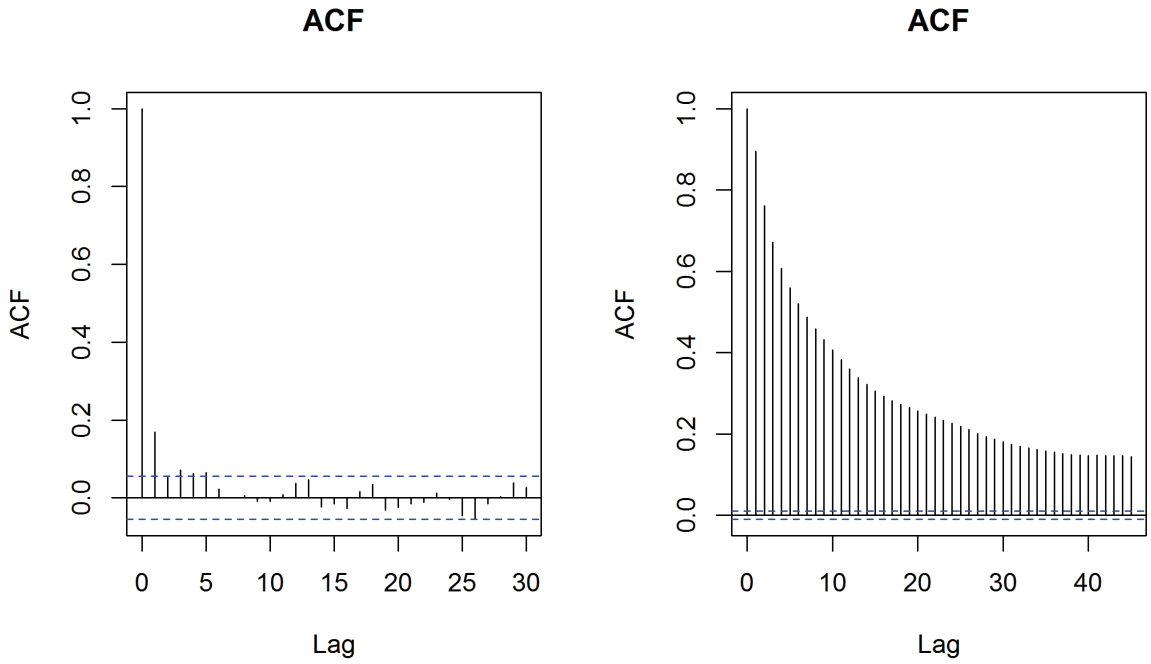

Figure 3.2.: Autocorrelations of the monthly maximum discharges (left) and daily discharges (right) of the Wechselburg/Zwickauer Mulde gauge. The daily discharges with the higher time-resolution show a much stronger dependence.

it holds that $\alpha(l) \leq \frac{1}{2}\beta(l)$, so that every absolutely regular process is likewise strong mixing. Though, β-mixing is weaker than Ψ- or Φ-mixing.

Typical examples for such absolutely regular processes are certain Markov chains or certain AR-processes. Nevertheless, more complex models like dynamical systems cannot be modelled by this concept of short-range dependence. For example Andrews (1984) has shown that even an $AR(1)$ process of independent Bernoulli innovations is no longer α-mixing since one can construct sets that are determined by the future process, no matter how far away. Gorodetskii (1978) even has been able to show that there exist linear processes with normal distributed innovations, whose coefficients decline too slowly, such that they are no longer mixing.
To cover all these processes, the so called near epoch dependence has been developed. It is based on the idea that although a random variable $X_t = f(\ldots, Z_{t-1}, Z_t, Z_{t+1}, \ldots)$, which is a functional of a mixing sequence, is not necessarily mixing it depends on the near epoch of Z_t. Therefore, some properties can be concluded, especially the validity of limit theorems.

Definition 3.2 (Near Epoch Dependence (NED))**.**
Let $((X_n, Z_n))_{n\in\mathbb{Z}}$ be a stationary process. $(X_n)_{n\in\mathbb{N}}$ is called L_1 near epoch dependent (NED) on the process $(Z_n)_{n\in\mathbb{Z}}$ with approximation constants $(a_l)_{l\in\mathbb{N}}$, if

$$\mathbb{E}\left|X_1 - \mathbb{E}\left(X_1|\mathcal{G}_{-l}^{l}\right)\right| \leq a_l \qquad l = 0, 1, 2, \ldots,$$

where $\lim_{l\to\infty} a_l = 0$ and $\mathcal{G}_{-l}^{l}$ is the σ-field generated by $Z_{-l}, \ldots, Z_l$.

Near epoch dependent processes are sometimes described with the term *approximating functionals*. One of the first applications of such kind of short-range dependent processes can be found in Ibragimov (1962). In the literature one often also finds L_2 or in general L_p near epoch dependence, where the L_1 norm is simply changed to the L_p norm, or the weaker form of P-NED (Dehling et al. (2016); Vogel and Wendler (2015)), which allows to consider processes with existing moments of lower order. The main difference between the different definitions of near epoch dependence are their assumptions on the existing moments.
The concept of near epoch dependence is especially useful in the case of an underlying mixing sequence, since in this case very helpful properties are inherited. More details on this and a detailed introduction to near epoch dependence can be found in Davidson (2002).
For the two examples given above and also given in Andrews (1984) and Gorodetskii (1978) Jenish and Prucha (2012) show that they are near epoch dependent.

3.1. Examples

A typical example of a model for short-range dependent data is a special case of the ARIMA(p,d,q)-model, which is an abbreviation for Auto-Regressive Integrated Moving Average. As indicated by the name it consists of an AR-part of order p as well as an MA-part of order q.

Definition 3.3. *A process $(X_t)_{t\in\mathbb{Z}}$ is called ARIMA(p,d,q)-process if*

$$X_t = \phi^{-1}(B)\theta(B)(1-B)^{-d}Z_t,$$

where $(Z_t)_{t\in\mathbb{Z}}$ is a White Noise series and d is an integer. The polynomials

$$\phi(z) = 1 - \phi_1 z - \ldots - \phi_p z^p$$

$$\theta(z) = 1 + \theta_1 z + \ldots + \theta_q z^q$$

have no common zeros and ϕ has no roots on the unit circle. The operator B is the so-called Backshift Operator defined by $BZ_t = Z_{t-1}$.

The parameter d is the integration parameter. It gives the times of differentiation needed to obtain a stationary time series.
A stationary ARIMA-series, that is $d = 0$, is strongly mixing.

Some of the widely used models when considering near epoch dependent processes are GARCH-processes (Generalized Autoregressive Conditional Heteroscedasticity) (Bollerslev (1986)), a generalisation of ARCH-processes. They are a common model for volatility clustering in financial data and are also used for example in hydrology (Wang et al. (2012)).

Definition 3.4. *A process $(X_t)_{t\in\mathbb{Z}}$ is called GARCH(p,q)-process, if*

$$X_t = \sigma_t Z_t,$$

where σ_t^2 is the positive conditional variance given by

$$\sigma_t^2 = \alpha_0 + \alpha_1 Z_{t-1}^2 + \ldots + \alpha_p Z_{t-p}^2 + \beta_1 \sigma_{t-1}^2 + \ldots + \beta_q \sigma_{t-q}^2,$$

where $\alpha_0, \ldots, \alpha_p, \beta_1, \ldots, \beta_q \in \mathbb{R}$ are non-negative with $\alpha_p \neq 0$ and $\beta_q \neq 0$ and $(Z_t)_{t\in\mathbb{Z}}$ is an i.i.d. sequence with mean zero and variance equal to one.

Hansen (1991) relaxes the assumptions on $(Z_t)_{t\in\mathbb{Z}}$, such that $(Z_t)_{t\in\mathbb{Z}}$ can be assumed to be α-mixing. He showed then that if $\left(\mathbb{E}[(\beta_1 + \alpha_1(\frac{X_t}{\sigma_t})^2)^r|\mathcal{F}_{t-1}]\right)^{1/5} \leq c < 1$ a.s. for all t, a GARCH(1,1)-process X_t is L_r-NED on the α-mixing process Z_t with approximation constants $a_l = c^l 2\alpha_0 c/(1-c)$ and $\mathcal{F}_t = \sigma(\ldots, Z_t)$.
An extension of the GARCH-model is the exponential GARCH (EGARCH) model proposed by Nelson (1991).

Definition 3.5. *The process $(X_t)_{t\in\mathbb{Z}}$ is called EGARCH(p,q)-process on the sequence $(Z_t)_{t\in\mathbb{Z}}$, if*

$$X_t = \sigma_t Z_t,$$

where σ_t^2 is the positive conditional variance given by

$$\log(\sigma_t^2) = \alpha_0 + \alpha_1 f(Z_{t-1}) + \ldots + \alpha_p f(Z_{t-p}) + \beta_1 \log(\sigma_{t-1}^2) + \ldots + \beta_q \log(\sigma_{t-q}^2),$$

where $\alpha_0, \ldots, \alpha_p, \beta_1, \ldots, \beta_q \in \mathbb{R}$ with $\alpha_p \neq 0$ and $\beta_q \neq 0$ and f is a measurable function which is linear in Z and given by

$$f(Z_t) = \theta Z_t + \lambda(|Z_t| - \mathbb{E}|Z_t|)$$

with parameters $\theta, \lambda \in \mathbb{R}$.

The term $\lambda(|Z_t| - \mathbb{E}|Z_t|)$ determines the size effect whereas θZ_t determines the sign effect of the shocks on volatility. It can be seen that $\mathbb{E}(f(Z_t)) = 0$.
One of the advantages of EGARCH-processes is that they do not have the non-negativity restriction of the GARCH-processes. We show that under similar assumptions as for the GARCH-process an EGARCH-process is near epoch dependent.

Theorem 3.1. *Let σ_1 be bounded and $\sum_{i=1}^{q} |\beta_i| < 1$. Moreover, assume that*

$$\sup_{t\in\mathbb{Z}} \mathbb{E}|Z_t| < \infty. \tag{3.1}$$

Then the EGARCH(p,q)-process on the sequence $(Z_t)_{t\in\mathbb{Z}}$ given by $X_t = \sigma_t Z_t$ is near epoch dependent.

Remark 3.1.

1. The assumption (3.1) in Theorem 3.1 is an analogue to the condition of Hansen (1991) for GARCH-processes to the EGARCH-case with arbitrary values p and q.

   Whether this condition is fulfilled depends on the existing moments of Z_i. For example, if $(Z_t)_{t\in\mathbb{Z}}$ is a White Noise process with variance σ (that is $\mathbb{E}|Z_t| \leq \sigma = 1$), the condition is fulfilled.

2. The boundedness of the conditional variance σ_1 is a common assumption for GARCH-processes (see Hansen (1991), Lee and Hansen (1994)). It results from the moment condition on σ, $\mathbb{E}|\sigma_1|^{1+\delta} < \infty$, which is needed in the following proof, and the Lipschitz-condition.

Proof. (Theorem 3.1)
Using an iterative expression of the term $\log(\sigma_t^2)$ we obtain

$$\log(\sigma_t^2) = \sum_{j=1}^{n} \sum_{\substack{k_1,\ldots,k_q\in\mathbb{N}_0,\\ k_1+\ldots+k_q=j-1}} \binom{j-1}{k_1,\ldots,k_q} \beta_1^{k_1}\cdot\ldots\cdot\beta_q^{k_q} \left(\alpha_0 + \sum_{k=1}^{p} \alpha_k f\left(Z_{t-k-(\sum_{i=1}^{q} ik_i)}\right)\right)$$
$$+ \sum_{\substack{k_1,\ldots,k_q\in\mathbb{N}_0,\\ k_1+\ldots+k_q=n}} \binom{n}{k_1,\ldots,k_q} \beta_1^{k_1}\cdot\ldots\cdot\beta_q^{k_q} \log\left(\sigma^2_{t-(\sum_{i=1}^{q} ik_i)}\right).$$

Now, considering the limit for $n\to\infty$, it is

$$\begin{aligned}
&\log(\sigma_t^2)\\
&= \lim_{n\to\infty} \sum_{j=1}^{n} \sum_{\substack{k_1,\ldots,k_q\in\mathbb{N}_0,\\ k_1+\ldots+k_q=j-1}} \binom{j-1}{k_1,\ldots,k_q} \beta_1^{k_1}\cdot\ldots\cdot\beta_q^{k_q} \left(\alpha_0 + \sum_{k=1}^{p} \alpha_k f\left(Z_{t-k-(\sum_{i=1}^{q} ik_i)}\right)\right)\\
&+ \lim_{n\to\infty} \sum_{\substack{k_1,\ldots,k_q\in\mathbb{N}_0,\\ k_1+\ldots+k_q=n}} \binom{n}{k_1,\ldots,k_q} \beta_1^{k_1}\cdot\ldots\cdot\beta_q^{k_q} \log\left(\sigma^2_{t-(\sum_{i=1}^{q} ik_i)}\right).
\end{aligned}$$

We want to show that the first term of the sum converges a.s. This is gained by the assumptions $\sup_t \mathbb{E}|Z_t| < \infty$ and $|\sum_{i=1}^{q} \beta_i| < 1$ and the linearity of the function f. With the Multinomial Theorem and the convergence of the geometric series we can apply the monotone convergence theorem to obtain the convergence of the series (see for example Proposition 3.1.1 of Brockwell and Davis (2006)).
For the second term we show that it converges to zero a.s., that is

$$\lim_{n\to\infty} \sum_{\substack{k_1,\ldots,k_q\in\mathbb{N}_0,\\ k_1+\ldots+k_q=n}} \binom{n}{k_1,\ldots,k_q} \beta_1^{k_1}\cdot\ldots\cdot\beta_q^{k_q} \log\left(\sigma^2_{t-(\sum_{i=1}^{q} ik_i)}\right) = 0 \quad \text{for all } t\in\mathbb{Z} \text{ a.s.}$$

By using the Multinomial Theorem we have

$$\begin{aligned}
&\lim_{n\to\infty} \sum_{\substack{k_1,\ldots,k_q\in\mathbb{N}_0,\\ k_1+\ldots+k_q=n}} \binom{n}{k_1,\ldots,k_q} \beta_1^{k_1}\cdot\ldots\cdot\beta_q^{k_q} \log\left(\sigma^2_{t-(\sum_{i=1}^q ik_i)}\right)\\
&\leq \sup_{k_1,\ldots,k_q} \log\left(\sigma^2_{t-(\sum_{i=1}^q ik_i)}\right) \lim_{n\to\infty} \sum_{\substack{k_1,\ldots,k_q\in\mathbb{N}_0,\\ k_1+\ldots+k_q=n}} \binom{n}{k_1,\ldots,k_q} |\beta_1|^{k_1}\cdot\ldots\cdot|\beta_q|^{k_q}\\
&= \sup_{k_1,\ldots,k_q} \log\left(\sigma^2_{t-(\sum_{i=1}^q ik_i)}\right) \lim_{n\to\infty} (|\beta_1|+\ldots+|\beta_q|)^n
\end{aligned}$$

and therefore the term converges a.s. to zero if

$$\sum_{i=1}^q |\beta_i| < 1.$$

Hence, we can write

$$\log(\sigma_t^2) = \sum_{j=1}^{\infty} \sum_{\substack{k_1,\ldots,k_q\in\mathbb{N}_0,\\ k_1+\ldots+k_q=j-1}} \binom{j-1}{k_1,\ldots,k_q} \beta_1^{k_1}\cdot\ldots\cdot\beta_q^{k_q} \left(\alpha_0 + \sum_{k=1}^p \alpha_k f\left(Z_{t-k-(\sum_{i=1}^q ik_i)}\right)\right).$$

This is a linear solution and for this reason the process $(\log(\sigma_t^2))_{t\in\mathbb{Z}}$ is near epoch dependent. Moreover,

$$\sigma_t = \sqrt{\exp(\log(\sigma_t^2))} = g(\log(\sigma_t^2))$$

with $g(x) = \sqrt{\exp(x)}$. This function g fulfils the Lipschitz-condition for all $x \in (-\infty, a]$, $a \in \mathbb{R}$. We can now apply Proposition 2.11 of Borovkova et al. (2001), where we need that σ_1 is bounded. Therefore, the process σ_t and hence $X_t = \sigma_t Z_t$ is near epoch dependent on the process $(Z_t)_{t\in\mathbb{Z}}$. □

3.2. Short-Range Dependence in Hydrology

Many time series in hydrology show a heteroscedastic behaviour. This can be caused by changing climate conditions but also by anthropogenic changes or other effects. More generally, almost all hydrological runoff-models assume the residuals to be heteroscedastic. More precisely it is assumed that for small discharges only small deviations in the simulation can occur, whereas for large discharges also large deviations can occur. Therefore, this behaviour can not fully be modelled by classical ARIMA-models. For example, Modarres and Ouarda (2013a) show that when using only an ARIMA-model for heteroscedastic data the residuals remain heteroscedastic. Now there are two possible solutions for this problem. The first one is the use of a Box-Cox or similar transformation before applying the ARIMA-model. On the other hand one can also apply a heteroscedastic model to the residuals and obtain for example an ARIMA-GARCH model. In hydrological time series it is often not sufficient to use a Box-Cox transformation only (Modarres and Ouarda (2013b)). Additionally, often the use of a model to describe the behaviour of the residuals is preferable because of the possible use of additional information (Evin et al. (2013)).

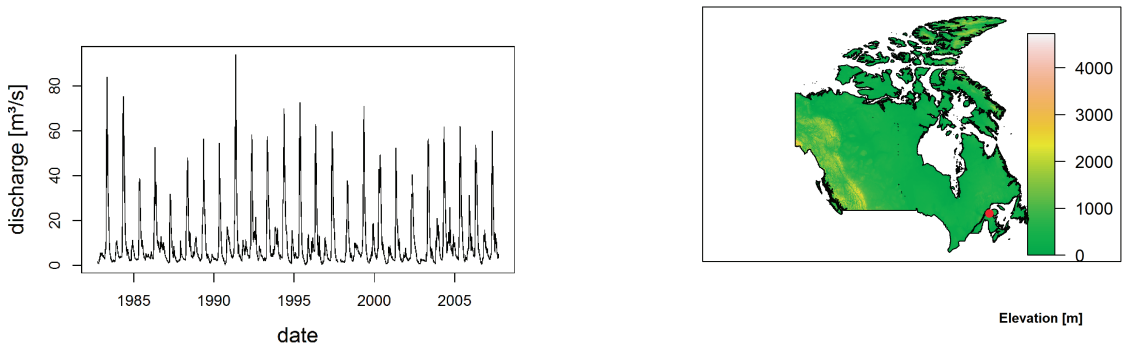

Figure 3.3.: Daily discharges [m^3/s] (left) and location of the considered gauge of the Matapedia river near Quebec, Canada (right).

Hence, a model is needed which takes into account this heteroscedasticity. In the considered case, the EGARCH-model proved to be superior to the other models (Modarres and Ouarda (2013a)).
We want to seize the data example of Modarres and Ouarda (2013a) and use it to apply the developed theory under short-range dependence later on. The observed data are daily discharges from a gauge of the Matapedia river near the basin Amqui in South-Eastern Canada with a catchment area of 558 km^2 (Figure 3.3).
The autocorrelation of this series shows significant dependence within the data (Figure 3.4).
To the logarithmised data an ARIMA(13,1,4)-model is fitted. The same order has also been chosen by Modarres and Ouarda (2013a) and to make the results comparable we adopt this parametrisation. The logarithmisation as well as the differentiation with $d = 1$ has been chosen to reduce the high persistence of the data.
When we have a look at the residuals of this model applied to the data we see a heteroscedastic behaviour (Figure 3.5). Therefore, Modarres and Ouarda (2013a) apply the Engle-test to test on autoregressive heteroscedastic behaviour (Engle (1982)). For all lags the p-value is almost zero and therefore a significant heteroscedastic behaviour is found (Figure 3.5).
The results stay the same when a Box-Cox-transformation is applied to the data (logarithmic or original) before fitting the ARIMA-model. Hence, a model is needed that can cope with this kind of behaviour. Modarres and Ouarda (2013a) try different kinds of heteroscedastic models (GARCH, Power Garch) but the EGARCH model covers the behaviour best. The parameters are chosen as $p = 3$ and $q = 1$ for the EGARCH-model. If we compare the observed residuals and the ones modelled by the EGARCH(3,1)-model in a QQ-plot we observe a very good fit (Figure 3.6) and also the results of the applied Goodness of Fit test (Vlaar and Palm (1993)) confirm this.
This data example emphasizes the necessity of complex dependence models like the EGARCH model which are not covered by the classical theory of dependent random variables using mixing assumptions.

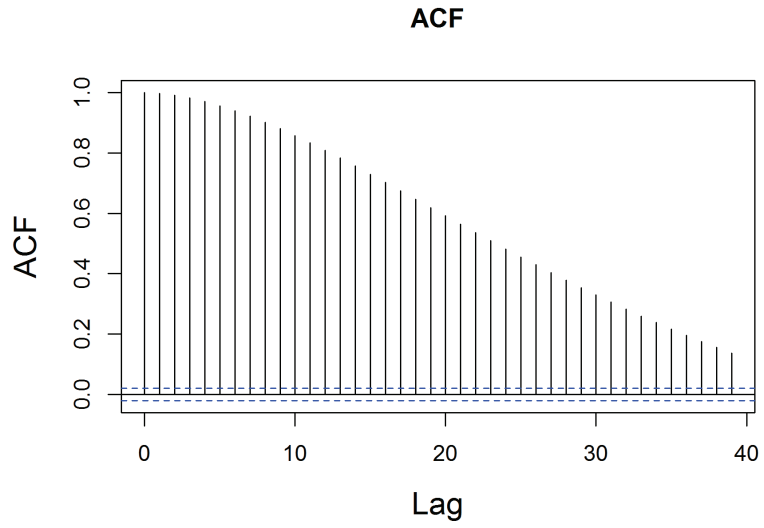

Figure 3.4.: Autocorrelationfunction (ACF) of the daily discharges [m^3/s] of the considered gauge of the Matapedia river near Quebec. A strong dependence of the data can be seen.

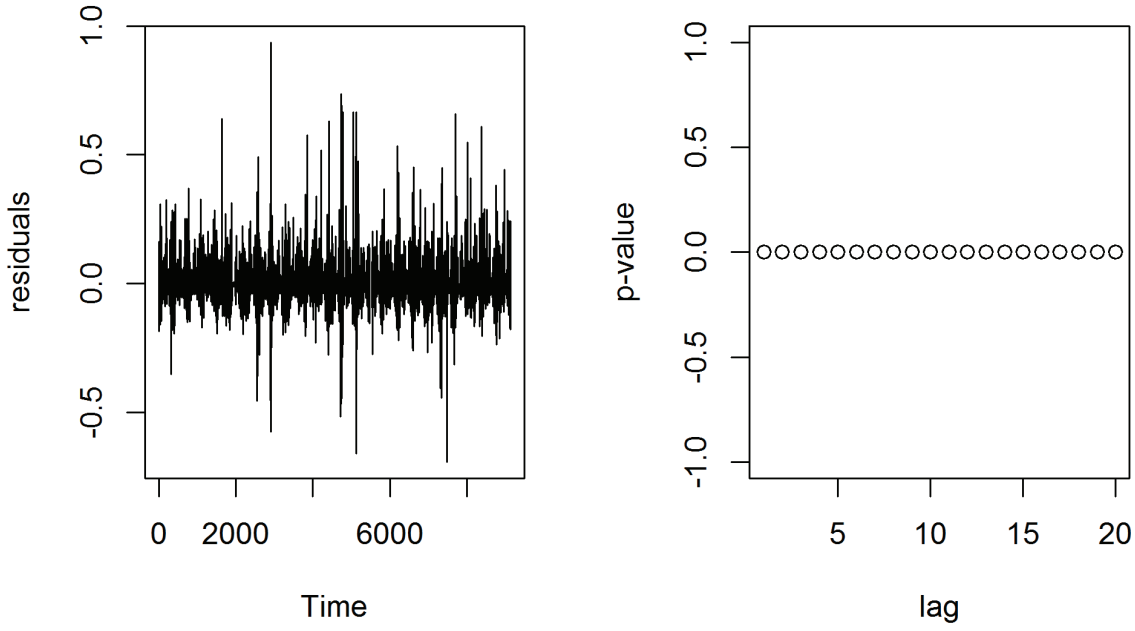

Figure 3.5.: Residuals resulting from the fitted ARIMA(13,1,4)-model (left) and p-values of the Engle-test (right) of the considered gauge of the Matapedia river near Quebec. This indicates a GARCH-behaviour remaining in the residuals.

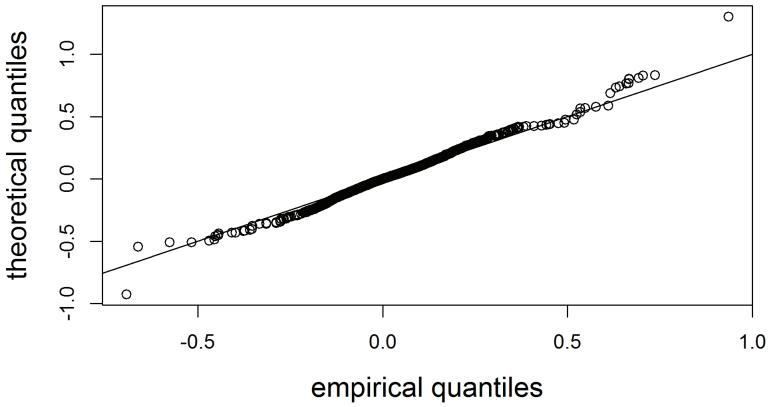

Figure 3.6.: QQ-plot of the observed residuals from the ARIMA(13,1,4)-model and theoretical residuals modelled with the EGARCH(3,1)-model of the considered gauge of the Matapedia river near Quebec. The model fits well to the data.

4. U-statistics, U-processes and U-quantiles

A central tool to cope with GL-statistics T_n, which are of particular interest in this thesis, are U-statistics and U-processes, since a representation of the error term $\sqrt{n}(T_n(x) - T(x))$ can be represented (under some conditions) via the first-order Gâteaux-differential, which is in fact a U-statistic. Besides, U-statistics are a common tool to prove asymptotic results of test-statistics as well as estimators since they are able to represent most of them. The simple form of expression together with some well known decomposition results make U-statistics and U-processes applicable in many situations. U-quantiles are also an often used class to represent quantile-based estimators.
We want to show three fundamental results:

1. a Central Limit Theorem for multivariate U-statistics
2. an invariance principle for multivariate U-processes
3. the convergence rate of the remaining term of the generalized Bahadur representation.

Let us first state some basic assumptions.
Let $X_1, \dots, X_n$ be a sequence of random variables with distribution function F. As mentioned above, here we will assume the random variables to be short-range dependent. Moreover, let F_n be the empirical distribution function of $X_1, \dots, X_n$, that is

$$F_n(x) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}_{[X_i \leq x]}, \quad -\infty < x < \infty.$$

U-statistics include a kernel $h(x_1, \dots, x_m)$, that is a measurable, symmetric and real-valued function. Symmetry in this case means invariance against permutation: $h(x_{\sigma_1}, \dots, x_{\sigma_m}) = h(x_1, \dots, x_m)$ for every permutation σ. The dimension of the kernel is m. In literature, mostly bivariate kernels are studied (cf. Borovkova et al. (2001), Wendler (2011b), Levy-Leduc et al. (2011)). This limits the number of possible estimators in this class very much, such that Fischer et al. (2016a) showed the Central Limit Theorem for multivariate ($m \geq 2$) kernels.
Analogously to the empirical distribution function of random variables an empirical distribution function H_n of the evaluations $h(X_{i_1}, \dots, X_{i_m})$ is given by

$$H_n(x) = \frac{1}{\binom{n}{m}} \sum_{1 \leq i_1 < \dots < i_m \leq n} \mathbb{1}_{[h(X_{i_1}, \dots, X_{i_m}) \leq x]}, \quad -\infty < x < \infty.$$

Sometimes H_n is also defined as

$$H_n(x) = \frac{1}{n(n-1) \cdot \ldots \cdot (n-m+1)} \sum \mathbb{1}_{[h(X_{i_1}, \dots, X_{i_m}) \leq x]},$$

where the sum is taken over all $n(n-1) \cdot \ldots \cdot (n-m+1)$ m-tuples $(i_1, \dots, i_m)$ of distinct elements from $\{1, \dots, n\}$, but because of the symmetry of h this does not make any difference.

H_F is defined as the distribution function of the kernel h with
$H_F(y) = \mathbb{P}_F(h(Y_1, \ldots, Y_m) \leq y)$ for independent copies $Y_1, \ldots, Y_m$ of X_1 and $0 < h_F < \infty$ is the related density. Please note that this implies that H_F is continuous. The index F refers to the fact that the distribution function of the original data $X_1, \ldots, X_n$ is F.
We define $h_{F;X_{i_2},\ldots,X_{i_k}}$ as the density of $h(Y_{i_1}, X_{i_2}, \ldots, X_{i_k}, Y_{i_{k+1}}, \ldots, Y_{i_m})$ for $2 \leq k \leq m$ and $i_1 < i_2 < \ldots < i_m$.

With these assumptions, the definition of U-statistics is now possible.

4.1. U-statistics

U-statistics form a class of statistics originally developed by Paul R. Halmos and Wassily Hoeffding (Halmos (1946); Hoeffding (1948)). The "U" stands for unbiased, since a U-statistic is an unbiased estimator. It is a very important class of statistics because of the relative simple form and also because many common estimators can be expressed as such a U-statistic. For more details we refer at this point to Serfling (1980) and Lee (1990).

Definition 4.1. *Let $h : \mathbb{R}^m \to \mathbb{R}$ be a measurable function. A U-statistic with kernel h is defined as*

$$U_n = \frac{1}{\binom{n}{m}} \sum_{1 \leq i_1 < \ldots < i_m \leq n} h(X_{i_1}, \ldots, X_{i_m}).$$

While examining U-statistics often the decomposing technique by Hoeffding is used (Hoeffding (1948)), which makes a separate consideration of the single terms possible.

Definition 4.2. *(Hoeffding decomposition)*
Let U_n be a U-statistic with kernel $h = h(x_1, \ldots, x_m)$. Then one can write U_n as

$$U_n = \theta + \sum_{j=1}^{m} \binom{m}{j} \frac{1}{\binom{n}{j}} S_{jn},$$

where

$$\begin{aligned}
\theta &= \mathbb{E}(h(Y_1, \ldots, Y_m)) \\
S_{jn} &= \sum_{1 \leq i_1 < \ldots < i_j \leq n} h_j(X_{i_1}, \ldots, X_{i_j}) \\
h_1(x_1) &= \tilde{h}_1(x_1) \\
h_2(x_1, x_2) &= \tilde{h}_2(x_1, x_2) - h_1(x_1) - h_1(x_2) \\
h_3(x_1, x_2, x_3) &= \tilde{h}_3(x_1, x_2, x_3) - \sum_{i=1}^{3} h_1(x_i) - \sum_{1 \leq i < j \leq 3} h_2(x_i, x_j) \\
&\ldots \\
h_m(x_1, \ldots, x_m) &= \tilde{h}_m(x_1, \ldots, x_m) - \sum_{i=1}^{m} h_1(x_i) - \sum_{1 \leq i_1 < i_2 \leq m} h_2(x_{i_1}, x_{i_2}) \\
&\quad - \cdots - \sum_{1 \leq i_1 < \ldots < i_{m-1} \leq m} h_{m-1}(x_{i_1}, \ldots, x_{i_{m-1}}) \\
\tilde{h}_j(x_1, \ldots, x_j) &= \mathbb{E}(h(x_1, \ldots, x_j, Y_{j+1}, \ldots, Y_m)) - \theta
\end{aligned}$$

for independent copies $Y_1, \ldots, Y_m$ of X_1.

We call $\frac{m}{n}\sum_{i=1}^{n} h_1(X_i)$ linear part, the remaining parts are degenerated.
For most of the results in this section we need a regularity condition for the kernel h. It is very similar to the Lipschitz-continuity and basically has been developed by Denker and Keller (1986). The same variation condition is also used in Fischer (2013) and the related paper Fischer et al. (2016a).

Definition 4.3. *A kernel h satisfies the variation condition, if there exists a constant L and an $\epsilon_0 > 0$, such that for all $\epsilon \in (0, \epsilon_0)$*

$$\mathbb{E}\left(\sup_{\|(x_1,\ldots,x_m)-(X_1',\ldots,X_m')\|\leq\epsilon} \left|h(x_1,\ldots,x_m) - h(X_1',\ldots,X_m')\right|\right) \leq L\epsilon,$$

where X_i' are independent with the same distribution as X_i and $\|\cdot\|$ is the Euklidean norm. A kernel h satisfies the extended variation condition, if there exists additionally a constant L' and a $\delta_0 > 0$, such that for all $\delta \in (0, \delta_0)$ and all $2 \leq k \leq m$

$$\mathbb{E}\left(\sup_{|x_{i_1}-Y_{i_1}|\leq\delta} \left|h(x_{i_1}, X_{i_2},\ldots,X_{i_k},Y_{i_{k+1}},\ldots,Y_{i_m}) - h(Y_{i_1}, X_{i_2},\ldots,X_{i_k},Y_{i_{k+1}},\ldots,Y_{i_m})\right|\right) \leq L'\delta,$$

for independent copies $(Y_n)_{n\in\mathbb{N}}$ of $(X_n)_{n\in\mathbb{N}}$ and all $i_1 < i_2 < \ldots < i_m$. If the kernel has dimension one, we note, that it satisfies the extended variation condition, if it satisfies the variation condition.

Remark 4.1. A Lipschitz-continuous kernel satisfies the variation condition.

Sometimes also a condition is needed which demands regularity in the L_2-space.

Definition 4.4. *A kernel h satisfies the L_2 variation condition, if there exists a constant L and an $\epsilon_0 > 0$, such that for all $\epsilon \in (0, \epsilon_0)$*

$$\mathbb{E}\left(\sup_{\|(x_1,\ldots,x_m)-(X_1',\ldots,X_m')\|\leq\epsilon} \left|h(x_1,\ldots,x_m) - h(X_1',\ldots,X_m')\right|\right)^2 \leq L\epsilon,$$

where X_i' are independent with same distribution as X_i and $\|\cdot\|$ is the Euklidean norm.

Remark 4.2. For bounded kernels the L_2 variation condition follows directly from the simple variation condition, since $(a-b)^2 \leq |a-b| \cdot (|a| + |b|)$.

Since a decomposition of U-statistics into the single kernels of the Hoeffding decomposition h_k, $k = 1,\ldots,m$, is used as a key tool in many proofs it is important to know, which properties of the kernel h can be assigned to the Hoeffding-kernels. Fischer (2013) respectively Fischer et al. (2016a) show that for example the extended variation condition is such a property.

Lemma 4.1. *(Fischer et al. (2016a))*
If the kernel h satisfies the extended variation condition, then the kernels h_k in Definition 4.2, $1 \leq k \leq m$, satisfy it as well.

Of course, also the boundedness of h is assigned, since h can only be bounded if every single part of its decomposition is bounded.

Now we state one of the main results: asymptotic normality of U-statistics under NED. We know this result from Wendler (2011a) for bivariate U-statistics, but our theorem admits arbitrary dimension m. Under independence one can find a Central Limit Theorem for U-Statistics for example in Serfling (1980), whereas Wendler (2011a) and Fischer (2013) respectively the related article Fischer et al. (2016a) show the same result for strongly mixing random variables. The proofs of the following theorems are given later on in this section.

Theorem 4.1. *Let $h : \mathbb{R}^m \to \mathbb{R}$ be a bounded kernel satisfying the extended variation condition. Moreover, let $(X_n)_{n\in\mathbb{N}}$ be L_1 NED with approximation constants $(a_l)_{l\in\mathbb{N}}$ on an absolutely regular process $(Z_n)_{n\in\mathbb{Z}}$ with mixing coefficients $(\beta(l))_{l\in\mathbb{N}}$. Let also a $\delta > 1$ exist, such that $\beta(l) = O\left(l^{-\delta}\right)$ and $a_l = O\left(l^{-\delta-2}\right)$. Then we have*

$$\sqrt{n}(U_n - \theta) \xrightarrow{D} N(0, m^2\sigma^2)$$

with $\sigma^2 = Var(h_1(X_1)) + 2\sum_{j=1}^{\infty} Cov(h_1(X_1), h_1(X_{1+j}))$.
If $\sigma = 0$, then the statement is meant as convergence to 0.

Although our focus in this thesis is on a bounded kernel h of the U-statistic, we also want to state the Central Limit Theorem for U-statistics with unbounded kernel. Again, we find an analogous result for kernels of dimension 2 in Wendler (2011a). The proof of this theorem then follows in the same way as for bounded kernels, simply interchanging the applied lemmata.

Theorem 4.2. *Let $h : \mathbb{R}^m \to \mathbb{R}$ be a kernel with uniform $(2+\gamma)$-moments, $\gamma > 0$, satisfying the extended variation condition. Moreover, let $(X_n)_{n\in\mathbb{N}}$ be L_1 NED with approximation constants $(a_l)_{l\in\mathbb{N}}$ on an absolutely regular process $(Z_n)_{n\in\mathbb{Z}}$ with mixing coefficients $(\beta(l))_{l\in\mathbb{N}}$. Let also a $\delta > \frac{2+\gamma}{\gamma}$ exist, such that $\beta(l) = O\left(l^{-\delta}\right)$ and $a_l = O\left(l^{-2\delta-1}\right)$. Then we have*

$$\sqrt{n}(U_n - \theta) \xrightarrow{D} N(0, m^2\sigma^2)$$

with $\sigma^2 = Var(h_1(X_1)) + 2\sum_{j=1}^{\infty} Cov(h_1(X_1), h_1(X_{1+j}))$.
If $\sigma = 0$, then the statement is meant as convergence to 0.

In general, if the distribution of the $(X_n)_{n\in\mathbb{N}}$ is not specified, the long-run variance σ^2 in Theorem 4.1 is unknown. Therefore, for applications an estimator of σ^2 is needed. For bivariate U-statistics or U-processes Vogel and Wendler (2015) and Dehling et al. (2016) give consistent estimators by using an empirical version of the first Hoeffding kernel and a weight function. Nevertheless, for the multivariate case such a result is not known. The multivariate extension to this estimator is

$$\hat{\sigma}^2 = \sum_{r=-(n-1)}^{n-1} \kappa\left(\frac{|r|}{b_n}\right)\hat{\rho}(|r|),$$

where κ is the weight function and b_n is the bandwidth.

$$\hat{\rho}(r) = \frac{1}{n}\sum_{i=1}^{n-r} \hat{h}_1(X_i)\hat{h}_1(X_{i+r})$$

is the empirical covariance for lag r, using the empirical version of the first Hoeffding kernel

$$\hat{h}_1(x) = \frac{1}{n^{m-1}} \sum_{1 \leq i_1 < \ldots < i_{m-1} \leq n} h(x, X_{i_1}, \ldots, X_{i_{m-1}}) - \frac{1}{n^m} \sum_{1 \leq i_1 < \ldots < i_m < n} h(X_{i_1}, \ldots, X_{i_m}).$$

As Dehling et al. (2016) have already shown we need some regularity conditions for κ and b_n to achieve consistency of the estimator. These conditions are similar to the assumption made in de Jong and Davidson (2000) and are fulfilled, for example, by the Bartlett kernel $\kappa(t) = (1 - |t|)1_{[|t| \leq 1]}$.

Assumption 4.1. *The function $\kappa : [0, \infty) \to [0, 1)$ is continuous at 0 and all but a finite number of points. Moreover, $|\kappa|$ is dominated by a non-increasing, integrable function and*

$$\int_0^\infty \Big| \int_0^\infty \kappa(t) \cos(xt) dt \Big| dx < \infty.$$

The bandwidth b_n satisfies $b_n \to \infty$ as $n \to \infty$ and $b_n / \sqrt{n} \to 0$.

With these considerations we are able to show that $\hat{\sigma}^2$ is a consistent estimator for the long-run variance.

Theorem 4.3. *Let $h : \mathbb{R}^m \to \mathbb{R}$ be a bounded kernel satisfying the extended variation condition. Moreover, let $(X_n)_{n \in \mathbb{N}}$ be NED with approximation constants $(a_l)_{l \in \mathbb{N}}$ on an absolutely regular process $(Z_n)_{n \in \mathbb{Z}}$ with mixing coefficients $(\beta(l))_{l \in \mathbb{N}}$ and let a $\delta > 11$ exist, such that $\sum_{l=1}^\infty l \beta^{2/(2+\delta)}(l) < \infty$ and $a_l = O\left(l^{-\delta-3}\right)$. The weight function κ and the bandwidth b_n shall fulfil Assumption 4.1. Then*

$$\hat{\sigma}^2 \xrightarrow{p} \sigma^2 \text{ for } n \to \infty,$$

where $\sigma^2 = Var(h_1(X_1)) + 2 \sum_{j=1}^\infty Cov(h_1(X_1), h_1(X_{1+j}))$.

Some results needed for the proofs are stated in the following section.

Preliminary results

For the proofs of the main theorems some lemmata are needed. The following results are similar to the case of strong mixing (Fischer (2013) and Fischer et al. (2016a)) but since the proofs need different arguments in some cases we state the proofs for the sake of completeness.

The first lemma is analogous to Lemma 4.2 in Fischer et al. (2016a) and an extension of Lemma 3.2.4 in Wendler (2011a).

Lemma 4.2. *Let $(X_n)_{n \in \mathbb{N}}$ be NED with approximation constants $(a_l)_{l \in \mathbb{N}}$ on an absolutely regular process $(Z_n)_{n \in \mathbb{Z}}$ with mixing coefficients $(\beta(l))_{n \in \mathbb{N}}$. Moreover, let be $A_L = \sqrt{2 \sum_{i=L}^\infty a_i}$ and let the kernel h be bounded and satisfy the extended variation condition. Then there exists for all $2 \leq k \leq m$ a constant C, such that for $r = \max\left\{i_{(2)} - i_{(1)}, i_{(2k)} - i_{(2k-1)}\right\}$ with $i_{(1)} \leq \ldots \leq i_{(2k)}$ follows*

$$\left| \mathbb{E}\left(h_k(X_{i_1}, \ldots, X_{i_k}) h_k(X_{i_{k+1}}, \ldots, X_{i_{2k}})\right) \right| \leq C \left(\beta\left(\left[\frac{r}{3} \right] \right) + A_{\left[\frac{r}{3}\right]} \right).$$

Proof.
Like Fischer et al. (2016a) for simplicity we only consider the case $i_1 < \ldots < i_{2k}$ and $r = i_2 - i_1 \geq i_{2k} - i_{2k-1}$. Using Corollary 2.17 of Borovkova et al. (2001) there exist sequences $(X'_n)_{n \in \mathbb{N}}$ and $(X''_n)_{n \in \mathbb{N}}$ with the properties

1. $(X'_n)_{n\in\mathbb{N}}$, $(X''_n)_{n\in\mathbb{N}}$ have the same distribution as $(X_n)_{n\in\mathbb{N}}$
2. $(X''_n)_{n\in\mathbb{N}}$ is independent of $(X_n)_{n\in\mathbb{N}}$
3. $\mathbb{P}\left(\sum_{i=r}^{\infty}|X_i - X'_i| > A_{\left[\frac{r}{3}\right]}\right) \leq A_{\left[\frac{r}{3}\right]} + \beta\left(\left[\frac{r}{3}\right]\right)$
4. $\mathbb{P}\left(\sum_{i=0}^{\infty}\left|X'_{-i} - X''_{-i}\right| > A_{\left[\frac{r}{3}\right]}\right) \leq A_{\left[\frac{r}{3}\right]}$.

Because h_k is degenerated it follows that $\mathbb{E}(h_k(X''_{i_1}, X_{i_2}, \ldots, X_{i_k})h_k(X_{i_{k+1}}, \ldots, X_{i_{2k}})) = 0$ and additionally

$$\mathbb{E}(h_k(X_{i_1}, X_{i_2}, \ldots, X_{i_k})h_k(X_{i_{k+1}}, \ldots, X_{i_{2k}})) = \mathbb{E}(h_k(X'_{i_1}, \ldots, X'_{i_k})h_k(X'_{i_{k+1}}, \ldots, X'_{i_{2k}})).$$

So we get by the triangular inequality

$$\begin{aligned}
&\left|\mathbb{E}(h_k(X_{i_1}, X_{i_2}, \ldots, X_{i_k})h_k(X_{i_{k+1}}, \ldots, X_{i_{2k}}))\right| \\
&= \left|\mathbb{E}(h_k(X'_{i_1}, \ldots, X'_{i_k})h_k(X'_{i_{k+1}}, \ldots, X'_{i_{2k}})) - \mathbb{E}(h_k(X''_{i_1}, X_{i_2}, \ldots, X_{i_k})h_k(X_{i_{k+1}}, \ldots, X_{i_{2k}}))\right| \\
&\leq \left|\mathbb{E}(h_k(X''_{i_1}, X_{i_2}, \ldots, X_{i_k})h_k(X'_{i_{k+1}}, \ldots, X'_{i_{2k}}) - h_k(X''_{i_1}, X_{i_2}, \ldots, X_{i_k})h_k(X_{i_{k+1}}, \ldots, X_{i_{2k}}))\right| \qquad (4.1) \\
&+ \left|\mathbb{E}(h_k(X'_{i_1}, \ldots, X'_{i_k})h_k(X'_{i_{k+1}}, \ldots, X'_{i_{2k}}) - h_k(X''_{i_1}, X_{i_2}, \ldots, X_{i_k})h_k(X'_{i_{k+1}}, \ldots, X'_{i_{2k}}))\right|. \qquad (4.2)
\end{aligned}$$

Because h is bounded and satisfies the extended variation condition this also holds for h_k with boundedness constant M and variation constant L, what we will utilize in the following, where, in contrast to Fischer (2013), an inequality for both the NED and the mixing coefficients is needed.

$$\begin{aligned}
(4.2) &= \left|\mathbb{E}((h_k(X'_{i_1}, \ldots, X'_{i_k}) - h_k(X''_{i_1}, X_{i_2}, \ldots, X_{i_k}))\underbrace{h_k(X'_{i_{k+1}}, \ldots, X'_{i_{2k}})}_{\leq M}\right| \\
&\leq M\left[\left|\mathbb{E}\left((h_k(X''_{i_1}, X_{i_2}, \ldots, X_{i_k}) - h_k(X'_{i_1}, \ldots, X'_{i_k}))\mathbb{1}_{\left[\left|X''_{i_1} - X'_{i_1}\right| \leq A_{\left[\frac{r}{3}\right]}, \ldots, \left|X_{i_k} - X'_{i_k}\right| \leq A_{\left[\frac{r}{3}\right]}\right]}\right)\right|\right. \\
&\left. + \left|\mathbb{E}\Big(\underbrace{(h_k(X''_{i_1}, X_{i_2}, \ldots, X_{i_k}) - h_k(X'_{i_1}, \ldots, X'_{i_k}))}_{\leq 2M}\mathbb{1}_{\left[\left|X''_{i_1} - X'_{i_1}\right| \leq A_{\left[\frac{r}{3}\right]}, \ldots, \left|X_{i_k} - X'_{i_k}\right| \leq A_{\left[\frac{r}{3}\right]}\right]^C}\Big)\right|\right] \\
&\leq ML\sqrt{k}A_{\left[\frac{r}{3}\right]} + M2M\mathbb{P}\left(\left|X''_{i_1} - X'_{i_1}\right| > A_{\left[\frac{r}{3}\right]}\right) + \ldots + M2M\mathbb{P}\left(\left|X_{i_k} - X'_{i_k}\right| > A_{\left[\frac{r}{3}\right]}\right) \\
&\leq ML\sqrt{k}A_{\left[\frac{r}{3}\right]} + 2kM^2A_{\left[\frac{r}{3}\right]} + 2(k-1)M^2\beta\left(\left[\tfrac{r}{3}\right]\right) \\
&\leq ML\sqrt{k}A_{\left[\frac{r}{3}\right]} + 2kM^2A_{\left[\frac{r}{3}\right]} + 2(k-1)M^2\beta\left(\left[\tfrac{r}{3}\right]\right) + 2M^2\beta\left(\left[\tfrac{r}{3}\right]\right) + ML\sqrt{k}\beta\left(\left[\tfrac{r}{3}\right]\right) \\
&\leq \left(2kM^2 + ML\sqrt{k}\right)\left(A_{\left[\frac{r}{3}\right]} + \beta\left(\left[\tfrac{r}{3}\right]\right)\right),
\end{aligned}$$

where we used properties 3. and 4. from above and $\beta(l) \geq 0$.
For the second term (4.1) we use similar arguments for getting the same result.

$$(4.1) \leq \left|\mathbb{E}((g_k(X'_{i_{k+1}}, \ldots, X'_{i_{2k}}) - g_k(X_{i_{k+1}}, \ldots, X_{i_{2k}}))g_k(X''_{i_1}, X_{i_2} \ldots, X_{i_k}))\right|$$

$$
\begin{aligned}
&\leq M\Bigg[\Bigg|\mathbb{E}\Bigg((g_k(X'_{i_{k+1}},\dots,X'_{i_{2k}}) - g_k(X_{i_{k+1}},\dots,X_{i_{2k}})) \\
&\qquad\qquad \cdot \mathbb{1}_{\left[\left|X'_{i_{k+1}}-X_{i_{k+1}}\right|\leq A_{\left[\frac{r}{3}\right]},\dots,\left|X'_{i_{2k}}-X_{i_{2k}}\right|\leq A_{\left[\frac{r}{3}\right]}\right]}\Bigg)\Bigg| \\
&+\Bigg|\mathbb{E}\Bigg((g_k(X'_{i_{k+1}},\dots,X'_{i_{2k}}) - g_k(X_{i_{k+1}},\dots,X_{i_{2k}})) \\
&\qquad\qquad \cdot \mathbb{1}_{\left[\left|X'_{i_{k+1}}-X_{i_{k+1}}\right|\leq A_{\left[\frac{r}{3}\right]},\dots,\left|X'_{i_{2k}}-X_{i_{2k}}\right|\leq A_{\left[\frac{r}{3}\right]}\right]^C}\Bigg)\Bigg|\Bigg] \\
&\leq ML\sqrt{k}A_{\left[\frac{r}{3}\right]} + M2Mk\left(A_{\left[\frac{r}{3}\right]} + \beta\left(\left[\tfrac{r}{3}\right]\right)\right) \\
&\leq \left(2kM^2 + ML\sqrt{k}\right)\left(A_{\left[\frac{r}{3}\right]} + \beta\left(\left[\tfrac{r}{3}\right]\right)\right).
\end{aligned}
$$

According to that, it is

$$(4.1) + (4.2) \leq C\left(A_{\left[\frac{r}{3}\right]} + \beta\left(\left[\tfrac{r}{3}\right]\right)\right).$$

□

The following Lemma generalises a result from Fischer (2013) (see also Fischer et al. (2016a)) to near epoch dependent processes as well as arbitrary constants $(c_{i,j})_{i,j\in\mathbb{N}}$. This extension is needed later on to show the consistency of the estimator $\hat{\sigma}^2$.

Lemma 4.3. *Let the kernel h be bounded and satisfy the extended variation condition. Moreover, let $(X_n)_{n\in\mathbb{N}}$ be NED with approximation constants $(a_l)_{l\in\mathbb{N}}$ on an absolutely regular process $(Z_n)_{n\in\mathbb{Z}}$ with mixing coefficients $(\beta(l))_{l\in\mathbb{N}}$ and $\sum_{l=0}^{n} l\left(\beta(l) + A_l\right) = O\left(n^{\gamma}\right)$ with $A_l = \sqrt{2\sum_{i=l}^{\infty} a_i}$ for a $\gamma < 1$. Then for all $2 \leq k \leq m$ and any constants $(c_{i,j})_{i,j\in\mathbb{N}}$*

$$
\begin{aligned}
&\sum_{i_1,\dots,i_{2k}=1}^{n} \left|\mathbb{E}(h_k(X_{i_1},\dots,X_{i_k})h_k(X_{i_{k+1}},\dots,X_{i_{2k}}))c_{i_1,i_{k+1}}\right| \\
&\qquad = \max_{i_1,i_{k+1}\in\{1,\dots,n\}} |c_{i_1,i_{k+1}}| O(n^{2k-2+\gamma}).
\end{aligned}
$$

This Lemma is analogous to Lemma 4.3 in Fischer et al. (2016a) and can be proved similarly.

Proof.
Again set $\{i_1,\dots,i_{2k}\} = \{i_{(1)},\dots,i_{(2k)}\}$ with $i_{(1)} \leq \dots \leq i_{(2k)}$. We can rewrite the above sum as

$$
\begin{aligned}
&\sum_{i_1,\dots,i_{2k}=1}^{n} \left|\mathbb{E}(h_k(X_{i_1},\dots,X_{i_k})h_k(X_{i_{k+1}},\dots,X_{i_{2k}}))c_{i_1,i_{k+1}}\right| \\
&\leq \max_{i_1,i_{k+1}\in\{1,\dots,n\}} |c_{i_1,i_{k+1}}| \sum_{i_1,\dots,i_{2k}=1}^{n} \left|\mathbb{E}(h_k(X_{i_1},\dots,X_{i_k})h_k(X_{i_{k+1}},\dots,X_{i_{2k}}))\right| \\
&= \max_{i_1,i_{k+1}\in\{1,\dots,n\}} |c_{i_1,i_{k+1}}| \\
&\qquad\qquad \cdot \sum_{l=0}^{n} \sum_{\substack{i_1,\dots,i_{2k}=1 \\ \max\left\{i_{(2)}-i_{(1)},i_{(2k)}-i_{(2k-1)}\right\}=l}}^{n} \left|\mathbb{E}(h_k(X_{i_1},\dots,X_{i_k})h_k(X_{i_{k+1}},\dots,X_{i_{2k}}))\right|
\end{aligned}
$$

$$\leq \max_{i_1,i_{k+1}\in\{1,\dots,n\}} |c_{i_1,i_{k+1}}| C \sum_{l=0}^{n} \sum_{\substack{i_1,\dots,i_{2k} \\ \max\{i_{(2)}-i_{(1)},i_{(2k)}-i_{(2k-1)}\}=l}} \left(\beta\left(\left[\frac{l}{3}\right]\right) + A_{\left[\frac{l}{3}\right]}\right),$$

by application of Lemma 4.2.
Now we want to simplify the expression by calculating the quantity of $(i_1,\dots,i_{2k})$ where $\max\{i_{(2)}-i_{(1)}, i_{(2k)}-i_{(2k-1)}\} = l$. Using combinatorial arguments we can see that there exist $(2k)!$ possibilities to obtain the same sequence $i_{(1)},\dots,i_{(2k)}$. We now fix $i_{(1)}$ and $i_{(2k)}$, having n^2 possibilities for this. Having in mind that $\max\{i_{(2)}-i_{(1)}, i_{(2k)}-i_{(2k-1)}\} = l$ and suppose $i_{(2)}-i_{(1)} = \max\{i_{(2)}-i_{(1)}, i_{(2k)}-i_{(2k-1)}\} = l$ then $i_{(2)}$ is automatically determined by the choice of $i_{(1)}$. Then, $i_{(2k-1)}$ can only take l distinct values. Supposing $i_{(2k)}-i_{(2k-1)} = \max\{i_{(2)}-i_{(1)}, i_{(2k)}-i_{(2k-1)}\} = l$ the same is valid. All remaining values of the k-tuple are arbitrary. Therefore, the quantity of the terms equals $(2k)! \cdot n^2 l n^{2k-4} = l \cdot (2k)! \cdot n^{2k-2}$ and

$$\begin{aligned}
&\sum_{i_1,\dots,i_{2k}=1}^{n} \left|\mathbb{E}(h_k(X_{i_1},\dots,X_{i_k})h_k(X_{i_{k+1}},\dots,X_{i_{2k}}))c_{i_1,i_{k+1}}\right| \\
&\leq \max_{i_1,i_{k+1}\in\{1,\dots,n\}} |c_{i_1,i_{k+1}}| C' n^{2k-2} \sum_{l=0}^{n} l \left(\beta\left(\left[\frac{l}{3}\right]\right) + A_{\left[\frac{l}{3}\right]}\right) \\
&= \max_{i_1,i_{k+1}\in\{1,\dots,n\}} |c_{i_1,i_{k+1}}| O(n^{2k-2+\gamma}).
\end{aligned}$$

□

These results serve as basis for proving the main theorems.

Proofs of the main results

In this segment we will assemble the results so far to the main proofs.

Proof of Theorem 4.1.
The proof makes use of the Hoeffding decomposition (Def. 4.2)

$$\sqrt{n}(U_n - \theta) = \sqrt{n} \sum_{j=1}^{m} \binom{m}{j} \frac{1}{\binom{n}{j}} S_{jn}.$$

We show that the linear part $\frac{m}{\sqrt{n}} \sum_{i=1}^{n} h_1(X_i)$ is asymptotically normal and that the remaining terms converge to 0 in probability.

We know that h_1 is bounded because h is bounded. Using Lemma 2.1.7 of Wendler (2011a) we also know that h_1 is NED with approximation constants $a'_l = C a_l^{\frac{1}{2}}$. In the near epoch dependent case we have to prove assumptions for the mixing coefficients as well as the approximating functions of the NED process. This is done in the following.
Together with the above assumptions it is

$$\sum_{l=1}^{\infty} \beta(l) < \infty, \qquad \sum_{l=1}^{\infty} a'_l < \infty$$

and $a_l' \leq \frac{C}{l^{\delta/2+1}} \leq \frac{C}{l \log l}$.
We have made these considerations to finally apply Theorem 2.3 of Ibragimov (1962) getting

$$\frac{m}{\sqrt{n}} \sum_{i=1}^{n} h_1(X_1) \xrightarrow{D} N(0, m^2\sigma^2)$$

with $\sigma^2 = \mathrm{Var}(h_1(X_1)) + 2\sum_{j=1}^{\infty} \mathrm{Cov}(h_1(X_1), h_1(X_{1+j})) < \infty$.
For the remaining terms we want to use Lemma 4.3 with constants $c_{i_1,i_{k+1}} = 1$ for all $2 \leq k \leq m$, needing $\sum_{l=0}^{n} l\,(\beta(l) + A_l) = O(n^\gamma)$ for a $\gamma < 1$.
Using

$$A_l = \left(2\sum_{i=l}^{\infty} a_i\right)^{\frac{1}{2}} \leq \left(2C\sum_{i=l}^{\infty} i^{-\delta-2}\right)^{\frac{1}{2}} = O\left(n^{-\frac{\delta+1}{2}}\right)$$

we get

$$\begin{aligned}
&\sum_{l=0}^{n} l\,(\beta(l) + A_l) \leq C\sum_{l=1}^{n} l\left(l^{-\delta} + n^{-\frac{\delta+1}{2}}\right) \leq C\sum_{l=1}^{n} l^{-\delta+1} + Cn^{-\frac{\delta+1}{2}}\sum_{l=1}^{n} l \\
&= O(n^\gamma) + O\left(n^{2-\frac{\delta+1}{2}}\right) = O(n^\gamma)
\end{aligned}$$

and so Lemma 4.3 is applicable.
So it is for all $2 \leq k \leq m$

$$\begin{aligned}
&\mathrm{Var}\left(\sqrt{n}\binom{m}{k}\binom{n}{k}^{-1} \sum_{1\leq i_1<\ldots<i_k\leq n} h_k(X_{i_1}, \ldots, X_{i_k})\right) \\
&\leq \frac{m^{2k}k^{\frac{k}{2}}}{n^{2k-1}} \sum_{1\leq i_1<\ldots<i_k\leq n} \sum_{1\leq i_{k+1}<\ldots<i_{2k}\leq n} \left|\mathbb{E}\left(h_k(X_{i_1}, \ldots, X_{i_k})h_k(X_{i_{k+1}}, \ldots, X_{i_{2k}})\right)\right| \\
&\leq \frac{m^{2k}k^{\frac{k}{2}}}{n^{2k-1}} \sum_{i_1,\ldots,i_{2k}=1}^{n} \left|\mathbb{E}\left(h_k(X_{i_1}, \ldots, X_{i_k})h_k(X_{i_{k+1}}, \ldots, X_{i_{2k}})\right)\right| \\
&= O(n^{2k-2+\gamma-(2k-1)}) = O(n^{-1+\gamma}).
\end{aligned}$$

And so

$$\mathrm{Var}\left(\sqrt{n}\binom{m}{k}\binom{n}{k}^{-1} \sum_{1\leq i_1<\ldots<i_k\leq n} h_k(X_{i_1}, \ldots, X_{i_k})\right) \overset{n\to\infty}{\longrightarrow} 0$$

and with the Chebychev inequality we have

$$\sqrt{n}\binom{m}{k}\binom{n}{k}^{-1} \sum_{1\leq i_1<\ldots<i_k\leq n} h_k(X_{i_1}, \ldots, X_{i_k}) \xrightarrow{P} 0 \quad \text{for } n \to \infty.$$

The Theorem of Slutsky completes the proof.

□

The proof of the consistency of the estimator is based on the proof of Dehling et al. (2016) but has to be generalised to the more complicated case of arbitrary dimension. In this case much more terms have to be considered and their asymptotic behaviour has to be investigated.

Proof of Theorem 4.3.
By decomposing the estimator into two parts

$$\begin{aligned}\hat{\sigma}^2 = &\sum_{r=-(n-1)}^{n-1} \kappa\left(\frac{|r|}{b_n}\right) \frac{1}{n} \sum_{j=1}^{n-|r|} \Big(\hat{h}_1(X_i)\hat{h}_1(X_{i+|r|}) \\ &\qquad - h_1(X_i)h_1(X_{i+|r|}) + h_1(X_i)h_1(X_{i+|r|})\Big) \\ = &\sum_{r=-(n-1)}^{n-1} \kappa\left(\frac{|r|}{b_n}\right) \frac{1}{n} \sum_{j=1}^{n-|r|} h_1(X_i)h_1(X_{i+|r|}) \\ &+ \sum_{r=-(n-1)}^{n-1} \kappa\left(\frac{|r|}{b_n}\right) \frac{1}{n} \sum_{j=1}^{n-|r|} \Big(\hat{h}_1(X_i)\hat{h}_1(X_{i+|r|}) - h_1(X_i)h_1(X_{i+|r|})\Big)\end{aligned}$$

we can apply the results of de Jong and Davidson (2000).
For the first term de Jong and Davidson (2000) have shown in their Theorem 2.1 that it converges to σ^2, where the assumptions of the theorem are fulfilled due to Assumption 4.1 and the NED-assumption. Therefore, it remains to show

$$\mathbb{E}\Big| \sum_{r=-(n-1)}^{n-1} \kappa\left(\frac{|r|}{b_n}\right) \frac{1}{n} \sum_{j=1}^{n-|r|} \Big(\hat{h}_1(X_i)\hat{h}_1(X_{i+|r|}) - h_1(X_i)h_1(X_{i+|r|})\Big) \Big| \longrightarrow 0.$$

Let us first expand $h_1(x) - \hat{h}_1(x)$ into single terms:

$$\begin{aligned}&h_1(x) - \hat{h}_1(x) \\ =&h_1(x) - \frac{1}{n^{m-1}} \sum_{1\le i_1<\ldots<i_{m-1}\le n} h(x, X_{i_1}, \ldots, X_{i_{m-1}}) \\ &+ \frac{1}{n^m} \sum_{1\le i_1<\ldots<i_m\le n} h(X_{i_1}, \ldots, X_{i_m}) \\ =&h_1(x) - \frac{1}{n^{m-1}} \sum_{1\le i_1<\ldots<i_{m-1}\le n} \Big(h_m(x, X_{i_1}, \ldots, X_{i_{m-1}}) + h_1(x) + \sum_{j=1}^{m-1} h_1(X_{i_j}) \\ &\qquad + \ldots + \sum_{1\le j_1<\ldots<j_{m-2}\le m-1} h_{m-1}(x, X_{i_{j_1}}, \ldots, X_{i_{j_{m-2}}}) \\ &\qquad + \sum_{1\le j_1<\ldots<j_{m-1}\le m-1} h_{m-1}(X_{i_{j_1}}, \ldots, X_{i_{j_{m-1}}})\Big) \\ &+ \frac{1}{n^m} \sum_{1\le i_1<\ldots<i_m\le n} \Big(h_m(X_{i_1}, \ldots, X_{i_m}) + \sum_{j=1}^{m} h_1(X_{i_j}) + \ldots\end{aligned}$$

$$
\begin{aligned}
&\quad + \sum_{1\le j_1<\ldots<j_{m-1}\le m} h_{m-1}(X_{i_{j_1}},\ldots,X_{i_{j_{m-1}}})\Big) \\
&= -\frac{1}{n^{m-1}} \sum_{1\le i_1<\ldots<i_{m-1}\le n} h_m(x,X_{i_1},\ldots,X_{i_{m-1}}) - (m-1)\frac{1}{n}\sum_{i=1}^n h_1(X_i) \\
&\quad - \ldots - 2\frac{1}{n^{m-2}} \sum_{1\le i_1<\ldots<i_{m-2}\le n} h_{m-1}(x,X_{i_1},\ldots,X_{i_{m-2}}) \\
&\quad - \frac{1}{n^{m-1}} \sum_{1\le i_1<\ldots<i_m\le n} h_{m-1}(X_{i_1},\ldots,X_{i_{m-1}}) \\
&\quad + \frac{1}{n^m} \sum_{1\le i_1<\ldots<i_m\le n} h_m(X_{i_1},\ldots,X_{i_m}) + m\frac{1}{n}\sum_{i=1}^n h_1(X_i) \\
&\quad + \ldots + 2\frac{1}{n^{m-1}} \sum_{1\le i_1<\ldots<i_{m-1}\le n} h_{m-1}(X_{i_1},\ldots,X_{i_{m-1}}) \\
&= \frac{1}{n}\sum_{i=1}^n h_1(X_i) - (m-2)\frac{1}{n}\sum_{i=1}^n h_2(x,X_i) + (m-2)\frac{1}{n^2} \sum_{1\le i<j\le n} h_2(X_i,X_j) \\
&\quad - \ldots - 2\frac{1}{n^{m-2}} \sum_{1\le i_1<\ldots<i_{m-2}\le n} h_{m-1}(x,X_{i_1},\ldots,X_{i_{m-2}}) \\
&\quad + 2\frac{1}{n^{m-1}} \sum_{1\le i_1<\ldots<i_{m-1}\le n} h_{m-1}(X_{i_1},\ldots,X_{i_{m-1}}) \\
&\quad - \frac{1}{n^{m-1}} \sum_{1\le i_1<\ldots<i_{m-1}\le n} h_m(x,X_{i_1},\ldots,X_{i_{m-1}}) + \frac{1}{n^m} \sum_{1\le i_1<\ldots<i_m\le n} h_{m-1}(X_{i_1},\ldots,X_{i_m}) \\
&= \frac{1}{n}\sum_{i=1}^n h_1(X_i) \\
&\quad - \sum_{k=2}^m \frac{(m-k)}{n^{k-1}} \left(\sum_{1\le i_1<\ldots<i_{k-1}\le n} h_k(x,X_{i_1},\ldots,X_{i_{k-1}}) + \frac{1}{n} \sum_{1\le i_1<\ldots<i_k\le n} h_k(X_{i_1},\ldots,X_{i_k}) \right).
\end{aligned}
$$

Using this representation we can split the expected value and handle the single terms separately.

$$
\begin{aligned}
&\mathbb{E}\Big| \sum_{r=-(n-1)}^{n-1} \kappa\left(\frac{|r|}{b_n}\right) \frac{1}{n} \sum_{j=1}^{n-|r|} \Big(h_1(X_j)h_1(X_{j+|r|}) - \hat{h}_1(X_j)\hat{h}_1(X_{j+|r|}) \Big) \Big| \\
&\le \mathbb{E}\Big| \sum_{r=-(n-1)}^{n-1} \kappa\left(\frac{|r|}{b_n}\right) \frac{1}{n} \sum_{j=1}^{n-|r|} \Big((h_1(X_j) - \hat{h}_1(X_j))h_1(X_{j+|r|}) \Big) \Big| \\
&+ \mathbb{E}\Big| \sum_{r=-(n-1)}^{n-1} \kappa\left(\frac{|r|}{b_n}\right) \frac{1}{n} \sum_{j=1}^{n-|r|} \Big((h_1(X_{j+|r|}) - \hat{h}_1(X_{j+|r|}))\hat{h}_1(X_j) \Big) \Big| \\
&\le \mathbb{E}\Big| \sum_{r=-(n-1)}^{n-1} \kappa\left(\frac{|r|}{b_n}\right) \frac{1}{n} \sum_{j=1}^{n-|r|} \frac{1}{n} \sum_{i=1}^{n} h_1(X_i)h_1(X_{j+|r|}) \Big|
\end{aligned}
$$

$$
\begin{aligned}
&+\mathbb{E}\Big|\sum_{r=-(n-1)}^{n-1}\kappa\left(\frac{|r|}{b_n}\right)\frac{1}{n}\sum_{j=1}^{n-|r|}(m-2)\frac{1}{n}\sum_{i=1}^{n}h_2(X_i,X_j)h_1(X_{j+|r|})\Big|\\
&+\mathbb{E}\Big|\sum_{r=-(n-1)}^{n-1}\kappa\left(\frac{|r|}{b_n}\right)\frac{1}{n}\sum_{j=1}^{n-|r|}(m-2)\frac{1}{n^2}\sum_{1\le i_1<i_2\le n}h_2(X_{i_1},X_{i_2})h_1(X_{j+|r|})\Big|\\
&+\dots\\
&+\mathbb{E}\Big|\sum_{r=-(n-1)}^{n-1}\kappa\left(\frac{|r|}{b_n}\right)\frac{1}{n}\sum_{j=1}^{n-|r|}\frac{1}{n^{m-1}}\sum_{1\le i_1<\dots<i_{m-1}\le n}h_m(X_j,X_{i_1},\dots,X_{i_{m-1}})h_1(X_{j+|r|})\Big|\\
&+\mathbb{E}\Big|\sum_{r=-(n-1)}^{n-1}\kappa\left(\frac{|r|}{b_n}\right)\frac{1}{n}\sum_{j=1}^{n-|r|}\frac{1}{n^{m}}\sum_{1\le i_1<\dots<i_{m}\le n}h_m(X_{i_1},\dots,X_{i_m})h_1(X_{j+|r|})\Big|\\
&+\mathbb{E}\Big|\sum_{r=-(n-1)}^{n-1}\kappa\left(\frac{|r|}{b_n}\right)\frac{1}{n}\sum_{j=1}^{n-|r|}\frac{1}{n}\sum_{i=1}^{n}h_1(X_i)\hat{h}_1(X_j)\Big|\\
&+\dots\\
&+\mathbb{E}\Big|\sum_{r=-(n-1)}^{n-1}\kappa\left(\frac{|r|}{b_n}\right)\frac{1}{n}\sum_{j=1}^{n-|r|}\frac{1}{n^{m}}\sum_{1\le i_1<\dots<i_{m}\le n}h_m(X_{i_1},\dots,X_{i_m})\hat{h}_1(X_j)\Big|.
\end{aligned}
$$

We denote the $4(m-1)+2$ terms with I_i, $i=1,\dots,4(m-1)+2$.
The first term I_1 containing the first term of the Hoeffding decomposition can be handled analogously to Dehling et al. (2016), using their Lemma D.9 (in the supplementary materials). The conditions of the lemma are fulfilled, since the boundedness of the kernel h replaces the Assumption 2.3. Note that the L_2 variation condition is fulfilled since the kernel is bounded. It remains to show that our definition of NED implies the required assumptions on the P-NED process. Therefore, we want to use Lemma A.1 of Dehling et al. (2016) saying that an L_1-near epoch dependent process on $(Z_n)_{n\in\mathbb{Z}}$ with approximating constants $(a_l)_{l\in\mathbb{N}}$ is P-NED on the same process $(Z_n)_{n\in\mathbb{Z}}$. If we then choose $s_k = Ck^{-6(1+\frac{2+\delta}{\delta})}$ and $\Phi(x)=x^{-1}$ we get

$$\Phi(\epsilon)s_k = \epsilon^{-1}Ck^{-6(1+\frac{2+\delta}{\delta})} \ge k^{-(\delta+3)}\epsilon^{-1} = a_k\epsilon^{-1}$$

and hence we know that $(X_n)_{n\in\mathbb{N}}$ is P-NED with approximating constants $(s_k)_{k\in\mathbb{N}}$ and function Φ on an absolutely regular process $(Z_n)_{n\in\mathbb{Z}}$ with mixing coefficients $(\beta(l))_{l\in\mathbb{N}}$. For these coefficients it holds $s_k\Phi(k^{-6}) = O(k^{-6(2+\delta)/\delta})$ and $\sum_{k=1}^{\infty}k\beta_k^{\delta/(2+\delta)} < \infty$ and so all assumptions needed for Lemma D.9 are fulfilled.
Let us now consider all the terms, which contain h_1 and $h_k(X_j,\dots)$, $k=1,\dots,m$. These are the terms I_{2k}, $k=1,\dots,m-1$.

$$
\begin{aligned}
I_{2k} &= \mathbb{E}\Big|\sum_{r=-(n-1)}^{n-1}\kappa\left(\frac{|r|}{b_n}\right)\frac{1}{n}\sum_{j=1}^{n-|r|}(m-(k+1))\frac{1}{n^k}\sum_{1\le i_1<\dots<i_k\le n}h_{k+1}(X_j,X_{i_1},\dots,X_{i_k})h_1(X_{j+|r|})\Big|\\
&= \mathbb{E}\Big|\frac{1}{n}\sum_{j_1=1}^{n}\frac{m-(k+1)}{n^k}\sum_{1\le i_1<\dots<i_{k+1}\le n}h_{k+1}(X_{i_1},\dots,X_{i_{k+1}})h_1(X_{j_1})\kappa\left(\frac{|i_1-j_1|}{b_n}\right)\Big|
\end{aligned}
$$

$$\leq \mathbb{E}\Bigg(\Bigg(\Bigg(\frac{1}{n}\sum_{j_1=1}^{n} h_1(X_{j_1})\frac{m-(k+1)}{n^k} \sum_{1\leq i_1<\ldots<i_{k+1}\leq n} h_{k+1}(X_{i_1},\ldots,X_{i_{k+1}})\kappa\left(\frac{|i_1-j_1|}{b_n}\right)\Bigg)^2\Bigg)^{\frac{1}{2}}\Bigg)$$

$$\leq \mathbb{E}\Bigg(\left(\frac{1}{n}\sum_{j_1=1}^{n} h_1(X_{j_1})^2\right)^{\frac{1}{2}} \left(\frac{1}{n}\sum_{j_1=1}^{n}\left(\frac{m-(k+1)}{n^k}\sum_{1\leq i_1<\ldots<i_{k+1}\leq n} h_{k+1}(X_{i_1},\ldots,X_{i_{k+1}})\kappa\left(\frac{|i_1-j_1|}{b_n}\right)\right)^2\right)^{\frac{1}{2}}\Bigg)$$

$$\leq \left(\mathbb{E}\left(\frac{1}{n}\sum_{j_1=1}^{n} h_1(X_{j_1})^2\right)\right)^{\frac{1}{2}} \left(\mathbb{E}\left(\frac{1}{n}\sum_{j_1=1}^{n}\left(\frac{m-(k+1)}{n^k}\sum_{1\leq i_1<\ldots<i_{k+1}\leq n} h_{k+1}(X_{i_1},\ldots,X_{i_{k+1}})\kappa\left(\frac{|i_1-j_1|}{b_n}\right)\right)^2\right)\right)^{\frac{1}{2}},$$

where we used the Cauchy-Schwarz inequality and in the last step the Hölder-inequality. Due to the boundedness of h and therefore also of h_1 the first term can simply be estimated by a constant C.

$$I_{2k} \leq \left(\mathbb{E}\left(\frac{1}{n}\sum_{j_1=1}^{n} h_1(X_{j_1})^2\right)\right)^{\frac{1}{2}} \left(\mathbb{E}\left(\frac{(m-(k+1))^2}{n^{2k+1}}\sum_{j_1=1}^{n}\left(\sum_{1\leq i_1<\ldots<i_{k+1}\leq n} h_{k+1}(X_{i_1},\ldots,X_{i_{k+1}})\kappa\left(\frac{|i_1-j_1|}{b_n}\right)\right)^2\right)\right)^{\frac{1}{2}}$$

$$\leq C\frac{m-(k+1)}{n^{k+\frac{1}{2}}}\left(\mathbb{E}\left(\sum_{j_1=1}^{n}\left(\sum_{1\leq i_1<\ldots<i_{k+1}\leq n} h_{k+1}(X_{i_1},\ldots,X_{i_{k+1}})\kappa\left(\frac{|i_1-j_1|}{b_n}\right)\right)^2\right)\right)^{\frac{1}{2}}$$

$$\leq C'\frac{1}{n^{k+\frac{1}{2}}}\Bigg(\mathbb{E}\Bigg(\sum_{i_1,\ldots,i_{2(k+1)}=1}^{n} h_{k+1}(X_{i_1},\ldots,X_{i_{k+1}})h_{k+1}(X_{i_{k+2}},\ldots,X_{i_{2(k+1)}}) \sum_{j_1=1}^{n}\kappa\left(\frac{|i_1-j_1|}{b_n}\right)\kappa\left(\frac{|i_{k+2}-j_1|}{b_n}\right)\Bigg)\Bigg)^{\frac{1}{2}}. \tag{4.3}$$

To show the convergence of I_{2k} to zero we finally want to apply Lemma 4.3. For this we have to show that

$$\sum_{l=0}^{n} l(\beta(l)+A_l) = O(n^\gamma)$$

with $A_l = \sqrt{2\sum_{i=l}^{\infty} a_i}$. From the assumption $\sum_{l=1}^{\infty} l\beta^{\delta/(2+\delta)}(l) < \infty$, which implies $l\beta^{\delta/(2+\delta)}(l) \to 0$ for $l \to \infty$, and the fact that the mixing coefficients $(\beta(l))_{l\in\mathbb{N}}$ are non-negative and monotone decreasing (therefore $\beta^{\delta/(2+\delta)}(l)$ is also monotone decreasing) we know that $l\beta_l^{\delta/(2+\delta)}$ is monotone decreasing and positive. Hence,

$$l\beta^{\delta/(2+\delta)(l)} = O\left(l^{-1}\right)$$

and so $\beta(l) = O\left(l^{-\eta}\right)$ for a $\eta > 2$. Analogously to the proof of Theorem 4.1, but now for $a_l = O\left(l^{-\delta-3}\right)$, we then can show

$$\sum_{l=0}^{n} l(\beta(l) + A_l) \leq C \sum_{l=1}^{n} l^{-\eta+1} + O(n^{2-\frac{\delta+2}{2}}),$$

where $2 - \frac{\delta+2}{2} < \frac{1}{2}$ since $\delta > 11$. Let us now have a closer look at the first term. We want to show that $C\sum_{l=1}^{n} l^{-\eta+1} = O(n^{\gamma})$ for a $0 < \gamma < 1/2$. It is

$$\frac{\sum_{l=1}^{n} l^{-\eta+1}}{n^{\gamma}} \leq \inf_{1\leq t\leq n} t^{-\gamma} \sum_{l=1}^{n} l^{1-\eta} \leq \sum_{l=1}^{n} l^{-1-\eta} l^{-\gamma} = \sum_{l=1}^{n} \left(\frac{1}{l}\right)^{-1+\eta+\gamma}.$$

This is the Dirichlet series and it converges for $-1+\eta+\gamma > 1$. Since $\eta > 2$ and $0 < \gamma < 1/2$ we have $\sum_{l=1}^{n} l^{-\eta+1} = O(n^{\gamma})$ and therefore

$$\sum_{l=0}^{n} l(\beta(l) + A_l) = O(n^{\gamma})$$

for a $0 < \gamma < 1/2$.
Now we can apply Lemma 4.3 with

$$c_{i_1,i_{k+2}} = \sum_{j_1=1}^{n} \kappa\left(\frac{|i_1 - j_1|}{b_n}\right) \kappa\left(\frac{|i_{k+2} - j_1|}{b_n}\right) = O(b_n)$$

to (4.3) and obtain

$$\begin{aligned} I_{2k} \leq C \frac{1}{n^{k+\frac{1}{2}}} \left(n^{2(k+1)-2+\gamma} b_n\right)^{\frac{1}{2}} &\leq C \left(\frac{n^{2k+\gamma}}{n^{2k+1}} b_n\right)^{\frac{1}{2}} \\ &= C \left(n^{\gamma-\frac{1}{2}}\right)^{\frac{1}{2}} \left(\frac{b_n}{\sqrt{n}}\right)^{\frac{1}{2}} \longrightarrow 0, \end{aligned}$$

because of Assumption 4.1 and $0 < \gamma < 1/2$.
Therefore, I_{2k} converges to zero for all $k = 1, \ldots, m-1$.

It remains to show the convergence of the remaining terms. The terms containing $h_1(\cdot)$ and $h_k(X_{i_1}, \ldots, X_{i_k})$, $k = 2, \ldots, m$, are denoted with I_{2k+1}, $k = 1, \ldots, m-1$.

$$I_{2k+1} = \mathbb{E}\left| \sum_{r=-(n-1)}^{n-1} \kappa\left(\frac{|r|}{b_n}\right) \frac{1}{n} \sum_{j=1}^{n-|r|} \frac{m-(k+1)}{n^{k+1}} \sum_{1\leq i_1<\ldots<i_{k+1}\leq n} h_{k+1}(X_{i_1}, \ldots, X_{i_{k+1}}) h_1(X_{j+|r|}) \right|$$

$$
\begin{aligned}
&\leq \left\| \frac{m-(k+1)}{n^{k+1}} \sum_{1\leq i_1<\ldots<i_{k+1}\leq n} h_{k+1}(X_{i_1},\ldots,X_{i_{k+1}}) \right\|_2 \\
&\quad \left\| \frac{1}{n} \sum_{j=1}^{n-|r|} h_1(X_{j+|r|}) \sum_{j_1=1}^{n} \kappa\left(\frac{|j-j_1|}{b_n}\right) \right\|_2 \\
&\leq C \frac{1}{n^{k+1}} \left(n^{2(k+1)-\gamma}\right)^{\frac{1}{2}} \frac{b_n}{\sqrt{n}} \\
&= C \frac{n^{k+\gamma/2}}{n^{k+1}} o(1) \longrightarrow 0,
\end{aligned}
$$

where we used the Hölder inequality, the boundedness of h_{k+1}, Lemma 4.3 and Assumption 4.1 as before.
The convergence of the remaining terms I_{2m}, I_{2k} and I_{2k+1} for $k = m+1,\ldots,2m-1$ can be shown analogously and is therefore omitted. □

A generalisation of U-statistics are U-processes. They are introduced in the following section.

4.2. U-processes

If the kernel h of a U-statistic should no longer be a fixed function but should have variable arguments, the extension from U-statistics to U-processes is necessary.

Definition 4.5. *Let $h : \mathbb{R}^{m+1} \to \mathbb{R}$ be a measurable and bounded function, symmetric in the first m arguments and non-decreasing in the last one. Suppose that for all $x_1,\ldots,x_m \in \mathbb{R}$ we have $\lim_{t\to\infty} h(x_1,\ldots,x_m,t) = 1$, $\lim_{t\to-\infty} h(x_1,\ldots,x_m,t) = 0$. We call the process $(U_n(t))_{t\in\mathbb{R}}$, which is simply a U-statistic with an additional, varying argument in the kernel, empirical U-distribution function. As U-distribution function we define $U(t) := \mathbb{E}\left(h(Y_1,\ldots,Y_m,t)\right)$ for independent copies $Y_1,\ldots,Y_m$ of X_1. Then the empirical process is defined as*

$$(\sqrt{n}(U_n(t) - U(t)))_{t\in\mathbb{R}}.$$

Remark 4.3. Without restriction we can choose the space of the parameter t as the compact interval $[0,1]$, since a transformation for example via the distribution function does not influence the dependence structure as it is Lipschitz, monotone etc.

We will mainly consider the function H_n of the empirical U-process in the following, where $U_n(t)$ has the kernel
$g(x_1,\ldots,x_m,t) = \mathbb{1}_{[h(x_1,\ldots,x_m)\leq t]}$. Therefore,

$$U(t) = \mathbb{E}\left(\mathbb{1}_{[h(Y_1,\ldots,Y_m)\leq t]}\right) = \mathbb{P}(h(Y_1,\ldots,Y_m) \leq t) = H_F(t)$$

and since H_F has density $h_F < \infty$ we have that H_F is Lipschitz-continuous.
For U-processes the Hoeffding decomposition is needed, too. In this case, additionally the parameter t has to be considered though the definition is very similar to the one in the case of U-statistics.

Definition 4.6. *(Hoeffding decomposition for U-processes)*
Let U_n be a U-process with kernel $h = h(x_1, \dots, x_m, t)$. Then one can write $U_n(t)$ as

$$U_n(t) = U(t) + \sum_{j=1}^{m} \binom{m}{j} \frac{1}{\binom{n}{j}} S_{jn},$$

where

$$\begin{aligned}
U(t) &= \mathbb{E}(h(Y_1, \dots, Y_m, t)) \\
S_{jn} &= \sum_{1 \le i_1 < \dots < i_j \le n} h_j(X_{i_1}, \dots, X_{i_j}, t) \\
h_1(x_1, t) &= \tilde{h}_1(x_1, t) \\
h_2(x_1, x_2, t) &= \tilde{h}_2(x_1, x_2, t) - h_1(x_1, t) - h_1(x_2, t) \\
h_3(x_1, x_2, x_3, t) &= \tilde{h}_3(x_1, x_2, x_3, t) - \sum_{i=1}^{3} h_1(x_i, t) - \sum_{1 \le i < j \le 3} h_2(x_i, x_j, t) \\
&\dots \\
h_m(x_1, \dots, x_m, t) &= \tilde{h}_m(x_1, \dots, x_m, t) - \sum_{i=1}^{m} h_1(x_i, t) - \sum_{1 \le i_1 < i_2 \le m} h_2(x_{i_1}, x_{i_2}, t) \\
&\quad - \dots - \sum_{1 \le i_1 < \dots < i_{m-1} \le m} h_{m-1}(x_{i_1}, \dots, x_{i_{m-1}}, t) \\
\tilde{h}_j(x_1, \dots, x_j, t) &= \mathbb{E}(h(x_1, \dots, x_j, Y_{j+1}, \dots, Y_m, t)) - U(t).
\end{aligned}$$

for independent copies $Y_1, \dots, Y_m$ of X_1.

The following proposition summarises two important properties of the kernel of a U-process that are very helpful for proving the NED case of the invariance principle.

Proposition 4.1. *Let $h : \mathbb{R}^{m+1} \longrightarrow \mathbb{R}$ be the kernel of a U-process. Then the following holds:*

(i) If for any random vector $(Z_1, \dots, Z_m)$

$$\mathbb{E}|h(Z_1, \dots, Z_m, t) - h(Z_1, \dots, Z_m, s)| \le C|t - s|$$

for an arbitrary constant $C > 0$, then $U(t)$ and $h_k(\cdot, \dots, \cdot, t)$ are Lipschitz continuous in t for every $1 \le k \le m$.

(ii) If for any random vector $(Z_1, \dots, Z_m)$, any constants $u_i \in \mathbb{R}$, $i = 1, \dots, m$, and any $t \in \mathbb{R}$ it holds that

$$\mathbb{E}\,|h(Z_1, \dots, Z_m, t) - h(Z_1 + u_i, \dots, Z_m + u_m, t)| \le C \sum_{i=1}^{m} |u_i|,$$

then for every $1 \le k \le m$ it holds that

$$\mathbb{E}\,|h_k(Z_1, \dots, Z_k, t) - h_k(Z_1 + u_i, \dots, Z_k + u_k, t)| \le C \sum_{i=1}^{k} |u_i|.$$

Proof. (i)
Remember the definition $U(t) = \mathbb{E}\left(h(Y_1, \ldots, Y_m, t)\right)$. Therefore, for any $s, t \in \mathbb{R}$

$$|U(t) - U(s)| = |\mathbb{E}\left(h(Y_1, \ldots, Y_m, t) - h(Y_1, \ldots, Y_m, s)\right)| \leq C|t - s|.$$

The remaining proof is given with the induction principle. First, let be $k = 1$.

$$\begin{aligned} |h_1(x,t) - h_1(x,s)| &= |\mathbb{E}\left(h(x, Y_2, \ldots, Y_m, t)\right) - U(t) - \mathbb{E}(h(x, Y_2, \ldots, Y_m, s)) + U(s)| \\ &\leq |\mathbb{E}(h(x, Y_2, \ldots, Y_m, t) - h(x, Y_2, \ldots, Y_m, s))| + C|t - s| \\ &\leq C'|t - s|, \end{aligned}$$

where we used the Lipschitz-continuity of U.
Now we choose k arbitrary and assume that the results hold for $(k-1)$.

$$\begin{aligned} &|h_k(x_1, \ldots, x_k, t) - h_k(x_1, \ldots, x_k, s)| \\ &= \Big|\mathbb{E}\left(h(x_1 \ldots, x_k, Y_{k+1}, \ldots, Y_m, t) - h(x_1 \ldots, x_k, Y_{k+1}, \ldots, Y_m, s)\right) \\ &- \sum_{i=1}^{k}(h_1(x_i, t) - h_1(x_i, s)) - \ldots - \sum_{1 \leq i_1 \ldots, i_k \leq k} (h_{k-1}(x_1, \ldots, x_{k-1}, t) - h_{k-1}(x_1, \ldots, x_{k-1}, s)) \\ &- U(t) + U(s)\Big| \\ &\leq C|t - s| + C_1 \sum_{i=1}^{k}|t - s| + \ldots + C_{k-1} \sum_{1 \leq i_1, \ldots, i_{k-1} \leq k} |t - s| + C'|t - s| \\ &\leq C''|t - s|, \end{aligned}$$

where we used the assumed Lipschitz-continuity of h_i, $i = 1, \ldots, k-1$, and the same argument as in the case $k = 1$.

(ii)
Using the same principle as in *(i)* we first show the results for $k = 1$.

$$\begin{aligned} &\mathbb{E}\,|h_1(Z_1, t) - h_1(Z_1 + u, t)| \\ &= \mathbb{E}\Big|\mathbb{E}\left(h(Z_1, Y_2, \ldots, Y_m, t)\right) - U(t) - \mathbb{E}(h(Z_1 + u, Y_2, \ldots, Y_m, t)) + U(t)| \\ &\leq \mathbb{E}|\mathbb{E}\left(h(Z_1, Y_2, \ldots, Y_m, t)\right) - \mathbb{E}(h(Z_1 + u, Y_2, \ldots, Y_m, t))\Big| \\ &\leq \mathbb{E}\left(\mathbb{E}|h(Z_1, Y_2, \ldots, Y_m, t) - h(Z_1 + u, Y_2, \ldots, Y_m, t)|\right) \\ &\leq \mathbb{E}(C|u|) = C|u|. \end{aligned}$$

For arbitrary k and assuming that the result is valid for $(k-1)$ we have

$$\begin{aligned} &\mathbb{E}|h_k(Z_1, \ldots, Z_k, t) - h_k(Z_1 + u_1, \ldots, Z_k + u_k, t)| \\ &= \mathbb{E}\Big|\mathbb{E}\left(h(Z_1, \ldots, Z_k, Y_{k+1}, \ldots, Y_m, t) - h(Z_1 - u_1, \ldots, Z_k - u_k, Y_{k+1}, \ldots, Y_m, t)\right) \\ &- \sum_{i=1}^{k}(h_1(X_i, t) - h_1(X_i, t)) \\ &- \ldots \end{aligned}$$

$$
\begin{aligned}
&- \sum_{1\le i_1\ldots,i_k\le k} (h_{k-1}(Z_{i_1},\ldots,Z_{i_{k-1}},t) - h_{k-1}(Z_{i_1}-u_{i_1},\ldots,Z_{i_{k-1}}-u_{i_k},t)) - U(t) + U(t)\Big| \\
&\le C\sum_{i=1}^{k}|u_i| + \sum_{i=1}^{k} C|u_i| + \ldots + \sum_{1\le i_1<\ldots<i_{k-1}\le k} C\sum_{j=1}^{k-1} C|u_{i_j}| \\
&\le C\sum_{i=1}^{k}|u_i|
\end{aligned}
$$

□

Remark 4.4. Note that if the kernel h is of the form $h(x_1,\ldots,x_m,t) = \mathbb{1}_{[h'(x_1,\ldots,x_m)\le t]}$ and h' has density h_F, it follows automatically from the condition in Proposition 4.1 *(i)* that h_F is bounded since

$$\mathbb{E}|h(Y_1,\ldots,Y_m,t) - h(Y_1,\ldots,Y_m,s)| = \mathbb{P}\left(t \le h'(Y_1,\ldots,Y_m) \le s\right) = |t-s| \sup_{x\in\mathbb{R}} h_F(x).$$

Therefore, a special case of this Proposition is:
Let h be of the special form $h(x_1\ldots,x_m,t) = \mathbb{1}_{[h'(x_1,\ldots,x_m)\le t]}$ and let H_F be the distribution function of the kernel h' with density h_F. If $h_{F;X_1,\ldots,X_k}$ and h_F are bounded, then $h_k(\cdot,\ldots,\cdot,t)$ is Lipschitz continuous in t for every $1 \le k \le m$.

Additionally, the extended variation condition has to be transformed. Fischer et al. (2016a) used the extended uniform variation condition, which has the same properties as the extended variation condition.

Definition 4.7. *We say h satisfies the extended uniform variation condition, if the extended variation condition holds for $h(x_1,\ldots,x_m,t)$ with a constant not depending on t.*

Please note, that many results given for the Hoeffding-kernels of U-statistics, e.g. Lemma 4.3, remain valid for the kernels of a U-process.

Now we want to establish an invariance principle for the U-process. For near epoch dependent sequences on absolutely regular processes it has already been shown by Borovkova et al. (2001) and Dehling and Philipp (2002), a result for strong mixing can be found in Wendler (2011b). Under independence one can find a strong invariance principle in Dehling et al. (1987). However, these results only consider the case of a bivariate kernel and therefore exclude such examples as are given in Section 1. The result given in the next theorem closes this gap.
From now on only consider the case where H_n is our empirical U-process, that is $U_n(t)$ has the kernel $g(x_1,\ldots,x_m,t) = \mathbb{1}_{[h(x_1,\ldots,x_m)\le t]}$.

Theorem 4.4. *Let h be a kernel with distribution function H_F and related density $h_F < \infty$. Moreover, let g_1 be the first term of the Hoeffding decomposition of H_n. Let $(X_n)_{n\in\mathbb{N}}$ be NED with approximation constants $(a_l)_{l\in\mathbb{N}}$ on an absolutely regular process $(Z_n)_{n\in\mathbb{Z}}$ with mixing coefficients $(\beta(l))_{l\in\mathbb{N}}$ with $\sum_{l=1}^{\infty} l^2\beta^{\frac{\delta}{2+\delta}}(l) < \infty$ for a $0<\delta<1$. Moreover, let $\sum_{l=1}^{\infty} l^2 a_l^{\frac{\delta}{2+2\delta}} < \infty$. Then*

$$\left(\frac{m}{\sqrt{n}}\sum_{i=1}^{n} g_1(X_i,t)\right)_{t\in\mathbb{R}} \xrightarrow{D} (W(t))_{t\in\mathbb{R}},$$

where W is a Gaussian process having continuous sample paths with probability 1.

Remark 4.5. Note that the condition $\sum_{l=1}^{\infty} l^2 \left(L\sqrt{2A_l}\right)^{\frac{\delta}{1+\delta}} < \infty$ of Dehling and Philipp (2002), where L is the variation constant of the kernel $g(x_1, \ldots, x_m, t) = \mathbb{1}_{[h(x_1,\ldots,x_m)\leq t]}$ and $A_l = \sqrt{2\sum_{i=l}^{\infty} a_i}$, follows directly from $\sum_{l=1}^{\infty} l^2 a_l^{\frac{\delta}{2+2\delta}} < \infty$, since

$$\begin{aligned}
&\sum_{l=1}^{\infty} l^2 \left(L\sqrt{2A_l}\right)^{\frac{\delta}{1+\delta}} = \sum_{l=1}^{\infty} l^2 \left(L\sqrt{2\sqrt{2\sum_{i=l}^{\infty} a_i}}\right)^{\frac{\delta}{1+\delta}} \\
&\leq \sum_{l=1}^{\infty} l^2 C \left(\sum_{i=l}^{\infty} a_i\right)^{\frac{\delta}{2(1+\delta)}} \leq \sum_{l=1}^{\infty} C l^2 \sum_{i=l}^{\infty} a_i^{\frac{\delta}{2(1+\delta)}} \\
&\leq \sum_{l=1}^{\infty} C l^2 \sum_{i=l}^{\infty} \frac{1}{i^2} i^2 a_i^{\frac{\delta}{2(1+\delta)}} \leq \sum_{l=1}^{\infty} C l^2 \frac{1}{l^2} \sum_{i=l}^{\infty} i^2 a_i^{\frac{\delta}{2(1+\delta)}} \\
&\leq C \sum_{l=1}^{\infty} l^2 a_l^{\frac{\delta}{2+2\delta}} < \infty.
\end{aligned}$$

Thus, in the NED case we can reduce the assumptions needed for the proof compared to the strong mixing case in Fischer et al. (2016a).

Theorem 4.4 can be proved in the same way as Theorem 4.8 of Dehling and Philipp (2002) by choosing $g_t(x) = g_1(x,t)$ and $G(t) = H_F(t)$ (preserving the properties of the single functions) and is therefore omitted.

Corollary 4.1. *Let $(X_n)_{n\in\mathbb{N}}$ be NED with approximation constants $(a_l)_{l\in\mathbb{N}}$ on an absolutely regular process $(Z_n)_{n\in\mathbb{Z}}$ with mixing coefficients $(\beta(l))_{l\in\mathbb{N}}$ with $\sum_{l=1}^{\infty} l^2 \beta^{\frac{\gamma}{2+\gamma}}(l) < \infty$, $0 < \gamma < 1$. Additionally, let be $\sum_{l=1}^{\infty} l^2 a_l^{\frac{\gamma}{2+2\gamma}} < \infty$. Moreover, let h be a Lipschitz-continuous kernel with distribution function H_F and related density $h_F < \infty$ and for all $2 \leq k \leq m$ let $h_{F;X_2,\ldots,X_k}$ be bounded. Then*

$$\sup_{t\in\mathbb{R}} \left|\sqrt{n}\left(H_n(t) - H_F(t)\right)\right| = O_p(1).$$

This Corollary is a straightforward consequence from Theorem 4.4 using Lemma 4.4 given at the end of this section. The assumptions on the coefficients a_l and β_l of Lemma 4.4 are automatically fulfilled by the assumptions of Theorem 4.4.

Proof.
Again, we use the Hoeffding decomposition getting

$$\begin{aligned}
&\sup_{t\in\mathbb{R}} \left|\sqrt{n}\left(H_n(t) - H_F(t)\right)\right| \\
&= \sup_{t\in\mathbb{R}} \left|\sqrt{n}\left(H_F(t) + \sum_{j=1}^{m} \binom{m}{j} \frac{1}{\binom{n}{j}} S_{jn,t} - H_F(t)\right)\right| \\
&= \sup_{t\in\mathbb{R}} \left|\frac{m}{\sqrt{n}} \sum_{i=1}^{n} g_1(X_i,t) + \sqrt{n}\frac{\binom{m}{2}}{\binom{n}{2}} \sum_{1\leq i<j\leq n} g_2(X_i,X_j,t)\right. \\
&\left. + \ldots + \sqrt{n}\frac{1}{\binom{n}{m}} \sum_{1\leq i_1<\ldots<i_m\leq n} g_m(X_{i_1},\ldots,X_{i_m},t)\right|
\end{aligned}$$

$$\leq \sup_{t\in\mathbb{R}} \left| \frac{m}{\sqrt{n}} \sum_{i=1}^{n} g_1(X_i,t) \right| + \sup_{t\in\mathbb{R}} \left| \sqrt{n} \frac{\binom{m}{2}}{\binom{n}{2}} \sum_{1\leq i<j\leq n} g_2(X_i,X_j,t) \right|$$
$$+ \ldots + \sup_{t\in\mathbb{R}} \left| \sqrt{n} \frac{1}{\binom{n}{m}} \sum_{1\leq i_1<\ldots<i_m\leq n} g_m(X_{i_1},\ldots,X_{i_m},t) \right|.$$

For the first term of the sum above we get, using Theorem 4.4 and the continuous mapping theorem,

$$\sup_{t\in\mathbb{R}} \left| \frac{m}{\sqrt{n}} \sum_{i=1}^{n} g_1(X_i,t) \right| \to \|W\|_\infty.$$

Thus, we have $\|W\|_\infty = O_p(1)$.

For the remaining terms we want to apply Lemma 4.4. Therefore, the kernel of the U-process $g(x_1,\ldots,x_m,t) = 1_{[h(x_1,\ldots,x_m)\leq t]}$ has to satisfy the extended uniform variation condition. We use the Lipschitz continuity of h for this.

$$\sup_{\|(x_1,\ldots,x_m)-(X'_1,\ldots,X'_m)\|\leq\epsilon} \left| \mathbb{1}_{[h(X'_1,\ldots,X'_m)\leq t]} - \mathbb{1}_{[h(x_1,\ldots,x_m)\leq t]} \right|$$
$$= \begin{cases} 1 & \text{, if } h(X'_1,\ldots,X'_m) \in (t-L\epsilon, t+L\epsilon) \\ 0 & \text{, else} \end{cases}$$

and so

$$\mathbb{E}\left(\sup_{\|(x_1,\ldots,x_m)-(X'_1,\ldots,X'_m)\|\leq\epsilon} \left| \mathbb{1}_{[h(X'_1,\ldots,X'_m)\leq t]} - \mathbb{1}_{[h(x_1,\ldots,x_m)\leq t]} \right| \right)$$
$$\leq \sup_{t\in\mathbb{R}} \left| \mathbb{E}\left(\mathbb{1}_{[h(X'_1,\ldots,X'_m)\in(t-L\epsilon,t+L\epsilon)]} \right) \right|$$
$$\leq \sup_{t\in\mathbb{R}} \left| \int_{t-L\epsilon}^{t+L\epsilon} h_F(x)dx \right| \leq 2L\epsilon(\sup_{x\in\mathbb{R}} h_F(x)) \leq L'\epsilon,$$

since h_F is bounded and therefore g fulfils the uniform variation condition.
Using the same arguments we can also show that g satisfies the extended uniform variation condition. For arbitrary $2 \leq k \leq m$ and $i_1 < i_2 < \ldots < i_m$ it is

$$\mathbb{E}\left(\sup_{|x_1-Y_{i_1}|\leq\epsilon} \left| \mathbb{1}_{[h(Y_{i_1},X_{i_2},\ldots,X_{i_k},Y_{i_{k+1}},\ldots,Y_{i_m})\leq t]} - \mathbb{1}_{[h(x_1,X_{i_2},\ldots,X_{i_k},Y_{i_{k+1}},\ldots,Y_{i_m})\leq t]} \right| \right)$$
$$\leq \sup_{t\in\mathbb{R}} \left| \int_{t-L\epsilon}^{t+L\epsilon} h_{F;X_{i_2},\ldots,X_{i_k}}(x)dx \right| \leq L\epsilon.$$

Applying Lemma 4.4 we get for $2 \leq k \leq n$

$$\sup_{t\in\mathbb{R}} \left| \sqrt{n} \frac{\binom{m}{k}}{\binom{n}{k}} \sum_{1\leq i_1,\ldots,i_k\leq n} g_k(X_{i_1},\ldots,X_{i_k},t) \right| \leq \sqrt{n} n^{-k} o_p(n^{k-\frac{1}{2}-\frac{1}{8}\frac{\eta-3}{\eta+1}}) = o_p(n^{-\frac{1}{8}\frac{\eta-3}{\eta+1}}).$$

With Slutsky's Theorem the proof is then completed. □

The used Lemma 4.4 is stated in the following section.

Preliminary results

The following lemma has already been proven by Wendler (2011a) in the case $m = 2$. We will modify the main idea of the proof to obtain a similar result for the degenerated terms of higher dimensional U-processes.

Lemma 4.4. *Let h be a kernel satisfying the extended uniform variation condition, such that the U-distribution function U is Lipschitz continuous. Moreover, let $(X_n)_{n\in\mathbb{N}}$ be NED with approximation constants $a_l = O(l^{-c})$, where $c = \max\{\eta + 3, 12\}$, on an absolutely regular process $(Z_n)_{n\in\mathbb{Z}}$ with mixing coefficients $\beta(l) = O(l^{-\eta})$ for an $\eta \geq 8$. Then, for all $2 \leq k \leq m$ and $\tau = \frac{\eta-3}{\eta+1}$, we have*

$$\sup_{t\in\mathbb{R}} \left| \sum_{1\leq i_1,\ldots,i_k\leq n} g_k(X_{i_1}, \ldots, X_{i_k}, t) \right| = o_p(n^{k-\frac{1}{2}-\frac{\tau}{8}}) \text{ a.s.}.$$

The lemma can be proved analogously to Fischer (2013) respectively Fischer et al. (2016a) but due to completeness and to show where the newly established results are needed we state the proof here.

Proof.
We define $Q_n^k(t) := \sum_{1\leq i_1,\ldots,i_k\leq n} g_k(X_{i_1}, \ldots, X_{i_k}, t)$.
For $l \in \mathbb{N}$ choose $t_{1,l}, \ldots, t_{s-1,l}$ with $s = s_l = O(2^{\frac{5}{8}l})$, such that

$$-\infty = t_{0,l} < t_{1,l} < \ldots < t_{s-1,l} < t_{s,l} = \infty$$

and $2^{-\frac{5}{8}l} \leq |U(t_{r,l}) - U(t_{r-1,l})| \leq 2 \cdot 2^{-\frac{5}{8}l}$ for all $r = 1, \ldots, s$. Since we required Lipschitz-continuity of U it follows that $2^{-\frac{5}{8}l} \leq C|t_{r,l} - t_{r-1,l}|$. Moreover,
$\mathbb{E}\left(h(Y_1, \ldots, Y_k, Y_{k+1}, \ldots Y_m, t)|Y_1 = X_{i_1}, \ldots, Y_k = X_{i_k}\right)$ is non-decreasing in t for all $2 \leq k \leq m$, because h is non-decreasing in t.
The case $k = 2$ has been treated by Wendler (2011a) and is therefore omitted here.

From now on suppose that the statement of the lemma is valid for $k - 1$.
Together with the above consideration we have for every $t \in [t_{r-1,l}, t_{r,l}]$ and $2^l \leq n < 2^{l+1}$

$$\begin{aligned}
&|Q_n^k(t)| \\
=&\Bigg| \sum_{1\leq i_1<\ldots<i_k\leq n} \big(\mathbb{E}(h(Y_1, \ldots Y_k, Y_{k+1}, \ldots, Y_m, t)|Y_1 = X_{i_1}, \ldots, Y_k = X_{i_k}) \\
&- g_1(X_{i_1}, t) - \cdots - g_1(X_{i_k}, t) - g_2(X_{i_1}, X_{i_2}, t) \\
&- \cdots - g_2(X_{i_{k-1}}, X_{i_k}, t) - \cdots - U(t)\big)\Bigg| \\
\leq& \max\Bigg\{\Bigg| \sum_{1\leq i_1<\ldots<i_k\leq n} \big(\mathbb{E}(h(X_{i_1}, \ldots, X_{i_k}, Y_{i_{k+1}}, \ldots, Y_m, t_{r,l}) \\
&- g_1(X_{i_1}, t_{r,l}) - \ldots - g_1(X_{i_k}, t_{r,l}) - g_2(X_{i_1}, X_{i_2}, t_{r,l}) - \cdots - g_2(X_{i_{k-1}}, X_{i_k}, t_{r,l}) \\
&- \cdots - U(t_{r,l}))\big)\Bigg|,
\end{aligned}$$

$$
\begin{aligned}
&\Bigg| \sum_{1\le i_1<\ldots<i_k\le n} \big(\mathbb{E}(h(X_{i_1},\ldots,X_{i_k},Y_{i_{k+1}},\ldots,Y_m,t_{r-1,l})) \\
&\quad - g_1(X_{i_1},t_{r-1,l}) - \ldots - g_1(X_{i_k},t_{r-1,l}) - g_2(X_{i_1},X_{i_2},t_{r-1,l}) - \cdots - g_2(X_{i_{k-1}},X_{i_k},t_{r-1,l}) \\
&\qquad - \cdots - U(t_{r-1,l})\big)\Bigg|\Bigg\} \\
&\quad + \binom{n-1}{k-1} \max\left\{ \left|\sum_{i=1}^n (g_1(X_i,t_{r,l}) - g_1(X_i,t)) \right|, \left|\sum_{i=1}^n (g_1(X_i,t) - g_1(X_i,t_{r-1,l})) \right| \right\} \\
&\quad + \binom{n-2}{k-2} \max\left\{ \left|\sum_{i=1}^n (g_2(X_{i_1},X_{i_2},t_{r,l}) - g_2(X_{i_1},X_{i_2},t)) \right|, \right. \\
&\qquad\qquad \left. \left|\sum_{i=1}^n (g_2(X_{i_1},X_{i_2},t) - g_2(X_{i_1},X_{i_2},t_{r-1,l})) \right| \right\} \\
&\quad + \cdots \\
&\quad + \binom{n}{k} |U(t_{r,l}) - U(t_{r-1,l})| \\
&\le \max\{|Q_n^k(t_{r,l})|, |Q_n^k(t_{r-1,l})|\} \\
&\quad + \binom{n-1}{k-1} \left|\sum_{i=1}^n (g_1(X_i,t_{r,l}) - g_1(X_i,t_{r-1,l})) \right| \\
&\quad + \binom{n-2}{k-2} \left| \sum_{1\le i_1<i_2\le n} (g_2(X_{i_1},X_{i_2},t_{r,l}) - g_2(X_{i_1},X_{i_2},t_{r-1,l})) \right| \\
&\quad + \cdots \\
&\quad + \binom{n-(k-1)}{k-(k-1)} \left| \sum_{1\le i_1<\ldots<i_{k-1}\le n} \big(g_{k-1}(X_{i_1},\ldots,X_{i_{k-1}},t_{r,l}) - g_{k-1}(X_{i_1},\ldots,X_{i_{k-1}},t_{r-1,l})\big) \right| \\
&\quad + \binom{n}{k} |U(t_{r,l}) - U(t_{r-1,l})|.
\end{aligned}
$$

Again we will treat the first, second and last of the terms separately.
For treating the first term we use that

$$
\begin{aligned}
&\mathbb{E}\left(\max_{n=2^l,\ldots,2^{l+1}-1} \max_{r=0,\ldots,s} |Q_n^k(t_{r,l})|^2 \right) \\
&\le \sum_{r=0}^s \mathbb{E}\left(\left(\sum_{d=0}^l \max_{i=1,\ldots,2^{l-d}} |Q^k_{2^l+i2^d}(t_{r,l}) - Q^k_{2^l+(i-1)2^d}(t_{r,l})| \right)^2 \right) \\
&\le \sum_{r=0}^s l \sum_{d=0}^l \sum_{i=1}^{2^{l-d}} \mathbb{E}\left(\left(Q^k_{2^l+i2^d}(t_{r,l}) - Q^k_{2^l+(i-1)2^d}(t_{r,l}) \right)^2 \right) \\
&\le \sum_{r=0}^s l \sum_{d=0}^l \underbrace{\sum_{i_1,\ldots,i_{2k}=1}^{2^{l+1}} |\mathbb{E}\left(g_k(X_{i_1},\ldots,X_{i_k},t) g_k(X_{i_{k+1}},\ldots,X_{i_{2k}},t)\right)|}_{=O((2^{l+1})^{2k-2+\gamma}), \text{ with Lemma 4.3}}
\end{aligned}
$$

$$\leq sl^2C2^{(2k-2)(l+1)} \leq C'l^2 2^{(2k-2+\frac{5}{8})l}.$$

For the first inequality we used the so called chaining technique: via the triangular inequality we split the term Q_n into two differences $Q_{2^l+i2^d} - Q_{2^l+(i-1)2^d}$.
Now we apply the Chebychev inequality getting for every $\epsilon > 0$

$$\begin{aligned}
&\sum_{l=1}^{\infty} \mathbb{P}\left(\max_{n=2^l,\dots,2^{l+1}-1} \max_{r=0,\dots,s} |Q_n^k(t_{r,l})| > \epsilon 2^{l(k-\frac{1}{2}-\frac{\tau}{8})} \right) \\
&\leq \sum_{l=1}^{\infty} \frac{1}{\epsilon^2 2^{l(2k-1-\frac{\tau}{4})}} \mathbb{E}\left(\max_{n=2^l,\dots,2^{l+1}-1} \max_{r=0,\dots,s} |Q_n^k(t_{r,l})|^2 \right) \\
&\leq \sum_{l=1}^{\infty} \frac{1}{\epsilon^2 2^{l(2k-1-\frac{\tau}{4})}} C'l^2 2^{(2k-2+\frac{5}{8})l} \leq \sum_{l=1}^{\infty} C' \frac{l^2}{\epsilon^2} 2^{\frac{-3+2\tau}{8}l} < \infty.
\end{aligned}$$

Then, with the Borel-Cantelli Lemma

$$\mathbb{P}\left(\max_{n=2^l,\dots,2^{l+1}-1} \max_{r=0,\dots,s} |Q_n^2(t_{r,l})| > \epsilon 2^{l(k-\frac{1}{2}-\frac{\tau}{8})} \text{ infinitely often} \right) = 0.$$

That is, $\max_{r=0,\dots,s} |Q_n^k(t_{r,l})| = o_p(n^{k-\frac{1}{2}-\frac{\tau}{8}})$.
Now we will treat the second term for which we want to apply Lemma 4.2.2 of Wendler (2011a). For $2^l \leq n < 2^{l+1}$ it follows that

$$\begin{aligned}
&\mathbb{E}\left(\sum_{i=1}^{n} (g_1(X_i, t_{r,l}) - g_1(X_i, t_{r-1,l})) \right)^4 \\
&\leq Cn^2(\log n)^2 \max\left\{ \mathbb{E}|g_1(X_i, t_{r,l}) - g_1(X_i, t_{r-1,l})|, Cn^{-\frac{3}{4}} \right\}^{1+\tau} \\
&\leq Cn^2(\log n)^2 (Cn^{-\frac{3}{4}})^{1+\tau}.
\end{aligned}$$

By use of the assumption $|U(t_{r,l}) - U(t_{r-1,l})| \geq 2^{-\frac{5}{8}l} \geq C2^{-\frac{3}{4}l} \geq Cn^{-\frac{3}{4}}$, the last term simplifies to

$$Cn^2(\log n)^2 |U(t_{r,l}) - U(t_{r-1,l})|^{1+\tau}.$$

All in all we get

$$\begin{aligned}
&\mathbb{E}\left(\max_{n=2^l,\dots,2^{l+1}-1} \max_{r=1,\dots,s} \binom{n-1}{k-1} \left| \sum_{i=1}^{n} (g_1(X_i, t_{r,l}) - g_1(X_i, t_{r-1,l})) \right| \right)^4 \\
&\leq \sum_{r=1}^{s} \mathbb{E}\left(\max_{n=2^l,\dots,2^{l+1}-1} n^{4(k-1)} \left| \sum_{i=1}^{n} (g_1(X_i, t_{r,l}) - g_1(X_i, t_{r-1,l})) \right| \right)^4 \\
&\leq \sum_{r=1}^{s} \max_{n=2^l,\dots,2^{l+1}-1} n^{4(k-1)} Cn^2(\log n)^2 |U(t_{r,l}) - U(t_{r-1,l})|^{1+\tau} \\
&\leq 2^{4(k-1)(l+1)} C(2^{l+1})^2 (\log 2^{l+1})^2 s \left(\max_{r=1,\dots,s} |U(t_{r,l}) - U(t_{r-1,l})| \right)^{1+\tau} \\
&\leq C' 2^{4l(k-1)} 2^{2l} (\log 2^{l+1})^2 s \left(2^{-\frac{5}{8}l} \right)^{1+\tau}
\end{aligned}$$

$$\leq C''(l+1)^2 2^{(4k-2-\frac{5}{8}\tau)l}.$$

Thereby we used Corollary 1 of Moricz (1983) and the assumption $s = O(2^{\frac{5}{8}l})$. With the generalized Chebychev inequality it is

$$\begin{aligned}
&\sum_{l=0}^{\infty} \mathbb{P}\left(\max_{n=2^l,\ldots,2^{l+1}-1} \max_{r=1,\ldots,s} \binom{n-1}{k-1} \left| \sum_{i=1}^{n} (g_1(X_i, t_{r,l}) - g_1(X_i, t_{r-1,l})) \right| > \epsilon 2^{(k-\frac{1}{2}-\frac{1}{8}\tau)l} \right) \\
&\leq \sum_{l=0}^{\infty} \frac{1}{\epsilon^4 2^{l(4k-2-\frac{1}{2}\tau)}} \mathbb{E}\left(\max_{r=1,\ldots,s} \binom{n-1}{k-1} \left| \sum_{i=1}^{n} (g_1(X_i, t_{r,l}) - g_1(X_i, t_{r-1,l})) \right| \right)^4 \\
&\leq \sum_{l=0}^{\infty} \frac{1}{\epsilon^4 2^{l(4k-2-\frac{1}{2}\tau)}} C(l+1)^2 2^{(4k-2-\frac{5}{8}\tau)l} = \sum_{l=0}^{\infty} \frac{Cl^2}{\epsilon^4 2^{\frac{1}{8}\tau}} < \infty.
\end{aligned}$$

Analogously to the above calculation we apply the Borel-Cantelli Lemma again getting

$$\binom{n-1}{k-1} \left| \sum_{i=1}^{n} (g_1(X_i, t_{r,l}) - g_1(X_i, t_{r-1,l})) \right| = o_p(n^{k-\frac{1}{2}-\frac{1}{8}\tau}).$$

For the last term, using the assumptions and the fact that $\tau < 1$, we have

$$\max_{r=0,\ldots,s} \binom{n}{k} |U(t_{r,l}) - U(t_{r-1,l})| \leq C n^k 2^{-\frac{5}{8}l} \leq C n^{k-\frac{5}{8}} < C n^{k-\frac{4}{8}-\frac{1}{8}\tau} = o_p(n^{k-\frac{4}{8}-\frac{1}{8}\tau}).$$

Now the terms including $g_2, \ldots, g_{k-1}$ remain. For these we assumed for $2 \leq j \leq k-1$

$$\sup_{t\in\mathbb{R}} \left| \sum_{1\leq i_1,\ldots,i_j\leq n} g_j(X_{i_1}, \ldots, X_{i_j}, t) \right| = o_p(n^{j-\frac{1}{2}-\frac{\tau}{8}})$$

and consequently

$$\begin{aligned}
&\binom{n-j}{k-j} \max_{r=1,\ldots,s} \Big| \sum_{1\leq i_1<\ldots<i_j\leq n} \big(g_j(X_{i_1}, \ldots, X_{i_j}, t_{r,l}) - g_j(X_{i_1}, \ldots, X_{i_j}, t_{r-1,l})\big) \Big| \\
&\leq n^{k-j} \left(\max_{r=1,\ldots,s} \Big| \sum_{1\leq i_1<\ldots<i_j\leq n} g_j(X_{i_1}, \ldots, X_{i_j}, t_{r,l}) \Big| + \max_{r=1,\ldots,s} \Big| \sum_{1\leq i_1<\ldots<i_j\leq n} g_j(X_{i_1}, \ldots, X_{i_j}, t_{r-1,l}) \Big| \right) \\
&\leq n^{k-j} o_p(n^{j-\frac{1}{2}-\frac{1}{8}\tau}) = o_p(n^{k-\frac{1}{2}-\frac{1}{8}\tau}).
\end{aligned}$$

So we could show for arbitrary k for all terms that they are of order $o_p(n^{k-\frac{1}{2}-\frac{1}{8}\tau})$. Using mathematical induction the proof is completed. □

4.3. U-quantiles

U-quantiles ξ_p with a given kernel h are defined as quantiles of the sample $(h(X_i, X_j))_{1\leq i,j\leq n}$, that is $\xi_p = \inf\{t | H_F(t) \geq p\}$. The empirical or sample version of this is then given by

$\hat{\xi}_p = \inf\{t|H_n(t) \geq p\}$. They are useful for expressing certain robust estimators based on quantiles of functions such as the well-known Hodges-Lehmann estimator. Among others, Serfling (1984), Choudhury and Serfling (1988) or Arcones (1996) study U-quantiles and their asymptotic properties. For strong mixing data Wendler (2011a) gives results about asymptotic normality and the Law of the Iterated Logarithm for bivariate kernels. For multivariate kernels and strong mixing data Fischer (2013) respectively Fischer et al. (2016a) study U-quantiles.

An often used technique to handle U-quantiles and their asymptotic is the Bahadur-representation established first by Bahadur (1966) for standard distribution functions. A generalisation to kernels is given by

$$\hat{\xi}_p = \xi_p + \frac{H_F(\xi_p) - H_n(\xi_p)}{h_F(\xi_p)} + R_n.$$

What remains unknown and is the point of interest in many studies is the behaviour of the remaining term R_n. For independent random variables Ghosh (1971) proves the rate of convergence. For strong mixing sequences and bivariate kernels Wendler (2011b) gives the asymptotic behaviour and for short-range dependent linear processes Wu (2005) states results. For strong mixing sequences of multivariate kernels Fischer (2013) respectively Fischer et al. (2016a) prove the rate of convergence.
In the case of near epoch dependence the following lemma is needed for the final proof of the convergence of the remaining term of the Bahadur-representation. It is a variance inequality which gives us the order of boundedness. Whereas for the strong mixing case the following result follows directly from a result in Doukhan et al. (2010) we need to prove it for the NED case.

Lemma 4.5. *Let $(X_n)_{n\in\mathbb{N}}$ be NED with approximation constants $(a_l)_{l\in\mathbb{N}}$ on an absolutely regular process $(Z_n)_{n\in\mathbb{Z}}$ with mixing coefficients $(\beta(l))_{l\in\mathbb{N}}$. Moreover, let be $\sum_{l=1}^{n} a_l < \infty$ and $\sum_{l=1}^{n} \beta(l) < \infty$. If X_1 is bounded and $\mathbb{E}X_i = 0$, then for a constant C it holds that*

$$\mathbb{E}\left(\sum_{i=1}^{n} X_i\right)^2 \leq C \cdot n.$$

Proof.
Because of the stationarity of the X_i we can write

$$\begin{aligned}
\mathbb{E}\left(\sum_{i=1}^{n} X_i\right)^2 &\leq \sum_{\substack{1\leq i,k\leq n \\ i+k\leq n}} |\mathbb{E}\left(X_i X_{i+k}\right)| \qquad (4.4)\\
&\leq \sum_{i=1}^{n}\sum_{k=i+1}^{n-i} \left(4\left\|X_1\right\|_\infty a_{\left[\frac{k}{3}\right]} + 2\left\|X_1\right\|_\infty^2 \beta\left(\left[\frac{k}{3}\right]\right)\right)\\
&\leq n\sum_{k=1}^{n} \left(4\left\|X_1\right\|_\infty a_{\left[\frac{k}{3}\right]} + 2\left\|X_1\right\|_\infty^2 \beta\left(\left[\frac{k}{3}\right]\right)\right)\\
&\leq C \cdot n,
\end{aligned}$$

where we used Lemma 2.18 of Borovkova et al. (2001) in line (4.4). □

Using this we can proof the following result for near epoch dependent processes and multivariate kernels of arbitrary dimension.

Theorem 4.5. *Let $(X_n)_{n\in\mathbb{N}}$ be NED with approximation constants $(a_l)_{l\in\mathbb{N}}$ on an absolutely regular process $(Z_n)_{n\in\mathbb{Z}}$ with mixing coefficients $(\beta(l))_{l\in\mathbb{N}}$ for which holds that $\beta(l) = O\left(l^{-\delta}\right)$ and $a_l = O\left(l^{-\delta-2}\right)$ for a $\delta > 1$. Moreover, let $h(x_1, \ldots, x_m)$ be a Lipschitz-continuous kernel with distribution function H_F and related density $0 < h_F < \infty$ and for all $2 \leq k \leq m$ let $h_{F;X_2,\ldots,X_k}$ be bounded. Then we have for the Bahadur representation with $\hat{\xi}_p = H_n^{-1}(p)$*

$$\hat{\xi}_p = \xi_p + \frac{H_F(\xi_p) - H_n(\xi_p)}{h_F(\xi_p)} + o_p\left(\frac{1}{\sqrt{n}}\right).$$

Proof.
For $t \in \mathbb{R}$, let us define $\xi_{nt} = \xi_p + tn^{-\frac{1}{2}}$, $Z_n(t) = \sqrt{n}\frac{H_F(\xi_{nt}) - H_n(\xi_{nt})}{h_F(\xi_p)}$ and
$V_n(t) = \sqrt{n}\frac{H_F(\xi_{nt}) - H_n(\hat{\xi}_p)}{h_F(\xi_p)}$.
We want to use that $|p - H_n(\hat{\xi}_p)| \leq \frac{1}{n}$ to obtain

$$\begin{aligned} V_n(t) &= \sqrt{n}\frac{H_F(\xi_{nt}) - p + p - H_n(\hat{\xi}_p)}{h_F(\xi_p)} \\ &= \underbrace{\sqrt{n}\frac{H_F(\xi_p + tn^{-\frac{1}{2}}) - p}{h_F(\xi_p)}}_{=:V_n'(t)} + \underbrace{\sqrt{n}\frac{\overbrace{p - H_n(\hat{\xi}_p)}^{=O(n^{-1})}}{h_F(\xi_p)}}_{=O(n^{-\frac{1}{2}})} \longrightarrow t. \end{aligned}$$

The next step is to show that $Z_n(t) - Z_n(0) \xrightarrow{P} 0$.
It is

$$\begin{aligned} &\mathrm{Var}(Z_n(t) - Z_n(0)) \\ &= \frac{n}{h_F^2(\xi_p)}\mathrm{Var}\left(\frac{1}{\binom{n}{m}}\sum_{1\leq i_1<\ldots<i_m\leq n} \mathbb{1}_{\left[h(X_{i_1},\ldots,X_{i_m})\leq \xi_p + tn^{-\frac{1}{2}}\right]} - \mathbb{1}_{\left[h(X_{i_1},\ldots,X_{i_m})\leq \xi_p\right]}\right). \end{aligned}$$

To find bounds for the right hand side, we define U_n and U_n' by

$$\begin{aligned} U_n &= \frac{1}{\binom{n}{m}}\sum_{1\leq i_1<\ldots<i_m\leq n} \mathbb{1}_{\left[h(X_{i_1},\ldots,X_{i_m})\leq \xi_p + tn^{-\frac{1}{2}}\right]} \\ &= \theta + \sum_{j=1}^{m}\binom{m}{j}\frac{1}{\binom{n}{j}}\sum_{1\leq i_1<\ldots<i_j\leq n} g_j(X_{i_1},\ldots,X_{i_k},\xi_p + tn^{-\frac{1}{2}}) \end{aligned}$$

$$\begin{aligned} U_n' &= \frac{1}{\binom{n}{m}}\sum_{1\leq i_1<\ldots<i_m\leq n} \mathbb{1}_{\left[h(X_{i_1},\ldots,X_{i_m})\leq \xi_p\right]} \\ &= \theta' + \sum_{j=1}^{m}\binom{m}{j}\frac{1}{\binom{n}{j}}\sum_{1\leq i_1<\ldots<i_j\leq n} g_j(X_{i_1},\ldots,X_{i_k},\xi_p). \end{aligned}$$

Therefore,

$$\begin{aligned}
&\sqrt{\operatorname{Var}\left(\frac{1}{\binom{n}{m}}\sum_{1\le i_1<\ldots<i_m\le n}\mathbb{1}_{\left[\xi_p<h(X_{i_1},\ldots,X_{i_m})\le\xi_p+tn^{-\frac{1}{2}}\right]}\right)}\\
&\le\sqrt{\underbrace{\operatorname{Var}(\theta)}_{=0}}+\sqrt{\underbrace{\operatorname{Var}(\theta')}_{=0}}+\sqrt{\operatorname{Var}\left(\frac{m}{n}\sum_{i=1}^{n}(g_1(X_i,\xi_p+tn^{-\frac{1}{2}})-g_1(X_i,\xi_p))\right)}\\
&+\sqrt{\operatorname{Var}\left(\frac{\binom{m}{2}}{\binom{n}{2}}\sum_{1\le i<j\le n}g_2(X_i,X_j,\xi_p+tn^{-\frac{1}{2}})\right)}+\sqrt{\operatorname{Var}\left(\frac{\binom{m}{2}}{\binom{n}{2}}\sum_{1\le i<j\le n}g_2(X_i,X_j,\xi_p)\right)}\\
&+\ldots+\sqrt{\operatorname{Var}\left(\frac{1}{\binom{n}{m}}\sum_{1\le i_1<\ldots<i_m\le n}g_m(X_{i_1},\ldots,X_{i_m},\xi_p+tn^{-\frac{1}{2}})\right)}\\
&+\sqrt{\operatorname{Var}\left(\frac{1}{\binom{n}{m}}\sum_{1\le i_1<\ldots<i_m\le n}g_m(X_{i_1},\ldots,X_{i_m},\xi_p)\right)}.
\end{aligned}$$

We have already shown in Theorem 4.1, for all $2\le k\le m$, that

$$\operatorname{Var}\left(\frac{\binom{m}{k}}{\binom{n}{k}}\sum_{1\le i_1<\ldots<i_k\le n}g_k(X_{i_1},\ldots,X_{i_k},\xi_p+tn^{-\frac{1}{2}})\right)=O(n^{-2+\gamma})$$

holds for a $\gamma<1$, if the kernel is bounded and satisfies the extended variation condition. This can be shown for the U-process kernels, too. Analogous to the proof of Corollary 4.1 we can use that $g(x_1,\ldots,x_m,\xi_p+tn^{-\frac{1}{2}})=\mathbb{1}_{\left[h(x_{i_1},\ldots,x_{i_m})\le\xi_p+tn^{-\frac{1}{2}}\right]}$ and $g(x_1,\ldots,x_m,\xi_p)=\mathbb{1}_{[h(x_{i_1},\ldots,x_{i_m})\le\xi_p]}$ satisfy the extended variation condition.
To replace the Lemma of Doukhan et al. (2010) in the strong mixing case we apply for the NED case Lemma 4.5 on $g_1(X_i,\xi_p+tn^{-\frac{1}{2}})-g_1(X_i,\xi_p)$ and have

$$\mathbb{E}\Big|\sum_{i=1}^{n}(g_1(X_i,\xi_p+tn^{-\frac{1}{2}})-g_1(X_i,\xi_p))\Big|^2\le Cn.$$

This is possible since g_1 and $f(x,y)=x-y$ fulfil the variation condition and are bounded (f when using bounded arguments). Additionally, $\sum_{l=1}^{\infty}\beta(l)<\infty$ and $\sum_{l=1}^{\infty}a_l'<\infty$ because of the assumptions.
This helps us to obtain

$$\begin{aligned}
&\sqrt{\operatorname{Var}\left(\frac{1}{\binom{n}{m}}\sum_{1\le i_1<\ldots<i_m\le n}1_{\left[\xi_p<h(X_{i_1},\ldots,X_{i_m})\le\xi_p+tn^{-\frac{1}{2}}\right]}\right)}\\
&\le\sqrt{\frac{m^2}{n^2}Cn}+2(m-1)\sqrt{O(n^{-2+\gamma})}\le\frac{Cm^2}{\sqrt{n}}+2(m-1)O(n^{-1+\gamma/2}),
\end{aligned}$$

where the constant C only depends on $\left\| g_1(X_i, \xi_p + tn^{-\frac{1}{2}}) - g_1(X_i, \xi_p) \right\|_3$.
Then,

$$\begin{aligned}
\mathrm{Var}(Z_n(t) - Z_n(0)) &\leq \frac{n}{h_F^2(\xi_p)} \left(\frac{Cm^2}{\sqrt{n}} + 2(m-1)O(n^{-1+\gamma/2}) \right)^2 \\
&\leq \frac{m^2}{h_F^2(\xi_p)} C^2 + \frac{4m^2(m-1)}{h_F^2(\xi_p)} C\sqrt{n} O(n^{-1+\gamma/2}) + 4(m-1)^2 O(n^{-2+\gamma}) \frac{n}{h_F^2(\xi_p)} \\
&\leq \frac{m^4}{h_F^2(\xi_p)} C^2 + \frac{4m^2(m-1)}{h_F^2(\xi_p)} C O(n^{-\frac{1}{2}+\gamma/2}) + \frac{4(m-1)^2}{h_F^2(\xi_p)} O(n^{-1+\gamma}).
\end{aligned}$$

Since $|g_1(X_i, \xi_p + tn^{-\frac{1}{2}}) - g_1(X_i, \xi_p)| \leq 1$ for all X_i and

$$|g_1(X_i) - g_1'(X_i)| \xrightarrow{P} 0,$$

the constant C converges to zero in probability and since $\gamma < 1$ we have that

$$\mathrm{Var}(Z_n(t) - Z_n(0)) \xrightarrow{P} 0.$$

Applying the Chebychev inequality we then get $Z_n(t) - Z_n(0) \xrightarrow{P} 0$.
Altogether we have for $t \in \mathbb{R}$ and every $\epsilon > 0$

$$\begin{aligned}
\mathbb{P}(\sqrt{n}(\hat{\xi}_p - \xi_p) \leq t, Z_n(0) \geq t + \epsilon) &= \mathbb{P}(Z_n(t) \leq V_n(t), Z_n(0) \geq t + \epsilon) \\
&\leq \mathbb{P}\left(|Z_n(t) - Z_n(0)| \geq \frac{\epsilon}{2} \right) + \mathbb{P}\left(|V_n(t) - t| \geq \frac{\epsilon}{2} \right) \longrightarrow 0
\end{aligned}$$

and analogously

$$\mathbb{P}(\sqrt{n}(\hat{\xi}_p - \xi_p) \geq t, Z_n(0) \leq t) \longrightarrow 0.$$

Using Lemma 1 of Ghosh (1971) the proof is completed. □

5. GL-statistics

Generalized L-statistics (GL-statistics) form a class of statistics that unifies a lot of commonly used other statistics, for example U-statistics, L-statistics and even statistics that cannot be classified in a certain class of statistics, for example the Hodges-Lehmann estimator. They have first been proposed by Serfling (1984), who proved a Central Limit Theorem under independence. Additionally, Fischer et al. (2016a) have shown an analogous result under strong mixing. In our context they are useful since a lot of robust estimator can be expressed as GL-statistic, detailed examples are given later on.
Like U-statistics every GL-statistic includes a kernel $h(x_1, \dots, x_m)$, that is a measurable, symmetric and real-valued function.
Using the same notation as before we now can define a GL-statistic as a functional of the empirical distribution function of the kernel.

Definition 5.1. *A generalized L-statistic with kernel h is given by*

$$\begin{aligned} T(H_n) &= \int_0^1 H_n^{-1}(t)J(t)dt + \sum_{i=1}^{d} a_i H_n^{-1}(p_i) \\ &= \sum_{i=1}^{n_{(m)}} \left[\int_{\frac{(i-1)}{n_{(m)}}}^{\frac{i}{n_{(m)}}} J(t)dt \right] H_n^{-1}\left(\frac{i}{n_{(m)}}\right) + \sum_{i=1}^{d} a_i H_n^{-1}(p_i), \end{aligned}$$

with a function $J(\cdot)$ and weights p_i and a_i,$i = 1, \dots, d$, where $n_{(m)} = n(n-1)\cdot \dots \cdot (n-m+1)$ and

$$H_n(x) = \frac{1}{\binom{n}{m}} \sum_{1 \le i_1 < \dots < i_m \le n} \mathbb{1}_{[h(X_{i_1}, \dots, X_{i_m}) \le x]}$$

as before. It is an estimator for $T(H_F)$.

A GL-statistics is determined by the choice of the function $J(\cdot)$ and the weights p_i and a_i with $i = 1, \dots, d$. J represents a smooth or continuous weighting in the first part in the definition above, whereas the coefficients a_i determine a discrete weighting. According to which kind of statistic or estimator should be expressed as a GL-statistic, a continuous or a discrete or even both weightings are used.
Compared to the representation of an L-statistic

$$T(F_n) = \int_0^1 F_n^{-1}(t)J(t)dt + \sum_{i=1}^{d} a_i F_n^{-1}(p_i) \tag{5.1}$$

we can see that the generalisation results from the changing of the empirical distribution function F_n to the empirical kernel distribution function H_n. This is due to the fact that we consider now $h(X_1, \dots, X_m)$ instead of $X_1, \dots, X_m$.

Remark 5.1. The general definition of an L-statistic is given by

$$T = \sum_{i=1}^{n} c_{i,n} F_n^{-1}(i/n)$$

with $c_{i,n} \in \mathbb{R}$ but among many others Serfling (1984) pointed out that nearly all relevant L-statistics can be expressed by eq. (5.1).

Let us now give some examples of well known estimators that can be expressed as GL-statistic.

5.1. Examples

In this section we want to show that some well known estimators can be expressed as GL-statistic. Here, the focus is laid on robust estimators, mainly with multivariate kernels.

Example 5.1. A U-statistic can be written as a GL-statistic by setting $d = 0$ and $J = 1$. Of course, not every U-statistic is robust but many robust estimators are included in this class.

Example 5.2. A well known L-statistic is the *α-trimmed mean*

$$\bar{X}_{(\alpha)} = \frac{1}{n - 2\,[n\alpha]} \sum_{i=[n\alpha]+1}^{n-[n\alpha]} X_{(i)},$$

where $X_{(i)}$ is the i-th value of the order statistic $X_{(1)} \leq X_{(2)} \leq \ldots \leq X_{(n)}$. To rewrite it as a GL-statistic we choose $J(t) = \frac{1}{1-2\alpha} \mathbb{1}_{[\alpha<t<1-\alpha]}$. As kernel we set $h(x) = x$ and let the sum vanish by the choice $d = 0$.

Example 5.3. The generalized *Hodges-Lehmann estimator*

$$\text{median}\left(\frac{1}{m} \left(X_{i_1} + \ldots + X_{i_m} \right),\ 1 \leq i_1, \ldots, i_m \leq n \right)$$

is neither a U-statistic nor an L-statistic, but it is possible to formulate it as a GL-statistic choosing the kernel $h(x_{i_1}, \ldots, x_{i_m}) = \frac{1}{m}(x_{i_1} + \ldots + x_{i_m})$ and setting $J = 0$, $d = 1$, and $a_1 = 1$. We get the median of the kernel by using the representation via the quantile function $H_n^{-1}(\frac{1}{2})$. Consequently, $p_1 = \frac{1}{2}$. The generalized Hodges-Lehmann estimator then is the GL-statistic

$$T(H_n) = H_n^{-1}\left(\frac{1}{2} \right).$$

Example 5.4. A robust measure for the variability is Gini's Mean Difference. In contrast to the mean deviation this estimator is not only robust but also almost as efficient as the classical measure for variability, the standard deviation (Gerstenberger and Vogel (2015)). It is given by

$$G_n = \frac{2}{n(n-1)} \sum_{1 \leq i < j \leq n} |X_i - X_j|$$

and can be written as a GL-statistic by choosing the kernel $h(x_i, x_j) = |x_i - x_j|$ (the kernel dimension is then of course $m = 2$) and a constant continuous function $J(t) = 1$. The discrete part of the GL-statistic vanishes since $d = 0$. When considering the expression Gini's Mean Difference using order statistics (see cf. Nair (1936)), that is

$$G_n = \frac{2}{n(n-1)} \sum_{1 \le i < j \le n} |X_i - X_j| = \frac{2}{n(n-1)} \sum_{1 \le i < j \le n} \left(X_{(j)} - X_{(i)}\right)$$
$$= \frac{2}{n(n-1)} \left(2 \sum_{i=1}^{n} i X_{(i)} - (n+1) \sum_{i=1}^{n} X_{(i)} \right) = \frac{2}{n(n-1)} \sum_{i=1}^{n} (2i - n - 1) X_{(i)}$$

with $X_{(i)}$ being the i-th order statistic of the sample $X_1, \ldots, X_n$, the kernel is chosen as identity function and $J(t) = \frac{4n}{n-1} t - \frac{2n}{n-1}$.

The following two examples of scale estimators can be found in Rousseeuw and Croux (1992) and can all be expressed as a GL-statistic.

Example 5.5. An estimator for the scale which is robust with an asymptotic breakdown point of 50% is given by

$$Q = \text{med}_{i<j<k} \{\min(|X_i - X_j|, |X_i - X_k|, |X_j - X_k|)\}.$$

It can be expressed as a GL-statistic by taking the three-dimensional kernel $h(x_1, x_2, x_3) = \min(|x_1 - x_2|, |x_1 - x_3|, |x_2 - x_3|)$ and the parameters $J = 0, d = 1, a_1 = 1$ and $p_1 = 1/2$. If this estimator is generalised to subsamples of order greater than three and other quantiles

$$Q_n = \{\min(|X_{i_l} - X_{i_k}|, 1 \le l < k \le m), 1 \le i_1 < \ldots < i_m \le n\}_{([\alpha \binom{n}{m}])},$$

where $([\alpha \binom{n}{m}])$ denotes the empirical α-quantile ($\alpha \in (0, 1)$), we can choose for arbitrary size m of the subsample $h(x_1, \ldots, x_m) = \min(|x_j - x_i|, 1 \le i < j \le m)$ with J, d and a_1 as before but $p_1 = \frac{[\alpha \binom{n}{m}]}{\binom{n}{m}}$.

Example 5.6. Another location-free scale estimator is given by

$$C_n^\alpha = c_\alpha |X_{(i+[\alpha n]+1)} - X_{(i)}|_{([n/2]-[\alpha n])},$$

$\alpha \in (0, 0.5)$, which takes the $[n/2] - [\alpha n]$ order statistic of the difference of the first and last order statistic of all (sorted) subsamples of length $[\alpha n] + 2$. The constant c_α makes the estimator Fisher-consistent under normality. By the choice of a kernel h of dimension $m = [\alpha n] + 2$ with $h(x_1, \ldots, x_m) = \max(x_1, \ldots, x_m) - \min(x_1, \ldots, x_m)$ and $J = 0, d = 1, a_1 = c_\alpha$ and $p_1 = \frac{1}{\binom{n}{m}}$ the representation by a GL-statistic can be obtained. A well-known special case of this estimator is the Least Median of Squares scale estimator

$$LMS_n = 0.7413 \min_i |X_{(i+[n/2])} - X_{(i)}|,$$

with $[\alpha n] = [n/2] - 1$ and $c_\alpha = \frac{1}{2\Phi^{-1}(0.75)} = 0.7413$.

We want to emphasize that both estimators of Rousseeuw and Croux (1992) use multivariate kernels, that are kernels with dimension larger than 2. In these cases results for bivariate GL-statistics would not be sufficient.

5.2. A General Central Limit Theorem

A main interest when considering GL-statistics or in general classes of statistics is their asymptotic behaviour. We have to know the limit distribution to establish confidence intervals or, when considering test statistics, the case of rejection of the hypothesis.
Serfling (1984) already has shown one of the key tools for proving a Central Limit Theorem for GL-statistics. The error term $T(H_n) - T(H_F)$ normalized with $\sqrt{n}$ can be approximated by the so called first-order Gâteaux-differential (cf. Serfling (1980))

$$d_1T(H_F; H_n - H_F) = -\int_{-\infty}^{\infty} (H_n(y) - H_F(y))\, J(H_F(y))dy + \sum_{i=1}^{d} a_i \frac{p_i - H_n(H_F^{-1}(p_i))}{h_F(H_F^{-1}(p_i))}.$$

In general the Gâteaux-differential of order k is defined as

$$d_kT(F; G - F) = \frac{d^k}{d\lambda^k} T(F + \lambda(G - F))\Big|_{\lambda=0}$$

and is a generalisation of the classical differential to infinite dimensional spaces.

Remark 5.2. Closely related to the Gâteaux-differential is the *influence curve*. Remember that the influence curve is defined as

$$IC(x; F, T) = \frac{d}{d\lambda} T(F + \lambda(\delta_x - F))\Big|_{\lambda=0},$$

where δ_x is the distribution function corresponding to a unit point mass at x (see Section 2). That is, we simply replace G by δ_x in the definition of the Gâteaux-differential. For functionals of the form

$$T(F) = \int_0^1 F^{-1}(t)J(t)dt$$

Serfling (1984) has shown that the influence curve simplifies to

$$IC(x; F, T) = -\int_{-\infty}^{\infty} \left(\mathbb{1}_{[x \le y]} - F(y)\right) J(F(y))dy.$$

We will see later on that also the asymptotic variance of GL-statistics can be expressed in terms of the influence curve.

To show that the approximation above is valid, Serfling (1984) used two main results. These are important, as they make the proof of the Central Limit Theorem possible. Therefore, we want to show in short steps how the approximation of Serfling (1984) works.
First, split the term $T(H_F)$ into its continuous and its discrete part,

$$T(H_F) = \int_0^1 H_F^{-1}(t)J(t)dt + \sum_{i=1}^{d} a_i H_F^{-1}(p_i) = T_1(H_F) + T_2(H_F)$$

and analogously for $T(H_n) = T_1(H_n) + T_2(H_n)$. Moreover, we define $\Delta_{i,n} = T_i(H_n) - T_i(H_F) - d_1T_i(H_F; H_n - H_F)$ with $i = 1, 2$ and $K(t) = \int_0^t J(y)dy$.

Then

$$\begin{aligned}\Delta_{1,n} &= T_1(H_n) - T_1(H_F) - d_1 T_1(H_F; H_n - H_F) \\ &= -\int_{-\infty}^{\infty} \underbrace{\left(\frac{\int_0^{H_n(y)} J(t)dt - \int_0^{H_F(y)} J(t)dt}{H_n(y) - H_F(y)} - J(H_F(y)) \right)}_{=:W_{H_n,H_F}(y)} (H_n(y) - H_F(y))\, dy \\ &= \begin{cases} -\int_{-\infty}^{\infty} W_{H_n,H_F}(y)\,(H_n(y) - H_F(y))\, dy & , \text{ falls } H_n \neq H_F \\ 0 & , \text{ falls } H_n = H_F, \end{cases}\end{aligned}$$

using Lemma 12 of Boos (1979). Now there are two possible ways of assessing $\Delta_{1,n}$, leading to two different assumptions in the Central Limit Theorem. Due to completeness we want to show both possibilities though only the first one is used in the proof.
Proving either

$$|\Delta_{1,n}| \leq \|W_{H_n,H_F}\|_{L_1} \cdot \|H_n - H_F\|_\infty$$

or

$$|\Delta_{1,n}| \leq \|W_{H_n,H_F}\|_\infty \cdot \|H_n - H_F\|_{L_1}.$$

and

(i) for $W_{H_n,H_F}(y) = \left(\frac{\int_0^{H_n(y)} J(t)dt - \int_0^{H_F(y)} J(t)dt}{H_n(y) - H_F(y)} - J(H_F(y)) \right)$ it holds that

$\|W_{H_n,H_F}\|_{L_1} = o_p(1)$ and additionally $\|H_n - H_F\|_\infty = O_p(n^{-\frac{1}{2}})$

or

(i)' $\|W_{H_n,H_F}\|_\infty = o_p(1)$ and $\|H_n - H_F\|_{L_1} = O_p(n^{-\frac{1}{2}})$,

respectively, we obtain

$$\sqrt{n}\Delta_{1,n} \xrightarrow{P} 0.$$

For the second term $\Delta_{2,n}$ we can use the assessment

$$\begin{aligned}\Delta_{2,n} &= T_2(H_n) - T_2(H_F) - d_1 T_2(H_F; H_n - H_F) \\ &= \sum_{i=1}^{d} a_i \left(\hat{\xi}_{p_i,n} - \xi_{p_i} - \frac{p_i - H_n(\xi_{p_i})}{h_F(\xi_{p_i})} \right),\end{aligned}$$

where ξ_{p_i} is the generalized p_i-quantile and $\hat{\xi}_{p_i,n}$ is the empirical p_i-quantile.
Serfling (1984) now uses the remaining term $R_{p_i,n}$ of the Bahadur-representation of a quantile $\hat{\xi}_{p_i,n}$ (Bahadur (1966))

$$\hat{\xi}_{p_i,n} = \xi_{p_i} + \frac{p_i - H_n(\xi_{p_i})}{h_f(\xi_{p_i})} + R_{p_i,n},$$

if $F'(\xi_{p_i}) = f(\xi_{p_i}) > 0$, with $R_{p_i,n} = \hat{\xi}_{p_i,n} - \xi_{p_i} + \frac{p_i - H_n(\xi_{p_i})}{h_f(\xi_{p_i})}$ to show the convergence of $\Delta_{2,n}$. If we can assume that

$$(ii)\ \sqrt{n}R_{p_i,n} \xrightarrow{P} 0 \quad \text{for all } i = 1, \ldots, d$$

and using the continuous mapping theorem we get

$$\sqrt{n}\Delta_{2,n} \xrightarrow{P} 0.$$

Therefore,

$$\sqrt{n}(T(H_n) - T(H_F) - d_1T_1(H_F; H_n - H_F)) \xrightarrow{P} 0,$$

if (ii) and (i) or (i)' are valid.

Now the reason for this approximation has been that the Gâteaux-differential $d_1T(H_F; H_n - H_F)$ has some desirable properties. It can be written as U-statistic with kernel

$$\begin{aligned} A(x_1, \ldots, x_m) = &- \int_{-\infty}^{\infty} \left(\mathbb{1}_{[h(x_1,\ldots,x_m)\leq y]} - H_F(y)\right) J(H_F(y))dy \\ &+ \sum_{i=1}^{d} a_i \frac{p_i - \mathbb{1}_{[h(x_1,\ldots,x_m)\leq H_F^{-1}(p_i)]}}{h_F(H_F^{-1}(p_i))}. \end{aligned}$$

Hence, if we can establish a Central Limit Theorem for U-statistics, we automatically have the convergence of $d_1T(H_F; H_n - H_F)$.

All these considerations made for $(T(H_n) - T(H_F))$ can be summarized in the Central Limit Theorem. We want to emphasize that up to now no assumptions on the dependence structure have been made. The approximations and convergences above can be proved without such assumptions and only in the proof of (i), (ii) or the Central Limit Theorem for U-statistics dependency will play a crucial role. This makes the following formulation of a general Central Limit Theorem possible.

Theorem 5.1. *Let $X_1, \ldots, X_n$ be random variables with distribution function F and let $h(x_1, \ldots, x_m)$ be a kernel with distribution function H_F and related density $h_F > 0$. Moreover, let H_n be the empirical distribution function of h and $T(H_n)$ a GL-Statistik with kernel h, bounded function J and weights p_i, a_i, $i = 1, \ldots, d$.*
If the following assumptions (i) or (i)', (ii) and (iii) are valid with

(i) For $W_{H_n,H_F}(y) = \left(\frac{\int_0^{H_n(y)} J(t)dt - \int_0^{H_F(y)} J(t)dt}{H_n(y) - H_F(y)} - J(H_F(y))\right)$ it holds that

$\|W_{H_n,H_F}\|_{L_1} = o_p(1)$ *and* $\|H_n - H_F\|_\infty = O_p(n^{-\frac{1}{2}})$.

or

(i)' It is $\|W_{H_n,H_F}\|_\infty = o_p(1)$ and $\|H_n - H_F\|_{L_1} = O_p(n^{-\frac{1}{2}})$.

(ii) For the remaining term of the generalized Bahadur-representation of an empirical quantile $R_{p_i,n} = \hat{\xi}_{p_i,n} - \xi_{p_i} + \frac{p_i - H_n(\xi_{p_i})}{h_f(\xi_{p_i})}$ it holds that

$$R_{p_i,n} = o_p(n^{-\frac{1}{2}}).$$

(iii) For a U-statistic with kernel

$$A(x_1, \ldots, x_m) = - \int_{-\infty}^{\infty} \left(\mathbb{1}_{[h(x_1,\ldots,x_m)\leq y]} - H_F(y)\right) J(H_F(y))dy$$

$$+\sum_{i=1}^{d} a_i \frac{p_i - \mathbb{1}_{\left[h(x_1,\ldots,x_m)\leq H_F^{-1}(p_i)\right]}}{h_F(H_F^{-1}(p_i))}$$

it holds that

$$\sqrt{n}(U_n(A) - \theta) \xrightarrow{D} N(0, \sigma^2),$$

where θ is as in Definition 4.2, then we have that

$$\sqrt{n}\left(T(H_n) - T(H_F)\right) \xrightarrow{D} N(0, \sigma^2).$$

Remark 5.3. Please notice that we do not specify the structure of the asymptotic variance σ. It depends much on the dependence structure of the random variables and is determined by the Central Limit Theorem of U-statistics. A specific notation can be found in the Central Limit Theorems in Section 5.3.

Now we concrete our assumptions on the dependence structure of the random variables.

5.3. Limit Theorems for GL-Statistics under Dependence

5.3.1. Properties of the Kernel A

The following lemma states three of the most fundamental properties of the kernel A in Theorem 5.1, which determines the asymptotic behaviour of a GL-statistic. The first two properties have already been shown by Fischer (2013) respectively Fischer et al. (2016a) and are only mentioned due to completeness.

Lemma 5.1. *Let $X_1, \ldots, X_n$ be a sequence of random variables with distribution function F and let $h(x_1, \ldots, x_m)$ be a Lipschitz-continuous kernel with distribution function H_F and related density $0 < h_F < \infty$ and for all $2 \leq k \leq m$ and all $i_1 < i_2 < \ldots < i_m$ let $h_{F;X_{i_2},\ldots,X_{i_k}}$ be bounded. Moreover, let J be a function with $J(t) = 0$ for $t \notin [\alpha, \beta]$, $0 < \alpha < \beta < 1$, and in $[\alpha, \beta]$ let J be bounded and a.e. continuous concerning the Lebesgue-measure and a.e. continuous concerning H_F^{-1}. Then the kernel A related to the GL-statistic $T(H_n)$ with*

$$A(x_1, \ldots, x_m) = -\int_{-\infty}^{\infty} \left(\mathbb{1}_{[h(x_1,\ldots,x_m)\leq y]} - H_F(y)\right) J(H_F(y))dy + \sum_{i=1}^{d} a_i \frac{p_i - \mathbb{1}_{[h(x_1,\ldots,x_m)\leq H_F^{-1}(p_i)]}}{h_F(H_F^{-1}(p_i))}$$

has the following three properties

(i) A is bounded

(ii) A fulfils the extended variation condition

(iii) A fulfils the L_2-variation condition.

Proof.
As mentioned above the first two properties have been shown by Fischer et al. (2016a) and are therefore omitted.

For property (iii) we use that

$$
\begin{aligned}
&\sqrt{\mathbb{E}\left(\sup_{\|(x_1,\dots,x_m)-(X_1',\dots,X_m')\|\leq\epsilon}|A(x_1,\dots,x_m)-A(X_1',\dots,X_m')|\right)^2}\\
&=\Bigg(\mathbb{E}\Bigg(\sup_{\|(x_1,\dots,x_m)-(X_1',\dots,X_m')|\leq\epsilon}\Bigg|-\int_{-\infty}^{\infty}(\mathbb{1}_{[h(x_1,\dots,x_m)\leq y]}-H_F(y))J(H_F(y))dy\\
&\qquad\qquad+\int_{-\infty}^{\infty}(\mathbb{1}_{[h(X_1',\dots,X_m')\leq y]}-H_F(y))J(H_F(y))dy\\
&\qquad\qquad+\sum_{i=1}^{d}a_i\frac{p_i-\mathbb{1}_{[h(x_1,\dots,x_m)\leq H_F^{-1}(p_i)]}}{h_F(H_F^{-1}(p_i))}-\sum_{i=1}^{d}a_i\frac{p_i-\mathbb{1}_{[h(X_1',\dots,X_m')\leq H_F^{-1}(p_i)]}}{h_F(H_F^{-1}(p_i))}\Bigg|\Bigg)^2\Bigg)^{\frac{1}{2}}\\
&\leq\sqrt{\mathbb{E}\left(\sup_{\|(x_1,\dots,x_m)-(X_1',\dots,X_m')\|\leq\epsilon}\left|\int_{-\infty}^{\infty}(\mathbb{1}_{[h(x_1,\dots,x_m)\leq y]}-\mathbb{1}_{[h(X_1',\dots,X_m')\leq y]})J(H_F(y))dy\right|\right)^2}\\
&+\sqrt{\mathbb{E}\left(\sup_{\|(x_1,\dots,x_m)-(X_1',\dots,X_m')\|\leq\epsilon}\left|\sum_{i=1}^{d}a_i\frac{\mathbb{1}_{[h(x_1,\dots,x_m)\leq H_F^{-1}(p_i)]}-\mathbb{1}_{[h(X_1',\dots,X_m')\leq H_F^{-1}(p_i)]}}{h_F(H_F^{-1}(p_i))}\right|\right)^2}.
\end{aligned}
$$

These terms can now be treated separately and analogous to the proof of the extended variation condition in Fischer et al. (2016a). For the first term we obtain

$$
\begin{aligned}
&\mathbb{E}\left(\sup_{\|(x_1,\dots,x_m)-(X_1',\dots,X_m')\|\leq\epsilon}\left|\int_{-\infty}^{\infty}(\mathbb{1}_{[h(x_1,\dots,x_m)\leq y]}-\mathbb{1}_{[h(X_1',\dots,X_m')\leq y]})J(H_F(y))dy\right|\right)^2\\
&\leq\mathbb{E}\left(\sup_{t\in\mathbb{R}}\left|\mathbb{1}_{[h(X_1',\dots,X_m')\in(t-\tilde{L}\epsilon,t+\tilde{L}\epsilon)]}\right|\left|\int_{-\infty}^{\infty}J(H_F(y))dy\right|\right)^2\\
&\leq\mathbb{E}\left(\sup_{t\in\mathbb{R}}\left|\mathbb{1}_{[h(X_1',\dots,X_m')\in(t-\tilde{L}\epsilon,t+\tilde{L}\epsilon)]}\right|C\right)^2\\
&\leq C\sup_{t\in\mathbb{R}}\left|\mathbb{E}(\mathbb{1}_{[h(X_1',\dots,X_m')\in(t-\tilde{L}\epsilon,t+\tilde{L}\epsilon)]})^2\right|\\
&\leq C\sup_{t\in\mathbb{R}}|\mathbb{P}(h(X_1',\dots,X_m')\in(t-\tilde{L}\epsilon,t+\tilde{L}\epsilon))|\\
&\leq C\left(\sup_{x\in\mathbb{R}}h_F(x)\right)2\tilde{L}\epsilon\leq L\epsilon
\end{aligned}
$$

using the boundedness of $|\int_{-\infty}^{\infty}J(H_F(y))dy|$ and h_F. The second term is treated analogously and omitted here.
Therefore,

$$
\sqrt{\mathbb{E}\left(\sup_{\|(x_1,\dots,x_m)-(X_1',\dots,X_m')\|\leq\epsilon}|A(x_1,\dots,x_m)-A(X_1',\dots,X_m')|\right)^2}\leq 2\sqrt{L\epsilon}
$$

and

$$
\mathbb{E}\left(\sup_{\|(x_1,\dots,x_m)-(X_1',\dots,X_m')\|\leq\epsilon}|A(x_1,\dots,x_m)-A(X_1',\dots,X_m')|\right)^2\leq L'\epsilon
$$

□

Now we are able to formulate the Central Limit Theorem.

5.3.2. Central Limit Theorem

Let us now specify the result of Theorem 5.1 for near epoch dependent random variables on absolutely regular processes. Under independence this result has been proved by Serfling (1984), some of the lemmata can also be found in Choudhury and Serfling (1988). Under strong mixing an analogous result can be found in Fischer et al. (2016a). We want to emphasize that the result is explicitly stated for multivariate kernels. Results for bivariate GL-statistics under strong mixing and near epoch dependence can also be found in Wendler (2012).

Theorem 5.2. *Let $h(x_1, \dots, x_m)$ be a Lipschitz-continuous kernel with distribution function H_F and related density $0 < h_F < \infty$ and for all $2 \leq k \leq m$ and all $i_1 < i_2 < \ldots < i_m$ let $h_{F;X_{i_2},\dots,X_{i_k}}$ be bounded. Moreover, let J be a function with $J(t) = 0$ for $t \notin [\alpha, \beta]$, $0 < \alpha < \beta < 1$, and in $[\alpha, \beta]$ let J be bounded and a.e. continuous concerning the Lebesgue-measure and a.e. continuous concerning H_F^{-1}. Additionally, let $X_1, \dots, X_n$ be NED with approximation constants $(a_l)_{l\in\mathbb{N}}$ on an absolutely regular process $(Z_n)_{n\in\mathbb{Z}}$ with mixing coefficients $(\beta(l))_{l\in\mathbb{N}}$ with $\sum_{l=1}^{\infty} l^2 \beta(l)^{\frac{\gamma}{2+\gamma}} < \infty$ for a $0 < \gamma < 1$. Moreover, let be $\sum_{l=1}^{\infty} l^2 a_l^{\frac{\gamma}{2+2\gamma}} < \infty$. Then for GL-statistics $T(H_n)$ it holds that*

$$\sqrt{n}\left(T(H_n) - T(H_F)\right) \xrightarrow{\mathcal{D}} N(0, \sigma_{GL}^2),$$

where

$$\begin{aligned}\sigma_{GL}^2 =& m^2\big(Var\left(\mathbb{E}\left(A(Y_1, \dots, Y_m)|Y_1 = X_1\right)\right) \\ &+ 2\sum_{j=1}^{\infty} Cov\left(\mathbb{E}\left(A(Y_1, \dots, Y_m)|Y_1 = X_1\right), \mathbb{E}\left(A(Y_1, \dots, Y_m)|Y_1 = X_{j+1}\right)\right)\big)\end{aligned}$$

with independent copies $Y_1, \dots, Y_m$ of X_1 and

$$\begin{aligned}A(x_1, \dots, x_m) =& -\int_{-\infty}^{\infty} \left(\mathbb{1}_{[h(x_1,\dots,x_m)\leq y]} - H_F(y)\right) J(H_F(y))dy \\ &+ \sum_{i=1}^{d} a_i \frac{p_i - \mathbb{1}_{\left[h(x_1,\dots,x_m)\leq H_F^{-1}(p_i)\right]}}{h_F(H_F^{-1}(p_i))}.\end{aligned}$$

Remark 5.4. A key assumption to prove this theorem is the boundedness of J on $[\alpha, \beta]$ and that the function vanishes outside of the interval. This makes it possible to prove assumption (i) given in Section 5.2. In this work we will limit the proofs to this case, since it is the relevant one in robust statistics, as we will prove in Section 5.4. Nevertheless, proving assumption (i)' instead of (i) and using the respective Central Limit Theorem for U-statistics with unbounded kernel (Theorem 4.2) makes it possible to extend the valid functions of J to unlimited functions continuous on $[0, 1]$ (see Serfling (1984)).

As in the case of U-statistics, to use this theorem for example for confidence intervals in applications like hydrology, the problem arises how to handle the asymptotic variance σ_{GL}^2. Normally, it is unknown due to the unknown conditional expected values and the unknown distribution. Therefore, it is necessary to find an estimator for σ_{GL}^2. The following corollary closes this gap. It is stated for NED but it is also possible to show an analogous result under strong mixing such that it is applicable for the theorem of Fischer et al. (2016a).

Corollary 5.1. *Let $h : \mathbb{R}^m \to \mathbb{R}$ be a Lipschitz-continuous kernel.*
Moreover, let $(X_n)_{n\in\mathbb{N}}$ be NED with approximation constants $(a_l)_{l\in\mathbb{N}}$ on an absolutely regular process $(Z_n)_{n\in\mathbb{Z}}$ with mixing coefficients $(\beta(l))_{l\in\mathbb{N}}$ and let a $\delta > 11$ exist, such that $\sum_{l=1}^{\infty} l\beta^{2/(2+\delta)(l)} < \infty$ and $a_l = O\left(l^{-\delta-3}\right)$. The weight function κ and the bandwidth b_n should fulfil Assumption 4.1. Then it holds for the long-run variance estimator $\hat{\sigma}^2_{GL}$ that

$$\hat{\sigma}^2_{GL} \xrightarrow{P} \sigma^2_{GL} \text{ for } n \to \infty$$

where

$$\hat{\sigma}^2_{GL} = \sum_{r=-(n-1)}^{n-1} \kappa\left(\frac{|r|}{b_n}\right) \frac{1}{n} \sum_{i=1}^{n-|r|} \hat{A}_1(X_i)\hat{A}_1(X_{i+|r|}),$$

with

$$\hat{A}_1(x) = \frac{1}{n^{m-1}} \sum_{1\le i_1<\ldots<i_{m-1}\le n} A(x, X_{i_1}, \ldots, X_{i_{m-1}}) - \frac{1}{n^m} \sum_{1\le i_1<\ldots<i_m\le n} A(X_{i_1}, \ldots, X_{i_m})$$

being the estimator for the first term of the Hoeffding-decomposition of A.

The proofs are given in the following section.

5.3.3. Proofs

Proof of Theorem 5.2.
Because of the considerations in Section 5.2 and the resulting general limit theorem for GL-Statistics (Theorem 5.1) it remains to show the three conditions in the special case of near epoch dependent data. That is:

(i) For $W_{H_n,H_F}(y) = \left(\frac{\int_0^{H_n(y)} J(t)dt - \int_0^{H_F(y)} J(t)dt}{H_n(y) - H_F(y)} - J(H_F(y))\right)$ holds
$\|W_{H_n,H_F}\|_{L_1} = o_p(1)$ and it is $\|H_n - H_F\|_\infty = O_p(n^{-\frac{1}{2}})$.

(ii) For the remainder term $R_{p_i,n} = \hat{\xi}_{p_i,n} - \xi_{p_i} + \frac{p_i - H_n(\xi_{p_i})}{h_f(\xi_{p_i})}$ of the Bahadur representation of an empirical quantile it holds that

$$R_{p_i,n} = o_p(n^{-\frac{1}{2}}).$$

(iii) For a U-statistic with kernel

$$A(x_1, \ldots, x_m) = -\int_{-\infty}^{\infty} \left(\mathbb{1}_{[h(x_1,\ldots,x_m)\le y]} - H_F(y)\right) J(H_F(y))dy + \sum_{i=1}^{d} a_i \frac{p_i - \mathbb{1}_{\left[h(x_1,\ldots,x_m)\le H_F^{-1}(p_i)\right]}}{h_F(H_F^{-1}(p_i))}$$

we have

$$\sqrt{n}(U_n(A) - \theta) \xrightarrow{D} N(0, \sigma^2).$$

As in the case of strong mixing condition (i) is fulfilled using Lemma 8.2.4.A of Serfling (1980) and Corollary 4.1. Condition (ii) can be proved by Theorem 4.5.

It remains to show that condition (iii) is satisfied. For this we apply Theorem 4.1. Again, the conditions on the coefficients a_l and $\beta(l)$ are fulfilled by the assumptions, similar to Corollary 4.1. The assumptions on the kernel are fulfilled because of Lemma 5.1.

□

Proof of Corollary 5.1.
As we have mentioned before, the error term $T(H_n)-T(H_F)$ can be approximated by a U-statistic with kernel A. This leads to the special structure of σ^2_{GL}, being similar to that of a long-run variance of U-statistics. Therefore, we want to apply Theorem 4.3 to prove this Corollary. It only remains to show the assumptions on the kernel A. These are fulfilled because of Lemma 5.1. □

5.4. Robustness of GL-statistics

In this section we want to show under which assumptions a GL-statistic is robust according to the influence function and the breakdown point.
We have already seen that the influence function can be obtained by replacing H_n with the pointmass in the Gâteaux differential. That is, the influence function of a GL-statistic is given by

$$\begin{aligned}
&IC(x_1,\dots,x_m;H_n,T)\\
&= \frac{d}{d\lambda}T(H_n+\lambda(\delta_x-H_n))\Big|_{\lambda=0+}\\
&= -\int_{-\infty}^{\infty}\left(\mathbb{1}_{[h(x_1,\dots,x_m)\le y]}-H_F(y)\right)J(H_F(y))dy+\sum_{i=1}^{d}a_i\frac{p_i-\mathbb{1}_{[h(x_1,\dots,x_m)\le H_F^{-1}(p_i)]}}{h_F(H_F^{-1}(p_i))}\\
&= A(x_1,\dots,x_m).
\end{aligned}$$

This is bounded under the conditions of the Central Limit Theorem 5.2 and therefore GL-statistics with bounded smooth weighting function J vanishing outside $(0,1)$ are robust.
The breakdown point of a GL-statistic depends much on the chosen coefficients. For example, for L-statistics it is known that they have a breakdown point equal to the order statistic closest to the minimum or maximum. Hence, the median has breakdown point 0 and the α-trimmed mean has a breakdown point of α (see e.g. Huber (1981)). Special forms of GL-statistics with kernels $h(x_1,x_2)=|x_1-x_2|$ or $h(x_1,x_2)=(x_1-x_2)^2$ corresponding to scale estimators have a breakdown point depending on the coefficients a_i. Rousseeuw and Croux (1992) show that in this case the highest breakdown points are obtained for coefficients $a_i=0$, if $i>\binom{[n/2]+1}{2}$, and if $\binom{[n/2]}{2}\le i\le\binom{[n/2]+1}{2}$ coefficients with $a_i>0$.

6. Asymptotics of Robust Estimators under Short-Range Dependence

After developing the theoretical background for a general class of statistics we want to give several certain applications. On the one hand, a concrete limit theorem for Linear-moments and trimmed Linear-moments is given, which allows us to calculate confidence intervals for parametric estimators used in hydrology. On the other hand, we want to show via statistical simulations the validity of the Central Limit Theorem for the robust estimators proposed in Section 5.

6.1. L-Moments and TL-Moments

Linear-moment based estimators are widely used in hydrology for the estimation of distribution parameters. A robust extension of this concept are the trimmed linear-moments. For both approaches limit theorems will be developed in the following.

6.1.1. L-Moments

The probability weighted moments (PWM) have been developed by Greenwood et al. (1979) to express parameters of easily invertible distributions by moments. As an advancement of the standard and the probability weighted moments Hosking (1990) suggested the so called Linear-moments (L-moments), which are estimated by a linear combination of order statistics (that is L-statistics). Used as an estimator they have the advantage of being similar to and for small samples sometimes even more efficient than the Maximum-Likelihood estimator and also more robust than the Moment estimators (Hosking (1990)), in the sense that they are less effected by single extreme large or small values. They exist in situations, where the classical moments do not exist and represent in contrast to the PWM the characteristic values of a sample such as mean or standard deviation straightforwardly. Therefore, they are used more frequently than the PWM but deliver the same results.
L-moments are among the most frequently used estimators in hydrology, since they serve the desire of having efficient estimators for small sample sizes as well as easy understandable and computationally fast expressions. Hence, in our context they are of special interest.
Let X be a real-valued random variable with cumulative distribution function F and quantile function $x(F)$. The r-th L-moment is defined as

$$\lambda_r = \frac{1}{r}\sum_{k=0}^{r-1}(-1)^k\binom{r-1}{k}\mathbb{E}X_{(r-k:r)}, \tag{6.1}$$

where $X_{(i:n)}$ is the i-th value of the order statistics of a sample $X_1,\dots,X_n$ drawn from F and $r = 1, 2, \dots$. For example, the first two L-moments are given by

$$\lambda_1 = \mathbb{E}X = \int_0^1 x(F)dF,$$

$$\lambda_2 = \frac{1}{2}\mathbb{E}\left(X_{(2:2)} - X_{(1:2)}\right) = \int_0^1 x(F)(2F-1)dF.$$

Hosking (1986) has shown that λ_r exists for $r = 1, 2, \ldots$ if and only if $\mathbb{E}|X|$ exists.

For the L-moments two different estimators have been proposed by Hosking (1990). The first one is the so-called sample L-moment defined by

$$\begin{aligned} l_r &= \frac{1}{\binom{n}{r}} \sum_{1\le i_1<\ldots<i_r\le n} r^{-1} \sum_{k=0}^{r-1} (-1)^k \binom{r-1}{k} x_{(i_{r-k}:n)} \\ &= \sum_{k=0}^{r-1} (-1)^{r-1-k} \binom{r-1}{k}\binom{r-1+k}{k} \frac{1}{n} \frac{1}{\binom{n-1}{k}} \sum_{j=k+1}^{n} \binom{j-1}{k} x_{(j:n)} \\ &=: \sum_{k=0}^{r-1} p^\star_{r-1,k} \frac{1}{n} \frac{1}{\binom{n-1}{k}} \sum_{j=k+1}^{n} \binom{j-1}{k} x_{(j:n)}. \end{aligned}$$

Analogously, we denote $b_k = \frac{1}{n}\frac{1}{\binom{n-1}{k}} \sum_{j=k+1}^{n} \binom{j-1}{k} x_{(j:n)}$ as the sample estimator for the r-th PWM β_r (Greenwood et al. (1979)) with

$$\beta_r = \mathbb{E}\left(X(F(X)^r\right).$$

It is easy to see that l_r is a U-statistic and therefore asymptotically normal under independence and short-range dependence. Additionally, l_r is of course an unbiased estimator for λ_r.
Another estimator is given by the use of the so-called plotting positions. Plotting positions $p_{i:n}$, $i = 1, \ldots, n$, are distribution-free estimators of $F(x_{(i:n)})$. The general formula is given by

$$p_{i:n} = \frac{i+\gamma}{n+\delta}, \quad i = 1, \ldots, n,$$

where $\delta > \gamma > -1$. There are a lot of different choices used for δ and γ. The classical one, also called Weibull-formula, is $p_{i:n} = i/(n+1)$. Hosking (1990) also proposes $p_{i:n} = (i - 0.35)/n$. These plotting positions are then used in the plotting-position estimator as follows.

$$\tilde{\lambda}_r = \frac{1}{n} \sum_{i=1}^{n} P^\star_{r-1}(p_{i:n}) x_{(i:n)},$$

with $P^\star_r(x) = \sum_{k=0}^{r} (-1)^{r-k} \binom{r}{k}\binom{r+k}{k} x^k = \sum_{k=0}^{r} p^\star_{r,k} x^k$ being the r-th shifted Legendre-polynomial. This estimator is an L-statistic with $c_{i,n} = P^\star_{r-1}(p_{i:n})$.
Again, we can also use an expression by the plotting position estimators for the PWM, $\tilde{\beta}_r = \frac{1}{n}\sum_{i=1}^{n} \left(\frac{i+\gamma}{n+\delta}\right)^r x_{(i:n)} = \frac{1}{n}\sum_{i=1}^{n} \frac{\sum_{k=0}^{r} g_{rk}(\gamma)(i-1)(i-2)\cdot\ldots\cdot(i-r+k)}{(n+\delta)^r} x_{(i:n)}, k = 1, \ldots, r-1,$

$$\tilde{\lambda}_r = \sum_{k=0}^{r-1} p^\star_{r-1,k} \tilde{\beta}_k.$$

Both estimators are closely related. In fact, $l_r - \tilde{\lambda}_r = O_p(1/n)$ if $\mathbb{E}X_i < \infty$ as pointed out by Hosking (1986). We want to give a short sketch of the proof of Hosking (1986) for this identity to show that it does not depend on the dependence structure of the X_i.

As we can see by the expressions of the L-moment estimators by the PWM-estimators it is sufficient to show that

$$b_r - \tilde{\beta}_r = O_p\left(\frac{1}{n}\right).$$

For this we want to express $\tilde{\beta}_r$ by a sum of b_s, $s = 1, \ldots, r$.
It is helpful to express the denominator of the plotting position as the sum over factorials

$$(i+\gamma)^r = \sum_{k=0}^{r} g_{rk}(\gamma)(i-1)(i-2)\cdot\ldots\cdot(i-r+k),$$

where g_{rk} is a function only depending on r and k, for that Hosking (1986) has shown that

$$g_{r0}(\gamma) = 1, \quad g_{r1}(\gamma) = \frac{1}{2}r(r+1) + r\gamma.$$

This can be used as follows

$$\begin{aligned}
\tilde{\beta}_r &= \frac{1}{n}\sum_{i=1}^{n}\left(\frac{i+\gamma}{n+\delta}\right)^r x_{(i:n)} = \frac{1}{n}\sum_{i=1}^{n}\frac{\sum_{k=0}^{r} g_{rk}(\gamma)(i-1)(i-2)\cdot\ldots\cdot(i-r+k)}{(n+\delta)^r}x_{(i:n)} \\
&= \frac{1}{(n+\delta)^r}\sum_{k=0}^{r} g_{rk}(\gamma)\frac{1}{n}\sum_{i=1}^{n}(n-1)(n-2)\cdot\ldots\cdot(n-r+k)\frac{\binom{i-1}{r-k}}{\binom{n-1}{r-k}}x_{(i:n)} \\
&= \frac{1}{(n+\delta)^r}\sum_{k=0}^{r} g_{rk}(\gamma)(n-1)(n-2)\cdot\ldots\cdot(n-r+k)b_{r-k}.
\end{aligned}$$

The term $N_k := \frac{(n-1)(n-2)\cdot\ldots\cdot(n-r+k)}{(n+\delta)^r}$ has the following asymptotic order

$$N_k = \begin{cases} 1 - (1/2r(r+1) + r\delta)n^{-1} + O(n^{-2}) & , \text{ if } k = 0 \\ n^{-1} + O(n^{-2}) & , \text{ if } k = 1 \\ O(n^{-k}) & , \text{ if } k \geq 2 \end{cases}.$$

Hence, since $\mathbb{E}(b_r) = \beta_r$,

$$\begin{aligned}
b_r - \tilde{\beta}_r &= b_r - \sum_{k=0}^{r} g_{rk}(\gamma)N_k b_{r-k} \\
&= (1 - g_{r0}(\gamma)N_0)b_r - \sum_{k=1}^{r} g_{rk}(\gamma)N_k b_{r-k} \\
&= (1 - (1 - (1/2r(r+1) + r\delta)n^{-1} + O(n^{-2})))b_r - \sum_{k=1}^{r} g_{rk}(\gamma)N_k b_{r-k}.
\end{aligned}$$

Note that all terms of order $k \geq 2$ are asymptotically negligible since their asymptotical order is smaller than that of the first two terms.
Finally,

$$\mathbb{E}\left(b_r - \tilde{\beta}_r\right)$$

$$= \left((1/2r(r+1) + r\delta)n^{-1} + O(n^{-2})\right)\beta_r - \left((1/2r(r+1) + r\delta)n^{-1} + O(n^{-2})\right)\beta_{r-1} + \dots$$
$$\leq C\frac{1}{n}\beta_r \leq C'\frac{1}{n},$$

since $\beta_r < \infty$.

A disadvantage of $\tilde{\lambda}_r$ is that it is not an unbiased estimator but it is consistent (Hosking (1986)). The plotting-position estimator is also useful when proving the asymptotic normality of a vector of the l_s, $s = 1, \dots, r$. Under independence this has already been done by Hosking (1990) and we want to establish an analogous result for NED sequences on absolute regular processes.

Theorem 6.1. *Let $X_1, \dots, X_n$ be NED with approximation constants $(a_l)_{l\in\mathbb{N}}$ on an absolutely regular process $(Z_n)_{n\in\mathbb{Z}}$ with mixing coefficients $(\beta(l))_{l\in\mathbb{N}}$ with $\sum_{l=1}^{\infty} l^2\beta^{\frac{\gamma}{2+\gamma}}(l) < \infty$ for a $0 < \gamma < 1$. Moreover, let be $\sum_{l=1}^{\infty} l^2 a_l^{\frac{\gamma}{2+2\gamma}} < \infty$. Then,*

$$\sqrt{n}\left((l_1 - \lambda_1), \dots, (l_r - \lambda_r)\right) \xrightarrow{D} \mathcal{N}(0, \Sigma)$$

with $\Sigma = (\rho_{i,j})_{1\leq i,j\leq r}$ and

$$\rho_{i,j} = Cov(l_i, l_j)$$
$$= \frac{1}{2}\Bigg(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left(P_{i-1}^{\star}(F(x))P_{j-1}^{\star}(F(y))\right) + \left(P_{i-1}^{\star}(F(y))P_{j-1}^{\star}(F(x))\right)$$
$$\cdot \left(F(x)(1-F(y)) + 2\sum_{j=1}^{\infty}\left(\mathbb{P}\left(X_1 \leq F(x), X_{j+1} \leq F(y) - F(x)F(y)\right)\right)\right)dxdy\Bigg).$$

Proof.
As mentioned above $\tilde{\lambda}_r$ can be expressed as an L-statistic and hence is asymptotically normal by Theorem 5.2. The assumptions of the theorem for the kernel h and the weighting function J are fulfilled since due to the expression as L-statistic $h(x) = x$ is the identity function and therefore Lipschitz continuous and H_F reduces to F, which is also continuous. The function J is chosen as $J = P_{r-1}^{\star}$ and hence is bounded on $(0,1)$ as well as continuous a.e. concerning the Lebesgue-measure and F^{-1} and can be defined as equal to 0 for all values not in $(0,1)$. Moreover, also the linear combination of the first r plotting-position estimators is an L-statistic

$$\sum_{i=1}^{r} a_i\tilde{\lambda}_r = \sum_{i=1}^{r} a_i P_{i-1}^{\star}(p_{j:n})x_{(j:n)} =: \sum_{i=1}^{r} P^{\star}(p_{j:n})x_{(j:n)}$$

and hence also converges to a normal distribution according to Theorem 5.2. Using the Cramer-Wold-device we then know that

$$\sqrt{n}\left((\tilde{\lambda}_1 - \lambda_1), \dots, (\tilde{\lambda}_r - \lambda_r)\right) \longrightarrow \mathcal{N}(0, \Sigma).$$

The entries of the covariance matrix Σ can be calculated by using the identity

$$\mathrm{Cov}(\tilde{\lambda}_r, \tilde{\lambda}_s) = \frac{1}{2}\left(\mathrm{Var}(\tilde{\lambda}_r + \tilde{\lambda}_s) - \mathrm{Var}(\tilde{\lambda}_r) - \mathrm{Var}(\tilde{\lambda}_s)\right). \tag{6.2}$$

For this we first need the asymptotic variance of L-statistics. Using the formula for the variance of Theorem 5.2 and the special form of it for L-statistics, we obtain

$$\begin{aligned}\sigma_L^2 =&\mathbb{E}\left(\left(\mathbb{E}\left(-\int_{-\infty}^{\infty}\left(\mathbb{1}_{[Y_1\leq y]}-F(y)\right)J(F(y))dy|Y_1=X_1\right)\right)^2\right)\\&+2\sum_{j=1}^{\infty}\mathrm{Cov}\Bigg(\mathbb{E}\left(-\int_{-\infty}^{\infty}\left(\mathbb{1}_{[Y_1\leq y]}-F(y)\right)J(F(y))dy|Y_1=X_1\right),\\&\qquad\mathbb{E}\left(-\int_{-\infty}^{\infty}\left(\mathbb{1}_{[Y_1\leq y]}-F(y)\right)J(F(y))dy|Y_1=X_{j+1}\right)\Bigg).\end{aligned}\tag{6.3}$$

The first term of eq. (6.3) can be simplified analogously to the independent case (see e.g. Lee (1990)) by

$$\begin{aligned}&\mathbb{E}\left(\left(\mathbb{E}\left(-\int_{-\infty}^{\infty}\left(\mathbb{1}_{[Y_1\leq y]}-F(y)\right)J(F(y))dy\Big|Y_1=X_1\right)\right)^2\right)\\&=\mathbb{E}\left(\left(\mathbb{E}\left(-\int_{-\infty}^{\infty}\mathbb{1}_{[Y_1\leq y]}J(F(y))dy\Big|Y_1=X_1\right)+\int_{-\infty}^{\infty}F(y)J(F(y))dy\right)^2\right)\\&=\mathbb{E}\Bigg[\left(\mathbb{E}\left(-\int_{-\infty}^{\infty}\mathbb{1}_{[Y_1\leq y]}J(F(y))dy\Big|Y_1=X_1\right)\right)^2\\&+2\mathbb{E}\left(-\int_{-\infty}^{\infty}\mathbb{1}_{[Y_1\leq y]}J(F(y))dy\Big|Y_1=X_1\right)\int_{-\infty}^{\infty}F(y)J(F(y))dy\\&+\left(\int_{-\infty}^{\infty}F(y)J(F(y))dy\right)^2\Bigg]\\&=\mathbb{E}\left(\mathbb{E}\left(-\int_{-\infty}^{\infty}\left(\mathbb{1}_{[Y_1\leq y]}\right)J(F(y))dy\Big|Y_1=X_1\right)\right)^2\\&-2\int_{-\infty}^{\infty}F(y)J(F(y))dy\int_{-\infty}^{\infty}\mathbb{E}\left(\mathbb{E}\left(\mathbb{1}_{[Y_1\leq y]}\Big|Y_1=X_1\right)\right)J(F(y))dy\\&+\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}F(x)F(y)J(F(x))J(F(y))dxdy\\&=\mathbb{E}\left(\mathbb{E}\left(-\int_{-\infty}^{\infty}\left(\mathbb{1}_{[Y_1\leq y]}\right)J(F(y))dy\Big|Y_1=X_1\right)\right)^2\\&-2\int_{-\infty}^{\infty}F(y)J(F(y))dy\int_{-\infty}^{\infty}\mathbb{E}\left(\mathbb{1}_{[Y_1\leq y]}\right)J(F(y))dy\\&+\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}F(x)F(y)J(F(x))J(F(y))dxdy\\&=\mathbb{E}\left(\mathbb{E}\left(-\int_{-\infty}^{\infty}\left(\mathbb{1}_{[Y_1\leq y]}\right)J(F(y))dy\Big|Y_1=X_1\right)\right)^2\\&-2\int_{-\infty}^{\infty}F(y)J(F(y))dy\int_{-\infty}^{\infty}F(y)J(F(y))dy\\&+\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}F(x)F(y)J(F(x))J(F(y))dxdy\\&=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}F(y)J(F(x))J(F(y))dxdy\end{aligned}$$

$$
\begin{aligned}
&- \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F(x)F(y)J(F(x))J(F(y))dxdy \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F(x)(1-F(y))J(F(x))J(F(y))dxdy.
\end{aligned}
$$

Also the second term of eq. (6.3) can be simplified by

$$
\begin{aligned}
&2\sum_{j=1}^{\infty}\mathrm{Cov}\Bigg(\mathbb{E}\left(-\int_{-\infty}^{\infty}\left(\mathbb{1}_{[Y_1\le y]}-F(y)\right)J(F(y))dy\Big|Y_1=X_1\right), \\
&\qquad\qquad \mathbb{E}\left(-\int_{-\infty}^{\infty}\left(\mathbb{1}_{[Y_1\le y]}-F(y)\right)J(F(y))dy\Big|Y_1=X_{j+1}\right)\Bigg) \\
=&2\sum_{j=1}^{\infty}\mathbb{E}\Bigg[\mathbb{E}\left(-\int_{-\infty}^{\infty}\left(\mathbb{1}_{[Y_1\le y]}-F(y)\right)J(F(y))dy\Big|Y_1=X_1\right) \\
&\qquad\qquad \cdot\,\mathbb{E}\left(-\int_{-\infty}^{\infty}\left(\mathbb{1}_{[Y_1\le y]}-F(y)\right)J(F(y))dy\Big|Y_1=X_{j+1}\right)\Bigg] \\
=&2\sum_{j=1}^{\infty}\mathbb{E}\Bigg[\left(\mathbb{E}\left(-\int_{-\infty}^{\infty}\left(\mathbb{1}_{[Y_1\le y]}J(F(y))dy|Y_1=X_1\right)\right)+\int_{-\infty}^{\infty}(F(y)J(F(y)))\,dy\right) \\
&\cdot\left(\mathbb{E}\left(-\int_{-\infty}^{\infty}\left(\mathbb{1}_{[Y_1\le y]}J(F(y))dy\Big|Y_1=X_{j+1}\right)\right)+\int_{-\infty}^{\infty}(F(y)J(F(y)))\,dy\right)\Bigg] \\
=&2\sum_{j=1}^{\infty}\Bigg[\mathbb{E}\left(\mathbb{E}\left(-\int_{-\infty}^{\infty}\left(\mathbb{1}_{[Y_1\le y]}J(F(y))dy\Big|Y_1=X_1\right)\right)\cdot\mathbb{E}\left(-\int_{-\infty}^{\infty}\left(\mathbb{1}_{[Y_1\le y]}J(F(y))dy\Big|Y_1=X_{j+1}\right)\right)\right) \\
&\qquad+\mathbb{E}\left(\mathbb{E}\left(-\int_{-\infty}^{\infty}\left(\mathbb{1}_{[Y_1\le y]}J(F(y))dy\Big|Y_1=X_1\right)\right)\cdot\int_{-\infty}^{\infty}(F(y)J(F(y)))\,dy\right) \\
&\qquad+\mathbb{E}\left(\mathbb{E}\left(-\int_{-\infty}^{\infty}\left(\mathbb{1}_{[Y_1\le y]}J(F(y))dy\Big|Y_1=X_{j+1}\right)\right)\cdot\int_{-\infty}^{\infty}(F(y)J(F(y)))\,dy\right) \\
&\qquad+\mathbb{E}\left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}F(x)F(y)J(F(x))J(F(y))dxdy\right)\Bigg] \\
=&2\sum_{j=1}^{\infty}\Bigg[\mathbb{E}\left(\mathbb{E}\left(-\int_{-\infty}^{\infty}\left(\mathbb{1}_{[Y_1\le y]}J(F(y))dy|Y_1=X_1\right)\right)\cdot\mathbb{E}\left(-\int_{-\infty}^{\infty}\left(\mathbb{1}_{[Y_1\le y]}J(F(y))dy|Y_1=X_{j+1}\right)\right)\right) \\
&\qquad+\mathbb{E}\left(\mathbb{E}\left(-\int_{-\infty}^{\infty}\left(\mathbb{1}_{[Y_1\le y]}J(F(y))dy\Big|Y_1=X_1\right)\right)\right)\int_{-\infty}^{\infty}(F(y)J(F(y)))\,dy \\
&\qquad+\mathbb{E}\left(\mathbb{E}\left(-\int_{-\infty}^{\infty}\left(\mathbb{1}_{[Y_1\le y]}J(F(y))dy\Big|Y_1=X_{j+1}\right)\right)\right)\int_{-\infty}^{\infty}(F(y)J(F(y)))\,dy \\
&\qquad+\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}F(x)F(y)J(F(x))J(F(y))dxdy\Bigg] \\
=&2\sum_{j=1}^{\infty}\Bigg[\mathbb{E}\left(\mathbb{E}\left(-\int_{-\infty}^{\infty}\left(\mathbb{1}_{[Y_1\le y]}J(F(y))dy\Big|Y_1=X_1\right)\right)\cdot\mathbb{E}\left(-\int_{-\infty}^{\infty}\left(\mathbb{1}_{[Y_1\le y]}J(F(y))dy\Big|Y_1=X_{j+1}\right)\right)\right)
\end{aligned}
$$

$$
\begin{aligned}
&- \mathbb{E}\left(\int_{-\infty}^{\infty} \mathbb{1}_{[Y_1 \leq y]} J(F(y))dy\right) \int_{-\infty}^{\infty} (F(y)J(F(y)))\, dy \\
&- \mathbb{E}\left(\int_{-\infty}^{\infty} \mathbb{1}_{[Y_1 \leq y]} J(F(y))dy\right) \int_{-\infty}^{\infty} (F(y)J(F(y)))\, dy \\
&+ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F(x)F(y)J(F(x))J(F(y))dxdy\Bigg] \\
=&2\sum_{j=1}^{\infty} \mathbb{E}\Bigg[\mathbb{E}\left(-\int_{-\infty}^{\infty}\left(\mathbb{1}_{[Y_1 \leq y]} J(F(y))dy \middle| Y_1 = X_1\right)\right) \\
&\cdot \mathbb{E}\left(-\int_{-\infty}^{\infty}\left(\mathbb{1}_{[Y_1 \leq y]} J(F(y))dy \middle| Y_1 = X_{j+1}\right)\right) \\
&- 2\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F(x)F(y)J(F(x))J(F(y))dxdy \\
&+ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F(x)F(y)J(F(x))J(F(y))dxdy\Bigg] \\
=&2\sum_{j=1}^{\infty}\Bigg[\int_{-\infty}^{\infty} \left(\mathbb{P}\left(X_1 \leq F(x), X_{j+1} \leq F(y)\right)\right) J(F(x))J(F(y))dxdy \\
&- \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F(x)F(y)J(F(x))J(F(y))dxdy\Bigg] \\
=&2\sum_{j=1}^{\infty}\left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left(\mathbb{P}\left(X_1 \leq F(x), X_{j+1} \leq F(y)\right) - F(x)F(y)\right) J(F(x))J(F(y))dxdy\right).
\end{aligned}
$$

Combining these two results we obtain for the asymptotic variance of an L-statistic under near epoch dependence

$$
\begin{aligned}
\sigma_L^2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \Bigg(&F(x)(1-F(y)) + 2\sum_{j=1}^{\infty}\left(\mathbb{P}\left(X_1 \leq F(x), X_{j+1} \leq F(y)\right) - F(x)F(y)\right)\Bigg) \\
&\cdot J(F(x))J(F(y))dxdy,
\end{aligned}
$$

which is consistent with the result of Puri and Tran (1980) for L-Statistics for strong mixing data.

Now we can use (6.2) and the fact that the sum of two plotting-position estimators is again an L-statistic to calculate the covariance of the plotting-position estimators. Remember that we have chosen $J(t) = P^{\star}_{i-1}(t)$ for $\tilde{\lambda}_i$.

$$
\begin{aligned}
&\mathrm{Cov}(\tilde{\lambda}_r, \tilde{\lambda}_s) \\
&= \frac{1}{2}\Bigg[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left(P^{\star}_{r-1}(F(x)) + P^{\star}_{s-1}(F(x))\right) \cdot \left(P^{\star}_{r-1}(F(y)) + P^{\star}_{s-1}(F(y))\right)
\end{aligned}
$$

$$\cdot \left(F(x)(1-F(y)) + 2\sum_{j=1}^{\infty} \left(\mathbb{P}\left(X_1 \le F(x), X_{j+1} \le F(y)\right) - F(x)F(y) \right) \right) dxdy$$

$$- \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} P_{r-1}^{\star}(F(x))P_{r-1}^{\star}(F(y))$$

$$\left(F(x)(1-F(y)) + 2\sum_{j=1}^{\infty} \left(\mathbb{P}\left(X_1 \le F(x), X_{j+1} \le F(y)\right) - F(x)F(y) \right) \right) dxdy$$

$$- \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} P_{s-1}^{\star}(F(x))P_{s-1}^{\star}(F(y))$$

$$\left(F(x)(1-F(y)) + 2\sum_{j=1}^{\infty} \left(\mathbb{P}\left(X_1 \le F(x), X_{j+1} \le F(y)\right) - F(x)F(y) \right) \right) dxdy \Bigg]$$

$$= \frac{1}{2}\Bigg[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left(P_{r-1}^{\star}(F(x)) + P_{s-1}^{\star}(F(x))\right) \cdot \left(P_{r-1}^{\star}(F(y)) + P_{s-1}^{\star}(F(y))\right)$$

$$\left(F(x)(1-F(y)) + 2\sum_{j=1}^{\infty} \left(\mathbb{P}\left(X_1 \le F(x), X_{j+1} \le F(y)\right) - F(x)F(y) \right) \right) dxdy \Bigg].$$

Using again $l_r - \tilde{\lambda}_r = O_p(1/n)$ the vector of the sample estimators has the same distribution as the plotting-position estimator, that is a straightforward extension of the independent case.

□

Using the Delta-Method it is now easy to calculate the asymptotic distribution of many parameter-estimators based on L-moments.

6.1.2. Trimmed L-Moments

Trimmed Linear-moments (TL-moments) are a robust modification of L-moments. They have been introduced by Elamir and Seheult (2003) and, due to their easy computation, are frequently used and developed for many different distributions (cf. Hosking (2007), Abdul-Moniem and Selim (2009), Ahmad et al. (2011)).
Let X be a real-valued random variable with cumulative distribution function F. The r-th TL-moment with trimming (t_1, t_2) is given by

$$\lambda_r^{(t_1,t_2)} = \frac{1}{r}\sum_{k=0}^{r-1}(-1)^k \binom{r-1}{k} \mathbb{E}\left(X_{(r+t_1-k:r+t_1+t_2)}\right),$$

where we can use the representation

$$\mathbb{E}\left(X_{(i:r)}\right) = \frac{r!}{(i-1)!(r-i)!}\int_0^1 x(F)F^{i-1}(1-F)^{r-i}dF$$

of the expected value of the i-th order-statistic of a sample with length r drawn from the distribution of X. $x(F)$ is the quantile function of the distribution F.
An unbiased estimate of $\lambda_r^{(t_1,t_2)}$ is

$$l_r^{(t_1,t_2)} = \frac{1}{r}\sum_{i=t_1+1}^{n-t_2} \frac{\sum_{k=0}^{r-1}(-1)^k \binom{r-1}{k}\binom{i-1}{r+t_1-k-1}\binom{n-i}{t_2+k}}{\binom{n}{r+t_1+t_2}} x_{(i:n)}. \tag{6.4}$$

The main idea of this kind of estimator to gain robustness is trimming by giving zero weight to the t_1 smallest and t_2 largest values.

Again, similar to L-moments, the plotting position estimator proves to be a useful tool for obtaining asymptotic results. It is defined as

$$\tilde{\lambda}_r^{(t_1,t_2)} = \frac{1}{n}\sum_{i=1}^{n} J_r^{(t_1,t_2)} x_{(i:n)}$$

with

$$J_r^{(t_1,t_2)}(x) = \frac{(r-1)!(r+t_1+t_2)!}{r(r+t_1-1)!(r+t_2-1)!} x^{t_1}(1-x)^{t_2} P_{r-1}^{\star(t_2,t_1)}(x)$$

and

$$P_r^{\star(t_2,t_1)}(x) = \sum_{j=0}^{r}(-1)^{r-j}\binom{r+t_2}{j}\binom{r+t_1}{r-j}x^j(1-x)^{r-j},$$

the shifted Jacobi polynomial. The Legendre polynomials are a special case of the Jacobi polynomials that are needed in this more general context.

Analogously to L-moments, which are in fact a special case of TL-moments with $t_1 = t_2 = 0$, we can state a result about asymptotic normality. The proof of it is analogous to the one for L-moments (see e.g. Hosking (2007) for the independent case) and therefore omitted here. Note that we limit the possible trimming to be smaller or equal to 2 to simplify the results. In practice, as we will see later on, all relevant cases are covered by this. For an idea of how the covariance looks like for values of trimming larger than 2 we refer to Hosking (2007).

Theorem 6.2. *Let $X_1, \dots, X_n$ be NED with approximation constants $(a_l)_{l\in\mathbb{N}}$ on an absolutely regular process $(Z_n)_{n\in\mathbb{Z}}$ with mixing coefficients $(\beta(l))_{l\in\mathbb{N}}$ with $\sum_{l=1}^{\infty} l^2\beta^{\frac{\gamma}{2+\gamma}}(l) < \infty$ for a $0 < \gamma < 1$. Moreover, let be $\sum_{l=1}^{\infty} l^2 a_l^{\frac{\gamma}{2+2\gamma}} < \infty$. Then for all $0 \le t_1, t_2 \le 2$*

$$\sqrt{n}\left((\hat{\lambda}_1^{(t_1,t_2)} - \lambda_1^{(t_1,t_2)}), \dots, (\hat{\lambda}_r^{(t_1,t_2)} - \lambda_r^{(t_1,t_2)})\right) \longrightarrow \mathcal{N}(0,\Sigma)$$

with $\Sigma = (\rho_{i,j})_{1\le i,j\le r}$ and

$$\begin{aligned}\rho_{i,j} &= Cov(\hat{\lambda}_i^{(t_1,t_2)}, \hat{\lambda}_j^{(t_1,t_2)}) \\ &= \frac{1}{2}\Bigg[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left(J_i^{(t_1,t_2)}(F(x)) + J_j^{(t_1,t_2)}(F(x))\right)\cdot\left(J_i^{(t_1,t_2)}(F(y)) + J_j^{(t_1,t_2)}(F(y))\right) \\ &\qquad \left(F(x)(1-F(y)) + 2\sum_{j=1}^{\infty}\mathbb{P}\left(X_1 \le F(x), X_{j+1} \le F(y)\right) - F(x)F(y)\right)dxdy\Bigg].\end{aligned}$$

In the following we present some specific estimators based on TL-moments for distributions that are important in our context.

Generalized Pareto Distribution (GPD)
For the Generalized Pareto Distribution the distribution function is given by

$$F(x) = 1 - \left(1 + \kappa \frac{x-\mu}{\beta}\right)^{-\frac{1}{\kappa}},$$

for $x > \mu$ if the shape parameter is $\kappa \geq 0$, and $\mu \leq x \leq \mu - \beta/\kappa$ otherwise, the scale parameter $\beta > 0$ and the threshold-parameter μ, often also notated as x_0. A special case of this distribution is the exponential distribution.
Therefore, the corresponding quantile is given by

$$x(F) = \mu - \frac{\beta}{\kappa}\left(1 - (1-F)^{-\kappa}\right).$$

Often, we want to have robustness against single large events for an estimator of the parameter of the GPD. The lower part of the sample is already censored by choosing a threshold x_0 and only considering values above this threshold. Since the GPD is bounded to the left, it seems sensible to only consider trimming to the right. Therefore, we chose $t_1 = 0$ and $t_2 = 1$. For the cases $t_1 = t_2 = 1$ and $t_1 = 1, t_2 = 0$ estimators for the GPD have already been developed (Abdul-Moniem and Selim (2009), Ahmad et al. (2011)), and Hosking (2007) developed TL(0,1)-estimators for the two parameter GPD. We want to consider the three-parameter GPD as well, which is why we present the explanation to the calculations below. Explicit representations of the first three TL(0,1)-moments are

$$\begin{aligned}
\lambda_1^{(0,1)} &= \mathbb{E}\left(X_{(1:2)}\right) \\
\lambda_2^{(0,1)} &= \tfrac{1}{2}\left(\mathbb{E}\left(X_{(2:3)}\right) - \mathbb{E}\left(X_{(1:3)}\right)\right) \\
\lambda_3^{(0,1)} &= \tfrac{1}{3}\left(\mathbb{E}\left(X_{(3:4)}\right) - 2\mathbb{E}\left(X_{(2:4)}\right) + \mathbb{E}\left(X_{(1:4)}\right)\right).
\end{aligned}$$

We can see that indeed the highest value of the order statistics is left out in the calculation. With the equations above, the first three TL(0,1)-moments can be written as

$$\begin{aligned}
\lambda_1^{(0,1)} &= \textstyle\int_0^1 x(F)(1-F)2dF \\
\lambda_2^{(0,1)} &= \tfrac{1}{2}\textstyle\int_0^1 x(F)(1-F)(6F-3(1-F))dF \\
\lambda_3^{(0,1)} &= \tfrac{1}{3}\textstyle\int_0^1 x(F)(1-F)(12F^2-24F(1-F)+4(1-F)^2)dF
\end{aligned}$$

and the ratio as

$$\tau_3^{(0,1)} = \frac{\lambda_3^{(0,1)}}{\lambda_2^{(0,1)}}.$$

By calculating the integrals above and using substitution we obtain the following estimators, where the TL(0,1)-moments are estimated via (6.4) and $t_3^{(0,1)} = l_3^{(0,1)}/l_2^{(0,1)}$,

$$\begin{aligned}
\hat{\kappa} &= \tfrac{36t_3^{(0,1)}-8}{9t_3^{(0,1)}+8} \\
\hat{\beta} &= \tfrac{2}{3}l_2^{(0,1)}(\hat{\kappa}-2)(\hat{\kappa}-3) \\
\hat{\mu} &= l_1^{(0,1)} + \tfrac{\hat{\beta}}{(\hat{\kappa}-2)}.
\end{aligned}$$

If μ is known, we get (analogously to Hosking (2007))

$$\begin{aligned}
\hat{\kappa} &= -\tfrac{3}{2}\tfrac{l_1^{(0,1)}-\mu}{l_2^{(0,1)}} + 3 \\
\hat{\beta} &= -\left(l_1^{(0,1)} - \mu\right)(\hat{\kappa}-2).
\end{aligned}$$

Since the choice of the trimming factors is crucial, especially in our context with several extreme values, it is necessary to consider not only TL(0,1)-moments but also other choices for the trimming factors. As mentioned above, a trimming in the lower part of the sample is not meaningful. However, we consider a higher trimming in the upper part of the sample, that is the TL(0,2)-moments. Analogously to the TL(0,1)-moments we can calculate

$$\begin{aligned}\lambda_1^{(0,2)} &= 3\int_0^1 x(F)(1-F)^2 dF\\ \lambda_2^{(0,2)} &= \tfrac{1}{2}\int_0^1 x(F)(1-F)^2(12F-4(1-F))dF\\ \lambda_3^{(0,2)} &= \tfrac{1}{3}\int_0^1 x(F)(1-F)^2(30F^2-40F(1-F)+5(1-F)^2)dF\end{aligned}$$

and

$$\begin{aligned}\hat{\kappa} &= \tfrac{30t_3^{(0,2)}-5}{6t_3^{(0,2)}+5}\\ \hat{\beta} &= \tfrac{1}{2}l_2^{(0,2)}(\hat{\kappa}-3)(\hat{\kappa}-4)\\ \hat{\mu} &= l_1^{(0,2)} + \tfrac{\hat{\beta}}{(\hat{\kappa}-3)}.\end{aligned}$$

Or, if μ is known,

$$\begin{aligned}\hat{\kappa} &= -\tfrac{1}{2}\tfrac{l_1^{(0,2)}-\mu}{l_2^{(0,2)}} + 4\\ \hat{\beta} &= -\left(l_1^{(0,2)} - \mu\right)(\hat{\kappa}-3).\end{aligned}$$

Generalized Extreme Value Distribution (GEV)
The distribution function of the GEV is given by

$$G(x) = \exp\left(-\left(1+\xi\frac{x-\mu}{\sigma}\right)^{-\frac{1}{\xi}}\right)$$

for $1+\xi(x-\mu)/\sigma > 0$, where $\xi \in \mathbb{R}$ is the shape parameter, $\sigma > 0$ the scale parameter and $\mu \in \mathbb{R}$ the location parameter. The special case $\xi = 0$ is called the Gumbel distribution and is given by

$$G(x) = \exp\left(-\exp\left(-\frac{x-\mu}{\sigma}\right)\right).$$

The GEV distribution plays a crucial role in extreme value theory and therefore is one of the most commonly used distribution functions in hydrology. Because of the Fisher-Tippett-Gnedenko-Theorem (Fisher and Tippett (1928)) the GEV has theoretical validity whenever using block maxima and under certain assumptions. We want to describe this in detail.

Theorem 6.3 (Fisher-Tippett-Gnedenko)**.** *Let $X_1, \ldots, X_n$ be independent and identically distributed random variables. Suppose there exists a sequence of constants $a_n > 0$ and $b_n \in \mathbb{R}$ such that*

$$\frac{\max(X_1,\ldots,X_n) - b_n}{a_n}$$

has a nondegenerate limit distribution for $n \to \infty$. Then this limit distribution is the extreme value distribution $G_\xi(ax+b)$ with $a > 0$ and $b \in \mathbb{R}$ given by

$$G_\xi(x) = \exp\left(-(1+\xi x)^{-\frac{1}{\xi}}\right),$$

for $1+\xi x > 0$ with $\xi \in \mathbb{R}$. If $\xi = 0$, the right-hand side can be interpreted as $\exp(-\exp^{-x})$.

It can be seen that the distribution given in this theorem is exactly the GEV distribution. The parameter ξ is called the extreme value index and defines the distribution type as well as the tail-behaviour. The most important case in flood statistics is $\xi > 0$, leading to an unbounded right part of the distribution and a heavy right tail.
For the Gumbel- and GEV distribution one can find approximative expressions of the parameter estimations, see Elamir and Seheult (2003) and Lilienthal (2013).
For the parameters of the GEV distribution the TL(0,1)-moment estimators are

$$z = \frac{10}{9}\left(\frac{1}{2 + l_3^{(0,1)}/l_2^{(0,1)}}\right) - \frac{2\log(2) - \log(3)}{3\log(3) - 2\log(4)}$$

$$\hat{\xi}_{TL(0,1)} = 8.567394 \cdot z - 0.675969 \cdot z^2$$

$$\hat{\sigma}_{TL(0,1)} = \frac{2}{3} l_2^{(0,1)} \frac{1}{\Gamma(\hat{\xi}_{TL(0,1)})} \left(\left(\frac{1}{3}\right)^{\hat{\xi}_{TL(0,1)}} - 2\left(\frac{1}{2}\right)^{\hat{\xi}_{TL(0,1)}} + 1\right)^{-1}$$

$$\hat{\mu}_{TL(0,1)} = l_1^{(0,1)} - \frac{\hat{\sigma}_{TL(0,1)}}{\hat{\xi}_{TL(0,1)}} - \hat{\sigma}_{TL(0,1)}\Gamma(\hat{\xi}_{TL(0,1)})\left(\left(\frac{1}{2}\right)^{\hat{\xi}_{TL(0,1)}} - 2\right),$$

where Γ is the Gamma-function, and for the Gumbel distribution

$$\hat{\sigma}_{\mathrm{TL}(0,1)} = \frac{l_2^{(0,1)}}{0.431}$$

$$\hat{\mu}_{\mathrm{TL}(0,1)} = l_1^{(0,1)} + 0.116\hat{\sigma}_{\mathrm{TL}(0,1)}.$$

For symmetric trimming, that can be sensible because of a possibly unbounded left tail or extremely small values, the TL(1,1)-moment estimators are

$$z = \frac{9}{20}\left(\frac{l_3^{(1,1)}}{l_2^{(1,1)}}\right) + \frac{\log(3) - 2\log(4) + \log(5)}{\log(2) - 2\log(3) + \log(4)}$$

$$\hat{\xi}_{\mathrm{TL}(1,1)} = 25.31711 \cdot z - 91.5507 \cdot z^2 + 110.0626 \cdot z^3 - 46.5518 \cdot z^4$$

$$\hat{\sigma}_{\mathrm{TL}(1,1)} = l_2^{(1,1)} \frac{1}{\Gamma(\hat{\xi}_{\mathrm{TL}(1,1)})} \frac{1}{3\left(\frac{1}{2}\right)^{\hat{\xi}_{\mathrm{TL}(1,1)}} - 6\left(\frac{1}{3}\right)^{\hat{\xi}_{\mathrm{TL}(1,1)}} + 3\left(\frac{1}{4}\right)^{\hat{\xi}_{\mathrm{TL}(1,1)}}}$$

$$\hat{\mu}_{\mathrm{TL}(1,1)} = l_1^{(1,1)} - \frac{\hat{\sigma}_{\mathrm{TL}(1,1)}}{\hat{\xi}_{\mathrm{TL}(1,1)}} - \hat{\sigma}_{\mathrm{TL}(1,1)}\Gamma(\hat{\xi}_{\mathrm{TL}(1,1)})\left(-3\left(\frac{1}{2}\right)^{\hat{\xi}_{\mathrm{TL}(1,1)}} + 2\left(\frac{1}{3}\right)^{\hat{\xi}_{\mathrm{TL}(1,1)}}\right)$$

for the GEV distribution and

$$\hat{\sigma}_{\mathrm{TL}(1,1)} = \frac{l_2^{(1,1)}}{0.353}$$

$$\hat{\mu}_{\mathrm{TL}(1,1)} = l_1^{(1,1)} - 0.459\hat{\sigma}_{\mathrm{TL}(1,1)}$$

for the Gumbel distribution.
These parametric estimators based on L- and TL-moments will find application in the flood statistic later on.

6.2. Simulations for Scale Estimators

In a simulation study we want to confirm the asymptotic normality of some of the estimators proposed in Section 5 under near epoch dependence.
For the following simulations we choose an EGARCH(1,1)-process with parameters $\alpha = 0.2$ and $\beta = 0.05$. The coefficients of the function f are chosen as $\theta = 0.9$ and $\lambda = 0.1$. These are common choices when simulating from an EGARCH-process. The value α describes the dependence of the variance on the former observation, whereas β describes the dependence on the former variance. Both values are relatively small indicating only a weak dependence on the former time step. What is special and corresponds to the case of a NED-process on an underlying absolutely regular process is the choice of Z_t as AR(1)-process with correlation coefficient $\rho = 0.8$, which means a rather high correlation at short time lags. Note that the assumptions of Theorem 3.1 are therefore fulfilled. We compare the three scale estimators of Section 5 with different sample lengths concerning their asymptotic normality using QQ-plots. For the estimators in Example 5.5 and 5.6 we use the special cases Q and LMS_n of the estimators. The results can be found in Figures 6.1-6.3.
For all estimators the asymptotic normality of these estimators is confirmed by the simulations, although a sample length of about $n = 1000$ is needed. The Gini's mean difference estimator proves to need the largest sample ($n = 2000$) of the three estimators to be approximated well by a normal distribution.

In a second scenario we want to increase the dependence in the EGARCH-process such that the AR(1) process as well as the EGARCH-process model very high correlation. For this we choose $\alpha = 0.8$ and $\beta = 0.1$. The results can be found in Figures 6.4-6.6. Because of the increased dependency within the EGARCH-process a larger sample size is needed to obtain a good approximation by the normal distribution. Similar to the first scenario, Gini's mean difference needs the largest sample sizes ($n = 5000$) for a good approximation, whereas the Q-estimator does not need many more data in the presence of strong dependence. The LMS_n estimator shows a good approximation already for $n = 500$, but surprisingly it gets worse for $n = 1000$. We do not have an explanation for this phenomenon but it remains for repeated simulations.

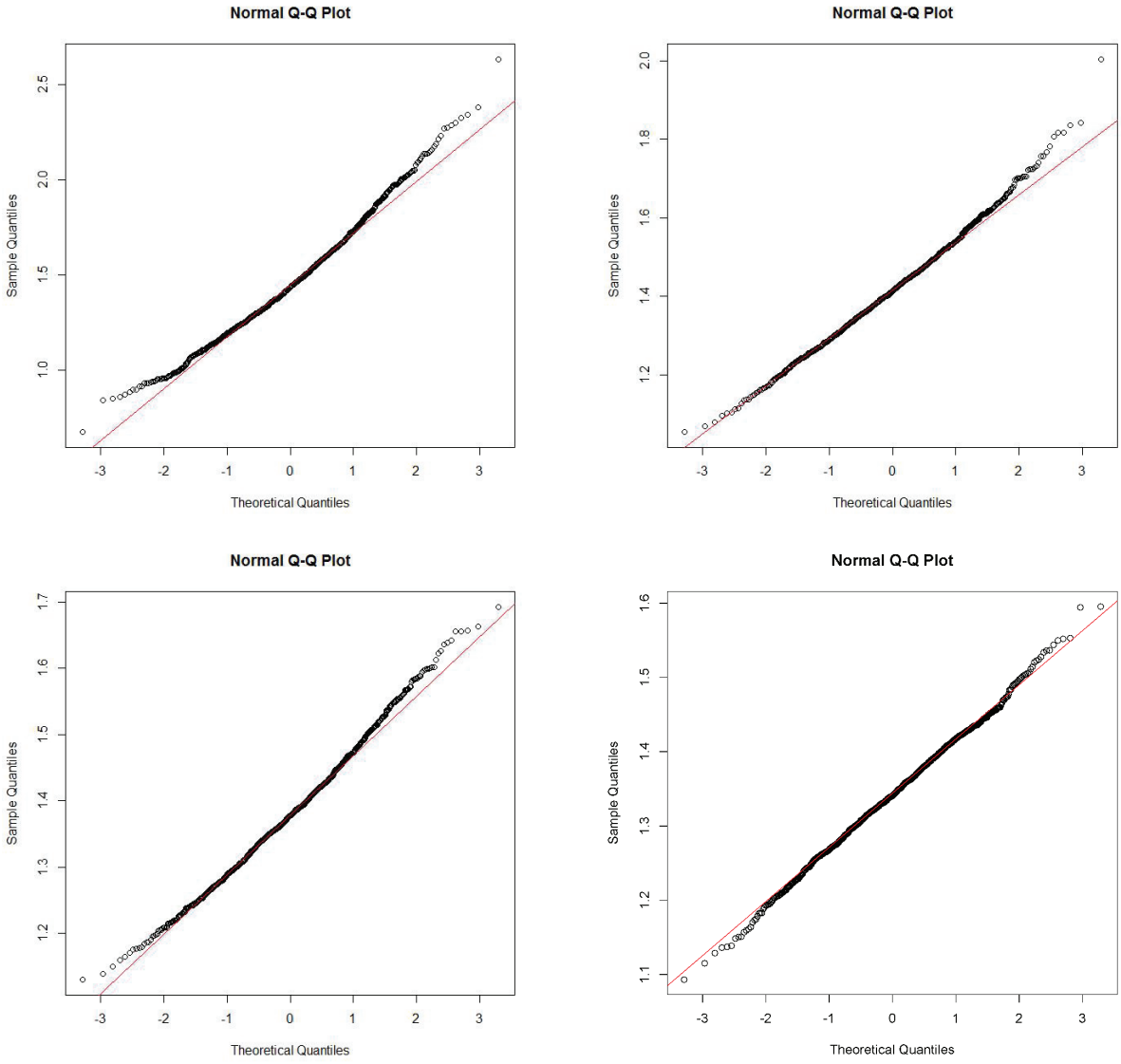

Figure 6.1.: Normal QQ-plot for Gini's Mean difference for $n = 100$ (top left), $n = 500$ (top right), $n = 1000$ (bottom left) and $n = 2000$ (bottom right) for strong dependence in the AR(1) process and moderate dependence in the EGARCH-process. From a sample size of $n = 2000$ a relatively good approximation by the normal distribution can be seen, although the upper tail is still not fitted well.

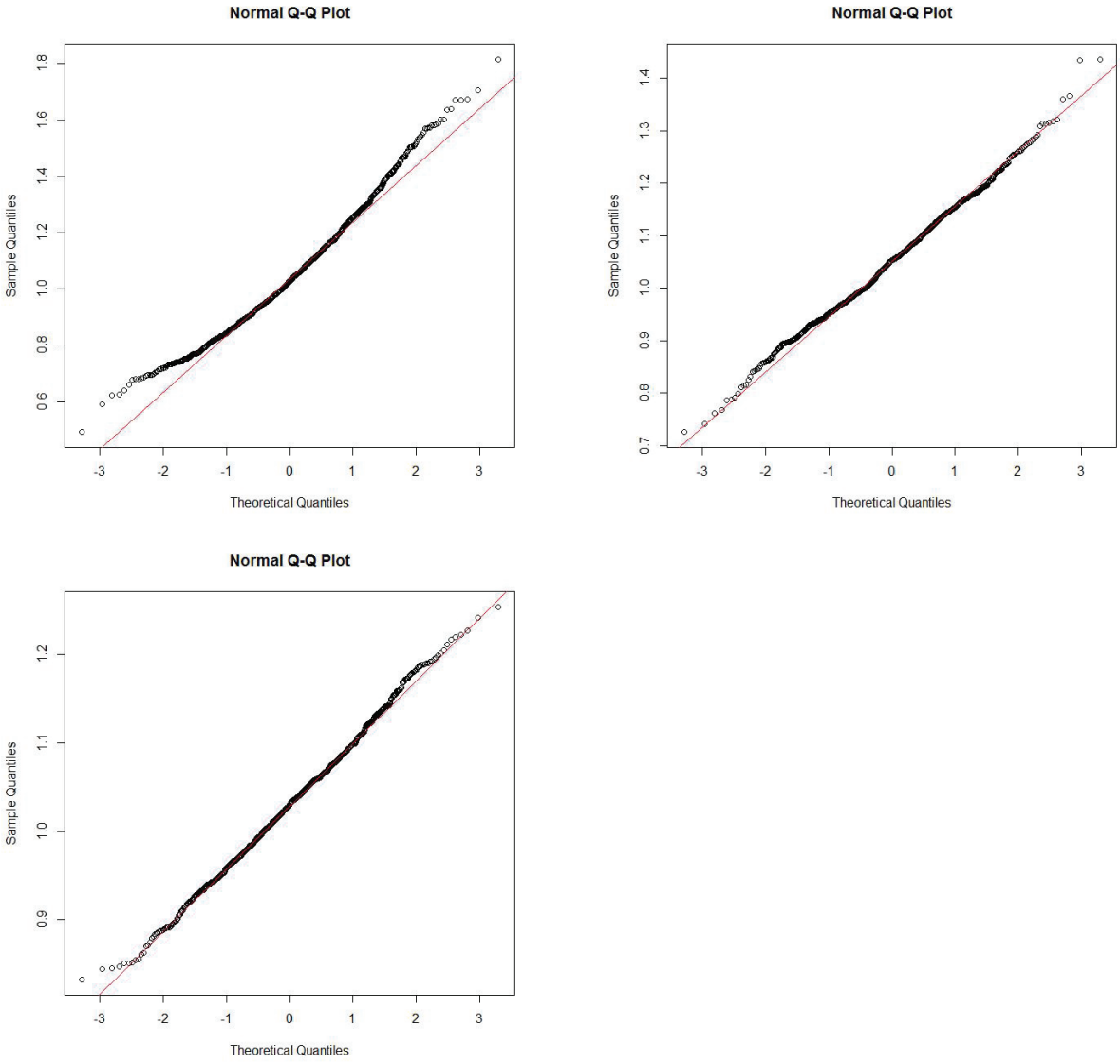

Figure 6.2.: Normal QQ-Plot for the LMS_n-estimator for $n = 100$ (top left), $n = 500$ (top right) and $n = 1000$ (bottom) for strong dependence in the AR(1) process and moderate dependence in the EGARCH-process. From a sample size of $n = 500$ a good approximation by the normal distribution can be seen as even the tails are fitted well.

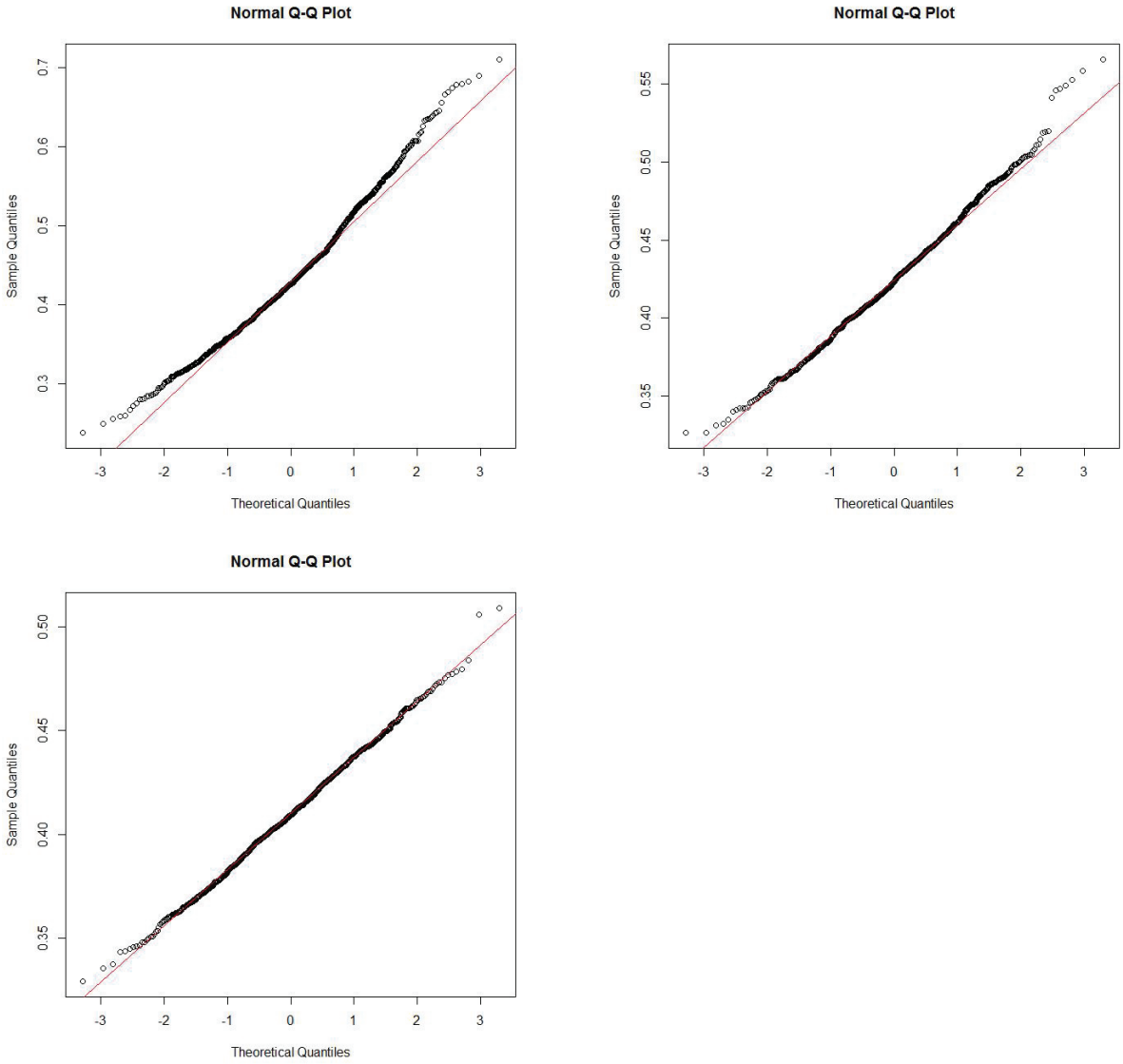

Figure 6.3.: Normal QQ-Plot for the Q-estimator for $n = 100$ (top left), $n = 500$ (top right) and $n = 1000$ (bottom) for strong dependence in the AR(1) process and moderate dependence in the EGARCH-process. For a sample size of $n = 1000$ a good approximation by the normal distribution can be seen as even the tails are fitted well.

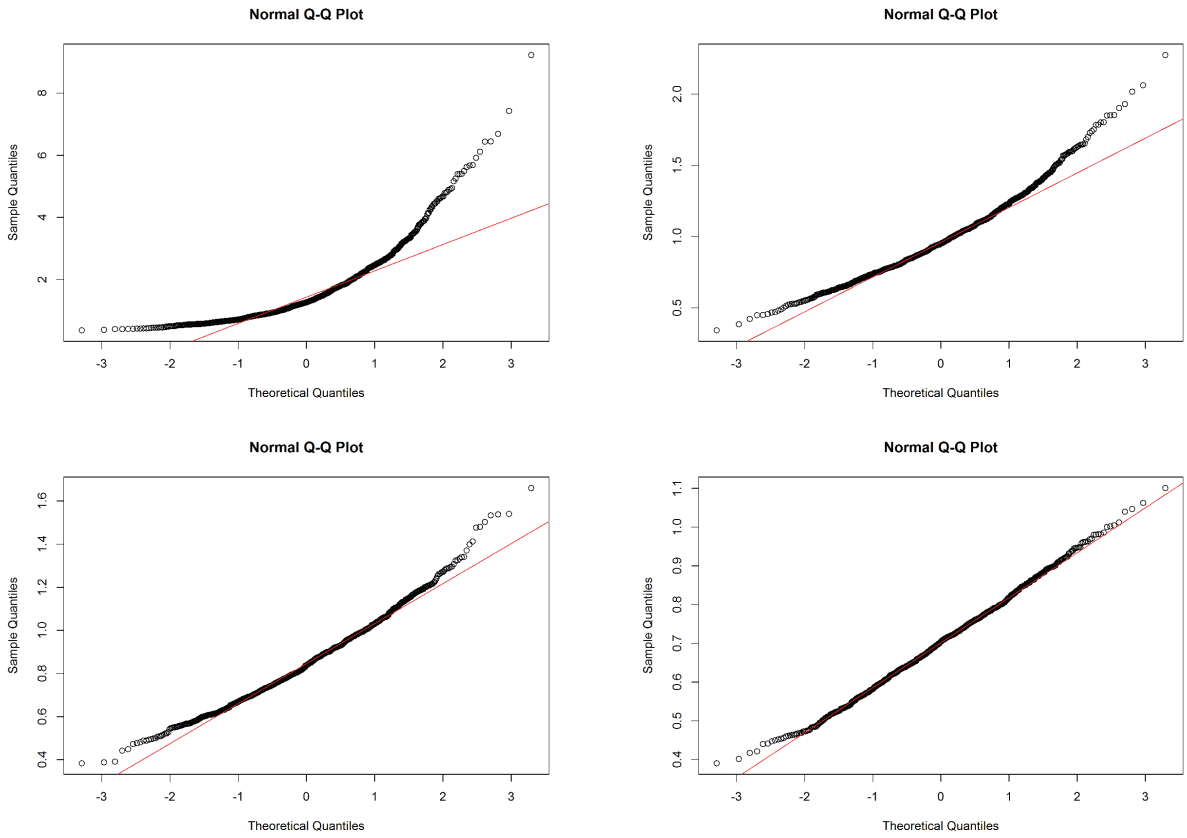

Figure 6.4.: Normal QQ-Plot for Gini's Mean difference for $n = 100$ (top left), $n = 1000$ (top right) and $n = 2000$ (bottom left) and $n = 5000$ (bottom right) for strong dependence in the AR(1) process as well as in the EGARCH-process. For a sample size of $n = 5000$ a good approximation by the normal distribution can be seen.

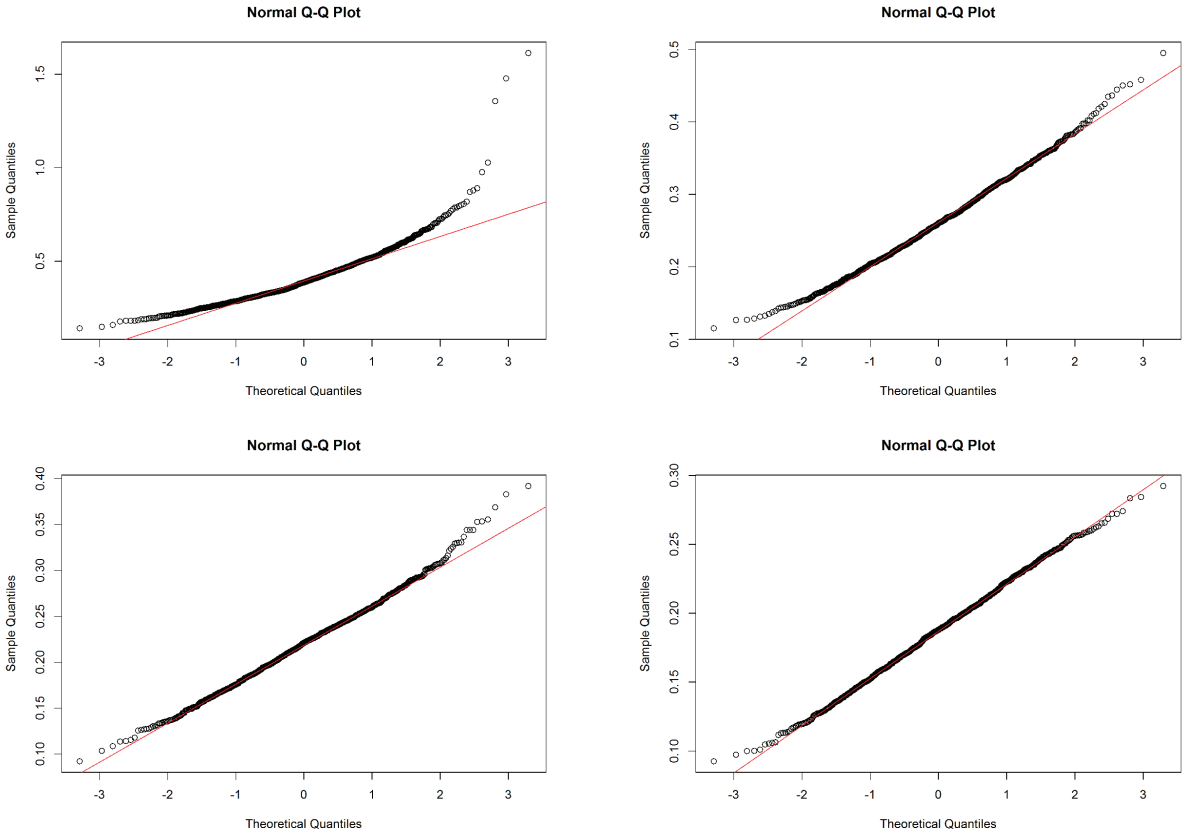

Figure 6.5.: Normal QQ-Plot for the LMS_n-estimator for $n = 100$ (top left), $n = 500$ (top right) and $n = 1000$ (bottom left) and $n = 2000$ (bottom right) for strong dependence in the AR(1) process as well as in the EGARCH-process. For a sample size of $n = 2000$ a good approximation by the normal distribution can be seen and even the tails are fitted well.

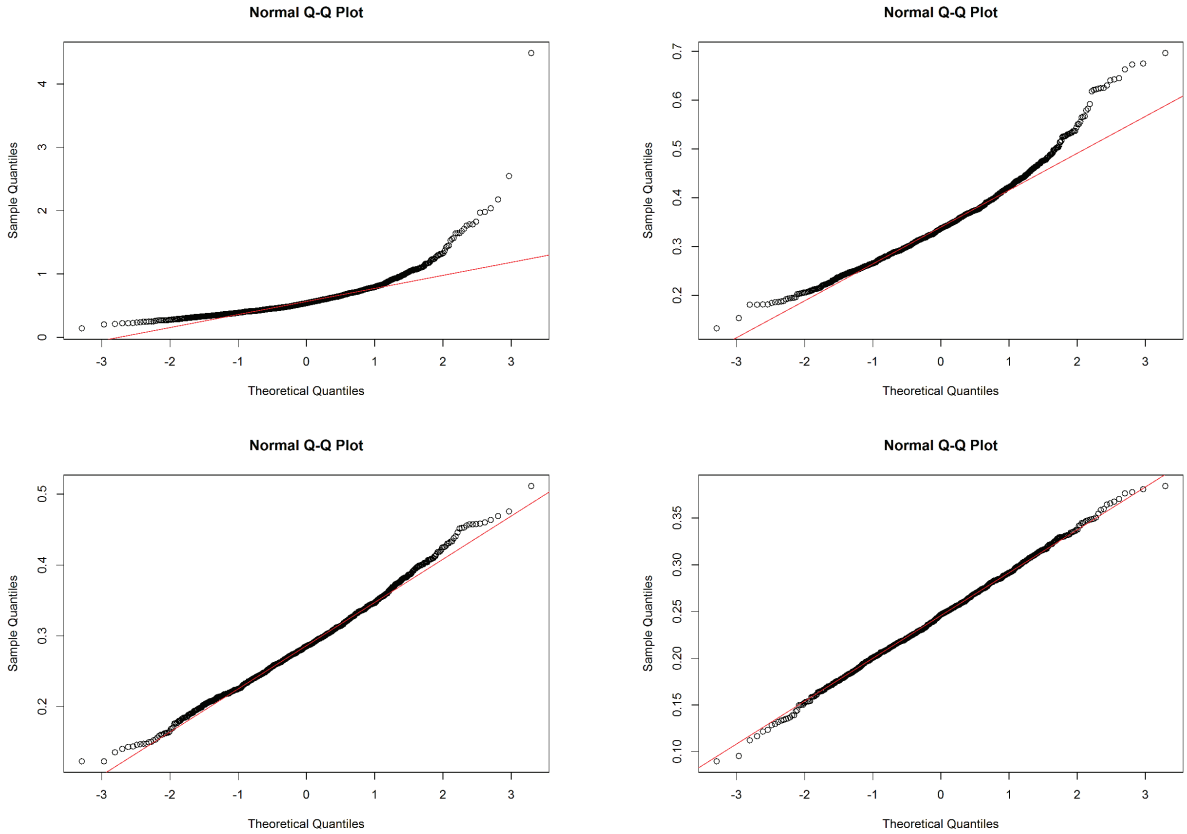

Figure 6.6.: Normal QQ-Plot for the Q-estimator for $n = 100$ (top left), $n = 500$ (top right), $n = 1000$ (bottom left) and $n = 2000$ (bottom right) for strong dependence in the AR(1) process as well as in the EGARCH-process. In case of a sample size of $n = 2000$ a good approximation by the normal distribution can be seen as even the tails are fitted well.

7. Robust Estimation in Flood Statistics

Flood statistic encompasses three essential aspects. On the one hand, design floods are needed for flood protection. When constructing a new flood protection system, e.g. a dam, a dyke or polder, the hydrological load has to be specified under the assumption that discharges exceeding this load result in critical situations or even failures. These facilities shall be secure up to certain discharges, for example floods with a return period of 1000 years. To obtain these floods from samples of not more than 100 years, statistical methods are used for extrapolation.

On the other hand, not only point-estimation (estimation at only one location) is of interest, but also the coherences between different gauges in a river basin. For example, the concurrence of two large flood events in two different tributaries downstream can lead to much higher damages downstream.

Additionally, the consideration of hydrologically similar gauged (homogeneous) catchments can enlarge the information spectrum and reduce the uncertainty in estimation due to the more detailed information on occurrence and magnitude or can be used for an estimation of flood statistics for ungauged catchments, that is a catchment where no discharge data are available. This transfer of information from homogeneous catchments to catchments where flood statistics need to be developed is called regionalisation.

The third main aspect is the modelling of discharge time series. This can be a time-dependent model as well as a model that takes into account climatic circumstances. These models then can be used to obtain information on long-term variabilities and behaviour of discharge series.

Before starting to introduce the basic statistics needed in flood statistic, we want to present the river basins studied here. In the second part of this chapter then general hydrological concepts in flood statistics are explained, that form a basis of the developed robust estimation methods, that are presented. In the third section we introduce a new classification method for flood events which is then used for regionalisation. Finally, an application of the developed Central Limit Theorem is given for an EGARCH-model used to model a daily discharge series. Parts of the results are published in Fischer et al. (2015), Fischer and Schumann (2016) and Fischer et al. (2016b).

7.1. Study Area

To apply our results concerning robust estimation in the hydrological context we consider gauges in two different regions in Germany. The Mulde river basin in the Eastern part of Germany is one of the fastest reacting river systems in Germany. The Harz region is dominated by the Brocken mountain and the rain shadow caused by it. Compared to many other river basins and regions in Germany, for these regions very detailed discharge data and information concerning land use and soil are available. Moreover, the Mulde river basin consists of many gauges with long periods of observation compared to other basins, such that the data are suitable for our purposes.

7.1.1. Mulde River Basin

We use runoff time series from 19 gauges in the Mulde river basin in the South-Eastern part of Germany. This basin almost completely drains the north side of a low mountain range, the Ore Mountains, and is part of the Elbe river basin which has been struggled by two extreme floods in the last twenty years. The Mulde basin consists of the watersheds of three main tributaries, the rivers Zwickauer Mulde, Zschopau and Freiberger Mulde, and the catchment of the Vereinigte (united) Mulde river downstream. An overview of the basin, the river network and runoff gauges is given in Figure 7.1. The catchment covers an area of approximately 7,400 km^2. The three main tributaries rise close to the mountain ridge in elevations between 760 and 1,125 m above sea level. The outlet (Bad Dueben gauge) is located at an elevation of 81 m a.s.l. The average elevation of the total catchment up to the Bad Dueben gauge is 436 m a.s.l. For each gauge catchment size and mean elevation of the catchment area are listed in Table 7.1 together with the mean annual runoff and the mean, standard variation and skewness of the annual flood series, that is the series of annual maximum flood events. All flood series have at least a length of 48 years, but most of them have been observed for more than 75 years. Like almost all parts of Eastern Germany, the Mulde river basin and its main tributary the Elbe river has been hit hard by the 2002 flood event, which has been the largest flood of the last 100 years (cover picture, DWA (2012)).

Figure 7.1.: Drainage basin of the Mulde river in the East of Saxony, Germany. In the Southern part the Ore Mountains are located.

Table 7.1.: Characteristic values and catchment area (AE) and statistical parameters for the annual maximum series (AMS) of the single gauges in the river basin Mulde.

Gauge	River	time series	AE [km^2]	mean elev. [m a.s.l.]	mean ann. runoff [m^3/s]	AMS		
						Mean	Std. dev.	Skewness
Aue 1	Schwarzwasser	1928-2013	362	737.6	6.32	69.2	51.88	2.38
Niederschlema	Zwickauer Mulde	1928-2013	759	697.4	12.64	114.0	90.91	2.80
Zwickau- Poelbitz	Zwickauer Mulde	1928-2013	1030	617.0	14.43	134.2	109.80	2.71
Harthau	Würschnitz	1965-2013	1367	438.1	1.51	32.5	24.11	2.22
Goeritzhain	Chemnitz	1910-2013	532	409.4	6.37	74.1	43.93	2.01
Wechselburg	Zwickauer Mulde	1910-2013	2107	489.1	26.31	225.5	168.56	2.79
Tannenberg	Zschopau	1960-2013	91	646.6	1.54	15.8	11.80	4.11
Streckewalde	Preßnitz	1921-2013	206	743.0	2.93	29.8	20.06	2.84
Hopfgarten	Zschopau	1911-2013	529	696.8	8.04	83.6	57.61	2.94
Rothenthal	Natzschung	1929-2013	75	752.2	1.37	15.3	11.18	3.66
Zoeblitz	Schwarze Pockau	1937-2013	129	698.7	2.28	23.6	19.60	4.99
Pockau	Flöha	1921-2013	385	679.5	5.86	68.3	43.66	2.34
Borstendorf	Flöha	1929-2013	644	653.6	9.20	95.4	69.53	3.54
Lichtenwalde	Zschopau	1910-2013	1575	612.2	21.77	227.1	164.24	3.15
Niederstriegis	Striegis	1926-2013	283	372.2	2.68	27.9	12.64	0.98
Berthelsdorf	Freiberger Mulde	1936-2013	244	594.1	3.56	36.9	43.60	5.65
Nossen	Freiberger Mulde	1926-2013	585	483.0	6.93	72.8	78.89	6.02
Erlln	Freiberger Mulde	1961-2013	2983	500.4	33.72	334.1	228.06	3.93
Golzern	Vereinigte Mulde	1911-2013	5442	477.3	62.30	520.7	335.15	3.07
Bad Dueben	Vereinigte Mulde	1961-2013	6171	436.4	64.80	513.2	394.06	2.46

7.1.2. Harz Region

The considered region of the Harz (2632 km^2) is located mainly in the German state Saxony-Anhalt and consists therefore mainly of the Eastern part of the Harz as well as a part of the Harz foreland. The Brocken mountain is the highest location in the region with 1141 m a.s.l. In the South-Western part we find foothills of the Central German Uphills, whereas the Northern and Western parts are more plain. The main rivers in this region are the Ilse, Bode, Holtemme, Selke and Wipper rivers. The Ilse river rises in the South-Eastern part of the Brocken mountain and runs to the North, whereas the Holtemme also rises in the South-Eastern part of the Brocken but runs to the North-East. The Selke river basin rises in the Brocken foreland and flows, like the Holtemme river, into the Bode river.
Caused by the altering water levels of the rivers, that are caused by the geographical as well as the climatic circumstances, 16 dams have been built in this region for water supply and flood protection, the largest one being the Rappbode-dam with 109.8 Mio. m^3 volume, lying next to the Trautenstein gauge. The foreland is dominated by agricultural use, whereas the higher parts are dominated by forests. Up to 700 m a.s.l. urbanised parts can be found mainly in the foreland.
Besides the rain shadow caused by the Brocken mountain, where only small amounts of precipitation are measured, the Harz region is known as one of the regions with highest precipitation in Germany. Quantities of more than 1400 mm can be measured in the mountainous parts, caused by high wind speeds and accumulation of rain.
Like the Mulde river basin the Harz region has been affected by the 2002 flood event, although the largest flood measured occurred for almost all catchments in 1994.

7.2. Annualities and Design Floods

Before developing robust estimation methods in context of flood statistics, we want to introduce some basic concepts of hydrology, first.
As mentioned before, in flood statistics annualities T are used to describe the return period of floods. This is the average time period in which a flood events exceeds or reaches the respective value. Therefore, for the annual maximum series it is simply a function of the non-exceedance probability $\mathbb{P}_U$:

$$T = \frac{1}{1 - \mathbb{P}_U}.$$

If we want to estimate a flood for a given annuality T, this is then simply the quantile for the respective probability $\mathbb{P}_U$.
Flood protection systems and other water engineering buildings are constructed to be secure for flood events of a certain annuality. For example, dams are designed with floods having an annuality of $T = 1000$ or $T = 10,000$ years.
The use of robust estimators in this context is expected to lead to a more stable estimation (see Section 2) of events for a given annuality. This is important, since the construction of a dam shall not depend on statistics and quantiles, whose calculation changes with every year. Moreover, neither a too large underestimation (damages) or overestimation (costs) is required. After a short introduction to the different possible models that can be used for estimating flood events, the application of robust estimators in the context of annualities and the related advantages and disadvantages are shown.

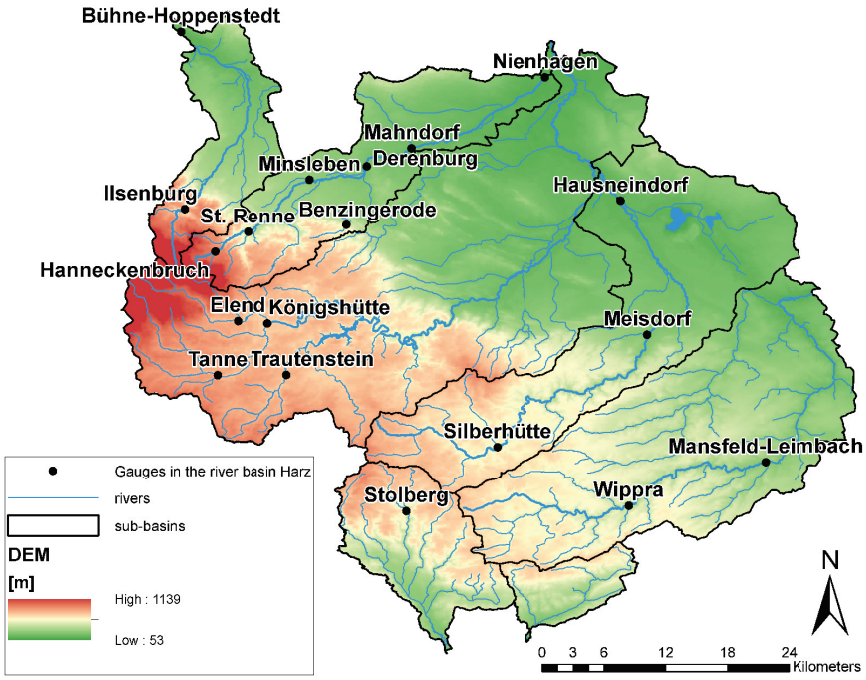

Figure 7.2.: The Harz region in the middle of Germany with the mountain Brocken located in the Eastern part.

7.2.1. How to estimate annualities

When estimating annualities the classical and probably easiest way to do this is the consideration of annual maximum discharges and the fit of a distribution to these. Nevertheless, this also limits the available information, since only one event per year can be used. Therefore, several models have been developed that enlarge the spectrum of information. Here, we want to focus on two specific models. First, the Peak-over-Threshold model that uses all discharge peaks above a given flood threshold. Second, we divide the annual series into seasons and flood types to obtain subclasses of homogeneous flood events. These models are introduced in the following, starting with the annual maximum series.

Annual maximum series

Flood statistics are often based on annual maximum series (AMS), where only the largest event in one year is taken. This series is assumed to be independent, since one cannot assume that a flood event has any influence on the flood in the following year. The estimation of distribution functions can be done with the method of moments, the maximum likelihood method or probability weighted moments (Hosking et al. (1985a)). Here, we apply these methods to adapt the GEV to flood series. In a number of European countries, the GEV is among the recommended choices of distribution functions (Salinas et al. (2014)). It is also the preferred distribution in our region of interest, Germany. Moreover, the GEV has been selected here, as the Fisher-Tippett-Gnedenko Theorem (Theorem 6.3) gives us theoretical validity for sufficiently large n. In hydrology, this limit distribution is widely applied (cf. N.E.R.C. (1975), Hosking et al. (1985a)) because of the flexibility of the distribution function with three parameters:

$$G(x) = \exp\left(-\left(1+\xi\frac{x-\mu}{\sigma}\right)^{-\frac{1}{\xi}}\right)$$

for $1+\xi(x-\mu)/\sigma > 0$, where $\xi \in \mathbb{R}$ is the shape parameter, $\sigma > 0$ the scale parameter and $\mu \in \mathbb{R}$ the location parameter (cf. Section 6.1). Most commonly, we have a positive shape parameter indicating a heavy right tail. Also the location parameter often is much larger than the scale parameter as we will see later on. To verify the Goodness of Fit of the GEV to the annual maximum series AMS we used the Anderson-Darling Test and the Akaike Information Criterion (AIC). The Anderson-Darling test (Anderson and Darling (1952)) has, in contrast to e.g. the Kolmogorov-Smirnov test, high power even when the parameters of the tested distribution are estimated from the sample and is therefore our choice in this case. The p-values of it as well as the values of the AIC applied to all considered gauges in the Mulde river basin for the GEV, the GPD (with threshold equal to the minimum annual maximum), the Gumbel (EVI) distribution and the PearsonIII distribution, which is an often used distribution function in flood statistics, are given in Table A1 in the appendix.
The distribution function of the PearsonIII distribution, which is a shifted Gamma-distribution, is given by

$$F(x) = \begin{cases} \frac{1}{\Gamma(\xi)} \int_0^{\frac{x-\mu}{\sigma}} t^{\xi-1} e^{-t} dt, & \text{for } \sigma > 0 \\ 1 - \frac{1}{\Gamma(\xi)} \int_0^{\frac{x-\mu}{\sigma}} t^{\xi-1} e^{-t} dt, & \text{for } \sigma < 0 \end{cases}$$

with parameters $\mu \in \mathbb{R}, \sigma \neq 0, \xi > 0$.

It can be seen that for the AMS the GEV is the distribution with the highest AIC values and the hypothesis of GEV-distributed data cannot be rejected. In fact, the p-values are very close to 1 for most gauges. It is not necessary to normalize the given data since asymptotic convergence to a non-degenerate random variable can be assumed. The GEV with several parameter estimation methods forms the benchmark to specify more robust estimations of flood quantiles. To differentiate from the POT approach that we apply, we use the abbreviation AMS when we fit a GEV to the annual maxima series and refer to the used estimation method when it is necessary.

Peak-over-Threshold (POT) models

In comparison to AMS, the POT method enlarges the information used for the fitting by considering not only the annual maxima but every (in our case monthly) maximum of discharge above a threshold, specifying a flood. For most gauges (including nearly all gauges in Germany) monthly maxima are the peak values observed with the smallest time-resolution. Daily maximum discharges are hardly available and would be useless in this case because of the hydrograph dynamics in a catchment.
Let us consider a data set $(X_1, \ldots, X_n)$ with $X_i = (X_i^{(1)}, \ldots, X_i^{(d)})$, that is, for example, a $n = 100$ year series of $d = 12$ monthly maximum discharges in a year. Of course, this model could also be applied with other numbers of realisations d per year. It is important to consider that not every monthly maximum corresponds to a flood peak, which has to be specified by a threshold. Since we are interested in flood statistics, we have to exclude such monthly maxima, which are not related to floods, to avoid average or low discharges influencing the flood statistics. For the threshold x_0 we use the minimum of the series of annual maxima

$$x_0 = \min_{1 \le i \le n} \left(\max_{1 \le k \le d} (X_i^{(k)}) \right)$$

to obtain a partial duration series. Also, the independence between the single flood events has to be ensured. Here we applied an approach, suggested by Malamud and Turcotte (2006), where a minimum time between two flood peaks is required to specify two independent flood events. For the watersheds in our data application with areas of several hundreds of km^2, we use a time span of at least seven days. For the application of POT we use a theorem from extreme value theory to find a suitable distribution, the Balkema-de Haan-Pickands Theorem (Balkema and de Haan (1974), Pickands (1975)). It says that the conditional excess distribution of $Y_1, \ldots, Y_n$ above a certain threshold x_0 converges to a generalized Pareto distribution (GPD) with distribution function

$$F(x) = 1 - \left(1 + \kappa \frac{x - x_0}{\beta} \right)^{-\frac{1}{\kappa}}$$

for $x > x_0$ if for the shape parameter $\kappa \ge 0$, and $x_0 \le x \le x_0 - \beta/\kappa$ otherwise, and the scale parameter $\beta > 0$ (see also Section 6.1).
For the conditional excess distribution, many other distribution functions can be used, especially the special case $\kappa = 0$ of the GPD, which is called the shifted exponential distribution

$$F(x) = 1 - \exp\left(-\frac{x - x_0}{\beta} \right),$$

for $x > x_0$. Rosbjerg et al. (1992) have shown that this distribution is preferable for modelling the exceedances in the case where $\kappa < 0.1$, as it gives a better approximation to the data. The

application of the statistical test to test the hypothesis of $\kappa = 0$ proposed by Hosking et al. (1985a) as well as calculation of the MSE show that for $\kappa < 0.1$ the exponential distribution is preferable to the correct GPD. Nevertheless, to ensure a high degree of flexibility and since most of the gauges have a parameter $\kappa > 0.1$, we use the GPD, which can be reduced to the special case of the exponential distribution.

Here, the threshold parameter x_0 is given by definition and does not need to be estimated. However, it is important to mention that the choice of it plays a crucial role in the behaviour of the estimates (cf. Begueria (2005)). As we are interested in annual return periods, one problem remains. We have to transform the results we get from the GPD, the distribution of the magnitude of excesses, into annual return periods. The relationship between annual maxima and the partial duration series can be specified as follows, where we denote the annual number of exceedances of x_0 by the monthly maxima with l. Using the total probability theorem, we obtain for the distribution function of the annual maxima F_a (cf. Shane and Lynn (1964))

$$F_a(x) = \sum_{k=0}^{\infty} \mathbb{P}(l = k) \left(F(x)\right)^k,$$

where $\mathbb{P}(l = k)$ is the probability that the annual number of exceedances of x_0 equals k. The most popular discrete distribution for describing the occurrence of rare events is the Poisson distribution with

$$\mathbb{P}_P(l = k) = \frac{\lambda^k}{k!} e^{-\lambda} \tag{7.1}$$

based on the assumption that the underlying process is a Poisson process. It seems, therefore, natural to use this distribution also in the case of the annual number of exceedances, and actually this is done in most cases (e.g. Cunnane (1973), Rosbjerg (1985), Stedinger et al. (1993)). The parameter λ represents both the mean and the variance of the distribution and is estimated by the mean of the annual number of exceedances

$$\hat{\lambda} = \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{d} \mathbb{1}_{[X_i^{(j)} > x_0]}.$$

Nevertheless, the application of the Poisson distribution also has some disadvantages in the case of partial duration series. One important point is that equation (7.1) specifies probability mass for every $k = 0, 1, \ldots$, even for $k > d$. Having in mind the example of $d = 12$ monthly maxima every year it is not possible that x_0 would be exceeded more than 12 times per year. Another problem, which has been discussed in the literature, is the assumption of equal mean and variance, which is the case for the Poisson distribution, but does not always hold (Taesombut and Yevjevich (1978), Cunnane (1979), Önöz and Bayazit (2001)). For this reason we consider different distributions and compare the results. Mathematically, the binomial distribution is applicable since it is a typical distribution for describing the number of exceedances of a threshold in a sample (Önöz and Bayazit (2001)). Its probability mass function is given by

$$\mathbb{P}_B(l = k) = \binom{r}{k} p^k (1 - p)^{r-k},$$

where p is the probability of an exceedance of the threshold and is estimated as

$$\hat{p} = \frac{1}{nd} \sum_{i=1}^{n} \sum_{j=1}^{d} \mathbb{1}_{[X_i^{(d)} > x_0]}.$$

By choosing $r = d$ we avoid the problem of having probability mass for $k > d$ since then the Binomial distribution function equals 0. Note that the Binomial distribution converges to the Poisson distribution for $r \to \infty$ and $p \to 0$ with $rp \to \lambda$.
As the third and last distribution, we use a distribution proposed by Gumbel and Schelling (1950)

$$\mathbb{P}_G(l = k) = \frac{\binom{nd}{m}m\binom{d}{k}}{d(1+n)\binom{d(n+1)-1}{m+k-1}},$$

where m is the rank of x_0 in the sample $(X_1, \ldots, X_n)$ (all monthly maximum values of all years) in decreasing order. If $n, d \to \infty$, the ratio m/n remains constant (m being approximately the median of the sample) and $n/d \to 1$, Gumbel and Schelling (1950) have shown that this distribution converges asymptotically to a normal distribution, which is a case of the central limit theorem. Otherwise, if $n, d \to \infty$ and m and k remain small, it converges to the Poisson distribution. This distribution has the advantage that we do not have to estimate any parameters and therefore have no uncertainties at this point. The only assumption needed is continuity of the data. The differences among these three distributions are shown by the example of a German flood series (Figure 7.3). The Poisson distribution has a broader but flatter behaviour, with probability mass even for $k > 12$, that is, an exceedance of monthly maximum discharges of a threshold more than 12 times a year. The Binomial and Gumbel-Schelling distributions have a similar behaviour, although the Gumbel-Schelling distribution has a larger skewness. The influence of the different shapes has been examined by comparing the three different POT models (Poisson, Binomial and Gumbel-Schelling) via their quantiles. The results are similar to those of Önöz and Bayazit (2001). The influence of the chosen weighting distribution is negligible, regardless of the estimated quantile. A possible reason for the similarity of all results could be that the Poisson distribution is the limit distribution of both the Binomial and the Gumbel-Schelling distributions. Thus, in the following we will use the Poisson distribution since the model is much easier to apply than the others. Combining the Poisson distribution with the GPD, we obtain the following distribution function of annual maxima

$$\begin{aligned} F_{a,P}(x) = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} \left(1 - \left(1 + \kappa \tfrac{x-x_0}{\beta}\right)^{-\frac{1}{\kappa}}\right)^k &= e^{-\lambda} \exp\left(\lambda \left(1 - \left(1 + \kappa \tfrac{x-x_0}{\beta}\right)^{-\frac{1}{\kappa}}\right)\right) \\ &= \exp\left(-\lambda \left(1 + \kappa \tfrac{x-x_0}{\beta}\right)^{-\frac{1}{\kappa}}\right). \end{aligned}$$

This is the GEV distribution with parameters $\xi = \kappa$, $\sigma = \beta\lambda^\kappa$ and $\mu = x_0 - \beta(1 - \lambda^\kappa)/\kappa$.

Saesonal Differentiation

A precondition of extreme value statistics is the assumption that the observations are homogeneous, i.e. subject to a common set of forces (Gumbel (1941)). The assumption of an "independent identically distributed random variable" has been critically discussed by several authors (e.g. Klemeš (2000)). One argument against this assumption are the different origins of floods. A flood peak with a certain size can result from several meteorological conditions and different combinations of hydrological processes. Often different flood types (e.g. long-rain floods, short-rain floods, flash floods, rain-on-snow floods and snowmelt floods) are mixed within a single series of annual maxima. In many cases individual types dominate the regional flood conditions (Merz and Blöschl (2003)), but the question remains open if rare and extraordinary extreme events belong into the same category as the majority of small and average floods. An attempt to consider this heterogeneity of flood types is seasonal statistics. As some flood types occur

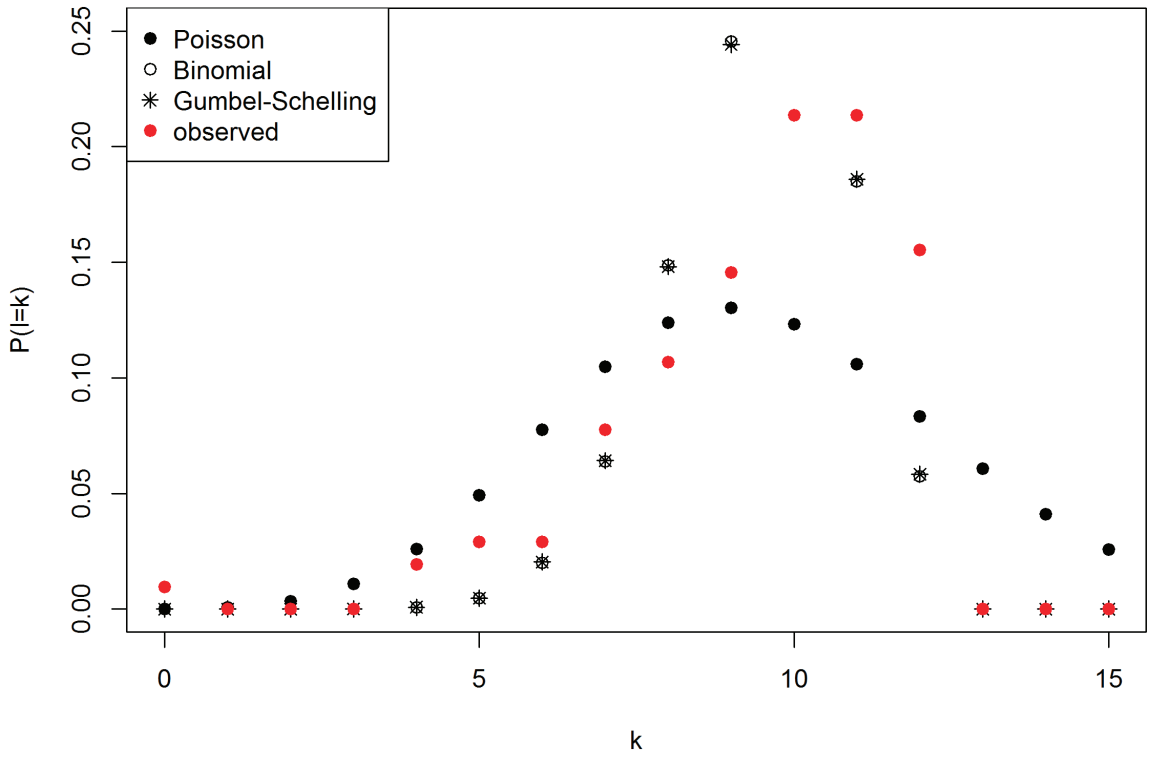

Figure 7.3.: Differences between the discrete distribution functions of the number of exceedances k of the threshold (minimum annual maximum) in a year for the Goeritzhain/Chemnitz gauge. The Binomial and Gumbel-Schelling distributions fit much better to the peak of the observed probabilities and also have zero probability for k-values larger than 12.

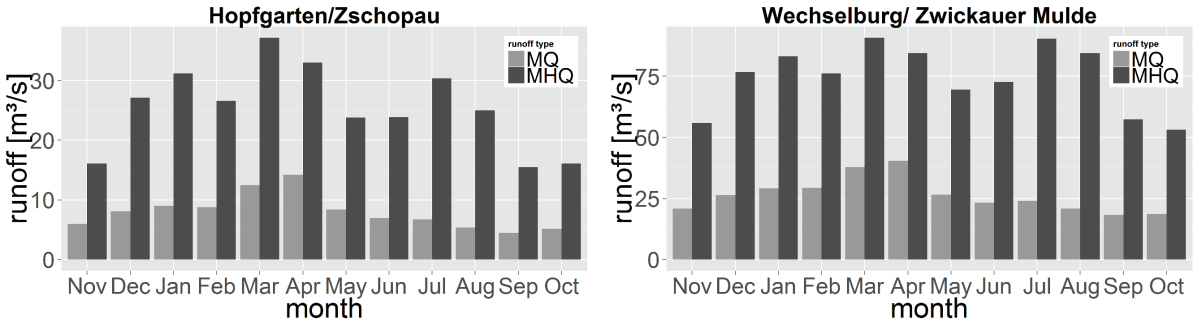

Figure 7.4.: Monthly means of discharge averages (MQ) and maxima (MHQ), derived from long series (years 1911 to 2010) for two gauges in the Mulde basin. A seasonality can be found with the values belonging to the months March and April, indicating a strong influence of snow melt caused by the close location to the Ore mountains.

only in a specific season of the year (e.g. snowmelt induced floods in winter but short-rain floods mainly in summer) seasonal flood statistics are an option to increase the homogeneity within the split samples. A subdivision of the year into two periods (winter and summer) has been proposed e.g. by Waylen and Woo (1982). The two fitted distribution functions (one for rainfall generated floods and the other for snowmelt floods) have been compounded. The basic idea of such an approach is based on the assumption that floods appearing in one season are identically distributed, which is essential in standard statistical inference (Allamano et al. (2011)). Several methods for depicting the flood seasonality have been developed. By means of directional statistics, Bayliss and Jones (1993) specified the annual mean day of flood (directional mean) and a flood variability measure. The last one is a measure of seasonality, describing how much annual maxima are concentrated within a period of the year. There are several options to subdivide a year into seasons. In some cases, e.g. in reservoir management, practical needs determine arbitrary selections, e.g. to specify a varying flood storage capacity within a year. Statistic based subdivisions have been proposed using directional mean and variance measures (Burn (1997)) of annual floods or relative frequencies of flood occurrences in months (Cunderlik et al. (2004)). Here, we want to introduce a seasonal model first, where the seasons are simply determined by the occurrence in the year. Nevertheless, it is applicable also for other distinction of subsets. The results presented here and in the following two sections are published in Fischer et al. (2016b).

The climatic seasonality of the runoff and flood regimes is shown in Figure 7.4 for two catchments. By the long-term averages of monthly maxima it becomes evident how the seasonal distribution of floods differs from the general seasonal runoff regime. The seasonal distribution of runoff has its maximum in late spring and its minimum in autumn. The flood peak averages are also high in spring (resulting from snowmelt and wet soil conditions) but there is often a maximum also in summer (July/August). All annual flood series contain winter and summer floods.

If we accept that flood series are heterogeneous and several distributions need to be combined, the question is how the yearly probabilities of a flood peak can be estimated. In the field of mathematical statistics a mixing distribution is defined as the probability distribution of a random variable gained by a combination of other random variables. In this case the cumulative distribution function of a finite mixture model can be calculated by a convex combination of the

distribution functions of these random variables

$$F(x) = \sum_{i=1}^{k} \omega_i F_i(x), \tag{7.2}$$

where $\omega_i > 0$ are the mixture weights with $\sum_{i=1}^{k} \omega_i = 1$, F_i are the mixture distributions and k is the number of mixture components. In the seasonal hydrological setting this means that we have for every season $i = 1, \ldots, k$ one distribution function F_i that has a certain influence ω_i on the series of annual maxima. Typically, the weights ω_i would be chosen as the probability that an annual maximum origins from season i.
It is not necessary that the F_i belong to the same distribution family or have the same number of parameters. If they are continuous, the mixing distribution is also continuous and the mixing density exists. A detailed introduction to finite mixing distributions can be found for example in McLachlan and Peel (2004). Rossi et al. (1984) have proposed a two component mixing distribution based on two different Gumbel distributions, which combines the different flood causes additively. For this model five parameters have to be assessed. If the distribution of the components is not known, it is rather difficult to estimate the mixture model. At least the number of mixing components should be known, otherwise an overestimation is possible (Boes (1966); Leroux (1992)).

If we consider seasonal components we can use more information about the data than the fact that they can be modelled stochastically. We can assume that summer and winter floods and therefore the random variables X_S and X_W are independent. Additionally, the annual maximum is in any case the maximum of the summer and the winter annual maximum. Vice versa, if the annual maximum X_A does not exceed some value x, then also the summer as well as the winter annual maximum must not exceed this value. Thus, the annual non-exceedance probability can be estimated multiplicatively (Todorovic and Rouselle (1971)):

$$\mathbb{P}(X_A \leq x) = \mathbb{P}(\max(X_S, X_W) \leq x) = \mathbb{P}(X_S \leq x, X_W \leq x) = \mathbb{P}(X_S \leq x)\mathbb{P}(X_W \leq x). \tag{7.3}$$

This approach has the advantage that one does not have to estimate a weighting factor (Waylen and Woo (1982)).
We applied both approaches, the statistical mixing (eq. (7.2)) and the maximum mixing (eq. (7.3)) approach, to flood data from the 19 gauges in the Mulde river basin. In the mixing approach, the seasonality of flood events has been handled in a similar model as proposed by Rossi et al. (1984). As above, the annual block maxima are modelled by a GEV distribution. Additionally, we tested the applicability of the GEV distribution to summer annual maxima, winter annual maxima and the annual maximum series using the AIC (compared to three other common distributions) and the Anderson-Darling test (Table A1 in the appendix). Although the PearsonIII distribution has lower AIC values, the p-values of the Anderson-Darling test are much higher for the GEV. For the PearsonIII distribution the p-values of the Anderson-Darling test sometimes even lead to rejection of the hypothesis ($p = 0.032$ for Zoeblitz). Additionally, the values for the AIC for GEV and PearsonIII do not differ much. Hence, for both models the marginal distributions of the summer and winter events are assumed to be GEV distributed. Therefore, three parameters have to be estimated for each season:

$$F_i(x) = \mathbb{P}(X_i \leq x) = \exp\left(-\left(1 + \xi_i \frac{x - \mu_i}{\sigma_i}\right)^{-\frac{1}{\xi_i}}\right),$$

for $1 + \xi_i(x - \mu_i)/\sigma_i > 0$ with shape parameter $\xi_i \in \mathbb{R}$, scale parameter $\sigma_i > 0$ and location parameter $\mu_i \in \mathbb{R}$ and $i = W, S$.
For parameter estimation the L-moments are used in every case. If we consider that floods do not occur in every season in every year, we have to apply censored series, where an extra parameter describes the probability that the seasonal maximum stays below a threshold x_0, which we use to censor the data. Here we replace the non-exceedance probability of the summer respectively winter events by

$$\mathbb{P}_{0,i} + (1 - \mathbb{P}_{0,i})\mathbb{P}(X_i \leq x)$$

with $\mathbb{P}_{0,i}$ being the probability that the event X_i is smaller than the threshold x_0 and $i = W, S$ indicating the summer or winter event. The question arises if this shortening of the data should be considered in the parameter estimation. Therefore, we verify in a simulation study that the shortening in the left tail has no significant influence on our studies. Since we are interested in high quantiles (90% and higher) we investigated the influence of the shortening on the estimation of these. To compare the typical hydrological setting with a more asymptotic one, samples of length $n = 100$ and $n = 1000$ are generated. The parameter vary between $\xi = \{0.1, 0.3, 0.6\}$, $\sigma = \{2, 8, 12, 20\}$ and $\mu = \{10, 30, 50, 100\}$, which represent most of the parameter sets of the considered gauges. We then calculate the quantile of the whole series and the standard error for a given annuality T

$$\begin{aligned} s_T^2 =& \left(\frac{\partial x_T}{\partial \sigma}\right)^2 \mathrm{Var}\,(\sigma) + \left(\frac{\partial x_T}{\partial \xi}\right)^2 \mathrm{Var}\,(\xi) + \left(\frac{\partial x_T}{\partial \mu}\right)^2 \mathrm{Var}\,(\mu) + 2\left(\frac{\partial x_T}{\partial \sigma}\right)\left(\frac{\partial x_T}{\partial \xi}\right) \mathrm{Cov}\,(\sigma, \xi) \\ &+ 2\left(\frac{\partial x_T}{\partial \sigma}\right)\left(\frac{\partial x_T}{\partial \mu}\right) \mathrm{Cov}\,(\sigma, \mu) + 2\left(\frac{\partial x_T}{\partial \xi}\right)\left(\frac{\partial x_T}{\partial \mu}\right) \mathrm{Cov}\,(\xi, \mu) \end{aligned}$$

with

$$\begin{aligned} \frac{\partial x_T}{\partial \xi} &= -\frac{\sigma}{\xi^2}\left[1 - \left\{-ln\left(1 - \frac{1}{T}\right)\right\}^{-\xi}\right] + \frac{\sigma}{\xi}\left[\left\{-ln\left(1 - \frac{1}{T}\right)\right\}^{-\xi} ln\left\{-ln\left(1 - \frac{1}{T}\right)\right\}\right] \\ \frac{\partial x_T}{\partial \sigma} &= -\frac{1}{\xi}\left[1 - \{-ln\,(1 - 1/T)\}^{-\xi}\right] \\ \frac{\partial x_T}{\partial \mu} &= 1 \end{aligned}$$

and

$$\begin{pmatrix} \mathrm{Var}\,(\mu) & \mathrm{Cov}\,(\mu, \sigma) & \mathrm{Cov}\,(\mu, \xi) \\ \mathrm{Cov}\,(\sigma, \mu) & \mathrm{Var}\,(\sigma) & \mathrm{Cov}\,(\sigma, \xi) \\ \mathrm{Cov}\,(\mu, \xi) & \mathrm{Cov}\,(\sigma, \xi) & \mathrm{Var}\,(\xi) \end{pmatrix} = \frac{1}{n}\begin{pmatrix} \sigma^2 w_{11} & \sigma^2 w_{12} & \sigma w_{13} \\ \sigma^2 w_{12} & \sigma^2 w_{22} & \sigma w_{23} \\ \sigma w_{13} & \sigma w_{23} & \sigma w_{33} \end{pmatrix}$$

for the L-moment estimators of the GEV (see Heo et al. (2001)), where the weights w_{ij} depend on ξ and are listed in Hosking et al. (1985a). The series is then shortened by the smallest p% values as long as the estimated quantile of the shortened series lies in the confidence interval of the quantile of the whole series. The results of the needed share of shortening to exceed the limits of the confidence interval are given in Table A2 in the appendix. As expected the removed share increases with increasing quantile and decreases with increasing sample length caused by the formula for the confidence interval that behaves alike. We can see that at least a shortening of 15% is needed to show a significant difference. For small samples like those used in

Table 7.2.: Annualities of the six largest annual maxima at the gauge Berthelsdorf/Freiberger Mulde, estimated with both mixing approaches. The extreme character of the 2002 event becomes evident as well as the differences of the tail behaviour of the approaches.

Flood peak HQ [m^3/s]	Year	Month	AMS	Statistics of summer events	Statistics of winter events	Statistical mixing approach	Maximum mixing approach
63.9	2006	March	10	13	40	24	13
68.2	1954	July	11	14	51	28	15
120	1958	July	39	35	597	97	47
122	1992	July	41	36	646	100	49
140	2013	July	55	44	1251	128	62
360	2002	August	476	187	185023	583	275

flood statistics we can remove 40% of the smallest data without a significant change. For higher quantiles this number increases. Since we never cut off more than 25% of the smallest data and have sample sizes smaller or equal to $n = 100$ the consideration of annual maxima above the flood threshold does not have a high influence on the estimation and the estimators do not have to be adjusted.

For the mixing model the distribution function of the annual maxima X_A is then as follows

$$\begin{aligned}\mathbb{P}(X_A \leq x) =& \mathbb{P}(X_A = X_S)(\mathbb{P}_{0,S} + (1 - \mathbb{P}_{0,S})\mathbb{P}(X_S \leq x)) \\ &+ \mathbb{P}(X_A = X_W)(\mathbb{P}_{0,W} + (1 - \mathbb{P}_{0,W})\mathbb{P}(X_W \leq x)).\end{aligned} \tag{7.4}$$

Since we assume $\mathbb{P}(X \in F_2) = 1 - \mathbb{P}(X \in F_1)$, the conditions on the weights are fulfilled. For the maximum mixing model we obtain

$$\mathbb{P}(X_A \leq x) = (\mathbb{P}_{0,S} + (1 - \mathbb{P}_{0,S})\mathbb{P}(X_S \leq x)) \cdot (\mathbb{P}_{0,W} + (1 - \mathbb{P}_{0,W})\mathbb{P}(X_W \leq x)). \tag{7.5}$$

The results of both models differ significantly. To give an example, the estimated annualities for the six largest annual maxima of the series 1936 to 2013 at the Berthelsdorf/Freiberger Mulde gauge are calculated with the two different approaches using only the winter and only summer series and compared with the GEV derived from the annual maximum series (Table 7.2). The parameters of the GEV are estimated with the L-moment method. Comparing the results of the mixing distributions to the estimated annualities of the series consisting only of the summer or only of the winter maxima gives an impression of how probable an event is to occur in summer or in winter. This has large influence on the mixing distribution.
Table 7.2 emphasises clear differences between summer and winter series. The observed extraordinary flood peaks in the first row occur seldom in winter time. Only one of these events with a peak of 63.9 m^3/s occurred in winter. It has a return period of 40 years in winter, but 13 years in summer. The highest flood peak (August 2002) has an extremely low exceedance probability in winter. In summer such a flood peak is much more probable. In this basin, the majority of all annual floods occurs in winter. As a result, the high return periods of large floods in the winter season affect the common distribution function (eq. (7.4) respectively (7.5)) strongly. The return periods of large floods are therefore higher in this example if we apply the statistical mixing approach (eq. (7.4)) as the low exceedance probabilities in winter have a higher impact on the results as in the maximum mixing approach (eq. (7.5)). These results make us assume

that the additional information, which has some influence on the maximum approach, captures the features of the annual maxima in a better way. Schumann (2005) already has shown that the maximum mixing approach delivers plausible exceedance probabilities even for extraordinary floods in our study region. In the following, we therefore will confine ourselves to the maximum mixing approach.

Differentiation of flood types according to timescales

A seasonal subdivision of flood series can improve the homogeneity of the resulting subsamples. However, several flood types possibly occur within these seasons. If we have in mind the large number of mid-size watersheds where rain floods can be caused by convective rain with short duration and high intensities as well as by synoptic rain with long duration, a subdivision of floods between winter and summer events does not seem to be sufficient. Therefore, a subdivision of different event types shall help to improve flood statistics by obtaining more homogeneous groups. The groups then can be combined in a new mixture model.
A new-developed concept of differentiating flood types is introduced in the following. It can also be found in Fischer et al. (2016b). For this the peak-volume relationship is used to classify the events. Then we analyse the coherences and differences of the event types and within the catchment.

A specification of flood types can be done by a number of process indicators, including the timing of the floods, storm duration, rainfall depths, snowmelt, catchment state, runoff response dynamics and spatial coherence (Merz and Blöschl (2003)). One problem of this approach consists in the availability of data and information that are needed to specify these indicators. In our case study, no information about the snow cover (height, snow density, areal extent) are available to specify the ablation process during winter floods. Only for two climate stations daily precipitation and temperature data are available. The series of these observations begin in the year 1951. The two stations are located at different elevations (1218 and 418 m a.s.l.). We have some information about the probable snow accumulation and snowmelt conditions from applications of the well-known degree-day method (Rango and Marinec (1995)), but a transfer of these results to catchments, which differ in their hypsometric conditions, would be rather uncertain. Especially the gradient of temperature with elevation remains unknown as well as the temporal changes of the degree-day ratio during a snowmelt period.

For this purpose, the event characteristic "flood timescale" (Gaál et al. (2012)) is applied to differentiate between floods of long and short duration. The flood timescale TQ (in hours) is defined as the ratio between the flood volume V (in mm) and the flood peak Q_P (in $\mathrm{mm} \cdot \mathrm{h}^{-1}$) of one event

$$TQ[h] = \frac{V}{Q_P}. \tag{7.6}$$

This ratio has almost a 1:1 relationship to the square root of the variance of the timing of runoff (i.e. the temporal dispersion of the runoff hydrograph) which has been discussed by Viglione et al. (2010a). Gaál et al. (2012) have used the timescales to specify the interplay of climatic and watersheds characteristics for flood generation in a wide range of nearly 400 watersheds that differ in size between 5 and $10,000$ km^2. The causal factors, controlling the relationship between flood peaks and volumes in a regional context, are discussed in greater detail by Gaál et al. (2015). The authors explain a weak dependence between peaks and volumes as a strong indicator for different flood types within a sample and propose a characterization of the heterogeneity of flood

generating processes by Spearman's correlation coefficient between flood peaks and volumes. A low coefficient of correlation is seen as an indicator for a large variety of flood generating processes. Here we used this assumption in the converse sense to specify flood types by estimation of linear regressions between flood peaks and volumes. However, flood timescales vary also between events of the same type of hydro-meteorological forcing. The duration of a flood is related to the duration of the meteorological input (rainfall, snowmelt) but depends also on hydrological processes within watersheds (flow-path lengths at hillslopes and in the river network). Temporal shifts between input (areal precipitation) and response (discharge at the outlet of a basin) depend on the area of the watershed (the length of flow paths) and such characteristics as hillslopes, drainage density, roughness, infiltration, soil storage and others (Viglione et al. (2010b)). Thus, flood timescales result from two groups of controls: from the type of precipitation (storm duration and spatial extent) and controls related to catchment processes (Gaál et al. (2012)). There are several empirical approaches to specify the lag times or the time of concentrations from watershed characteristics, considering mainly hydraulic relationships (e.g. Fang et al. (2008)). Besides their regional validity, the temporal characteristics of floods are also affected by the variability of hillslope processes (e.g. Robinson and Sivapalan (1997), Tucker and Bras (1998)) between flood events. Contributions of sub-surface and surface flow components may change between events according to the rainfall intensity and the initial moisture state of the watershed. The catchment specific controls disturb the relationship between hydro-meteorological drivers, flood type and timescales. E.g. the timescales of short-rain floods scatter, depending on the location of the rain cell and the remaining length of flow paths to the basin outlet. Nevertheless, the relationship between peak and volume has an explanatory power to differentiate flood events into groups.
As described above, two flood characteristics are combined in the timescale:

- the flood peak denoted by x_i
- the corresponding volume denoted by y_i.

According to eq. (7.6) the timescale TQ is the quotient of both: $TQ_i = y_i/x_i$. From a flood series consisting of n samples, (x_i, y_i), $1 \leq i \leq n$, we obtain a sample of n TQ-values. Consider that not every summer maximum of discharge values exceeds the threshold x_0 (see Section 7.2.1). Events, where the peak stays beneath this threshold, are removed from the sample.
Although timescales can be used to differentiate between events of different duration, the former distinction into seasons is crucial, since the timescales also underlie different season-depending processes. The differences in meteorological forcing and hydrological processes result in different hydrographs for summer and winter events. Snowmelt induced floods e.g. have in general larger timescales than rain floods. The seasonal variation of TQ-values becomes evident if we compare TQ-values of summer and winter maxima by directional statistics in a Burn-diagram (Burn (1997)). The day of the flood peak within the hydrological year (1 to 365.25) is transferred into the radian, the size of TQ-values is considered by the distance from the origin.
For this purpose all timescales TQ_i are normalized by the following equation

$$TQ'_i = \frac{TQ_i - \min\limits_i (TQ_i)}{\max\limits_i (TQ_i) - \min\limits_i (TQ_i)}.$$

In Figure 7.5 seasonal distributions of timescales are shown for two catchments, the mid-size catchment of the Hopfgarten gauge at the Zschopau river with a watershed area of 529 km^2 and the large-size catchment of the Wechselburg gauge at the Zwickauer Mulde river, which has a watershed of 2107 km^2. For both catchments, the highest TQ-values are concentrated in the

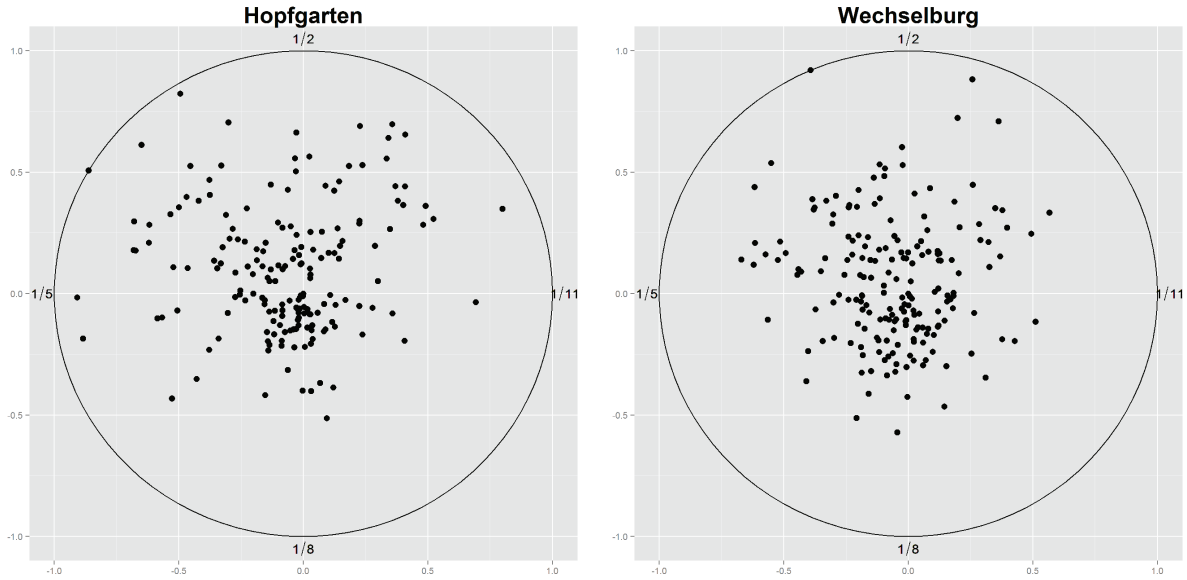

Figure 7.5.: Burn-diagram of TQ-values for two gauges within the Mulde basin, where 1/11 denotes the first of November, 1/2 the first of February and so on. A seasonal dependence of the size of the TQ-values can be seen where the largest TQ-values occur in spring.

winter half-year (floods between November and April). Also the spread of the timescales (spread of the distances from the origin) is much higher in winter, which is an indicator for a higher variety of flood processes. TQ-values in summer are much smaller, indicating that the summer floods are relatively short compared to the winter events.
In Table 7.3 the statistical values (mean, standard deviation and skewness) of timescales for winter and summer periods of all gauges are listed. It becomes evident that the means of timescales are higher in winter than in summer. However, the coefficients of variation are smaller in winter than in summer for 8 of the 14 watersheds with an area smaller than 800 km^2, while this occurs only for one catchment with an area larger than 1000 km^2.

Table 7.3.: Mean, Standard Deviation, Coefficient of Variation (CV) and Skewness of the summer and winter TQ-values

Gauge	AE [km^2]	EL [m a.s.l.]	TQ winter				TQ summer			
			Mean	Std. dev.	CV	Skewness	Mean	Std. dev.	CV	Skewness
Aue 1	362	737.6	61.3	44.52	0.73	4.79	38.3	23.83	0.62	3.86
Niederschlema	759	697.4	62.4	24.26	0.39	0.88	43.9	16.00	0.36	0.61
Zwickau-Poelbitz	1030	617.0	74.0	35.11	0.47	1.30	48.4	16.92	0.35	0.75
Harthau	135.7	438.1	44.4	20.04	0.45	0.49	23.6	10.56	0.45	0.50
Goeritzhain	532	409.5	52.5	18.12	0.34	0.40	27.7	15.07	0.54	0.89
Wechselburg	2107	489.1	74.8	32.23	0.43	0.71	52.8	22.02	0.42	0.68
Tannenberg	90.6	646.6	45.8	20.13	0.44	0.93	30.9	12.53	0.41	0.83
Streckewalde	206	743.0	65.3	42.50	0.65	2.09	32.1	18.50	0.58	1.68
Hopfgarten	529	696.8	60.3	26.19	0.43	0.45	38.8	22.12	0.57	1.61
Rothenthal	75	752.2	45.4	19.94	0.44	0.91	29.6	15.50	0.52	1.26
Zoeblitz	129	698.7	59.0	24.32	0.41	0.59	35.9	19.27	0.54	1.06
Pockau	385	679.5	55.4	25.79	0.47	1.61	36.6	22.29	0.61	2.96
Borstendorf	644	653.6	63.6	27.31	0.43	1.07	39.4	18.57	0.47	1.47
Lichtenwalde	1575	612.2	82.9	38.72	0.47	0.87	54.3	27.30	0.50	1.04
Niederstriegis	283	372.2	51.9	19.70	0.38	1.34	33.3	14.32	0.43	0.69
Berthelsdorf	244	594.1	69.0	34.69	0.50	1.39	51.7	32.65	0.63	1.38
Nossen	585	483.0	62.2	29.48	0.47	1.57	43.4	18.39	0.42	1.05
Erlln	2983	500.4	104.9	59.28	0.57	2.91	64.4	25.44	0.40	0.82
Golzern	5442	477.3	86.7	37.21	0.43	0.65	63.9	23.97	0.38	0.79
Bad Dueben	6170.8	436.4	140.6	75.18	0.53	1.27	83.4	32.98	0.40	0.65

In the next step the flood events for each catchment are grouped into different classes according to their timescales. For this purpose a statistical procedure, described below, is developed. To specify flood events by their timescales into two groups of long and short events, a threshold of the timescale TQ_0 is needed. If the value TQ_i is greater or equal than TQ_0, the corresponding event i, represented by the tuple of peak and volume (x_i, y_i), belongs to a long event, otherwise it is a short one. To find an appropriate threshold TQ_0 we use the information contained in the sample of TQ-values. We start with the order statistic of the triple (x_i, y_i, TQ_i) based on the TQ value. After sorting the events accordingly to $TQ_{(1)} \leq \ldots \leq TQ_{(n)}$ we denote the peak corresponding to the i-th order statistic of the TQ as $x_{|TQ_{(i)}}$ and the volume $y_{|TQ_{(i)}}$ likewise. It can be assumed that the relationship between peak and volume of rain-floods is linear (cf. Gaál et al. (2015)). A well-known measure for the Goodness of Fit of a linear regression $y = \beta x$ to a sample (x, y) is the coefficient of determination R^2, which is defined as

$$R^2 = \frac{\sum_{i=1} (\hat{y}_i)^2}{\sum_{i=1} (y_i)^2},$$

$\hat{y}_i$ being the regression estimate of y_i. We assume the intercept in the linear regression to be zero since a peak of zero determines automatically a volume of zero. Therefore, the coefficient of determination does not include the mean of the y_i (cf. Eisenhauer (2003)).
To specify the threshold TQ_0, the following formula is used

$$TQ_0 = TQ_{(\tau)}$$

with

$$\tau = \arg\max(R^2(1,k) + R^2(k+1,n); [n\varepsilon] \leq k \leq n - [n\varepsilon] + 1).$$

Here, $R^2(i,j)$ is the coefficient of determination of the linear regression of the sample $(x_{|TQ_{(i)}}, y_{|TQ_{(i)}}), \ldots, (x_{|TQ_{(j)}}, y_{|TQ_{(j)}})$. The parameter $\varepsilon \in (0,1)$ determines the minimum quota of data in one subsample. If the length of one subsample would be too small, the coefficient of determination falsifies the results. To ensure that the use of the coefficient of determination makes sense even for the smallest possible sample size n, $\epsilon = 0.25$ (a subsample should contain not less than 25 percent of the data) is chosen. We are aware that this coefficient of determination can be strongly influenced by single data points (outliers) which are located far away from the centre of the remaining data (concerning the value of x and y). In cases when such single events occurred in a sample, we also calculated the coefficient of determination after removing these data points. We rarely found a difference in the specification of TQ_0. Of course, it is also possible to use a robust coefficient of determination based on M-estimators proposed for example by Renaud and Victoria-Feser (2010). The developed procedure can also easily be generalized to a distinction into an arbitrary number of groups by simply choosing more thresholds and optimize R^2 for these groups. Nevertheless, the number of groups depends on the given sample size and must be known before the distinction. In our case, the sample size allows no distinction into more than three groups. Otherwise, the number of observations in one group is too small resulting in too high uncertainty. For illustration, Figure 7.6 depicts the distinction into long- and short-rain events for floods in winter and summer for the Hopfgarten/Zschopau gauge.

The thresholds of timescales can be estimated from the slopes of the regression lines and vice versa. For all gauges it becomes evident that the coefficients of determination of the linear regressions in winter are smaller than in summer. This is an indicator for a larger variety of

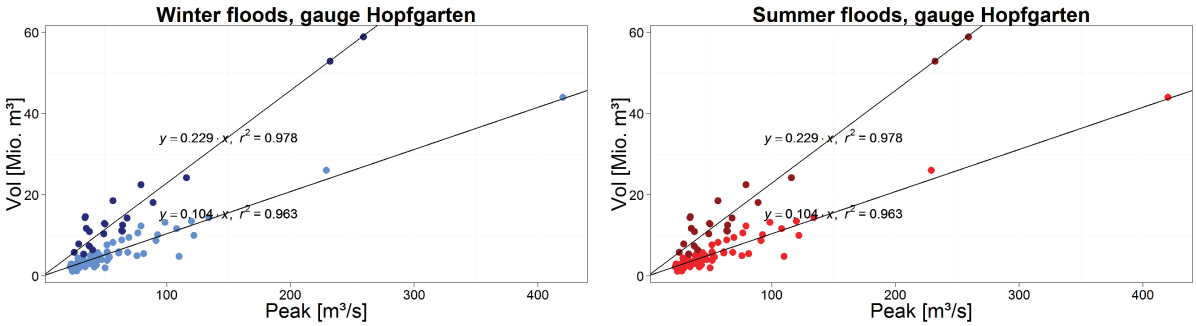

Figure 7.6.: Relationship between peak and volume and division into long and short events for winter and summer floods using linear regression for the Hopfgarten/Zschopau gauge; the upper line specifies floods with long, the lower one with short timescales.

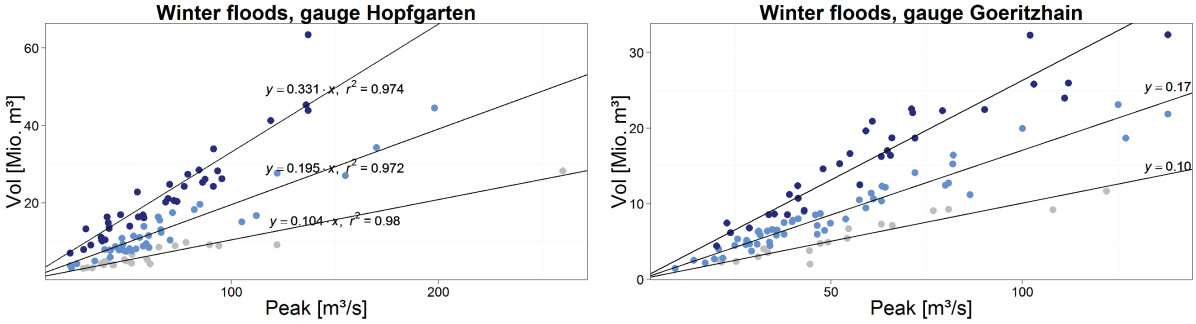

Figure 7.7.: Relationship between peak and volume and division into three groups of events for winter floods using for the gauges Hopfgarten/Zschopau and Goeritzhain/Chemnitz; the upper line specifies floods with very long, the mid line with long and the lower one with short timescales.

hydro-meteorological conditions of flood formation and process controls in winter, when the hydro-meteorological drivers of floods are rainfall and snowmelt. Also the hydrological processes differ more between events. The amount of runoff from melting snow e.g. depends on the previous accumulation of snow, the warming, the state of soil and the input of energy by parallel rainfall. The highest volumes of winter floods are estimated for rain-on-snow events in late March and April after a period of snow accumulation and (presumed) frozen soil. To consider this high variety of flood generating processes, the procedure of subdivision has been repeated for a division of winter events into three groups. In Figure 7.7 an example of the results is given for the gauges Hopfgarten (area 529 km^2, average elevation 697 m a.s.l.) and Goeritzhain/Chemnitz river (area 532 km^2, average elevation 409 m a.s.l). Both catchments have nearly the same size but differ in their elevations. Comparing the slopes of the two regression lines for large timescales in winter, the floods, originating in the catchment with the higher mean elevation (Hopfgarten), have on average 25% higher flood volumes than floods with the same peak, that originate from the lower catchment. The differences in the volumes of long and short flood events between both catchments are smaller (7% higher volume for long floods, but 12% less volume for short floods in Hopfgarten). The last effect can be explained by a higher degree of urbanization (24% in the catchment of the Goeritzhain gauge, but only 6% in Hopfgarten).
Now, the results of the subdivision are compared in more details. After subdivision of winter

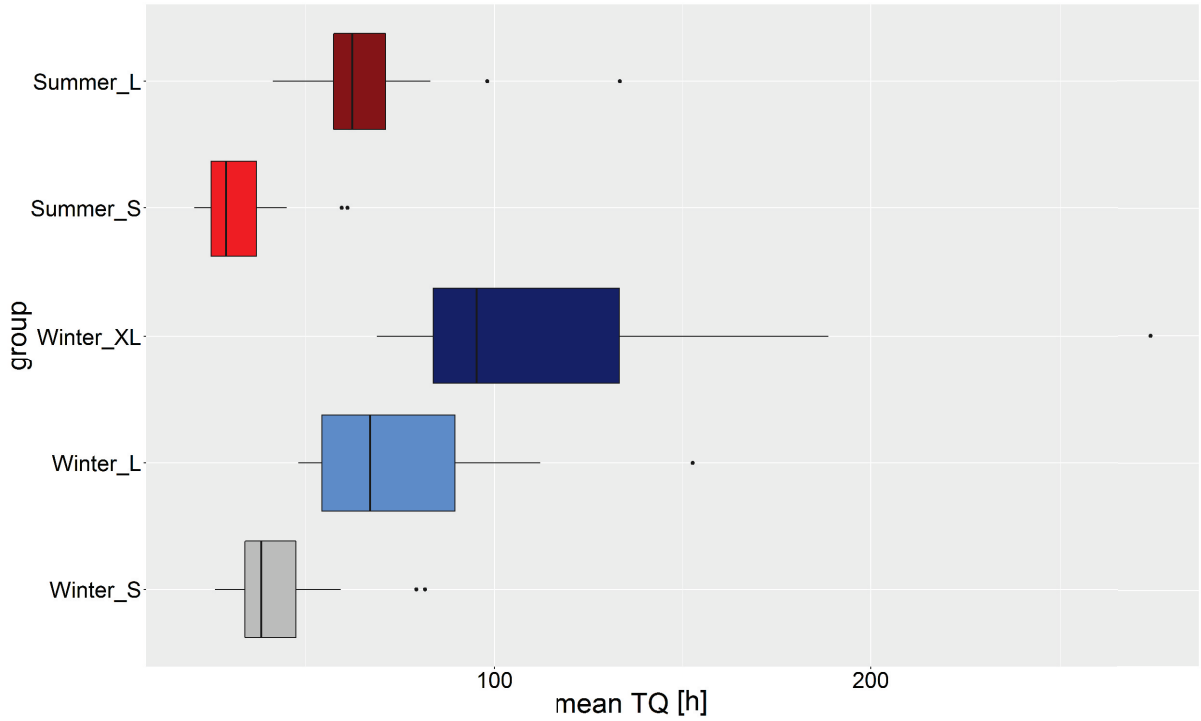

Figure 7.8.: Box-Whisker plots of the mean TQs of all 20 catchments in the Mulde basin for different flood types. A similar behaviour of short winter and short summer floods as well as long winter and long summer floods can be detected.

floods into three and of summer floods into two groups, the seasonal differences of the relationships between flood peaks and volumes are compared. In Figure 7.8 Box-Whisker-Plots of the slopes of the regressions (the mean TQ values) are shown. The large variety of slopes of events with extra-long timescales in winter (called further on as XL-events) becomes evident. It can be seen that short events in summer have a smaller slope (this means on average a smaller flood volume) than short events in winter. This can be explained by higher rainfall intensities in summer, which increase the peak and lower initial soil moisture, which reduces the volumes. The slopes of long floods in summer and winter are comparable. The application of Fisher's least significant difference (LSD) procedure (Fisher (1935)) has shown no significant differences between the means of slope of short floods in winter and summer at the 95% confidence level. Also the volume-peak-relationships for long floods do not show any significant differences in their means between winter and summer.

In Table 7.4 the coefficients of correlation between the means of TQs of all event groups are shown, derived from the sample of 13 catchments which are smaller than 800 km^2. This selection of smaller catchments is necessary to avoid the dominating impact of large catchments. The second value in each location of the table is a p-value derived from Steiger's Z-Test which tests the statistical significance of the estimated correlations (Steiger (1980)). p-values below 0.05 indicate statistical significant non-zero correlations at the 95% confidence level. All correlations among slopes in winter are statistical significant. Also the slopes of long floods in summer are statistically significantly correlated with the slopes of short, long and extra-long events in winter. In difference, short events in summer show no statistically significant correlation of their peak-volume relationships to any other group.

Table 7.4.: Correlation coefficients between the mean TQs for catchments with areas smaller than 800 km^2; the colour indicates the significance of the correlation with red being highly significant and white being not significant

	Mean TQ short events winter	Mean TQ long events winter	Mean TQ XL events winter	Mean TQ short events summer	Mean TQ long events summer
Mean TQ long events summer	0.6464 (0.0125)	0.6573 (0.0106)	0.6451(0.0127)	0.3607 (0.2052)	-
Mean TQ short events summer	0.1746 (0.5505)	-0.0025 (0.9932)	0.0620 (0.8332)	-	0.3607(0.2052)
Mean TQ XL events winter	0.7592 (0.0016)	0.8537 (0.0001)	-	0.0620 (0.8332)	0.6451(0.0127)
Mean TQ long events winter	0.9132 (0.000)	-	0.8537 (0.0001)	-0.0025 (0.9932)	0.6573(0.0106)
Mean TQ short events winter	-	0.9132 (0.000)	0.7592 (0.0016)	0.1746 (0.5505)	0.6464 (0.0125)

In contrast, for the six large catchments (areas above 800 km^2), the slopes of peak-volume relationships of short and long events in summer are correlated (correlation coefficient is 0.875; $p = 0.026$). This difference between large and small catchments is interesting. In general the differences of timescales between long and short events in summer decline with the increase of the drainage areas. To investigate the coherence of the peak-volume-relationships and the catchment size, a measure for the difference between long and short floods has to be defined. For this purpose we calculate the average distance of the long-event volumes to the peak-volume regression line of the short events. To obtain a scale-independent measure, we additionally divide it by the average volume (estimated by the median) of long events at the gauge. This yields the following formula

$$D = \frac{\frac{1}{n-r_0} \sum_{j=r_0}^{n} \left(y_{|TQ_{(j)}} - \beta_s \cdot x_{|TQ_{(j)}} \right)}{\mathrm{med}(y_{|TQ_{(j)}} | r_0 \leq j \leq n)},$$

where r_0 is the rank of TQ_0 and β_S is the regression coefficient of the linear regression of the sample of short events $(x_{|TQ_{(j)}}, y_{|TQ_{(j)}})$, $1 \leq j < r_0$. Taking the whole series, that is $r_0 = 1$, and choosing β_S as the slope of the regression of the whole series we obtain for D the normalised standard error of the regression. So D is the normalised standard error of the long events if they are modelled by a regression of the short events. The normalisation is necessary to make a comparison between different gauges (and therefore different scales of floods) possible. To extent this measure to more than two groups one could simply calculate D pairwise for two groups concerning the size of the threshold of consecutive groups. To take the Goodness of Fit of the regression into account, we multiply D with the factor $1 - R^2(1, r_0)$ and denote it by D_n. A test on correlation shows that the hypothesis of independence between D and $1 - R^2(1, r_0)$ cannot be rejected and an artificial coherence based on this factor is excluded. The relationship between D_n for summer floods and the catchment size is shown in Figure 7.9 for the 20 gauges. It becomes evident how the difference measure declines with the catchment size. There are

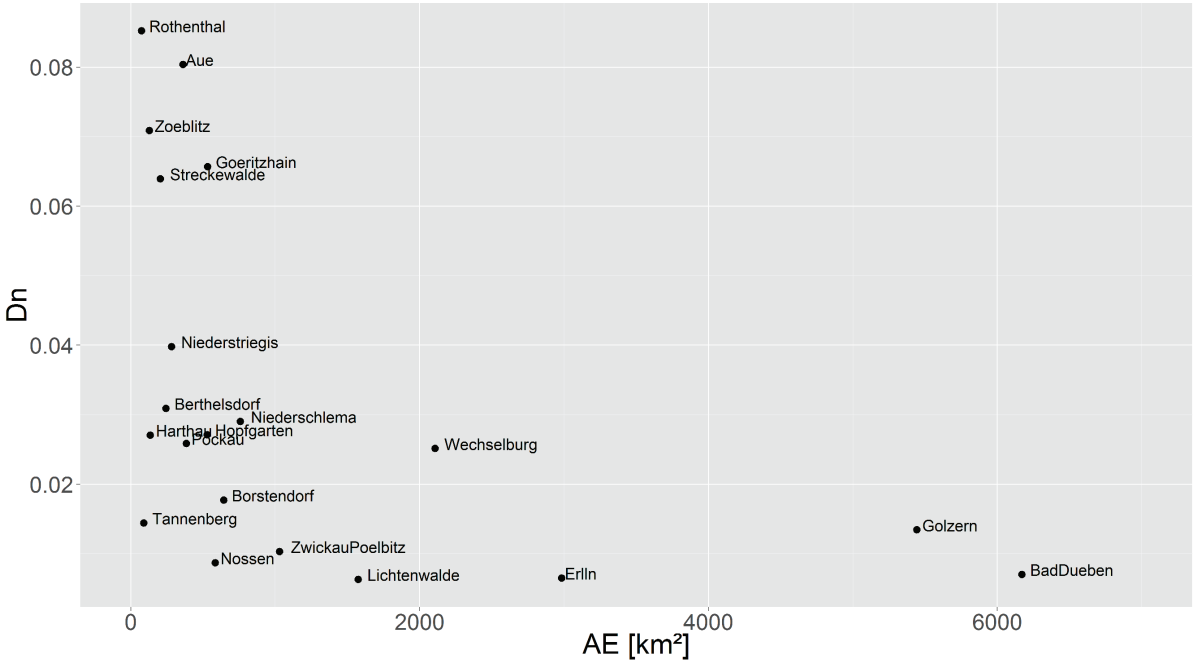

Figure 7.9.: D_n in relation to the catchment size of 20 gauges in the river basin Mulde. The difference of long and short events clearly depends on the catchment size. For large catchment areas no extreme short events with high peaks occur because these events are attenuated by the large catchment.

several hydrological explanations for this result. The impact of convective storm cells that are limited in their areal extent but result in floods with high peaks and small volumes is smaller for large watersheds, where only a part of the catchment is affected. Also the effects of runoff concentration may compensate differences of the rain durations up to a certain degree if runoff is formed only at the upper part of the catchment and the peak is alleviated by the flood wave propagation downstream. Otherwise, small watersheds react faster and high rainfall intensities result in steep hydrographs with high peaks. Hence, it takes a larger volume of direct runoff to reproduce a comparable high peak during a long rain event. The spread of the differences for small and medium sized watersheds is an indicator for the variability of hydrological processes among flood events, caused by different durations of rainfall.

Furthermore, the existence of a third group of events with extra-long timescales in winter has to be discussed. On average 25% of all winter floods belong into this category. Nearly 60% of all extra-long events happen in March and April, but less than 15% in November and December. The volumes of such extra-long floods are on average 60% higher than the volumes of floods with the same peak belonging into the category of long events. By detailed analyses of events with very high timescales using the day-degree snowmelt assessment, specific conditions of flood formation become evident. Such events result mainly from rain on snow under the following conditions: a long period of snow accumulation and frozen soils in combination with a warm front, accompanied by a rain period of several days. Such conditions occur in the Mulde basin mostly in late March or April. In contrast, long winter floods are typical snowmelt events that occur mostly in early spring (March). Here the flood volumes are smaller. This can be caused by unfrozen soils and less rainfall. The monthly frequencies of flood types of all analysed winter

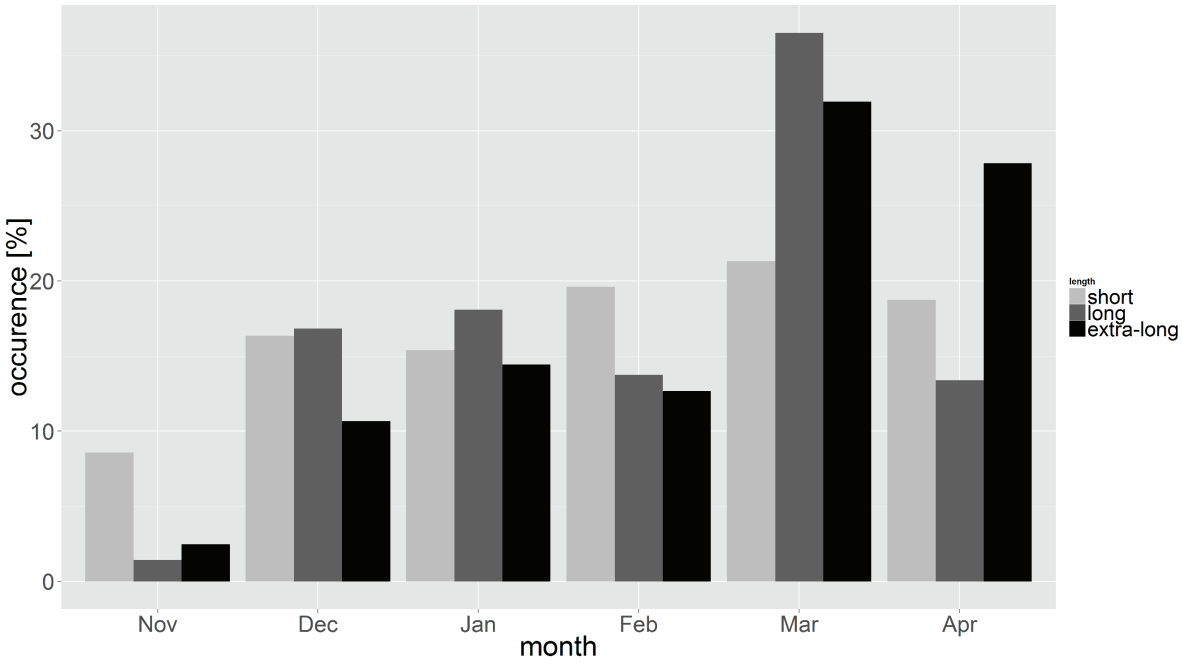

Figure 7.10.: Distribution of flood types over the months in the winter half-year of all considered winter events. The snowmelt induced extra long events mainly occur in spring whereas the short events occur during the whole winter.

floods is shown in Figure 7.10. It can be seen that short events are more equally distributed among the months in winter than long or extra-long events.

In Table 7.5 the results of analyses of timescales are summarized. The two groups of watersheds consists of 13 catchments with an area between 75 and 759 km^2 (small basins) and 6 catchments with areas from 1030 up to 6171 km^2 (large basins). The timescales of the small basins show in relation to the timescales of the large catchments a smaller dispersion. On average, long summer events have twice the volume of short events that have the same peak. For smaller basins, this relationship is a bit higher (2.24) than for large basins (2.03). In winter such differences between small and large basins practically do not exist. The relationship of the volumes of long and short events is 1.69 (small basins) or 1.7. Between long and extra-long events this relationship is 1.53 respectively 1.59. The shares of long and short events differ in summer between small and large basins more than in winter. This can be explained by higher rainfall intensities in summer, falling on a smaller area.

As a result of this event analyses, we decide to differentiate between event types only in the summer flood statistics. The existence of at least three different flood types in winter allows no sufficient specification of probability distribution functions of flood peaks separated into these types at least for our region of interest. Especially the extraordinary character of the relative small part of XL-floods (extreme long timescales) cannot be considered in a statistical mixing model (see e.g. Section 7.2.1) sufficiently. Nevertheless, if the observed runoff series is longer, also a mixing model with three groups is possible.

The results of the analyses of flood timescales for winter and summer half-years can be summarised as follows:

Table 7.5.: Means and coefficients of variation (CV) of flood timescales in hours for winter and summer floods, differentiated by event groups and small and large basins

Summer	short events			long events					
	share	mean	CV	share	mean	CV			
small basins	0.681	26.1	0.172	0.319	58.5	0.119			
large basins	0.542	33.9	0.464	0.458	68.7	0.417			
Winter	short events			long events			extra-long events		
	share	mean	CV	share	mean	CV	share	mean	CV
small basins	0.443	36.0	0.173	0.314	60.9	0.200	0.243	93.5	0.224
large basins	0.502	44.2	0.466	0.280	75.0	0.461	0.218	119.0	0.525

- Flood timescales can be applied to differentiate floods into event types. In summer we found two groups of events, which differ significantly in their peak-volume-relationships. In winter the grouping is more difficult. In our study region we have to differentiate at least between three different groups of winter floods.
- Short and long floods in winter and summer are relatively similar in their timescales. One group of events exists only in the winter half-year. These extra-long events result from specific rain on snow cases.
- For large basins, the relationships between flood volumes and peaks in summer are correlated. This is an indicator for similar underlying climate processes. This is not the case for catchments with an area size smaller than 800 km^2. Here, the slopes of the relationships are not correlated. This indicates a flood control by different types of rainfall events.

This new information on the heterogeneity of floods will be used in a new mixture model. For this, the obtained information on the different event types has to be combined in a statistical model for the distribution function. The development of such a model is given in the following section.

Considering seasonal aspects and flood types with the maximum mixing approach

In the considerations above it becomes obvious that besides the distinction between summer and winter also a distinction between long and short summer events is necessary. We have shown that the summer runoff series consists of two different types of events: short and long floods. These can be separated as described above by using the TQ-threshold TQ_0. We therefore want to extend the seasonal mixing model from Section 7.2.1 to a model which considers the different distributions of winter and summer floods as well as the differences between short and long summer events. Since we want to consider several flood types (long and short floods) in summer, we need a statistical mixing model to specify the resulting flood probabilities in summer. We are also interested in the annual flood probabilities, so we have to compound the probabilities of summer floods with the probabilities of floods in winter. Using the maximum model for both steps, eq. (7.5) has to be extended to a nested maximum approach. It would of course be possible to apply a statistical mixing distribution similar to the approach of Rossi et al. (1984) as well, but the results above make the maximum model more recommendable. For summer we have

two mixing components (floods with long or short timescales). The weights in the statistical mixing model could be chosen as the probability that the random variable belongs to the first or the second distribution, respectively. These weights are unknown. The maximum mixing model benefits from the fact that the non-exceedance of a summer flood peak X demands that it is not exceeded by any short or long flood event in this half-year. So the multiplicative approach of Todorovic and Rouselle (1971) can be applied and we have

$$\mathbb{P}(X_A \leq x) = \left(\mathbb{P}(X_{S_S} \leq x)\mathbb{P}(X_{S_L} \leq x)\right)\mathbb{P}(X_W \leq x),$$

where X_{S_L} are the maxima of the long summer runoffs and X_{S_S} are maxima of the short summer runoffs.
If a flood event does not happen every year in summer, we have to incorporate also correction factors due to the censoring of the data. Here it is advisable to correct the summer events jointly. If a summer event lies beneath the threshold, also the respective short and long summer event has to lie beneath the threshold. This modifies the model into the following representation of the distribution of the annual maximum discharges

$$\mathbb{P}(X_A \leq x) = \left(\mathbb{P}_{0,S} + (1 - \mathbb{P}_{0,S})\left(\mathbb{P}(X_{S_S} \leq x)\mathbb{P}(X_{S_L} \leq x)\right)\right)\left(\mathbb{P}_{0,W} + (1 - \mathbb{P}_{0,W})\mathbb{P}(X_W \leq x)\right).$$

Still, one problem remains. Unlike the distinction between summer and winter annual maxima, a distinction between short and long summer floods does not result in representative series of both groups of events. As only a member from one of both groups is the maximum of the summer half-year, we only know the highest summer flood per year which can be characterized by a long or by a short timescale. Smaller values are overlaid. A reconstruction of these observations is not possible since, as mentioned above, monthly maximum values are the highest time-resolution for peaks. In the observed subsamples only large values of the flood types are included since all small events are overlaid by higher events belonging to the other flood type. If we now want to fit a distribution to these subsamples, we do not have the full information included in the series but only the information belonging to the right tail. Nothing is known about the small events, in fact, simply fitting a distribution to these subsamples would totally ignore the presence of small events. This leads to a distortion of parameters, especially of location and skewness. To estimate the distributions of long and short flood events, we have to correct the statistical parameters of both samples. This is a problem already recognized by McCuen and Beighley (2003), who propose a Maximum-Likelihood based approach to estimate the parameters. Since this approach is only valid for a normal distribution and we can certainly not assume normality of the data, it is not applicable in this case. We therefore had to develop a new approach for reconstruction of the overlaid information. In the following we propose two approaches, one based on censored survival statistics and the other one being a heuristic approach. Both approaches are presented first and compared later on in this section.

We want to present the methods by using the example of summer annual maxima.
Suppose, we have a series over n years of summer annual maximum discharges $(X_{S1}, \ldots, X_{Sn})$. It is known that this time series contains short as well as long events. Each X_{Si} is therefore the maximum of both types of events in one year, such that the maximum event of the other type is overlaid as its peak is smaller. We have no possibility to obtain the overlaid events since the data basis does not contain these. By using the TQ-values we can, however, determine which of the X_{Si} belongs to which type of event. This leads to two subsamples $(X_{S_L1}, \ldots, X_{S_Ln_1})$ and $(X_{S_S1}, \ldots, X_{S_Sn_2})$ consisting of different quantities of long, respectively short events, where $n_1 + n_2 = n$. The unknown full series of long and short summer series are analogously denoted

with $(X_{S_L1}, \ldots, X_{S_Ln})$ and $(X_{S_S1}, \ldots, X_{S_Sn})$. The two different approaches to reconstruct the overlaid information are presented in the following and are compared later on.

Censored Survival Statistic

Survival data, that are data measuring the time to an event, are used in several disciplines, especially in medicine and in biology, where case studies with probates (humans or animals) are used to analyse the effect of certain diseases, medicine etc. In this context often the problem of censored series occurs, for example when a probate already died before a study started and the exact date is unknown (left-censoring) or if the time of the study expires although the desired effect did not occur (right-censoring). For more details on this topic we want to refer to Klein and Moeschberger (2003).

Since this scenario of right-censored series is very similar to our case of overlaid series of flood types it seems natural to use methods of survival analysis to estimate the parameters of the censored series. We therefore want to describe the scenario in more detail.

With the notation used above we have $X_{Si} = \max(X_{S_Si}, X_{S_Li})$, $i = 1, \ldots, n$, and define the function $\delta_i = \mathbb{1}_{[X_{S_Si} \leq X_{S_Li}]}$ to indicate whether the i-th summer annual maximum is a short or a long event. Let us denote with h_θ the probability density of (X_S, δ) with unknown parameters θ. To construct estimators for the parameters of the distribution function of the X_{S_S} (and also X_{S_L}) we want to use a Maximum-Likelihood approach. For this we can use a result given by Kremer et al. (2014) for the log-Likelihood function

$$\begin{aligned}\log L_n(\theta, X_S, \delta) &= \sum_{i=1}^{n} \log h_\theta(X_{Si}, \delta_i) \\ &= \sum_{i=1}^{n} \left(\delta_i \log f(X_{Si}) + (1-\delta_i) \log F(X_{Si}) + \delta_i \log G(X_{Si}) + (1-\delta_i) \log g(X_{Si})\right), \end{aligned} \tag{7.7}$$

where F is the distribution function of X_{S_S} with related density f and G is the distribution function of X_{S_L} with density g. Of course, the δ_i are not known but based on the distinction of the events described above. That means, actually the uncertainty of this distinction has to be taken into account. Nevertheless, we want to assume the δ_i to be known, since we can proceed on the assumption that the distinction is correct. The uncertainty in that point is therefore negligible.

We want to concentrate on estimating the distribution function of the short summer events, since the estimation of the long summer events is totally analogous and does not have to be shown in detail. Therefore, for the Maximum-Likelihood estimation only the first term of eq. (7.7) is needed. For the distribution function of the short (as well as the long) summer events we choose, motivated by the Goodness of Fit tests above, the GEV with parameters μ, σ and ξ.

Therefore, the Likelihood function can be reduced to

$$\begin{aligned}\log L_n^\star = &\sum_{i=1}^{n} \left[\delta_i \log \left(\frac{1}{\sigma} \left(1 + \xi \left(\frac{X_{Si} - \mu}{\sigma}\right)\right)^{(-\frac{1}{\xi}-1)} \exp\left(-\left(1 + \xi \left(\frac{X_{Si} - \mu}{\sigma}\right)\right)^{-\frac{1}{\xi}}\right)\right) \right. \\ &\left. + (1-\delta_i) \log \left(\exp\left(-\left(1 + \xi \left(\frac{X_{Si} - \mu}{\sigma}\right)\right)^{-\frac{1}{\xi}}\right)\right) \right] \\ = &\sum_{i=1}^{n} \left(-\delta_i \log \sigma + \delta_i (-\frac{1}{\xi} - 1) \log \left(1 + \xi \left(\frac{X_{Si} - \mu}{\sigma}\right)\right) - \left(1 + \xi \left(\frac{X_{Si} - \mu}{\sigma}\right)\right)^{-\frac{1}{\xi}} \right). \end{aligned}$$

Table 7.6.: RMSEs of the estimates of the parameters using Survival Analysis for sample sizes $n = 100$ and $n = 1000$. High values for the RMSE for small sample sizes indicate a lack of efficiency of the ML-estimator.

Number	Scenario	RMSE for $n = 100$			$n = 1000$		
		μ_{S_S}	σ_{S_S}	ξ_{S_S}	μ_{S_S}	σ_{S_S}	ξ_{S_S}
3	$\mu_{S_S} = 20$, $\sigma_{S_S} = 5$, $\xi_{S_S} = 0.4$ $\mu_{S_L} = 40$,$\sigma_{S_L} = 2$, $\xi_{S_L} = 0.1$	20.517	13.317	0.449	8.061	3.787	0.158
4	$\mu_{S_S} = 20$, $\sigma_{S_S} = 2$, $\xi_{S_S} = 0.6$ $\mu_{S_L} = 20$,$\sigma_{S_L} = 2$, $\xi_{S_L} = 0.1$	1.132	0.379	0.160	0.0896	0.0988	0.0414

Number	Scenario	RMSE for $n = 100$			$n = 1000$		
		μ_{S_L}	σ_{S_L}	ξ_{S_L}	μ_{S_L}	σ_{S_L}	ξ_{S_L}
3	$\mu_{S_S} = 20$, $\sigma_{S_S} = 5$, $\xi_{S_S} = 0.4$ $\mu_{S_L} = 40$,$\sigma_{S_L} = 2$, $\xi_{S_L} = 0.1$	0.237	0.178	0.0984	0.101	0.121	0.0104
4	$\mu_{S_S} = 20$, $\sigma_{S_S} = 2$, $\xi_{S_S} = 0.6$ $\mu_{S_L} = 20$,$\sigma_{S_L} = 2$, $\xi_{S_L} = 0.1$	0.3242	0.335	0.114	0.0876	0.104	0.0348

This function is then optimised to find the parameters.

We want to test the validity of this approach by a simulation study.

Therefore, two samples of length $n = 1000$ of long and short summer floods with parameters of the GEV-distribution μ_{S_S}, σ_{S_S}, ξ_{S_S} and μ_{S_L}, σ_{S_L}, ξ_{S_L}, respectively, are chosen. For these two samples the maximum of each data pair is used to obtain the annual summer maximum series. Applying the ML-estimation proposed above we then calculated the RMSE of the estimates of the single parameters. This has been repeated 1000 times per scenario. Two scenarios given in Table 7.6 are taken into account that mirror parameter sets assumed to be typical for this kind of hydrological time series.

Since in hydrology we have very limited sample sizes with seldom more than 100 years of observation we also want to test the applicability of the ML-estimation to small samples of size $n = 100$. Although we use a parametric estimation the problem of censoring could lead to a high uncertainty for small sample sizes, especially if one event dominates and therefore the other event type only has a small share on the summer annual maxima. Hence, the simulation done above is repeated with a sample size $n = 100$. The results can be found in Table 7.6.

Whereas the results for large samples ($n = 1000$) lead to a rather small RMSE, the RMSEs for $n = 100$ are rather high for the parameters of the short events. For the second scenario they may be tolerable in both cases and both event types. This is the scenario where both samples only differ in their shape parameter and thus have nearly equal share on the summer maxima. That is, even for a sample size of $n = 100$ we can expect about $n_1 \approx n_2 \approx 50$ observations. The real

disadvantage of the ML-estimation approach can be seen for the first scenario. Here, because of the choice of the parameters, we can expect a much higher share of the long summer events on the summer maxima. When we estimate the parameters for the short summer events, the observed number of short events is very small. In this case, the ML-estimation approach has large disadvantages, especially when considering that the first scenario is the most probable in hydrology. As for the most cases in Germany, and nearly for all of our gauges in the Mulde basin, the scale and/or shape for the short summer events is much higher than for the long summer events, whereas these have the larger or equal location parameter. This is due to the fact that the short summer floods are often very small but also consist of some single very large flood events. The long summer events are on average higher but without having these single extreme floods, see also the discussions in Section 7.2.1. Hence, this approach does not seem to be applicable in our context. This has been the reason for us to develop a new approach to fill up the overlaid information, which is presented in the following paragraph. It ensures a sufficient sample size for both subsamples and makes an estimation less uncertain. Nevertheless, if the sample size is sufficient large for both subsamples, the ML-approach is definitely worth to consider.

Heuristic approach: "The Filling method"
Analogously to the censored series used above, in this approach we use the information that the missing events are definitely smaller than the overlaying events, otherwise they would not have been overlaid. In contrast to the ML-approach now the gaps in the subsamples are filled using this information.
Let us first have a look at the mode of the GEV distribution, that is the value, where the density and therefore the probability mass function takes the highest value. For a GEV(μ, σ, ξ) distribution with $\xi > 0$ (as is typical for discharge series, corresponding to a bounded left tail and therefore only positive discharges and a heavy right tail) the mode is given by

$$m_{GEV} = \mu + \sigma \frac{(1+\xi)^{-\xi} - 1}{\xi}.$$

The second part $\frac{(1+\xi)^{-\xi}-1}{\xi}$ only containing the parameter ξ decreases with increasing ξ, converging to zero for $\xi \longrightarrow 0$ and being equal to $-1/2$ for $\xi = 1$ (see Figure 7.11). Since for our flood series the location parameter μ is mostly much larger than the scale parameter σ (see Figure 7.12), the second term in the mode is very small and only has small influence on the mode. Coming back to the question of dominance of one sample it is obvious that the dominance is closely related to the difference of the modes. The higher the difference of the modes, the more dominant is the sample with the distribution of higher mode since for these (much) higher values are more probable. The most critical case is the one, where we have one sample with only slightly smaller scale but much heavier tail. In this case, it is not clear, which sample dominates in a maximum series. Nevertheless, since both modes are close together the samples are treated as more or less equal in share. Summarising these considerations we can conclude that the dominance of samples in a maximum series mainly corresponds to the difference of location, since this is the dominating term in the mode.

Remark 7.1. It is also possible to estimate the probability of overlay by calculating the intersection of both densities of the subsamples. For this the intersections of both functions have to be calculated. In practice this is of course not possible since the true parameters are not known and cannot be estimated. Therefore, this consideration is left out here.

Statistical simulations reported below confirm the theoretical assumptions concerning the location and show that maxima series resulting from two different distributions are always dominated

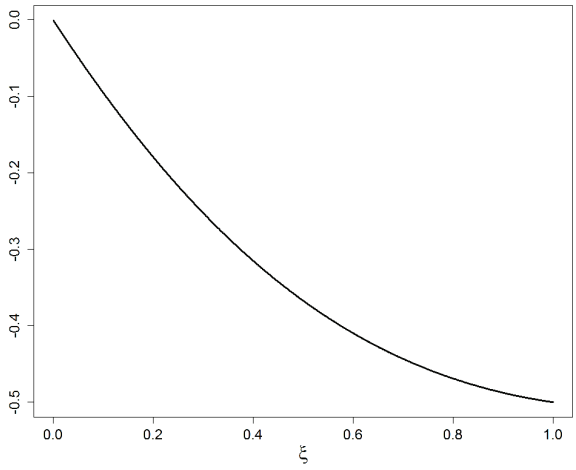

Figure 7.11.: The function $f(\xi) = \frac{(1+\xi)^{-\xi}-1}{\xi}$ representing the second term of the mode of the GEV distribution related to the shape parameter ξ. A convergence to zero for $\xi \longrightarrow 0$ becomes evident.

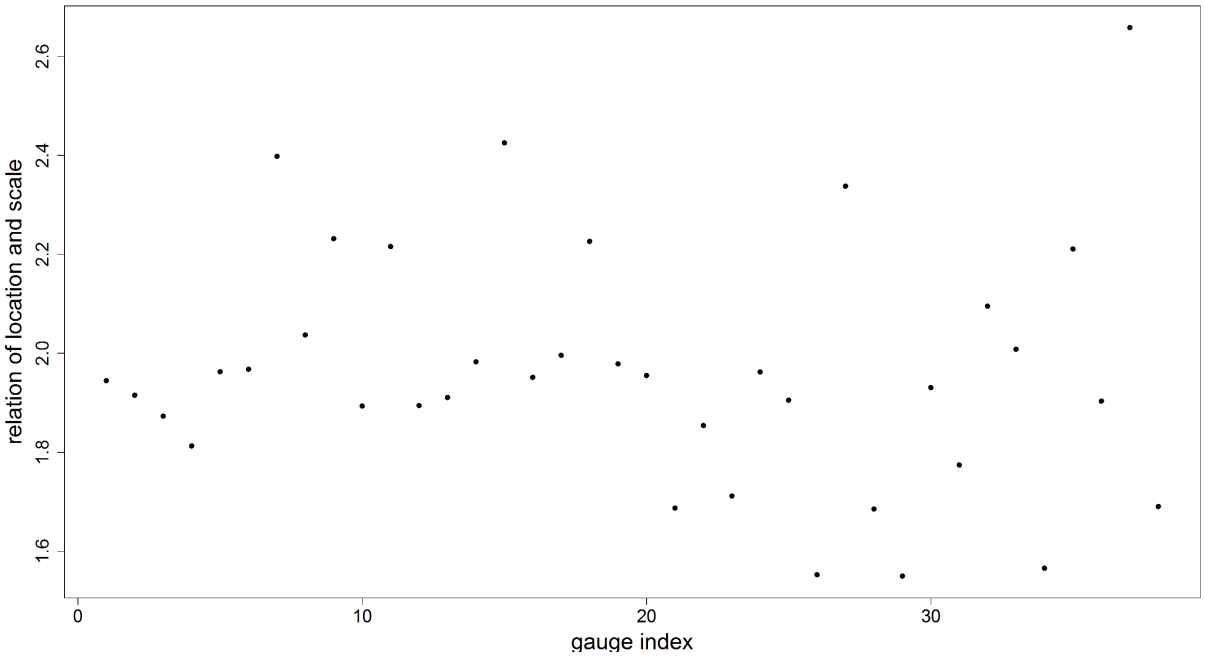

Figure 7.12.: Relation of location and scale parameter of the GEV for all considered gauges in the Mulde and Harz region. One can see that for the Mulde gauges (index 1-20) the relation of location and scale is always larger than 1.8. For the Harz region (21-38) it is always larger than 1.5.

by the distribution, which has the larger location parameter.
For these simulations we want to compare two samples of GEV-distributed data samples $(X_1, \ldots, X_n)$ and $(Y_1, \ldots, Y_n)$ of size $n = 1000$. For both samples the pairwise maximum $Z_i = \max(X_i, Y_i)$, $i = 1, \ldots, n$, is taken and then the share of all X_i on Z_i is calculated. This is repeated 1000 times. We take all combinations of parameter values of the following sets into account

- the location parameter $\mu \in \{10, 30, 50, 100\}$
- the scale parameter $\sigma \in \{2, 8, 12, 20\}$
- the shape parameter $\xi \in \{0.1, 0.3, 0.6\}$,

obtaining $4 \cdot 4 \cdot 3 \cdot 4 \cdot 4 \cdot 3 = 2304$ different combinations of parameters. The chosen parameter spaces represent the considered parameter space of the Mulde river basin well.
Since the results are too comprehensive to display here we show the results in a 3D-plot (Figure 7.13). On the axes the difference of the parameters of the two distributions is given, where large (positive) differences indicate a much higher parameter for the first sample. The colour of the single dots represents the share of the first sample on the maxima.
If all other parameters (scale and shape) are equal, shares of this dominating series of about 97% can be gained. If the other series has a larger scale, this relation is diminished but still the series with the larger location parameter dominates the maxima series (about $70 - 80\%$). The shape parameter only has a small influence on the shares of the two distribution functions in this series.

For the same simulation setting we also check if there is a certain pattern in the non-overlaid part by considering the ranks of the non-overlaid part of the samples related to the whole sample. One could suppose that a certain part of the upper tail is the part which is non-overlaid. That would give us additional information on the overlaid part and would help us to reconstruct it. Nevertheless, in the simulations no certain pattern is detectable. Of course, the rank of the non-overlaid part of the series depends much on the parameters, since we have shown in Figure 7.13 that often one sample dominates the maximum series. A similarity in the non-overlaid parts is that always the highest or at least the second-highest value of a sample is also contained in the maximum series. This is not surprising. What is somehow surprising is the rank of the smallest value contained in the maximum series. It varies from 1 (the smallest value in the single sample) up to the second largest rank. And also in the ranks of the non-overlaid part of the sample there is no certain pattern, that is they are not in consecutive order. Therefore, we can obtain no additional information by using the ranks of the non-overlaid part of the sample.

The knowledge about the influence of the parameters on the share leads to the following procedure to adjust the statistical parameters of the subsamples. The shares are estimated by counting the number of events belonging to the one or the other distribution, which occur in the maxima series, respectively. First, one estimates the location parameter μ_S of the whole summer annual maxima series. We assume that the event type with higher share in this sample determines this parameter. We therefore need the relation between the relative shares of the subsamples of both types in the whole series:

$$\alpha = \frac{\#(X_{S_S} | X_{S_S} \in X_S)}{\#(X_{S_L} | X_{S_L} \in X_S)} = \frac{a_{S_S}}{a_{S_L}}.$$

If $\alpha > 1$, then the share a_{S_S} of the short events is larger and μ_S is mainly determined by these. If $\alpha < 1$, the same is valid for the long events, whereas, if $\alpha \approx 1$, the share of both types of

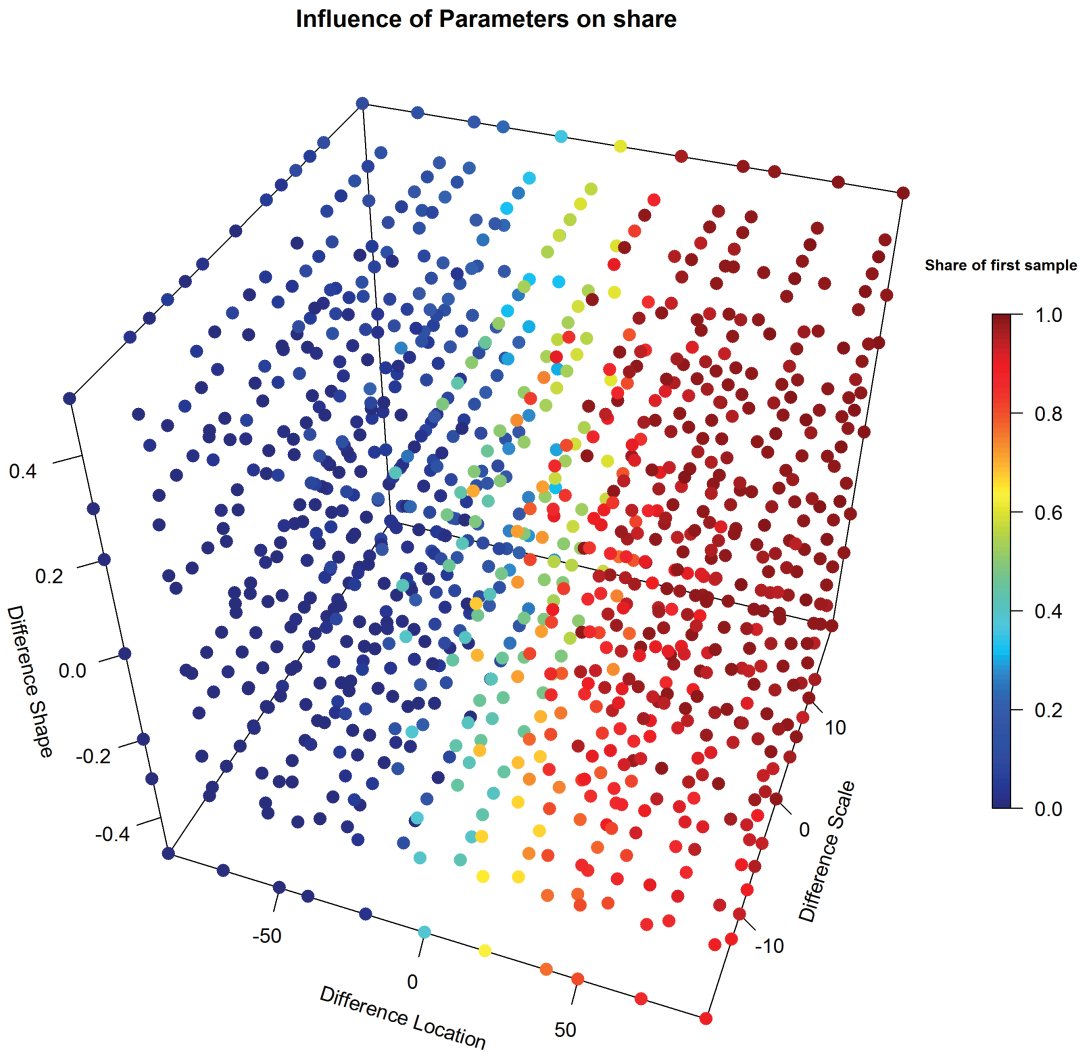

Figure 7.13.: Share of the first sample on the maximum series of two GEV-distributed samples as a function of the difference between the parameters of the distributions. It can be seen that mainly the location parameter has influence on the share, where a larger location parameter leads to a larger share.

events is nearly the same and we can assume that both have nearly the same location parameter μ_S. For the subsample with a smaller share we use the information that its location parameter is smaller than that of the dominating series, otherwise this series would be dominated in another way as we have stated above. Therefore, the location parameter of the smaller subsample has to be reduced.
To adjust the location parameters we introduce two parameters ρ_{S_S} and ρ_{S_L}. Depending on the relationships of shares α, these parameters, which are used for a multiplicative correction of the location parameters μ_{S_S} and μ_{S_L}, are specified as follows:

$$\begin{array}{rrcl} \rho_{S_S} = a_{S_S} + \frac{1}{2}, & \rho_{S_L} = 1 & , & \alpha < 1 \\ \rho_{S_S} = 1, & \rho_{S_L} = 1 & , & \alpha = 1 \\ \rho_{S_L} = a_{S_L} + \frac{1}{2}, & \rho_{S_S} = 1 & , & \alpha > 1 \end{array} .$$

The more equal the shares of both event types within the summer series, the more equal both location parameters are to the location parameter of the summer series. The smaller the share of one type, the more deviates its location parameter from the summer series, since the distribution with higher location parameter dominates the summer events. At the same time, samples with the same location parameter will lead to more or less equal share (see the simulation above) and therefore the location parameter is not corrected much for both samples. The definition of the correction factors ρ_i guarantees that the location parameter is not reduced by more than 50%. This is needed since we do not want to obtain a support which is negative and far away from 0.
To replace the missing values we can use the fact that for the subseries, which dominates the summer annual maxima, the missing values scatter around the parameter μ_S, since all overlaid events do not belong to the sample of large events we observe. Probably, no events from the right tail occur in the series of missing values (these are already included in the non-overlaid series) and so we can assume that these are distributed symmetrically. There is no option to estimate the variance of this subsample directly. Relatively large events are possibly overlaid by large events, belonging to the other subsample. That is why we consider here a spectrum around the location parameter to introduce an artificial variance. We assume that this spectrum lies symmetrically in distance of the assumed variance of the respective event type around μ_S, so that we include larger as well as smaller events in keeping the mean μ_S. Under the consideration that we only know a subsample, we use the robust estimator MAD (median absolute deviation) for estimation of this variability from the known part of this series

$$MAD(X_1, \ldots, X_m) = C \cdot \text{med}(|X_i - \text{med}(X_1, \ldots, X_m)|, i = 1, \ldots, m),$$

where $C = \frac{1}{\Phi^{-1}(3/4)} \approx 1.4826$ ensures the consistency under normality (cf. Huber (1981)). Since we only use a part of the series, we can assume that large deviation from normality will not occur. As example we show a QQ-plot of the non-overlaid part of a simulated series of $n = 1000$ data points ($\sim GEV(20, 5, 0.4)$) compared with a normal distribution (Figure 7.14), where we only see a small deviation from normality on the lower tail and for the largest value. Here, again we want to use the property of stability in estimation of robust estimators, even for small samples. Of course, also other normalisation factors have been used. Nevertheless, smaller choices showed a strong underestimation of the left tail, whereas larger values of the normalisation constant led to a larger overestimation of the right tail.
Please note that the MAD estimator can also be expressed as a GL-statistic. It has been constructed by Huber (1981) especially for symmetric distributions, which has been the reason for Rousseeuw and Croux (1992) to develop more general estimators, also efficient for asymmetric

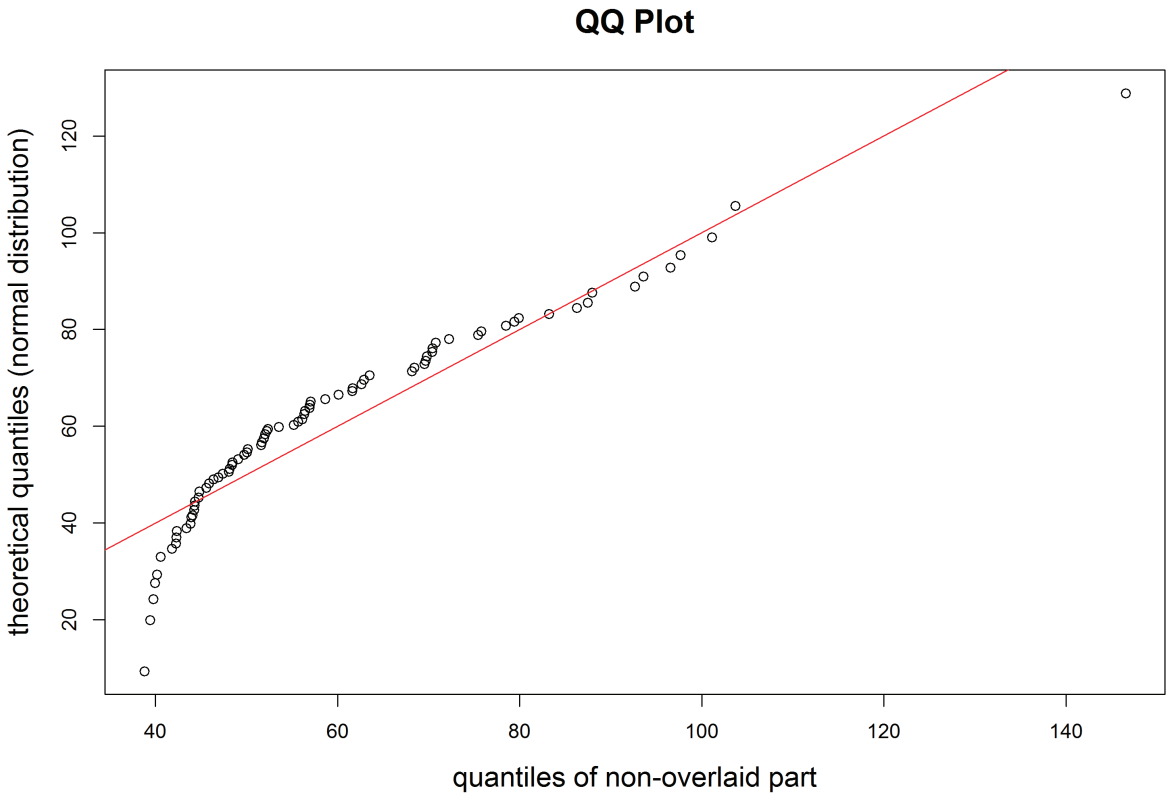

Figure 7.14.: QQ-plot of the non-overlaid part of a sample of $n = 1000$ $GEV(20, 5, 0.4)$-distributed data compared with the normal distribution. The maximum series consisting of the non-overlaid parts is constructed with a second series of $n = 1000$ $GEV(40, 2, 0.1)$-distributed data. A small deviation from normality can be found in the left tail.

distributions. But we have seen that symmetry can be assumed here, so the MAD estimator is sufficient.
The robustness of MAD has the effect that for a large scale parameter of the annual summer maxima series (that means many large events in the sample) the assessment of the variance will result in a large value, too. This is important, since one can assume that for a large scale also the overlaid events scatter much. However, if the shape parameter is large, single extremely large events will occur in the mixed series of summer annual maxima. Since one can assume here that the overlaid events only represent the left tail of the distribution, the large shape parameter should have no influence on the spectrum of the "filling" data. By using the robust estimator this requirement is fulfilled.
After estimating the location as well as the variance of the overlaid sample it remains to simulate the missing values. For this, a distribution has to be chosen. Since the right tail is already represented by the non-overlaid part of the subsample, a symmetric distribution seems to be favourable. As mentioned before the left tail should be bounded and close to the theoretical left bound of the support, $\mu - \sigma/\xi$.
To generate the missing values in the two subsamples we therefore choose the uniform distribution, for which the bounds are known. Of course, also other symmetric distributions like the normal distribution would be possible, but this has unbounded support. As mentioned above, the skewness has no influence on the overlaid data and the use for the calculation of the left bound would instead only increase the uncertainty due to the uncertain estimation. In fact, these parts of the subsamples will mainly determine the location of the whole sample, such that a uniform distribution over an interval +/- the distance of the variance to the adjusted location is sensible. Moreover, this is approximately symmetric around the mode.
Therefore, the missing values in both samples are simulated from the following two distributions:

$$\begin{aligned} X_{S_S} &\sim \mathcal{U}[\rho_{S_S}\mu_S - MAD_{S_S}, \rho_{S_S}\mu_S + MAD_{S_S}] \\ X_{S_L} &\sim \mathcal{U}[\rho_{S_L}\mu_S - MAD_{S_L}, \rho_{S_L}\mu_S + MAD_{S_L}], \end{aligned}$$

with MAD_i denoting the estimated standard deviation of the known part of the series i.
As an alternative one could also use the $kMAD$, the MAD for asymmetric distributions given by

$$kMAD(F,k) = \inf\{t > 0 | F(m+kt) - F(m-t) \geq 1/2\},$$

where F is the cumulative distribution function and m the median of F (see e.g. Ruckdeschel and Horbenko (2010)). Nevertheless, in this case we would have to choose a tuning parameter k leading to a similar problem as for the classical MAD. Hence, as mentioned above, we can assume symmetric data such an extension is not necessary in this case.

To explain this procedure in greater detail, we present the following example. Let $(Y_1, \ldots, Y_n)$ with $n = 100$ be a GEV distributed time series with the parameters $\mu_1 = 20$ (location), $\sigma_1 = 5$ (scale) and $\xi_1 = 0.4$ (shape). Moreover, let $(Z_1, \ldots, Z_n)$ be GEV distributed with $\mu_2 = 30$, $\sigma_2 = 3$ and $\xi_2 = 0.2$. The overall series of maxima is then
$(X_1, \ldots, X_n) = (\max(Y_1, Z_1), \ldots, \max(Y_n, Z_n))$. The estimated parameters (using L-moments) of this series are $\hat{\mu} = 31.53$, $\hat{\sigma} = 4.137$ and $\hat{\xi} = 0.168$.
The share of the Y_i in the sample X_i is 21%, the share of the Z_i therefore 79%. The relationship between the numbers of the members of Y_i and of Z_i in X_i is $\alpha = 0.266$.
We now use the samples Y_i' and Z_i', consisting only of data that are part of the series X_i, to construct new time series Y_i and Z_i. To fill in the missing values, we simulate values from a

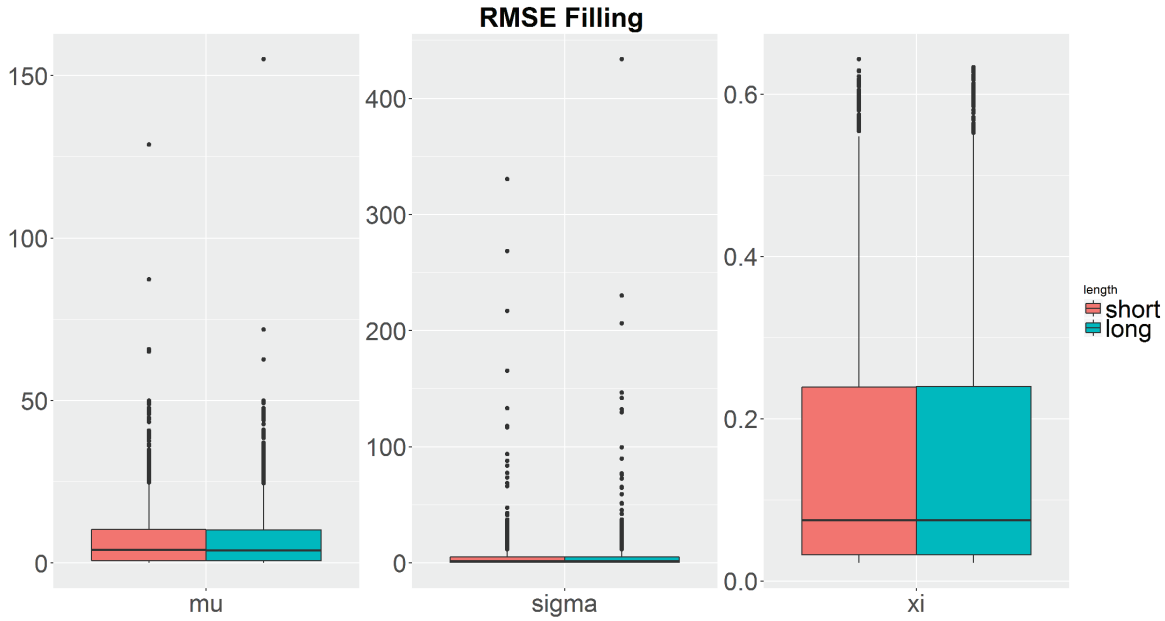

Figure 7.15.: Boxplots for the RMSEs of the estimated parameters for long and short summer series using the "filling method" under the simulation scenario. The location is estimated well for all samples, whereas the scale has some variance in the estimation for the sample of short floods. Estimating the shape parameter of the GEV always leads to a high RMSE such that this is not surprising.

uniform distribution. For the Y_i, that have the smaller share, the interval used for the uniform distribution is $[\rho_1\hat{\mu} - \text{MAD}(Y'), \rho_1\hat{\mu} + \text{MAD}(Y')]$ with $MAD(Y') = 6.510$, $\hat{\mu} = 31.53$ and $\rho_1 = 0.21 + 0.5 = 0.71$. The series of the Z_i' is filled analogously, though we multiply $\hat{\mu}$ with $\rho_2 = 1$. The parameters of the GEV are then estimated as $\hat{\mu}_1 = 22.181$, $\hat{\sigma}_1 = 4.739$ and $\hat{\xi}_1 = 0.366$ as well as $\hat{\mu}_2 = 30.330$, $\hat{\sigma}_2 = 2.577$ and $\hat{\xi}_2 = 0.119$. If one would estimate these parameters only based on the events occurring in the total series of the X_i one would get $\tilde{\mu}_1 = 40.313$, $\tilde{\sigma}_1 = 6.745$ and $\tilde{\xi}_1 = -0.212$ as well as $\tilde{\mu}_2 = 30.582$, $\tilde{\sigma}_2 = 2.845$ and $\tilde{\xi}_2 = 0.102$. The less represented subsample (Y_i) is estimated markedly skewed to the left and has a larger location parameter, since small events are overlaid and not considered in this case. An assessment of the statistical characteristics of both subsamples based on the results within the total series gets better if the shares of the two event types are more similar. In such cases, also the location parameters are more similar. To show that this filling method can be applied, several simulation studies have been done. We choose two samples of length $n = 1000$ of long and short summer floods with parameters of the GEV-distribution μ_{S_S}, σ_{S_S}, ξ_{S_S} and μ_{S_L}, σ_{S_L}, ξ_{S_L}, respectively. For these two samples the maximum of each data pair is used to obtain the annual summer maximum series. Applying the "filling method" to the summer maxima we then calculated the RMSE of the estimates of the single parameters.

This has been done 1000 times for each choice of parameters for the simulation scenario already stated above. We want to show the results of the RMSE in a boxplot (Figure 7.15).

The results show in general a good estimation of the parameters by having small RMSE. For the estimation of the location we see a median RMSE of 4 and only a few very large values. Since we vary the location parameter between 10 and 100, the very small box in the boxplot indicates

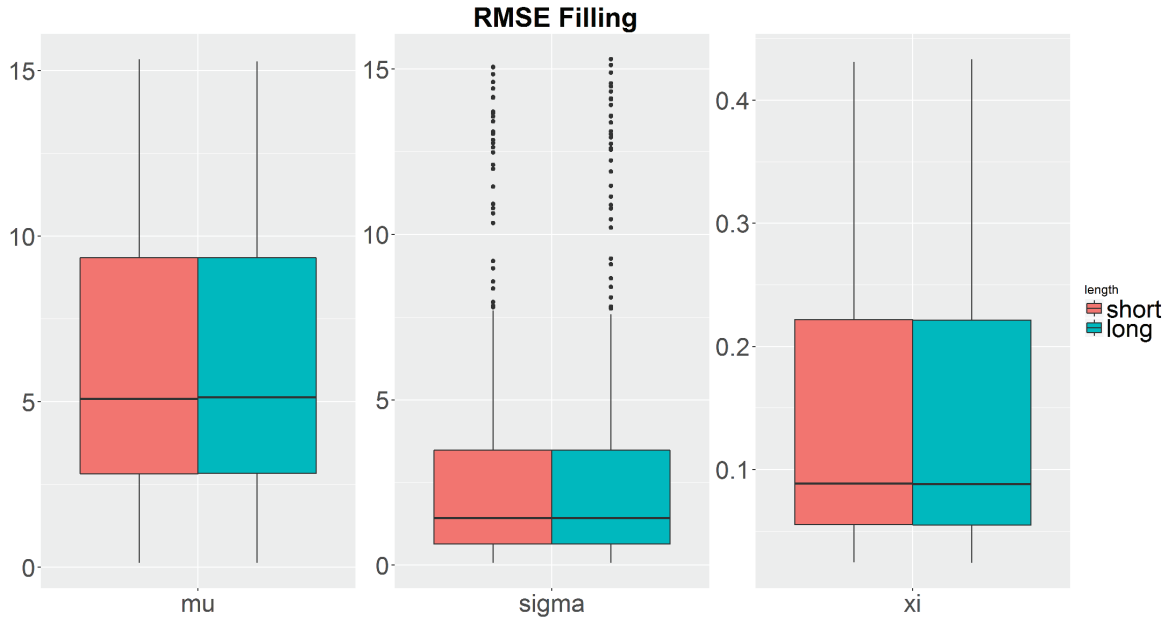

Figure 7.16.: Boxplots for the RMSEs of the estimated parameters for long and short summer series using the "filling method" under the simulation scenario (7.2.1), where the share of the short respectively long series on the maximum is at least 25%. The RMSE for both, the scale and the location, is relatively small indicating a good estimation. Estimating the shape parameter of the GEV always leads to a high RMSE such that this is not surprising.

really good estimation. The estimation of the shape parameter though has a very high RMSE compared to the assumed values, but this can be expected since the estimation of the shape parameter is always a source of high uncertainty. What is striking is the very high variability in the estimation of the scale parameter. Although the main part of the RMSEs of the estimation lies in a small range around 2, with median 1.6, for single samples it reaches values of more than 400. Now we have to take into account that the choices of the parameters made in the simulation can lead to shares of short or long samples of less than 1%. This is a result that is excluded in real practice by our method of dividing the groups, which demands a minimum share of ϵ, mostly chosen as $\epsilon = 0.25$. This would avoid samples, where the information of one subsample is so small such that the information obtained from it is more or less not existing. In this case almost all information is obtained by the Filling method and therefore the uncertainty in the estimation can be extremely high in some samples. To give a fair comparison we thus also want to present the RMSEs of all samples, where the share of the short respectively the long subsample is at least 25%. This can be found in Figure 7.16. Here, we see that when having a minimum amount of data in each subsample the Filling method delivers good estimates of all the parameters.

For a few of the scenarios that are assumed to be most relevant in practice we want to give a detailed analysis, including both scenarios already used for the ML-estimation above (now Scenario 3 and 4). The results can be found in Table 7.7.

As a comparison we also calculated the RMSE when estimating the parameters of a GEV-distributed sample of length $n = 1000$ with parameters $\mu = 20$, $\sigma = 6$, $\xi = 0.4$ using L-moments and 1000 runs. The results can be found in Table 7.8.

Table 7.7.: RMSEs of the estimates of the parameters using the "filling method" for sample size $n = 1000$

Number	Scenario	RMSE for					
		μ_{S_S}	μ_{S_L}	σ_{S_S}	σ_{S_L}	ξ_{S_S}	ξ_{S_L}
1	$\mu_{S_S} = 20$, $\sigma_{S_S} = 6$, $\xi_{S_S} = 0.4$ $\mu_{S_L} = 20$,$\sigma_{S_L} = 6$, $\xi_{S_L} = 0.4$	1.44	1.46	0.255	0.255	0.0456	0.0455
2	$\mu_{S_S} = 20$, $\sigma_{S_S} = 3$, $\xi_{S_S} = 0.2$ $\mu_{S_L} = 20$,$\sigma_{S_L} = 3$, $\xi_{S_L} = 0.2$	2.94	2.95	0.640	0.654	0.1197	0.1171
3	$\mu_{S_S} = 20$, $\sigma_{S_S} = 5$, $\xi_{S_S} = 0.4$ $\mu_{S_L} = 40$,$\sigma_{S_L} = 2$, $\xi_{S_L} = 0.1$	0.72	0.098	3.130	0.079	0.2126	0.0259
4	$\mu_{S_S} = 20$, $\sigma_{S_S} = 2$, $\xi_{S_S} = 0.6$ $\mu_{S_L} = 20$,$\sigma_{S_L} = 2$, $\xi_{S_L} = 0.1$	0.91	0.50	0.766	0.312	0.2330	0.0319
5	$\mu_{S_S} = 5$, $\sigma_{S_S} = 6$, $\xi_{S_S} = 0.1$ $\mu_{S_L} = 20$,$\sigma_{S_L} = 2$, $\xi_{S_L} = 0.1$	5.45	0.077	2.077	0.076	0.0745	0.0246
6	$\mu_{S_S} = 60$, $\sigma_{S_S} = 8$, $\xi_{S_S} = 0.1$ $\mu_{S_L} = 40$,$\sigma_{S_L} = 8$, $\xi_{S_L} = 0.2$	0.45	5.57	0.319	1.285	0.0249	0.0436

Table 7.8.: RMSEs for the estimates of the parameters of $n = 1000$ simulated GEV(20,6,0.4)-distributed random variables

RMSE for		
μ	σ	ξ
0.233	0.241	0.0459

The Scenarios 1 and 2 are examples where both distributions, short and long summer, are equal. In this case, also the RMSEs are equal. For the second scenario, where scale and shape have higher values, the RMSEs also increase. This is a typical behaviour of estimators when estimating the GEV-parameters. Nevertheless, for both scenarios all RMSE values are very small and not much larger as if we would have estimated the parameters directly from the whole sample. The Scenarios 3 and 4 are again simulated as probable flood scenarios. For these two scenarios we can observe a much higher share of the long summer events than of the short summer events. This is also mirrored in the RMSEs. We can see that the estimation of the shape parameter of the short events has a higher RMSE than before and also the RMSE of the scale parameter for Scenario 3 is higher. Remember that the shape parameter used is chosen on basis of the actual observed values of short respectively long summer events and only changed due to the additional values from the uniform distribution. Nevertheless, all RMSEs are still relatively small compared to the RMSEs when we have a full sample without overlaid data (Table 7.8) and even the RMSE of the shape parameter is not atypical when estimating this parameter. In fact, the estimation of the shape parameter is always the source of high uncertainty as can be seen in Table 7.8. The Scenarios 5 and 6 are unlikely to really occur in the gauges we considered. However, due to completeness we also want to show the results for these scenarios. With decreasing share of the flood type on the whole annual summer maximum sample the RMSE increases. Especially in Scenario 5 the RMSE of the estimated location parameter for short events is very high. This is a scenario where the long summer events have a share of about 80%. Therefore, nearly no information of the short flood series can be found in the annual summer maximum series and the estimation is very uncertain. A comparison with an estimation only based on the values occurring in the annual summer maximum series though showed that the results are still much better.
Since the first approach using censored statistics mainly showed high RMSE values for small sample sizes we want to give a fair comparison by repeating the simulations above with $n = 100$ (Table 7.9). To make a comparison easier the critical results of the ML-estimation method of the Survival Analysis are repeated here (Table 7.10).
Although the RMSEs increase with decreasing sample size, in no scenario they are as high as for the ML-estimation in Scenario 3. Thus, the filling of the overlaid series makes a good estimation (concerning RMSE) possible even for small sample sizes and possible extreme shares. Even for Scenario 4, where the ML-estimator gave good results for both sample sizes, our filling approach is comparable concerning the RMSE. Since these are the scenarios that are most interesting for us in this context and since no extremely high RMSE occurred during simulation for any considered scenario we decide to use the filling approach from now on.

For the scenarios 3 and 4, which are the most important for the considered Mulde river basin, we also want to show the results for the Bias, using boxplots in Figures 7.17 and 7.18.
The boxplots show that the RMSEs result mainly from the bias in the estimation. We also tested the influence of the adjusting factor in the calculation of the ρ_i. Therefore, we used the factor 0.25 and 0.6 instead of 1/2. For both choices RMSE and bias of the estimates got worse leading to the conclusion that 1/2 is a good choice.
Additionally, for Scenario 3 we compare the overlaid part of the short and long summer samples with the simulated series used to fill the overlaid part. The results are given in a QQ-plot (Figure 7.19). We can see that the main part of the data is reconstructed well, only the right tail is overestimated. Nevertheless, this is not a problem for the filled series, since the right tail of the simulated series forms only a small part of the filled series and consists of values lying in the mean of the whole filled short or long summer sample. Note, that the different sample sizes

Table 7.9.: RMSEs of the estimates of the parameters using the "filling method" for sample size $n = 100$

Number	Scenario	RMSE for μ_{S_S}	μ_{S_L}	σ_{S_S}	σ_{S_L}	ξ_{S_S}	ξ_{S_L}
1	$\mu_{S_S} = 20$, $\sigma_{S_S} = 6$, $\xi_{S_S} = 0.4$ $\mu_{S_L} = 20$,$\sigma_{S_L} = 6$, $\xi_{S_L} = 0.4$	1.404	1.471	0.462	0.482	0.0996	0.0997
2	$\mu_{S_S} = 20$, $\sigma_{S_S} = 3$, $\xi_{S_S} = 0.2$ $\mu_{S_L} = 20$,$\sigma_{S_L} = 3$, $\xi_{S_L} = 0.2$	3.159	3.155	1.328	1.435	0.162	0.162
3	$\mu_{S_S} = 20$, $\sigma_{S_S} = 5$, $\xi_{S_S} = 0.4$ $\mu_{S_L} = 40$,$\sigma_{S_L} = 2$, $\xi_{S_L} = 0.1$	2.424	0.262	5.019	0.200	0.222	0.0827
4	$\mu_{S_S} = 20$, $\sigma_{S_S} = 2$, $\xi_{S_S} = 0.6$ $\mu_{S_L} = 20$,$\sigma_{S_L} = 2$, $\xi_{S_L} = 0.1$	0.985	0.829	0.947	0.410	0.259	0.0899
5	$\mu_{S_S} = 5$, $\sigma_{S_S} = 6$, $\xi_{S_S} = 0.1$ $\mu_{S_L} = 20$,$\sigma_{S_L} = 2$, $\xi_{S_L} = 0.1$	5.690	0.251	2.908	0.197	0.183	0.0825
6	$\mu_{S_S} = 60$, $\sigma_{S_S} = 8$, $\xi_{S_S} = 0.1$ $\mu_{S_L} = 40$,$\sigma_{S_L} = 8$, $\xi_{S_L} = 0.2$	1.061	6.140	0.781	3.606	0.0818	0.149

Table 7.10.: RMSEs of the estimates of the parameters using Survival Analysis for sample sizes $n = 100$. High values for the RMSE for small sample sizes indicate a lack of efficiency of the ML-estimator.

Number	Scenario	RMSE for $n = 100$			$n = 1000$		
		μ_{S_S}	μ_{S_L}	σ_{S_S}	σ_{S_L}	ξ_{S_S}	ξ_{S_L}
3	$\mu_{S_S} = 20$, $\sigma_{S_S} = 5$, $\xi_{S_S} = 0.4$ $\mu_{S_L} = 40$,$\sigma_{S_L} = 2$, $\xi_{S_L} = 0.1$	20.517	0.237	13.317	0.178	0.449	0.0984
4	$\mu_{S_S} = 20$, $\sigma_{S_S} = 2$, $\xi_{S_S} = 0.6$ $\mu_{S_L} = 20$,$\sigma_{S_L} = 2$, $\xi_{S_L} = 0.1$	1.132	0.324	0.379	0.335	0.160	0.114

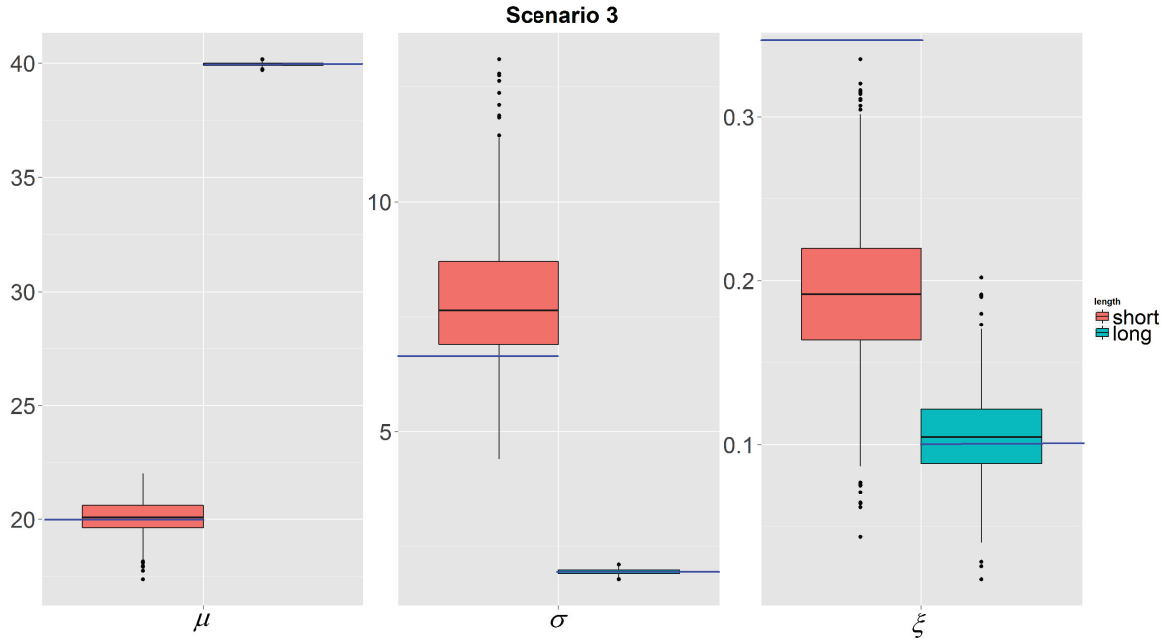

Figure 7.17.: Boxplots for the estimates of the parameters for long and short summer series using the "filling method" under Scenario 3 with the blue lines denoting the true parameter. The location is estimated well for both samples, whereas the scale has some variance in the estimation for the sample of short floods. Estimating the shape parameter of the GEV always leads to a high RMSE such that this is not surprising.

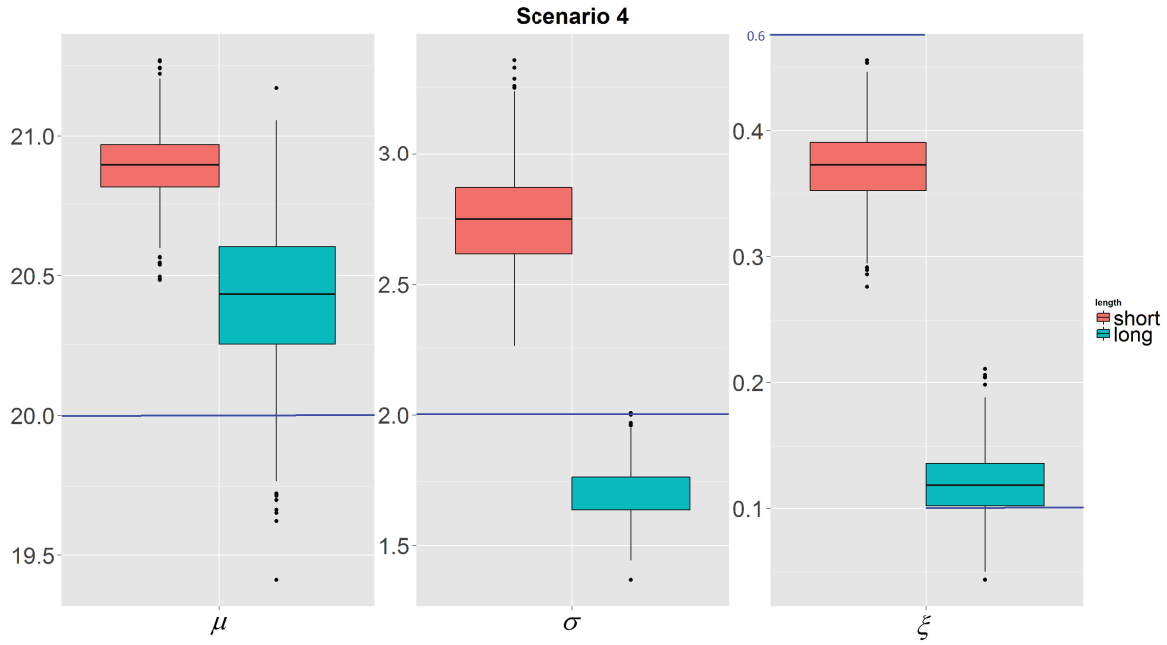

Figure 7.18.: Boxplots for the estimates of the parameters for long and short summer series using the "filling method" under Scenario 4 with the blue lines denoting the true parameter. For this scenario a smaller RMSE occurs for all estimates, although the shape parameter for the short events is underestimated.

result because of the different share of the short and long samples on the summer sample due to the chosen parameters.

The need for this data correction is shown here for the Berthelsdorf gauge considered above (in Table 7.5). As a threshold to distinguish between short and long summer events a TQ-value of 36 hours is estimated. The share of short events is $\alpha_{S_S} = 0.415$ and hence long events have a share of $\alpha_{S_L} = 0.585$. The quotient of both is $\alpha = 0.709$. We apply the method to fill up the series. The differences between the GEV parameters are estimated compared to an estimation based on the non-overlaid part of the series only (Table 7.11). How this filling procedure influences the distribution functions and the calculated annualities of long and short summer events is shown in Figure 7.20 and Table 7.12. As expected, the filling of data series of subsamples reduces the exceedance probabilities and increases the return periods.

Table 7.11.: Estimated parameters for the Berthelsdorf gauge with and without the method of the filled series. The adjustment of the location parameter μ and the scale parameter σ becomes evident, whereas the shape parameter ξ remains the nearly same.

	short events in summer		long events in summer	
	half-years maxima only	filled series	half-years maxima only	filled series
μ	14.664	13.846	17.119	13.910
σ	6.808	4.302	7.736	4.116
ξ	0.755	0.718	0.494	0.336

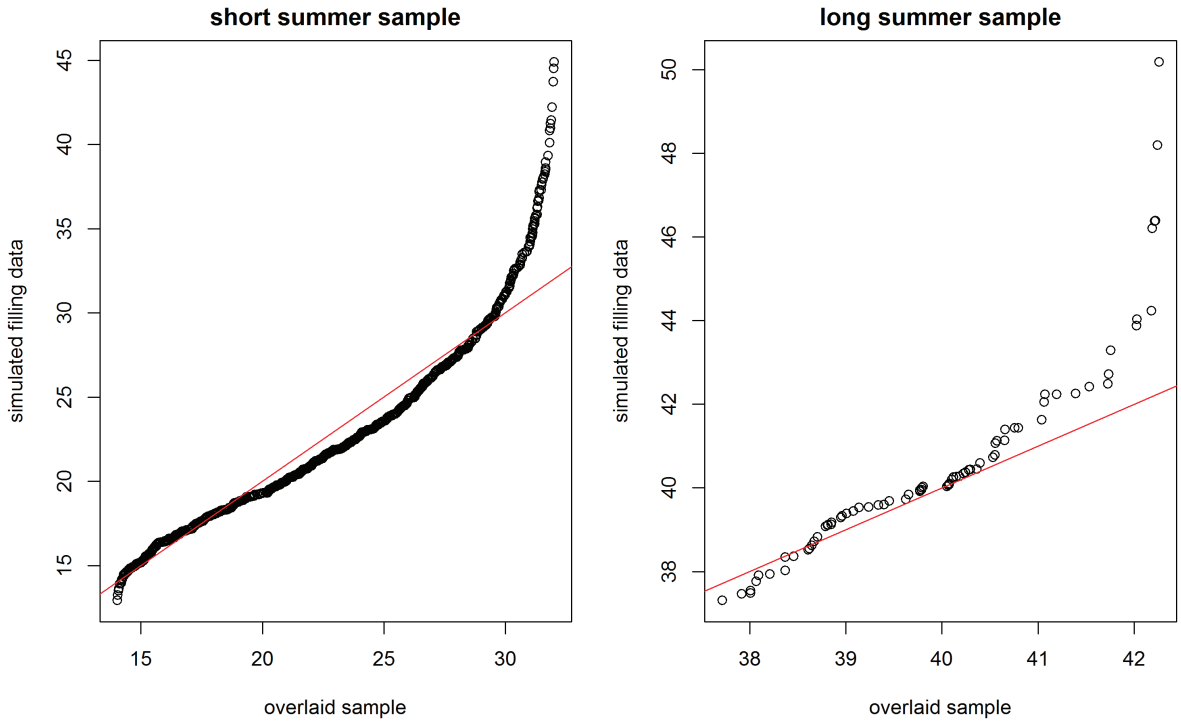

Figure 7.19.: QQ-plot of the overlaid part of the short (left) and long (right) summer series and the reconstructed part using the filling method. The underlying parameters are chosen as Scenario 3. A good accordance for the main part of the samples can be found, only the right tail is overestimated.

Table 7.12.: Estimated quantiles for the filled series and the maxima only.

annuality T	short events in summer		long events in summer	
	half-years maxima only	filled series	half-years maxima only	filled series
100	296.33	170.78	153.41	59.13
200	497.15	276.33	215.73	74.26
500	988.45	526.77	338.64	100.48
1000	1664.88	861.72	476.45	126.42
2000	2806.36	1412.63	670.49	159.15

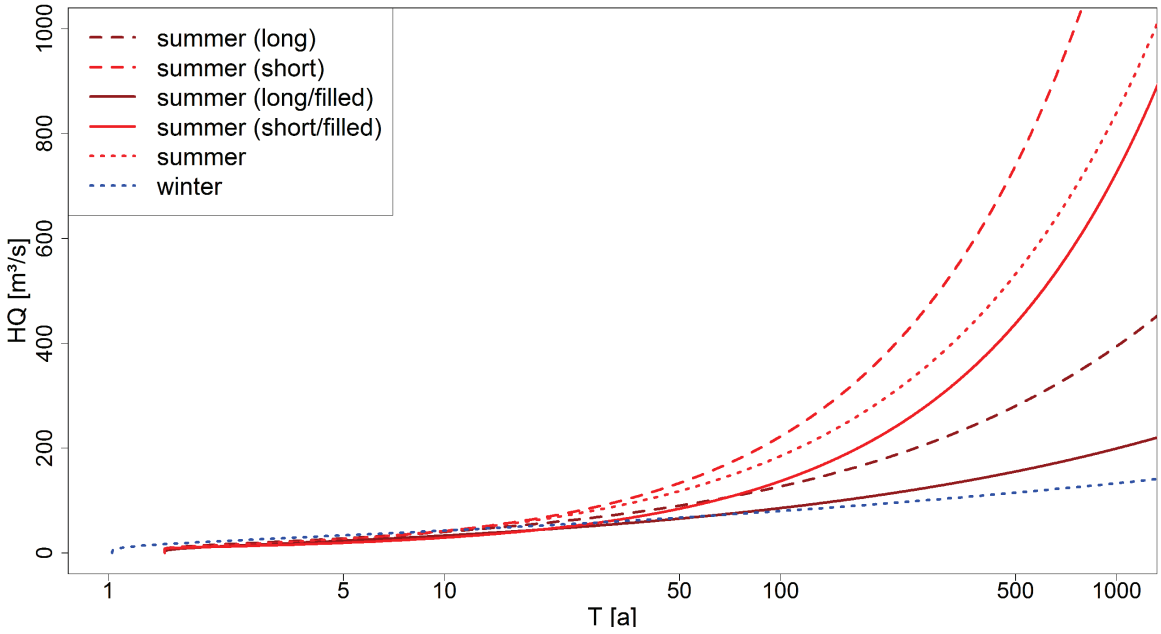

Figure 7.20.: Quantiles of the series of short and long summer annual maxima under use of the filling method or the non-overlaid subsamples for the Berthelsdorf/Freiberger Mulde gauge. The lesser slope of the "filled" series becomes obvious caused by taking into account the overlaid small values.

Summarizing the results so far we are now able to distinguish between different event types and reconstruct the missing information of the event types caused by the data situation. This forms the basis for the now developed new mixture models that takes into account the different event types.

We can see that by using the mixture distribution of short and long summer events the lower tail of the summer floods is represented mainly by the long summer events whereas the right tail, that is the extreme events, mostly consists of short events (Figure 7.21). This implies that the floods with high peaks are mainly caused by short but intensive (convective) rain events. A distinction into these groups of different duration is therefore necessary.
A seasonal model, which only differs between summer and winter (in the following specified as the WS (winter-summer) maximum approach) can use the series of summer maxima directly. If we distinguish additionally between long and short events within the series of summer events (WST (winter-summer-types) maximum approach) we have to apply a method of "filling" the data series of the two additional subsamples. The proposed filling procedure causes an additional uncertainty in estimation of flood probabilities. To evaluate this, we compare both models with regard to the differences in their estimations of annualities for the gauges of the Mulde river basin. For the Berthelsdorf gauge we have already compared summer and winter annual maximum discharges. As a threshold specifying a flood event among the series of half-year maxima, we choose the minimum annual maximum discharge (9.27 m^3/s). The series of summer events is subdivided into short and long floods by the threshold of the timescale of 36 h. The calculated return periods of the six largest floods at this gauge in the discharge series from 1936 to 2013 are listed in Table 7.13. Here we compare the classical annual maxima series (AMS) using the GEV distribution

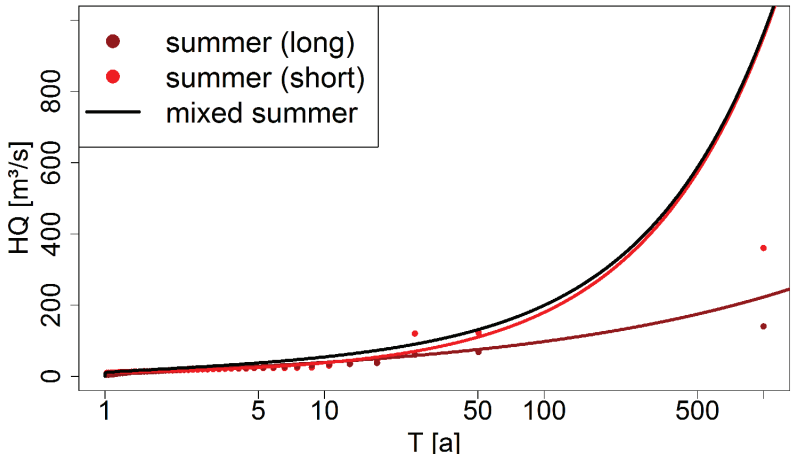

Figure 7.21.: Short and long summer events with their estimated annuality and the fitted mixing distributions. It can be seen that the right tail of the mixing distributions is mainly influenced by the short summer events.

Table 7.13.: Annualities at the Berthelsdorf/Freiberger Mulde gauge for the six largest floods using three different models

HQ[m$^3/s$]	AMS (GEV)	Maximum approach WS	Maximum approach WST
63.9	10	13	13
68.2	11	15	15
120	35	47	53
122	41	49	54
140	55	62	71
360	476	275	342

with results of the two maximum mixing approaches WS and WST. A graphical representation of the results can be found in Figure 7.22.
The Berthelsdorf gauge has a relative small catchment area of 244 km^2. As discussed above, this leads to a large difference between short and long events. Whereas the short events are characterised by a large peak with relative small volume, the long events have relative small peaks even if they have high volumes.
As shown above, in large catchments this distinction is less clear. The Wechselburg/Zwickauer Mulde gauge has a large catchment area of 2107 km^2 and only small differences between short and long summer annual maxima (Figure 7.23). Here the TQ-threshold TQ_0 is calculated as 55 h. The threshold of flood peaks is 59 m$^3/s$ and $\alpha = 1.421$.
An overview of the estimated parameters is given in Table 7.14.
It is obvious that the value of skewness in the series of the half-year annual maxima is much lower than in the previous example. In addition, the difference between the distributions of long and short summer floods is much smaller.
As in the example of the Berthelsdorf gauge before we estimated the distribution functions and the return periods with the different approaches (Figure 7.24 and Table 7.15).
The differences in return periods of the highest observed flood at this gauge between the seasonal maximum approach (WS) and the utilization of AMS are not as high as before. This results from

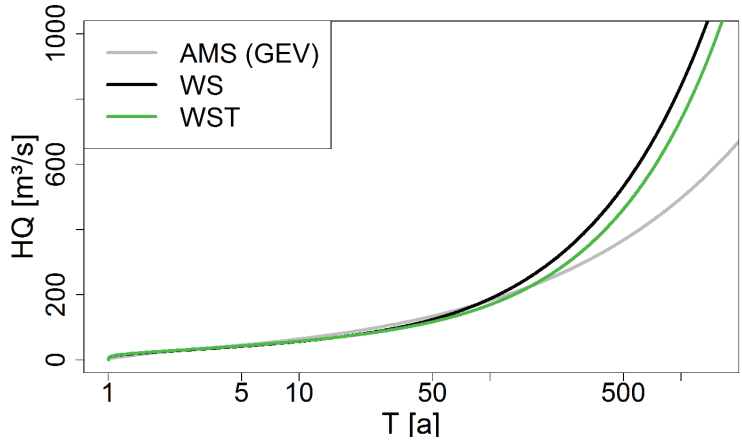

Figure 7.22.: Quantiles calculated for the Berthelsdorf/Freiberger Mulde gauge using different statistical approaches. The WST model lies between the classical WS model and the AMS-fitted GEV since it is not influenced by single extraordinary large summer events like the WS model but takes into account the different distributions of the seasons.

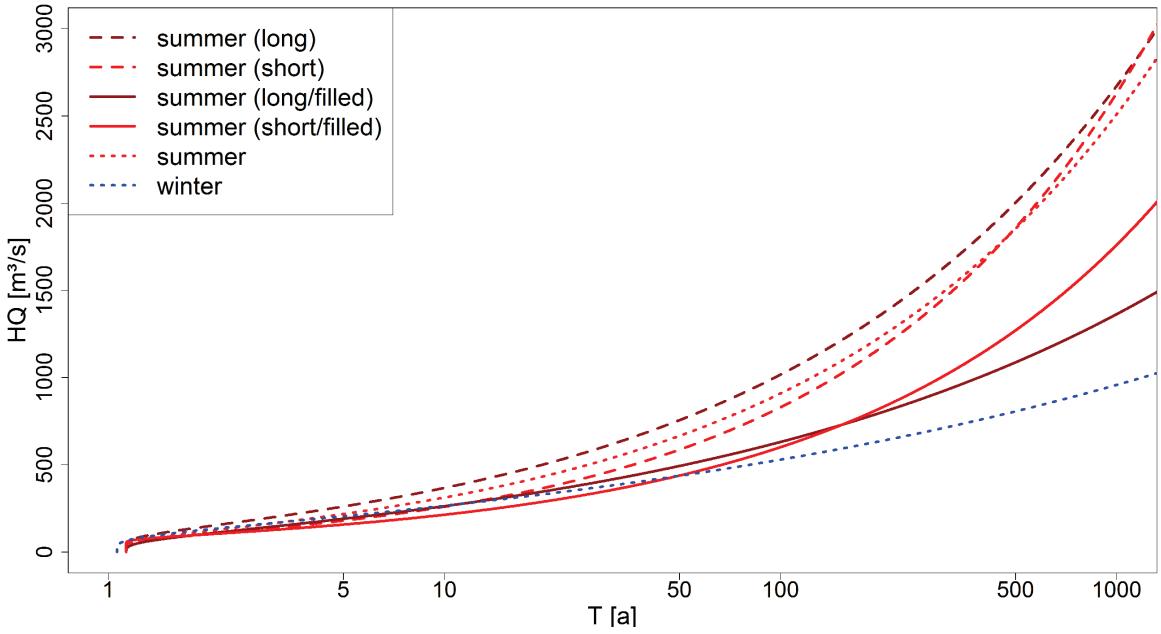

Figure 7.23.: Distribution functions of the series of short and long summer annual maxima at the Wechselburg/Zwickauer Mulde gauge under use of the filling method or the subsamples. The lesser slope of the "filled" series becomes obvious caused by taking into account the overlaid small values.

Table 7.14.: Estimated parameters of the GEV for the Wechselburg/Zwickauer Mulde gauge.

	ξ	μ	σ
AMS	0.305	152.56	77.54
Winter	0.216	115.07	53.83
Summer	0.433	110.48	58.27
Summer (short)	0.503	94.94	43.50
Summer (long)	0.41	136.16	68.50
Summer (short/filled)	0.486	92.56	32.04
Summer (long/filled)	0.281	95.29	59.62

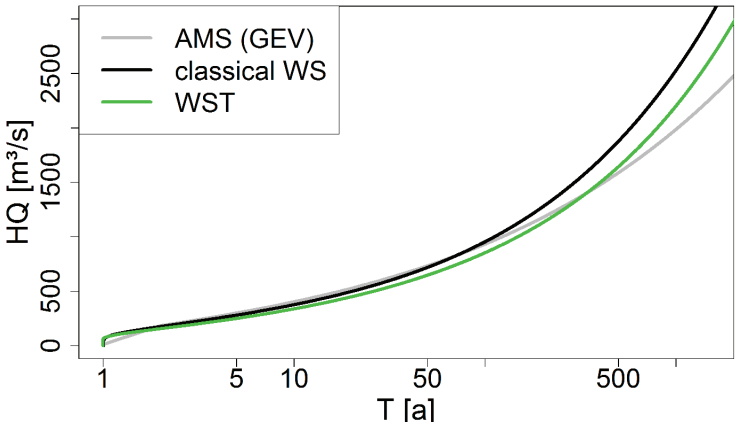

Figure 7.24.: Distribution functions calculated for the Wechselburg/Zwickauer Mulde gauge using different statistical approaches. The WST model lies between the classical WS model and the AMS-fitted GEV since it is not influenced by single extraordinary large summer events like the WS model but takes into account the different distributions of the seasons.

Table 7.15.: Annualities of the highest observed flood peaks at the Wechselburg/Zwickauer Mulde gauge calculated with different statistical approaches.

HQ[m^3/s]	AMS (GEV)	Maximum approach WS	Maximum approach WST
544	22	25	27
587	27	30	34
633	33	36	42
915	95	90	117
991	120	109	145
1000	123	111	149

the smaller difference between short and long summer events in large catchments as mentioned above. Only for very high annualities or large floods a more distinct difference can be seen. Additionally to this detailed presentation we estimated annualities of the 2002 and 2013 event for all gauges with the different approaches. The results can be found in Table 7.16. The 2002 event can be classified as a short summer event, the 2013 event is a long summer event. Both events are the largest flood events since the beginning of the discharge measurements for nearly all gauges . Together with the return periods, estimated from AMS, the seasonal approach WS and the WST approach, the annualities of both events, derived from the statistics of summer events in three versions are shown. The GEV is used as the distribution of all summer events (column "Summer"). And also the WST maximum mixing model is used to combine short and long summer events (column "Summer WST"). The return periods of both events, estimated from the samples of long and short summer events and the filling approach described above are shown in the columns "Summer short" and "Summer long". A comparison of the return periods of long and short summer events shows that the short event 2002 would be an extraordinary event if the same peak would occur with a long summer flood. The flood 2013 has a smaller probability if we consider that it is a long flood. This extraordinary character becomes not evident if we estimate the return periods from the AMS, WS or WST models. Comparing the return periods of WS and WST, the WST approach results in annualities which are higher than in the case where the WS model is used. This results from the differentiation between long and short floods, reducing the exceedance probability of a flood peak belonging to one of both categories only.

Table 7.16.: Estimated annualities of the 2002 and 2013 events of the gauges in the Mulde basin using AMS, WS and WST approach, respectively. "—" indicates that no data are available for this event, ∞ is a number greater than 1 Million and too large to display

Gauge	2002-event (short)							2013-event (long)						
	AMS	WS	WST	Summer	Summer WST	Summer short	Summer long	AMS	WS	WST	Summer	Summer WST	Summer short	Summer long
Aue 1	111	104	148	136	207	223	1749	39	40	50	54	71	85	319
Niederschlema	52	54	57	74	86	180	140	22	24	25	34	36	74	60
Zwickau- Poelbitz	43	56	62	69	79	84	601	41	53	58	65	74	79	555
Harthau	64	56	66	62	82	81	2168	18	20	21	25	31	35	175
Goeritzhain	153	96	170	103	201	265	790	174	104	190	111	222	288	915
Wechselburg	123	112	157	124	156	274	312	120	110	153	121	152	269	304
Tannenberg	292	210	245	219	244	238	1940	22	28	30	43	45	50	172
Streckewalde	262	196	304	199	322	319	9752	56	59	83	63	89	100	536
Hopfgarten	202	184	259	223	295	364	1039	50	55	65	79	92	119	281
Rothenthal	547	269	271	308	305	299	∞	9	10	10	15	16	16	1621
Zoeblitz	1232	298	387	299	340	316	∞	27	26	27	28	34	34	401
Pockau	196	180	236	235	315	306	∞	22	25	29	41	60	61	1087
Borstendorf	396	259	426	314	657	612	701963	34	34	43	49	71	77	389
Lichtenwalde	220	217	260	270	409	1059	507	56	64	70	91	115	275	147
Niederstriegis	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Berthelsdorf	1218	512	669	276	360	315	3989	220	165	220	65	82	85	302
Nossen	726	303	398	303	390	539	859	92	79	97	79	97	172	164
Erlln	300	194	216	196	199	155	3421	–	–	–	–	–	–	–
Golzern	480	236	289	242	255	389	543	180	124	138	133	132	222	248
Bad Dueben	64	87	96	106	121	133	347	39	54	57	70	76	84	214

For all gauges the maximum approach WST results in a smaller slope of the distribution function than that of the distribution derived from the WS approach. In the maximum approach WS, the short summer events dominate the estimated skew of the whole summer annual maximum series and lead to a higher skewness of the whole distribution. By separating the summer events due to their timescales the high influence of the short summer events with large skewness is reduced. In comparison with the AMS based estimations, we see that both seasonal approaches are influenced in the lower quantile range by the winter series, but the higher quantile range is determined by summer events. An AMS based estimation seeks a compensation between both series. In some regions in Germany the most severe floods are rare summer events. Often this flood type is under-represented in an AMS. As a result, the estimation of their probabilities is strongly affected by distributions of winter events. The maximum approach WS is in this case a useful model to consider the differences between winter and summer floods explicitly. However, summer floods do neither constitute a homogeneous sample. The series of short events has in many cases a higher skewness than the series of long events. However, the number of short flood events and the distribution of their peaks is a catchment characteristic. Depending on the size of the watershed, high summer flood peaks can be caused by short events. As a result, the statistics of summer floods can be strongly affected in the upper quantile range by short events. The proposed WST approach considers this complexity and reduces the impact of single flood types by considering the probabilities of their occurrence.
We also want to present the results of a Goodness of Fit test of the statistical models. For this we compare the annual maximum discharges with the theoretical models. Again, we use the Anderson-Darling test and additionally show the results in QQ-plots. The p-values of the Anderson-Darling test are given in Table 7.17 for each gauge and all three approaches. It can be seen that the hypothesis of a fitting distribution cannot be rejected for either of the three approaches. Nevertheless, the GEV gives the highest p-values closely followed by the WST approach, whereas the classical mixing model WS seems to have problems in representing the data. A possible reason can be seen in the QQ-plots (Figs. 7.25 and 7.26).
We see that the largest event is overestimated in the AMS model, where it is mostly underestimated in the WS model. The WST model seems to find a balance here and also the quantiles in the area of 80-90% are estimated well, that is not the case for the WS model. For the Streckewalde gauge, where the p-value of the Anderson-Darling has the smallest overall value for the WST-model ($p = 0.166$) we can see in Figure 7.27 that this is mainly because of the underestimation of the right tail. In this case the WST-model seems to reduce the influence of the right tail too much.

7.2.2. Comparing robust estimators to non-robust ones in the hydrological context

We have already stated in Section 2 that the model uncertainty in flood statistics and the occurrence of extraordinary large events demands special needs of the robust estimators. Classical studies comparing robust estimators with non-robust ones may not be sufficient in this context, since they do not consider the above-mentioned features. A first step when using robust estimators in hydrology is therefore a simulation study with hydrologically fitting assumptions on the data to compare these estimators with common ones concerning robustness and efficiency. We choose the data basis to represent an annual maximum series, which is the most frequently used flood series. The results presented here are published in Fischer et al. (2015).
For the simulations we first analyse Gumbel-distributed data with location parameter $\mu = 100$ and scale parameter $\sigma = 10$, and in a second step GEV-distributed data with the same location

Table 7.17.: p-values of the Anderson-Darling test for a Goodness of Fit of the three models GEV, WS and WST to the annual maximum discharges for all considered gauges in the Mulde basin. The GEV gives the highest p-values closely followed by the WST model.

Gauge	GEV	WS model	WST model
Aue	0.976	0.565	0.941
Niederschlema	0.950	0.876	0.874
ZwickauPoelbitz	0.933	0.807	0.421
Harthau	0.994	0.944	0.983
Goeritzhain	0.974	0.654	0.312
Wechselburg	0.923	0.515	0.496
Tannenberg	0.853	0.892	0.833
Streckewalde	0.996	0.819	0.166
Hopfgarten	0.991	0.792	0.889
Rothenthal	0.993	0.558	0.286
Zoeblitz	0.492	0.649	0.649
Pockau	0.692	0.309	0.191
Borstendorf	0.949	0.522	0.639
Lichtenwalde	0.685	0.263	0.365
Niederstriegis	0.827	0.619	0.544
Berthelsdorf	0.799	0.886	0.787
Nossen	0.585	0.426	0.405
Erlln	0.841	0.533	0.705
Golzern	0.919	0.650	0.482
Bad Dueben	0.998	0.911	0.831

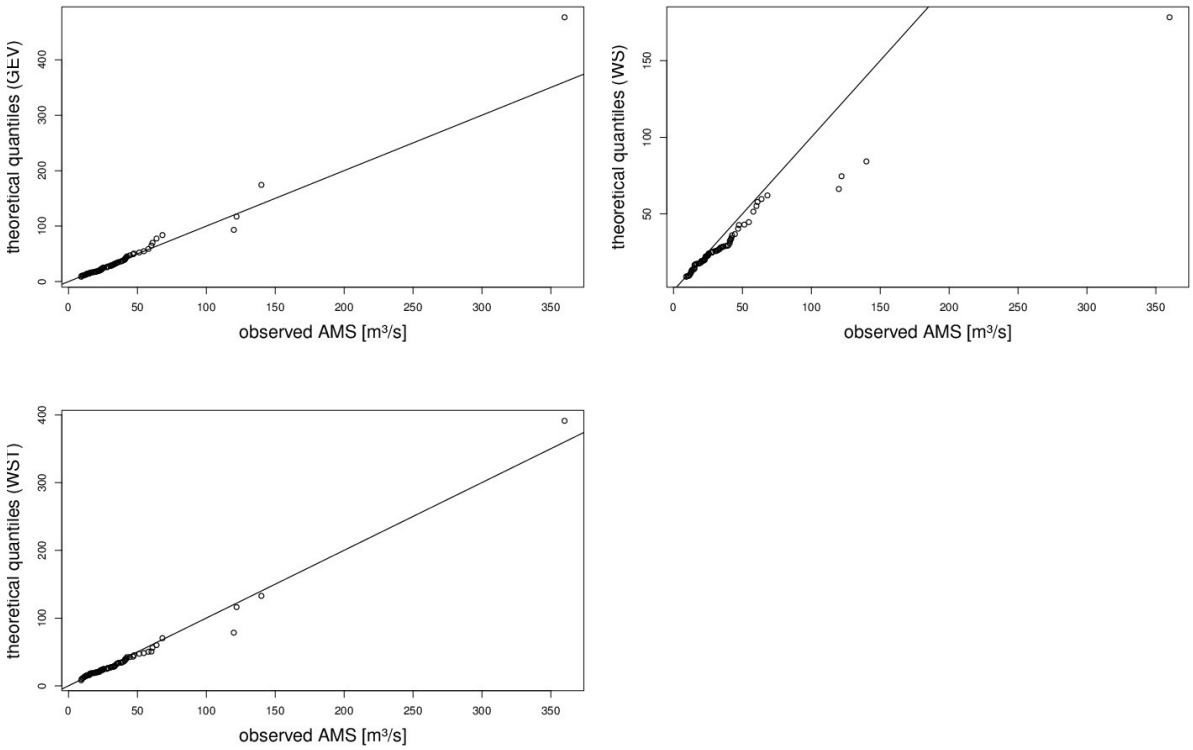

Figure 7.25.: QQ-Plot of the annual maxima of the Berthelsdorf/Freiberger Mulde gauge compared to the AMS, WS and WST model. The WST model shows the best fit, especially in the right tail.

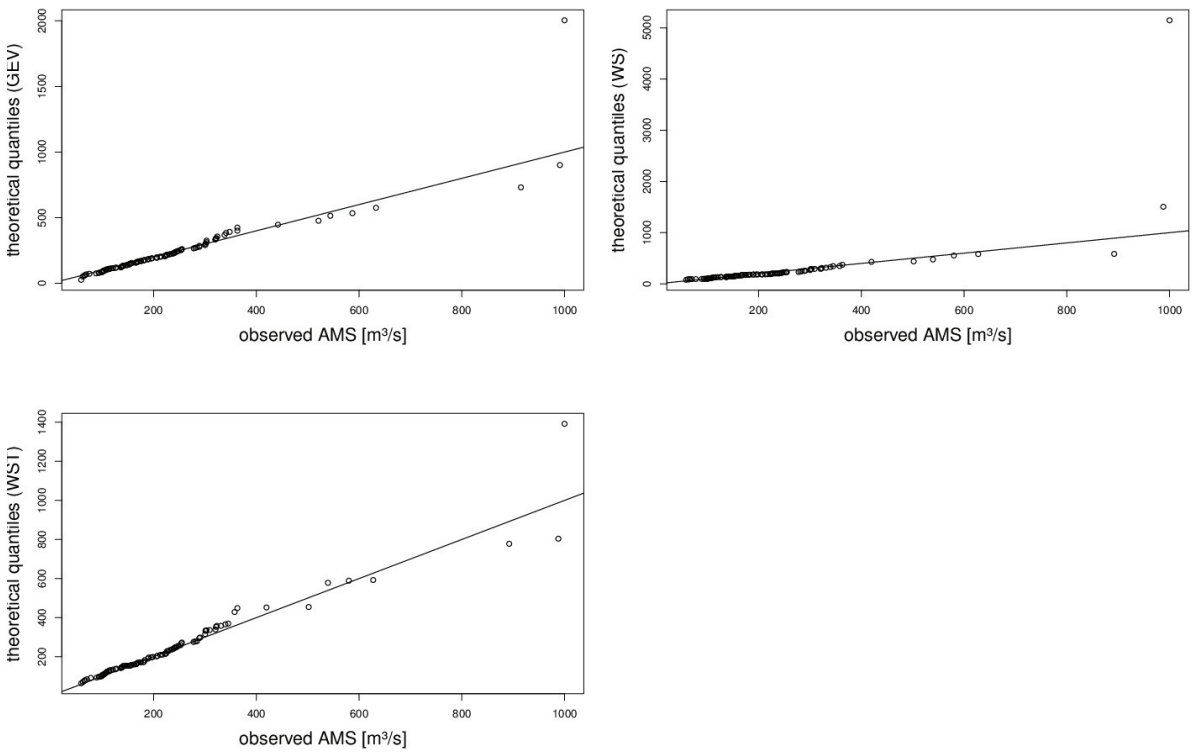

Figure 7.26.: QQ-Plot of the annual maxima of the Wechselburg/Zwickauer Mulde gauge compared to the AMS, WS and WST model. The WST model has the overall best fit, although also the AMS-fitted GEV and the WS model show good fits except of the one extraordinary extreme event.

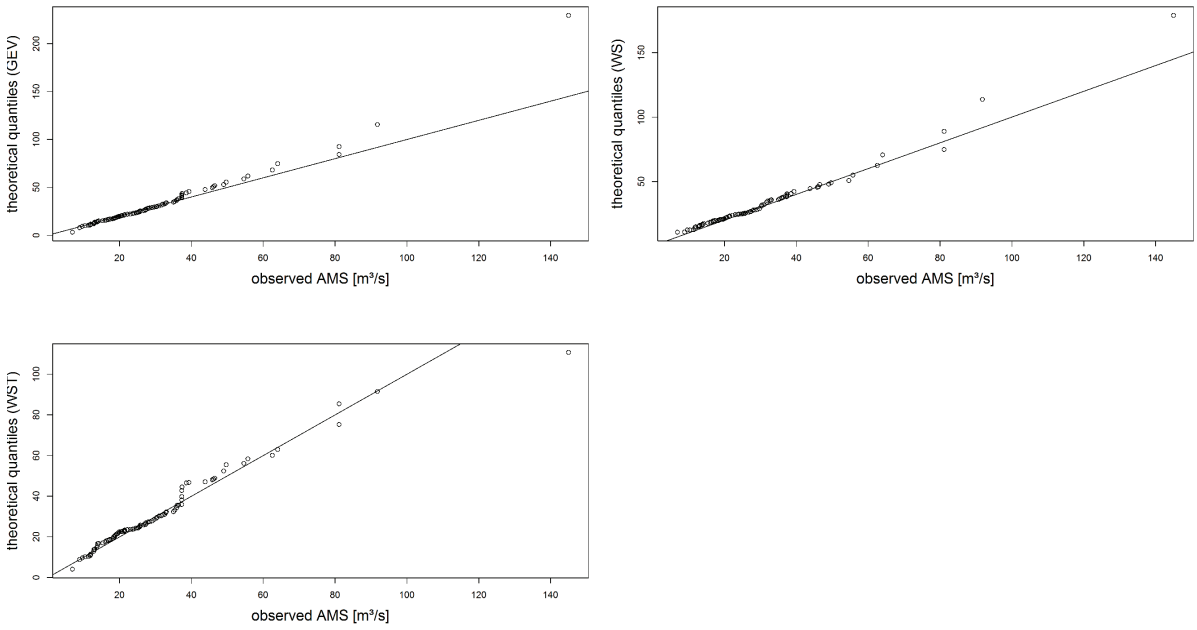

Figure 7.27.: QQ-Plot of the annual maxima of the Streckewalde/Preßnitz gauge compared to the AMS, WS and WST model. Here, the WST shows a strong underestimation of the right tail.

and scale parameter and shape $\xi_1 = 0.1$ (GEV1). These are representative choices fitting such distributions to maximum discharges (cf. Madsen et al., 1997). For comparison we also consider a larger shape parameter $\xi_2 = 0.2$ in the model, to which we refer as GEV2. This should mirror the possibility of extraordinary events. We seen in the section before that for distinct seasons even higher parameters can be sensible. To model the AMS this choice nevertheless should be sufficient.
As a justification for the choice of these parameters we fitted the GEV distribution via L-moments (described later on) to 33 series of annual maximum discharges of gauges in different river basins in Saxony in Germany. The histogram of estimated shape parameters of the GEV (rounded to one decimal figure) can be found in Figure 7.28. We can see some very large values for ξ, which indicate a significant deviation from the Gumbel distribution. To consider also the robustness against model-misspecification (see Section 2.2) we choose the three- as well as two-parametric distribution and fit the "wrong" model to the data.
In the simulations for each of the three distribution functions mentioned above two scenarios are considered. The independent, identically Gumbel respectively GEV distributed random variables are modified in one of the following ways:

1. No modification: independent identically distributed random variables.

2. We include extraordinary extreme values in these time series, which equal the 99.9%-quantile of the underlying distribution. For this, randomly chosen 2% of the data (rounded up) are replaced by the value of the 99.9%-quantile.

First of all, we fit both the Gumbel and the GEV distribution, respectively, to compare fits by

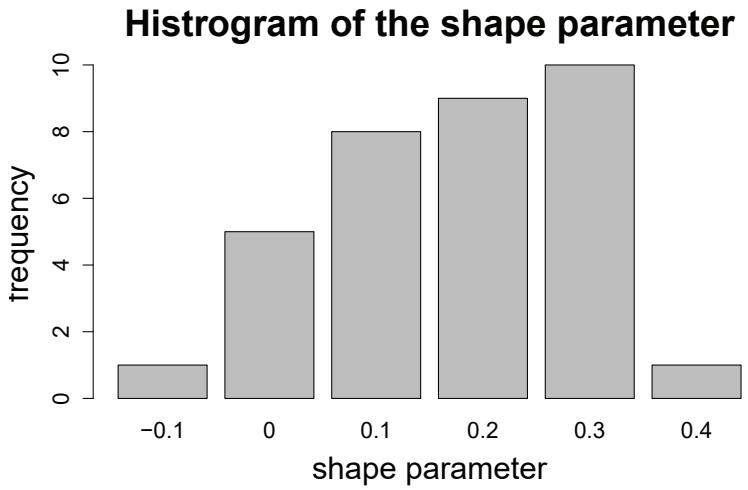

Figure 7.28.: Histogram of the estimated shape parameter for annual maxima of three river basins. The typical range of the shape parameter for the considered basins is between 0 and 0.3.

two and three parametric distributions. This is done by calculating the 99%- and the 99.9%-quantiles of the fitted distributions and considering the bias and the RMSE for the corresponding quantiles of the assumed true distribution. In both scenarios this true distribution is the one without modification, which has the following quantiles:

$$Q_{0.99;\mathrm{Gumb}} = 146.0,\ Q_{0.999;\mathrm{Gumb}} = 169.1$$

$$Q_{0.99;\mathrm{GEV1}} = 158.4,\ Q_{0.999;\mathrm{GEV1}} = 199.5$$

$$Q_{0.99;\mathrm{GEV2}} = 175.5,\ Q_{0.999;\mathrm{GEV2}} = 249.0.$$

In the simulation we consider 1000 repetitions for each of the different sample lengths equal to $n = 30, 50, 100, 200$. Annual series with a length of more than 100 years are very rare in hydrology and therefore an upper length of 200 seems to be sufficient.
We compare the following five different estimators.

- Maximum-Likelihood-Estimation
- L-moments
- Trimmed L-moments with asymmetric trimming (0,1)
- Trimmed L-moments with symmetric trimming (1,1)
- Minimum Distance estimator

  The Minimum Distance estimator for the parameter vector θ using the Cramer-von-Mises distance is given by

$$\hat{\theta} = \arg\min_{\theta} \int_{-\infty}^{\infty} (F_n(x) - G_\theta(x))^2 \, dx,$$

see Dietrich and Hüsler (1996). Here, F_n is the empirical distribution function of the sample and G_θ the distribution function with parameter(vector) θ to be fitted. In our case we need to minimize (for the GEV distribution)

$$\int_{-\infty}^{\infty} \left(F_n(x) - \exp\left(-\left(1 + \xi \left(\frac{x-\mu}{\sigma} \right) \right)^{-\frac{1}{\xi}} \right) \right)^2 dx \tag{7.8}$$

or (for the Gumbel distribution)

$$\int_{-\infty}^{\infty} \left(F_n(x) - \exp\left(-\exp\left(-\frac{x-\mu}{\sigma} \right) \right) \right)^2 dx, \tag{7.9}$$

which is done numerically.

Of course it is also possible to use other distance measures, for example the Hellinger-distance. Nevertheless, the results do not differ much so that we decide to use the easier to handle Cramer-von-Mises distance.

Remark 7.2. The third classical non-robust estimator, the standard sample moments, is excluded here, since the moment estimators do not exist for a shape parameter $\xi > 1/3$, and therefore do not seem to be suitable in this hydrological context, since there can be samples with a larger shape parameter as confirmed by Figure 7.28.

The choice of estimation methods includes classical non-robust estimators used in hydrology (ML-estimation and L-moments) as well as less known robust estimators. The bias and RMSEs for the estimated quantiles for all estimators and sample lengths are discussed in the following.

Evaluation for data without disturbances

In Tables A3, A5 and A7 in the appendix we can find the results for bias and RMSE of the estimated quantiles for the fitting to independent, identically distributed random variables, following a Gumbel, GEV1- or GEV2-distribution.
For i.i.d. Gumbel data (Table A3) we see that a Gumbel-fitting with Maximum Likelihood leads to the lowest RMSE for all sample sizes n and both quantiles considered here. Concerning the RMSE, TL(0,1) estimation is second best, followed by TL(1,1) and L-moments, which do not differ much. The Minimum Distance estimator performs worst, only in the case of $n = 100$ it is better than classical L-moments. The bias of all Gumbel-based fittings is very small, even for small sample sizes. Fitting a GEV-distribution to Gumbel-distributed data with non-robust estimators (ML and L-moments) roughly doubles the RMSE. Robust estimation (TL(0,1), TL(1,1), and MD) worsens the results even more as both the RMSE and the bias become larger for all n and both quantiles. Robust estimators aim at reducing biases which are due to using only approximately valid models, but it seems that in this case estimation error increases when fitting an unnecessary third parameter by one of these robust estimators.
In case of i.i.d. GEV1-distributed data with a shape parameter of $\xi = 0.1$ (Table A5), fitting a Gumbel distribution causes a substantial negative bias, which dominates the RMSE. Again, the L-moments give the best results, but the differences are not large, neither to the robust TL(1,1)-moments. It is striking that the RMSE is nevertheless much smaller for a Gumbel-fitting than for a GEV-fitting if the sample size is small, and about equal for large sample sizes. This is due to the much smaller variability of the two-parametric fits. A non-robust GEV-fitting causes a large

positive bias only for the 99.9%-quantile, but nevertheless the RMSE is large for both quantiles and all sample sizes due to the large variability. The results for the ML and the L-moments are very similar for large sample sizes ($n = 200$), though the L-moments deliver results with less deviation from the assumption in small samples. A robust fitting of a GEV-distribution leads, in comparison to the same estimators with a Gumbel-fitting, to a positive bias and higher variability and therefore to larger RMSEs for all sample sizes and both quantiles.
When the shape parameter is further increased to $\xi = 0.2$ (GEV2; Table A7) the previous remarks remain qualitatively valid. Fitting a Gumbel distribution to such data both bias and RMSE is nearly twice (2.5 times for the 99.9%-quantile) the one as before. Very striking are the results of the L-moment estimates for small sample sizes. They produce a large positive bias resulting in a large RMSE, which do not fit to the results for the other sample sizes. In case of at most 50 observations, fitting a Gumbel distribution with TL(1,1)-moments seems best, closely followed by ML, MD and TL(0,1). In case of at least 100 observations, GEV fitting becomes worthwhile with L-moments and ML performing best, followed by TL(1,1)- and TL(0,1)-moments.

Evaluation in the presence of extraordinary extreme events

In Tables A4, A6 and A8 in the appendix data from scenario 2 containing extraordinary extreme events are considered.
For Gumbel-distributed data and a Gumbel-fitting we see that the smallest values for the RMSE are given for the TL(0,1)-moments followed by the MD-estimator, which even has a smaller bias. For both estimators the RMSE is not much higher than in the case of data without disturbances. The non-robust estimators are substantially more biased, with the L-moments behaving worst. The RMSE of the ML-estimator is comparable to that of the TL(1,1)-moments, though ML has larger bias. If a GEV-fitting is used, the bias increases rapidly and therefore also the RMSE is large.
Somewhat surprisingly, the results are similar to this when the data follow a GEV-distribution with $\xi = 0.1$. The occurrence of extraordinary extreme events apparently reduces the negative bias and the RMSE of the estimations based on a Gumbel fit in this situation. Fitting a Gumbel distribution by L-moments works best for all sample sizes, followed by ML, TL(1,1)- and TL(0,1)-moments. For the 99.9%-quantile the difference becomes even larger. So the robust estimators have a higher RMSE and are no longer better than the non-robust ones, where the ML-estimator behaves best overall. Note that the results for a GEV-fitting are much worse. If the value of the shape parameter is increased, the RMSE and bias results increase by the factor 2.

Since it seems that the value of the shape parameter has a large influence on the estimation, we want to investigate this further. Therefore, we fit a Gumbel and a GEV-distribution to GEV-distributed data with increasing shape parameter and location and scale parameter as before. In the simulations above the L-moments proved to be the estimator with the smallest RMSE for the GEV-fitting for almost all sample sizes and are also not too bad for the Gumbel-fitting. That is why we choose this estimator for the calculations. The results can be found in Figure 7.29. We can see that for a shape parameter near zero the Gumbel-fitting has smaller RMSE for all sample sizes. For the hydrological relevant case of a positive shape parameter (we can seldom find negative shape parameter in maxima series) the Gumbel-fitting is even better for all values of ξ and sample sizes up to $n = 100$. Only for a sample size of $n = 200$ and $\xi \geq 0.2$ the GEV-fitting is preferable concerning the RMSE. We now want to have a closer look at the range, where the Gumbel-fitting is much better than the GEV-fitting, and also have a look at the bias there (Figure 7.30). It can be seen that the bias in the Gumbel-fitting increases with

increasing shape parameter, causing also an increase of the RMSE. The bias of the GEV-fitting remains stable near zero, though the RMSE increases. For increasing sample sizes both fitting approaches have more and more similar RMSE.

The simulation results show that the choice of the number of fitted parameters is in fact a crucial question. Surprisingly, we get better results for most cases, except the i.i.d. GEV case with $\xi = 0.2$ and $n \geq 100$, when fitting only a two-parametric distribution, regardless of whether there are rare events or not. This is confirmed by Figure 7.29, where we also see that for increasing shape parameter and increasing sample size the GEV-Fitting is better and should be preferred for a sample size of at least $n = 100$ and $\xi \geq 0.2$. Remembering the recommendations given by the DWA (DWA (2012)) (see Section 2), recommending a two-parametric distribution function for samples with less than 30 years and a three-parametric one for samples with more than 50 years, the results confirm the recommendation and we can add that for estimated values for the shape parameter smaller than 0.2 and large sample sizes also a two-parametric fitting should be preferred and also the range of the sample size for a two-parametric fitting could be increased to $n = 50$. Of course, it has to be considered that the parameter value is estimated and not the theoretical value. Thus, these limit cases should be investigated further. Concerning the non-robust estimators the results of Hosking et al. (1985a) are confirmed. For small sample sizes the L-moments in the GEV-fitting have smaller RMSE than the ML-estimator. Nevertheless, it becomes obvious that the size of the shape parameter plays a crucial role. The larger the shape parameter the more differ the RMSEs of these two estimators. Fewer observations are needed to make the ML superior to L-moments when just two parameters need to be estimated as in the case of a Gumbel distribution.

When extraordinary or rare extreme events occur in our data, the robust Minimum Distance estimator or the Trimmed L-moments offer a smaller bias and RMSE for the higher quantile, but they have the disadvantage of having larger RMSE compared to the classical estimators ML and L-moments, when no such extraordinary extreme events occur. This is the common lack of efficiency of many robust estimators, particularly in small samples. Among the robust estimators considered here, the TL-moments are preferable, since they have the smallest RMSE when extraordinary extreme events occur and not too large RMSE when there are no such events - especially for estimation of the higher quantile. Of course also other trimming is possible and could be investigated further, but we will see later on that a trimming of 1 (symmetric or not) is sufficient in our case of application.

Based on the results given, a two-parametric fitting with the Maximum-Likelihood estimator is recommended for a sample size smaller than $n = 100$ whereas a three-parametric fitting using L-moments seems to be preferable otherwise. The Trimmed L-moments seem to be a recommendable choice when extraordinary events occur in the sample. They do not have a tendency of overestimating the quantile, even if the return period of the event is much larger than the sample length. Additionally, they seem to be rather efficient, which is inherited from the ordinary L-moments. The Minimum Distance estimator considered here is apparently not efficient enough for the small sample sizes given in hydrology. If there are no obvious extraordinary events in the sample, L-moments are recommended having the largest efficiency for the small sample sizes occurring in hydrology.

Moreover, further hydrological phenomena can occur. Therefore, we also investigated another scenario, representing an uncertainty in the measurement of the data. We wanted to model the situation, when a rating curve does not consider the overflowing of river beds (too small discharge values are assumed) or backwater in river increases the water level (too high discharges are assumed). This is done by cutting off the 20% highest data of the simulated distribution and

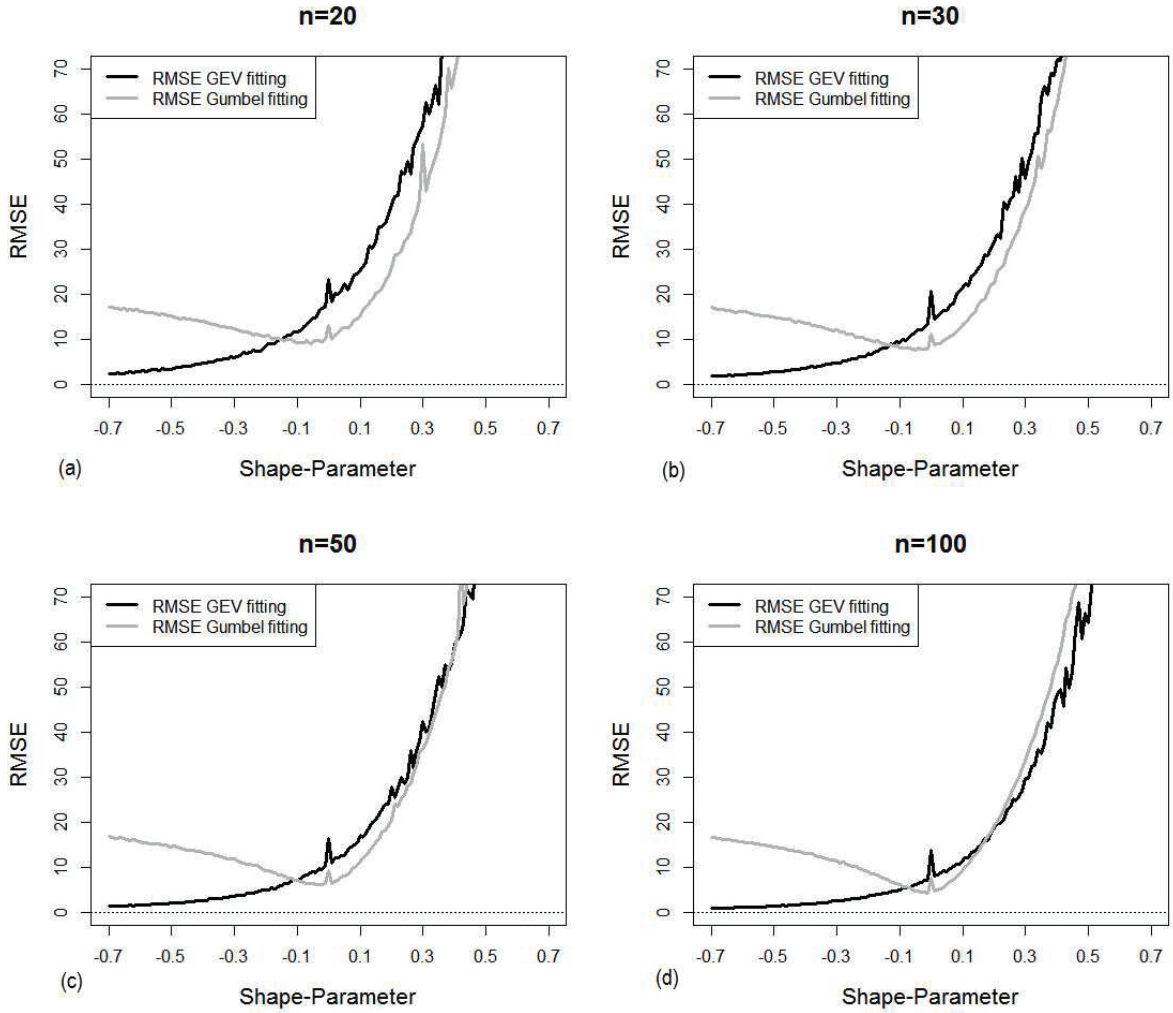

Figure 7.29.: Fitting of the GEV (black) and Gumbel (grey) distribution to a GEV distribution with increasing shape parameter via L-moments for different sample lengths ($n = 20$ (a), 30 (b), 50 (c), 100 (d)). For values of the shape parameter smaller than -0.1 the RMSE of the GEV fitting is much smaller. For small sample sizes this changes at that point such that the Gumbel-fitting has much lesser RMSE whereas for large sample sizes ($n > 30$) Gumbel- and GEV-fitting are nearly the same concerning the RMSE.

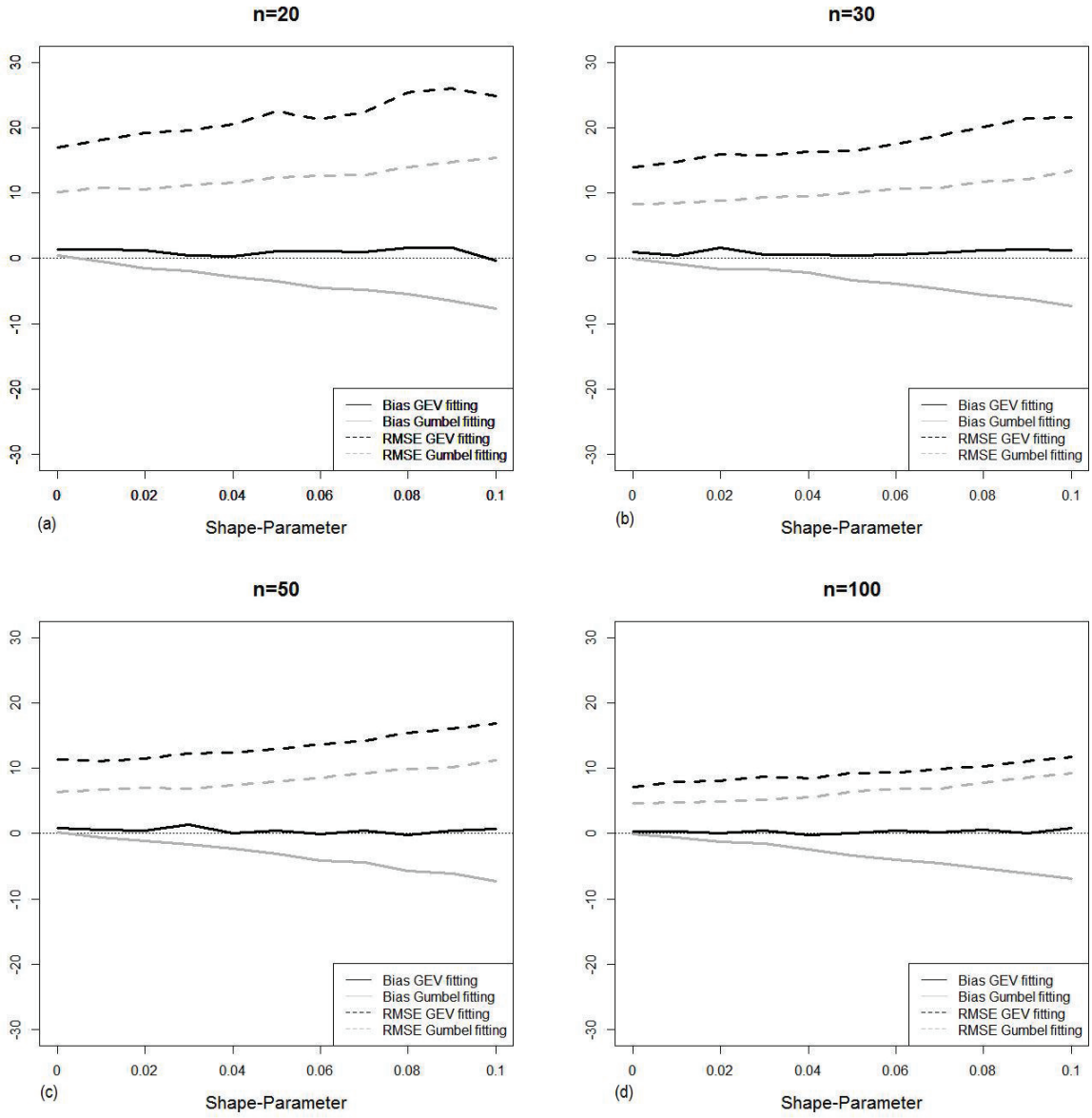

Figure 7.30.: Bias and RMSE of the Gumbel and GEV fitting to i.i.d. GEV-distributed random variables with varying shape parameter and different sample lengths. For all cases the Gumbel-fitting has the smaller RMSE but the higher Bias.

replacing them by data representing an error between 0 and 30%. For this scenario all estimators failed and are not able to cope with the misspecification of the model. Therefore, detailed results are omitted. Also the case of a mixing distribution is considered (see eq. 7.3) here. If in this case a simple GEV or Gumbel distribution is fitted, all estimation methods fail. Therefore, it does seem to be possible to estimate the underlying model with these classical methods.
In the recommendations made above the question arises, how to detect extraordinary events. A possibility to make this choice could be the sensitivity curve. For non-robust estimators one could use this tool to examine the influence of a single event to the estimation and therefore decide, whether to use a robust estimator or not. We have already seen in Section 2, how this could be done.
Additionally, the results give the impression that the choice of the number of parameters is crucial and should depend on the sample size and on the value of the shape parameter. The recommendations of the DWA (2012) are confirmed for the scenarios considered here but we have by far not covered all relevant cases where a two-parametric distribution function might be preferred.

7.2.3. Application of robust estimators on the estimation of annualities

After showing the potential of the use of robust estimators in the context of flood analysis we want to apply robust estimators to the hydrological models introduced in Section 7.2.1.

Robust POT with TL-moments

By the application of a robust estimator for parameter estimation of the GPD in the POT model, the influence of extreme events is reduced in two ways: by using more information (more data) resulting in a downweighting of the influence of extremes, and by a robust estimation of the parameters themselves. As a robust estimator we choose based on the results of Section 7.2.2 the trimmed L-moments (TL-moments), the robust extension of the classical L-moments (Elamir and Seheult (2003)). Here we apply a trimming of the upper part of the sample only by giving zero weight to the most extreme value, in our case the highest value (TL(0,1)-moments), since the distribution is bounded in the lower tail by a threshold. Because of the small degree of trimming, we do not loose much information about the extremes. For details and a concrete derivation of the TL(0,1)-moments for the GPD, see Section 6.1. TL-moments have not been commonly used for the analysis of flood data, but in the few cases where the approach has been applied (cf. Asquith (2007)) and in the studies above, it has been found to be promising. We will refer to the combination of POT with TL-moments in the following as the robust POT with TL-moments (robust POT) approach. Most of the results presented here are published in Fischer and Schumann (2016) and are supplemented by new results concerning a robust estimation of the AMS.
The choice of trimming is a crucial question, especially in the presence of several extraordinary events. Therefore, we not only considered TL(0,1)-moments but also a higher trimming in the upper part of the sample using TL(0,2)-moments (for details see section 6.1). The results for TL(0,2)-moments showed that a higher trimming does not result in higher stability and therefore no more robust quantile estimates. However, the trimming TL(0,2) resulted in a prolongation of the period of stabilization for the estimated quantiles, indicating a lower efficiency. Considering these results, then, we applied a trimming degree of (0,1).
To obtain a fair comparison we also want to compare the robust POT with TL-moments approach with the classical AMS approach based on fitting a GEV distribution to the data, where we use

robust estimation of the parameters with the TL(0,1)-moments, too (see Section 6.1.2). By this we can compare the advantage (or disadvantage) of robustness, of POT and a combination of both.
To compare the different approaches AMS, AMS (robust), POT and robust POT with TL-moments, we analyse flood series from 15 gauges in Germany, located in the Mulde river basin in Saxony (Figure 7.1) with an observation period length of at least 75 years and floods recorded up to the year 2013. The catchment size varies between 75 and 5433 km^2. All data series contain several extraordinary extreme events. One of these events occurred in August 2002.
We are aware that there could be some dependence in the time series when considering monthly maximum discharges. Nevertheless, we recognized in our calculations that the estimators used are able to cope with slight dependencies of monthly maxima. For the considered monthly maximum discharges it is also ensured that they do not result from the same flood event. That is, if the maximum occurs at the beginning of a month and the month before the maximum is at the end, it is ensured that these peaks do not belong to the same flood event. In this case for the second month we chose the maximum of a second event.
Additionally, the choice of a discharge threshold specifying monthly maxima as flood peaks reduces these dependencies to a negligible amount. As a threshold to specify floods among the monthly maxima, we choose the minimum value of annual flood peaks of each series. For all gauges we compare POT and robust POT with the classical AMS concerning robustness and Goodness of Fit. For this purpose, we fit a GEV distribution to the AMS sample

$$\left(\max_{1\leq k\leq d}(X_1^{(k)}), \ldots, \max_{1\leq k\leq d}(X_n^{(k)}) \right)$$

with L-moments and TL(0,1)-moments and the POT model to the whole sample $(X_1, \ldots, X_n)$ using the L-moment estimation and the TL(0,1)-moments.

To characterize the overall agreement between the estimated distribution $\hat{F}_l$ at site l with sample length n_l and observations $X_i(l)$, $i = 1, \ldots, n_l$, the index pval (Renard et al. (2013)) is used. This reliability index is calculated as:

$$\text{pval}_i(l) = \hat{F}_l \left(\max_{1\leq k\leq d}(X_i^{(k)}(l)) \right).$$

Under the hypothesis of a reliable estimation ($\hat{F}_l = F_l$) the vector $(\text{pval}_i(l))_{i=1,\ldots,n_l}$ is uniformly distributed on the interval $[0, 1]$ for every gauge l. We estimate the annual distribution function $\hat{F}_l$ via the different approaches (AMS, POT and robust POT with TL-moments) for a sample of a length of 50 years and apply it on the annual maxima at the gauge. By application of the inverse distribution function, a QQ-plot can be used as a graphical tool to demonstrate the Goodness of Fit of pval for each fitting approach (Figure 7.31). Often also a probability-probability-plot is applied, but since a QQ-plot illustrates the extreme domains in a better way, we choose this presentation.

We can see that all three approaches give a good estimation of the distribution of the whole sample. In the area of the upper quantiles, POT seems to give the best Goodness of Fit, and all in all it does not differ much from the AMS statistic. For both robust approaches, robust AMS and robust POT (both using TL-moments), a larger deviation exists for some gauges. Nevertheless, the deviation of all four approaches does not seem so large that we have to reject any of them. We are aware that the criterion pval could be affected by single extreme observations, which are

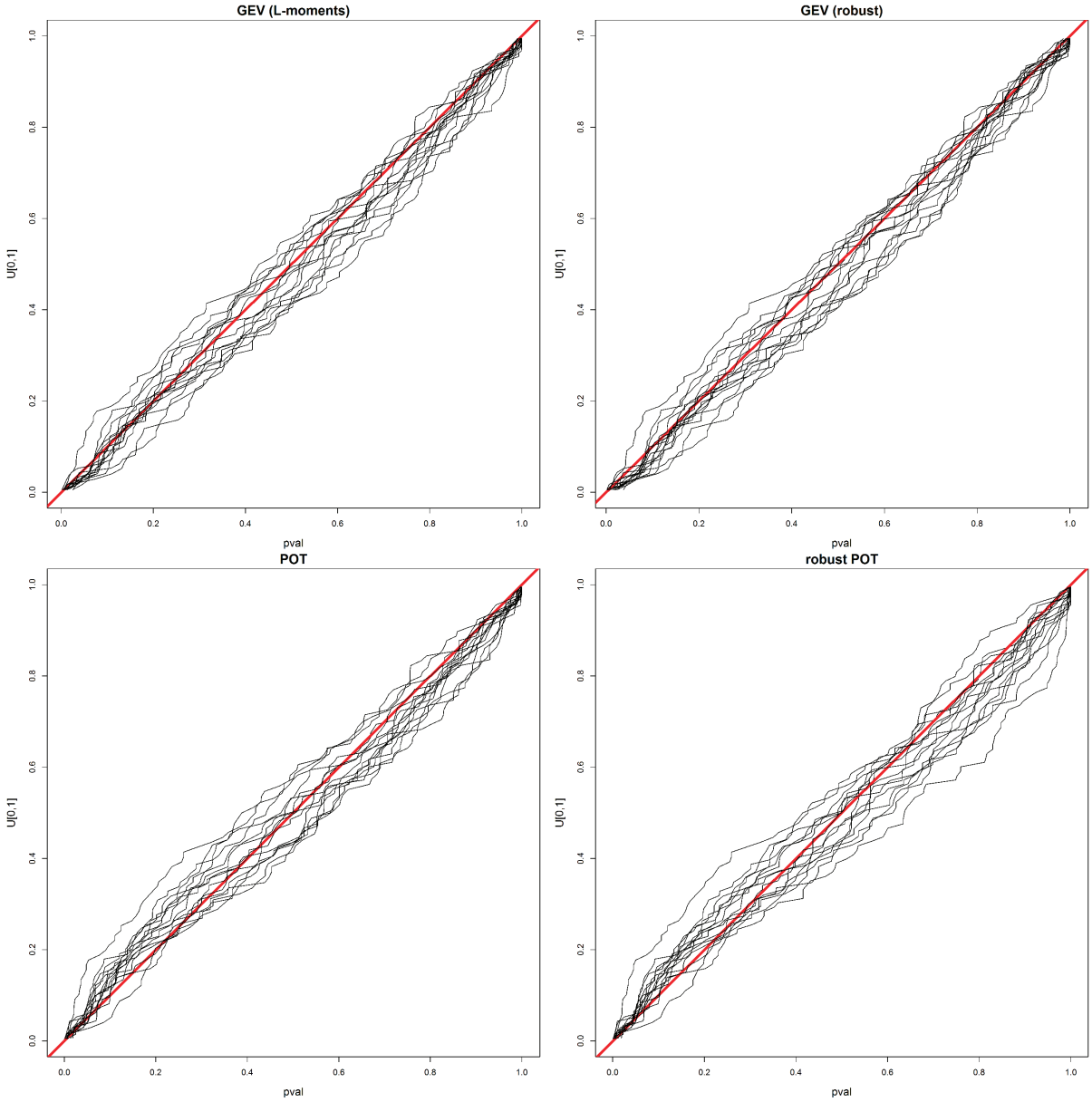

Figure 7.31.: QQ-plot of pval for AMS (top left), AMS (robust with TL-moments) (top right), POT (bottom left) and robust POT with TL-moments (bottom right) approaches for gauges $l = 1, \ldots, 15$. All four approaches show a good fitting to the data with the robust POT with TL-moments being the one with the largest differences in the fit.

most relevant for the fit of $\hat{F}_l$, but in the same way for the evaluation of this fit. To consider this we also applied cross-validation, that is

$$\text{pval'}_i(l) = \hat{F}_{l;(i)}\left(\max_{1\le k\le d}(X_i^{(k)}(l))\right),$$

where $\hat{F}_{l;(i)}$ is the fitted distribution function for gauge l based on all data except $\max_{1\le k\le d}(X_i^{(k)}(l))$. The results we obtained from this approach are very similar to those of the original pval, so we can expect that the above-mentioned problem did not influence our results substantially. To investigate the fitting in the higher quantiles, where a difference in the fitting seems to be largest, we look closely at the QQ-plot (Figure 7.32) for the Nossen gauge. Here, it becomes evident how much the classical AMS model is influenced by the single extreme event, leading to a worse fitting to the higher quantiles. The POT model uses additional data values, which are mostly located in the lower domain. That is why it has a better fitting in the central quantiles. Thus, the larger information spectrum leads to a more stable fitting. When using robust estimators for the AMS approach the influence of the extreme large values is reduced and a better fit is obtained for most of the observed data except that of the right tail. If we combine the POT approach with the application of the TL(0,1)-moment estimator, we can see a good fitting of all quantiles except of the single extreme event.
In the next step, the criterion $SPAN_T$ (see Section 2.3.3) is calculated for all 15 gauges for quantiles with annual return periods of 100, 200 and 1000 years. In Figure 7.33 the empirical distributions of these vectors are plotted. The GEV approach with L-moment estimators is most sensitive to changes in the data, having clearly higher $SPAN_T$ scores for each of the three annual return periods. In general, the maximum value of $SPAN_T$ increases with increasing T, indicating that the estimates become more sensitive for higher quantiles. This makes sense due to the small quantity of data in these high domains. The robust POT seems to be advantageous in the lower quantile areas, being much smaller than the other two approaches. In general, the main part of the calculated $SPAN_T$ values for the robust POT with TL-moments approach are smaller, and therefore it is the most robust approach in this sense.
To have a closer look at these results, we analysed for one gauge (Nossen) values of $SPAN_T$ that are derived from Monte Carlo simulations. We drew 1000 random samples of $n = 50$ years (for the POT approach this means 50 times 12 values per year) from the observed series. The 99.9%-quantile, derived from the different models, is compared via $SPAN_T$ with the quantile of the remaining 37 years of this series. It becomes evident that the robust POT approach with TL-moments is most robust, whereas the classical GEV model with L-moments estimates is worst. We considered also the GEV model with Maximum Likelihood (ML) estimates, which is surprisingly more robust than the GEV-L-moment model. Nevertheless, for single gauges it resulted in the highest values of $SPAN_T$ (≈ 2), indicating that for distinct random samples where high flood peaks are concentrated in one subperiod, the estimated quantiles differ significantly from estimates from the other subsample. For samples in which the values within the two subperiods are similarly distributed, the high efficiency of the ML-estimator leads to small values of $SPAN_T$. Overall, this leads to smaller $SPAN_T$ values than for the GEV-L-moment approach. In total, the robustness of the POT and robust POT approaches is less affected by the order of occurrence of the events in a data series.

Remark 7.3. A comparison with the third classical estimation method, the method of moments, is not possible since the random samples can lead to a shape parameter $\gamma > 1/3$, where the third moment and therefore the moment estimators no longer exist.

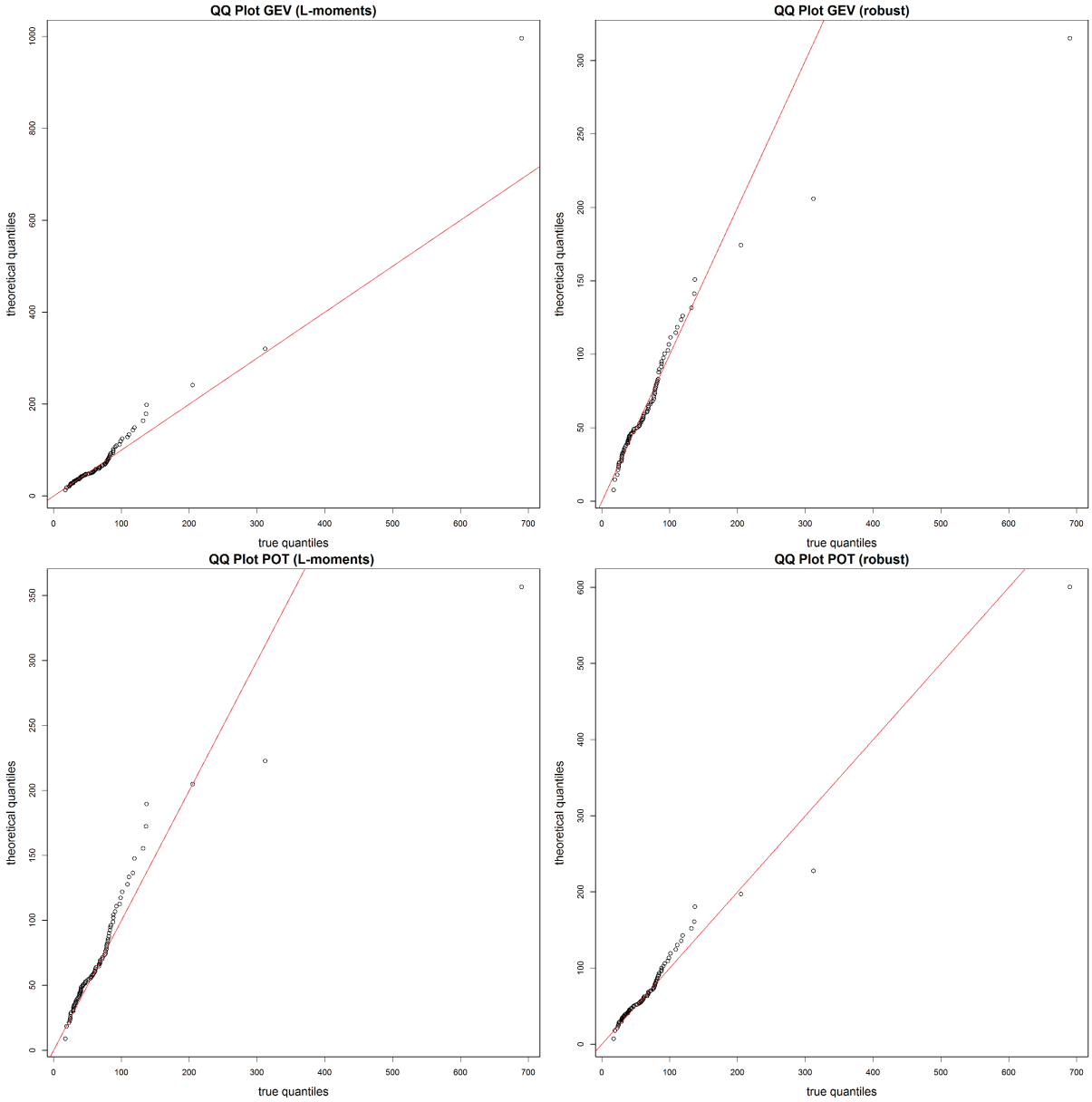

Figure 7.32.: QQ-plots for the annual maximum discharges of the Nossen gauge and the estimates using AMS (top left), robust AMS with TL-moments (top right), POT (bottom left) and robust POT with TL-moments (bottom right). The robust POT with TL-moments approach fits best to the main part of the data since it is not influenced by the extraordinary extreme event.

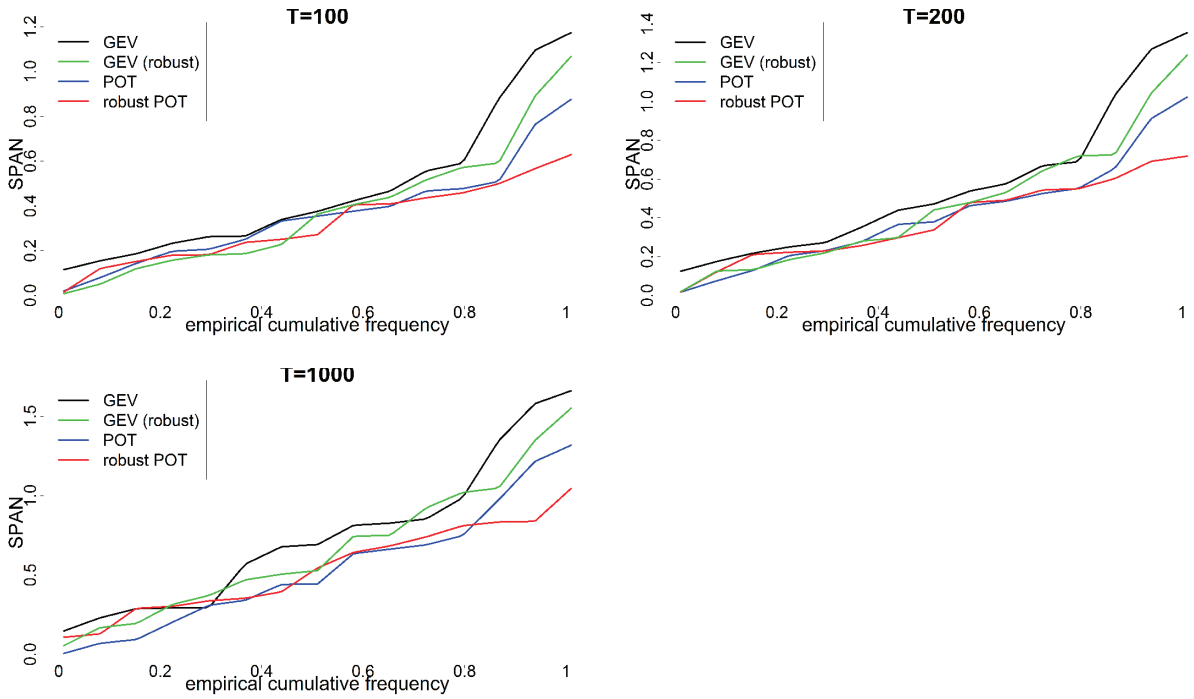

Figure 7.33.: Empirical distribution of $SPAN_T$ for annual return periods of $T = 100$ (top left), $T = 200$ (top right) and $T = 1000$ (bottom) years and all gauges. If the estimation is stable, the relative distance of the quantiles has to be close to zero and therefore also $SPAN_T$, consisting of the values of all single gauges, should be close to zero. The smallest values are obtained for the robust POT with TL-moments approach.

Since we are interested in the robustness of quantile estimations in prolonged series, the third criterion we used compares the annual return periods of yearly values, which are added step-by-step to the analysed data series. Here we estimate the absolute deviation between the return period of an event, which is estimated on the basis of a series ending one year before it, and the return period of the same event after integrating it into the analysed series of observations. The first value, using all previous events, is the "predicted return period"; it is compared with the return period using the prolonged series, which we name as the "observed return period". This approach is applied for an increasing sample length, starting with a minimum of ten years. The differences between predicted and observed return periods depend on the length of previous observations, the statistical characteristics of the time series and the size of the added value. Thus, it is not comparable between gauges but is a means to compare the robustness of quantile estimators. The evolution of quantiles can be compared with the annual return period of the most extreme floods estimated from the whole sample by the AMS approach.

We analyse the robustness of quantiles of year-by-year increased time series for two gauges (Nossen and Wechselburg) and estimate the absolute deviation between the predicted return period and the observed return period (Tables 7.18 and 7.19). It becomes evident that after a certain time of stabilisation, which is needed for an adequate modelling, the POT and especially the robust POT approach with TL-moments provide much more stable estimates of the annual return period than the classical annual maxima approach AMS. Whereas the AMS approach is highly influenced by the occurrence of extreme events, the two POT models are less responsive. Additionally, the use of robust estimators in the AMS approach also leads to more stable estimation, but it is not as stable as the combination of the POT model with the TL(0,1)-moments.

At Nossen gauge the flood in the years 2002 has been so extreme that all methods estimate return periods much higher than 10,000 years. An interesting property of both robust methods also becomes evident. In the year 2013 another extreme event occurred. Because of the very short time span since the previous extreme event in the year 2002, the non-robust AMS approach estimates a higher quantile (already estimating a very high quantile in the step before), fitting well in this case. The robust AMS as well as robust POT handled the former extreme event as a type of outlier. When such an extreme event occurs a second time, however, some probability mass is given to this large flood. Robust AMS, POT and robust POT result in a larger difference now. Comparing the evolving 99%-quantiles from a series with an increasing length (Tables 7.20 and 7.21) with estimations based on the total series the deviations from the quantiles of the total series are smaller for the robust approaches and for the POT approaches. In most of the cases, the estimated return periods are reduced when the total series is used and are often close to the robust estimations. This shows the difficulties when estimating high quantiles from short data series with extreme events. When the series length increases, the estimation changes considerably. The use of robust estimators as done in the robust AMS model leads to a higher stability of the estimations in this case. The increase of data by using for example the POT increases this behaviour.

The behaviour discussed above leads to the assumption that the robust POT with TL-moments results in an estimation of high quantiles that is stable over time and is less influenced by single extreme floods. Nevertheless, very extreme floods are still identified as extremely rare events. To demonstrate this, we compare the 99%-quantiles of flood series with increasing sample lengths. As before, we started with the first 10 years of observations and estimated the 99%-quantile from this sample. Then the sample length is increased step by step by one year and the 99%-quantile is estimated. This is done up to the point where all recorded floods are included, in our case the year 2013. Plotting these results, the typical shape of the time series of quantiles is a sawtooth

Table 7.18.: Absolute differences between predicted and observed return period of extraordinary extreme floods (Gauge Wechselburg/Zwickauer Mulde, records beginning in the year 1910) with the four different fitting approaches. The return period of the observed flood is estimated with the time series ending in the years before the occurrence of this event. Only differences of more than 5 years are shown. The character ∞ stands for cases where it is not possible to fit a suitable distribution to the given data

Series used for prediction	Year of the event	AMS	robust AMS	POT	robust POT
1910-1922	1923	3	7	34	∞
1910-1923	1924	66	18	57925	∞
1910-1925	1926	10	10	68	∞
1910-1931	1932	91	8	75	45
1910-1953	1954	354	33	444	88
1910-1974	1975	19	15	16	3
1910-2001	2002	172	99	124	38
1910-2012	2013	71	78	70	49

Table 7.19.: Absolute differences between predicted and observed return period of extraordinary extreme floods (Gauge Nossen/Freiberger Mulde, records beginning in the year 1926) with the four different fitting approaches. The return period of the observed flood is estimated with the time series ending in the years before the occurrence of this event. Only differences of more than 5 years are shown. The character ∞ stands for cases where the model is not able to fit a suitable distribution to the given data

Series used for prediction	Year of event	AMS	robust AMS	POT	robust POT
1926-1930	1931	30	0	1	∞
1926-1953	1954	18	4	6	4
1926-1957	1958	750	92	65	29
1926-2001	2002	19.5 Mio	10247	2.7 Mio	29608
1926-2012	2013	31	202	127	353

Table 7.20.: Flood peaks (in m^3/s) with a return period of 100 years at Wechselburg/Zwickauer Mulde gauge, derived from a series with a growing length, and mean absolute deviation (MAD) from the results of the total series with five different approaches.

Series	AMS (ML)	AMS (L-moments)	robust AMS	POT	robust POT
1910-1940	1516	687	1013	690	824
1910-1950	649	613	874	619	728
1910-1960	814	792	826	717	716
1910-1970	755	714	711	684	702
1910-1980	829	766	823	725	729
1910-1990	753	713	718	717	733
1910-2000	757	720	732	709	726
1910-2010	822	800	785	754	726
1910-2013 (all)	953	919	868	791	708
MAD	231.9	193.4	95.4	89.1	34.5
MAD (%)	24.3	21.0	11.0	11.3	4.9

Table 7.21.: Flood peaks (in m^3/s) with a return period of 100 years at Nossen/Freiberger Mulde gauge, derived from a series with a growing length, and mean absolute deviation (MAD) from the results of the total series with four different approaches.

Series	AMS (ML)	AMS (L-moments)	robust AMS	POT	robust POT
1926-1956	170	165	193	192	206
1926-1966	241	206	259	225	239
1926-1977	242	195	241	203	211
1926-1986	210	180	226	188	190
1926-1996	193	173	212	171	196
1926-2006	280	305	211	232	190
1926-2013 (all)	316	337	234	251	225
MAD	93.3	133	21	49.2	14.3
MAD (%)	29.5	39.5	8.97	19.6	7.4

curve as shown in Figure 2.1. Every large event causes a jump in the quantile in the year of its occurrence. Afterwards, during periods with "normal" floods, these differences decrease slowly until the next large event occurs. Overall, we normally have a slight increase of the values of quantiles with growing sample length as more extreme events become more probable with the increase of the length of observations. An example of such a moving quantile estimation is shown in Figure 7.34 for the Wechselburg and Nossen gauges. For the classical AMS approach (GEV), we see the typical "sawtooth" curve as mentioned above. Every large event leads to an abrupt increase of quantiles. The jumps are reduced much when using the robust TL-moments, although still the influence of single events can be detected. The jumps are reduced to a similar height, however, when POT is used. Although there are still jumps in the estimated values, their magnitude is much smaller, and all in all we have a smaller variability of the quantiles and a smoother increase. For robust POT, when a robust parameter estimation with TL-moments is additionally used, we can see that the estimation is no longer affected by any jumps. However, at the beginning of the time series, the variability of quantiles estimated from the robust POT approach is high, an effect caused by the low efficiency of the TL(0,1) estimator for small samples, which is a typical characteristic of robust estimators.

One can also see that all four approaches need a sample length of at least 30 years to deliver stable results, the robust AMS even longer since it has much less data available than the POT approach.

Additionally, we analysed the flood series for the other 13 gauges and calculated the coefficient of variation of the series of quantile estimations for the increasing sample lengths starting with 40 years (Figure 7.35). As expected, the coefficient of variation for the POT is generally smaller than if we apply the classical AMS approach. The coefficient of variation for robust POT with TL-moments is a bit higher than the others. This results from the needed time for stabilisation of the estimations. When a certain minimum length of the given series is ensured, the variation of results for robust POT decreases on average faster than that of the other approaches. Small quantiles are less affected by extremes than high quantiles. This is shown in Figure 7.36 with the examples of the 80% and 99.5%-quantiles of the whole samples. Since not the distinct gauge but the overall behaviour is of interest here, we spare the names of the gauges and use numbers instead. To compare the results, the ratios of the estimated floods between the AMS and the robust AMS with TL-moments, the POT and robust POT with TL-moments are used. One

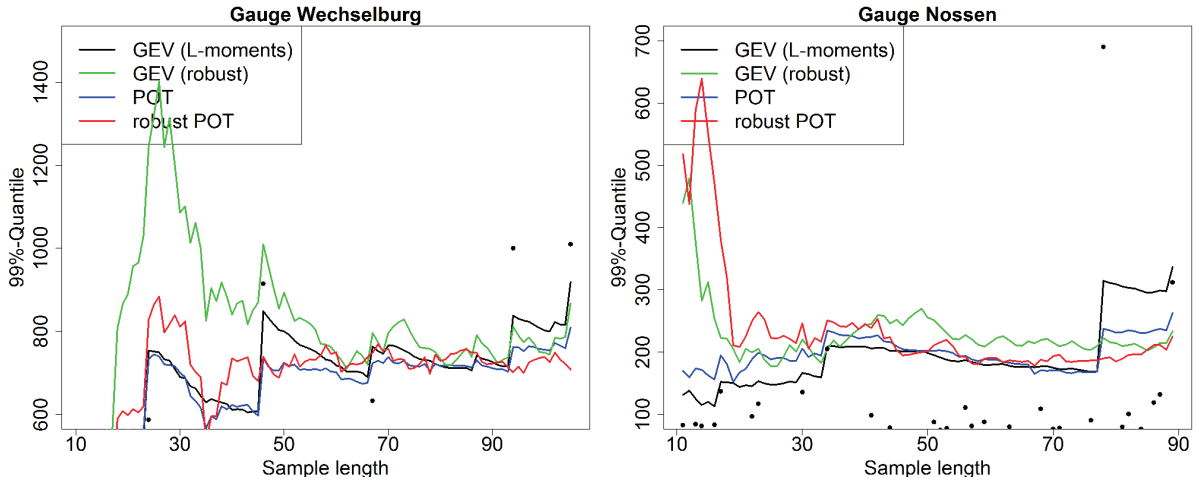

Figure 7.34.: Estimation of the 99%-quantile of the Wechselburg (left) and Nossen (right) gauges for growing sample length with four different approaches. The black dots are the annual maxima of the single years. After a time of stabilisation the robust POT with TL-moments approach is the most stable one and does not change much when an extraordinary extreme event occurs, whereas the non-robust GEV approach shows large jumps in these cases.

can see that the results of AMS estimations for a return period of $T = 200$ years are in general higher than the POT results (ratios larger than 1). Averaging all 15 gauges, this ratio is about 1.3. This means that on average, the quantiles derived from AMS are 30 percent higher than using POT. When using the robust TL-moment estimators (robust AMS) the estimated 80%-quantile differ a lot from the ones estimated with the classical AMS approach, but they show no distinct direction of deviation. For some gauges, the AMS quantiles are higher, for others this is not the case. Using robust POT, these differences to the AMS have a clear direction. Here the flood quantiles are on average 50% smaller than those of AMS. The differences among quantiles with lower return periods, e.g. $T = 5$, are much smaller, regardless of the estimation approach. So, using robust estimators as well as an enlargement of the data basis only affects the higher quantiles by giving less weight to the influence of extreme events. This supports the results concerning the reliability of POT or robust POT with TL-moments.

As has been shown in Figure 2.1 and 7.34, flood quantiles change with the length of the flood series and the occurrence of extreme floods. The robust POT approach (after a certain period needed for its stabilisation) results in a low variability of extreme quantiles. Compared to the AMS-estimated quantiles derived from the total series of observations, these "final" quantiles are much smaller, since the extreme flood in the year 2013 has less impact (Figure 7.36). However, the robust POT quantiles for the series up to the year 2013 are similar to the estimated quantiles of all shorter recorded series (Figure 7.34). This is not the case for the AMS approach. To show the differences in the results between the two approaches, we calculated the frequencies of the 99% flood quantiles for year-by-year increasing series of observations starting with a series of at least 40 years of observation and ending in the year 2013 (Figure 7.37).

The results are shown by boxplots of both methods for two gauges. The flood range where the highest frequencies are concentrated is quite similar for both methodologies, but robust POT avoids very high estimations of these quantiles. For robust POT, all estimated quantiles are

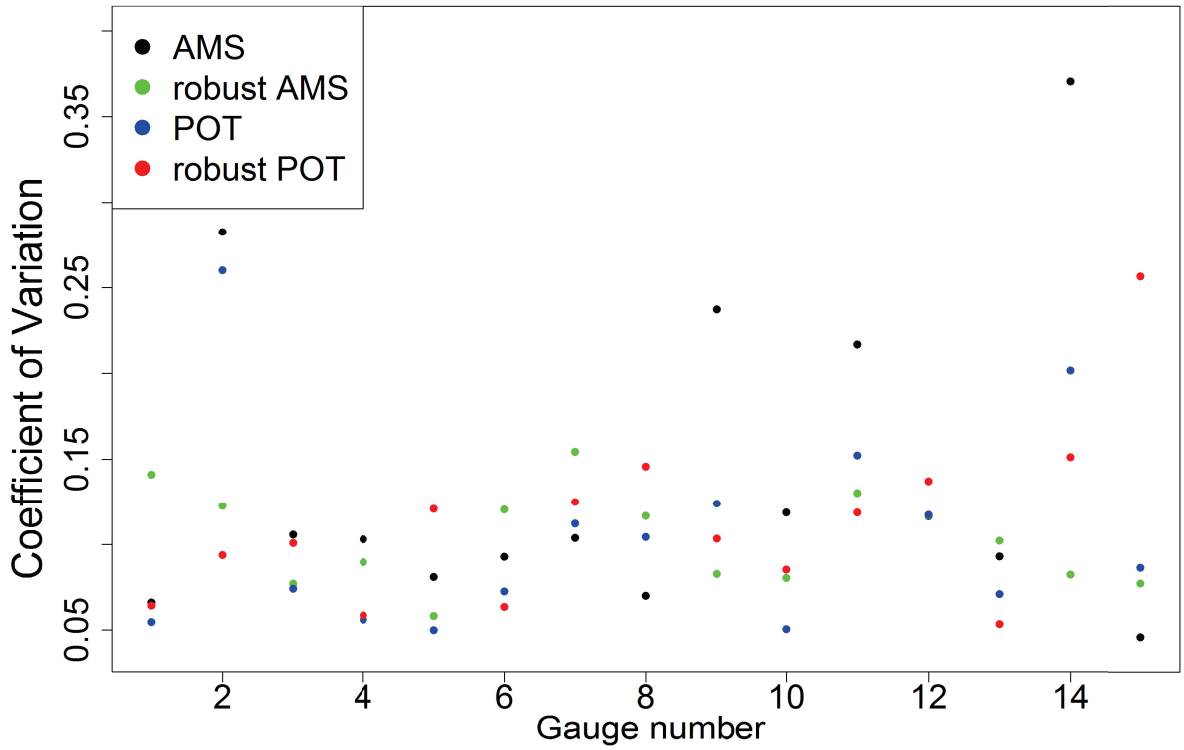

Figure 7.35.: Coefficient of variation of the 99%-quantile estimation for increasing sample length with four different approaches for 15 gauges of the Mulde basin. Overall the robust POT approach has the smallest variation.

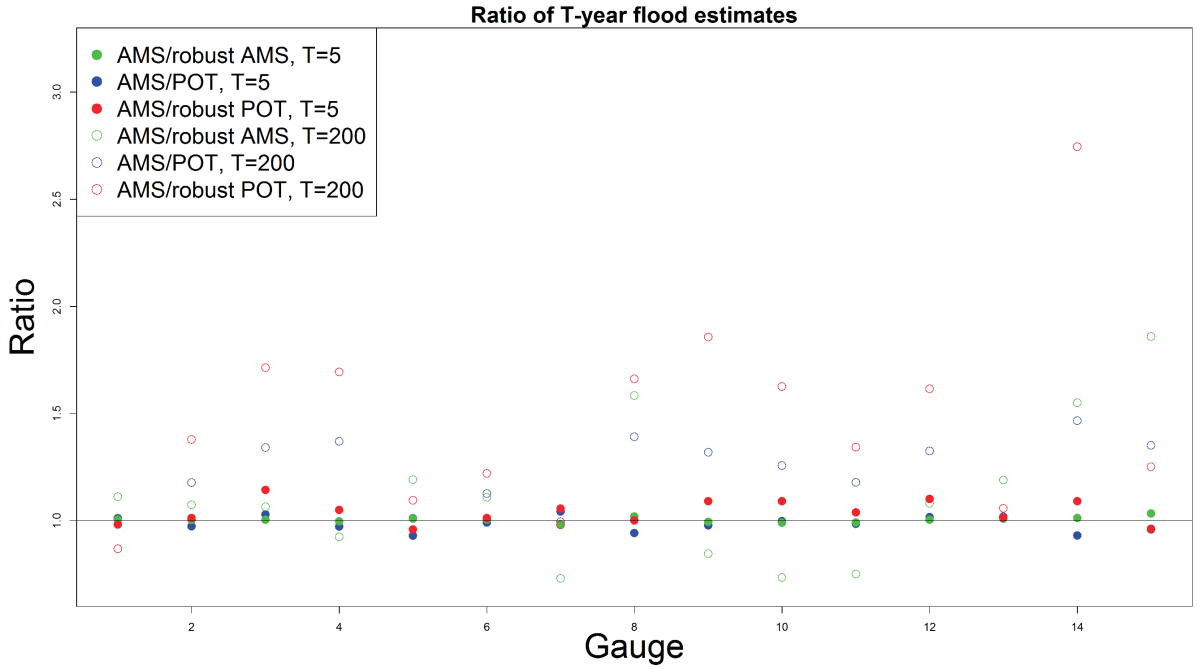

Figure 7.36.: Comparison of the AMS and the robust AMS (green), the POT (blue) and the robust POT with TL-moments (red) by the ratio between the flood quantiles 0.8 (filled) and 0.995 (blank). It can be seen that for small return periods all models do not differ much in their estimation. Nevertheless, for extreme quantiles the classical AMS estimation is up to twice as high as the robust estimation.

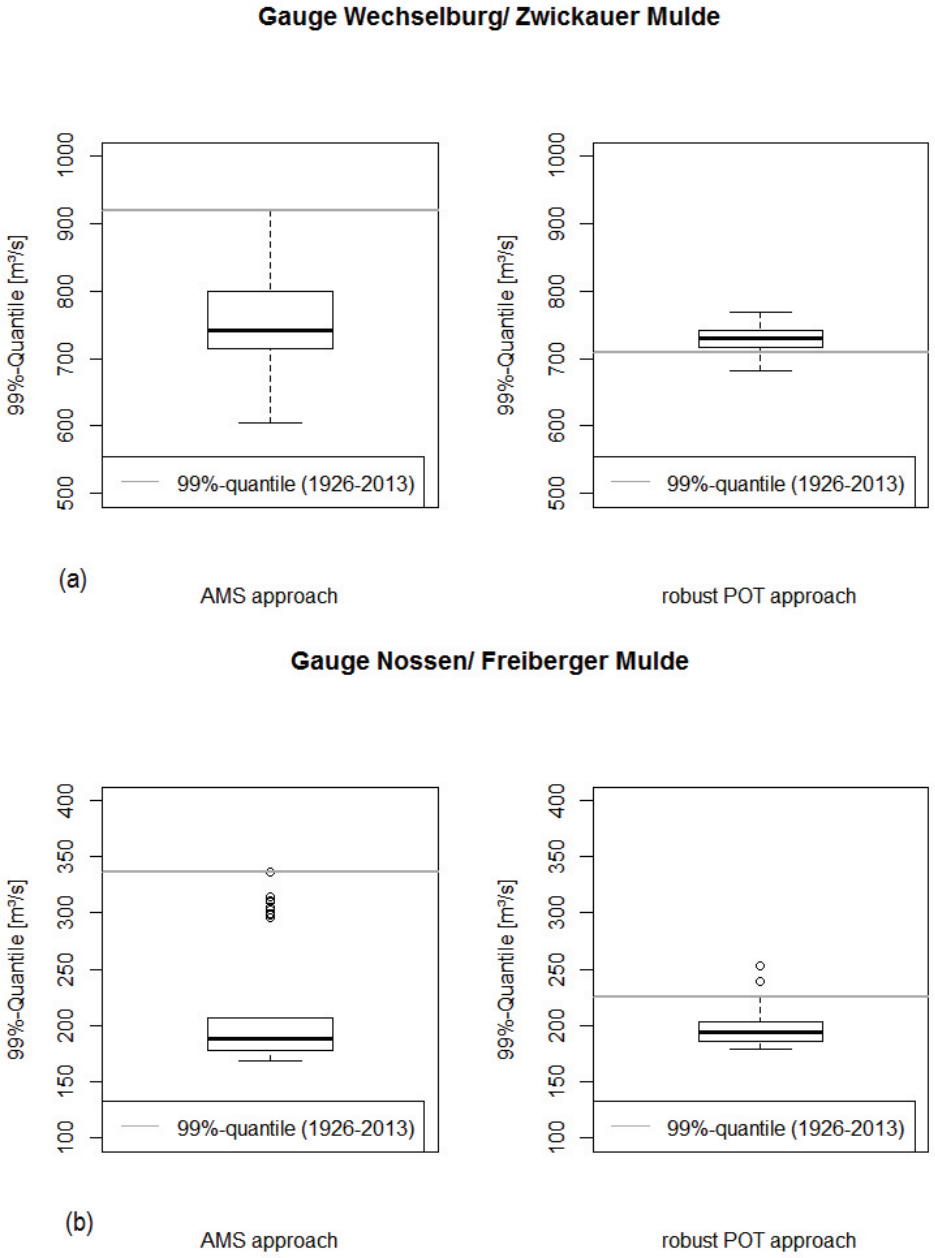

Figure 7.37.: Boxplots of all estimated 99%-quantiles with the AMS or robust POT with TL-moments approaches for year-by-year increasing sample length starting with a length of 40 for the (a) Wechselburg and (b) Nossen gauges. The estimated 99%-quantiles for the fully recorded series for AMS and robust POT are also shown (grey line). The robust POT with TL-moments approach has only small variation and estimates the quantile well.

grouped close together, having small deviations. The quantile for the total series up to 2013 is located close to the centre of the boxplot, that is the median of all estimated quantiles. In contrast, the range of the boxplot for AMS-based quantiles is broader, and the quantile calculated from the total series is located far from the centre of the boxplot. For the Nossen gauge it is even in the range of the most extreme values. Comparing the mean and median of the estimated quantiles, they do not differ much for both approaches (about $5 - 10\%$), therefore one can say that over time, both approaches in general come to the same results, although the robust POT with TL-moments leads to less variation.

From the results given in this chapter we can conclude that the use of robust estimators in the context of flood statistic stabilises the estimated extreme quantiles, that are needed for design floods. While the small quantiles do not change compared to the classical AMS approach, where a GEV distribution is fitted to the annual maxima using L-moments, the extreme quantiles (90% and higher) are less affected by extraordinary large events. If additionally the data basis and therefore the used information is extended by a POT approach, this behaviour increases. The influence of extraordinary extreme events in short time series, that could lead to an overestimation, is reduced in this case on the one hand by enlarging the time series and on the other hand by down-weighting the tail events.

A natural question arising with the results of the robust POT approach is the probable underestimation. Having a look at the results so far one cannot conclude, if the robust POT with TL-moments approach maybe simply has a too large bias, whereas the classical approaches result in high quantiles that are nevertheless correct. Therefore, we want to prove in the following that after a long period of observation the estimated quantile of the AMS series converges to that of the POT series. That is, the large values estimated by the AMS approach are caused by extraordinary large events right at the beginning of the series, that have a large influence on the estimation. Hence, the estimation is falsified. The robust POT approach can cope with this.

Of course, the recorded discharge series are not long enough to validate this assumption. Instead, we simulate a series of 12000 GEV-distributed data. These should be the monthly maximum discharges of 1000 years. The location parameter is chosen as $\mu = 50$, scale as $\sigma = 5$ and shape as $\xi = 0.6$. To guarantee extraordinary extreme events at the beginning of the series, the largest two values are set on the position 50 and 500. Again, the 99%-quantile of the prolonged series is calculated (see Figure 7.38).

As assumed the convergence of the estimated quantile of AMS as well as the POT approach to the quantile estimated by the robust POT approach can be confirmed by the simulation. We can conclude that the estimation using the AMS approach is strongly influenced by the extraordinary extreme events at the beginning of the series whereas this influence is reduced the longer the sample is. The robust POT with TL-moments is not influenced so much by the single events and therefore is able to estimate the limit quantile from the beginning on.

To validate the results we repeated the estimation given in Figure 7.38 1000 times and calculated the mean deviation between the AMS and the robust POT 99%-quantile for the last 100 estimates of length $n = 901, \ldots, 1000$ years. A boxplot of the results is given in Figure 7.39. The main part of the estimated deviation lies in the range of 50 and 250 m^3/s. The large deviation can be explained by having in mind that the mean of the last 100 years is considered. Therefore, the convergence might not have reached the limit during the whole time span and higher values have influence on the mean. All in all, we can conclude that the robust POT with TL-moments and the AMS approach both lead to the same estimation of the quantiles after a sufficient time period.

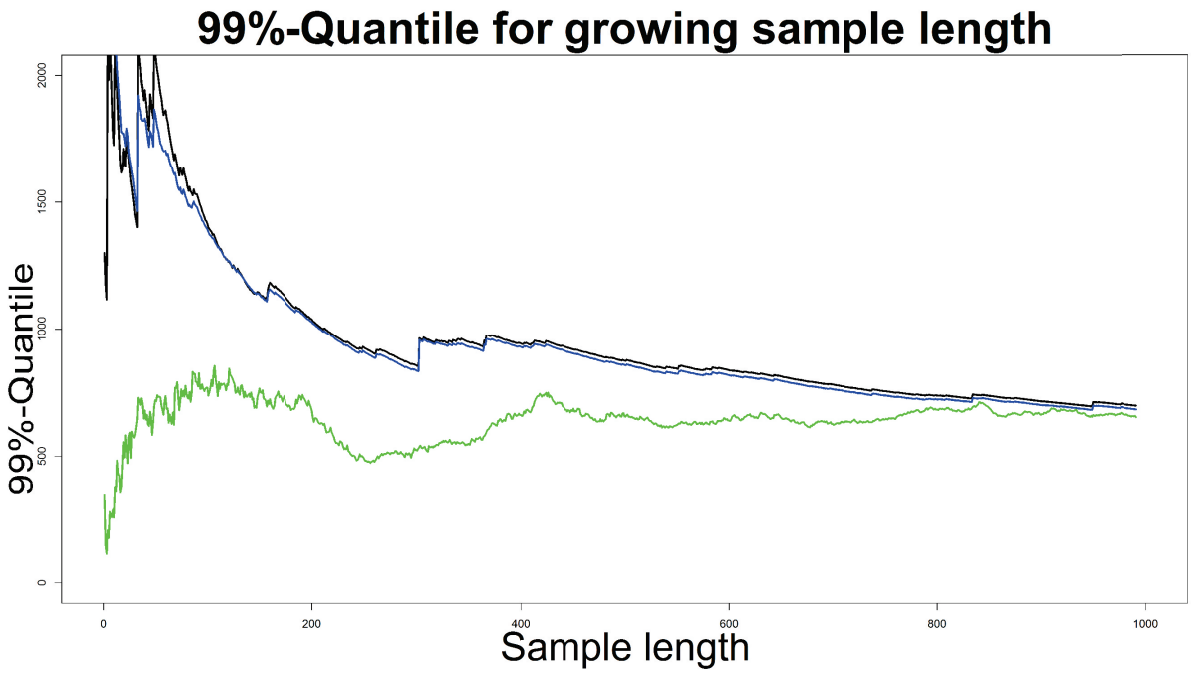

Figure 7.38.: Comparison of the AMS (black), the POT (blue) and the robust POT with TL-moments (green) approach by the estimated 99%-quantile for a year-by-year prolonged series. It can be seen that all approaches converge to the same value for a sufficient large sample length.

Robust flood-type mixing approach

Since the application of robust estimators in hydrological models seems to be very promising according to the results in Section 7.2.3 we also want to extend the seasonal model of section 7.2.1 to robust estimators. Based on the results before, again the robust TL-moments shall be used and replace the L-moments. Since the model is based on the GEV distribution, no new estimators have to be introduced.

In the following we compare the robust AMS, robust WS and robust WST (all with TL-moments) with the initial AMS model based on L-moments. First, we want to get an idea of how much the use of robust estimators changes the quantiles. Therefore, for the Wechselburg gauge we calculated the 90, 99, 99.5 and 99.9%-quantile with all four approaches. The results can be found in Table 7.22. Additionally, the detailed shape of the distribution function can be found in Figure 7.40. As expected, the robust approach reduces the influence of the extreme events in the years 1954, 2002 and 2013 such that the robust AMS as well as the robust WST lead to smaller estimates than the non-robust AMS. Compared to the progression of the distribution with non-robust estimators (Figure 7.40), here it seems to be subdued and the WST approach only results in larger quantiles for much larger values of p. Surprisingly, this is not the case for the robust WS approach. When we estimate the robust approach with TL-moments, this leads to even higher quantiles than the non-robust WS model for almost all annualities which is not the case for any of the other models (see Figure 7.24). On the first glance it seems that the WS model is not suitable for the use of robust estimators. We want to investigate this further.

First, let us have a look at the influence of single annual observations on the estimation. As before, the year-by-year prolonged series of the Wechselburg gauge is considered and the different

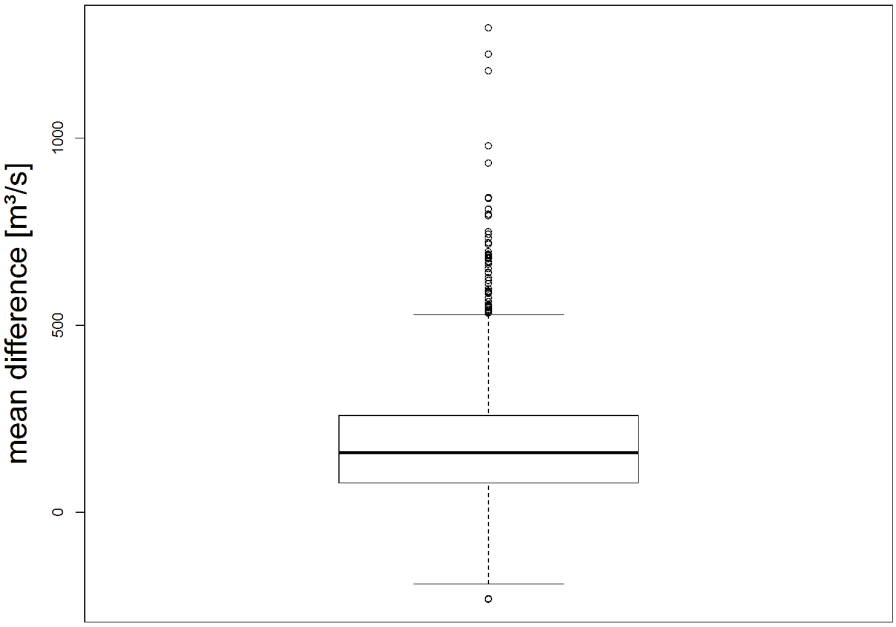

Figure 7.39.: Boxplot of the mean difference of the last 100 estimates of a year-by-year prolonged series calculated with the AMS and the robust POT with TL-moments and repeated 1000 times. A mean deviation of about 160 m^3/s can be seen with some extremes in the right tails.

Table 7.22.: Estimated quantiles (in m^3/s) for the Wechselburg/Zwickauer Mulde gauge with the different models. The robust WST with TL-moments results in the smallest estimated quantiles (for quantiles larger than 99%), whereas the robust WS model with TL-moments results in very high quantiles.

Quantile	AMS (L-moments)	AMS (robust)	robust WS	robust WST
90%	403	249	403	336
99%	919	868	1191	774
99.5%	1176	1015	1671	988
99.9%	1986	1819	3754	1783

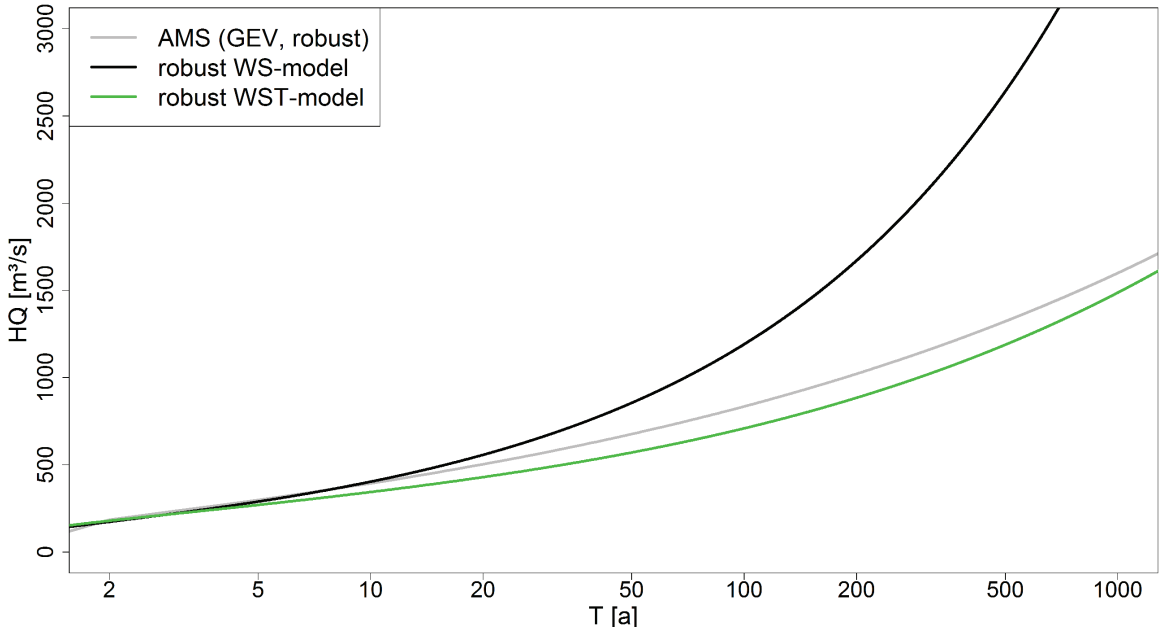

Figure 7.40.: Distribution functions calculated for the Wechselburg/Zwickauer Mulde gauge using different robust statistical approaches. The robust AMS as well as the robust WST approaches behave alike and only for very high quantiles $T \approx 2000$ the robust WST approach (all three with TL-moments) results in higher quantiles. The robust WS model estimates much higher quantiles.

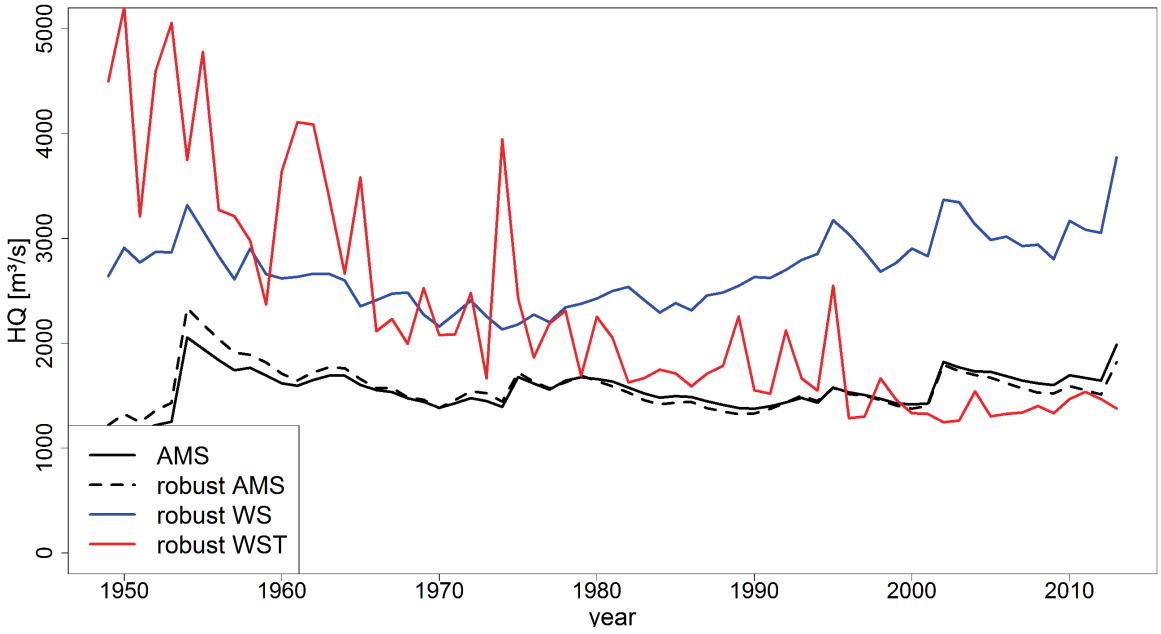

Figure 7.41.: Estimation of the 99%-quantile of the Wechselburg gauge for growing sample length with four different approaches. A high fluctuation for the robust WST model can be seen, although both the AMS models and the robust WST model converge to same value.

approaches are used to estimate the 99%-quantile (Figure 7.41).

We can see a high influence of the single extraordinary extreme events on both, AMS and robust AMS series. Additionally, the WS model proves to be not compatible with the robust estimators again. The deviation from the original estimation is much too high and cannot be explained. What is very striking is the high variance in the estimation by the robust WST model. The fluctuation in every time step is much higher than for the AMS. Only for a sample length of 90 years the estimation becomes more stable. This fluctuation can be explained by the estimation technique. By using the filling method, at every time step small deviations in the estimation of the parameters occur. In very high quantiles these deviations have large influence. Moreover, robust estimation is known to be less efficient for small samples as we have seen above. The multiple use of robust estimators (for winter as well as short and long summer) increases the uncertainty and leads to such deviations. Nevertheless, the QQ-plots still show the best fit of the robust WST model with TL-moments (Figure 7.42). To obtain a better idea of how much the use of robust estimators influences the estimation, we want to simulate a very long time series to evaluate.

Hence, we simulated three series of $n = 1000$ GEV-distributed random variables. The parameters are chosen according to the example of the Wechselburg gauge, that is for the short summer series we have $\mu_{S_S} = 92$, $\sigma_{S_S} = 32$, $\xi_{S_S} = 0.5$, for the long summer series $\mu_{S_L} = 95$, $\sigma_{S_L} = 60$, $\xi_{S_L} = 0.3$ and for the winter series $\mu_W = 115$, $\sigma_W = 54$, $\xi_W = 0.2$. The summer series as well as the AMS are then calculated with the pairwise maximum. To obtain an extraordinary large value right at the beginning of the series we interchange the value in the short summer series at time $i = 100$ with the maximum of the series. The results of the 99.9%-quantile for the AMS, the robust

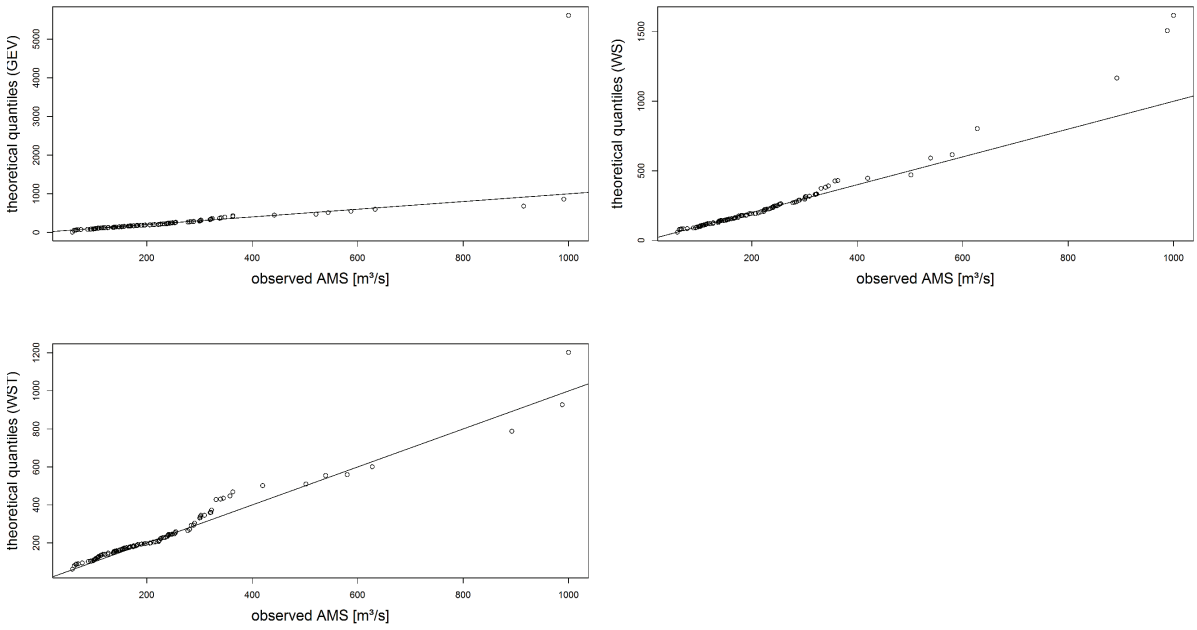

Figure 7.42.: QQ-plot of the annual maximum series of the Wechselburg gauge compared with the robust AMS, the robust WS and the robust WST model. All models except the robust WST model overestimate the highest flood in the series.

AMS, the robust WS and the robust WST model with growing sample size can be found in Figure 7.43. We can see that again the AMS/robust AMS model (which result in almost the same estimates) with TL-moments converges to the value of the robust WST approach with TL-moments, although the AMS method is much influenced by the extraordinary extreme value at the beginning of the sample. For the robust WST this is not the case. Also, the fluctuation of the WST estimate is much reduced with increasing sample length. The WS model again proves to be not compatible with robust estimators.
The results show that the use of robust estimators in hydrological models proves to be a valuable tool when needing to obtain stable estimation. The occurrence of extraordinary extreme events at the beginning of a time series leads in the non-robust case to a large overestimation which reduces only slowly in time. But the simple use of robust estimators in annual maximum series is often not sufficient. Instead, the enlargement of the information spectrum helps much to obtain stable estimation. Additionally, not all hydrological models are suitable for the use of robust estimators as can be seen for the WS model. Therefore, before applying robust estimators a careful investigation should be done.

7.3. Homogeneous Groups and Regionalisation

The number of gauges measuring discharges is very limited. For example, in the Mulde river basin there exist 59 active gauges in a catchment area of more than 6200 km^2. Moreover, many of these gauges have been installed not long ago and therefore long series discharges are even more rare. Therefore, the information given by only a few gauges has to be transferred to unobserved locations to obtain information on these. This is called regionalisation. Of course, not the infor-

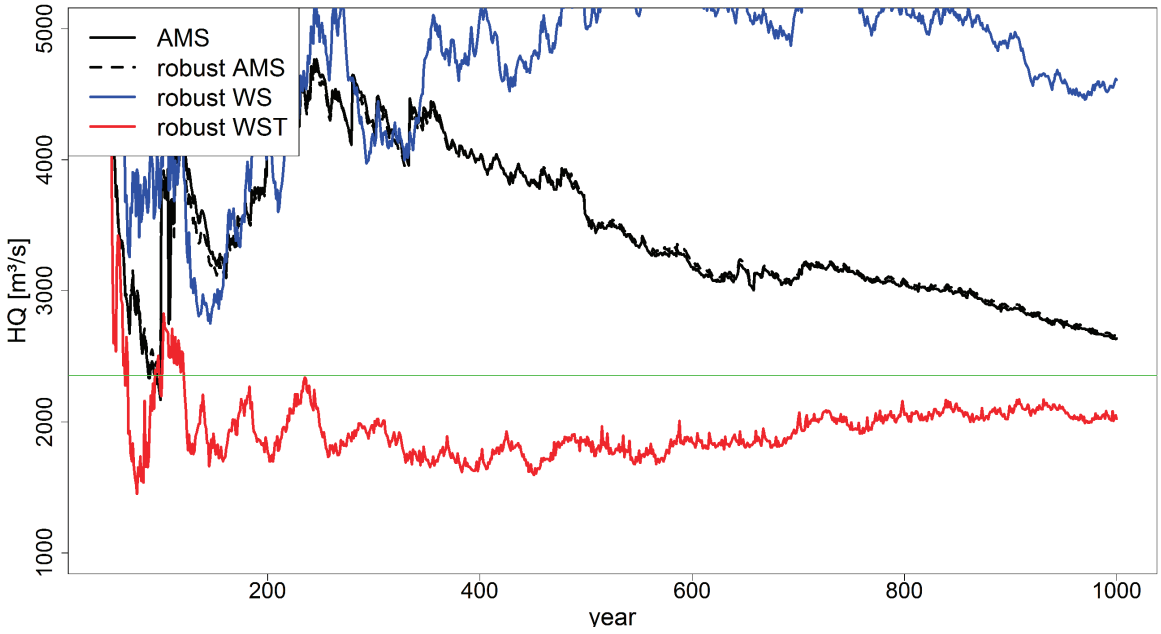

Figure 7.43.: Estimation of the 99%-quantile of the simulated series for growing sample length with four different approaches. The green line indicates the theoretical value of the WST model with the given parameters.

mation of any gauge can be transferred to any ungauged position since discharge and especially flood behaviour depends much on the topographical as well as climatological circumstances. Gauges that can be assumed to behave alike are called homogeneous. One of the main aspects in regionalisation is the classification of homogeneous groups. Besides classical approaches like the Index-Flood approach (Hosking and Wallis (1997)) also geostatistical approaches (Merz and Blöschl (2005)) or topological Kriging (Skøien et al. (2006)) can be used.
We want to show that also the classification into alert steps can result into homogeneous groups since for such a group the alert steps are the same. The origin of this work has been the desire of municipal organisations of water management to obtain easy understandable alert steps that can be explained to publicness. Therefore, to make a broad applicability possible, very basic statistical methods shall be used.

Classification of floods into alert steps

In the last two decades the once-in-a-century-floods accumulated.[1]
This has been the statement of the (German) Federal Agency for Civic Education (Bundeszentrale für politische Bildung) after the large floods in June 2013 affected many parts of Germany. This quotation clarifies the problems of comparing flood events with the help of statistical return periods or annualities. Only experts know that annualities (or the related return periods) are reciprocals of exceedance-probabilities of the peak discharge. Many laymen do not know what the word "statistical" does mean in connection with (statistical) return periods and ask themselves why there occur two floods with return periods of more than 100 years in a period of only 11

[1] „In den vergangenen zwei Jahrzehnten haben sich die „Jahrhunderthochwasser“ gehäuft.“ Bundeszentrale für politische Bildung (2013)

years (2002 and 2013). Practical experience shows that a distinction of probability-based design floods and a flood scale usable for public is needed.

For other natural dangers such as earth quakes and storms such different scales exist for a long time. For earth quakes in media often the Richter scale is used, having its origin in the thirties of the last century. The Richter scale is a logarithmic magnitude scale that characterises the energy release of earth quakes by the amplitude of a seismographic record. It is easily understandable that an earth quake of 5 on the Richter scale is smaller than one of 7 and much smaller than one of 9. But for constructional earth quake assessment for skyscrapers and other high buildings a reference value of the soil acceleration for specific earth quake zones is used that corresponds to an exceedance probability of 0.1 within a time span of 50 years. This probability would correspond to an annuality of 475 years. Likewise the wind speed is described by the Beaufort scale, distinguishing between calm and hurricane force in 13 steps. In construction wind loads are specified by a wind speed pressure (mean wind speed) and squall speeds. These two examples show that a distinction between ordinal and constructional (mostly probabilistic) scales is common use. Nevertheless, for the characterisation of floods probabilities are used for design floods as well as for public event classification.
The use of the return period, that is the annuality, for the characterisation of the floods has been proposed by Gumbel (1941) in the context of extreme value theory. For the design of dams and other buildings this probabilistic approach then has been borrowed. Concerning the duration of use and the permissible risk of exceedance the used design events are based on more and more high annualities. For example, the DIN 19700-11 sets the annualities used for design events at 1000 up to 10,000.
In the following we propose a method for the classification of floods based on alert steps.

How to choose an ordinal scale for flood classification

To assess a flood event under the use of an ordinal scale, hydrological assumptions on an area should be summarised and the application of probabilities should be avoided to prevent from misinterpretation.
One possibility is a classification according to flood damages. In general, public is less interested in the hydrological characteristics of a flood event but in the consequences and the frequency of damages. This could be for example flooded areas, affected inhabitants or pecuniary damages. However, such criteria are not appropriate to classify the magnitude of a flood. Although the damages increase with increasing magnitude of a flood, there is still a large variability caused by locally different flood protection and random effects. These problems can be summarised in three counterarguments:

- Flood damages are in general distributed differently in space according to the distinct event. For example, a crevasse causes local damages that can vary according to the location of appearance. It is difficult to aggregate several of such local damages in a river basin or an administration authority since different criteria have to be considered.

- Flood damages occur when the hydrological burden (discharge) exceeds the resilience of the respective system (dike, water profile etc.). However, the resilience varies in time due to driftwood jam or moisture penetration. From the fail of flood protection systems caused by constructional failure of single parts due to ageing or human failure one cannot conclude the magnitude of the hydrological burden.

Table 7.23.: Possible classification of flood events with annualities.

flood class	qualitative property	annuality T of the peak discharge
1	Small flood	>2a
2	Medium flood	>5a
3	Large flood	>15a
4	Very large flood	>30a

- In contrast to storms and earth quakes floods have a much more local connection. The possibility of designation of flood areas analogous to earth quake areas or wind zones is not given in this context. Additionally, floods cannot be characterised by single impacts for large areas. Flood damages result from the water level or flow velocity, that are hydrostatic, hydrodynamical and buoyant forces. Besides, also flotsam and erosion can be decisive. All these influences depend much on the local circumstances.

In contrast to natural phenomena like earth quakes or storms having impact on whole areas a flood event cannot be classified according to the strongly varying local damages. In the following we therefore propose a classification based on the frequency (rarity) of events and on probabilities without using them explicitly. It is a connection between the real occurring and by public received floods and the, demanded by the EU-HWRM-RL (European Parliament and the Council (2007)), description of dangers according to annualities and therefore unreal floods.
Often large floods occur at several gauges at the same time. Depending on the type of precipitation (convective or stratiform) large or small areas of a catchment area are affected. Also the reaction of the area plays a crucial role. The classification of floods can therefore also help to detect homogeneous groups of gauges within a catchment area by comparing the single flood events according to their classification.

Classification

Starting point for the development of a classification for floods is the assumption that the effects of flood events are larger the more seldom they occur. More seldom events result in larger flooded areas and the consequences are more serious. To avoid declarations concerning probabilities, annualities of flood peaks have to be translated into qualitative flood classes. An example is given in Table 7.23. Extreme floods therefore occur in mean every 30 years, large floods in mean every 10 years and so on.
The classification into four flood classes is based on the common four flood alert steps in Germany. In contrast to the proposed classes the alert steps are fitted to locally needed flood protection systems and not connected directly to the annualities of discharges. They often vary strongly within a river basin. Therefore, it can be useful to vary the proposed thresholds for the flood classes according to the catchment areas and consider the alert steps 3 and 4 in this case.

To obtain the threshold of the flood classes based on the annualities given in Table 7.23 the use of the common flood statistic is not advisable. The flood-statistical evaluation of past extreme flood events is problematic since these events change the statistical parameters and therefore the distribution function. Hence, large differences between estimates before and after the occurrence of such an extreme event could be obtained (see Section 7.2.3). Additionally, one should avoid that for different gauges of the same river basin for every event a flood-statistical analysis has to be done, whose results can differ much between the gauges depending on the used distribution

and the sample length. This is the reason why we propose a much simpler approach based on the Chebychev-inequality.

Mathematically, for a given sample $X_1, \ldots, X_n$ we need a value k, such that $\mathbb{P}(X \geq k) \leq \beta$, where β is the threshold to be determined. Since this β is mostly based on a chosen annuality T, we denote the k calculated with the given T as k_T.
There exist several possibilities to calculate this k_T. First of all, one could fit a distribution function and estimate the corresponding quantiles. But, as mentioned above, this leads to a large uncertainty concerning the choice of the distribution and the efficiency of the estimator, especially for small samples. An empirical estimation of the quantiles is limited to the largest observation of the sample as largest quantile and therefore for small samples this is also not suitable.
This is the reason why we want to estimate k_T based on statistical characteristics such as mean, standard deviation and skewness.
The commonly known inequality for the problem above is the Chebychev-inequality. This inequality can be sharpened by further assumptions on the random variables that are made because of the hydrological nature of the data. We can assume that the random variables are unimodal. That means that there exists exactly one m such that the distribution function $F(x)$ is convex for $x < m$ and concave for $x > m$. If the mode is arbitrary, Dharmadhikari and Joag-Dev (1986) generalise a result of Vysochanskij and Petunin (1980) to arbitrary r-th moments, $r > 0$:

$$\mathbb{P}(|X| \geq k_T) \leq \max\left\{\left(\frac{r}{(r+1)k_T}\right)^r \mathbb{E}|X|^r, \frac{s}{(s-1)k_T^r}\mathbb{E}|X|^r - \frac{1}{s-1}\right\},$$

where s is a constant with $s > r+1$ and $s(s-r-1)^3 = r^r$. For further details on inequalities of this type we refer to Savage (1961) and Sellke (1996).

In the case considered here we can restrict ourselves to the first component of the maximum. Due to the calculation of k_T for $T \geq 2$ it can be shown that for large T ($T > 5$) the maximum takes exactly this value and also for smaller annualities no significant difference occurs for this restriction, but the calculation becomes much easier. This leads us to the case of the so called generalised Gauss- or Camp-Meidell inequality but without the assumed symmetry of the data, which cannot be assumed for flood peaks.
Since we only have one data sample with unknown moments, these have to be estimated. We are aware that the use of estimates instead of the theoretical moments causes an uncertainty that can change the threshold (see Kabán (1996)). But these differences are negligible for sufficiently large sample sizes ($n \geq 40$). To obtain a sufficient exactness and to take into account the properties of the data such as skewness, we calculate the Camp-Meidell bound with the third moments.
We obtain

$$\mathbb{P}(|X| \geq k_T) \leq \left(\frac{3}{4}\right)^3 \frac{1}{k_T^3}\mathbb{E}|X|^3$$

and, since the considered data are discharges and hence positive, the absolute value vanishes. As mentioned above the right hand side of the inequality should equal $\frac{1}{T}$. By transformation to k_T we have

$$k_T = \frac{3}{4}\left(\frac{\mathbb{E}(X)^3}{1/T}\right)^{\frac{1}{3}}.$$

Now the non-centred third moment has to be estimated on the basis of the sample $x_1, \ldots, x_n$ with the related order statistic $x_{(i)}, i = 1, \ldots, n$. Commonly, the third moment is estimated by

the mean

$$\hat{\mu}_3 = \frac{1}{n}\sum_{i=1}^{n} x_i^3.$$

But this estimator is not robust and therefore not stable over time, see Section 2.2. Hence, we again want to use a robust estimator to obtain a time-stable estimation even for small sample sizes. Therefore, the mean is replaced by the trimmed mean (see Example 5.2) applied to the third power. In detail, 10% of the data (rounded down), that is 5% symmetrical in the upper and lower tail respectively, are trimmed. Thus, the calculated thresholds are stable even in the presence of new occurring extraordinary large or small events. The k_T-values are then estimated as follows from the AMS:

$$k_T = \frac{3}{4}\left(\frac{\frac{1}{n-2[n\cdot\alpha]}\sum_{i=1+[n\cdot\alpha]}^{n-[n\cdot\alpha]} x_{(i)}^3}{\frac{1}{T}}\right)^{\frac{1}{3}}.$$

Here, α is the degree of the one-sided trimming, that is $\alpha = 0.05$.

A "very large flood" therefore has for the main part of gauges in a region a peak discharge that is greater or at least equal to the specific k_{30}-threshold. The annuality-thresholds chosen in Table 7.23 can be increased or decreased depending on local flood conditions. The number of exceedances is in that case a very important property for the distinction of the choice of the thresholds within catchment areas.

To make an objective choice of the annuality-thresholds possible we propose an optimisation approach. For this two basic assumptions are made:

- Within a (regional) group of gauges the same event should be classified to the same class.
- The empirical return period should be equal to the theoretical return period.

Under these assumptions we develop the following optimisation criterion:

Suppose, we want to determine annuality-thresholds $T_1, \ldots, T_m$ for m flood classes. If we consider a gauge i, $i = 1, \ldots, k$, with sample length n_i within a group of gauges, define $n_{i;T_j}$ as the number of annual maximum discharges of this gauge, whose annuality is larger than the threshold T_j and smaller than T_{j+1} (that is they are larger than k_T and smaller than k_{T+1}). This threshold should be chosen such that approximately the number of these events in the series of n_i years equals the expected value of the number of events within the class $(n_i/T_j - n_i/T_{j+1})$, that is the difference (in absolute value) of the both quotients $n_i/n_{i;T_j}$ and $n_i/T_j - n_i/T_{j+1}$ (respectively between $n_i/n_{i;T_m}$ and n_i/T_m for the largest class) should be minimal. Hence, the optimisation criterion is given by the minimum sum of the absolute variation coefficients of the single classes

$$\min_{(T_1,\ldots,T_m | T_1<\ldots<T_m)\in\mathbb{N}} \sum_{j=1}^{m} |VarC_{T_j}|,$$

where

$$VarC_{T_j} = \begin{cases} \dfrac{\sqrt{\mathrm{Var}\left(\frac{n_i}{n_{i;T_j}} - \left(\frac{n_i}{T_j} - \frac{n_i}{T_{j+1}}\right)\right)}}{\left(\frac{1}{k}\sum_{i=1}^{k}\left(\frac{n_i}{n_{i;T_j}} - \left(\frac{n_i}{T_j} - \frac{n_i}{T_{j+1}}\right)\right)\right)} & \text{for } j < m \\ \dfrac{\sqrt{\mathrm{Var}\left(\frac{n_i}{n_{i;T_m}} - \frac{n_i}{T_m}\right)}}{\left(\frac{1}{k}\sum_{i=1}^{k}\left(\frac{n_i}{n_{i;T_m}} - \frac{n_i}{T_m}\right)\right)} & \text{for } j = m \end{cases}.$$

Here, the threshold values defined in Table 7.23 are used as starting values.

To confirm the applicability of this approach we show the frequency of the floods for different classes and gauges sorted by river basins in the appendix. Additionally, we give the results of the classification highlighted by different colours (Tables A9-A11). It is clear to see that very large and large floods are classified the same for several gauges. With this classification regional groups can be identified.
Please note that an objective distinction into homogeneous groups is only possible if the optimisation is made for the whole basin. Otherwise, the choice of groups would influence the results. Nevertheless, for small basins like the Mulde or Harz this is possible and therefore also a distinction into homogeneous groups is possible.
For the classification of the gauges in the Mulde river basin a clear coherence can be detected for the years 2002 and 2013, where for all gauges at least a very large event is classified (2002) respectively for all gauges of the subcatchment Zwickauer Mulde and all but two gauges in the subcatchment Freiberger Mulde (2013). These two years are in fact the years, where the two extraordinary flood events occurred that caused incredible damages. Moreover, also for the year 1954 in the subcatchment Zwickauer Mulde as well as for the year 1958 in the Freiberger Mulde events are classified as very large and large. This corresponds to the flood behaviour for these years. In 1954 extraordinary long and heavy rainfall especially in the area of the Ore mountains effected mainly the subcatchment Zwickauer Mulde and led to large floods, for example in the area of Zwickau. In the year 1958 the centroid of the heavy rainfalls has been located near Dresden and therefore the Eastern part of the Mulde, and so the subcatchment Freiberger Mulde has been much more affected. Additionally to these special events, also small cluster of gauges behaving alike can be detected. For example in the subcatchment Freiberger Mulde the gauges Streckewalde and Hopfgarten as well as Berthelsdorf, Nossen and Niederstriegis show very similar behaviour for several years. These gauges are located very close together respectively and would probable be assumed to behave alike because of the topographical similarity.
In the Harz region the 2002 event has been one of the largest flood events in this area, too. It has only been exceeded by the event in 1994. In this year, the Vb weather, which is very unusual in this region, led to heavy and long rainfall in an area that normally is located at the lee side of the Brocken mountain. Both events are classified here as extreme large events for many gauges. Since the Harz basin consists of many different river systems that do not enter each other, it cannot be expected that for all gauges these events are classified the same.
We can see that with this classification method we can detect, which gauges behave alike, and therefore homogeneous groups can be found.

Moreover, we are able to detect groups of events of two or more gauges that belong to the same class. This makes a distinction into local and catchment-wide events possible. As an example

we want to use the two gauges Hopfgarten and Lichtenwalde, both belonging to the same subcatchment Zschopau. Since Hopfgarten is located upstream to Lichtenwalde, we can expect similar events for both gauges. Nevertheless, also the influence of the discharging distributary Flöha has to be taken into account. For both gauges we have an annual maximum series of the length $n = 102$ years (in Hopfgarten the observation of one year is missing and cannot be used for comparison). For these, 65 pairs of events (one pair each year) belong to the same class and thus are assumed to belong to a homogeneous group. Likewise, 37 events are assumed to be inhomogeneous. In both groups we can find extraordinary extreme events, for example the events of 1954 and 1958 in the inhomogeneous groups whereas the events of 2002 and 2013 are classified the same.
In common regionalisation methods these homogeneous groups are then used to derive regional distribution functions or quantiles. One of the most common methods is the so called index flood approach. Here, it is assumed that there exists a regional valid distribution function that is approximately the same for all gauges belonging to the homogeneous group and only differs in one parameter, the index flood parameter (Dalrymple (1960)). As index flood parameter often the mean annual maximum or the first PWM/L-moment is used, depending on the parameter estimation. A special case of this model is the regional GEV-model, where a constant shape parameter ξ is assumed for all gauges and the index flood parameter is chosen as ratio of location and scale for each gauge respectively. If we want to model floods for ungauged catchments, the index flood parameter can be estimated by taking into account the specific values of the catchments for example in a (non-)linear regression.

In the following we want to show how the classification into alert steps can help to find homogeneous groups. If we estimate the parameters of a GEV distribution for both gauges using $TL(0,1)$-moments (the advantages of these estimators in this context were shown by Lilienthal et al. (2016)), we obtain as shape parameter $\xi_1 = 0.350$ for the Hopfgarten gauge and $\xi_2 = 0.209$ for the Lichtenwalde gauge. These parameters differ significantly and therefore both gauges would not be assumed to belong to a homogeneous group. But if we use the classification method proposed above, we can estimate the parameters for the group of events that are classified alike and obtain $\xi_{1;h} = 0.387$ and $\xi_{2;h} = 0.396$. Here, the assumption of a constant shape parameter is not violated and a regionalisation model could be used. For the part of events not belonging to the same class still different shape parameters are estimated with $\xi_{1;i} = 0.410$ and $\xi_{2;i} = 0.354$. This is also emphasized by the Figures 7.44 and 7.45, where we can see the high correlation of the homogeneously classified events. The inhomogeneously classified events do not show such a strong coherence. The potential use of such a classification is shown in the following section.

7.3.1. A Regional Mixture Model

The classification of events of gauges belonging to the same catchment gives us several additional information that can be used in flood statistics. We have seen that events at one gauge are often coherent to events of an upstream gauge. But also inflows of subcatchments or local circumstances can have influence on the flood events of one gauge. Using the classification method stated above we now want to determine which annual maximum discharges of the Lichtenwalde gauge are similar to events of the both upstream gauges Hopfgarten and Borstendorf (each belonging to a different subcatchment) and which are caused by local circumstances. These information then shall be used in a mixture model of regional estimation to obtain quantiles for the Lichtenwalde gauge.
These three gauges have a common length of observation of $n = 84$ years. Now, $n_h = 53$

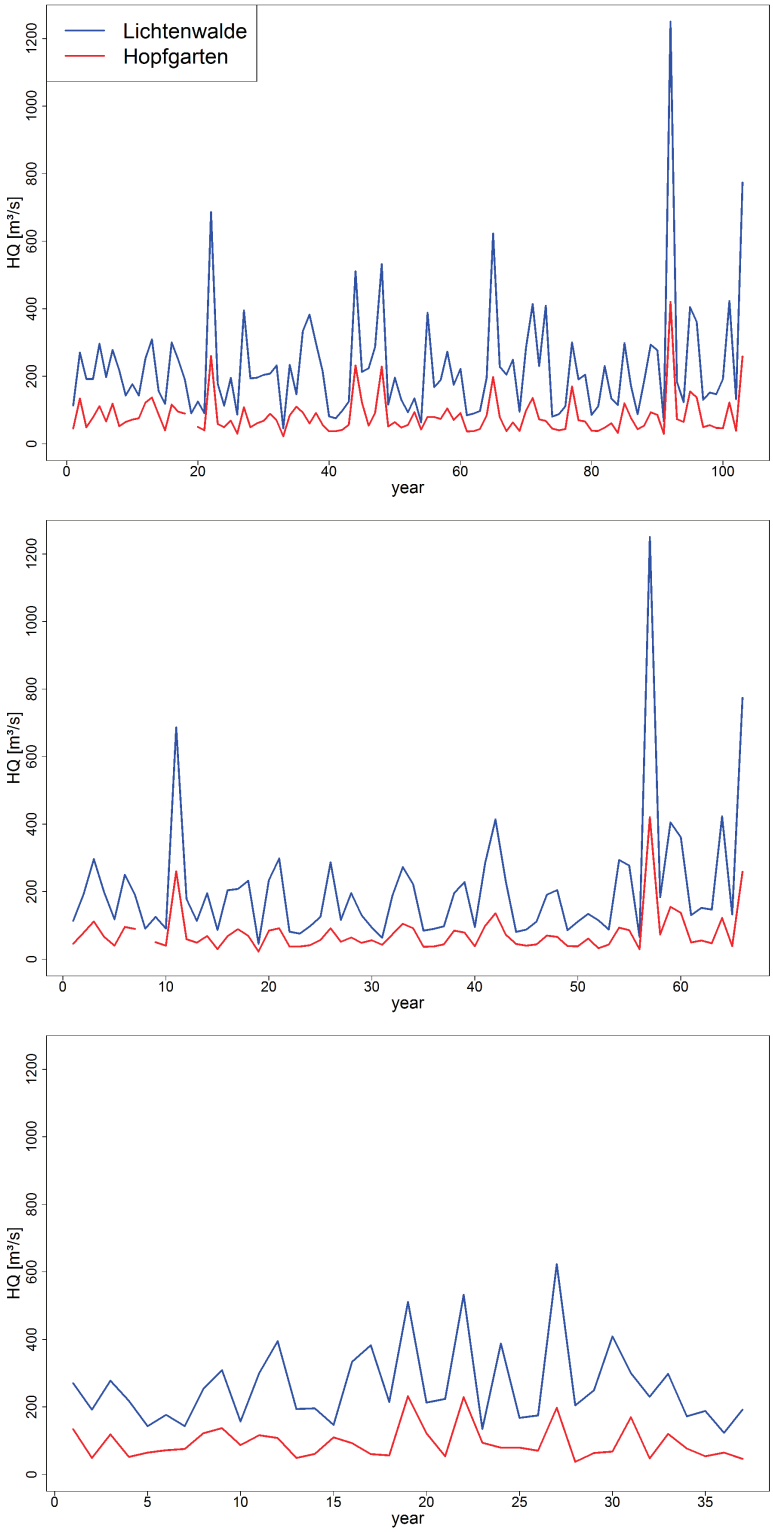

Figure 7.44.: Annual maximum series (top) and events classified alike (middle) and different (bottom) of the Hopfgarten and Lichtenwalde gauges. The similar behaviour of events of the same class becomes obvious whereas the events of different classes show a different behaviour that distorts the homogeneity of the annual maximums series.

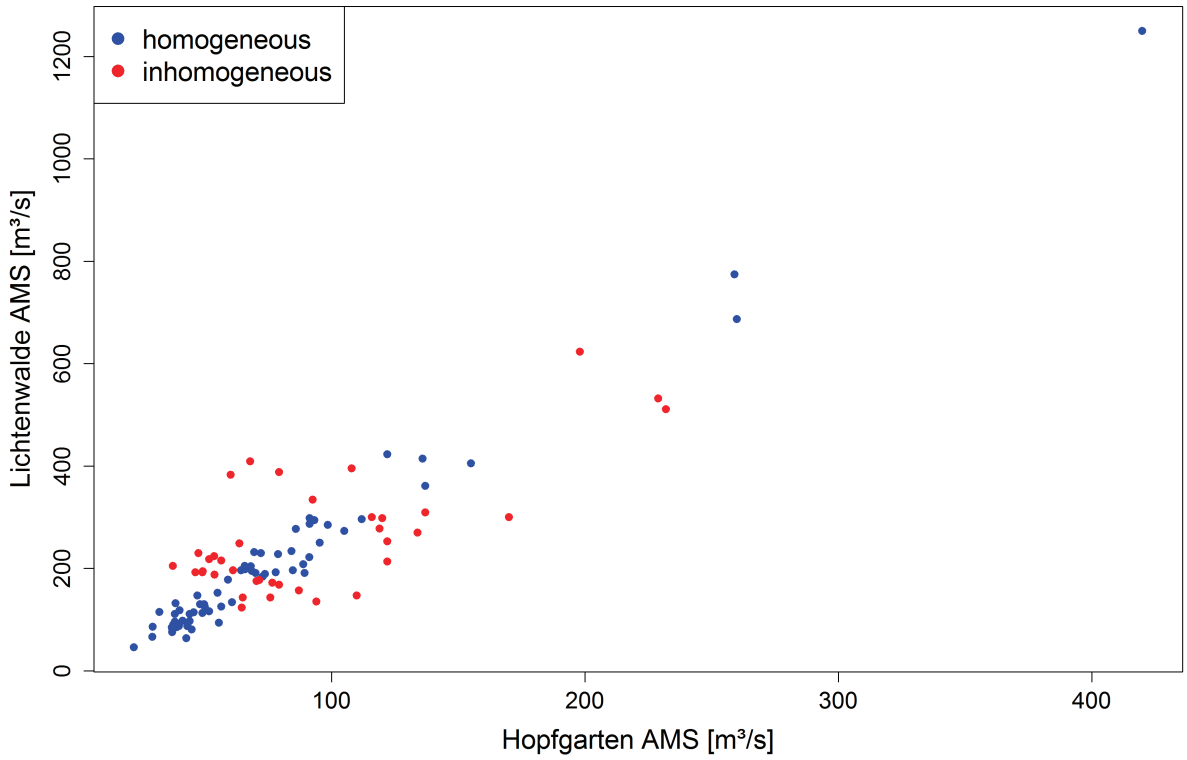

Figure 7.45.: Homogeneous and imhomogeneous classes of the annual maximum series for the Hopfgarten and Lichtenwalde gauges. Whereas the homogeneously classified events show a strong correlation, the inhomogeneously classified events scatter much more.

events of these are classified the same for all three gauges. For these, the values of the shape parameter of the GEV lie within the range of the standard errors. We call these homogeneous events. $n_{h_1} = 5$ events are classified the same for the Hopfgarten and Lichtenwalde gauges, whereas $n_{h_2} = 17$ events are classified alike for the Lichtenwalde and Borstendorf gauges. Thus, $n_i = 9$ events at the Lichtenwalde gauge are not classified the same as for any of the two gauges. These events are called local since they are assumed to only depend on local circumstances. For these four components we now want to apply a mixture model. In contrast to Section 7.2.1, here, a multiplicative mixture model is not sensible since we do not have the information that at least one event-type (homogeneous or local) has to occur in a year. Therefore, the additive mixture model in eq. (7.2) is used. As mentioned before we have four different components: the homogeneous group consisting of events of two different gauges that are homogeneous to the events of the Lichtenwalde gauge, the simple-homogeneous groups with Hopfgarten and Borstendorf consisting only of events of one gauge that are homogeneous to the events of the Lichtenwalde gauge and the local group consisting of local events of the Lichtenwalde gauge. As before we use the GEV distribution to model the annual maximum series. Due to the very small sample sizes, a robust estimation by e.g. the trimmed L-moments is not sensible. The results of Section 7.2.2 imply the use of the probability weighted moments (PWM) or L-moments here.
For the first group a regionalisation approach is needed to obtain quantiles for the Lichtenwalde gauge. We want to apply the index flood approach and estimate the parameters for the distribution function of the flood series of the Lichtenwalde gauge by the moments of the homogeneous part of the AMS of the Hopfgarten and Borstendorf gauges. For this, the algorithm proposed by Hosking et al. (1985b) is used. For each gauge $i = 1, 2$ the PWMs $\beta_{r;i}$ are estimated and scaled by

$$\begin{aligned}\hat{\beta}'_{0;i} &= \hat{\beta}_{0;i}/\hat{\beta}_{0;i} = 1\\ \hat{\beta}'_{1;i} &= \hat{\beta}_{1;i}/\hat{\beta}_{0;i}\\ \hat{\beta}'_{2;i} &= \hat{\beta}_{2;i}/\hat{\beta}_{0;i}.\end{aligned}$$

The regional moments are then the weighted means of these

$$\begin{aligned}\hat{\beta}^{\star}_{0;i} &= 1\\ \hat{\beta}^{\star}_{1;i} &= \frac{\sum_{i=1}^{2} n\hat{\beta}'_{1;i}}{2n}\\ \hat{\beta}^{\star}_{2;i} &= \frac{\sum_{i=1}^{2} n\hat{\beta}'_{2;i}}{2n}.\end{aligned}$$

To estimate the parameters the approximation of Hosking et al. (1985a) is used. With these a regional quantile $\hat{q}_T$ for the annuality T can be estimated which has to be multiplied with the index flood $\hat{\beta}_0$, the first PWM estimated for the AMS of the Lichtenwalde gauge.
For the simple-homogeneous groups, that are the events of the Lichtenwalde gauge that only coincide with one of the other gauges concerning the event class, the same approach as before is used, although we do not have to use weighted means of the moments, since we only have one gauge per group. What remains to estimate are the local effects, that are events of the AMS of the Lichtenwalde gauge not coinciding with any of the other gauges concerning their class. Of course, for ungauged basins these local effects would be unknown and have to be estimated for example by precipitation-runoff-models. Nevertheless, we want to use them in this model to investigate their influence on the estimated quantiles. The resulting mixed probability

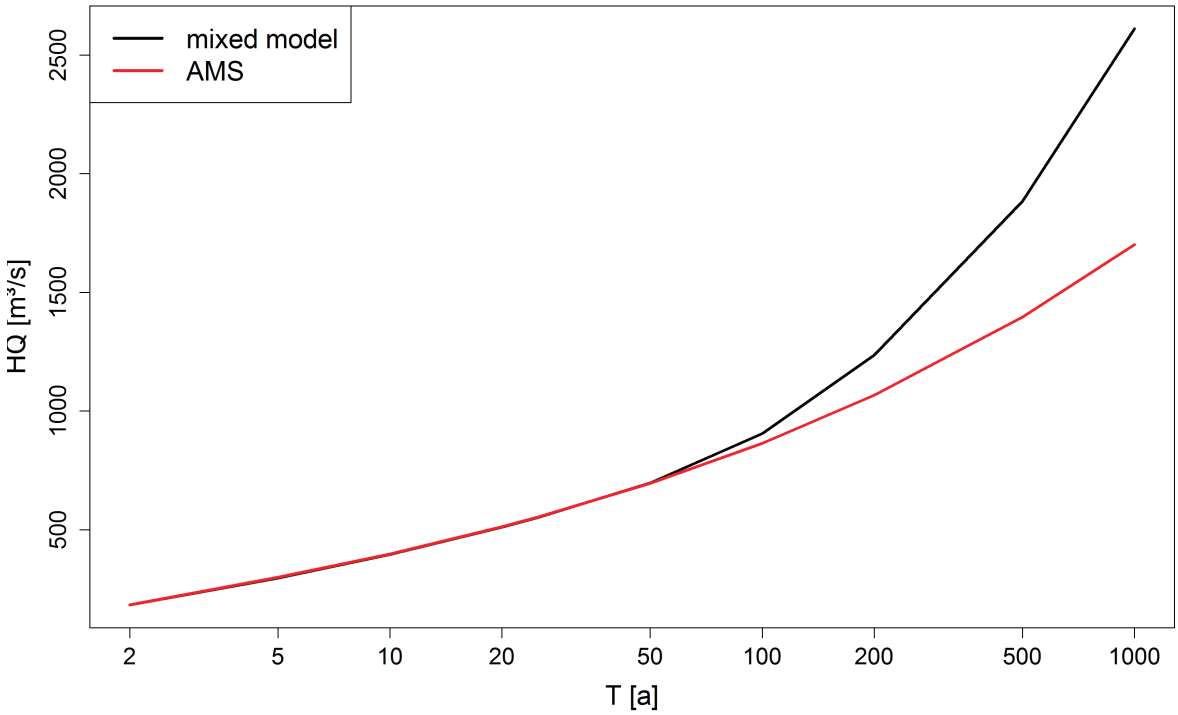

Figure 7.46.: Regional mixture model for the Lichtenwalde gauge compared to the AMS model. A deviation of both models for high quantiles can be seen.

distribution is then as follows

$$F_{LW}(x) = \frac{53}{84}F_{reg}(x) + \frac{17}{84}F_{hom;B}(x) + \frac{5}{84}F_{hom;H}(x) + \frac{9}{84}F_{loc}(x),$$

with F_{reg} being the regional distribution function of the homogeneous group, $F_{hom;B}$ being the distribution function of the single-homogeneous group of the Borstendorf gauge ($F_{hom;H}$ likewise for the Hopfgarten gauge) and F_{loc} being the distribution function of the local events. Please note that the index flood is taken implicitly into account in the homogeneous samples as a normalising constant.
The resulting quantiles of this mixture model are shown in Figure 7.46. We can see an almost perfect accordance with the AMS model in the lower quantiles, whereas we have a deviation of both models for the high quantiles (annualities larger than $50a$). If we compare the single components of the model (Figure 7.47) we can see that the local effects only have a high influence for the small quantiles. Also the homogeneously classified events for the Borstendorf gauge do not influence the right tail of the mixture model much. The components having the largest influence on this tail are the overall homogeneously classified events as well as the Hopfgarten-homogeneous events. These effects are not modelled by the AMS model. We have already seen in Section 7.2.1 that the AMS model seeks a compensation of all mixture components. The same holds true here, where the regional mixture model follows mainly the components with the largest quantile values.

The classification of the annual maximum series in combination with a regional mixture model

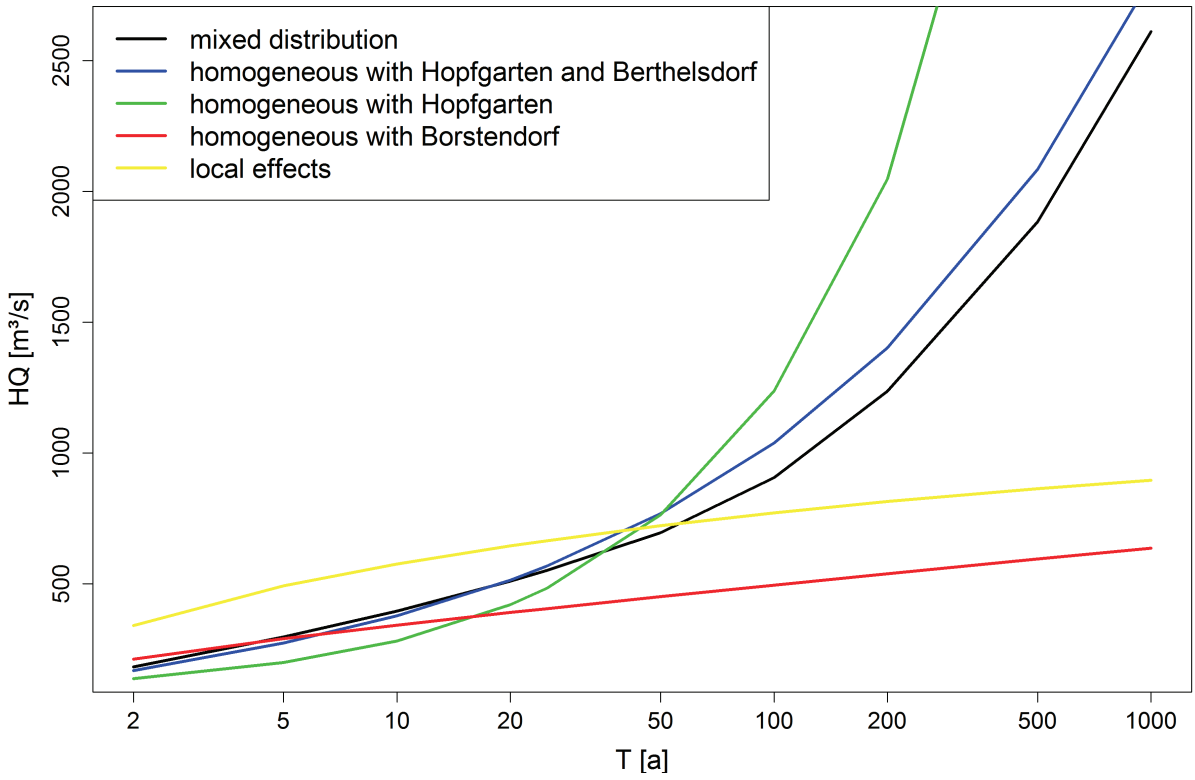

Figure 7.47.: Single components of the regional mixture model for the Lichtenwalde gauge. Only small influence of the local effects and the Borstendorf-homogeneous events for the high quantiles can be seen.

makes a differentiated analysis of the single flood-generating components at a gauge possible. The influence of single subcatchments as well as local circumstances can be extracted and evaluated. It also offers a possibility to estimate design floods for ungauged basins when the local effects can be assumed to have no influence on the high quantiles, as we have seen in the example above, or can be estimated by catchment characteristics. It also can be used as a starting point to develop a conditional model, where the annualities of events of a upstream gauge can be estimated on the basis of the events of the upstream gauge. But this is not the aim of this work and can only be given as an outlook to future work.

7.4. Modelling Time Series

There exist several different possibilities to model discharge series statistically. Often, ARIMA or special cases of ARIMA models are used to describe the dependence structure within a discharge series. GARCH processes or models based on Markov chains are used also.
We have already introduced the example of daily discharges at the gauge in Quebec (Section 3.2), that can be modelled by a combination of an ARMA and EGARCH model. The EGARCH model has been used to describe the heteroscedastic behaviour of the residuals. For these, we now want to apply the robust scale estimators shown before. In detail, we want to use Gini's mean difference G_n (Example 5.4) to estimate the scale and calculate asymptotic confidence intervals with the variance estimator (Theorem 4.3).
The sample size is sufficient such that the asymptotic normality is valid according to the simulations in Section 6.2. Gini's mean difference of the residuals of the ARIMA model applied to the Matapedia discharge series is estimated as $\hat{G}_n^2 = 0.0728$. Now we know according to Theorem 5.2 that Gini's mean difference applied to NED data converges in distribution to a normal distribution with mean $\theta = \mathbb{E}|X - Y|$ and variance $4\sigma/n$, with σ^2 being the variance of the GL-statistics with kernel $h(x_1, x_2) = |x_1 - x_2|$.
We obtain for the estimator of σ that $\hat{\sigma} = 0.0196$ and hence as confidence interval for Gini's mean difference with confidence level $\alpha = 0.05$

$$G_n \in [\hat{G}_n - z_{1-\alpha/2} 2\hat{\sigma}^2/\sqrt{n}, \hat{G}_n + z_{1-\alpha/2} 2\hat{\sigma}^2/\sqrt{n}] = [0.0723, 0.0733],$$

where z_α is the α quantile of the standard normal distribution.
Of course, one could also use special estimators for the parameters of the margins, e.g. of a GEV distribution. In this case the asymptotic theory for L- respectively TL-moments can be used to obtain confidence intervals.

8. Summary and Outlook

In this thesis asymptotic results for classes of robust estimators under dependence are developed. The class of GL-statistics proves to be a very general class, including many of the well known robust estimators, such as Gini's mean difference or the Q_n-estimator of Rousseeuw and Croux (1992). Asymptotic results concerning the asymptotic normality or the consistency of a long-run variance estimator can be achieved by using an approximation by U-statistics. A challenge lies in the use of multivariate kernels, which is necessary for the given robust estimators. Therefore, new limit theorems and invariance principles have to be developed for U-statistics and U-processes for these multivariate kernels. The dependence model is chosen as near epoch dependent process on absolutely regular random variables. This is a very general concept and it has been shown that, besides others, also the well known EGARCH-processes can be represented by such a process. With these results, concrete limit distributions for L-moments as well as trimmed L-moments could be calculated. L-moments form a basis of many parametric estimators, mainly used in hydrological applications. Trimmed L-moments are the robust extension of these estimators and use a trimming of the data to deliver stable estimations.

All these theoretical results form a basis for the application of robust estimators in flood statistics. The aim of flood statistics is the estimation of design floods for given annualities. For this, many different models were developed, all focussing on different aspects of floods. For example, the peak-over-threshold (POT) approach does not only take annual maxima into account but uses all flood peaks above a certain threshold. In this thesis a new model is developed that combines floods of different genesis in a mixture model. For the distinction of these floods the flood timescale is used. Also, a method to reconstruct overlaid events had to be developed, since only one maximum event is observed in a season.
Challenges that have to be faced in flood statistics are the very limited number of observed events as well as the influence of single extraordinary large events in these short time series. Here, robust estimators were applied to obtain estimates stable in time. Stabilisation of estimates is a necessary property of models in flood statistic since the design flood used e.g. for the building of a dam shall not change with every new event. After a simulation study, which takes into account these challenges in hydrology, the TL-moments were found to be the most suitable robust estimators. Moreover, the combination of the POT model with these robust estimators then delivered the desired stable results. Additionally, the convergence of the estimated quantiles of the annual maximum series (the classical method) to the POT-results has been shown for increasing sample length. Hence, one can see that the classical approach is very sensible to extraordinary large floods at the beginning of the observation periods, respectively in short time series. Similar results were shown for the new developed mixing model. But robust estimators were not only used directly in the calculation of design floods, they also proved to be a valuable tool when defining classes of floods, e.g. alert steps. These alert steps then were used to develop a regional mixture model.
The necessity of a consideration of short range dependence becomes apparent when having discharge data of higher time-resolution than monthly maxima. In this case, the developed results for GL-statistics could be used to calculate confidence intervals of the variance of an EGARCH-

process, that has been used to model the residuals of an ARMA model of daily discharges.

Here, also some potential future research becomes evident. Daily discharges are not always short-range dependent. As we have stated in this thesis, also long-range dependence could appear in the data. For this case, no theoretical results for GL-statistics exist so far. Nevertheless, Levy-Leduc et al. (2011) show the invariance principle of U-processes under long-range dependence. Theses results can be extended to multivariate kernels and also used to achieve a Central Limit Theorem for U-statistics. This could be a starting point for the development of a Central Limit Theorem for GL-statistics under long-range dependence and hence for all the estimators used in this thesis.

Bibliography

I. B. Abdul-Moniem and Y. M. Selim. TL-Moments and L-Moments Estimation for the Generalized Pareto Distribution. *Applied Mathematical Sciences*, 3(1):43–52, 2009.

U. N. Ahmad, A. Shabrii, and Z. A. Zakaria. Trimmed L-moments (1,0) for the generalized Pareto distribution. *Hydrological Sciences Journal*, 56(6):1053–1060, 2011.

P. Allamano, F. Laio, and P. Claps. Effects of disregarding seasonality on the distribution of hydrological extremes. *Hydrology and Earth System Sciences*, 15(10):3207–3215, 2011.

T. W. Anderson and D. A. Darling. Asymptotic theory of certain "goodness of fit" criteria based on stochastic processes. *The Annals of Mathematical Statistics*, 23(2):193–212, 1952.

D. W. K. Andrews. Non-strong mixing autoregressive processes. *Journal of Applied Probability*, 21(4):930–934, 1984.

M. Arcones. The Bahadur-Kiefer representation for U-quantiles. *Annals of Statistic*, 24:1400–1422, 1996.

F. Ashkar. *On the Statsitical Frequency Analysis of Hydrological Extremes*, pages 485–505. Springer, 1993.

W. H. Asquith. L-Moments and TL-Moments of The Generalized Lambda Distribution. *Computational Statistics & Data Analysis*, pages 4484–4496, 2007.

R. R. Bahadur. A note on quantiles in large samples. *Annals of Mathematical Statistic*, 37: 577–580, 1966.

A. A. Balkema and L. de Haan. Residual life time at great age. *The Annals of Probability*, 2(5): 792–804, 1974.

A. Bárdossy and S. K. Singh. Robust estimation of hydrological model parameters. *Hydrology and Earth System Sciences*, 12(6):1273–1283, 2008.

A. C. Bayliss and R. C. Jones. Peaks-over-threshold flood database: summary statistics and seasonality. In *Flood Studies Report*. N.E.R.C., 1993.

S. Begueria. Uncertainties in partial duration series modelling of extremes related to the choice of the threshold value. *Journal of Hydrology*, 303(1-4):215–230, 2005.

M. F. P. Bierkens and F. C. van Geer. Stochastic Hydrology. Website, 2008. Online available at `http://www.earthsurfacehydrology.nl/wp-content/uploads/2012/01/Syllabus_Stochastic-Hydrology.pdf`, last visit: 18.11.2016.

D. C. Boes. On the Estimation of Mixing Distributions. *The Annals of Mathematical Statistics*, 37(1):177–188, 1966.

T. Bollerslev. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3):307–327, 1986.

D. D. Boos. A Differential for L-Statistics. *The Annals of Statististics*, 7:955–959, 1979.

S. Borovkova, R. Burton, and H. Dehling. Limit theorems for functionals of mixing processes with applications to U-statistics and dimension estimation. *Transactions of the American Mathematical Society*, 353(11):4261–4319, 2001.

R.C. Bradley. *Introduction to Strong Mixing Conditions*. Kendrick Press, 1st edition, 2007.

P. Brigode, L. Oudin, and C. Perrin. Hydrological model parameter instability: A source of additional uncertainty in estimating the hydrological impacts of climate change? . *Journal of Hydrology*, 476(0):410–425, 2013.

P. J. Brockwell and R. A. Davis. *Time Series: Theory and Methods*. Springer, New Delhi, 2006.

Bundeszentrale für politische Bildung. Hochwasser in Deutschland, 2013. last visited: 21.07.2016.

D. H. Burn. Catchment similarity for regional flood frequency analysis using seasonality measures. *Journal of Hydrology*, 202(1-4):212–230, 1997.

J. Choudhury and R. J. Serfling. Generalized Order Statistics, Bahadur Representations, and Sequential Nonparametric Fixes-Width Confidence Intervals. *Journal of Statistical Planning and Inference*, 19:269–282, 1988.

J. M. Cunderlik, T. B. M. J. Ouarda, and B. Bobee. Determination of flood seasonality from hydrological records. *Hydrological Sciences Journal*, 49 (3):511–526, 2004.

C. Cunnane. A particular comparison of annual maxima and patial duration series methods of flood frequency prediction. *Journal of Hydrology*, 18:257–271, 1973.

C. Cunnane. A note on the Poisson assumption in partial duration series models. *Water Resources Research*, 15(2):489–494, 1979.

T. Dalrymple. Flood Frequency Analysis. Technical Report 1543-A, US Geolleogical Survey, 1960.

J. Davidson. *Stochastic Limit Theory: Advanced Texts in Econometrics*. Oxford University Press, Oxford, 2002.

R. M. de Jong and J. Davidson. Consistency of Kernel Estimators of Heteroscedastic and Autocorralted Covariance Matrices. *Econometrica*, 68(2):407–424, 2000.

H. Dehling and W. Philipp. Empirical Process Techniques for Dependent Data. In *Empirical Process Techniques for Dependent Data*, pages 3–115, 2002.

H. Dehling, M. Denker, and W. Philipp. The almost Sure Invariance Principle for the Empirical Process of U-Statistic structure. *Annales de l'Institut Henri Pointcare (B): Probability and Statistics*, 23:349–382, 1987.

H. Dehling, D. Vogel, M. Wendler, and D. Wied. Testing for Changes in Kendall's Tau. *Econometric Theory*, 2016. doi: 10.1017/S026646661600044X.

M. Denker and G. Keller. Rigorous Statistical Procedures for Data from Dynamical Systems. *Journal of Statistical Physics*, 44:67–93, 1986.

S. Dharmadhikari and K. Joag-Dev. The Gauss–Tchebyshev Inequality for Unimodal Distributions. *Theory of Probability & Its Appllications*, 30(4):867–871, 1986.

D. Dietrich and J. Hüsler. Minimum distance estimators in extreme value distributions. *Communications in Statistics - Theory and Methods*, 25(4):695–703, 1996.

DIN 19700-11. Dam plants - Part 11: Dams, 2004.

DIN 4049-1. Hydrologie - Grundbegriffe, 1992.

P. Doukhan, O. Klesov, and G. Lang. Rates of convergence in some SLLN under weak dependence conditions. *Acta Scientarium Mathematicarum*, 76:683–695, 2010.

DWA. Ermittlung von Hochwasserwahrscheinlichkeiten, 2012.

J.G. Eisenhauer. Regression through the origin. *Teaching Statistics*, 25(3):76–80, 2003.

E. A. Elamir and A. H. Seheult. Trimmed L-Moments. *Computational Statistics and Data Analysis*, 43(3):299–314, 2003.

R. F. Engle. Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, 50(4):987–1007, 1982.

European Parliament and the Council. Assessment and management of flood risks. *Official Journal of the European Union*, 2007/60/EC, 2007.

G. Evin, D. Kavetskii, M. Thyer, and G. Kuczera. Pitfalls and improvements in the joint inference of heteroscedasticity and autocorrelation in hydrological model calibration. *Water Resources Research*, 49(7):4518–4524, 2013.

X. Fang, D. B. Thompson, T. G. Cleveland, P. Pradhan, and R. Malla. Time of Concentration Estimated Using Watershed Parameters Determined by Automated and Manual Methods. *Journal of Irrigation and Drainage Engineering*, 134(2):202–211, 2008.

S. Fischer. Generalisierte L-Statistiken für unabhängige und stark mischende Zufallsvariablen. Master's thesis, Ruhr-Universität Bochum, Germany, 2013.

S. Fischer and A. Schumann. Robust flood statistics: comparison of peak over threshold approaches based on monthly maxima and TL-moments. *Hydrological Sciences Journal*, 61(3): 457–470, 2016.

S. Fischer, R. Fried, and A. Schumann. Examination for robustness of parametric estimators for flood statistics in the context of extraordinary extreme events. *Hydrology and Earth System Sciences Discussions*, 12:8553–8576, 2015.

S. Fischer, R. Fried, and M. Wendler. Multivariate Generalized Linear-Statistics of Short Range Dependent Data. *Electronic Journal of Statistics*, 68(2):407–424, 2016a.

S. Fischer, A. Schumann, and M. Schulte. Characterisation of seasonal flood types according to timescales in mixed probability distributions. *Journal of Hydrology*, 539:38 – 56, 2016b.

R. A. Fisher. *Design of Experiments*. Oliver and Boyd, London, 1935.

R. A. Fisher and L. H. C. Tippett. Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Mathematical Proceedings of the Cambridge Philosophical Society*, 24(02):180–190, 1928.

L. Gaál, J. Szolgay, S. Kohnová, J. Parajka, R. Merz, A. Viglione, and G. Blöschl. Flood timescales: Understanding the interplay of climate and catchment processes through comparative hydrology. *Water Resources Research*, 48(4), 2012.

L. Gaál, J. Szolgay, S. Kohnová, K. Hlavčová, J. Parajka, A. Viglione, R. Merz, and G. Blöschl. Dependence between flood peaks and volumes: a case study on climate and hydrological controls. *Hydrological Sciences Journal*, 60(6):968–984, 2015.

F. Garavaglia, M. Lang, E. Paquet, J. Gailhard, R. Garçon, and B. Renard. Reliability and robustness of rainfall compound distribution model based on weather pattern sub-sampling. *Hydrology and Earth System Sciences*, 15(2):519–532, 2011.

C. Gerstenberger and D. Vogel. On the efficiency of Gini's mean difference. *Statistical Methods and Applications*, 24(4):569–596, 2015.

J. K. Ghosh. A new Proof of the Bahadur Representation of Quantiles and an Application. *Annals of Mathematical Statistics*, 42:1957–1961, 1971.

V. V. Gorodetskii. On the strong mixing property for linear sequences. *Theory of Probability & Its Applications*, 22(2):411–413, 1978.

J. A. Greenwood, J. M. Landwehr, N. C. Matalas, and J. R. Wallis. Probability weighted moments: Definition and relation to parameters of several distributions expressable in inverse form. *Water Resources Research*, 15.5:1049–1054, 1979.

F. E. Grubbs. Procedures for detecting outlying observations in samples. *Technometrics*, 11: 1–21, 1969.

J.-L. Guerrero, I. K. Westerberg, S. Halldin, L.-C. Lundin, and C.-Y. Xu. Exploring the hydrological robustness of model-parameter values with alpha shapes. *Water Resources Research*, 49(10):6700–6715, 2013.

E. J. Gumbel. The Return Period of Flood Flows. *The Annals of Mathematical Statistics*, 12(2): 163–190, 1941.

E. J. Gumbel. *Statistics of Extremes*. Echo Point Books and Media, New York, 1958.

E. J. Gumbel and H. von Schelling. The Distribution of the Number of Exceedances. *The Annals of Mathematical Statistics*, 21(2):247–262, 1950.

P. R. Halmos. The theory of unbiased estimation. *The Annals of Mathematical Statistics*, 17(1): 34–43, 1946.

F. R. Hampel. A general qualitative definition of robustness. *Annals of Mathematical Statistics*, 42:1887–1896, 1971.

B. E. Hansen. GARCH(1, 1) processes are near epoch dependent. *Economics Letters*, 36(2):181 – 186, 1991.

J.-H. Heo, J. D. Salas, and K.-D. Kim. Estimation of confidence intervals of quantiles for the Weibull distribution. *Stochastic Environmental Research and Risk Assessment*, 15(4):284–309, 2001.

W. Hoeffding. A Class of Statistics with Asymptotically Normal Distribution. *Annals of Mathematical Statistics*, 19:293–325, 1948.

J. R. M. Hosking. The theory of probability weighted moments. IBM Research Report, IBM, 1986.

J. R. M. Hosking. Some theory and practical uses of trimmed L-moments. *Journal of Statistical Planning and Inference*, 137(9):3024–3039, 2007.

J. R. M. Hosking and J. R. Wallis. *Regional Frequency Analysis: An Approach based on L-moments*. Cambridge Unversity Press, 1997.

J. R. M. Hosking, J. R. Wallis, and E. F. Wood. Estimation of the generalized extreme-value distribution by the method of probability-weighted moments. *Technometrics*, 27(3):251–261, 1985a.

J. R.M. Hosking. L-moments: Analysis and estimation of distributions using linear combinations of order statistics. *Journal of the Royal Statistical Society. Series B (Methodological)*, 52(1): 105–124, 1990.

J.R.M. Hosking, J. R. Wallis, and E. F. Wood. An appraisal of the regional flood frequency procedure in the UK Flood Studies Report. *Hydrological Sciences Journal*, 30(1):85–109, 1985b.

P. J. Huber. *Robust Statistics*. John Wiley and Sons, 2nd edition edition, 1981.

I. A. Ibragimov. Some limit theorems for stationary processes. *Theory of Probabability & Its Applications*, 7:349–382, 1962.

N. Jenish and I. R. Prucha. On spatial processes and asymptotic inference under near-epoch dependence. *Journal of Econometrics*, 170(1):178–190, 2012.

J. Jeong, E. Park, W. Shik Han, K. Kim, S. Choung, and Il Moon Chung. Identifying outliers of non-Gaussian groundwater state data based on ensemble estimation for long-term trends . *Journal of Hydrology*, 548:135 – 144, 2017.

A. Kabán. Non-parametric detection of meaningless distances in high dimensional data. *Statistics and Computing*, 22(2):375–385, 1996.

J. P. Klein and M. L. Moeschberger. *Survival Analysis - Techniques for Censored and Truncated Data*. Springer, New York, 2003.

V. Klemeš. Dilettantism in hydrology: Transition or destiny? *Water Resources Research*, 22: 177–188, 1986.

V. Klemeš. Tall Tales about Tails of Hydrological Distributions. *Journal of Hydrologic Engineering*, 5(3):227–231, 2000.

K. Kochanek, B. Renard, P. Arnaud, Y. Aubert, M. Lang, T. Cipriani, and E. Sauquet. A data-based comparison of flood frequency analysis methods used in France. *Natural Hazards and Earth System Science*, 14(2):295–308, 2014.

A. Kremer, R. Weißbach, and F. Liese. Maximum likelihood estimation for left-censored survival times in an additive hazard model. *Journal of Statistical Planning and Inference*, 149:33 – 45, 2014.

W. Kruskal, T. S. Ferguson, J. W. Tukey, E. J. Gumbel, and F. J. Anscombe. Discussion of the Papers of Messrs. Anscombe and Daniel. *Technometrics*, 2(2):157–166, 1960.

A.J. Lee. *U-statistics: Theory and Practice*. CRC press Taylor and Francis Group, 1990.

S.-W. Lee and B. E. Hansen. Asymptotic Theory for the GARCH(1,1) Quasi-Maximum Likelihood Estimator. *Econometric Theory*, 10:29–52, 1994.

B. G. Leroux. Consistent estimation of a mixing distribution. *The Annals of Statistics*, 20(3): 1350–1360, 1992.

C. Levy-Leduc, H. Boistard, E. Moulines, M. S. Taqqu, and V. A. Reisen. Asymptotic properties of U-processes under long-range dependence. *Annalls of Statististics*, pages 1399–1426, 2011.

J. Lilienthal. Regionalisierung von Hochwasserdaten mit unterschiedlichen statistischen Verfahren. Master's thesis, Technical University Dortmund, Dortmund, 2013.

J. Lilienthal, R. Fried, and A. Schumann. Homogeneity testing for skewed and cross-correlated data in regional flood frequency analysis. *SFB 823 Discussion Papers*, 16/2016, 2016.

H. Madsen, P. F. Rasmussen, and D. Rosbjerg. Comparison of annual maximum series and partial duration series methods for modeling extreme hydrologic events: 1. at-site modeling. *Water Resources Research*, 33(4):747–757, 1997.

B. D. Malamud and D. L. Turcotte. The applicability of power-law frequency statistics to floods. *Journal of Hydrology*, 322(1-4):168–180, 2006.

R. H. McCuen. *Modelling Hydrologic Change: Statistical Methods*. CRC Press, Boca Raton, 2003.

R. H. McCuen and R. E. Beighley. Seasonal flow frequency analysis. *Journal of Hydrology*, 279 (1–4):43 – 56, 2003.

G. McLachlan and D. Peel. *Finite mixture models*. John Wiley & Sons, 2004.

R. Merz and G. Blöschl. A process typology of regional floods. *Water Resources Research*, 39 (12):1–20, 2003.

R. Merz and G. Blöschl. Floof Frequency Regionalisation - Spational Proximity vs. Catchment Attributes. *Journal of Hydrology*, 302(1-4):283–306, 2005.

R. Modarres and T. B. M. J. Ouarda. Modelling Heteroscedasticity of Streamflow Time Series. *Hydrological Sciences Journal*, 58(1):54–64, 2013a.

R. Modarres and T. B. M. J. Ouarda. Generalized autoregressive conditional heteroscedasticity modelling of hydrologic time series. *Hydrological Processes*, 27(22):3174–3191, 2013b.

F. Moricz. A General Moment Inequality for the Maximum of the Rectangular Partial Sums of Multiple Series. *Acta Mathematica Hungaria*, 41:337–346, 1983.

U. S. Nair. The Standard Error of Gini's Mean Difference. *Biometrika*, 28(3/4):428–436, 1936.

D. B Nelson. Conditional heteroskedasticity in asset returns: A new approach. *Econometrica: Journal of the Econometric Society*, 59(2):347–370, 1991.

N.E.R.C., editor. *Flood Studies Report*, volume 1. Institute of Hydrology Wallingford, London, 1975.

B. Önöz and M. Bayazit. Effect of the occurrence process of the peaks over threshold on the flood estimates. *Journal of Hydrology*, 244:86–96, 2001.

J. III Pickands. Statistical inference using extreme order statistics. *The Annals of Statistics*, 3 (1):119–131, 1975.

M. L. Puri and L. T. Tran. Empirical distribution functions and functions of order statistics for mixing random variables. *Journal of Multivariate Analysis*, 10(3):405 – 425, 1980.

A. Rango and A. Marinec. Revisiting the degree-day method for snowmelt computations. *Water Resources Bulletin*, 31(4), 1995.

B. Renard, K. Kochanek, M. Lang, F. Garavaglia, E. Paquet, L. Neppel, K. Najib, J. Carreau, P. Arnaud, Y. Aubert, F. Borchi, J.-M. Soubeyroux, S. Jourdain, J.-M. Veysseire, E. Sauquet, T. Cipriani, and A. Auffray. Data-based comparison of frequency analysis methods: A general framework. *Water Resources Research*, 49(2):825–843, 2013.

O. Renaud and M. P. Victoria-Feser. A robust coefficient of determination for regression. *Journal of Statistical Planning and Inference*, 140(7):1852–1862, 2010.

J. S. Robinson and M. Sivapalan. Temporal scales and hydrological regimes: Implications for flood frequency scaling. *Water Resources Research*, 33 (12):2981–2999, 1997.

D. Rosbjerg. Estimation in partial duration series with independent and dependent peak values. *Journal of Hydrology*, 76:183–195, 1985.

D. Rosbjerg, H. Madsen, and P. F. Rasmussen. Prediction in partial duration series with generalized Pareto-distributed exceedances. *Water Resources Research*, 28(11):3001–3010, 1992.

F. Rossi, M. Fiorentino, and P. Versace. Two-Component Extreme Value Distribution for Flood Frequency Analysis. *Water Resources Research*, 20(7):847–856, 1984.

P. J. Rousseeuw and C. Croux. *Explicit scale Estimators with high Breakdown Point*, pages 77–92. L1 Statistical Analysis and Related Methods. Elsevier Science Ltd, 1992.

P. Ruckdeschel and N. Horbenko. Robustness Properties for Generalized Pareto Distributions. Technical Report 182, ITWM, 2010.

J. D. Salas, J. W. Delleur, V. Yevjevich, and W. L. Lane. *Applied Modelling of Hydrological Time Series*. Water Resources Publications, Colorado, 1980.

J. L. Salinas, A. Castellarin, A. Viglione, S. Kohnová, and T. R. Kjeldsen. Regional parent flood frequency distributions in Europe – Part 1: Is the GEV model suitable as a pan-European parent? *Hydrology and Earth System Sciences*, 18(11):4381–4389, 2014.

I. R. Savage. Probability inequalities of the Tchebycheff type. *Journal of Research of the National Bureau of Standards-B. Mathematics and Mathematical Physics B*, 65:211–222, 1961.

A. Schumann. Hochwasserstatistische Bewertung des Augusthochwassers 2002 im Einzugsgebiet der Mulde unter Anwendung der saisonalen Statistik. *Hydrologie und Wasserbewirtschaftung*, 49(4):200–206, 2005.

T. Sellke. Generalized Gauss-Chebyshev inequalities for unimodal distributions. *Metrika*, 43(1): 107–121, 1996.

R. Serfling. *Approximation Theorems of Mathematical Statistics*. John Wiley and Sons, 1980.

R. Serfling. Generalized L-, M-, and R-Statistics. *The Annals of Statistics*, 12.1:76–86, 1984.

R. M. Shane and W. R. Lynn. Mathematical model for flood risk evaluation. *Journal of the Hydraulics Division*, 90(6):1–20, 1964.

J. Skøien, R. Merz, and G. Blöschl. Top-kriging - geostatistics on stream networks. *Hydrology and Earth System Sciences*, 10:277 – 287, 2006.

C. S. Spencer and R. H. McCuen. Detection of outliers in Pearson type III data. *Journal of Hydrologic Engineering*, 1:2–10, 1996.

J. R. Stedinger, R.L. Vogel, and E. Foufoula-Georgiou. Frequency analysis of extreme events. In David R. Maidment, editor, *Handbook of Hydrology*. McGraw-Hill, New York, NY, 1993.

J. H. Steiger. Tests for Comparing Elements of a Correlation Matrix. *Psychological Bulletin*, 87: 245–251, 1980.

Hydrological Subcommittee. Guidelines for Determining Flood Flow Frequency. Technical Report 17B, U.S. Department of the Interior Geological Survey, 1981.

V. Taesombut and V. Yevjevich. Use of partial series for estimating the distribution of maximum annual flood peak. In *Hydrology Paper*, volume 97, Fort Collins, Colorado, 1978. Colorado State University.

P. Todorovic and J. Rouselle. Some Problems of Flood Analysis. *Water Resources Research*, 7 (5):1144–1150, 1971.

G. E. Tucker and R. L. Bras. Hillslope processes, drainage density, and landscape morphology. *Water Resources Research*, 34 (10):2751–2764, 1998.

J. W. Tukey. A Survey of Sampling from Contaminated Distributions. *Contributions to probability and statistics*, 2:448–485, 1960.

A. Viglione, G. B. Chirico, J. Komma, R. Woods, M. Borga, and G. Blöschl. Quantifying space-time dynamics of flood event types. *Journal of Hydrology*, 394(1-2):213–229, 2010a.

A. Viglione, G. B. Chirico, R. Woods, and G. Blöschl. Generalised synthesis of space–time variability in flood response: An analytical framework. *Journal of Hydrology*, 394(1-2):198–212, 2010b.

P. J. G. Vlaar and F. C. Palm. The message in weekly exchange rates in the european monetary system: Mean reversion, conditional heteroscedasticity, and jumps. *Journal of Business and Economic Statistics*, 11(3):351–360, 1993.

D. Vogel and M. Wendler. Studentied Sequential U-quantiles under Dependence with Applications to Change-point Analysis. *ArXiv*, 1503.04161v2, 2015.

V. A. Volkonskii and Y. A. Rozanov. Some Limit Theorems for Random Functions I. *Theory of Probability & Its Applications*, 4(2):178–197, 1959.

D. F. Vysochanskij and Y. Petunin. Proof of the 3 σ rule for unimodal distributions. *Theory of Probability and Mathematical Statistics*, 21:25–36, 1980.

H. Wang, X. Gao, L. Qian, and S. Yu. Uncertainty analysis of hydrological processes based on ARMA-GARCH model. *Science China Technological Sciences*, 55(8):2321–2331, 2012.

Q. J. Wang. LH moments for statistical analysis of extreme events. *Water Resources Research*, 33:2841–2848, 1997.

P. Waylen and M.-K. Woo. Prediction of annual floods generated by mixed processes. *Water Resources Research*, 18(4):1283–1286, 1982.

M. Wendler. *Empirical U-Quantiles of Dependent Data*. PhD thesis, Ruhr-University Bochum, 2011a.

M. Wendler. Bahadur Representation for U-Quantiles of Dependent Data. *Journal of Multivariate Analysis*, 102:1064–1079, 2011b.

M. Wendler. U-processes, U-quantile processes and generalized linear statistics of dependent data. *Stochastic Processes and their Applications*, 122(3):787–807, 2012.

W. B. Wu. On the Bahadur-Representation for sample U-quantiles for dependent sequences. *Annals of Statistic*, 33(4):1934–1963, 2005.

V. Yevjevich. *Probability and Statistics in Hydrology*. Water Resources Publications, Colorado, 1972a.

V. Yevjevich. *Stochastic Processes in Hydrology*. Water Resources Publications, Colorado, 1972b.

Appendices

A. Appendix

Here, complementary tables to the results stated in the sections are given.

Table A1.: Goodness of Fit of the GEV, GPD, Gumbel (EVI) and Pearson III distribution to the short and long summer maxima, the whole summer maxima and the winter and annual maxima with the Anderson-Darling test and the AIC. Only values above the flood-threshold are considered, which is taken as threshold parameter for the GPD. The Pearson III distribution has the lowest AIC values for nearly all short and long summer maxima and summer maxima, although in general the AIC does not show great differences between the GEV and Pearson III for all gauges. For all samples the GEV hast the highest p-value for the Anderson-Darling test. The Gumbel distribution is not fitting.

		summer short				summer long				summer				winter				AMS			
		GEV	GPD	EVI	PIII	GEV	GPD	EVI	PIII	GEV	GPD	EVI	PIII	GEV	GPD	EVI	PIII	GEV	GPD	EVI	PIII
Aue	AD	0.946	0	0.002	0.565	0.959	0.94	0.044	0.265	0.921	0	0	0.453	1	0	0	0.484	0.976	0	0	0.33
	AIC	489	482	511	465	238	234	246	209	722	712	753	706	717	727	738	723	843	859	861	851
Niederschlema	AD	0.825	0.736	0.003	0.302	0.998	0.97	0.031	0.193	0.872	0.738	0	0.369	0.899	0	0	0.281	0.95	0	0	0.241
	AIC	470	464	485	460	300	298	311	286	765	761	795	754	789	803	811	800	917	932	938	927
Zwickau-Poelbitz	AD	0.903	0.973	0.002	0.37	0.99	0.738	0.08	0.078	0.934	0.85	0	0.31	0.995	0.345	0	0.705	0.933	0	0	0.501
	AIC	568	563	594	552	178	178	184	158	742	738	775	726	801	801	820	801	936	943	961	942
Harthau	AD	0.981	0.988	0.024	0.254	0.954	0.975	0.349	0.039	0.906	0.984	0.01	0.409	0.998	0	0.003	1	0.994	0	0.003	0.608
	AIC	209	206	229	186	97	94	98	49	300	295	323	276	321	318	323	318	368	375	374	371
Goeritzhain	AD	0.633	0	0	0.237	0.966	0.136	0.034	0.12	0.756	0	0	0.14	0.983	0	0	0.974	0.974	0	0	0.602
	AIC	709	723	718	715	249	249	253	245	958	977	977	968	945	957	945	944	1029	1062	1031	1034
Wechselburg	AD	0.905	0.766	0	0.001	0.95	0.522	0.003	0.087	0.966	0.877	0	0.003	0.945	0	0	0.927	0.923	0	0	0.431
	AIC	613	606	650	545	467	468	479	454	1085	1075	1132	1025	1138	1147	1150	1139	1275	1282	1294	1279
Tannenberg	AD	0.545	0.413	0.004	0.012	0.976	0.784	0.212	0.049	0.745	0.287	0.001	0.019	0.985	0.629	0	0.008	0.853	0	0	0.425
	AIC	158	157	180	127	51	49	55	23	205	204	232	181	301	295	303	280	348	349	360	350
Streckewalde	AD	0.964	0.753	0	0.015	0.993	0.877	0.067	0.14	0.937	0.609	0	0.033	0.991	0	0	1	0.998	0	0	0.762
	AIC	446	441	468	428	170	167	173	154	613	607	640	598	639	639	640	636	741	757	752	746
Hopfgarten	AD	0.999	0	0	0.123	0.977	0.691	0.028	0.138	1	0	0	0.176	0.998	0	0	0.677	0.991	0	0	0.366
	AIC	566	563	597	547	216	216	228	206	778	777	822	774	926	934	936	928	1018	1040	1039	1028
Rothenthal	AD	0.902	0.599	0.001	0.308	0.999	0	0.05	0.097	0.928	0.366	0	0.375	0.969	0	0	0.474	0.993	0	0	0.832
	AIC	393	390	413	380	115	113	115	107	504	502	530	501	497	514	504	502	574	592	583	580
Zöblitz	AD	0.818	0.27	0.001	0.11	0.785	0.92	0.134	0.032	0.88	0	0	0.097	0.984	0	0	0.975	0.492	0	0	0.184
	AIC	317	319	345	311	127	122	127	116	448	450	476	442	501	506	502	500	565	586	581	578
Pockau	AD	0.95	0	0	0.295	0.958	0	0.1	0.072	0.825	0	0	0.089	0.837	0	0	0.782	0.692	0	0	0.858
	AIC	577	573	603	559	201	198	202	182	784	774	813	740	841	843	852	839	903	911	909	902
Borstendorf	AD	0.519	0.449	0.001	0.219	0.949	0.994	0.164	0.019	0.546	0.743	0	0.3	0.936	0.474	0	0.975	0.949	0	0	0.811
	AIC	470	466	491	458	177	174	178	143	674	666	698	652	799	796	812	795	877	883	886	878
Lichtenwalde	AD	0.851	0.644	0.01	0.118	0.584	0	0.003	0.349	0.852	0	0	0.232	0.644	0	0	0.927	0.685	0	0	0.606
	AIC	283	283	299	268	570	562	589	550	847	840	884	822	1143	1137	1162	1136	1261	1262	1277	1260
Niederstriegis	AD	0.257	0	0.001	0.316	0.976	0.994	0.128	0.001	0.175	0	0	0.392	0.843	0	0	0.975	0.827	0	0	0.964
	AIC	322	318	351	296	118	114	116	83	441	433	467	415	591	583	591	585	667	663	666	662
Berthelsdorf	AD	0.981	0.955	0.011	0.006	0.79	0	0.002	0.272	0.952	0	0	0.159	0.436	0	0	0.75	0.585	0	0	0.195
	AIC	153	152	197	134	460	460	483	451	614	614	682	606	756	749	756	750	851	865	885	867
Nossen	AD	0.978	0	0.082	0.896	0.822	0.951	0.131	0.068	0.93	0	0.013	0.314	0.835	0	0	0.993	0.841	0	0	0.811
	AIC	213	211	222	212	163	160	164	146	372	367	386	351	612	605	613	607	658	660	664	658
Erlln	AD	0.984	0.962	0.003	0.337	0.876	0.943	0.006	0.323	0.743	0.884	0	0.416	0.762	0	0	0.986	0.919	0	0	0.842
	AIC	532	529	557	512	567	560	582	544	1072	1066	1108	1057	1333	1331	1335	1327	1385	1404	1392	1389
Golzern	AD	0.749	0.809	0.105	0.216	0.989	0.958	0.055	0.876	0.846	0.987	0.009	0.7	0.992	0	0	0.921	0.998	0	0.001	0.838
	AIC	255	252	265	225	208	204	222	207	459	455	486	459	659	663	665	659	726	728	739	727
Bad Dueben	AD	0.749	0.809	0.105	0.216	0.989	0.958	0.055	0.876	0.846	0.987	0.009	0.7	0.992	0	0	0.921	0.998	0	0.001	0.838
	AIC	255	251	265	225	207	204	222	207	458	455	486	459	658	663	665	659	726	727	738	727

Table A2.: Share (in pct.) of the smallest values of a GEV(μ, σ, ξ) distributed sample that can be removed without a significant change of the estimated quantile for different parameter choices and sample lengths and the 90%- and 99%-quantiles. For all samples we can remove at least 40% of the smallest data until a significant deviation from the estimated quantile occurs.

GEV parameters			$n = 100$		$n = 1000$	
μ	σ	ξ	90%	99%	90%	99%
100	20	0.6	43.255	74.977	17.181	40.852
100	20	0.3	50.493	77.962	21.9	41.032
100	20	0.1	56.238	82.963	26.411	42.397
100	12	0.6	42.709	75.529	17.307	41.463
100	12	0.3	50.246	76.637	22.611	42.118
100	12	0.1	57.522	82.371	26.717	42.266
100	8	0.6	43.602	76.785	17.445	40.938
100	8	0.3	50.249	77.463	22.075	41.541
100	8	0.1	57.653	82.095	26.793	42.165
100	2	0.6	43.273	74.884	17.325	41.112
100	2	0.3	50.247	78.835	22.273	42.239
100	2	0.1	57.625	81.527	26.3	42.059
50	20	0.6	42.722	76.016	17.052	41.986
50	20	0.3	50.374	78.576	22.183	41.471
50	20	0.1	57.441	81.981	26.922	42.081
50	12	0.6	43.99	77.433	16.954	41.819
50	12	0.3	50.673	78.197	22.416	41.252
50	12	0.1	56.882	82.366	26.467	41.91
50	8	0.6	42.511	76.018	17.32	41.11
50	8	0.3	50.472	78.601	22.534	42.55
50	8	0.1	56.884	81.587	26.51	42.355
50	2	0.6	43.16	75.973	17.166	41.599
50	2	0.3	50.417	78.32	22.112	42.479
50	2	0.1	56.909	82.199	26.76	42.31
30	20	0.6	42.293	76.539	17.541	40.81
30	20	0.3	50.358	77.821	22.065	42.122
30	20	0.1	58.839	81.66	26.558	41.556
30	12	0.6	43.104	77.054	16.485	40.684
30	12	0.3	51.273	78.19	22.137	42.206
30	12	0.1	57.212	81.988	26.864	42.733
30	8	0.6	42.434	77.124	16.931	39.987
30	8	0.3	50.902	77.38	22.031	42.361
30	8	0.1	57.057	81.538	26.409	42.558
30	2	0.6	42.775	76.656	17.128	40.767
30	2	0.3	50.986	77.851	22.296	42.48
30	2	0.1	56.821	80.678	27.022	42.201
10	20	0.6	42.906	76.226	16.713	41.267
10	20	0.3	50.772	78.376	22.115	41.534
10	20	0.1	58.641	81.878	26.541	42.11
10	12	0.6	42.803	77.119	17.412	41.114
10	12	0.3	50.938	77.43	22.363	41.145
10	12	0.1	57.427	81.835	26.403	42.556
10	8	0.6	42.736	75.944	16.808	40.671
10	8	0.3	50.792	80.237	22.063	42.088
10	8	0.1	57.454	82.087	26.408	42.727
10	2	0.6	43.761	76.209	17.145	40.819
10	2	0.3	50.543	78.02	22.28	42.195
10	2	0.1	56.815	82.507	26.861	41.902

Table A3.: Estimation of the 99%- and the 99.9%-quantile for independent, identically Gumbel (100, 10)-distributed random variables with sample size 30, 50, 100 and 200.

n=30	Gumbel-Fitting				GEV-Fitting			
	99%-quantile		99.9%-quantile		99%-quantile		99.9%-quantile	
Estimator	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
ML	-1.14	7.39	-1.78	10.7	0.797	18.4	9.62	77.3
L-moments	-0.398	8.08	-0.609	11.7	0.136	14.0	4.58	35.5
TL(1,1)-moments	-0.114	8.27	-0.139	12.1	4.14	22.3	21.2	74.8
TL(0,1)-moments	0.418	7.77	0.626	11.2	9.86	36.9	37.7	124
MD	-0.816	8.67	-1.28	12.7	6.47	32.7	36.4	168

n=50	Gumbel-Fitting				GEV-Fititng			
	99%-quantile		99.9%-quantile		99%-quantile		99.9%-quantile	
Estimator	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
ML	-0.295	5.61	-0.52	8.07	-0.9	11.1	1.00	26.8
L-moments	0.078	6.45	0.108	9.37	0.593	10.6	3.27	25.0
TL(1,1)-moments	-0.176	6.47	0.283	9.54	3.08	16.4	13.3	48.2
TL(0,1)-moments	0.155	5.99	0.243	8.66	3.46	21.4	13.4	57.1
MD	-0.362	6.60	-0.58	9.67	3.54	20.8	17.7	75.5

n=100	Gumbel-Fitting				GEV-Fitting			
	99%-quantile		99.9%-quantile		99%-quantile		99.9%-quantile	
Estimator	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
ML	-0.24	4.01	-0.39	5.76	0.355	8.11	2.39	19.2
L-moments	0.101	4.71	0.103	6.86	0.23	7.78	1.83	18.3
TL(1,1)-moments	0.256	4.71	0.38	6.97	0.755	10.2	4.19	26.8
TL(0,1)-moments	0.117	4.32	0.159	6.22	1.206	14.1	5.14	33.8
MD	-0.01	4.48	-0.03	6.55	1.37	12.7	6.59	34.6

n=200	Gumbel-Fitting				GEV-Fitting			
	99%-quantile		99.9%-quantile		99%-quantile		99.9%-quantile	
Estimator	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
ML	-0.192	2.95	-0.31	4.24	-0.07	5.24	0.373	11.7
L-moments	-0.061	3.24	-0.09	4.72	-0.08	5.16	0.476	11.8
TL(1,1)-moments	0.029	3.32	0.022	4.86	0.747	7.42	2.95	18.3
TL(0,1)-moments	0.042	3.03	0.044	4.38	1.49	9.73	4.51	21.8
MD	-0.035	3.29	-0.06	4.81	0.565	8.57	2.3	21.5

Table A4.: Estimation of the 99%- and the 99.9%-quantile for independent, identically Gumbel (100, 10)-distributed random variables with sample size 30, 50, 100 and 200 and simulated extreme events.

n=30	Gumbel-Fitting				GEV-Fitting			
	99%-quantile		99.9%-quantile		99%-quantile		99.9%-quantile	
Estimator	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
ML	6.29	9.79	8.90	14.0	27.7	34.8	81.7	123
L-moments	11.8	14.2	17.2	20.7	26.7	29.3	74.1	84.6
TL(1,1)-moments	4.67	10.7	6.84	15.7	18.9	34.9	66.4	135
TL(0,1)-moments	3.04	8.83	4.32	12.7	15.4	31.1	47.8	99.1
MD	2.80	9.75	3.96	14.3	25.1	53.2	103	315

n=50	Gumbel-Fitting				GEV-Fitting			
	99%-quantile		99.9%-quantile		99%-quantile		99.9%-quantile	
Estimator	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
ML	3.81	6.89	5.37	9.84	16.2	19.9	42.7	57.0
L-moments	6.81	8.90	10.1	13.0	15.7	18.6	41.1	50.5
TL(1,1)-moments	3.06	7.73	4.46	11.3	10.2	22.2	31.7	70.0
TL(0,1)-moments	2.35	6.85	3.34	9.84	8.60	20.5	25.0	58.4
MD	1.70	7.32	2.39	10.7	12.3	29.9	43.9	117

n=100	Gumbel-Fitting				GEV-Fitting			
	99%-quantile		99.9%-quantile		99%-quantile		99.9%-quantile	
Estimator	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
ML	3.87	5.60	5.50	8.01	14.2	16.0	35.0	40.7
L-moments	6.85	8.01	10.1	11.8	15.2	16.7	38.1	42.9
TL(1,1)-moments	3.25	6.09	4.78	8.95	12.0	17.8	32.5	50.8
TL(0,1)-moments	2.30	4.92	3.26	7.10	8.50	15.3	21.7	40.4
MD	2.14	5.33	3.05	7.75	9.54	18.3	26.6	52.5

n=200	Gumbel-Fitting				GEV-Fitting			
	99%-quantile		99.9%-quantile		99%-quantile		99.9%-quantile	
Estimator	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
ML	4.04	4.96	5.78	7.09	13.9	14.8	33.0	35.7
L-moments	6.92	7.60	10.2	11.1	14.9	15.6	36.7	39.1
TL(1,1)-moments	3.37	4.81	4.95	7.10	11.5	14.5	29.0	38.3
TL(0,1)-moments	2.23	3.87	3.14	5.55	8.73	12.5	20.8	31.3
MD	2.12	4.01	3.04	5.82	7.63	12.8	18.7	32.9

Table A5.: Estimation of the 99%- and the 99.9%-quantile for independent, identically GEV (0.1, 100, 10)-distributed random variables with sample size 30, 50, 100 and 200.

n=30	Gumbel-Fitting				GEV-Fitting			
	99%-quantile		99.9%-quantile		99%-quantile		99.9%-quantile	
Estimator	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
ML	-10.6	13.8	-28.1	29.2	2.68	27.3	21.9	103
L-moments	-7.08	13.3	-22.7	28.1	0.292	21.3	8.34	65.9
TL(1,1)-moments	-8.74	13.3	-25.0	29.0	5.63	33.3	37.2	138
TL(0,1)-moments	-11.0	14.0	-28.6	31.1	7.11	34.5	36.9	133
MD	-11.2	14.9	-28.7	32.1	11.6	58.3	86.5	550

n=50	Gumbel-Fitting				GEV-Fitting			
	99%-quantile		99.9%-quantile		99%-quantile		99.9%-quantile	
Estimator	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
ML	-10.7	12.6	-28.1	29.7	0.939	18.2	8.94	56.0
L-moments	-6.64	11.0	-22.0	25.5	-0.05	17.5	4.75	51.6
TL(1,1)-moments	-9.37	12.3	-26.0	28.4	3.92	23.3	20.9	83.6
TL(0,1)-moments	-10.9	12.9	-28.5	30.1	2.97	23.1	16.1	79.2
MD	-11.0	12.1	-28.4	29.3	6.93	34.1	39.8	167

n=100	Gumbel-Fitting				GEV-Fitting			
	99%-quantile		99.9%-quantile		99%-quantile		99.9%-quantile	
Estimator	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
ML	-9.98	10.5	-27.1	27.5	0.718	12.1	4.67	33.9
L-moments	-6.95	8.14	-22.5	23.3	-0.17	11.8	2.13	32.8
TL(1,1)-moments	-9.19	9.93	-25.7	26.3	2.01	15.6	10.2	47.2
TL(0,1)-moments	-10.9	11.9	-28.4	29.2	2.17	15.1	9.41	44.2
MD	-11.0	12.1	-28.4	29.3	1.84	17.9	11.1	56.7

n=200	Gumbel-Fitting				GEV-Fitting			
	99%-quantile		99.9%-quantile		99%-quantile		99.9%-quantile	
Estimator	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
ML	-9.98	10.5	-27.1	27.5	0.167	8.23	1.77	21.9
L-moments	-6.95	8.14	-22.5	23.3	0.301	8.36	1.90	22.4
TL(1,1)-moments	-9.19	9.93	-25.7	26.3	1.06	10.3	4.83	29.9
TL(0,1)-moments	10.2	11.6	-28.8	29.1	1.19	10.6	4.89	30.2
MD	-10.7	11.3	-28.0	28.5	2.07	12.3	8.33	36.3

Table A6.: Estimation of the 99%- and the 99.9%-quantile for independent, identically GEV (0.1, 100, 10)-distributed random variables with sample size 30, 50, 100 and 200 and simulated extreme events.

n=30	Gumbel-Fitting				GEV-Fitting			
	99%-quantile		99.9%-quantile		99%-quantile		99.9%-quantile	
Estimator	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
ML	-0.363	8.75	-13.3	18.2	46.5	60.4	172	283
L-moments	11.0	14.9	4.02	15.1	43.8	47.4	150	168
TL(1,1)-moments	-3.08	12.0	-16.7	23.8	31.7	56.5	135	284
TL(0,1)-moments	-7.09	11.5	-22.9	26.3	27.0	51.6	105	221
MD	-6.92	13.1	-22.6	27.8	38.4	95.3	216	917

n=50	Gumbel-Fitting				GEV-Fitting			
	99%-quantile		99.9%-quantile		99%-quantile		99.9%-quantile	
Estimator	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
ML	-4.12	8.19	-18.7	21.3	25.1	31.7	79.4	110
L-moments	3.85	9.03	-6.57	13.6	27.1	30.4	85.6	98.5
TL(1,1)-moments	-5.33	10.1	-20.0	23.6	17.3	32.6	61.9	123
TL(0,1)-moments	-8.75	11.3	-25.3	27.3	15.3	33.0	53.2	117
MD	-8.68	11.7	-25.1	27.5	18.7	41.8	75.9	186

n=100	Gumbel-Fitting				GEV-Fitting			
	99%-quantile		99.9%-quantile		99%-quantile		99.9%-quantile	
Estimator	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
ML	-3.98	6.28	-18.4	19.7	23.3	26.3	68.8	80.4
L-moments	3.69	6.95	-6.80	10.9	26.2	28.4	81.3	90.1
TL(1,1)-moments	-4.79	7.53	-19.2	21.0	18.2	26.8	59.1	92.2
TL(0,1)-moments	-8.73	9.97	-25.3	26.2	14.7	23.6	44.3	74.0
MD	-8.46	10.2	-24.7	26.0	13.8	27.3	47.0	96.3

n=200	Gumbel-Fitting				GEV-Fitting			
	99%-quantile		99.9%-quantile		99%-quantile		99.9%-quantile	
Estimator	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
ML	-3.71	5.10	-18.0	18.7	21.6	23.1	60.9	66.2
L-moments	3.84	5.63	-6.59	8.92	25.1	26.2	75.6	79.6
TL(1,1)-moments	-5.08	6.49	-19.6	20.5	18.8	23.2	58.0	74.1
TL(0,1)-moments	-8.18	8.93	-24.5	25.0	14.5	19.7	41.4	58.2
MD	-8.37	9.29	-24.6	25.3	11.8	19.2	35.3	60.3

Table A7.: Estimation of the 99%- and the 99.9%-quantile for independent, identically GEV (0.2, 100, 10)-distributed random variables with sample size 30, 50, 100 and 200.

n=30	Gumbel-Fitting				GEV-Fitting			
	99%-quantile		99.9%-quantile		99%-quantile		99.9%-quantile	
Estimator	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
ML	-23.6	26.4	-71.8	73.8	10.5	68.1	101	1233
L-moments	73.0	78.0	314	343	0.435	33.7	18.8	145
TL(1,1)-moments	-21.7	25.2	-68.5	70.9	5.94	45.7	53.5	226
TL(0,1)-moments	-26.9	28.7	-76.5	77.8	7.56	44.7	51.6	205
MD	-25.7	27.9	-74.7	76.4	23.7	82.8	172	607

n=50	Gumbel-Fitting				GEV-Fitting			
	99%-quantile		99.9%-quantile		99%-quantile		99.9%-quantile	
Estimator	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
ML	-23.7	25.3	-72.0	73.1	2.50	28.3	22.0	111
L-moments	43.5	49.0	173	178	0.009	24.5	9.69	93.5
TL(1,1)-moments	-21.8	23.9	-68.7	70.0	5.03	32.7	33.9	132
TL(0,1)-moments	-26.6	27.7	-76.0	76.8	5.69	34.7	34.74	141
MD	-26.2	27.5	-75.3	76.2	11.0	48.2	69.3	248

n=100	Gumbel-Fitting				GEV-Fitting			
	99%-quantile		99.9%-quantile		99%-quantile		99.9%-quantile	
Estimator	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
ML	-23.3	24.1	-71.3	71.8	1.98	19.2	11.2	63.8
L-moments	-17.2	18.9	-62.0	63.1	-0.65	17.9	2.84	60.5
TL(1,1)-moments	-22.2	23.1	-69.1	69.8	2.74	21.8	16.5	78.5
TL(0,1)-moments	-26.6	27.2	-76.1	76.5	2.23	22.9	14.5	79.8
MD	-25.6	26.3	-74.3	74.8	5.27	28.8	29.2	115

n=200	Gumbel-Fitting				GEV-Fitting			
	99%-quantile		99.9%-quantile		99%-quantile		99.9%-quantile	
Estimator	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
ML	-23.2	23.6	-71.1	71.5	0.874	12.8	5.67	41.9
L-moments	-17.1	18.1	-61.9	62.6	0.11	13.3	2.98	44.1
TL(1,1)-moments	-22.3	22.7	-69.3	69.7	1.92	16.1	11.1	57.6
TL(0,1)-moments	-26.6	26.9	-76.0	76.2	1.69	15.2	8.49	50.8
MD	-25.8	26.2	-74.7	74.9	2.05	18.5	11.4	65.8

Table A8.: Estimation of the 99%- and the 99.9%-quantile for independent, identically GEV (0.2, 100, 10)-distributed random variables with sample size 30, 50, 100 and 200 and simulated extreme events.

n=30	Gumbel-Fitting				GEV-Fitting			
	99%-quantile		99.9%-quantile		99%-quantile		99.9%-quantile	
Estimator	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
ML	-7.16	13.6	-47.9	50.7	84.6	122	436	1254
L-moments	11.4	17.6	-19.8	27.9	73.0	78.0	314	343
TL(1,1)-moments	-14.8	20.1	-58.2	61.6	51.9	95.3	292	646
TL(0,1)-moments	-21.6	24.3	-68.9	70.7	43.4	79.5	201	419
MD	-22.0	25.1	-69.1	71.4	63.3	146	439	1504

n=50	Gumbel-Fitting				GEV-Fitting			
	99%-quantile		99.9%-quantile		99%-quantile		99.9%-quantile	
Estimator	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
ML	-13.6	16.3	-57.2	58.6	40.6	53.8	160	239
L-moments	0.530	11.5	-35.9	39.5	43.5	49.0	173	201
TL(1,1)-moments	-17.5	20.2	-62.3	64.0	28.5	53.6	131	257
TL(0,1)-moments	-23.8	25.1	-72.0	72.9	22.6	47.4	95.1	206
MD	-23.4	24.9	-71.1	72.3	28.9	65.1	145	347

n=100	Gumbel-Fitting				GEV-Fitting			
	99%-quantile		99.9%-quantile		99%-quantile		99.9%-quantile	
Estimator	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
ML	-13.7	15.2	-57.3	58.1	36.1	42.1	131	160
L-moments	0.290	7.97	-36.2	38.0	42.6	45.6	162	176
TL(1,1)-moments	-16.9	18.3	-61.3	62.1	32.6	45.6	132.2	197
TL(0,1)-moments	-22.9	23.6	-70.7	71.2	24.4	37.7	91.3	149
MD	-23.8	24.6	-71.8	72.4	20.8	39.3	84.2	165

n=200	Gumbel-Fitting				GEV-Fitting			
	99%-quantile		99.9%-quantile		99%-quantile		99.9%-quantile	
Estimator	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
ML	-13.6	14.3	-57.2	57.5	34.4	37.5	119	133
L-moments	-0.489	5.93	-37.3	38.3	41.9	43.4	157	164
TL(1,1)-moments	-16.2	17.0	-60.3	60.7	31.8	38.1	120	151
TL(0,1)-moments	-23.0	23.3	-70.8	71.1	23.4	30.3	81.0	109
MD	-23.6	24.1	-71.5	71.8	16.9	28.1	61.0	105

Table A9.: Classification of the gauges in the subcatchment Zwickauer Mulde/Chemnitz of the Mulde river basin according the method proposed in Section 7.3. It can be seen that large and extreme large events (magenta and red) are detected simultaneously for nearly all gauges.

Gauge-ID	563290	562012	563880	563790	562031	562040	564160	562070	564201	564300	564530	564620	564400	564410	562115
Name of gauge	**Sachsen-grund**	**Schön-heide 3**	**Markers-bach 1**	**Aue 1**	**Aue 3**	**Nieder-schlema**	**Wolfers-grün**	**Zwickau-Pölbitz**	**Nieder-mülsen 1**	**Nieder-lungwitz**	**Altchem-nitz 1**	**Har-thau**	**Chem-nitz 1**	**Goeritz-hain**	**Wechsel-burg 1**
Begin	1971	1983	1974	1928	1986	1928	1965	1928	1966	1965	1985	1965	1918	1910	1910
End	2013	2013	2013	2013	2013	2013	2013	2013	2013	2012	2013	2013	2013	2013	2013
Length [a]	43	31	40	86	28	86	49	86	48	48	29	49	96	104	104
Length net. [a]	40	31	40	86	28	85	49	85	45	46	25	44	86	104	104
Median [m^3/s]	2.0	29.7	3.7	50.7	77.5	79.5	2.1	101.0	6.4	19.8	20.4	26.3	49.0	63.3	180.5
Variance [m^3/s]	2.0	25.6	4.8	47.9	67.4	80.7	1.7	98.6	6.9	17.5	23.3	22.6	37.9	43.7	152.4
3. moment	34.37	100381.37	362.41	528819.53	1909110.40	1880403.66	33.28	3080433.26	1202.40	26845.12	64835.11	50071.57	253285.02	516803.58	14506055.85
k(T=2)	3.07	43.74	6.72	76.06	116.63	116.05	3.04	136.78	10.02	28.19	37.81	34.70	59.53	75.48	229.14
k(T=5)	4.16	59.34	9.12	103.20	158.25	157.45	4.12	185.58	13.60	38.25	51.30	47.07	80.76	102.41	310.90
k(T=15)	6.00	85.54	13.15	148.76	228.11	226.96	5.94	267.51	19.60	55.14	73.95	67.86	116.42	147.63	448.15
k(T=30)	7.56	107.80	16.57	187.48	287.48	286.03	7.48	337.13	24.70	69.49	93.20	85.52	146.72	186.05	564.79
number T>2	12	5	7	23	8	25	12	28	13	13	6	13	27	41	37
number T>5	5	3	6	13	4	17	8	15	10	6	3	7	12	21	17
number T>15	4	2	5	6	2	5	4	5	3	3	2	4	6	4	7
number T>15 and T<30	2	0	3	0	1	1	4	1	1	2	0	2	3	1	2
number T>30	2	2	2	6	1	4	0	4	2	1	2	2	3	3	5
Medium [%]	0.18	0.06	0.03	0.12	0.14	0.09	0.08	0.15	0.07	0.15	0.12	0.14	0.17	0.19	0.19
Large [%]	0.03	0.03	0.03	0.08	0.07	0.14	0.08	0.12	0.16	0.07	0.04	0.07	0.07	0.16	0.10
Very large [%]	0.05	0.00	0.08	0.00	0.04	0.01	0.08	0.01	0.02	0.04	0.00	0.05	0.03	0.01	0.02
Extreme [%]	0.05	0.06	0.05	0.07	0.04	0.05	0.00	0.05	0.04	0.02	0.08	0.05	0.03	0.03	0.05
AMS [m^3/s]															
1910														37.7	139
1911														43	99.3
1912														51.4	157
1913														66	226
1914														57.4	153
1915														100	278
1916														60.5	167
1917														71.3	230
1918													23.1	47.2	93.4
1919													13.8	19.5	109
1920													38.1	45.9	138
1921													41	48.3	137
1922													36.2	63.1	168
1923													73.5	90.1	254
1924													51.4	93.4	363
1925													13.6	33.9	136

1926												78.7	104	363
1927												46.5	63.1	289
1928			39	54.8		76.2						22.5	33.9	95.1
1929			27	42.9		55						17.6	20.8	109
1930			36.8	52.4		82.6						30.2	31.5	126
1931			200	189		122						24.5	41.5	102
1932			205	388								126	138	587
1933			51.2	84		89.6						26.9	44.3	107
1934			32.5	82.2		86.5						21.8	35.2	107
1935			62.3	99.4		105						45.5	81.9	165
1936			27.3	51		62						12.1	42.1	103
1937			105	157		158						47.1	78.9	212
1938			34.2	67.1		87.1						52.5	79.9	197
1939			41.8	64		70.7						34.6	77	121
1940			58.1	111		142						37.4	63.2	215
1941			76.1	179		230						56.5	122	303
1942			28.2	93.4		70.3						46.5	29.2	241
1943			13.6	27.4		44.5						11	9.24	59.1
1944			87.3	172		205						79.4	71.9	348
1945			54.7	73		107							48.7	148
1946			73.6	156		165							127	322
1947			34.8	63.7		71.7							41.4	152
1948			68.2	133		157							71.9	237
1949			32.8	67.9		58.1							42.8	117
1950			17.9	43.6		52.9						29.4	34	87.9
1951			25.7	48.7		45.8						29.4	34.6	65.1
1952			41.5	70.5		69.7						21.4	109	102
1953			31.8	73		53.6						30.4	52.2	113
1954			224	585		683						140	241	915
1955			94.3	102		148						75.1	117	223
1956			38.2	64.2		60.3						61.1	81.8	206
1957			107	174		175						44.8	42.1	206
1958			96.2	173		216						75.6	82.8	338
1959			33	56.1		65.3						53.6	56.2	166
1960			60.3	128		136						68.7	87.8	241
1961			47.4	111		192						48.9	64.7	303
1962			41.7	61.6		57.2						20.1	19.6	74.8
1963			35.2	76.3		46.8						96.8	30.9	68.9
1964			23.2	22.7		29.9						24.7	20.1	61.5
1965			82.4	107	2.08	167		22.2		35.2		56.2	71	190
1966			85.7	198	2.56	237	4.36	19.8		26.2		55.3	57	284
1967			41.7	77.5	1.08	98.2	4	10.8		17.8		51.6	53.8	138
1968			61.7	112	1.87	147	4.6	11.2		36.4		70.5	80.7	255
1969			54.7	84.9	3.72	101	22	22.2		18.9		47	51.4	153
1970			73.6	118	5.31	157	13.2	15.3		19.8		52.6	54.6	244
1971	1.67		27.7	48.6	1.43	50.4	3.04	10.8		11.1		45.6	42.5	103
1972	2.76		36.4	73.8	1.31	73.7	2.83	17.6		12.8		45.2	51.4	99.8

1973	1.88			46.4		65.4	3.42	102	3.67	12.5		16.3	31.6	51.8	147
1974	3.81		4.98	44.5		88.5	6.74	133	9.59	25.7		31.1	64.5	78.6	239
1975	6.69		13.2	194		402	2.35	465	14.1	37.6		69.9	121	138	633
1976	1.77		5.9	74.3		136	1.6	177	3.93	18.4		24.3	51.3	60.8	249
1977	1.28		3.67	31.5		71.2	1.48	81.9	3.53	23.4		32.1	70.1	78.6	183
1978	10.7		6.09	71.9		162	5.17	249	13.6	38		59.1	100	127	442
1979	1.77		3.67	34.6		59.5	2.49	74.4	6.87	13.1		22.4	37.7	47.2	113
1980	3.65		4.98	97.5		185	2.42	218	5.75	27				129	324
1981	2.9		8.41	125		172	4.39	222	3.41	14.7				112	320
1982	1.26		19.1	57.7		72.8	1.48	76.9	2.82	12.3				64.2	189
1983	1.95	22.3	2.77	45.9		77.8	1.44	95.6	8.33					125	246
1984	1.34	35.3	3.57	59.3		78.7	7.09	103	6.44					109	180
1985	0.91	26.7	1.98	29.7		35.5	0.771	34.3	1.48	7.68	5.46	10.1	14.8	33.7	66.4
1986	1.19	14	2.52	32.4	60.5	59.5	1.62	61.2	2.48	10.8	11.6	26.2	47.2	59.1	141
1987	2.49	51.2	12	118	158		1.81	230	2.31	19.7	38.8	46.5	78.6	102	281
1988	1.58	27	3.87	55.5	77.5	79.5	2.42	81	2.38	9.61	30.9	30.8	52.3	65.9	175
1989	1.42	20.7	2.77	50.1	59.2	73.6	0.926	80.3	1.68	13.2	18.2	19.5	46.8	76.7	170
1990	1.34	16.9	2.19	31.2	77.5	68.7	0.926	73.6	1.91	10.2	15.1	13.1	27.4	44.2	141
1991	0.91	20.7	3.02	81.7	56.6	41.7	0.721	64.5	1.27	9.04	15.4	11.6	32.4	40.5	114
1992	1.26	16.5	4.18	48.4	56.6	50.9	0.721	52.5		14.5	13.7	21	55.1	55	98.6
1993	0.843	12.9	1.88	44.2	68.7	43.6	0.683	45.2		9.61	12.8	5.62	26	35	93.7
1994	1.26	15.4	1.79	31.2	49.3	42.4	1.6	51.4		26.1	13.7	21	67	86.3	195
1995		84.8	5.77	144	206	226	4.52	296	17.3	37.9	60.1	53.7	103	129	544
1996	1.95	31.1	13.9	64.9	82.6	114	4.05	152	6.57	16.9	22.7	30.7	53.2	100	233
1997	3.1	18.2	2.08	38.1	34	44.5	1.73	54.4	6.05	26.3	11.3	19.3	21.8	57.5	157
1998	3.96	41.2	3.02	61.9	87.5	91.3	1.73	106	4.96	17.3	18	15.4	36.8	54.9	181
1999	3.96	31.6	2.89	72.7	80.4	93.1	1.93	110	4.13	21.1	28.8	26.3	55.9	80.7	226
2000	3.66	29.7	3.43	70.3	91.1	101	2.72	131	11.2	36.1	35.8	37.4	75.4	63.3	321
2001	1.76	30.2	1.79	37	50.3	49.4	1.5	60.4	8.07	15.9		15.8	34.2	46.9	127
2002	6.9	152	29	315	370	400	6.27	500	26.5	98.5	110	115	233	250	1000
2003	2.26	24.4	2.49	45.7	69.9	77.3	3.39	110	7.32	34.2	15.7	26.4	58	79.1	222
2004	1.5	19.3	4.62	41.1	58.7	57.7	2.1	79.8	12.7	52	19.7	30.3	95.9	77.2	175
2005	2.49	38	6.47	111	145	158	2.72	181	16.6	53.8	38.5	53	112	125	341
2006	2.37	38	4.29	83.7	117	124	2.65	128	9.96	36.3	47.6	42.1	91.7	103	253
2007	4.43	40.6	2.26	43.6	75.9	84.1	2.5	109	9.17	25.9	25.5	25.7	49.1	63.7	223
2008	3.1	48	3.25	55	84.5	91.1	2.18	127	16.1	60.7	22.4	27.6	64.3	108	300
2009		21	3.75	47.7	66.7	70.2	1.57	79.3	6.8	32.5	20.4	27.6	52.6	66.3	155
2010	2.26	42.2	5.38	114	164	195	2.57	265	16.3	64.8		120	187	179	521
2011	1.67	35.9	5.17	82.9	127	138	2.69	164	14.1	40.9		38.1	65.4	111	301
2012		15.7	2.32	30.3	49.7	53.5	1.27	60.5	2.92	18.2		17.7	45	61.1	134
2013	9.49	112	15.2	213	267	285	7.15	487	27.1		95.2	73.2	203	257	1010

Table A10.: Classification of the gauges in the subcatchment Freiberger Mulde/Flöha of the Mulde river basin according the method proposed in Section 7.3. It can be seen that large and extreme large events (magenta and red) are detected simultaneously for nearly all gauges.

Gauge-ID	567400	567700	567850	567420	568350	568400	568160	567451	567470	567000	566010	566040	567320	566100
Name of Gauge	**Tannen-berg**	**Wiesa**	**Strecke-walde**	**Hopf-garten**	**Rothen-thal**	**Zöblitz**	**Borsten-dorf**	**Lichten-walde 1**	**Krieb-stein UP**	**Wolfs-grund**	**Berthels-dorf**	**Nossen 1**	**Nieder-striegis 1**	**Erlln**
Begin	1960	1961	1921	1911	1929	1937	1929	1910	1933	1921	1936	1926	1926	1961
End	2013	2013	2013	2013	2013	2013	2013	2013	2013	2012	2013	2013	2012	2012
Length [a]	54	53	93	103	85	77	85	104	81	92	78	88	87	52
Length net. [a]	54	53	93	102	85	76	85	104	81	92	78	88	86	52
Median [m^3/s]	13.2	12.0	25.6	68.1	13.0	20.3	80.8	192.0	209.0	5.8	25.3	59.7	25.1	297.5
Variance [m^3/s]	9.4	9.0	18.6	52.2	9.9	14.7	61.8	149.9	152.0	3.8	26.9	52.9	17.4	193.7
3. moment	4767	4378	33905	705661	4562	143010	1094635	14595844	16889563	374	44776	384225	32066	43321064
k(T=2)	15.86	15.41	30.47	83.73	15.62	22.86	96.91	229.61	241.05	6.80	33.43	68.39	29.91	329.86
k(T=5)	21.51	20.91	41.34	113.61	21.20	31.02	131.49	311.53	327.05	9.22	45.35	92.79	40.58	447.55
k(T=15)	31.01	30.14	59.59	163.76	30.56	44.71	189.54	449.07	471.44	13.29	65.38	133.75	58.50	645.13
k(T=30)	39.08	37.99	75.10	206.38	38.51	56.35	238.87	565.95	594.13	16.75	82.39	168.56	73.72	813.03
number T>2	16	17	33	36	29	31	30	35	30	35	26	36	35	21
number T>5	8	9	14	18	17	13	15	15	16	12	12	13	14	10
number T>15	3	2	6	7	3	3	6	6	3	5	4	5	4	1
number T>15 and T<30	2	1	2	2	1	1	4	2	0	3	1	2	3	0
number T>30	1	1	4	5	2	2	2	4	3	2	3	3	1	1
Medium	0.15	0.15	0.20	0.18	0.14	0.24	0.18	0.19	0.17	0.25	0.18	0.26	0.24	0.21
Large	0.09	0.13	0.09	0.11	0.16	0.13	0.11	0.09	0.16	0.08	0.10	0.09	0.12	0.17
Very large	0.04	0.02	0.02	0.02	0.01	0.01	0.05	0.02	0.00	0.03	0.01	0.02	0.03	0.00
Extreme	0.02	0.02	0.04	0.05	0.02	0.03	0.02	0.04	0.04	0.02	0.04	0.03	0.01	0.02
AMS [m^3/s]														
1910								83.5						
1911				45.7				114						
1912				134				270						
1913				49.1				192						
1914				78				192						
1915				112				296						
1916				66				198						
1917				119				278						
1918				51.8				218						
1919				65				143						
1920				71.5				177						
1921			32.9	75.9				143		4.5				
1922			32.9	122				253		15				
1923			35.4	137				309		11				
1924			25.1	87.2				157		11.5				
1925			22.2	40.1				118		2.42				

1926			36.4	116				300		8.35		92.4	50.9
1927			26.7	95.4				250		8.84		61.7	34.9
1928			25.8	89.4				191		11.9		40.2	18.1
1929			15.4		5.8		35	90		2.7		29.9	20.7
1930			19.8	50	8.6		39.1	125		2.62		40.1	25.8
1931			13.1	40	8.98		36.6	91.1		4.7		38.3	26.5
1932			64	260	28.8		235	687		8.33		87.7	34.1
1933			23	59.2	13		77.3	178	201	10.1		35.5	24.8
1934			21.4	49.1	8.84		44.6	113	125	6.92		77.9	19.1
1935			30.4	68.6	14.4		81.2	195	173	5.54		82.9	46.2
1936			11.8	29.5	9.94		52.9	86.2	83.3	1.89	12.9	28.8	16.6
1937			32.6	108	19.3	37.7	192	395	352	8.83	36.2	84.3	52
1938			21.4	49.2	7.57	14.5	62	194	209	5.79	25.5	81.6	46.8
1939			27.2	61.2	16.5	23.4	99.8	196	185	3.28	23.6	42.8	16.6
1940			31.7	68.2	15.4	23	106	204	217	7.67	28	83.6	28.9
1941			25.3	88.9	10.9	24.9	92.8	208	263	4.64	34	137	
1942			23.7	69.5	6.06	17.8	79.9	232	119	4.67	23.7	54.4	23.4
1943			10.4	22.1	2.79	6.06	24.4	45.7	54.6	2.25	9.55	19.2	6.9
1944			37.3	84.3	11.8	23.9	122	234	241	7.58	39.8	71.5	36.2
1945			49	110	5.87	21.4	74.3	147	178	2.95	15.3	25.6	16.7
1946			38.6	92.6	12.3	22.5	137	334	328	7.5	42.5	97.2	44.6
1947			37.3	60.2	5.68	17.5	93.9	383	341	9.85	51.4	117	37.2
1948			39.3	91.4	16.3	26.3	112	298	303	8.59	35.4	68.6	23.7
1949			24.6	56.5	13.3	19.9	89.2	215	246	6.75	25.3	61.1	14.4
1950			14.2	37.2	8.73	13.5	33	81.4	88.2	3.73	10.6	50.4	17
1951			16.8	37.2	5.87	10.8	28.4	75.4	80.3	3.57	12.1	24.8	12.6
1952			20.3	41.3	8.29	14.4	50	97.6	90.6	2.67	13.2	32.1	11.8
1953			19.9	56.5	19.7	17.9	89.2	125	151	4.13	18.2	55.5	22.1
1954			81.1	232	14	37.4	147	511	449	8.42	68.2	136	61.8
1955			62.5	122	12.1	28.5	104	213	237	7.41	28.7	57.6	16.8
1956			18.7	53.7	5.54	11.6	54.4	224	248	6.42	25.3	68.6	34
1957			46.5	91.4	25.3	34	158	287	269	6.58	30.4	57.6	16.4
1958			81.1	229	17.3	34	208	532	428	8.59	120	205	69.1
1959			18.9	51.7	5.67	11.6	46.5	116	137	3.81	13.5	29.1	20
1960	12.4		25.6	64.4	9.13	20.7	80.9	196	222	4.93	28.2	55.5	39.2	
1961	12.7	6.96	18.4	48.2	6.21	13.1	50	130	163	4.01	18.9	37	20.2	214
1962	9.97	9.51	25.6	55.6	10.8	12.6	39.7	93.5	105	4.89	19.9	25.1	10.7	128
1963	6.14	6.64	11.8	94	3.85	8.43	22.8	135	54.6	4.19	14.2	38.1	15	168
1964	9.51	7.44	12.8	42.7	6.03	10.3	32.5	63.4	69.4	4.71	23.2	22.6	14.8	188
1965	15	11.8	33	79.3	15.4	26.3	143	388	367	9.05	61	98.6	37.4	542
1966	13.5	14.1	28	79.3	13.5	18.3	75	168	214	5.12	17	30	21.9	232
1967	16.2	13.5	28.5	73.7	22.2	19	80.9	189	221	6.41	31.2	43	19.5	242
1968	17.9	15.5	43.8	105	19.4	23.8	100	273	307	6.68	41.3	78.8	35.2	383
1969	8.04	12.1	31	70.4	13.9	19.3	72.6	175	184	8.7	22.2	38.1	21.3	226
1970	14	16.6	36	91.3	15.7	23.4	92.7	222	233	8.1	40.8	51.9	24.9	304
1971	5.8	6.74	12	37	6.85	9.55	41.3	84.6	103	3.94	19.5	29.2	14.2	135
1972	8.68	12.1	13.6	37.6	10.3	15.8	39.5	89.9	106	3.1	15.2	23.9	10	144

1973	10.6	12.8	16	44	8.04	13.7	36.7	97.2	103	7.29	20.2	34.7	14.6	168
1974	16.3	18.8	29.1	84.8	15.6	23	81.9	196	206	5.77	32.8	67	32.7	320
1975	29.7	28.3	54.6	198	44	40.2	233	623	595	8.83	54.7	87.8	30.5	610
1976	17.1	15.8	29.7	78.9	17.4	18.9	100	228	242	5.77	39.1	75.2	29.9	360
1977	8.06	7.51	17.3	37.4	24.8	23	113	205	257	6.87	31.2	77.7	36.6	387
1978	16.1	9.69	27.3	63.7	24	25.5	100	249	262	5.87	21.1	59.1	32.7	404
1979	9.54	15.6	13.9	38.2	10.6	12.5	42.7	95.6	127	2.4	23.8	44.3	28.7	215
1980	19.9	12.8	32.1	98.5	21.3	37.1	127	285	355	5.63	44.6	111	38.8	481
1981	15.6	27.6	45.8	136	23.6	27.8	134	414	371	6.02	42.1	81.4	33.7	471
1982	15.8	12.1	27.3	72.2	9.13	20.6	78.5	230	226	7.15	40.9	61.3	20.8	366
1983	13.2	8.27	37.3	67.9	26.4	29.1	127	409	426	7.76	37.9	87.8	59.7	569
1984	10.6	8.27	20.5	44.8	11.3	20	69.5	80.6	157	4.85	12.7	24.4	21.3	251
1985	10.2	5.92	9.65	39.8	4.59	9.38	41.8	87.5	157	2.78	16.2	32.6	22.3	210
1986	8.06	8.47	18.6	44	24	9.13	59	111	122	4.45	22.9	40	24.3	210
1987	25.6	30	49.7	170	37.8	35.6	156	300	436	5.24	33.4	80.2	47.3	488
1988	9.97	13.3	37.3	70.1	13.8	23.8	72.8	191	200	5.44	32.3	60.2	25.3	292
1989	13.5	9.27	24	65.8	17	20.9	80.8	205	231	5.44	23.8	46.7	28.1	303
1990	8.69	9.93	13.1	39	7.75	13.8	42.7	85.5	91	2.63	11.4	28.9	14.9	153
1991	8.69	6.93	16.8	38.2	11.3	20.9	59	111	100	2.19	9.27	17.4	19	111
1992	12.2	20.5	18.1	47.5	28.5	20.9	109	230	244	18	122	109	15.2	292
1993	13.2	10.2	17.3	60.7	12.4	19.9	59	134	136	5.06	22.4	39.9	13.5	199
1994	6.86	6.29	13.9	32.1	6.32	7.94	44.5	115	145	3.76	21.5	76.4	36.9	396
1995	26.4	12.9	35.9	120	17.4	23.4	108	298	298	5.19	23.8	78.4	35.8	453
1996	14.2	9.93	27.9	76.7	10.1	16.8	62.1	172	206	7.43	15.1	38.8	20.8	241
1997	6.27	6.49	8.86	43.1	9.4	15.8	52	87.5	103	6.21	15.8	33.5	15.2	141
1998	12.7	8.09	16.3	53.9	12.1	17.8	77.3	188	204	4.45	24.7	62.6	23.9	305
1999	17.5	17.4	55.7	93.3	13.5	84.6	116	294	315	8.07	41.7	72.4	30.5	400
2000	15.1	19.4	30.3	86	22.6		118	277	281	7.62	47.5	90.7	32.4	417
2001	6.38	5.94	6.91	29.3	6.89	9.91	33.1	66.3	86.7	2.91	15.5	31.2	14.8	162
2002	85	65	145	420	88	160	540	1250	1246	29.9	360	690	173	1659
2003	18.7	12	19.2	72.8	14.4	19.8	95.1	184	230	14.4	46.9	68.2	31.4	330
2004	15	9.79	19.5	64.6	12.3	14.5	81	123	190	4.32	21.9	47.2	36.9	234
2005	31.3	29.7	37.4	155	29.4	38.4	156	405	399	12.2	58	79.9	47.8	476
2006	24.5	24.5	46.1	137	27.4	40.7	160	361	356	15.2	63.9	101	47.4	444
2007	10.6	21.6	21.3	49.7	19.3	23.1	68.7	130	140	3.83	25	40.5	14.5	162
2008	10.3	11.6	20.7	55	13.1	17.3	70	152	192	3.52	24.5	76.5	31.9	328
2009	13.5	8.12	11.4	47.1	9.19	15.4	70	147	161	5.88	22.8	66.6	27.8	201
2010	13.2	9.01	12.9	46.3	13.1	15.4	89.3	192	243	4.03	34.9	119	52.6	412
2011	30.3	22.6	35	122	20.8	36.5	159	423	435	12.3	60.2	132	49.9	537
2012	10.3	8.12	11.9	38.5	10.1	24	73.7	132	140	6.03	25.7	46.7	23.7	190
2013	34.9	33.6	91.8	259	26.3	54.3	257	774	746		25.7	312		

Table A11.: Classification of the gauges in the Harz region according the method proposed in Section 7.3. It can be seen that large and extreme large events (magenta and red) are detected simultaneously for nearly all gauges.

Name of gauge	Ilsenburg	Hoppenstedt	Stolberg	Wippra	Mansfeld-Leimbach	Tanne	Königshütte	Elend	Trautenstein	Silberhütte	Meisdorf	Hausneindorf	Hanneckenbruch	Steinerne Renne	Mahndorf	Nienhagen	Minsleben	Benzigerode	Derenburg
Begin	1954	1930	1954	1937	1960	1961	1951	1951	1951	1949	1921	1981	1962	1971	1972	1983	1970	1976	1969
End	2014	2014	2014	2013	2013	2013	2013	2013	2013	2014	2014	2014	2013	2014	2014	2014	2013	2013	2013
Length brutto [a]	61	85	61	77	54	53	63	63	63	66	94	34	52	44	43	32	44	38	45
Length netto [a]	60	82	58	75	54	53	61	63	59	66	85	34	52	43	43	31	44	38	45
Median [m^3/s]	6.11	11.95	4.14	11.2	11.7	20.8	23.7	12.1	9.9	9.0	11.8	12.0	5.1	6.6	11.3	13.7	0.7	0.2	0.6
Variance [m^3/s]	5.7	9.9	3.8	10.3	10.1	20.5	22.5	12.4	6.3	8.6	11.7	12.2	5.5	4.6	9.9	8.4	1.0	0.2	0.9
3. moment	700.47	5303.15	273.86	3756.73	3956.89	39189.29	43151.68	6336.72	1956.24	1958.54	5646.95	9109.17	387.84	464.14	4547.04	6128.94	1.86	0.03	1.97
2	8.37	16.42	6.12	14.64	14.90	31.97	33.01	17.43	11.79	11.79	16.77	19.67	6.88	7.30	15.61	17.24	1.16	0.30	1.18
8	13.28	26.06	9.71	23.23	23.64	50.73	52.39	27.66	18.70	18.71	26.62	31.21	10.91	11.58	24.76	27.35	1.84	0.47	1.88
14	16.00	31.40	11.70	27.99	28.48	61.12	63.12	33.32	22.53	22.54	32.07	37.60	13.14	13.95	29.84	32.96	2.22	0.57	2.27
37	22.12	43.40	16.17	38.69	39.37	84.48	87.24	46.06	31.14	31.15	44.32	51.97	18.17	19.29	41.24	45.55	3.07	0.79	3.13
number T> 2	17	30	19	23	18	17	20	16	21	22	29	11	17	14	15	10	10	11	11
number T>8	9	10	9	8	5	7	9	8	5	10	9	4	5	5	5	4	5	4	6
number T>14	4	3	3	5	4	4	4	5	2	4	6	2	3	3	4	2	4	3	4
number T>14 and T<37	2	2	2	3	3	3	2	3	1	2	3	1	1	1	2	2	2	3	2
number T>37	2	1	1	2	1	1	2	2	1	2	3	1	2	2	2	0	2	0	2
Medium	0.13	0.24	0.17	0.2	0.24	0.18	0.18	0.13	0.27	0.18	0.24	0.21	0.23	0.21	0.23	0.19	0.11	0.18	0.11
Large	0.08	0.08	0.10	0.04	0.02	0.06	0.08	0.05	0.05	0.09	0.04	0.06	0.04	0.05	0.02	0.06	0.02	0.03	0.04
Very large	0.03	0.02	0.03	0.04	0.05	0.06	0.03	0.05	0.02	0.03	0.04	0.03	0.02	0.02	0.05	0.06	0.05	0.08	0.04
Extreme	0.03	0.01	0.01	0.023	0.02	0.02	0.03	0.03	0.02	0.03	0.04	0.03	0.04	0.05	0.05	0.00	0.05	0.00	0.04
JahresHQ [m^3/s]																			
1921											9.84								
1922																			
1923																			
1924																			
1925																			
1926																			
1927											22.3								
1928											17								
1929											12.4								
1930		6.84									8.57								
1931		13.6									11.2								
1932		21.8									7.4								
1933		6.16																	

1934		6.16								4.99									
1935		17.7																	
1936		5.45								12.5									
1937		12.1		13.8						9.48									
1938		23.6		23						23.2									
1939		8.26		13.2						14.7									
1940		29.6		24.4															
1941		29.4		15.4						15.7									
1942		28.4		29.1						19.9									
1943		3.2		2.85						4.14									
1944		8.39		10.7						16.3									
1945																			
1946		24.9		13.5						39.6									
1947										20.6									
1948		18.1		55						65.7									
1949		8.26		5.42					4.97	6.83									
1950		9.03		14.5					5.61	10.9									
1951		3.97		13			15.2	10.5	4.31	6.85	9.93								
1952		5.25		10.3			13.3	9.61	3.07	4.37	10.7								
1953		4.44		12.1			13	4.09	7.39	14.6	23.1								
1954	3.4	6.36	0.99	3.71			15.6	31	7.11	6.85	5.65								
1955	15.2	22.2	19.6	14.8			40.5	27.1	18.6	21	29.4								
1956	4.29	13	5.66	16.1			23	5.17		12.5	54.3								
1957	6.51	13.7	1.76	3.5			22.4	16.2		5.74	11.3								
1958	32	31	3.14	12.7			23.7	39.9	12.5	9.32	20.4								
1959	3.61	1.91	1.04	2.33			7.06	4.88	7.97	4.02	7.65								
1960	6.51	7.7	1.55	4.62	3.3		22.8	18.9	16.5	12.1	19.5								
1961	20.8	16.3	6.53	14.7	13.8	15.3	34.7	83.3	25.7	16	22.3								
1962	11.5	9.03	3.54	13.4	13.8	12.9	19.8	16.2		8.2	11.8	5.23							
1963	2.48	5.57	1.96	3.86	3.3	11	19.8	2.17		4.49	6.01	0.68							
1964	6.62	4.15	1.89	6.36	7.8	10.3	16.1	10.5	7.97	5.22	10.4	2.26							
1965	3.71	6.71	1.68	11.2	13.9	13.7	22.5	7.16	7.61	10.5	16.2	1.58							
1966	13.6	14.2	11.2	38	21.4	54.5	37.6	24.4	14.5	47	30.2	5.93							
1967	7	12.5	2.4	15.1	12.9	18.1	17.4	8.74	6.62	8.85	11.3	2.35							
1968	4.28	18.1	5.02	6.48	10.9	20.8	26.3	8.74	11.5	4.87	17.4	2.99							
1969	5.21	11.7	2.56	35.8	18.2	18	24.3	14.3	10.7	12.2	11.5	3.67							3.5
1970	4.91	11.5	7.52	22.9	20.3	29	26.8	14.9	9.9	12.8	16.5	4.67				1.01			0.73
1971	3.82	5.36	7.64	8.83	10.2	23.8	20.7	9.17	8.46	20.3	12.6	3.87	4.72			0.23			1.36
1972	3.82	6.95	1.9	5.96	6.23	14.9	11.2	5.48	21.7	3.09	5.05	3.5	3.8	8.9		0.71			1.82
1973	2.91	3.35	2	3.83	5.33	8.6	7.59	3.85	5.54	4.65	4.9	1.73	2.16	7.42		0.76			0.66
1974	3.34	10.8	4.09	10.9	7.99	16.4	19	9.17	9.9	7.3	8.85	6.68	6.9	8.46		0.33			0.69
1975	10.3	11.1	6.28	10.9	21.6	40.1	29.3	33.1	15.7	9.15	13.8	11.5	13	17.3		1.88			0.7
1976	8.12	10.5	3.4	9.3	10.5	21.4	25.3	12.1	10.7	9.45	11.8	5.45	5.48	12.7		0.8	0.103		0.3
1977	5.84	6.01	1.65	12.6	15.5	10.7		6.46	17.1	5.85	7.85	3.62	4.72	9.34		0.92	0.141		0.77
1978		6.48	2.2	7.56	5.99	46.2	8.79	4.6	9.16	7.15	6.73	4.26	4.48	7.04		0.67	0.114		0.48
1979	3.86	6.95	2.8	7.12	9.7	28.9	5.43	3.4	8.24	9.45	8.45	4.4	3.42	5.9		0.67	0.183		0.38
1980	10.7	25.8	2.4	8.32	7.16	47.2	26.5	13.1	46.8	3.88	5.2	7.96	6.04	18.8		1.6	0.074		4.22

1981	11.5	26.3	7.25	11.1	21.4	64.1	37.8	22.2	17.5	11.6	11.8	13	9.31	8.1	31.4		2.75	0.256	1.55
1982	7.5		10.1	11.7	18.5	18.5	16.4	8.66	8.24	26.9	17.7	16.8	8.4	7.5	15.2		1.01	0.13	0.57
1983	14	8.72	5.3	4.26	5.15	17.4	23.7	17.4	8.69	5.25	9.45	7.15	11.3	8.44	25.8	24.7	1.48	0.399	0.55
1984	5.51	19.4	2.55	7.24	9.56	16.9		15.3	8.46	6.7	8.65	8.6	9.79	8.1	18.2	13.7	0.88	0.222	1.32
1985	7.3	18.8	4.28	7.12	11	15.9	22.9	8.09	10.7	7.86	7.45	8.55	7.54	6.56	11.4	13.9	0.57	0.242	0.59
1986	9.85	21.2	6.84	12.6	14	34.6	36.9	11.3	8.52	13.8	21.3	17.3	10.3	8.48	16.1	15.5	1.1	0.171	0.57
1987	6.38	22.4	12.2	17.7	18.8	33.8	51.4	11.3	13.9	13.8	21.3	19.8	8.4	9.56	18.7	18.1	5.51	0.306	1.1
1988	2.69	17.3	5.94	11.9	16.7	9.25	16.3	6.05	6.66	11.7	17.4	17.6	2.2	2.1	9.3		1.1	0.415	0.75
1989	4.91	18.3	3.03	5.84	5.4	18	27.9	9.07	7.58	5.82	5.85	6.35	6.73	5.89	11.3	12.1	0.714	0.079	0.115
1990	5.59	6.44	4.92	16.2	14.3	14	23.5	12.1	8	7.01	8.71	10.1	3.28	3.98	6.16	6.45	0.598	0.25	0.272
1991	11.8	16.9		12.9	10.1	20.9	33.4	19.6	9.66	9.76	14.5	12.3	3.28	6.74	11.3	8.25	0.582	0.173	0.115
1992	5.41	10.4		7.13	5.19	26.2	27.2	7.91	6.77	7.01	7.11	5.68	5	5.63	10	9.98	0.442	0.055	0.059
1993	5.07	8.6	3.67	9.62	7.43	26.9	37.8	12.5	10.1	6.6	7.11	7.75	7.83	4.88	8.87	11	0.234	0.048	0.15
1994	17.3	34	11.4	79.8	83.3	66.2	69.1	44.6	13.9	74	85.7	60.3	27.4	37.2	49.5	34	6.21	0.618	2.58
1995	8.21	14.6	6.17	17.8	19.9	52	57.8	15.1	9.94	17.1	18.1	21	6.92	8.34	15.3	15.8	0.596	0.132	0.294
1996	4.43	11.8	1.05	5.17	7.12	7.96	10.4	7.36	4.96	5.43	6.68	6.96	6.64	3.78	7.56	12	0.947	0.588	2.82
1997	14.2	28.5	3.03	10.3	12.4	18	19.3	30.9	6.6	9.26	12.2	15.8	15.9	15.6	16.7	18.5	0.392	0.106	0.331
1998	13.9	30.2	6.77	14.5	13.8	137	163	56	18.5	23.5	22.8	24	8.17	19.5	31	28.1	1.17	0.279	0.781
1999	6.52	23	10.7	23.6	30.8	33	57.9	11.3	15.2	20.6	20.8	23.6	4	7.29	19.9	20.6	0.678	0.211	0.313
2000	7.33	16	4.3	7.4	8.18	52	57.9	23.1	11.3	6.93	8	8.06	6.82	8.96	14.8	14	0.599	0.18	0.295
2001	5.53	9.24	2.13	5.07	7.27	26.2	34.8	16.6	9.38	5.91	6.9	8.43	4.4	6.98	8.49	11.4	0.392	0.139	0.175
2002	36.6	50.1	11.8	20.1	20.1	70.1	106	37	19.1	20	21	23	45.8		56	34.9	2.75	0.578	1.97
2003	4.74	21.5	9.12	23.1	29.1	33.8	45.3	11	15.7	18.9	24.5	37.4	4.54	6.69	19	22.9	1.25	0.384	2.03
2004	4.89	8.88	3.43	4.55	5.95	19.7	21.8	13.8	10.2	6.76	6.48	7.35	9.46	6.38	6.68	8.22	0.278	0.075	0.202
2005	5.05	8.88	4.19	6.32	10.4	19.5	24.9	12.1	12.4	8.79	12.2	10.4	6.11	5.8	8.87	8.9	0.479	0.562	0.453
2006	3.71	8.52	6.23	10.9	13.8	27.9	50.7	12.1	12.6	12.3	15.7	11.6	4.98	7.31	8.11	9.97	0.244	0.183	0.343
2007	8.17	31.7	11.1	26.9	35.5	38.9	56.4	19.8	16	20	35.2	38	10.2	12.4	23.8	29	1.13	0.376	1.4
2008	5.43	13.3	10.3	6.83	7.66	47.4	57.9	17.6	14.2	9.29	10.9	9.87	6.11	7.2	11.2	11.7	0.761	0.168	0.407
2009	2.94	4.79	1.68	5.3	6.56	7.48	9.68	4.99	2.34	4.36	4.44	3.6	2.33	2.11	4.82	4.45	0.278	0.058	0.197
2010	4.54	13.5	5.91	15.3	19	15	16.4	7.32	8.25	10.9	26.5	25.1	1.45	4.23	8.43	9.77	0.574	0.286	0.818
2011	7.49	20.2		19.1	24.4	41.6	65.9	12.8	19	17.6	37.3	37.6	0.771	5.78	12.8	15.4	0.725	0.338	0.791
2012	3.08	4.7	3.15	4.93	5.55	17.4	22.2	5.67	6.02	7.86	9.29	9.75	3.45	3.71	6.39	8.14	0.303	0.115	0.319
2013	10.6	25.5	4.75	15.8	16.2	17.1	20.5	12.5	5.72	14.9	30.7	29.4	4.27	6.88	20.1	21.3	1.17	0.449	1.17
2014	11.5	25	3.75							3.85	7	7.37		4.91	8.16	11			