Institute of Earth and Environmental Sciences

Improving flood frequency analysis by integration of empirical and probabilistic regional envelope curves

Cumulative Dissertation

for the degree of
Doctor of Natural Sciences (Dr. rer. nat.)
in Geoecology

submitted to the
Faculty of Mathematics and Natural Sciences
at the University of Potsdam, Germany

by

Björn Felix Guse

Potsdam, July 2010

Published online at the
Institutional Repository of the University of Potsdam:
URL http://opus.kobv.de/ubp/volltexte/2010/4926/
URN urn:nbn:de:kobv:517-opus-49265
http://nbn-resolving.org/urn:nbn:de:kobv:517-opus-49265

Contents

List of Figures

List of Tables

Abstract

Flood design necessitates discharge estimates for large recurrence intervals. However, in a flood frequency analysis, the uncertainty of discharge estimates increases with higher recurrence intervals, particularly due to the small number of available flood data. Furthermore, traditional distribution functions increase unlimitedly without consideration of an upper bound discharge. Hence, additional information needs to be considered which is representative for high recurrence intervals.

Envelope curves which bound the maximum observed discharges of a region are an adequate regionalisation method to provide additional spatial information for the upper tail of a distribution function. Probabilistic regional envelope curves (PRECs) are an extension of the traditional empirical envelope curve approach, in which a recurrence interval is estimated for a regional envelope curve (REC). The REC is constructed for a homogeneous pooling group of sites. The estimation of this recurrence interval is based on the effective sample years of data considering the intersite dependence among all sites of the pooling group.

The core idea of this thesis was an improvement of discharge estimates for high recurrence intervals by integrating empirical and probabilistic regional envelope curves into the flood frequency analysis. Therefore, the method of probabilistic regional envelope curves was investigated in detail. Several pooling groups were derived by modifying candidate sets of catchment descriptors and settings of two different pooling methods. These were used to construct PRECs. A sensitivity analysis shows the variability of discharges and the recurrence intervals for a given site due to the different assumptions. The unit flood of record which governs the intercept of PREC was determined as the most influential aspect.

By separating the catchments into nested and unnested pairs, the calculation algorithm for the effective sample years of data was refined. In this way, the estimation of the recurrence intervals was improved, and therefore the use of different parameter sets for nested and unnested pairs of catchments is recommended.

In the second part of this thesis, PRECs were introduced into a distribution function. Whereas in the traditional approach only discharge values are used, PRECs provide a discharge and its corresponding recurrence interval. Hence, a novel approach was developed, which allows a combination of the PREC results with the traditional systematic flood series while taking the PREC recurrence interval into consideration. An adequate mixed bounded distribution function was presented, which in addition to the PREC results also uses an upper bound discharge derived by an empirical envelope curve. By doing so, two types of additional information which are representative for the upper tail of a distribution function were included in the flood frequency analysis. The integration of both types of additional information leads to an improved discharge estimation for recurrence intervals between 100 and 1000 years.

Zusammenfassung

Abschätzungen von Abflüssen mit hohen Wiederkehrintervallen werden vor allem für die Bemessung von Extremhochwässern benötigt. In der Hochwasserstatistik bestehen insbesondere für hohe Wiederkehrintervalle große Unsicherheiten, da nur eine geringe Anzahl an Messwerten für Hochwasserereignisse verfügbar ist. Zudem werden zumeist Verteilungsfunktionen verwendet, die keine obere Grenze beinhalten. Daher müssen zusätzliche Informationen zu den lokalen Pegelmessungen berücksichtigt werden, die den Extrembereich einer Verteilungsfunktion abdecken.

Hüllkurven ermitteln eine obere Grenze von Hochwasserabflüssen basierend auf beobachteten maximalen Abflusswerten. Daher sind sie eine geeignete Regionalisierungsmethode. Probabilistische regionale Hüllkurven sind eine Fortentwicklung des herkömmlichen Ansatzes der empirischen Hüllkurven. Hierbei wird einer Hüllkurve einer homogenen Region von Abflusspegeln ein Wiederkehrintervall zugeordnet. Die Berechnung dieses Wiederkehrintervalls basiert auf der effektiven Stichprobengröße und berücksichtigt die Korrelationsbeziehungen zwischen den Pegeln einer Region.

Ziel dieser Arbeit ist eine Verbesserung der Abschätzung von Abflüssen mit großen Wiederkehrintervallen durch die Integration von empirischen und probabilistischen Hüllkurven in die Hochwasserstatistik. Hierzu wurden probabilistische Hüllkurven detailliert untersucht und für eine Vielzahl an homogenen Regionen konstruiert. Hierbei wurden verschiedene Kombinationen von Einzugsgebietsparametern und Variationen von zwei Gruppierungsmethoden verwendet. Eine Sensitivitätsanalyse zeigt die Variabilität von Abfluss und Wiederkehrintervall zwischen den Realisationen als Folge der unterschiedlichen Annahmen. Die einflussreichste Größe ist der maximale Abfluss, der die Höhe der Hüllkurve bestimmt.

Eine Einteilung in genestete und ungeneste Einzugsgebiete führt zu einer genaueren Ermittlung der effektiven Stichprobe und damit zu einer verbesserten Abschätzung des Wiederkehrintervalls. Daher wird die Verwendung von zwei getrennten Parametersätzen für die Korrelationsfunktion zur Abschätzung des Wiederkehrintervalls empfohlen.

In einem zweiten Schritt wurden die probabilistischen Hüllkurven in die Hochwasserstatistik integriert. Da in traditionellen Ansätzen nur Abflusswerte genutzt werden, wird eine neue Methode präsentiert, die zusätzlich zu den gemessenen Abflusswerten die Ergebnisse der probabilistischen Hüllkurve – Abfluss und zugehöriges Wiederkehrintervall - berücksichtigt. Die Wahl fiel auf eine gemischte begrenzte Verteilungsfunktion, die neben den probabilistischen Hüllkurven auch eine absolute obere Grenze, die mit einer empirischen Hüllkurve ermittelt wurde, beinhaltet. Damit werden zwei Arten von zusätzlichen Informationen verwendet, die den oberen Bereich einer Verteilungsfunktion beschreiben. Die Integration von beiden führt zu einer verbesserten Abschätzung von Abflüssen mit Wiederkehrintervallen zwischen 100 und 1000 Jahren.

Acknowledgements

First of all, I thank my principal supervisor Bruno Merz for being an excellent supervisor and supporting me during all steps of my work with a lot of engagement and patience. He was always interested in discussing my ideas and results and he gave me many precious recommendations.

I am also thankful to Axel Bronstert, Jan Szolgay and Günter Meon for the acceptance to read and review my thesis.

For the discussions of probabilistic regional envelope curves and his precious recommendations while writing two papers, I am thankful to Attilio Castellarin.

Annegret Thieken helped me to find the objective of my thesis and to focus my research. I thank her for our discussions.

I thank Thomas Hofherr for the collaboration in the CEDIM group "Synopsis of the natural hazards in Saxony" and for the discussion of upper bound functions.

During my thesis, I had several discussions with Steffi Uhlemann. Maybe she cannot remember a day without at least one short question. I thank her for her engagement to find answers immediately and for finding solutions for confusing English sentences.

For the enjoyable atmosphere during the 3 ½ years, I thank all my colleagues from the section "Hydrology" at the GFZ. Particularly, I like to thank Benjamin Creutzfeldt, Knut Günther, Manuela Nied and Theresa Blume. Every one of them helped to me to write this thesis. And I thank Uli Herrmann for his GIS support which avoided many frustrating hours.

I also thank the colleagues from the Institute of Earth and Environmental Sciences at the University of Potsdam and from the Center of Disaster Management and Risk Reduction Technology (CEDIM) for helpful advices and the pleasant collaboration. I thank CEDIM and GFZ for financing my work.

I am grateful to Benjamin Gaede and Christopher Melching for reading my manuscripts in a short period of time and improving the English spelling.

Several institutions contributed data for this research. This is gratefully acknowledged.

I am thankful to my parents for supporting me during my studies of geoecology and this PhD work.

And, of course, I am deeply grateful to Christiane for being a lovely partner at my side and supporting me with all her capabilities while I was writing my thesis. She was a wonderful listener and encouraged me in times of frustrations caused by any paper-related problems. I thank her so much for taking care of our son Jannes, particularly during the last months. And I thank Jannes for being the self content small and yet already distinct person who allowed me to finish my work.

CHAPTER 1

INTRODUCTION

1 Introduction

1.1 Motivation

1.1.1 Flood risk analysis

The flood in August 2002 along the Elbe and the Danube rivers emphatically revealed that extreme floods are a relevant natural hazard in Germany. More than one hundred dikes breached and created large inundations along the Elbe river and its tributaries (e.g. Vorogushyn et al., 2010), which caused damages of a severity never observed before in Germany (damage costs of 11.6 billion EUR and 21 fatalities, e.g. Thieken et al., 2005).

In Germany, the 2002 flood had an especially large effect on the federal state of Saxony. The highest ever measured maximum discharges, also called floods of record, were exceeded at several gauges, especially in the Mulde and in the western tributaries to the Elbe, and discharges up to six times larger than the former floods of record were observed.

A probability of a specific discharge occurring and the resulting inundation area which is affected by this flood magnitude are designated as flood hazards. In a flood risk analysis not only the flood hazard, but also the vulnerability are considered, which includes the exposure of the given assets and the damage which is caused due to a certain degree of inundation (susceptibility) (see Fig. 1.1, Apel et al., 2004; Merz and Thieken, 2004; Merz, 2006).

Floods are one among several natural hazards which might evoke severe damages. To consider the risk reduction potential efficiently, the damages which are caused by different natural hazards needs to be compared (Durham, 2003). Grünthal et al. (2006) compared the risks of storms, floods and earthquakes for Cologne/ Western Germany and estimated that earthquakes are the dominant natural hazard for recurrence intervals larger than 200 years. For smaller recurrence intervals, floods caused the largest loss.

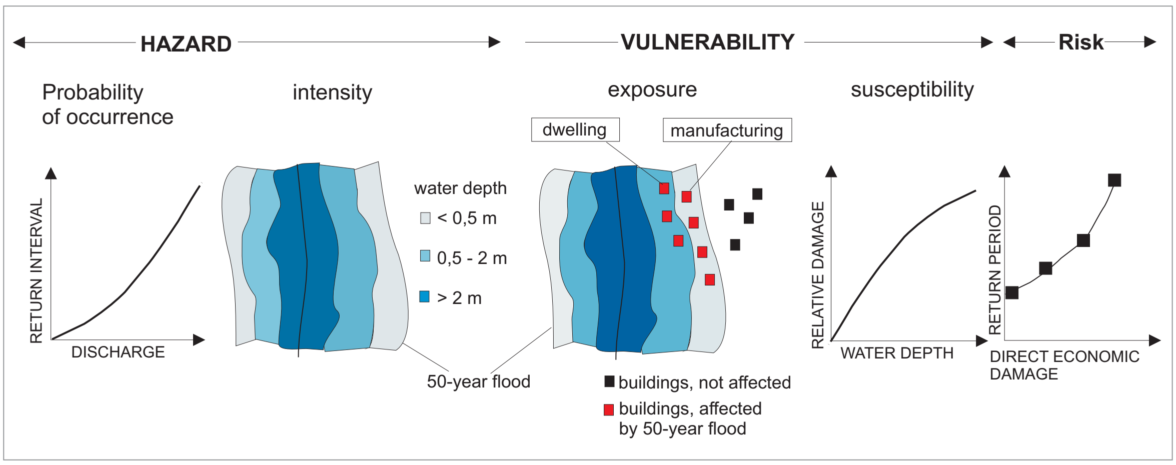

Fig. 1.1: Flood risk chain from (Merz et al., 2007).

Based on this, the CEDIM project "Synopsis of natural hazards in Saxony" was realised, in which the risks caused by earthquakes, winter storms and floods for each Saxon community were compared. This thesis is embedded in the flood part of this project and is only focused on flood frequency analyses, the first step of flood hazard estimation (see Fig. 1.1). Merz et al. (2002) and Merz et al. (2008) pointed out that the largest uncertainty in flood risk estimations is related to the flood frequency analysis when estimating floods with a large recurrence interval.

1.1.2 Flood frequency analysis – Limits in the discharge estimation for large recurrence intervals

In a flood frequency analysis, the probability of a given flood magnitude is estimated based on the available measured discharges at a gauge. Traditionally, the annual maximum series (AMS), i.e. the largest discharges of each hydrological year, is used and illustrated as plotting positions (empirical probabilities) (e.g. Cunnane, 1978). An appropriate distribution function provides flood quantiles, i.e. pairs of a discharge and the corresponding recurrence interval, along the whole range of recurrence intervals for this site under study (Fig. 1.2) (e.g. Stedinger et al., 1993).

The accuracy of the flood quantile estimates depends on the available number of flood data. The uncertainty of discharge estimates increases with large recurrence intervals (e.g. Merz and Thieken, 2005; Chbab et al., 2006). The example of a flood frequency analysis (Fig. 1.2) illustrates the effect of the selection of a distribution function: With increasing recurrence intervals, the four adequate distribution functions spread out. Hence, the discharge estimations of larger T are considerably affected by the selection of the distribution function. Furthermore, it is shown that the largest flood is not very well represented by the proposed distribution functions.

There are 30 – 100 years of observations for the majority of the gauges. However, for flood design estimates, discharges which are representative for extreme events (T = 1000 - 10,000) are of major interest and need to be calculated (e.g. Büchele et al., 2006; Petrow et al., 2006). Hence, the available data length contrasts with target recurrence intervals of at least hundred years and in several cases thousand or even more years.

When estimating very large recurrence intervals, it is worth mentioning that traditional distribution functions increase unlimitedly without consideration of an upper bound. This implies that the distribution function may estimate discharges for large recurrence intervals which can be assumed as unrealistically large due to the hydro-meteorological situation in the catchment of interest (Enzel et al., 1993), especially in the case of a high skewness.

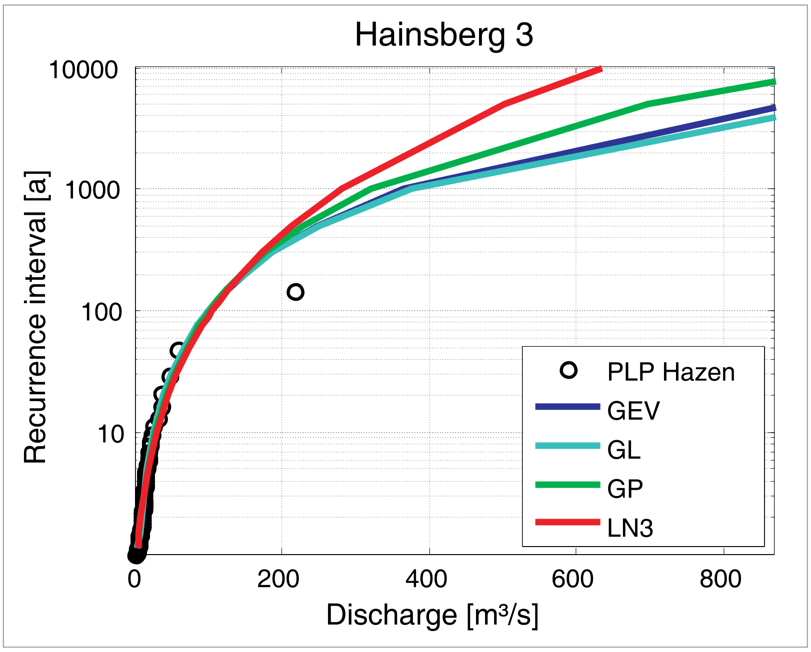

Fig. 1.2: Example of a flood frequency analysis for the site Hainsberg 3 in Saxony. Four distribution functions are shown (GEV = Generalised Extreme Value, GL = Generalised Logistic, GP = Generalised Pareto, LN3 = Log-Normal with three parameters). The observed flood data are illustrated as Hazen plotting position (PLP Hazen).

The critical point of an unlimited increase of unbounded distribution functions was taken into consideration by using a distribution function with an upper bound. A couple of bounded distribution functions have been presented (Kanda, 1981; Eliasson, 1994; Francés and Botero, 2003; Jensen et al., 2004). These distribution functions differ from the traditional ones by introducing an upper bound as additional information, which prevents the function from an unlimited increase. However, by using an upper bound, no representative flood data for recurrence intervals in the order of 1000 years is included.

In this context, it is interesting to compare the flood frequency analysis with the storm and earthquake research. For storm frequency analysis, there is a comparable number of data available which also leads to an increasing uncertainty for larger recurrence intervals (Pandey et al., 2001). In contrast, the earthquake research is concentrated on very large recurrence intervals in the order of 500 to 10,000 years (Grünthal and Wahlström, 2006; Grünthal et al., 2006). For a consistent risk comparison with earthquake estimates in the CEDIM project, quantile estimates in the flood and storm research are required for up to 1000-years events.

It hence becomes apparent that for discharge estimations of large recurrence intervals, additional information besides the at-site flood data needs to be included in a flood frequency analysis. This is required both for flood design estimations and for a comparison with other natural hazards. Furthermore, it is especially important to integrate information which is representative for the target recurrence interval, i.e. in our case for T between 100 and 1000 years.

1.1.3 Flood regionalisation

In hydrologic research, three types of additional information to extend the available data are separated. These are spatial (flood regionalisation), temporal (historical data) and causal (flood generation processes) information (Merz and Blöschl, 2008a, b). In this thesis, only spatial additional information is considered. In a flood regionalisation, flood data from neighbouring sites is used in addition to the at-site data based on the principle of „trading space for time“ (Stedinger et al., 1993).

A flood regionalisation includes two steps (e.g. GREHYS, 1996b). First, a homogeneous group of sites is formed (pooling group). This means that sites which are characterised by similar hydrologic conditions as the site of interest are selected. Therefore, relevant hydrologic parameters need to be determined, controlling catchment behaviour. A similarity in these parameters leads to the assumption that these sites could be collected together in a pooling group with the aim of a larger amount of flood data available (e.g. Acreman and Sinclair, 1986; Burn, 1997; Castellarin et al., 2001).

Second, a suitable regionalisation method which controls the transfer of information from the neighbouring sites to the site of interest is applied for this pooling group (e.g. GREHYS, 1996a, b). During the last decades, a couple of pooling grouping methods (e.g. cluster analysis, Region of Influence) and regionalisation approaches (e.g. index flood, multiple regressions, geostatistical approaches (Top-kriging)) have been developed and several studies considered their use for flood regionalisation (e.g. Mosley, 1981; Burn, 1990b; GREHYS, 1996a, b; Ouarda et al., 2001; Merz and Blöschl, 2005; Rao and Srinivas, 2006; Skoien et al., 2006; Ouarda et al., 2008). The benefit of flood regionalisation in comparison to at-site flood statistics was shown in several studies. It was noted that even when only a small number of neighbouring sites is available or in the case of slightly heterogeneous regions, regional estimates outperform at-site estimates (e.g. Cunnane, 1988; Stedinger and Lu, 1995; GREHYS, 1996a; Madsen and Rosbjerg, 1997b; Kjeldsen and Rosbjerg, 2002).

When adding data from neighbouring sites, the novel information is not necessarily identical with the data length. A neighbouring site can include the same flood events in its AMS. In this case, the additional gain of information is less than the length of the added flood series due to the cross-correlations among the flood series. The effective data length is lower than the total data length and needs to be considered to express the real information content of the data (Matalas and Langbein, 1962). Hence, intersite dependence among the sites needs to be taken into account in a flood frequency analysis (e.g. Stedinger, 1983; Madsen and Rosbjerg, 1997a; Castellarin et al., 2008).

The selection of adequate pooling grouping procedures and flood regionalisation methods depends on the research question. This study aims at estimating discharges for large recurrence intervals. Hence, a regionalisation method which is focused on the estimation of large flood quantiles needs to be selected.

1.1.4 Empirical and probabilistic regional envelope curves

Envelope curves are one possibility to regionalise maximum flood discharges. The method of envelope curves estimates an upper bound discharge (Jarvis, 1925). For this, the floods of record are normalised by their catchment sizes and related to the catchment size. Then, an upper line is drawn above all floods of record and an upper bound discharge is derived for all sites of a specific region (e.g. Crippen and Bue, 1977; Herschy, 2002; Castellarin et al., 2005).

Empirical envelope curves are based on observed discharge data only, and their magnitude is determined by the largest unit flood of record within the data. Due to the limited amount of available data, it is doubtful if the envelope curve really represents an upper bound discharge which might not be exceeded in future. In fact, it is recommended to update the empirical envelope curves at regular intervals (Crippen and Bue, 1977; Matalas, 1997).

There are two possibilities to resolve this methodical limit. First, an exceedance probability, i.e. the inverse of the recurrence interval, needs be assigned to the envelope curve. The lack of an exceedance probability is considered as a limit of empirical envelope curves (Castellarin et al., 2005; Gaume et al., 2010).

Alternatively, a larger flood data set from a wider geographical context can be used. Further, causal information needs to be considered which helps to explain that the envelope curve based on a larger data set can be assumed as an absolute upper bound for the study region.

Recently, the first of these limits was intensively considered, and in this way the concept of envelope curves was significantly improved by proposing a probabilistic approach for regional envelope curves (Castellarin et al., 2005). In this method, a recurrence interval is assigned to the regional envelope curve (REC).

The method of probabilistic regional envelope curves (PRECs) is based on a stricter definition of regional homogeneity within the pooling group of sites. To adequately consider the neighbouring concept of flood regionalisation, a pooling group is required which is homogeneous according to the index flood concept, meaning that the flood series of all sites of the pooling group normalised by the mean of their AMS can be represented by the same distribution function (Dalrymple, 1960; Robson and Reed, 1999).

The determination of the recurrence interval is based on the effective data length of the pooling group. The core point of the calculation of the effective data length is an accurate estimation of the intersite dependence among all sites of the pooling group (Castellarin et al., 2005; Castellarin, 2007).

Finally, a PREC derives a pair consisting of a discharge and its corresponding recurrence interval (PREC flood quantile) for each site of the pooling group. Since this PREC flood quantile is representative of a large recurrence interval, this method is related to discharge estimates of large recurrence intervals. Whereas a distribution function provides discharges for all recurrence intervals (with increasing uncertainty for larger T), a PREC derives the discharge of one specific recurrence interval only. Its adequacy as a flood regionalisation method was confirmed by Castellarin (2007) by comparing PREC flood quantiles with the traditional index flood method.

1.2 Research question and objectives

Based on the introducing statements, the core research question of this thesis is emphasised. It was highlighted that discharge estimates for large recurrence intervals are uncertain. However, accurate estimates are required for flood design as well as for a comparison of natural hazards. To improve the estimations, additional information which is adequate for large recurrence intervals needs to be considered. Hence, this thesis aims at resolving the following research question:

Can the estimation of large flood quantiles be improved by the integration of empirical and probabilistic regional envelope curves into the flood frequency analysis?

For this, six sub-questions were selected, which were separately considered in two parts. First, probabilistic regional envelope curves are examined in detail. In the second part, PREC flood quantiles as well as an upper bound discharge derived from an empirical envelope curve are inserted into the flood frequency analysis.

First part:

Whereas empirical envelope curves are a well-known method, probabilistic regional envelope curves were only applied in one initial study with fixed homogeneous regions (Castellarin, 2007). The effect of different constitutions of the regions on PREC flood quantiles was not investigated. To resolve the research question, the PREC concept needs to be examined in detail to assess the value of the PREC flood quantiles as additional spatial information for a FFA. It is necessary to examine the influence of the most relevant methodical aspects and to estimate the variability of the probability estimations of the PRECs using different assumptions.

For this, PRECs are investigated in two steps according to the PREC concept. The first step is the construction of a REC which is mostly affected by the formation of the pooling group, i.e. the determination of the neighbouring sites which were used to construct the REC.

The second step is an assignment of the recurrence interval to the REC. It is particularly related to the estimation of the effective sample of years of data which considers the effect of intersite dependence and which is a prerequisite for the calculation of the recurrence interval of a PREC. Thus, it is relevant to examine the cross-correlations among the sites accurately to express the real information content of the data of the pooling group.

Thus, three sub-questions which consider the major aspects of the two PREC steps need be resolved:

1. **Which pooling method is appropriate to construct RECs?**

2. **How large is the effect on the flood quantiles estimated by probabilistic regional envelope curves when modifying the constitution of the pooling group?**

3. **How large is the effect on the recurrence interval of PREC when refining the estimation of the effective sample years of data?**

These points were examined in chapters two and three, which correspond to the first and second paper, respectively:

In the first paper, the first step of the PREC concept, i.e. the construction of the regional envelope curve, is the focus of research. Here, the influence of the formation of homogeneous regions on the PREC flood quantiles is examined. Hence, two pooling methods (cluster analysis, Region of Influence) are applied. For both methods, pooling groups are constructed based on candidate sets of suitable catchment descriptors. By using different candidate sets, several pooling groups are formed for each site. This approach makes it possible to determine the effect of the construction of the pooling groups on the derivation of PREC flood quantiles for the sites under study.

The second paper is tailored to the second step of the PREC concept, i.e. the determination of the effective sample years of data and the recurrence interval, which is assigned to the REC. Its estimation considers the intersite dependence and therefore requires an accurate estimation of the cross-correlation. The role of cross-correlations among flood series within the PREC concept is hence investigated. Therefore, all pairs of catchments are subdivided into nested and unnested pairs of catchments, based on the assumption that nested catchment structures are higher correlated than unnested ones. In this chapter the nested-unnested approach is compared with the traditional one, which uses only one global cross-correlation function by calculating the effective sample years of data and the recurrence interval of PRECs for both approaches. By keeping all aspects constant except for the parameter sets of the cross-correlation function, the influence of this single aspect on the flood quantile estimation were considered in isolation.

Second part:

Flood quantile estimates should benefit from using the PREC flood quantiles as additional information for a flood frequency analysis. Since PRECs provide a discharge with a large recurrence interval, a PREC flood quantile presents additional information for the upper tail of a distribution function. This part of a distribution function in particular is generally not well represented by available flood data.

Since PREC flood quantiles have not yet been used as additional information, it is required to identify a method to integrate them into a flood frequency analysis. In traditional flood frequency analysis, discharge values, i.e. the AMS, without an assigned recurrence interval are used exclusively. In contrast to that, the PREC flood quantile consists of a discharge and its corresponding recurrence interval. Hence, a distribution function which uses the PREC flood quantiles in addition to the systematic time series of observed flood data needs to be selected. In addition to PREC flood quantiles, an upper bound discharge derived by an empirical envelope curve should also be integrated into the selected distribution function.

Thus, these three sub-questions are considered in the second part:

4. **How can PREC flood quantiles be combined with measured flood data using the additional information of the recurrence interval?**

5. **Which distribution function is suitable to integrate the PREC flood quantiles as well as an upper bound discharge?**

6. **How large is the effect of the integration of PREC flood quantiles on the determination of a discharge with a recurrence interval of 1000 years?**

These three points were investigated in chapter four, which corresponds to the third paper:

This paper aims at inserting PREC flood quantiles into a flood frequency analysis. A method allowing this was developed. In this way, the additional information of the recurrence interval of PREC is used. A mixed bounded distribution function which in addition to PREC flood quantiles also considers an upper bound derived by an empirical envelope curve of a wider geographical context is presented.

1.3 Structure of this thesis and author's contributions

This dissertation thesis is written cumulatively. Three papers are presented in the chapter's two to four (see Fig. 1.3). They were published in or submitted to, respectively, international peer-reviewed journals. These three chapters correspond to the three papers. A few exceptions were made. A final reference list for all chapters together is presented at the end of this thesis and the citations are adapted to this reference list. Figure sizes were adapted to the layout of this thesis. To keep the design consistent, some aspects such as the abbreviation of tables were modified.

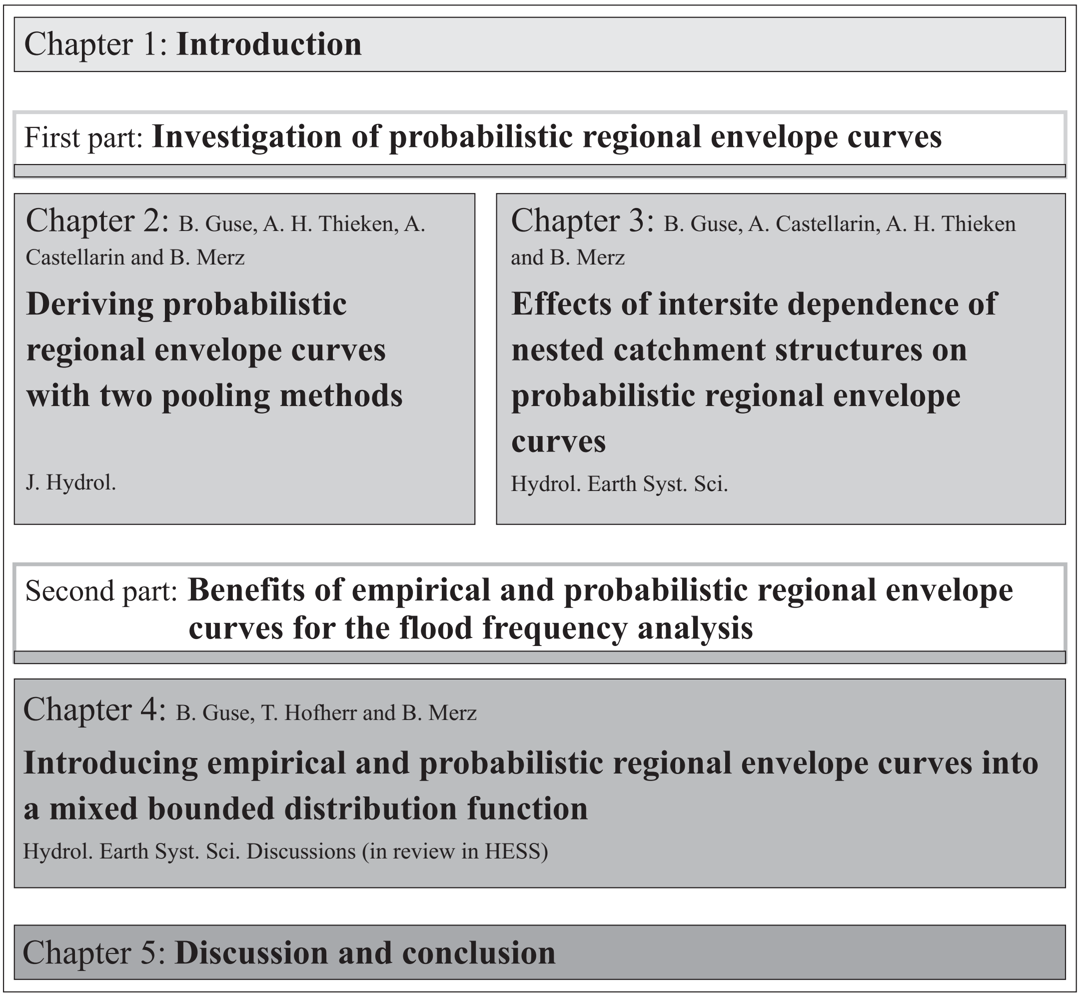

Fig. 1.3: Structure of this thesis.

The mentioned co-authors contributed to the conceptual design of the papers and to the discussion of the results as well as to the final formulations of the manuscript.

The PREC method was developed by Attilio Castellarin. There was an intense exchange of the PREC method and its calculation algorithms during the work on this thesis. Furthermore, Attilio Castellarin significantly contributed to the introduction of chapter two. The adequacy test of the GEV for the selected flood series was done by Attilio Castellarin.

Thomas Hofherr developed the mixed bounded distribution function, which was used in a refined version for the flood research (chapter four).

CHAPTER 2

DERIVING

PROBABILISTIC REGIONAL

ENVELOPE CURVES

WITH TWO POOLING METHODS

2 Deriving probabilistic regional envelope curves with two pooling methods

Abstract

A probabilistic regional envelope curve (PREC) assigns a recurrence interval to a regional envelope curve. A central point of this method is the determination of homogeneous regions according to the index flood hypothesis. A flood discharge associated with the recurrence interval (PREC flood quantile) is estimated for each gauge of a homogeneous region. In this study, the influence of two pooling methods on PREC for a large group of catchments located in the south-east of Germany is investigated. Firstly, using cluster analysis, fixed homogeneous regions are derived. Secondly, the Region of Influence (RoI) approach is combined with PREC. The sensitivity of PREC flood quantiles with respect to pooling groups is evaluated. Different candidate sets of catchment descriptors are used to derive pooling groups for both pooling methods. Each pooling group is checked by a homogeneity test. PRECs are then constructed for all homogeneous regions. The ensemble of PREC realisations reveals the sensitivity of the PREC flood quantiles. A comparison with the traditional index flood method ascertains the suitability of the pooling methods. A leave-one-out jackknifing procedure points out a similar performance of cluster analysis and RoI. Furthermore, a comparison of different degrees of heterogeneity for deriving pooling groups reveals that the performance of PREC for ungauged catchments decreases in more heterogeneous pooling groups.

Published as:

Guse, B., Thieken, A. H., Castellarin, A., and Merz, B., 2010. Deriving probabilistic regional envelope curves with different pooling groups, J. Hydrol., 380(1-2), 14-26.

2.1 Introduction

For flood risk analyses and estimations of design floods it is fundamental to accurately quantify the discharges of rare events, i.e. flood events with recurrence intervals of 100 years or more. The well-established methods of flood frequency analysis (FFA) are hampered by the uncertainty that occurs particularly for estimates of high recurrence intervals due to limited observation data (e.g. Robson and Reed, 1999; Merz and Thieken, 2005). Regional Flood Frequency Analysis (RFFA) is widely employed in the estimation of design floods when dealing with data record lengths that are too short compared to the recurrence interval of interest (e.g. Hosking and Wallis, 1997). Still, most methods of FFA and RFFA do not consider an upper bound of the flood discharges.

Regional envelope curves (RECs) are a traditional, deterministic method for representing the upper bound of the maximum floods observed in a distinct region. A REC bounds the largest floods of each gauge, termed floods of record, of a region. Since their first introduction (Jarvis, 1925), RECs have been applied to different regions and scales. Traditionally, they refer to administrative units (e.g. China and USA (Costa, 1987), Europe and World (Herschy, 2002)). RECs have also been constructed for hydro-meteorological regions with different climatic conditions and, consequently, different flood regimes (e.g. 17 regions in the USA (Crippen and Bue, 1977); north-western and western Greece (Mimikou, 1984)).

A main criticism on RECs relates to their deterministic view and their need to be checked routinely for being exceeded by recent events (e.g. Crippen and Bue, 1977; Castellarin et al., 2005). The applicability of RECs to engineering problems, such as flood design, is limited by the lack of an exceedance probability (or a recurrence interval) that can be assigned to the envelope curves. To overcome this deficiency, Castellarin et al. (2005, 2007), Castellarin (2007), and Vogel et al. (2007) proposed a probabilistic interpretation of RECs which, besides the magnitude, also considers the frequency of a REC.

Probabilistic regional envelope curves (PRECs) are based on the well-known index flood method (Dalrymple, 1960), which is often applied in flood regionalisation studies (e.g. GREHYS, 1996b; Hosking and Wallis, 1997; Robson and Reed, 1999). Only if a region is homogeneous as defined by the index flood hypothesis, a PREC can be constructed and an exceedance probability can be assigned to the curve. A flood discharge associated with the exceedance probability, termed PREC flood quantile, was derived for each site of a homogeneous region.

According to Castellarin et al. (2005), the estimation of the exceedance probability of a PREC further requires the evaluation of the overall sample years of the underlying data which in turn depends on the intersite or cross correlation amongst the annual maximum series (AMS) of flood flows observed at different gauges. It is important to emphasise that the exceedance probability of a PREC always differs from zero which highlights the difference between PRECs and Probable Maximum Floods. A PREC provides one recurrence interval without an extrapolation and, in principle, enables one to estimate the design flood at ungauged sites as a function of the drainage area (e.g. Castellarin, 2007) or of a set of suitable physiographic and climatic catchment descriptors (e.g. Castellarin et al., 2007). PRECs should be seen as complements to RFFA. They can provide additional information on plausible values of extreme floods and the corresponding exceedance probability in gauged and ungauged basins. A leave-one-out jackknifing approach has shown that PREC flood quantiles have a similar reliability as the traditional index flood method (Castellarin, 2007).

Similarly to RFFA, the construction of a PREC requires the identification of hydrologically homogeneous regions or pooling groups (GREHYS, 1996b; Castellarin et al., 2001). Catchments with similar hydrological behaviour can be classified into one group, and the hydrometric information collected at all gauges that belong to the pooling group can be used to improve the accuracy of the design flood estimates for all gauges of the group. The homogeneity of a pooling group can be assessed by statistical tests (e.g. Viglione et al., 2007; Castellarin et al., 2008).

The requirement of homogeneity and the need for sufficient data within a group are often conflictive. On the one hand, a larger number of observations reduces the uncertainty in estimating high recurrence intervals (Robson and Reed, 1999). On the other hand, a larger number of gauges in the pooling group generally results in a higher hydrological heterogeneity of the group. Several studies highlight the relevance of regional homogeneity for RFFA (e.g. Lettenmaier et al., 1987; Stedinger and Lu, 1995) and, more recently, for PRECs (Castellarin, 2007). Therefore, an appropriate classification technique is required for the identification of pooling groups.

Flood regionalisation studies propose two approaches for deriving pooling groups: the delineation of a subdivision of the study area into fixed homogeneous regions and the neighbourhood approach or Region of Influence approach (RoI) (Burn, 1990b; GREHYS, 1996b; Ouarda et al., 2001). In fixed homogeneous regions, each gauging station definitely belongs to one and only one region. A traditional approach to identify fixed homogeneous regions is a separation in administrative units, where all gauging stations are geographically connected, e.g. in adjacent sub-catchments. This method has been replaced by others that enhance the hydrological similarity within a fixed region (Acreman and Sinclair, 1986). Cluster analysis is an objective procedure that can be applied to subdivide the study area into clusters of catchments (fixed regions) on the basis of a suitable set of climatic and physiographic catchment descriptors (predictor variables). The goal of the procedure is to maximise the similarity within a cluster and the dissimilarity between the clusters (e.g. Mosley, 1981). The catchments of one cluster are not necessarily geographically connected.

The RoI approach identifies a pooling group separately for each gauging station (site of interest) without explicit spatial connection within the RoI (Burn, 1990b). Gauging stations for a RoI are selected according to their similarity to the site of interest using a suitable set of predictor variables (Zrinji and Burn, 1994). In a hybrid RoI approach, the RoI is derived by considering the geographical distance between the sites in addition to the predictor variables (Eng et al., 2007).

Up to now, PRECs were applied in northern Italy with a relatively limited number of gauging stations grouped into three different fixed homogeneous regions (Castellarin, 2007). This paper presents the application of the PREC approach in Germany, considering a rather large number of sites. The main aim of the study is to verify, whether the utilisation of the RoI approach in the formation of homogeneous pooling groups may improve the reliability of the design flood estimates that can be retrieved from PRECs for ungauged sites. To address this issue, we construct PRECs for the study area using fixed homogeneous regions and RoIs. In particular, we form several PRECs for each gauging site on the basis of the data collected in homogeneous fixed regions and RoIs with different sizes and catchment descriptors. A sensitivity analysis enables us to consider the sensitivity of PREC flood quantiles to different constitutions of the pooling group. By means of "leave-one-out" cross-validation procedure, we simulate the ungauged conditions at all considered sites during the construction of each PREC as proposed by Castellarin (2007). All flood estimates are compared with the corresponding estimates (i.e. flood quantiles associated with the same values of the recurrence interval) obtained by applying a traditional regionalisation approach. The comparison enables us to better understand and quantify (1) the suitability of the two different pooling methods (i.e. cluster analysis and RoI) in the context of probabilistic regional envelope curves, and (2) the accuracy of flood quantiles retrieved from PRECs for ungauged basins.

2.2 Methods

Since the construction of pooling groups is a prerequisite for the application of PREC, it is advisable to quantify the sensitivity of PREC to the formation of pooling groups. For both pooling methods (cluster analysis and RoI), the sensitivity of PREC results was determined by considering several variations of pooling groups derived in a three-step-procedure.

1. Formation of candidate sets of catchment descriptors.
2. Construction of homogeneous regions using two pooling methods.
3. Test on homogeneity of each pooling group.

Finally a specific PREC was constructed for each homogeneous region. To compare the different results some performance measures were analysed. Each step of the procedure is described in the remainder of this section.

2.2.1 Candidate set of catchment descriptors

Different catchment descriptors were used as predictor variables to derive homogeneous regions. In a first step all catchment descriptors were standardised to a mean value of zero and a standard deviation of one. This standardisation allows a comparison between the predictor variables and avoids the influence of different value scales (see e.g. Nathan and McMahon, 1990).

The catchment descriptors were combined by summing up the standardised values for each site. This approach is only applicable, if all standardised variables have a positive correlation with the unit index flood, i.e. the index flood normalised by the catchment size. In order to get only positive correlations, standardised variables with a negative correlation to the unit index flood were multiplied with -1. This implies, for instance, that the fraction of the area, which is not covered by arable land, was used instead of the fraction of arable land for selecting candidate sets of catchment descriptors.

A full enumeration approach was used to consider all possible subsets of the catchment descriptors with one to three predictor variables. A larger number of catchment descriptors within one candidate set could provide small additional information, but could also lead to multi-collinearity (Merz and Blöschl, 2005). Thus variants with more than three predictor variables were not taken into account.

With regard to the selection of suitable sets of predictor variables, it is worth noting that we were interested in assessing the sensitivity of PRECs and of flood quantiles derived from these PRECs with respect to different pooling groups. To this aim, we looked for several good combinations of predictor variables rather than the optimal set. It was assumed that, next to the best subset of catchment descriptors, other 'good subsets' have a similar explained variance. Since PREC is based on the assumption of a scaling of the index flood (mean of the annual maxima series), it seemed reasonable to perform a preliminary identification of candidate sets of catchment descriptors by looking at the explained variance of the empirical index flood values. Therefore, candidate sets of catchment descriptors were identified on the basis of this criterion.

The correlation coefficient between a subset of catchment descriptors and the unit index flood was used as goodness-of-fit criterion, as in other studies (e.g. Burn, 1990b; Uhlenbrook et al., 2000) under the assumption that a high correlation is a good indicator for a sufficient explained variance of the selected subset (Merz and Blöschl, 2004).

All subsets of catchment descriptors were selected that showed a correlation coefficient of more than 0.60. This threshold was assumed as sufficient, because the correlation coefficient was only used for a pre-selection of subsets of catchment descriptors.

All selected subsets were checked for multicollinearity between the catchment descriptors using the variance inflation factor (VIF) (Hirsch et al., 1992) (Eq. (2.1)).

$$VIF_k = \frac{1}{1 - r_k^{\,2}} \tag{2.1}$$

r_k^2 stands for a multiple correlation coefficient, which was calculated by a regression of variable k using all other variables as predictor variables. To avoid multi-collinearity, all subsets with VIF > 5 were omitted. Montgomery et al. (2001) and Eng et al. (2005) recommended a threshold between 5 and 10.

2.2.2 Formation of homogeneous regions

To assess the influence of homogeneous regions on PREC, two different approaches for the derivation of pooling groups were applied. These methods were fixed homogeneous regions derived by a cluster analysis and the Region of Influence (RoI) method. To ensure an appropriate comparison of both methods, the same candidate sets of catchment descriptors were used.

Fixed homogeneous regions using cluster analysis

Fixed homogeneous regions were derived by cluster analysis with the K-means algorithm, which had already been used in flood frequency analysis (e.g. Burn, 1989; Burn and Goel, 2000) and very recently in a flood seasonality study (Beurton and Thieken, 2009). The cluster analysis was performed allowing three to seven clusters, and was therefore applied five times to each subset of predictor variables. The different number of clusters considers the trade-off between the homogeneity within a cluster and the number of sites within one group.

Region of Influence (RoI)

The approach "Region of Influence" (Burn, 1990b) constructs an individual region (group of gauging sites) for each gauge by finding stations that are similar to the characteristics of the station under study (site of interest). The RoI was determined by the similarity of gauging stations in the physiographical space of the selected catchment descriptors. Similarity was assessed by the Euclidean distance between each site and the site of interest in the physiographical space. The Euclidean distance has been used in several RoI approaches (e.g. Zrinji and Burn, 1994; Castellarin et al., 2001; Gaál et al., 2008), although other similarity measures are possible (see e.g. Cunderlik and Burn, 2006).

All gauging stations which are closer to the site of interest than a specific threshold of the Euclidean distance in the physiographical space were assigned to the RoI of the site of interest. The higher the threshold, the larger is the number of sites within a region (Burn, 1990b). Different similarity measure thresholds to derive RoIs were investigated by Gaál et al. (2008). To account for the sensitivity of the results to the threshold, three thresholds for the similarity measure (0.5, 1 and 2) were applied in this study. In contrast to RoI approaches in frequency analysis (Burn, 1990b), the sites were not weighted according to their closeness to the site of interest in the physiographical space. The original RoI method was varied, because the intercept of PREC is only determined by one pair of unit flood of record and drainage area (see "Probabilistic regional envelope curve"). Consequently, a weighting scheme would not affect the magnitude of the regional envelope curve.

Traditionally, a fixed number of sites is targeted at when deriving a RoI (Burn, 1997). This target number is a function of the aspired return period. In our case a target number of sites cannot be determined, since the recurrence interval *T* associated with the PREC is not known a priori. Therefore, the maximum number of sites in the RoI was identified on the basis of the hydrological affinity with the site of interest.

2.2.3 Homogeneity test

Each pooling group was checked for homogeneity by applying the heterogeneity measure of Hosking and Wallis (1997) (Tab. 2.1). The H_1-test calculates the variability of the L-coefficient of variation (L-CV). The sample L-CV is compared with an expected value for a homogeneous region obtained by a Monte-Carlo simulation. The second and third heterogeneity measures H_2 and H_3 consider the L-CV and the L-skewness as well as the L-skewness and the L-kurtosis, respectively. A more detailed explanation of L-moments and the heterogeneity measure is given by Hosking and Wallis (1997).

Since the homogeneity test for the L-CV (H_1) is a more significant test than the tests with higher moments (H_2 and H_3) (Castellarin et al., 2001, 2007; Hosking and Wallis, 1997), this study focused on the H_1-test using the hw.test (Viglione, 2008, implemented in R). All regions with a H_1 value lower than 2 were used for deriving a PREC.

Tab. 2.1: Interpretation of the heterogeneity measure (Hosking and Wallis, 1993; Robson and Reed, 1999).

Heterogeneity measure	Interpretation	Review
< 1	Homogeneous	Not required
1 – 2	Possibly heterogeneous	Optional
2 – 4	Heterogeneous	Desirable
> 4	Strongly heterogeneous	Essential

2.2.4 Probabilistic regional envelope curve

The method of probabilistic regional envelope curves (PREC) is based on two principles. In the first place, all gauging stations of a region have to be homogeneous in terms of the index flood hypothesis. Secondly, the index flood μ_X (mean of the annual maxima series) is related to the drainage area A (Eq. (2.2), adopted from Castellarin, 2007). Under these assumptions the index flood scales with the drainage area and depends only on the drainage area (Castellarin, 2007):

$$\mu_X = C * A^{b+1} \tag{2.2}$$

To derive a regional envelope curve, all floods of record Q_{FOR} of a region are normalised by their corresponding catchment area A to the unit flood of record q_{FOR} and are related to A in a double-logarithmic scale (Eq. (2.3), adopted from Castellarin et al., 2005). The regional envelope curve bounds all unit floods of record of a region and is defined by its slope b and the intercept a:

$$\log\left(\frac{Q_{FOR}}{A}\right) = a + b * \log(A) \qquad (2.3)$$

The slope b is derived by a regression of the unit index flood against the drainage area (Fig. 2.1). The intercept a is determined by a parallel upshift of the regression until the envelope curve bounds all unit floods of record (Castellarin et al., 2005). In a homogeneous region the index floods of all gauges are close to the regression line. In this study, a PREC was determined for each region with at least four sites. It was assumed that a lower number of sites is not representative for a regression analysis.

An exceedance probability is assigned to that particular data pair of unit flood of record and its drainage area that determines the intercept of the envelope curve. This exceedance probability is valid for the range of catchment sizes covered in the pooling group. For this, the AMS of all gauging stations of that region were considered. The total number of sample years of data was reduced to an effective number of sample years of data, by accounting for cross-correlated sites (Castellarin, 2007). Several studies have shown that the correlation of annual maximum series decreases with the distance of the catchments (see e.g. Hosking and Wallis, 1988; Troutman and Karlinger, 2003). Under these assumptions, a regional cross-correlation function by Tasker and Stedinger (1989) (Eq. (2.4), from Castellarin, 2007) was optimised using the distances between catchment centroids, the correlation coefficients between the AMS and the lengths of overlapping time series.

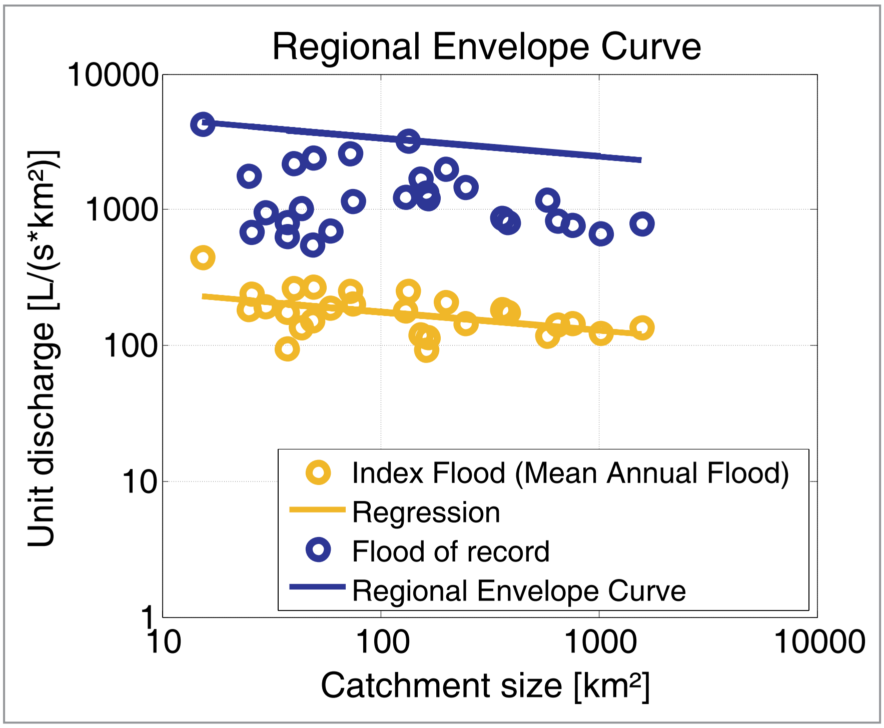

Fig. 2.1: Example of a Regional Envelope Curve.

$$\rho_{i,j} = \exp\left(- \frac{\lambda_1 d_{i,j}}{1 + \lambda_2 d_{i,j}} \right) \quad (2.4)$$

d is the distance between catchment centroids, ρ the correlation coefficient, λ_1, λ_2 the parameters, and i,j are the index denoting pairs of catchments.

In comparison to Castellarin (2007), the method for considering intersite correlations was changed in this paper due to the larger number of catchments available and the presence of numerous nested catchments, i.e. gauging stations along the same river. Troutman and Karlinger (2003) emphasised that the correlation between the AMS of nested catchments was higher than for unnested catchments. Guse et al. (2009) showed that distinct parameter sets for nested and unnested catchments led to a reduction of the recurrence interval of PRECs due to larger correlations between nested catchments. Hence, specific parameters of the cross-correlation function were used for nested and unnested catchments.

Considering the intersite correlation, the overall effective sample years of data n_{eff} were calculated by an empirical relationship, which was determined by Castellarin et al. (2005) and Castellarin (2007) in Monte-Carlo simulations (Eq. (2.5)). This approach is based on the average correlation coefficient ρ (see Eq. (2.4)). Castellarin (2007) proposed an algorithm that can be applied for real world datasets with Y years, in which the record lengths of the gauges varies. In the first step of the algorithm, the number of years n_1 was identified in which only one gauging station had a measured discharge. These observations n_1 were reasonably effective. The remaining years Y-n_1 were divided in $Y_{sub} \leq (Y - n_1)$ subsets with the same gauging stations L_s and the length l_s. Next, for each subset s of l_s years, the effective number of observations $n_{eff,s}$ was calculated separately. Finally, the effective samples for all subsets were summed up. The number of effective sample years of data for the whole regional data set n_{eff} includes n_1, the years with one observations, and the sum of $n_{eff,s}$ (Eq. (2.5), adopted from Castellarin, 2007).

$$n_{eff} = n_1 + \sum_{s=1}^{Y_{sub}} n_{eff,s} = n_1 + \sum_{s=1}^{Y_{sub}} \frac{L_s l_s}{1 + \left[\overline{\rho^{\beta}}\right]_{L_s} (L_s - 1)}$$

$$\text{with } \beta := 1.4 \frac{(L_s l_s)^{0.176}}{\left[(1-\rho)^{0.376}\right]_{L_s}} \quad (2.5)$$

In this way the effective sample years of data is equivalent to the number of independent observations. This reduction of the regional plotting position determines the information content of the collected data (Castellarin, 2007).

The next step is a selection of an appropriate plotting position depending on an adequate distribution function to estimate the recurrence interval of the PREC. Castellarin (2007) recommended the use of the Hazen plotting position (Eq. (2.6), from Castellarin (2007)) in order to get unbiased flood quantiles, when the Generalised Extreme Value (GEV) distribution is a suitable parent distribution. Its suitability for the case study is reported in “Study area and data”. As a result, the recurrence interval T_{PREC} is twice as high as the number of effective observations n_{eff}.

$$T_{PREC} = 2 * n_{eff} \quad (2.6)$$

The exceedance probability is greatly influenced by the formation of homogeneous regions. Adding or removing only one gauging station to/from a homogeneous group modifies the effective sample years of data and hence the exceedance probability of the PREC.

The discharge associated with the exceedance probability for a specific site is determined by the intercept of the drainage area and the regional envelope curve. It is worth noting that the gauging stations within a region have a different influence on the exceedance probability of the PREC. Due to the fact that the intercept of the PREC is determined by the data pair of the highest unit flood of record and its drainage area, this gauging station is the most decisive. This aspect highlights the particular importance of a consistent assignment of gauging stations to pooling groups.

A discharge Q_{PREC} and a recurrence interval T_{PREC} were derived for all gauging stations of a region. T_{PREC} is constant for all gauging stations in the region. Since the PREC was only calculated for homogeneous regions, the number of PREC realisations is different for the gauging stations. It depends on the number of homogeneous regions in which the specific gauging station is included.

2.2.5 Sensitivity analysis

The effect of pooling groups on PREC flood quantiles (Q_{PREC}, T_{PREC}) was examined by a sensitivity analysis. Pooling groups of both pooling methods were derived for all candidate sets of catchment descriptors with a correlation coefficient to the unit index flood >0.60. For each candidate set of catchment descriptors, cluster analysis was applied five times (allowing three to seven clusters) and the Region of Influence approach three times (with different thresholds in the physiographical space) (see "Formation of homogeneous regions"). These predefined number of clusters and thresholds in the physiographical space led to several candidate solutions of pooling groups. Ultimately all pooling groups with a heterogeneity measure $H_1 < 2$ were used to derive a PREC. Each PREC realisation led to a pair of Q_{PREC} and recurrence interval T_{PREC} (PREC flood quantile) for each gauge of the pooling group.

The rationale behind this scheme is that different constitutions of the regions lead to different realisations of PREC. The application of several candidate sets of catchment descriptors allows a quantification of the sensitivity of the PREC results in terms of the pooling method and the selected subset of catchment descriptors. However, it is worth noting that the uncertainty of the ensemble of PRECs results is not estimated by this procedure.

2.2.6 Performance criteria

The performance of PREC flood quantiles was evaluated by comparing them with a traditional index flood approach.

The index flood method is based on the assumption that a regional growth curve is valid for all sites of a pooling group. For this, the AMS was normalised by the index flood μ_X. To calculate the T-year flood X(T), a regional quantile x_T was scaled to at-site conditions by the index flood μ_X (Eq. (2.7)).

$$X(T) = \mu_X * x_T \tag{2.7}$$

The GEV was also used for the index flood approach. The parameters were estimated with regional L-moments, by weighting at-site L-moments of all gauges according to the data length (Robson and Reed, 1999).

In order to assess the accuracy of PREC for ungauged catchments, a cross-validation procedure was applied. The PREC was recalculated following a leave-one-out jackknifing algorithm (Castellarin, 2007; Castellarin et al., 2007), termed PREC-JK: (1) A pooling group with M sites, which fulfilled the homogeneity criteria, was selected. (2) A site m was excluded from this pooling group. (3) For the remaining M-1 stations the PREC-JK was calculated and the recurrence interval of PREC-JK ($T_{PREC\text{-}JK}$) was determined. (4) The discharge of PREC-JK $Q_{PREC\text{-}JK}$ was evaluated for the given drainage area of the site m. Since site m was not included in the calculation, the PREC-JK result was considered as ungauged. (5) The index flood method was applied for the same pooling group. In this case the site m was included. The flood quantile for the given recurrence interval $T_{PREC\text{-}JK}$ was calculated by the index flood method (Q_{IF} ($T_{PREC\text{-}JK}$)). In this context the index flood method was assumed as the 'true' result. To get a perfect estimator for ungauged conditions, $Q_{PREC\text{-}JK}(T_{PREC\text{-}JK})$ was compared with Q_{IF} ($T_{PREC\text{-}JK}$) (Eq. (2.8), adopted from Castellarin, 2007).

$$\varepsilon_{PREC-JK} = \frac{Q_{PREC-JK}(m, T_{PREC-JK}) - Q_{IF}(m, T_{PREC-JK})}{Q_{IF}(m, T_{PREC-JK})} \tag{2.8}$$

The cross-validation was performed for all homogeneous regions. It was repeated M-times for all sites within a cluster. In the case of a RoI, the jackknifing approach was only applied once for the site of interest. The relative error of PREC-JK in comparison to the index flood method enables us to compare the two pooling methods.

2.3 Study area and data

The study area is the federal state of Saxony in the south-east of Germany (Fig. 2.2). Saxony is characterised by the mountain range of the *Erzgebirge* in the south-west with elevation up to 1214 m above sea level *(Fichtelberg)* and a mean annual precipitation up to 1244 mm (at the synoptic station *Carlsfeld*). The highest monthly precipitation occurs in summer (Flemming, 2001). The river Elbe with a drainage area of about 52,000 km² at the gauge Dresden is the biggest river in Saxony. Several feeder rivers originating in the *Erzgebirge* flow into the Elbe, the most important one is the river Mulde (Fig. 2.2). The mountain range east of the Elbe has a lower elevation than the *Erzgebirge*. Towards the north the elevation flattens. The north-western and north-eastern parts of Saxony are influenced by mining activities.

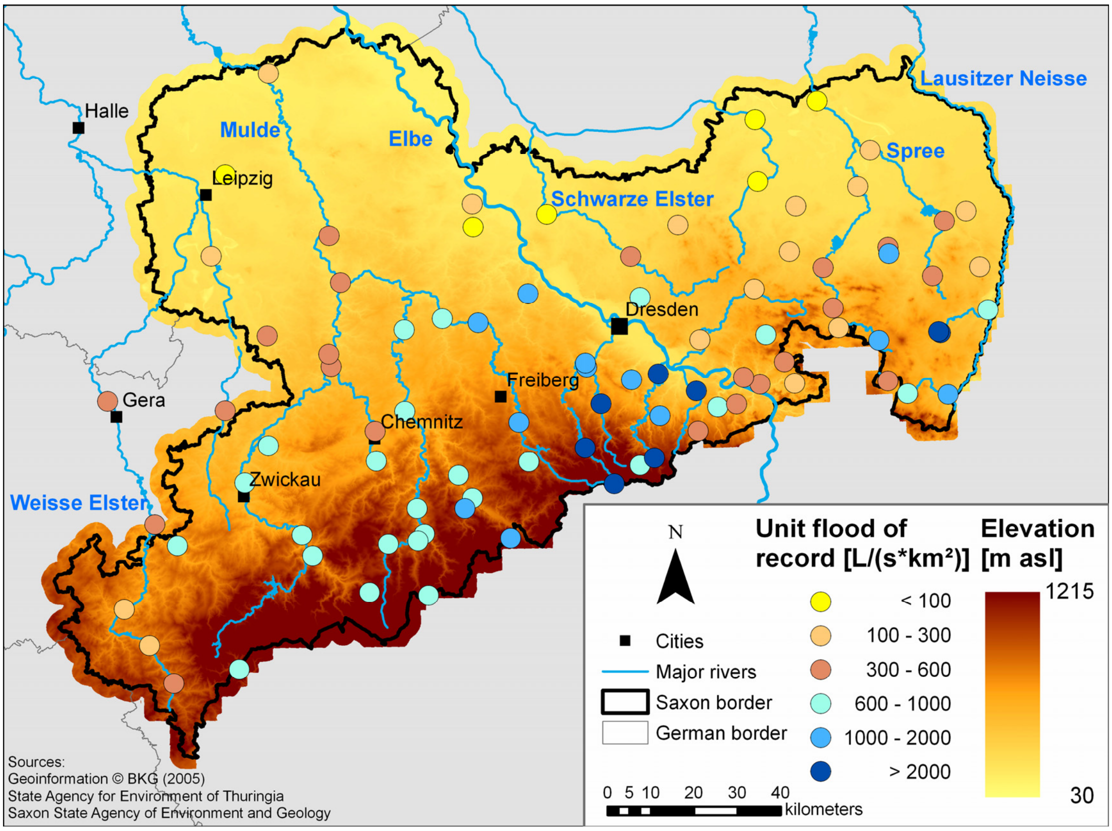

Fig. 2.2: Study area: Elevation above sea level in the federal state of Saxony, Germany, and available discharge gauging stations coloured by the unit flood of record.

In Saxony, several severe floods occurred in the past. Ulbrich et al. (2003) distinguished between flash floods along the tributaries of the rivers Elbe and Mulde and slowly rising river floods along the Elbe. The *Erzgebirge* was affected by local (e.g. in 1927, 1957) and regional floods (e.g. in 1954, 1958, 2002) (Pohl, 2004; Thieken et al., 2007). Among the regional floods, especially the recent destructive flood of 2002 along the rivers Elbe and Mulde and their tributaries from the *Erzgebirge* is still present in people's minds. During this event a record-breaking daily precipitation of 312 mm/day was measured at the synoptic station *Zinnwald-Georgenfeld*, which is located in the upper stream of the Müglitz (Ulbrich et al., 2003). For the 2002 flood, IKSE (2004) estimated recurrence intervals up to 200 - 500 years at some tributaries of the Elbe river, e.g. at the rivers Mulde, Müglitz and Weisseritz.

One hundred and seventeen discharge gauging stations from all over Saxony with the maximum discharges for each month were provided by Saxon authorities. For the catchment of the *Weisse Elster,* which is only partly located in Saxony, additional data was provided by authorities of Thuringia and Saxony-Anhalt. The gauging stations are evenly distributed throughout the area of this study (Fig. 2.2). All major rivers are included in the data set. Observation periods range from 20 to 150 years with a mean length of 50 years. This data set includes extreme floods with local as well as regional spatial extent. The highest unit discharges were observed in the western tributaries of Elbe (i.e. at the rivers Gottleuba and Müglitz) and in the river Pliessnitz, a tributary of the Lausitzer Neisse near the German-Polish border (Fig. 2.2). Due to a few very extreme floods, the series of annual maximum floods show a high skewness, especially in the *Erzgebirge* (Petrow et al., 2007).

Since the index flood hypothesis requires a strong homogeneity within a region, only gauging stations were used that represented the regional hydrological situation. Thus, the available data set was reduced, i.e. gauges of heavily influenced rivers due to mining activities (four sites), gauging stations directly downstream of a dam (two sites) and very small catchments (<10 km^2) (four sites) were discarded. Furthermore, only gauging stations with at least 30 years of data were used. Due to these restrictions the number of gauging stations was reduced to 95.

The construction of pooling groups (see “Formation of homogeneous regions”) requires the derivation of different catchment descriptors. These predictor variables were pre-selected based on a literature review (e.g. Wiltshire, 1986; Pitlick, 1994; GREHYS, 1996b; Castellarin et al., 2004; Merz and Blöschl, 2005). Those catchment descriptors were applied, which have yielded good results in flood regionalisation studies (Tab. 2.2). All catchment descriptors are catchment averages.

Tab. 2.2: List of catchment descriptors.

Abbreviation	Catchment descriptors
MAP	Mean annual precipitation (mm)
MAXDAY	Maximum daily precipitation (mm)
P50	Annual frequency of days with precipitation of more than 50 mm/d (%)
MAX5DAY	Maximum precipitation in 5 days (mm)
PAMS	Mean of the annual maximum series of daily precipitation (mm)
ELEV	Mean elevation of the catchment (m asl)
SLOPE	Mean slope of the catchment (%)
RANGE_NORM	Range of catchment elevation, normalised with the catchment size ($10^{-3}m^{-1}$)
ARABLE	Fraction of arable land coverage (%)
URBAN	Fraction of urban land coverage (%)
MINING	Fraction of mining activities (%)
BEDROCK	Fraction of bedrock areas (%)
KF_LOW	Fraction of low permeability areas (%)

Precipitation data with a daily resolution in and around Saxony was provided by the German Weather Service (DWD). Precipitation indices were derived on the basis of 453 stations with at least 30 years of data in order to ensure a sufficient sample size. The second constraint was that the time series endured at least up to 2002. This year was selected because of the severe flood event in August 2002. In order to optimise the spatial distribution of precipitation stations, 23 stations with an observation period of less than 30 years were additionally used to derive the maximum daily precipitation and the 5-day-precipitation sum. These stations were added, because the year of the maximum daily precipitation coincided with the flood of record of the downstream gauging station. In these cases, it was assumed that the maximum precipitation was representative for this catchment. The precipitation values were interpolated for the different precipitation indices using ordinary kriging. In the next step the catchment boundaries were superimposed on the precipitation map and the mean value was derived for each catchment.

The mean elevation of the catchments was derived from a digital elevation model for Saxony with a grid size of 25 m. Outside Saxony the SRTM-DEM (Jarvis et al., 2008) with a grid size of 90 m was resampled to 25 ms. A mean slope was derived from this combined DEM. The DEM also provided the catchment centroids, from which the distances between the catchments were calculated, which were then used to optimise the theoretical cross-correlation function (see Eq. (2.4)). The digital landscape model ATKIS (BKG GeoDataCentre, 2005) was used to derive landscape parameters such as the fraction of urban area. The hydrogeological map HÜK200 (1:200,000) of the Saxon State Agency of Environment and Geology provided the fraction of bedrock and low permeability area. The hydrogeological map HÜK200 distinguished between bedrock and unconsolidated rock. Permeability was classified in eleven classes. Low permeability was assessed for all rocks with permeability $<10^{-7}$ (AG Boden, 1994).

Soil parameters were not used in this study, since for example Merz (2006) has emphasised the low performance of soil parameters in multiple regressions without a hydrological soil classification such as the Hydrology of Soil Types (HOST) classification in the United Kingdom (Boorman et al., 1995). The drainage area itself was not used as variable, because it is already included in the concept of regional envelope curves.

Among the available data for the catchment descriptors only the DEM covered the catchments outside of Saxony. Therefore, catchments with insufficient information for the other catchment descriptors were omitted. This led to a further reduction of the data set. In total, all thirteen catchment descriptors listed in Tab. 2.2 were determined for 89 gauging stations shown in Fig. 2.2. Their catchment size varies between 13 (Rennersdorf 2/ Pliessnitz river) and 6170 km² (Bad Düben/ Mulde river).

For each of the 89 gauges the flood of record Q_{FOR} was determined. In a further step, the annual maximum series (AMS), which contain the highest discharge for each hydrological year (1st November to 31th October), was calculated. Independence between flood events in the AMS was ensured by a time gap of at least 7 days between consecutive annual maxima (GREHYS, 1996b). L-moment ratio diagram (see e.g. Vogel and Fennessey, 1993; Peel et al., 2001) clearly indicates that the GEV is a suitable parent distribution function for the whole study area.

2.4 Results

2.4.1 Suitable candidate sets of catchment descriptors

Considering the 13 catchment descriptors listed in Tab. 2.2, 13 subsets with one, 78 with two and 286 with three catchment descriptors resulted. Among the 377 possible subsets of one, two and three catchment descriptors, 39 subsets have a correlation coefficient to the unit index flood higher than 0.6. All subsets with three catchment descriptors were checked for redundancy compared with the subsets with two catchment descriptors. The rationale behind this approach was that an additional parameter ought to lead to a higher proportion of explained variance. Consequently, subsets with three catchment descriptors were only used (a) if they did not include two catchment descriptors, which formed one of the selected subsets with two catchment descriptors, or (b) if the correlation coefficient was higher than this subset with two catchment descriptors. This procedure reduced the number of subsets from 39 to 20. The test of multi-collinearity by the VIF-test resulted in no further reduction.

Table 2.3 illustrates that the correlation coefficient to the unit index flood of the 20 subsets differed between 0.60 and 0.70. All 20 subsets were considered as candidate set and were used to form homogeneous regions and to derive a PREC. The selected subsets contain two or three catchments descriptors. Among the catchment descriptors precipitation and topographic indices have a higher explanatory power than land use and geologic parameters. The maximum of the 5-day-precipitation sum (MAX5DAY), the range of elevation within the catchment (RANGE_NORM) and the fraction of urban land coverage (URBAN) were most often included.

Tab. 2.3: Selected subsets of catchment descriptors (CD) and the correlation coefficient (COR) to the unit index flood of all gauging stations.

CD1	CD2	CD3	COR
MAX5DAY	ELEV	RANGE_NORM	0.70
MAX5DAY	RANGE_NORM	URBAN	0.69
MAP	MAX5DAY	RANGE_NORM	0.69
MAX5DAY	RANGE_NORM		0.68
MAX5DAY	ELEV	URBAN	0.68
ELEV	RANGE_NORM	URBAN	0.66
PAMS	RANGE_NORM	URBAN	0.64
MAX5DAY	ELEV		0.64
ELEV	RANGE_NORM		0.64
MAP	MAX5DAY	URBAN	0.64
MAP	MAX5DAY		0.62
MAP	RANGE_NORM		0.62
PAMS	RANGE_NORM		0.62
P50	RANGE_NORM	URBAN	0.61
MAX5DAY	ARABLE	URBAN	0.61
MAXDAY	RANGE_NORM	URBAN	0.61
MAX5DAY	URBAN	BEDROCK	0.61
MAX5DAY	PAMS	URBAN	0.61
RANGE_NORM	URBAN	BEDROCK	0.60
RANGE_NORM	BEDROCK		0.60

2.4.2 Results for the best subset of catchment descriptors

The best subset of predictor variables contains MAX5DAY, the mean elevation (ELEV) and RANGE_NORM with a correlation coefficient of 0.70 (Tab. 2.3). The pooling groups derived by cluster analysis are illustrated in Tab. 2.4, using the solution with seven clusters as an example. The heterogeneity measure of the cluster analysis shows that there are four ($H_1 < 2$) homogeneous regions (clusters 1, 2, 4, and 6) (Tab. 2.4). Clusters 3 and 7 are strongly heterogeneous. The H_1-test was not applied for cluster 5, because there are only two sites in this cluster. For these three regions the assumptions of PREC are not fulfilled. Thus a PREC was only calculated for the clusters 1, 2, 4 and 6.

The RoI approach provides one region for each of the 89 gauging stations. As outlined in "Formation of homogeneous regions", three different thresholds of the similarity measure were applied. The total number of PREC realisations is lower than 89, because in several cases the number of sites in the RoI is lower than four (Tab. 2.5). Only for 50 sites, there are at least four sites in the physiographical space with a Euclidean distance lower than 0.5. It becomes apparent that, also for the RoI approach, the method of PREC is not applicable for all gauging stations.

In summary, with both pooling methods heterogeneous regions were constructed, for which it was impossible to calculate a PREC. As mentioned before, this deficiency could partly be compensated by the use of different subsets of catchment descriptors.

Tab. 2.4: Results of heterogeneity measure and of PREC method for the best subset of catchment descriptors, derived by cluster analysis for the seven-cluster solution.

Cluster	1	2	3	4	5	6	7
Number of gauges	7	24	10	18	2	8	20
H_1	0.6	0.8	7.8	1.5		-1.4	9.3
Number of observations	277	1471		1326		498	
Effective number of observations	160	483		403		202	
Recurrence interval [a]	320	966		805		403	

Tab. 2.5: Number of sites below and above the threshold (H_1=2) of the heterogeneity measure for the best subset of catchment descriptors constructed by the Region of Influence approach using different thresholds of the Euclidean distance.

Threshold	$H_1 < 2$	$H_1 > 2$	Sum
0.5	28	22	50
1	27	49	76
2	14	74	88

2.4.3 Analysis of homogeneous regions for different candidate sets of catchment descriptors

Since 20 subsets of catchment descriptors were selected and the cluster analysis was performed five times (number of clusters from 3 to 7), altogether 500 regions were constructed and checked for homogeneity by the Hosking-Wallis test. The fraction of homogeneous regions ($H_1 < 2$) is in the range between 43%

(3 cluster) and 54% (7 cluster) for the different numbers of clusters (Tab. 2.6).

With the RoI approach, one region was formed for each gauging station and each subset of catchment descriptors. The fraction of homogeneous regions is strongly influenced by the threshold of the Euclidean distance in the physiographical space. The number of homogeneous regions decreases from 54% for a threshold of 0.5 to 12% for a threshold of 2. As expected, both methods reveal that the fraction of homogeneous regions increases with a decreasing number of gauging stations (higher number of clusters, lower RoI-threshold).

The distribution of the relative number of homogeneous regions shows a spatial pattern for both pooling methods (Fig. 2.3). The gauging stations in the *Erzgebirge* are mostly grouped in homogeneous regions. In contrast, there are no or only a low number of homogeneous regions for several gauges in the Weisse Elster subbasin and east of the Elbe. The relative number of homogeneous regions is larger for the cluster analysis than for the RoI approach. This can be explained by the low number of homogeneous regions that were constructed for a threshold of two in the RoI approach (Tab. 2.6).

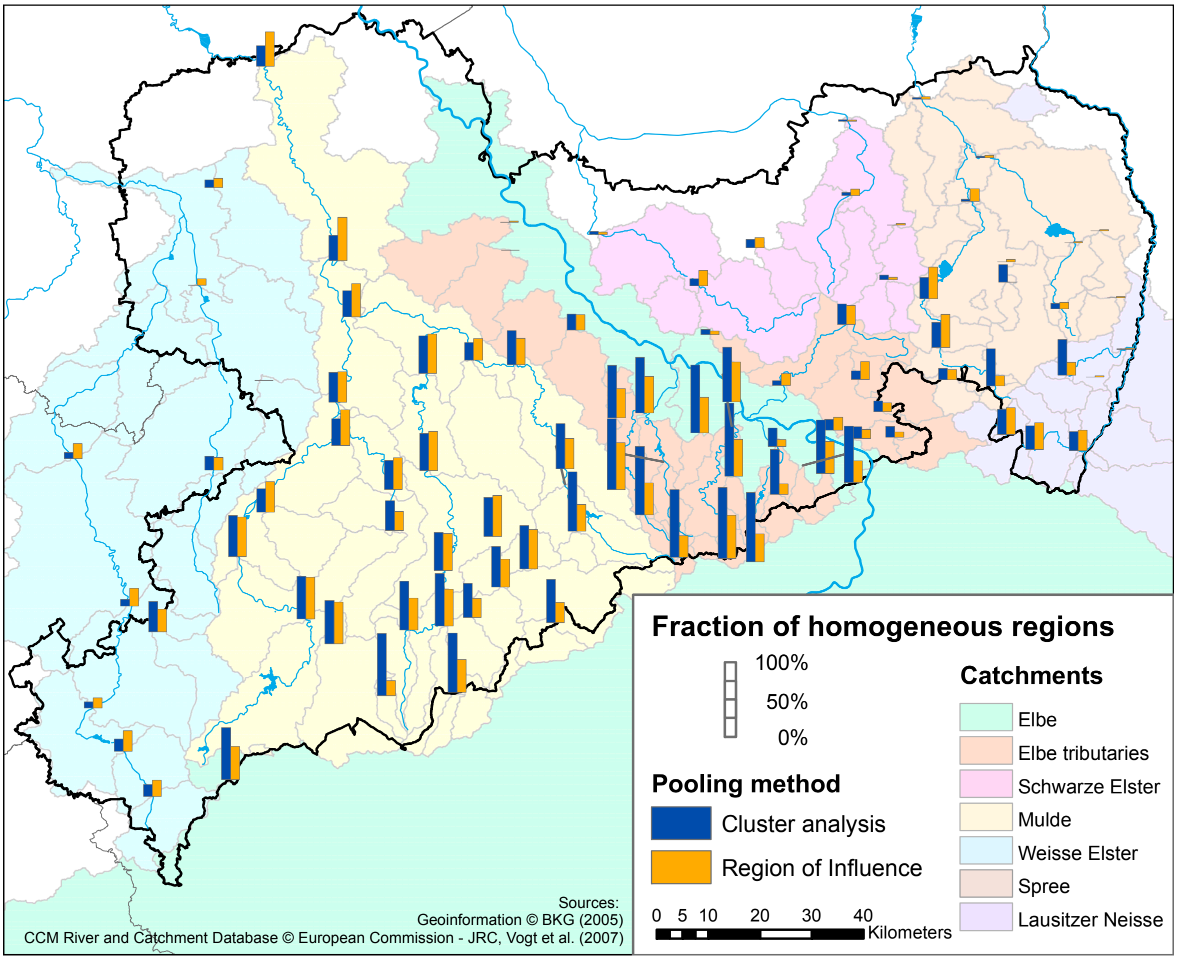

Fig. 2.3: Fraction of homogeneous regions (H_1<2) [%] by cluster analysis and Region of Influence for the gauging stations in the study area.

Tab. 2.6: Number of homogeneous regions derived by cluster analysis and Region of Influence (RoI).

Number of clusters	$H_1 < 2$	$H_1 > 2$	$H_1 < 2$ [%]
3	26	34	43.3
4	35	44	44.3
5	48	47	50.5
6	58	55	51.3
7	67	58	53.6
RoI-threshold			
0.5	575	493	53.8
1	628	1002	38.5
2	212	1539	12.1

The H_1-test was not applied for pooling groups with less than four sites.

2.4.4 PREC results for candidate sets of catchment descriptors

Due to the fact that one PREC is provided for each homogeneous region, it is not possible to show all PREC realisations for all sites. All PREC realisations for the gauging station Dohna/Müglitz are shown as an example in Fig. 2.4. In addition, the pairs of the unit flood of record and the drainage area, which determine the intercept of PREC, are highlighted by black circles. The site itself is indicated separately. Both figures illustrate the influence of different subsets of catchment descriptors and pooling methods on the results of PREC.

Besides the slope and the intercept, also the range of the catchment size that is covered by the PREC depends on the constitution of the pooling group. As expected, the slope decreases with catchment size with two exceptions for RoI. In the example shown in Fig. 2.4 four sites govern the intercept of PREC including the selected site itself for both pooling methods.

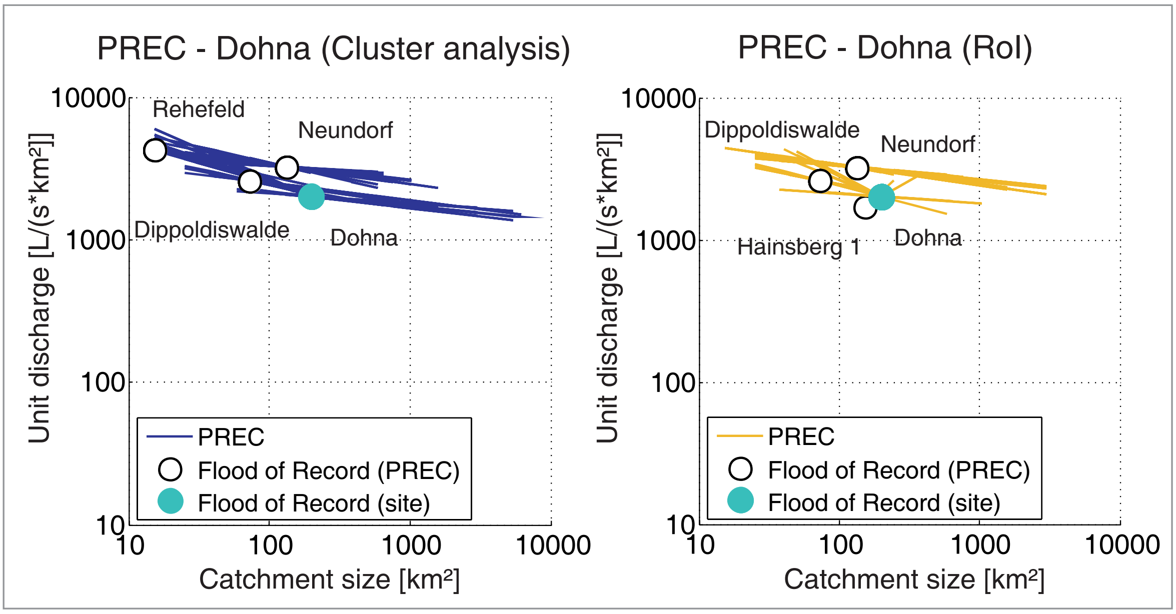

Fig. 2.4: All PREC realisations for the gauge Dohna in homogeneous regions derived by cluster analysis (left) and RoI (right).

As illustrated in Fig. 2.5, the results of PREC for the gauge Dohna differ in discharge (400-630 m³/s) and recurrence interval (300-1200 years) for the two pooling schemes, as well as for different subsets of catchment descriptors. As expected, the discharge augments with increasing recurrence interval. The site itself has only a minor influence on the recurrence interval, because all AMS of the region are collected together (overall sample years of data).

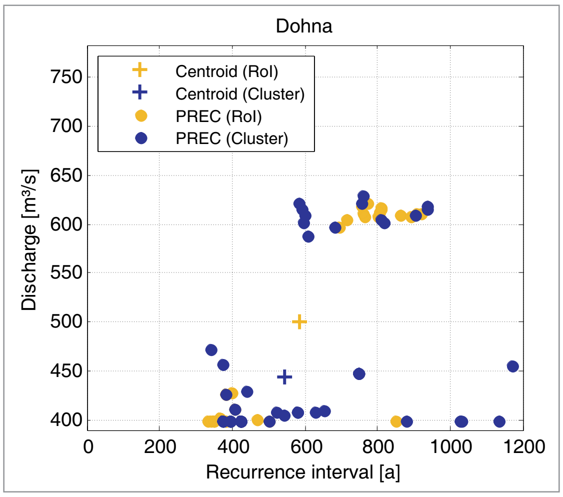

Fig. 2.5: Pairs of discharges and recurrence intervals for all PREC realisations of the gauge Dohna.

Both pooling methods show the influence of the pair of the unit flood of record and drainage area, which determines the intercept of PREC. All discharges are at least 400 m³/s, which is the flood of record at the gauge Dohna. In this example the PREC results of both methods are scattered in three groups. In the first group, the gauge Dohna itself determines the intercept of PREC. The gauge Dippoldiswalde and Rehefeld or Hainsberg 1 in the case of the cluster analysis or RoI, respectively, have the highest unit flood of record for the PREC realisations of the second group, where the discharge varies between 400 and 480 m³/s (Figs. 2.4 and 2.5).

In the third group, the discharge of PRECs for the gauge Dohna is between 580 and 630 m³/s. The intercept of these PRECs is determined by the gauge Neundorf and in two cases for the cluster analysis also by Rehefeld. The range is caused by the different slopes of the PRECs, which were derived for pooling groups with different combinations of gauges. The higher the difference in the catchment size (e.g. Rehefeld (15 km²) and Dohna (198 km²), (see Fig. 2.4)), the larger is the PREC discharge affected by a variation of the slope.

The three groups of PREC realisations show that the inclusion of a gauge with a high unit flood of record (here: Neundorf) results in an upshift of the PREC. The extent of the upshift depends on the difference between the unit flood of record of the site of interest and the highest unit flood of record in the homogeneous group. It is important to highlight that Dohna and Neundorf have a relatively high unit flood of record. For a gauging station with a lower unit flood of record, the difference between the unit flood of record and the regional envelope curve discharge might be significantly higher, if the PREC is also determined by Neundorf.

2.4.5 Performance evaluation of PREC

The reliability of the PREC was evaluated by a leave-one-out jackknifing procedure (PREC-JK). The relative error of the PREC-JK to the index flood method was calculated for each gauging station (see Eq. (2.8)). In Fig. 2.6, only those gauging stations were considered, which had at least eight PREC-JK realisations. This criterion was fulfilled for 68 (Cluster analysis) and 61 sites (RoI), with on average 44 and 21 PREC-JK realisations, respectively.

The PREC-JK approach for both pooling methods illustrates that the median of the relative error is in most cases positive (Fig. 2.6). A high positive relative error indicates a high over-estimation of the discharge of PREC-JK for this recurrence interval in comparison to the index flood method. A negative relative error occurs for the gauging stations which determine the intercept of REC or which are close to the REC (see Fig. 2.7). Comparing the pooling methods, the relative errors (median of the box) as well as the scatter (size of the box) are similar for cluster analysis and RoI (Fig. 2.6).

The relative error between PREC-JK and the index flood method depends on the position of the gauging station in the 'unit discharge-area plot' (Fig. 2.7). If the unit flood of record q_{FOR} of a gauging station is close to the regional envelope curve, the unit discharge $q_{PREC\text{-}JK}$ derived from the regional envelope curve for this station is similar to or lower than that of the index flood method. In contrast, the higher the difference between the regional envelope curve q_{PREC} and the flood of record discharge q_{FOR} for a gauging station, the higher the relative error of PREC-JK in comparison to the index flood method. This relationship has a correlation coefficient of 0.73 (see Fig. 2.7).

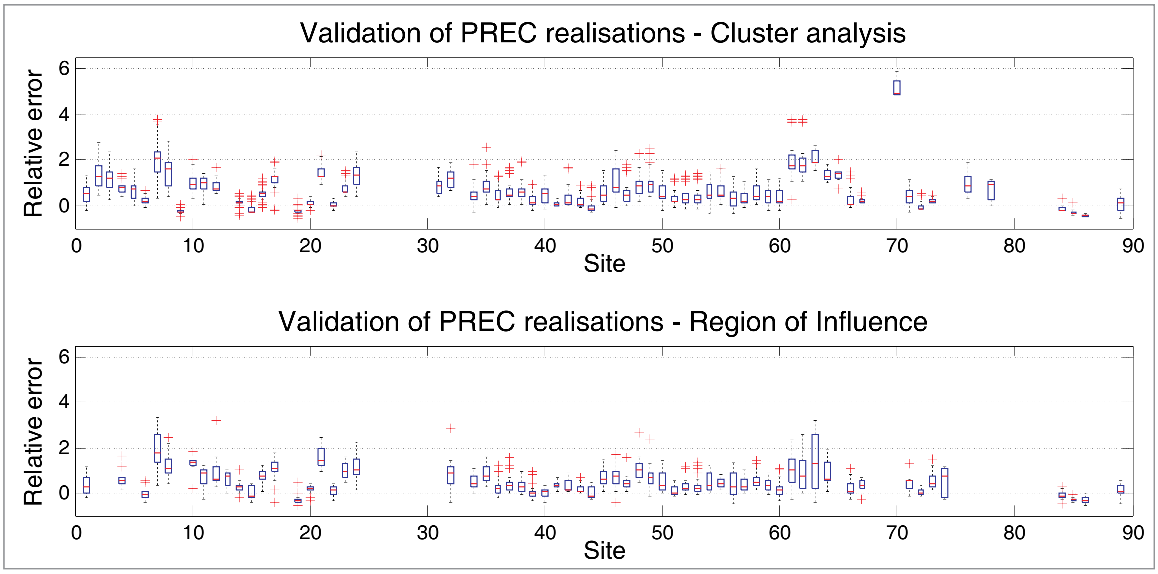

Fig. 2.6: Relative error of the PREC realisations for the two pooling methods cluster analysis (top) and Region of Influence (bottom) for the 89 sites of the study area. The boxplot edges are formed by the 25th and 75th percentiles. Outliers are illustrated with red crosses.

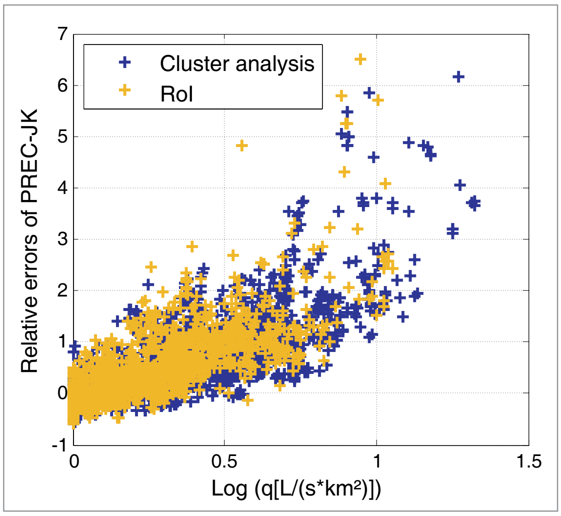

Fig. 2.7: Relative error of PREC-JK versus the distance of the unit flood of record q_{FOR} to q_{PREC} for pooling groups identified by cluster analysis and the Region of Influence approach.

2.4.6 Assessing the effect of the threshold of the heterogeneity measure

The homogeneity of a pooling group is a fundamental assumption of PREC. The influence of the degree of homogeneity on PREC was determined by varying the threshold of the heterogeneity measure. In order to consider the influence of the threshold on PREC, the sensitivity analysis was repeated for stronger ($H_1 < 1$) and weaker ($H_1 < 4$) thresholds of the Hosking-Wallis test. Following the classification of Hosking and Wallis (1997), a threshold of $H_1 < 1$ means that 'possibly homogeneous regions' ($1 < H_1 < 2$) are excluded (Tab. 2.1). By increasing the threshold to four, also 'slightly heterogeneous regions' ($2 < H_1 < 4$) are included. In this case only 'strong heterogeneous regions' ($H_1 > 4$) are excluded. The influence of the relative number of homogeneous regions for different thresholds of the Hosking-Wallis test has been discussed by Cunderlik and Burn (2002). An increase of H_1 from 2 to 4 results in a larger number of homogeneous regions (Fig. 2.8). This is especially relevant for those gauging stations, which were only seldom grouped in a homogeneous region when applying the strict definitions of homogeneity.

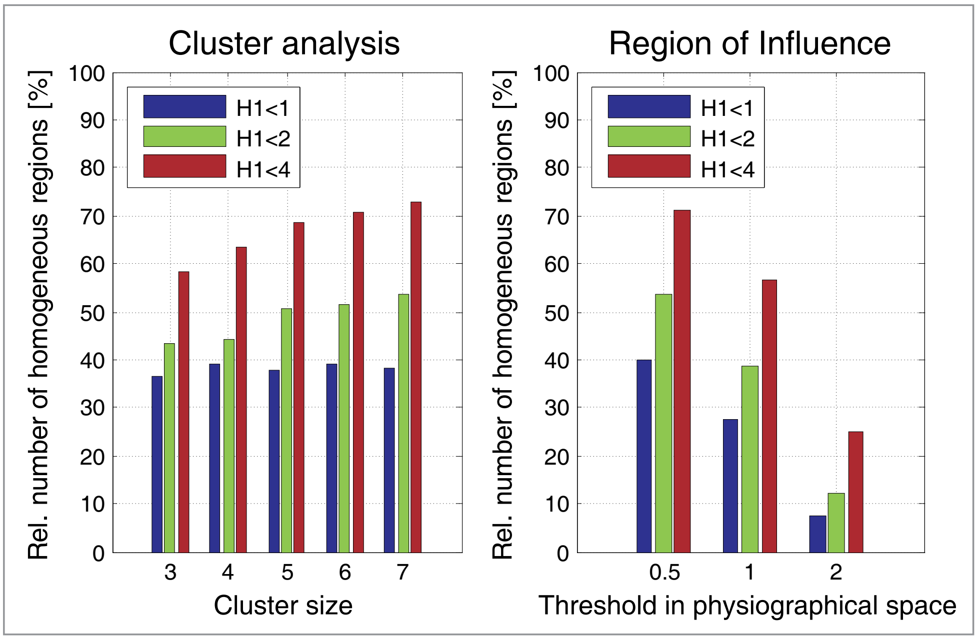

Fig. 2.8: Relative number of homogeneous regions for different thresholds of heterogeneity for cluster analysis and Region of Influence.

A comparison of the mean absolute relative error for the three thresholds illustrates that an increase in the degree of heterogeneity leads to a higher mean absolute relative error for most of the gauging stations and for both pooling methods (Fig. 2.9, Tab. 2.7). In addition, there are more results of the mean absolute relative error for $H_1 < 4$ because of the higher number of PREC realisations.

Considering that the relative error was calculated with the index flood method as reference, it is necessary to mention that the index flood estimate is subject to a higher uncertainty due to the higher degree of heterogeneity.

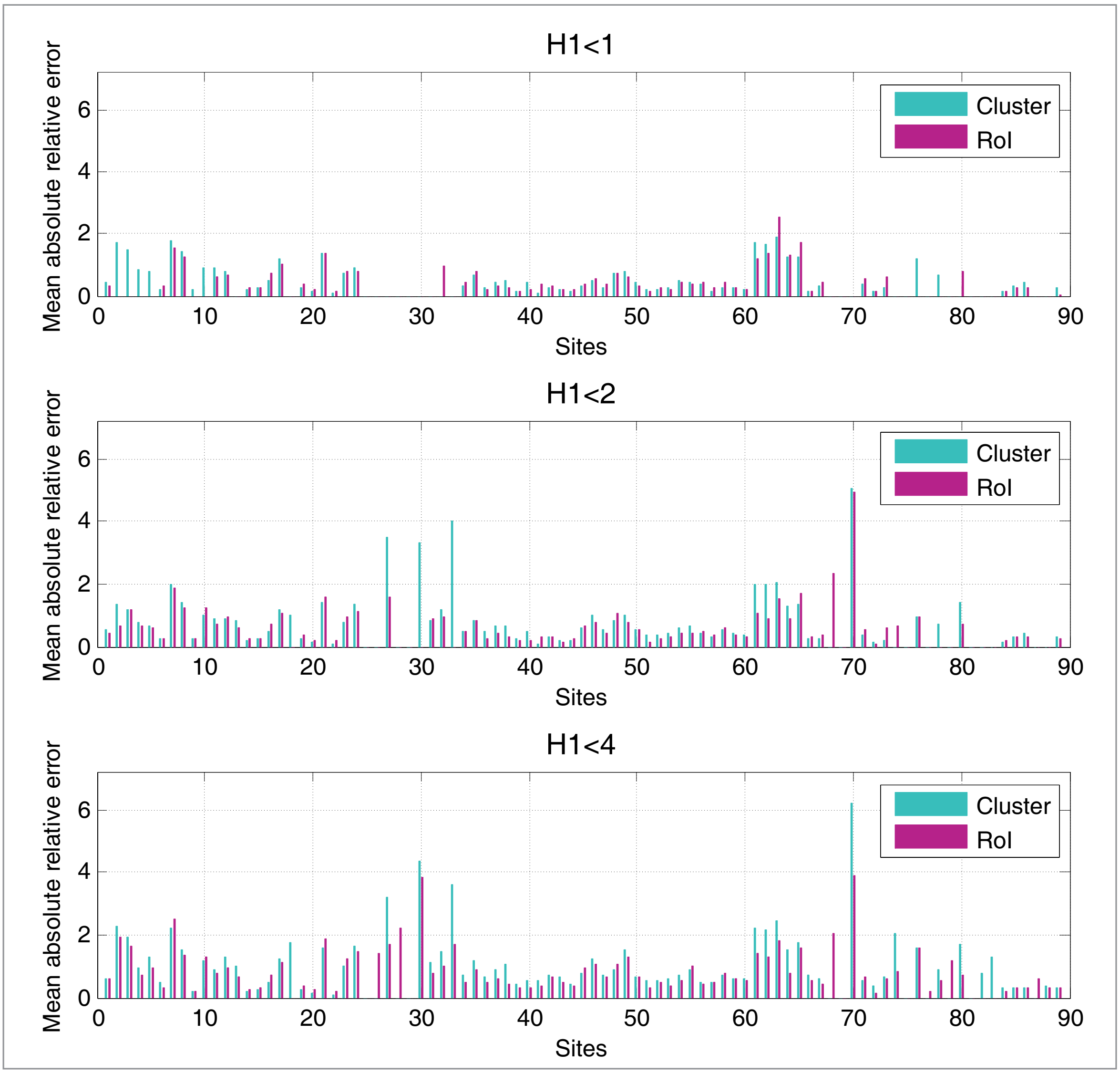

Fig. 2.9: Mean absolute relative error of PREC-JK for both pooling methods (cluster analysis, Region of Influence (RoI)) using different thresholds of the heterogeneity measure H_1. The mean absolute relative error is illustrated for sites with at least four PREC-JK realisations.

Tab. 2.7: Overall performance indices of the jackknifing procedure for both pooling methods and the different thresholds of the heterogeneity measure.

	Cluster analysis	Region of Influence
$H_1 < 1$: n = 57		
Mean of the mean absolute relative error	0.54	0.54
Mean of the standard deviation of absolute relative error	0.21	0.26
$H_1 < 2$: n = 70		
Mean of the mean absolute relative error	0.81	0.69
Mean of the standard deviation of absolute relative error	0.36	0.40
$H_1 < 4$: n = 75		
Mean of the mean absolute relative error	1.12	0.88
Mean of the standard deviation of absolute relative error	0.56	0.53

n = Number of sites with at least four PREC realisations

An overall performance indice was calculated as follows. All sites were selected which had at least four realisations for both pooling methods (see Tab. 2.7). The mean and the standard deviation of the absolute relative errors were calculated for all PREC realisations of these sites (n in Tab. 2.7). Both were averaged over the n sites. These performance indices increase with a higher degree of heterogeneity (Tab. 2.7). The result emphasises the relevance of the homogeneity criteria for PREC. The two performance indices are similar for the cluster analysis and RoI for the three thresholds of heterogeneity.

2.5 Discussion

The method of probabilistic regional envelope curves (PREC) derives a flood discharge and its recurrence interval for a homogeneous group of discharge gauges. One main assumption is its applicability in a homogeneous region in terms of the index flood method.

By using different subsets of catchment descriptors and two pooling methods (cluster analysis and RoI), a large number of homogeneous regions, which fulfilled the heterogeneity measure of Hosking and Wallis (1993), was derived for the mountainous catchments in Saxony. In contrast, the gauges located in the lowlands were mostly grouped in heterogeneous regions, which mean that the method of PREC could not be applied.

The reliability of PREC was assessed by a cross-validation procedure and a comparison with the index flood method. For a better understanding of the cross-validation results, it is worth emphasising an important difference between the index flood method and the PRECs. The index flood method represents the mean flood behaviour in a homogeneous region by a regional growth curve. Under this assumption it is expected that there are very small differences between the at-site flood behaviour and the regional distribution function in a homogeneous region. In contrast, the regional envelope curve is governed by the highest flood of record in a homogeneous region. Under the assumption that the estimation of the flood of record is more uncertain than the estimation of the index flood, the PREC is more sensitive to gauging stations with a high difference of an at-site flood of record to PREC than the index flood estimation.

The results of the PREC can be compared with a traditional at-site flood frequency analysis. The example of Dohna shows that most of the PREC realisations are close to the GEV distribution function (Fig. 2.10). This fact enhances the accuracy of the flood quantile estimates for high recurrence intervals. If there were large deviations between PREC and at-site flood frequency analysis, a more detailed consideration of the hydrologic situation at this gauge would be required.

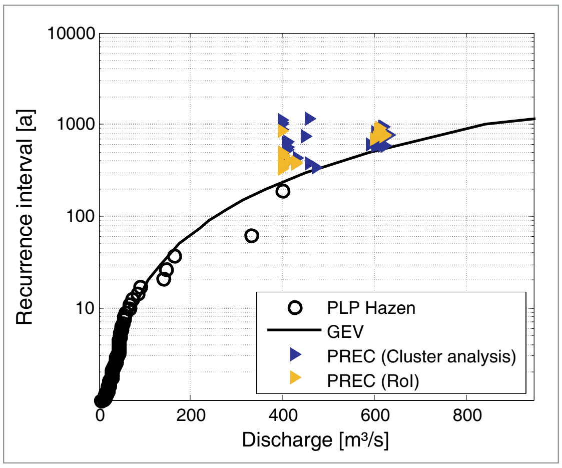

Fig. 2.10: Comparison of PREC results of both pooling methods with at-site flood frequency analysis (GEV) for Dohna.

It is important to highlight an essential difference of the PREC in comparison to other regionalisation methods. The magnitude of the recurrence interval of a PREC is mainly governed by one data point, i.e. the pair of the maximum unit flood of record and its drainage area. Castellarin et al. (2005) emphasised that a discordant site might reduce the use of the PREC method, since the recurrence interval is governed by the largest standardised maximum flood.

In other flood regionalisation methods (e.g. index flood, multiple regressions) commonly all sites have the same influence or their influence is weighted according to a selected weighting scheme. Sites, which are closer to other stations in a real or physiographical space, have higher weights. Consequently, the effect of a discordant site could be reduced by weighting the sites according to their similarity to the considered site or by averaging the values for all sites of a region. However, in the PREC concept weighting or averaging of sites is not possible when deriving the intercept of the PREC. Thus, in the PREC concept, the site that determines the intercept, plays an exceptional role. Because of that, an appropriate construction of homogeneous pooling groups is extremely important for PRECs.

The explicit estimation of a recurrence interval in the PREC scheme is another difference to traditional regional flood frequency methods. Whereas a target recurrence interval might be predefined in traditional approaches, the recurrence interval of PREC could only be approximately approached by the number of sites within a pooling group.

2.6 Conclusion

In this study the method of probabilistic regional envelope curves (PREC) was applied for the first time outside the original study area in Italy. It was shown that the transfer of this method to another region with different geographical conditions is possible. The goal of this paper was to quantify the influence of the pooling methods on PREC and to determine the sensitivity of PREC flood quantiles within different pooling groups. A combination of PREC and the RoI approach was introduced and compared with fixed homogeneous regions.

The main outcomes of this study are:

(1) The number of homogeneous regions strongly depends on the physiographic conditions of the catchment. The application of both pooling methods with different candidate sets of catchment descriptors leads to a high number of homogeneous regions for the mountainous catchments and to a lower number for gauges in the lowlands and the eastern part of Saxony.

(2) A sensitivity analysis illustrates that PREC flood quantiles change in discharge as well as in the assigned recurrence interval depending on the constitution of the pooling group. It is thus recommended to compare different subsets as demonstrated in this study instead of using only the best subset of predictions.

(3) A leave-one-out jackknifing approach for ungauged conditions emphasises a similar relative error of the PREC results for both pooling methods (cluster analysis, RoI). An overall performance indice also affirms an increasing absolute relative error for different degrees of heterogeneity.

Acknowledgements

This work is part of the Center for Disaster Management and Risk Reduction Technologies (CEDIM) (http://www.cedim.de), a joint venture between the Helmholtz Centre Potsdam - GFZ German Research Centre for Geosciences, the Technical University of Karlsruhe (TH) and the Forschungszentrum Karlsruhe. We thank CEDIM and the GFZ for the financial support. Attilio Castellarin acknowledges the Italian Ministry of Education for the grant no. 2007HBTS85_003.

We thank the State Agency of Environment and Geology of the Free State of Saxony for the permission to use the discharge data and the hydrogeological map (HÜK 200) (http://www.umwelt.sachsen.de/umwelt/geologie/7703.htm, 05.02.2007). Furthermore we thank the State Agency of Flood Prevention and Water Management of Saxony-Anhalt and the State Agency for Environment of Thuringia for additional discharge data. We also thank the Potsdam Institute for Climate Impact Research (PIK) and the German Weather Service (DWD) for the precipitation data as well as the Federal Agency for Cartography and Geodesy of Germany (BKG) for the ATKIS-Basis-DLM and the digital elevation model for Saxony. We also thank the European Commission – Joint Research Centre (JRC) for the CCM River and Catchment Database. The SRTM Digital Terrain Model was downloaded from (http://srtm.csi.cgiar.org/SELECTION/inputCoord.asp, 19.05.2008).

We acknowledge the very helpful recommendations of an anonymous reviewer and Alberto Montanari.

CHAPTER 3

EFFECTS OF

INTERSITE DEPENDENCE OF

NESTED CATCHMENT STRUCTURES

ON PROBABILISTIC REGIONAL

ENVELOPE CURVES

3 Effects of intersite dependence of nested catchment structures on probabilistic regional envelope curves

Abstract

This study analyses the intersite dependence of nested catchment structures by modelling cross-correlations for pairs of nested and unnested catchments separately. Probabilistic regional envelope curves are utilised to derive regional flood quantiles for 89 catchments located in Saxony, in the Southeast of Germany. The study area has a nested structure and the intersite correlation is much stronger for nested pairs of catchments than for unnested ones. Pooling groups of sites (regions) are constructed based on several candidate sets of catchment descriptors using the Region of Influence method. Probabilistic regional envelope curves are derived on the basis of flood flows observed within the pooling groups. Their estimated recurrence intervals are based on the number of effective sample years of data (i.e. equivalent number of uncorrelated data). The evaluation of the effective sample years of data requires the modelling of intersite dependence. We perform this globally, using a cross-correlation function for the whole study area as well as by using two different cross-correlation functions, one for nested pairs and another for unnested pairs. In the majority of the cases, these two modelling approaches yield significantly different estimates for the effective sample years of data, and therefore also for the recurrence intervals. The reduction of the recurrence interval when using two different cross-correlation functions is larger for larger pooling groups and for pooling groups with a higher fraction of nested catchments. A separation into nested and unnested pairs of catchments gives a more realistic representation of the characteristic river network structure and improves the estimation of regional information content. Hence, applying two different cross-correlation functions is recommended.

Published as:

Guse, B., Castellarin, A., Thieken, A. H., and Merz, B., 2009. Effects of intersite dependence of nested catchment structures on probabilistic regional envelope curves, Hydrol. Earth Syst. Sci., 13(9), 1699-1712, doi:10.5194/hess-13-1699-2009.

3.1 Introduction

The estimation of flood quantiles is a major topic in hydrologic research and engineering practise. Due to the uncertainty in the estimation of flood discharges with large recurrence intervals T (e.g. T>100 years) by an at-site flood frequency analysis, several gauges may be pooled together in a pooling group following the principle 'trading space for time' (e.g. Stedinger et al., 1993; Robson and Reed, 1999). Therefore, it is assumed that analysis results, e.g. results of a regional flood frequency analysis (RFFA), are valid for all gauges of a specific pooling group. RFFA aims at improving the estimation of flood quantiles by using the larger number of flood data. However, an improvement can only be reached by increasing the effective sample years of data (i.e. the number of independent observations). The increase in the effective sample years of data when adding a new site to a pooling group can be assessed by considering the intersite correlations or cross-correlations among all gauges in a pooling group (Matalas and Langbein, 1962).

A pooling group comprises catchments of similar hydrologic behaviour. In flood regionalisation studies, fixed homogeneous regions are traditionally used, whereby each site is explicitly assigned to one region, e.g. through cluster analysis (e.g. Acreman and Sinclair, 1986; Nathan and McMahon, 1990; Rao and Srinivas, 2006). In contrast, the Region of the Influence (RoI) approach (e.g. Burn, 1990a,b; Zrinji and Burn, 1994) constructs a separate pooling group for each site in the region under study.

Several methods and many studies on regional flood frequency analysis have been presented (e.g. Cunnane, 1988; GREHYS, 1996a,b; Robson and Reed, 1999; Merz and Blöschl, 2005; Ouarda et al., 2008). A common application is the widely used index flood approach, which assumes that a regional growth curve is representative for all sites of a homogeneous region. The at-site flood quantiles vary only in the scale factor index flood (e.g. Dalrymple, 1960; Stedinger and Lu, 1995; Robson and Reed, 1999). Linear regression models relate catchment descriptors (e.g. drainage area, precipitation indices) to a predefined flood quantile (e.g. Rosbjerg and Madsen, 1995; Kroll and Stedinger, 1998; Robson and Reed, 1999; Reis et al., 2005). Recently, geostatistical methods (e.g. Top-Kriging) were introduced to regionalise flood quantiles (Merz and Blöschl, 2005; Skoien et al., 2006).

Regional envelope curves (REC) are a variant of linear regression models which only use the size of the drainage area to estimate the maximum flood discharge (e.g. Crippen and Bue, 1977; Herschy, 2002). A shortcoming of the traditional REC method is, however, that no recurrence interval can be assigned to the maximum discharge. Therefore, Castellarin et al. (2005) proposed the method of probabilistic regional envelope curves (PREC), which enhance the traditional REC approach with a probabilistic interpretation. The method of PREC requires a pooling group, which fulfils the homogeneity criteria of the index flood method. The recurrence interval of PREC is directly related to the effective sample years of data. Hence, its calculation algorithm explicitly considers the effect of cross-correlated sites in a pooling group of data.

Several studies have demonstrated the relevance of intersite correlation for regional flood estimates (e.g. Stedinger, 1983; Hosking and Wallis, 1988; Madsen and Rosbjerg, 1997a; Vogel et al., 2001). Matalas and Langbein (1962) introduced the concept of regional information content to determine the effect of intersite correlation within flood sequences. The regional information content expresses the number of independent discharge observations. The authors showed that the variance of the regional mean increases for cross-correlated sites.

Kuczera (1983) assessed that a low number of observations and the presence of intersite correlation leads to a larger uncertainty of an empirical Bayes estimator. Stedinger (1983) demonstrated that the variance of the regional variance and skewness increases due to intersite correlation. Hosking and Wallis (1988) pointed out that cross-correlation among sites leads to less accurate estimates of regional flood quantiles; however, the influence of regional heterogeneity is more significant. Applying hydrologic linear regression models, Stedinger and Tasker (1985) introduced intersite correlation by extending the weighted least square (WLS) to the generalised least square (GLS) method, which explicitly considers the impact of cross-correlated sites. Several studies confirmed that the GLS estimator outperforms the WLS or the ordinary least squares (OLS) estimator for the application of linear regression models in the case of cross-correlated sites (see e.g. Stedinger and Tasker, 1985, 1986; Kroll and Stedinger, 1998; Reis et al., 2005).

The impact of intersite correlation on regional estimates was analysed for Partial Duration Series by Madsen and Rosbjerg (1997a, b) and for Annual Maxima Series by Kjeldsen and Rosbjerg (2002) and Kjeldsen and Jones (2006). Madsen and Rosbjerg (1997a) pointed out that intersite correlation needs to be considered to accurately assess the uncertainty of the regional estimator. It has recently been demonstrated by Castellarin et al. (2008) that intersite correlation affects the heterogeneity measure of Hosking and Wallis (1993), which estimates the hydrologic heterogeneity of a region.

The distance between two catchments is generally assumed to be the most important factor for intersite correlation resulting in different cross-correlation models. In these models, the correlation coefficient decreases as a function of the distance between the catchments (see e.g. Tasker and Stedinger, 1989; Troutman and Karlinger, 2003).

The effects of the river network structure and mutual location of catchments were considered by Troutman and Karlinger (2003). They pointed out that peak flows between nested catchments, i.e. catchments along the same stream, are more correlated than peak flows between unnested catchments. In terms of flood regionalisation methods Skoien et al. (2006) demonstrated the better performance of Top-Kriging, which considers the effect the river network structure, in comparison to a traditional Ordinary Kriging approach, which is based only on the distances between the catchments.

Castellarin et al. (2005) developed an empirical function by using a Monte-Carlo simulation to reveal the reduction of the overall sample years of data in a pooling group due to intersite correlation, and to obtain the effective number of sample years of data for estimating the recurrence interval of a PREC. This is equivalent to the number of independent data associated with the concept of information content by Matalas and Langbein (1962). Castellarin (2007) examined the accuracy of PREC flood quantiles by comparing different cross-correlation functions for an Italian data set. Owing to the small number of nested catchments, different cross-correlation functions for nested and unnested catchment relationships were not estimated.

In this study, we assess the impact of different approaches to model regional cross-correlation structure with respect to their impact on the effective number of observations and the recurrence interval of probabilistic regional envelope curves (PREC). First, a global approach is considered, in which the cross-correlation function is identified for the whole study area. Second, the method of PREC as described by Castellarin et al. (2005) and Castellarin (2007) is extended by deriving two different cross-correlation functions, one for nested pairs of catchments and one for unnested ones. While applying both approaches we did not vary any other aspect of flood regionalisation (e.g. selection of catchment descriptors, pooling method, etc.), since our investigation mainly focuses on the correlation structure for nested and unnested pairs of catchments. Significant factors, which influence the effect of intersite correlation on PREC, are determined. The study region, Saxony in south-eastern Germany, includes several pairs of nested catchments and enables us to examine in detail the effect of nested catchment structures on PREC flood quantiles, whose importance was not adequately acknowledged in previous studies.

3.2 Methods

3.2.1 Regional information content and number of effective observations

The regional information content (IC) can be defined as the ratio of the effective sample years of data n_{eff} to the total sample years of data n. The effective sample years of data represents the equivalent number of independent observations within a pooling group (Eq. 3.1).

$$IC = \frac{n_{eff}}{n} \tag{3.1}$$

The core idea of regional information content (Matalas and Langbein, 1962) is that a correlated site gives a lower degree of additional information to the site being studied than an uncorrelated site. Hence, the additional information decreases for a higher intersite correlation. An IC of 1 means that these sites are completely uncorrelated (independent), implying that the total flood sequence gives additional information ($n_{eff} = n$). In contrast, a small value of IC indicates that there is only low additional information within the time series (Matalas and Langbein, 1962).

On the basis of the regional information content, Castellarin et al. (2005) and Castellarin (2007) estimated the exceedance probability of a regional envelope curve. The effective sample years of data n_{eff}, hereafter also referred to as the number of effective observations, were calculated by reducing the total sample years of the AMS of all gauges in two steps (Castellarin, 2007). First, the intersite correlation between the different AMS was modelled as a function of the distance between the catchment centroids. Second, the results of the cross-correlation function were used to estimate the number of effective observations.

(1) A regional cross-correlation function (Eq. 3.2, from Castellarin, 2007), proposed by Tasker and Stedinger (1989), was applied, which estimates the cross-correlation as a function of the distance.

$$\rho_{i,j} = \exp\left(-\frac{\lambda_1 d_{i,j}}{1+\lambda_2 d_{i,j}}\right) \tag{3.2}$$

d= distance between catchment centroids, ρ= correlation coefficient by Pearson, λ_1, λ_2= parameters, i, j= index denoting pairs of catchments

Therefore, empirical cross-correlation coefficients between the AMS were related to the distances between the catchment centroids. By using catchment centroids the river network structure is incorporated in the distance calculation (e.g. Troutman and Karlinger, 2003). The parameters λ_1 and λ_2 of the cross-correlation function were fitted by a weighted optimisation, in which empirical coefficients were weighted proportionally to the length of the overlapping time series.

(2) The number of effective observations was calculated by an empirical relationship, which incorporated the theoretical average cross-correlation values from Eq. (3.2). The data set of a specific pooling group comprises M times AMS with a variable length, but not more than Y years. The length of the AMS varies due to missing observations or different observations periods. In a first step, all years n_1 with only one observation among the M discharge time series were considered separately. In these years, all other (M-1) gauging stations have no discharge measurements. The n_1 observations are certainly effective, because there is only one discharge value within the pooling group for this year (Castellarin, 2007).

After this, the Y–n_1 remaining years were analysed. These years were split into Y_{sub} subsets with $Y_{sub} \leq (Y-n_1)$. In each of these subsets (denoted as s), there were L_s flood sequences with an equal length of years l_s with $L_s \leq M$. The number of effective observations of a subset $n_{eff,s}$ was calculated for each subset s separately. Then, the numbers of effective observations $n_{eff,s}$ were summed up for all subsets. In the last step, the number of effective sample years of data n_{eff} comprises the n_1 years with one observation and the sum of the numbers of effective observations $n_{eff,s}$ for all Y_{sub} subsets.

The number of effective observations was calculated by an empirical relationship derived by Castellarin et al. (2005) and Castellarin (2007) from Monte-Carlo simulations (Eq. (3.3), adopted from Castellarin, 2007).

$$n_{eff} = n_1 + \sum_{s=1}^{Y_{sub}} n_{eff,s} = n_1 + \sum_{s=1}^{Y_{sub}} \frac{L_s l_s}{1 + \left[\overline{\rho^{\beta}}\right]_{L_s} (L_s - 1)}$$

$$\text{with } \beta := 1.4 \frac{(L_s l_s)^{0.176}}{\left[\overline{(1-\rho)^{0.376}}\right]_{L_s}}, \qquad (3.3)$$

where the term $[\;]_{L_s}$ denotes that the average values $\overline{\rho^{\beta}}$ and $\overline{(1-\rho)^{0.376}}$ are calculated for the L_s annual flood sequences.

It becomes apparent that different parameter values for λ_1 and λ_2 directly affect the results of Eq. (3.2) and therefore also of Eq. (3.3). Consequently, the number of effective observations n_{eff} is affected by using different parameter sets for the cross-correlation function. Equation (3.3) illustrates that the magnitude of n_{eff} depends on the size of the available data set and their cross-correlation characteristics.

3.2.2 Probabilistic Regional Envelope Curves

The calculation of the number of effective sample years of data is a fundamental step towards the estimation of the exceedance probability of a regional envelope curve. A regional envelope curve (REC) is determined by relating all floods of record Q_{FOR} normalised by the drainage area A to A (Eq. (3.4), adopted from Castellarin, 2007). The flood of record is the largest discharge of each time series.

$$\log\left(\frac{Q_{FOR}}{A}\right) = a + b * \log(A) \qquad (3.4)$$

a= intercept of REC, b = slope of REC.

The concept of the probabilistic regional envelope curve (PREC) requires that two basic assumptions are fulfilled: Firstly, PREC is based on the index flood hypothesis. The index flood method (Dalrymple, 1960) requires that all selected flood series constitute a homogeneous region. These flood series are identically distributed, i.e. have the same growth curve, except for the scale parameter, the index flood (e.g. Robson and Reed, 1999). In this study, the mean of the AMS was used as index flood. Secondly, there is a scaling of the index flood μ_X to the drainage area (A) (Eq. (3.5), from Castellarin et al., 2005). The index flood depends on the drainage area alone.

$$\mu_X = C * A^{b+1} \qquad (3.5)$$

A regional envelope curve can be derived in two steps. First, the slope b is estimated by a regression analysis (orange line in Fig. 3.1). The second step is a parallel upshift of the regression line up to the intercept a. Then all floods of record are bounded by REC (blue line in Fig. 3.1) (Castellarin et al., 2005).

Since the PREC method is based on the index flood hypothesis, the derivation of a pooling group which fulfils the homogeneity criteria of the index flood hypothesis is an essential step in the PREC concept. In this work, PREC was applied for all regions with at least four sites.

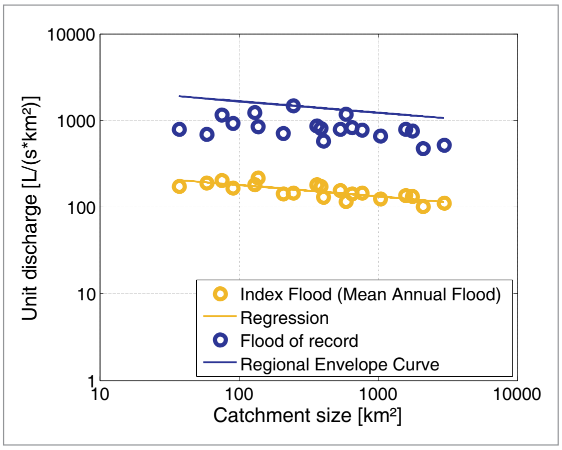

Fig. 3.1: Example of a Regional Envelope Curve (REC)

The core idea of PREC is the assignment of an exceedance probability to a REC. The exceedance probability is estimated for that particular data pair (i.e. the unit flood of record and associated drainage area) which determines the intercept of REC. This is the exceedance probability of the largest unit flood of record in the region.

For this purpose, the plotting position of the maximum unit flood of record was used, which was determined by the number of effective observations n_{eff} (Eq. 3.3) and the Hazen function (Eq. (3.6), from Castellarin, 2007).

$$T = 2 * n_{eff} \tag{3.6}$$

Castellarin (2007) showed that the Hazen function is a suitable quantile unbiased plotting position when the Generalised Extreme Value (GEV) distribution is an adequate parent distribution. The suitability of the GEV for this study is discussed in section 3.3. Equation (3.6) implies that the reduction effect of intersite correlations on n_{eff} directly affects the estimation of T. The recurrence interval T, i.e. the inverse of the exceedance probability, is derived for the entire pooling group and therefore is identical for all gauges. Its validity is restricted to the range of the catchment size within the pooling group, i.e. from the smallest to the largest catchment size. Thus, the use of different parameter values for the cross-correlation function affects T in the same way as n_{eff}. We referred to Castellarin et al. (2005), Castellarin (2007) and Castellarin et al. (2007) for more detailed information of the PREC concept.

3.2.3 Pooling scheme

The method of PREC is based on the index flood hypothesis. This implies the need of pooling groups fulfilling the homogeneity criteria of the index flood method. In this case, the PREC concept is valid for all sites of the pooling group. This study is tailored to assess the impact of different approaches to the modelling of the regional cross-correlation structure on several PREC applications. For this purpose several pooling groups, derived using the Region of Influence (RoI) method, are needed. We derived several candidate sets of catchment descriptors instead of one ‘best subset’, because the use of a ‘best subset’ neglects that different subsets of catchment descriptors could have a similar performance. The pooling groups were constructed by the following six steps:

(1) we selected meaningful predictor variables which were standardised (mean=0, std=1) to allow a comparison between them. We combined the standardised catchment descriptors to create all possi-

ble subsets with one, two and three catchment descriptors.

(2) We defined only those subsets of catchment descriptors as candidate sets, which had the largest correlation to the empirical unit index floods. A correlation coefficient of 0.6 was selected as threshold. Index flood values were used as explained variable because the PREC method is based on a scaling of the index flood values with the drainage area (see Eq. 3.5).

(3) We then checked all candidate sets with three catchment descriptors on redundancy compared to the selected subsets of two catchment descriptors. We only maintained candidate sets with three catchment descriptors which led to a larger proportion of explained variance (higher correlation coefficient).

(4) We checked all subsets on multicollinearity by the Variance Inflation Factor (VIF) (Hirsch et al., 1992) and removed them if the VIF was larger than five.

(5) Each remaining candidate set was used to derive a Region of Influence (RoI) (Burn, 1990a,b). The RoI method identifies a specific pooling group of sites (region in the widest sense) for each gauge (site of interest). The rationale behind this approach is that the specific hydrologic conditions of the site of interest are considered to select hydrologically similar gauges. Instead of the geographical distance, a physiographical space was formed by the catchment descriptors of a selected subset. By selecting sites close to the site of interest in the physiographical space, adequate sites were determined for constructing a RoI. We used the Euclidean distance in the physiographical space between the sites to evaluate their similarity to the site of interest (Burn, 1990a,b). RoIs were formed by assessing three different thresholds (0.5, 1 and 2) of the Euclidean distance. The different thresholds reflect the trade-off between the size and the regional homogeneity of a pooling group (e.g. Burn, 1990a, Castellarin et al., 2001). In a preliminary analysis, another RoI variant was also applied as proposed by Gaál and Kyselý (2009). Thereby, we started with a RoI which includes the ten most similar sites. The size of the RoI was determined by iteratively adding sites to the RoI until the threshold of $H_1<2$ was exceeded or in the case of an initially heterogeneous RoI by removing sites until the H_1-test falls below the threshold (step-wise approach) (see e.g. Zrinji and Burn, 1994, Castellarin et al., 2001, Gaál and Kyselý, 2009). Since we found no significant variations in the results, we only report here the results of the first RoI variant.

(6) The pooling groups constructed by RoI were tested on homogeneity by the heterogeneity measure (H-test) of Hosking and Wallis (1993). The H-test compares the regional heterogeneity of a pooling group in terms of the variability of L-moment ratios with simulated synthetic time series calculated by a Monte-Carlo

simulation. The H_1-test focuses on the sample variability of the L-coefficient of variation (L-CV). Since synthetic time series generated in the test are independent by definition, intersite correlation introduces some bias in the H_1-test results (Castellarin et al., 2008). Hosking and Wallis (1997) mentioned that a very low value of their heterogeneity measure (H_1<-2) indicates a high intersite correlation. All regions with H_1-values lower than 1 are acceptably homogeneous and with H_1-values between 1 and 2 are possibly heterogeneous (Hosking and Wallis, 1997). This means that a modification of the region is not required or optional, respectively. The H_1-test was performed with the hw.test (Viglione, 2008, implemented in R). We used each RoI with H_1<2 to form a pooling group and to derive a PREC. Hence, the number of PREC realisations was identical with the number of homogeneous RoIs.

3.2.4 Application and interpretation of different cross-correlation functions

The number of effective observations was calculated using the cross-correlation function (see Eq. 3.2) with separately optimised parameter sets. In a first approach, the number of effective observations (see Eq. 3.3) was calculated by using one cross-correlation function for the whole study area (global approach, termed: $n_{eff,G}$). Second, the cross-correlation function was applied with different parameter sets for nested and unnested catchments (nested approach, termed: $n_{eff,N}$). The parameter set for nested structures was used for the pairs of catchments which are in an upstream-downstream relationship, termed P_N. For all of the others, the unnested parameter set was employed. The numbers of the effective observations $n_{eff,G}$ and $n_{eff,N}$ were compared for the same pooling groups. Therefore, the information content (IC), i.e. the fraction of the effective observations n_{eff} to the total observations n, was calculated according to Eq. (3.1) for the global case (IC_G) and the nested-unnested (IC_N) approach. In a next step, the ratio R_N was calculated as a function of the differences between $n_{eff,N}$ and $n_{eff,G}$ using $n_{eff,G}$ as reference (Eq. 3.7).

$$R_N = \frac{n_{eff,N} - n_{eff,G}}{n_{eff,G}} * 100 = \frac{T_N - T_G}{T_G} * 100 \qquad (3.7)$$

The ratio R_N facilitates the interpretation of the influence of the different parameter sets on the number of effective observations. Furthermore, the study focuses on the recurrence interval T of PREC, which, according to Eq. (3.6), is twice as high as the number of effective observations. Consequently, the ratio R_N is identical when using T instead of n_{eff} (Eq. 3.7).

The effect of the nested structure on T was investigated by calculating a degree of nesting D_N (Eq. 3.8). It is defined as the ratio between nested catchment pairs P_N and all pairs of catchments P in a pooling group.

$$D_N = \frac{\sum_{k=1}^{Y-n_1} \frac{P_{N_k}}{P_k}}{Y - n_1} \qquad (3.8)$$

The nested catchment structure was estimated separately for each year. The years with a single observation (n_1 in Eq. 3.3) were not considered in this approach. Ultimately, the mean degree of nesting D_N for the Y-n_1 years (see Sect. 3.2.1) was calculated for each RoI.

Finally, since the estimation of flood quantiles derived by PRECs is affected by intersite correlation (through the different parameter sets of the cross-correlation function), but also by regional heterogeneity (through the threshold adopted for the heterogeneity measure H_1), we analysed the influence of different thresholds of H_1 on R_N. As regional homogeneity is a fundamental prerequisite for applying PREC (see Castellarin et al., 2005; Castellarin 2007), we considered two different thresholds, $H_1<1$ and $H_1<2$, that according to Hosking and Wallis (1997) refer to 'acceptably homogeneous' and 'possibly heterogeneous' regions, in this order. Following the main hypotheses for applying PREC, we did not consider larger thresholds (i.e. larger heterogeneity degrees).

3.3 Study area and data

The federal state of Saxony in the south-eastern Germany has a size of about 18.400 km² and is characterised by higher elevations in the Southwest (*Erzgebirge*) and lower elevations in the northern parts (Fig. 3.2). The Elbe is the largest river with a catchment size of about 52,000 km² at Dresden gauge. There are five large catchments in Saxony (Weisse Elster, Mulde, Schwarze Elster, Spree and Lausitzer Neisse, from west to east) and several tributaries to the River Elbe (Western and Eastern tributaries) (Fig. 3.2).

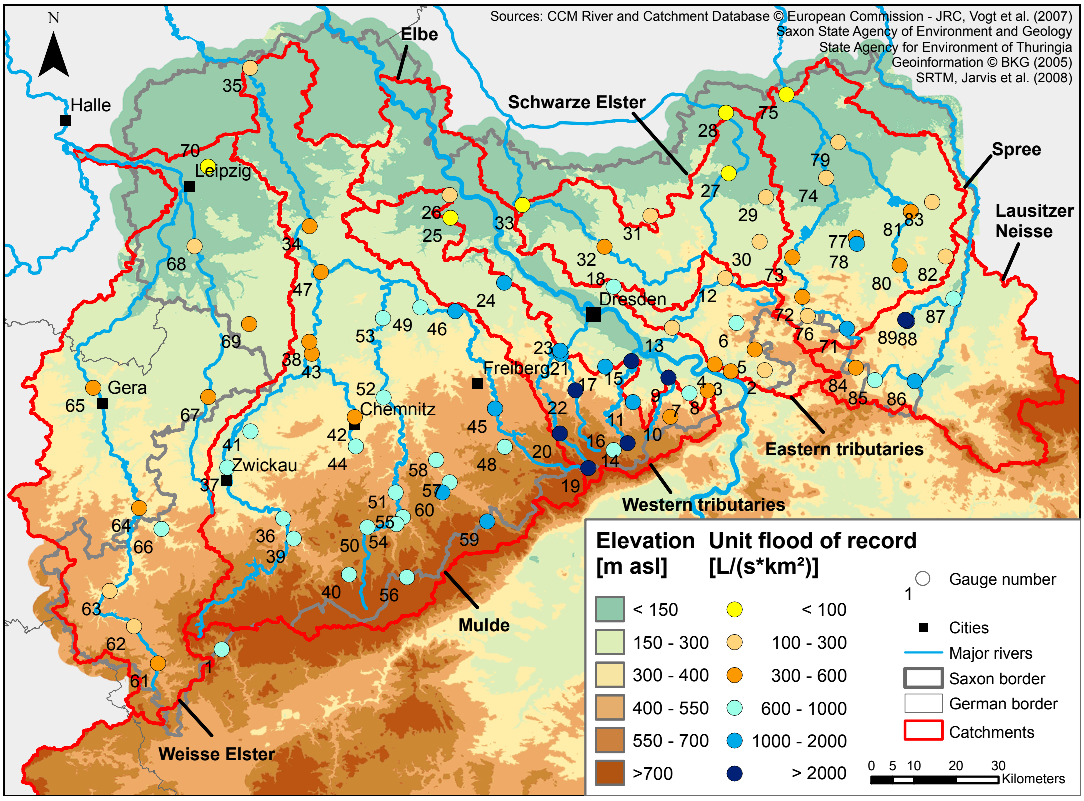

Fig. 3.2: Elevation and gauging stations in Saxony, Germany. The colour scale indicates the unit flood of record at each gauge.

We only used gauges that (1) had a time series of more than 29 years, (2) were not strongly influenced by mining activities, (3) had a catchment size larger than 10 km², and (4) were not located directly downstream of a dam. Furthermore, we omitted gauges whose catchments were mostly outside of Saxony. Ultimately, we considered 89 gauges as indicated in Fig. 3.2. Most of the gauges have a few nested catchment relationships. Nested catchment structures are especially located in the Mulde catchment. All tributaries of the Mulde catchments originating in the *Erzgebirge* are related to the two most downstream gauges (sites 34-35 in Fig. 3.2).

We derived the annual maxima series (AMS) as well as the highest observed discharge, the flood of record Q_{FOR}, for all gauges. The suitability of the GEV as parent distribution for the 89 gauges was checked by a L-moment ratio diagram (see e.g. Vogel et al., 1993; Peel et al., 2001). It clearly stated that the use of the GEV was adequate.

Climatic, geologic and land-use data were used to derive catchment descriptors as basis for pooling catchments into homogeneous regions. Precipitation data was provided by the German Weather Service (DWD). We estimated precipitation indices from 453 stations in and around Saxony which had a record length of at least 30 years and still existed in 2002. We selected this year because of a severe wide-spread flood which occurred in 2002, in particular along the Elbe and Mulde. During this flood the highest daily precipitation ever recorded in Germany was measured. Therefore it is important to include the precipitation values of this year, e.g. to calculate the maximum daily precipitation. Additional precipitation stations were used to calculate the maximum daily precipitation and the maximum five-day precipitation sum. To better cover the spatial variability of precipitation, we improved the spatial resolution of precipitation stations by adding precipitation time series shorter than thirty years when the flood of record of the downstream gauge occurred during the period covered by the shorter precipitation time series. This led to 23 additional precipitation stations (476 in total) which could be used to calculate the maximum daily precipitation and the maximum five-day precipitation sum. All precipitation indices were interpolated by ordinary kriging.

Mean elevation, mean slope and catchment centroids were derived from digital elevation models. In Saxony a grid size of 25 meters was used, whereas the digital elevation model from the Shuttle Radar Topographic Mission (SRTM) with a grid size of 90 meters (Jarvis et al., 2008) was resampled to a grid size of 25 meters for the areas outside of Saxony. Catchment centroids were required for the optimisation of the cross-correlation function (see Eq. 3.2). Furthermore, landscape parameters were derived from the digital landscape model ATKIS (BKG GeoDataCentre, 2005) and hydrogeological parameters were taken from the hydrogeological map (HÜK200) by the Saxon State Agency of Environment and Geology. Altogether, 13 catchment descriptors were selected (Tab. 3.1).

Tab. 3.1: List of catchment descriptors

Abbreviation	Catchment descriptors
MAP	Mean annual precipitation [mm]
MAXDAY	Maximum daily precipitation [mm]
P50	Annual frequency of days with precipitation of more than 50 mm/d [%]
MAX5DAY	Maximum precipitation in five days [mm]
PAMS	Mean of the annual maximum series of daily precipitation [mm]
ELEV	Mean elevation of the catchment [m]
SLOPE	Mean slope of the catchment [%]
RANGE_NORM	Range of catchment elevation, normalised with the catchment size [$10^{-3}m^{-1}$]
ARABLE	Fraction of arable land coverage [%]
URBAN	Fraction of urban land coverage [%]
MINING	Fraction of mining activities [%]
BEDROCK	Fraction of bedrock areas [%]
KF_LOW	Fraction of low permeability areas [%]

3.4 Results

3.4.1 Intersite correlation in the study area

Figure 3.3 illustrates the variability of empirical correlation coefficients for pairs of annual flood sequences in the study area. The heterogeneity of the correlation pattern becomes apparent when comparing empirical correlation coefficients higher than 0.8 (e.g. Mulde gauges, sites 34-60 in Fig. 3.3), as well as very low correlation coefficients (e.g. Mulde vs. Spree gauges (sites 71-83)). The gauges of the Mulde catchment and the western tributaries to the Elbe River originating in the *Erzgebirge* are characterised by large empirical correlation coefficients also beyond their catchment boundaries. The correlation coefficients of neighbouring catchments are larger than coefficients across the catchment boundaries. This correlation pattern demonstrates that AMS of neighbouring catchments are more correlated. Among the 3916 possible pairs of catchments, there are 179 nested (5%) and 3737 unnested ones.

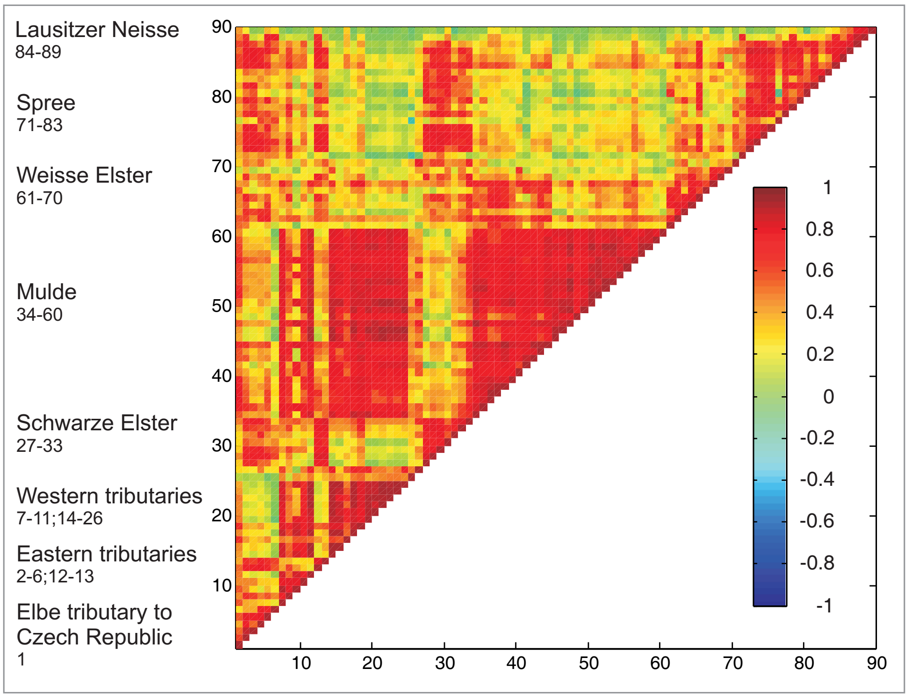

Fig. 3.3: Empirical cross-correlation coefficients for AMS of Saxon gauges.

3.4.2 Cross-correlation functions

The cross-correlation function (Eq. 3.2) was optimised for global, nested and unnested catchment relationships. The different parameter sets for the cross-correlation function are given in Tab. 3.2. The parameters for the global and unnested cases are similar, whereas the parameters for nested catchments are noticeably different.

The relationship of the correlation coefficient to the distance of the catchment centroids for all pairs of sites shows that the correlation coefficients vary between -0.25 and 1 (Fig. 3.4). As expected, the correlation decreases with increasing distance. Due to the structure of the river network in Saxony, all distances between the centroids of nested catchments are lower than 50 km, whereas unnested catchment relationships reach up to a distance of more than 200 km. Figure 3.4 illustrates that the cross-correlation functions for the global and the unnested case are very similar, whereas the nested cross-correlation function strongly differs from them.

The global and the unnested cross-correlation functions clearly decrease up to a distance of about 50 km between the catchment centroids. The slope of the functions decreases slightly for larger distances. The differentiation in nested and unnested catchments shows a remarkable difference in terms of average cross-correlation. As expected, the cross-correlation function for nested catchments yields higher correlations than the function for unnested ones, with differences of up to 0.2. To give an example, at a distance of 40 km, there is a correlation coefficient of 0.7 for nested catchments, but only 0.5 for unnested catchments (Fig. 3.4).

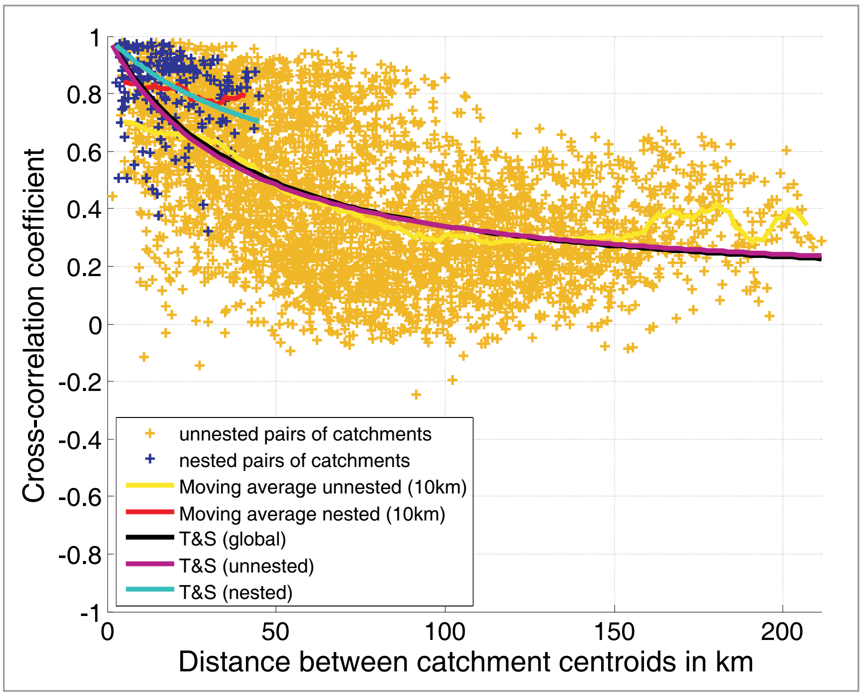

Fig. 3.4: Cross-correlation functions fitted to global, unnested and nested catchment structures (T&S: cross-correlation function by Tasker and Stedinger (1989)).

Tab. 3.2: Parameters (λ_1, λ_2) of the cross-correlation function by Tasker and Stedinger (1989) and available sample size (m) for different catchment structures

	Global	Nested	Unnested
λ_1	0.021	0.012	0.022
λ_2	0.009	0.012	0.011
m	3916	179	3737

The scattering in the correlation-distance plot (Fig. 3.4) illustrates that the distance between the catchment centroids is not the only relevant explanatory variable (see Troutman and Karlinger, 2003). However, the distance has a high explanatory power for this study area due to the significant decrease of the correlation coefficients with increasing distance.

3.4.3 Region of Influence

20 candidate sets listed in Tab. 3.3 met the criteria of the pooling scheme (see Sect. 3.2.3 (1)-(4)). They were used to construct Regions of Influence (RoI) and to derive the corresponding probabilistic regional envelope curves.

The RoI approach was applied to each of the 89 gauges separately, using the 20 candidate sets of catchment descriptors and the three different thresholds in the physiographical space. This led to a maximum possible number of 5340 pooling groups. The maximum number was not reached, since regions which were heterogeneous (H_1>2) or had a small number of sites within a RoI (n<4) were omitted. Ultimately, 1415 pooling groups with on average 13 sites fulfilled the assumption of the PREC concept and were used further.

Tab. 3.3 Subsets of catchment descriptors (CD) and the correlation coefficient (COR) to the index flood of the annual maxima series of all gauges.

CD1	CD2	CD3	COR
MAX5DAY	ELEV	RANGE_NORM	0.70
MAX5DAY	RANGE_NORM	URBAN	0.69
MAP	MAX5DAY	RANGE_NORM	0.69
MAX5DAY	RANGE_NORM		0.68
MAX5DAY	ELEV	URBAN	0.68
ELEV	RANGE_NORM	URBAN	0.66
PAMS	RANGE_NORM	URBAN	0.64
MAX5DAY	ELEV		0.64
ELEV	RANGE_NORM		0.64
MAP	MAX5DAY	URBAN	0.64
MAP	MAX5DAY		0.62
MAP	RANGE_NORM		0.62
PAMS	RANGE_NORM		0.62
P50	RANGE_NORM	URBAN	0.61
MAX5DAY	ARABLE	URBAN	0.61
MAXDAY	RANGE_NORM	URBAN	0.61
MAX5DAY	URBAN	BEDROCK	0.61
MAX5DAY	PAMS	URBAN	0.61
RANGE_NORM	URBAN	BEDROCK	0.60
RANGE_NORM	BEDROCK		0.60

3.4.4 Influence of intersite correlation on information content

The number of effective observations was calculated for all 1415 pooling groups with the global parameter set ($n_{eff,G}$) for the cross-correlation function as well as with the separate parameter sets for nested and unnested catchment structures ($n_{eff,N}$). Figure 3.5 illustrates that, as it is expected, the number of effective observations is lower than the number of total observations for all pooling groups. It further indicates that the ratio of the number of effective observations to the number of total observations - information contents IC_G (global approach) and IC_N (nested-unnested approach) - decreases as the number of total observations increases. To give an example, while the information content is about 0.5 for data sets with 600 total observations, it decreases to only 0.3 in the case of 2000 total observations.

These results show how the information content decreases when an additional site is added. The larger the number of sample years of data, the lower is the additional gain of information by adding one site to the pooling group. Furthermore, the additional gain of information is lower for nested catchments. Hence, the reduction effect of cross-correlated sites on n_{eff} becomes larger as the number of total observations n increases.

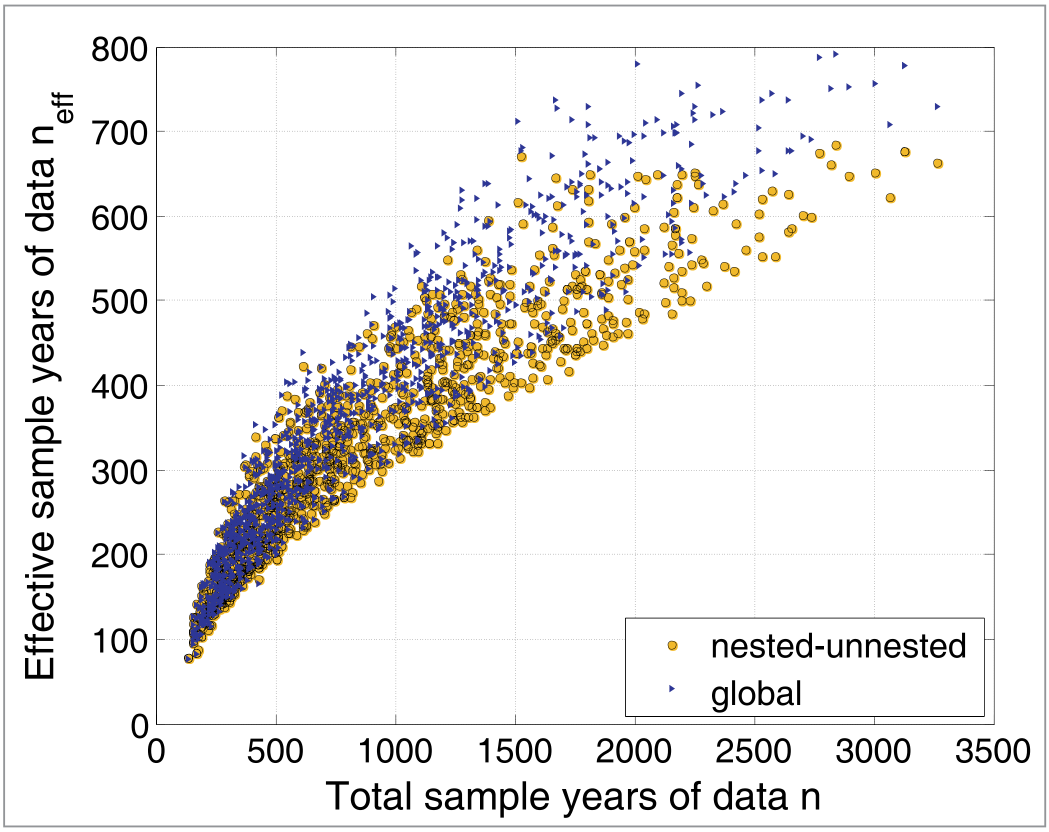

Fig. 3.5: Number of effective observations vs. sample years of data within the pooling groups for a global cross-correlation function and separate cross-correlation functions for nested and unnested catchments.

3.4.5 Recurrence interval

While a comparison of the effective sample years of data to the total sample years of data already illustrates the effect of intersite correlation on the information content, the recurrence interval T of PREC shows this effect more clear, since T is directly related to flood quantile estimates (see Eq. 3.6).

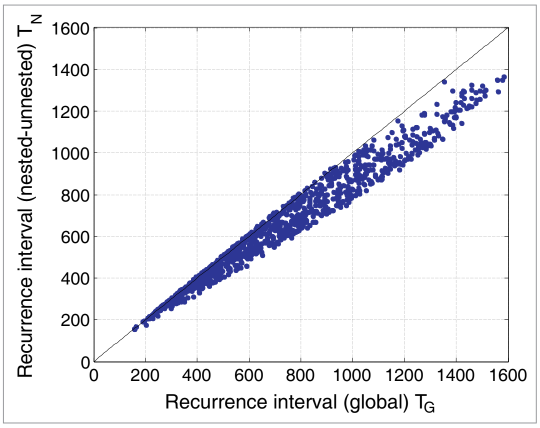

Fig. 3.6: Difference in recurrence intervals between a single cross-correlation function (global) and separated cross-correlation functions (nested-unnested).

A comparison of T_G and T_N reveals that the recurrence interval is higher in most cases when the global cross-correlation function is used (Fig. 3.6). The range of the ratio R_N is between -23 and 3% (Fig. 3.7a). In other words, the recurrence interval is up to 23% lower when using separate parameter sets for nested and unnested catchment relationships. The difference increases with increasing recurrence intervals (Fig. 3.6), but the ratio R_N does not show a distinct relation to the recurrence interval (Fig. 3.7a).

3.4.6 Degree of nesting

The calculation of T_G differs from T_N only in the parameter set for the cross-correlation function. Since there are large differences between the parameter sets for nested catchments on the one hand and for unnested catchments and the global approach on the other hand (see Tab. 3.2), it is interesting to look at the nested catchment structure in the study area. Therefore the ratio R_N is related to the degree of nesting D_N, i.e. the relative number of nested catchments within a pooling group (see Eq. 3.8). It is expected that R_N is mainly affected in pooling groups with a large degree of nesting.

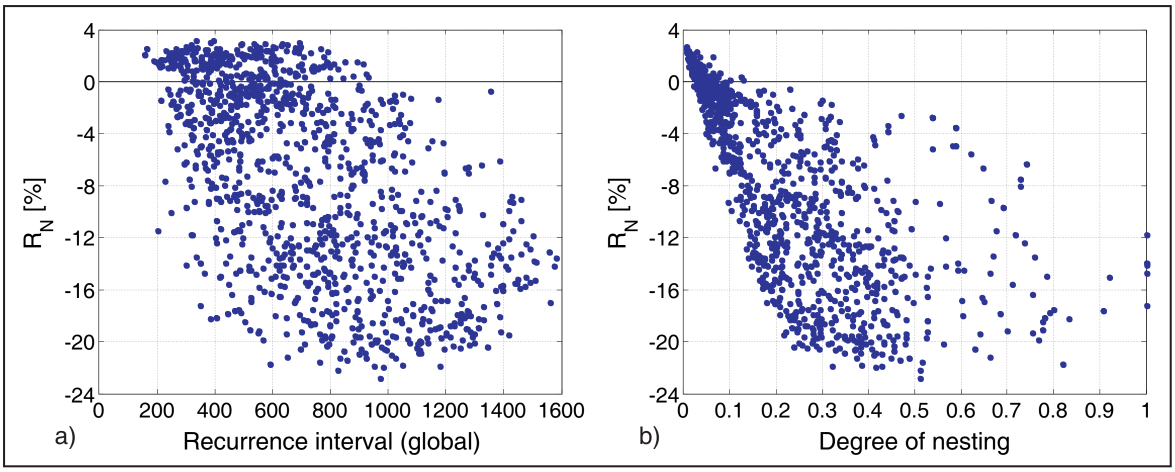

Fig. 3.7: Ratio of recurrence interval R_N vs. the recurrence interval estimated by a global cross-correlation function T_G (a) and degree of nesting D_N (b).

R_N decreases with a higher degree of nesting (Fig. 3.7b). There is a particularly strong decrease of R_N for D_N between 0 and 0.2. This implies that even a small degree of nesting affects the recurrence interval of the PREC appreciably. However, a certain degree of nesting is required to estimate large differences between the recurrence intervals for the global T_G and the nested-unnested approach T_N. For example, R_N<-10 is observed for D_N>0.1. R_N decreases up to a degree of nesting of about 0.4.

Positive values of R_N are observed for a degree of nesting lower than 0.15 and therefore for pooling groups without or with only a few nested catchments. In the most extreme case (D_N=0), the parameter set for unnested catchment relationships is always used in the nested-unnested approach. Figure 3.4 has shown that the cross-correlation function for unnested catchments leads to the smallest correlation values. Consequently the correlation among sites is lower if only unnested catchments were used compared with the global approach using all catchments. Hence, the lower correlation between unnested catchments leads to a higher recurrence interval (T_N>T_G) resulting in a positive value of R_N (see Eq. 3.7).

3.4.7 Different thresholds of the heterogeneity measure

The threshold of the heterogeneity measure was varied to investigate the effect on the formation of pooling groups and, in particular, on the recurrence interval of the PREC. The procedure for H_1<2 was repeated for H_1<1. A lower threshold of the heterogeneity measure leads to less pooling groups and thus to less realisations of PRECs.

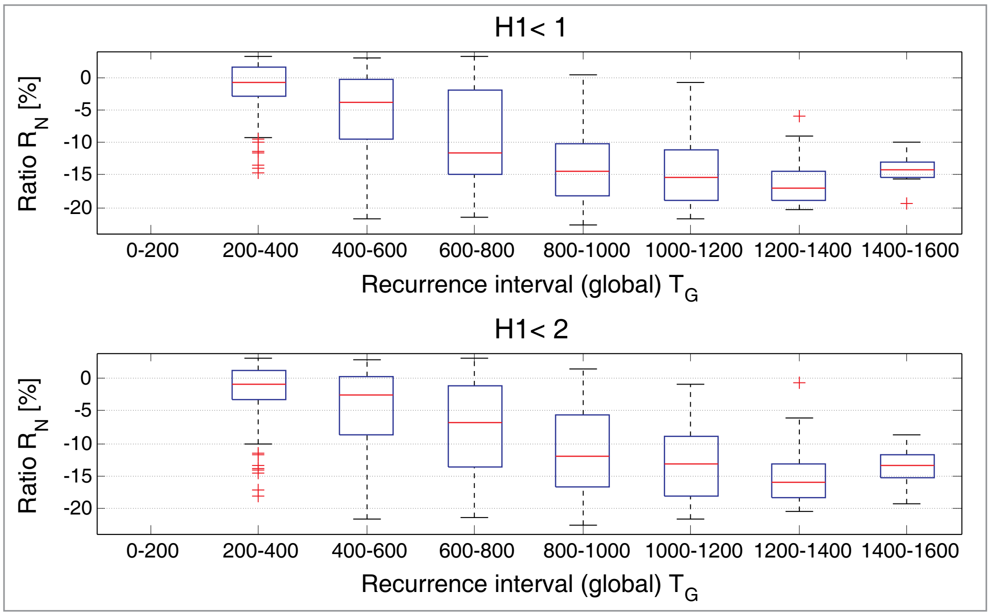

Fig. 3.8: Ratio of recurrence interval R_N vs. the recurrence interval estimated by a global cross-correlation function T_G for different thresholds of the heterogeneity measure.

The comparison of the ratio R_N with T_G reveals that R_N decreases with increasing T_G for both thresholds of the heterogeneity measure (Fig. 3.8). This relationship is illustrated in equidistant intervals of the recurrence intervals. We only show the results for cases with more than ten pooling groups within the equidistant interval. The number of pooling groups decreases for a lower threshold because of the stricter homogeneity criterion. While different thresholds for the heterogeneity measure lead to a different number of pooling groups, and therefore affect the results, they do not influence the general statements.

3.5 Discussion

Our goal was to estimate the impact of the intersite dependence of nested catchment structures on the effective sample years of data and the recurrence interval of PREC. Therefore, nested and unnested pairs of catchments were treated with separate parameter sets for the cross-correlation function. This enabled us to compare the nested approach with the traditional one using a global cross-correlation function.

There are three interesting aspects to discuss. These are (1) the differences between the three cross-correlation function applications (global, nested, unnested); (2) the link of a differentiation in nested and unnested catchments to different hydrologic situations, and (3) the impact of the two different approaches of the cross-correlation function (global vs. nested-unnested) on the effective sample years of data and the recurrence interval derived by PRECs and its relevance for this study area.

The heterogeneity in the correlation matrix (see Fig. 3.3) leads to a scattering of the empirical correlation coefficients in relationships to the distance (see Fig. 3.4). Madsen and Rosbjerg (1997a) determined a scattering due to the heterogeneity of the region. Whereas we separated the catchment relationships into nested and unnested catchment relationships, Madsen and Rosbjerg (1997a) divided the study area in two regions and estimated separate regional correlation functions for each region. By doing so, they estimated that a separate consideration of two regions led to a larger average intersite correlation than an overall approach. In our study, we also estimated larger intersite correlations for the nested-unnested approach than for the global approach for most of the pooling groups. A lower intersite correlation is found for the nested-unnested approach in specific cases (e.g. a low degree of nesting)

Figure 3.4 demonstrates that there is a large decrease of the cross-correlation functions for the unnested and global approaches up to a distance of about 50 km. Merz and Blöschl (2003) assumed that catchments whose centroids have a distance less than 50 km are frequently affected by the same event, resulting in a relatively large correlation between their flood sequences. Catchments with larger distances are affected by different events, and consequently the discharge time series are less correlated.

The relevance of separate parameter sets for the cross-correlation function for nested and unnested catchments depends on the spatial extent of floods and consequently on the prevailing flood regime. Large-scale precipitation events may lead to larger intersite correlations than local convective rainfalls. In regions that are mainly influenced by long precipitation events, widespread floods may occur at neighbouring gauges across catchment boundaries, independently of the catchment structure. In this case, the gauges might be correlated beyond catchment boundaries (Merz and Blöschl, 2003), and it is expected that there are only limited differences between the correlation relationships within and across catchment boundaries. It is assumed that especially large floods across wide areas lead to a large correlation between catchments (Hosking and Wallis, 1988), implying that high flood quantiles are affected stronger by intersite correlation. This statement coincides with the decrease in the regional information content with increasing sample years of data (see Fig. 3.5).

An opposite situation is given for flood regimes that are dominated by local convective precipitation events with small spatial extent. A local precipitation event might evoke a flash flood only along the river. Then, only a few catchments, in particular nested catchments, are affected by the same flash flood and low correlation relationships across catchment boundaries are expected. In this case, the impact of a separation in nested and unnested catchment relationships might be strong.

In our study area, Saxony, both local floods (e.g. in 1927, 1957) as well as regional widespread floods (e.g. in 1954, 1958, 2002) occurred in the past (e.g. Pohl, 2004, Petrow et al., 2007). The rivers of the *Erzgebirge*, specifically the headwaters of the Mulde river and in particular the western tributaries of the Elbe river, were affected by flash floods (e.g. Ulbrich et al., 2003), which in Saxony occur mostly in July and August. These floods produce the highest unit flood discharges in the study area. Due to the fast catchment response in the *Erzgebirge*, downstream gauges are directly affected by flash floods. In this context it is necessary to mention that the western tributaries of the Elbe are relatively small tributaries with only up to three gauges, whereas there are several nested relationships among the gauges of the Mulde catchment (see Fig. 3.2).

Since no gauges located at the River Elbe, the largest river in the study area, are included in the analysis, the differences between the catchment sizes of nested catchments are not too large. This aspect is especially important for this study area, since most of the largest floods occurred in the western tributaries of the River Elbe. These rivers flow into the Elbe upstream of the gauge Dresden. Because of their relatively small catchment sizes (<200 km²) in comparison to the Dresden gauge (52.000 km²), it is not expected that the mean discharge at gauge Dresden is significantly influenced by a local flood in one of the western tributaries only.

In the study area, there are only 5% of pairs of nested catchments. As expected, this study has shown significantly larger correlation among nested catchments than unnested ones. The effect of a distinction in nested and unnested cross-correlation functions might be even larger in regions with a larger number of nested catchment relationships. However, 5% of pairs of nested catchments lead to a significant reduction in the recurrence interval of PREC.

In this study, only one specific point of the PREC method is assessed. It is clear that the recurrence interval of PREC is affected by all steps of the PREC method. However, to determine the influence of one particular step in the PREC concept, it is necessary that all other aspects are constant. This was realised in this study by emphasising the selection of the parameter sets for the cross-correlation function in a hydrologically more comprehensive way.

The introduction of the nested structure to the PREC concept results in a reduction of the recurrence interval of up to 23% (see Fig. 3.7). Therefore, it is recommended to use different cross-correlation functions for nested and unnested catchments, in particularly for pooling groups with a large degree of nesting (see Fig. 3.7b). In this study, there is a relevant effect for a degree of nesting larger than 0.15.

3.6 Conclusions

This study focused on the modelling of intersite dependence when estimating the recurrence interval of a probabilistic regional envelope curve (PREC). A correct representation of the intersite dependence is fundamental for quantifying the regional information content of a pooling group, and therefore also for identifying the effective sample years of data, which is a key step of the PREC concept. The regional information content is defined as the ratio between the effective sample years of data (i.e. equivalent number of independent observations) and the overall sample years of data in the regional sample.

The analysis clearly shows that the intersite correlation for nested pairs of catchments is significantly larger than for unnested pairs, suggesting separate cross-correlation functions for nested and unnested pairs of catchments. A separation into nested and unnested pairs of catchments while modelling the intersite dependence represents an innovation and a refinement of the existing approach.

The study adopts a cross-correlation function whose parameters are identified for the whole study area (traditional approach) as well as differentiated between nested and unnested catchment pairs (proposed approach). The main outcomes can be summarised as follows:

1. the differentiation in cross-correlation functions for nested and unnested pairs of catchment enables one to improve the estimates of the number of effective observations;
2. in most of the cases, the number of effective observations and, therefore, the recurrence interval of PREC, are reduced by modelling the intersite dependence for

pairs of nested and unnested catchments separately;

3. the reduction of the estimated recurrence interval increases with the size of the pooling group, or, evidently, with a higher degree of nesting in the pooling group of sites;
4. the results of the analysis are valid for different degrees of heterogeneity of the pooling group of sites. Defining the heterogeneity of the pooling groups in terms of H_1-values as proposed by Hosking and Wallis (1993), the study shows that the same considerations that are valid for possibly heterogeneous pooling groups of sites ($H_1<2$) still hold for acceptably homogeneous groups ($H_1<1$).

Because of the effect of nested catchment structures on the recurrence interval of PREC, we recommend to apply different cross-correlation functions for nested and unnested catchments in PREC studies. Our study points out that the effect of nested structure becomes relevant for regions in which the number of nested pairs of catchments is larger than 15% of the total number of pairs. Separate cross-correlation functions reflect the characteristic catchment structure and incorporate this structure in the estimation of flood quantiles.

Acknowledgements

We thank the three anonymous referees for their very useful comments and their helpful recommendations to our manuscript in HESSD.

This work is part of the Center for Disaster Management and Risk Reduction Technology (CEDIM) (www.cedim.de), a joint venture between the Helmholtz Centre Potsdam - GFZ German Research Centre for Geosciences, the Technical University of Karlsruhe (TH) and the Forschungszentrum Karlsruhe. We thank CEDIM and the GFZ and the University of Bologna for the financial support. Attilio Castellarin acknowledges the Italian Ministry of Education for the grant no. 2007HBTS85_003.

We thank the State Agency of Environment and Geology of the Free State of Saxony for the permission to use the discharge data and the hydrogeological map (HÜK 200) (http://www.umwelt.sachsen.de/umwelt/geologie/7703.htm, 05.02.2007). Furthermore, we thank the State Agency of Flood Prevention and Water Management of Saxony-Anhalt and the State Agency for Environment of Thuringia for additional discharge data. We also thank the Potsdam Institute for Climate Impact Research (PIK) and the German Weather Service (DWD) for precipitation data as well as the Federal Agency for Cartography and Geodesy of Germany (BKG) for the ATKIS-Basis-DLM and the digital elevation model for Saxony. We also thank the European Commission – Joint Research Centre (JRC) for the CCM River and Catchment Database. The SRTM Digital Terrain Model was downloaded from (http://srtm.csi.cgiar.org/SELECTION/inputCoord.asp, 19.05.2008).

CHAPTER 4

INTRODUCING EMPIRICAL AND PROBABILISTIC REGIONAL ENVELOPE CURVES INTO A MIXED BOUNDED DISTRIBUTION FUNCTION

4 Introducing empirical and probabilistic regional envelope curves into a mixed bounded distribution function

Abstract

A novel approach to consider additional spatial information in flood frequency analyses, especially for the estimation of discharges with recurrence intervals larger than 100 years, is presented. For this purpose, large flood quantiles, i.e. pairs of a discharge and its corresponding recurrence interval, as well as an upper bound discharge, are combined within a mixed bounded distribution function. Large flood quantiles are derived using probabilistic regional envelope curves (PRECs) for all sites of a pooling group. These PREC flood quantiles are introduced into an at-site flood frequency analysis by assuming that they are representative for the range of recurrence intervals which is covered by PREC flood quantiles. For recurrence intervals above a certain inflection point, a Generalised Extreme Value (GEV) distribution function with a positive shape parameter is used. This GEV asymptotically approaches an upper bound derived from an empirical envelope curve. The resulting mixed distribution function is composed of two distribution functions, which are connected at the inflection point.

This method is applied to 83 streamflow gauges in Saxony/Germany. Our analysis illustrates that the presented mixed bounded distribution function adequately considers PREC flood quantiles as well as an upper bound discharge. The introduction of both into an at-site flood frequency analysis improves the quantile estimation. A sensitivity analysis reveals that, for the target recurrence interval of 1000 years, the flood quantile estimation is less sensitive to the selection of an empirical envelope curve than to the selection of PREC discharges and of the inflection point between the mixed bounded distribution function.

Published as:

Guse, B., Hofherr, T., and Merz, B., 2010. Introducing empirical and probabilistic regional envelope curves into a mixed bounded distribution function, Hydrol. Earth Syst. Sci. Discuss., 7, 4253-4290 (in review in Hydrol. Earth Syst. Sci.).

4.1 Introduction

Flood frequency analysis provides flood quantiles, i.e. discharges and their corresponding recurrence intervals. Especially for recurrence intervals T >100 years, flood quantile estimates are very uncertain, due to the limited length of the measured flood series and the low number of representative data for extreme floods (e.g., Cohn and Stedinger, 1987; Merz and Thieken, 2005; Reis Jr. and Stedinger, 2005).

To reduce the estimation uncertainty of an at-site flood frequency analysis, it is recommended to use more information than the observed flood series (e.g. Hosking and Wallis, 1986a; Stedinger and Cohn, 1986; Merz and Blöschl, 2008a,b; Merz and Thieken, 2009). Since the quantile estimates become less precise with higher recurrence intervals, additional information becomes increasingly important in these cases (e.g., Hosking and Wallis, 1986a). Additional information can be classified into three groups: causal, temporal (historic floods) and spatial (flood regionalisation) information (Merz and Blöschl, 2008a,b). First, process understanding can be incorporated as causal information into a flood frequency analysis. For example, Merz and Blöschl (2008a) illustrated that an investigation of event runoff coefficients helps to explain the generation processes of extreme floods and therefore to describe the upper tail behaviour of a distribution function.

Second, systematic time series can be extended by integrating historic floods as non-systematic data (Stedinger and Cohn, 1986). These historic extreme floods lead to more data for the estimation of large quantiles (e.g., England Jr. et al., 2003b; Benito et al., 2004). Historic observations contain considerable measurement errors, but due to the short systematic observation period, such additional information is useful (e.g., Hosking and Wallis, 1986b), and an increase of the effective record length leads to a better estimation of flood quantiles (Condie and Lee, 1982; Stedinger and Cohn, 1986; Cohn and Stedinger, 1987).

Third, flood regionalisation aims at improving flood quantile estimates by using information from gauges with similar hydrologic characteristics. In this way, the limited length of flood series is compensated by using regional flood series, following the principle of ‘trading space for time’ (Stedinger et al., 1993). Gutknecht et al. (2006) proposed to combine local and regional methods within a ‘multi-pillar’-approach to reduce the uncertainty of flood quantile estimates for large recurrence intervals.

The selection of a distribution function which is suitable to estimate extreme floods is difficult (e.g., Merz and Thieken, 2005; El Adlouni et al., 2008). Parameter estimation methods mostly concentrate on the central parts of the distribution function. The upper tail which is the most relevant for extreme events and is subject to the largest uncertainty is often not adequately described (Moon et al., 1993). Hence, for the estimation of large flood quantiles, it is recommended to concentrate on extreme floods and to derive as much information as possible from them (Naghettini et al., 1996).

Hydrological characteristics, e.g. generation mechanisms of extreme floods, might be different compared to those of high-frequency floods (e.g., Chbab et al., 2006; Gutknecht et al., 2006; Merz and Blöschl, 2008b). Therefore, the use of a single distribution function to represent the flood behaviour across the complete spectrum of recurrence intervals is critical (England Jr. et al., 2003a), which is why, mixed distribution functions are recommended. The two-component extreme value (TCEV) distribution (Rossi et al., 1984) includes two different distribution functions for normal and extreme events, respectively (e.g., Francés, 1998; Fernandes and Naghettini, 2008). The idea of mixed distribution functions is also the basis of the gradex approach (Guillot and Duband, 1967), in which the traditional flood frequency curve is used up to a recurrence interval, at which the corresponding discharge leads to catchment saturation. Above that threshold, the flood frequency curve follows the rainfall frequency curve, assuming that the rainfall records are longer and more precise than flood series (e.g., Naghettini et al., 1996; Gutknecht et al., 2006; Merz et al., 2008).

Traditional distribution functions with three parameters, such as the Generalised Extreme Value (GEV) or General Logistic (GL), are unbounded or only bounded in specific cases (e.g. GEV with a shape parameter $k > 0$). This implies that the increase of the frequency curve is unlimited and that a non-zero exceedance probability for unrealistic large flood discharges is estimated (Enzel et al., 1993).

Distribution functions were developed which asymptotically approach an upper bound (e.g. the extreme value distribution with four parameters (EV4) Kanda, 1981; Francés and Botero, 2003). Francés and Botero (2003) combined non-systematic and systematic data with a bounded distribution function in their application of the EV4.

Upper bound discharges can be derived, on the one hand, by estimating a probable maximum flood (PMF). To estimate a PMF, a probable maximum precipitation (PMP) is transformed into a PMF. Therefore, the most extreme meteorological and hydrological conditions for a given region are derived (e.g., Costa, 1987; Houghton-Carr, 1999; Fernandes et al., 2010). On the other hand, envelope curves provide upper bound discharges. Envelope curves bound all regional unit floods of record, i.e. the maximum unit flood discharges, by relating them to their catchment sizes. The method of empirical envelope curves (ECs) is a simple method which is not based on physical assumptions (Crippen, 1982). ECs are traditionally constructed for an administrative region (e.g., China and USA, Costa, 1987, Europe and the World, Herschy, 2002). Merz and Thieken (2009) enlarged the European data set of Stanescu (2002) by German floods of record from the last years and derived an EC which was used as additional information to constraint the selection of distribution functions.

Castellarin et al. (2005) and Castellarin (2007) extended the traditional method of envelope curves. They introduced the method of probabilistic regional envelope curves (PREC) which provides a large flood quantile, i.e. a pair consisting of a PREC discharge and its corresponding recurrence interval, for each gauge of a homogeneous pooling group of sites. In contrast to empirical envelope curves, probabilistic regional envelope curves (PREC) assign a non-zero exceedance probability to the regional envelope curve.

This study aims at improving flood frequency estimates for large recurrence intervals T by using additional information provided by empirical and probabilistic regional envelope curves. Since this study aims at integrating both, a distribution function needs to be selected which considers an upper bound discharge as well as large flood quantiles derived from PRECs. By doing so, for the first time, PREC flood quantiles are inserted into a flood frequency curve.

This study is structured as follows: In Sect. 4.2, study area, Saxony/ Germany, and data are presented. The methods of empirical envelope curves and probabilistic regional envelope curves are briefly explained in Sect. 4.3. Here, we also present the results of previous studies, in which PREC flood quantiles were derived for Saxon gauges (Guse et al., 2009, 2010). The novel method to improve the flood frequency estimates is described in Sect. 4.4. It is explained how large flood quantiles and an upper bound discharge can be introduced into a suitable distribution function. In Sect. 4.5, we show the results of our method and evaluate the sensitivity of relevant choices when estimating discharges with the presented mixed bounded distribution for a target T of 1000 years.

4.2 Study area and data

The study area is the federal state of Saxony which is located in South-Eastern Germany. The south-western part is covered by the mountain range of the Erzgebirge, which has the largest altitudes in Saxony (Fig. 4.1). The Elbe is the largest river in the investigation area.

The largest unit floods of record were observed at the western tributaries of the River Elbe coming from the Erzgebirge (e.g. gauges 9 and 15 in Fig. 4.1) and at a tributary of the Lausitzer Neisse (gauges 82 and 83). In the observation period, both local and regional floods are included which affected in particular the Erzgebirge (Pohl, 2004). Extreme floods in Saxony belong to two flood types: Small tributaries in the mountain range of the Erzgebirge are affected by flash floods, while, riverine floods along the River Elbe are characterised by a slow rise of the water level (Ulbrich et al., 2003; Petrow et al., 2006). An extreme event in 2002 led to severe flood damages at almost all tributaries originating in the Erzgebirge and along the rivers Elbe and Mulde (e.g., Ulbrich et al., 2003; Thieken et al., 2005). Particularly due to this flood, several Saxon flood time series are very skewed (Petrow et al., 2007). The 2002 flood led to large modifications of the frequency curve and especially of the shape parameter at several gauges in Saxony (Schumann, 2004, 2005), and revealed the uncertainty of at-site flood frequency estimates without additional information. This confirmed the need for representative extreme events within the data series.

The discharge gauges are distributed along all relevant rivers and tributaries in the investigation area. We used 83 gauges, including two from Thuringia (gauges 61 and 62). We selected gauges with observation periods >29 years and catchment sizes >10 km² and without large effects due to mining activities or dams. The annual maxima series (AMS) as well as the maximum observed discharge, i.e. the flood of record, were derived for all 83 gauges.

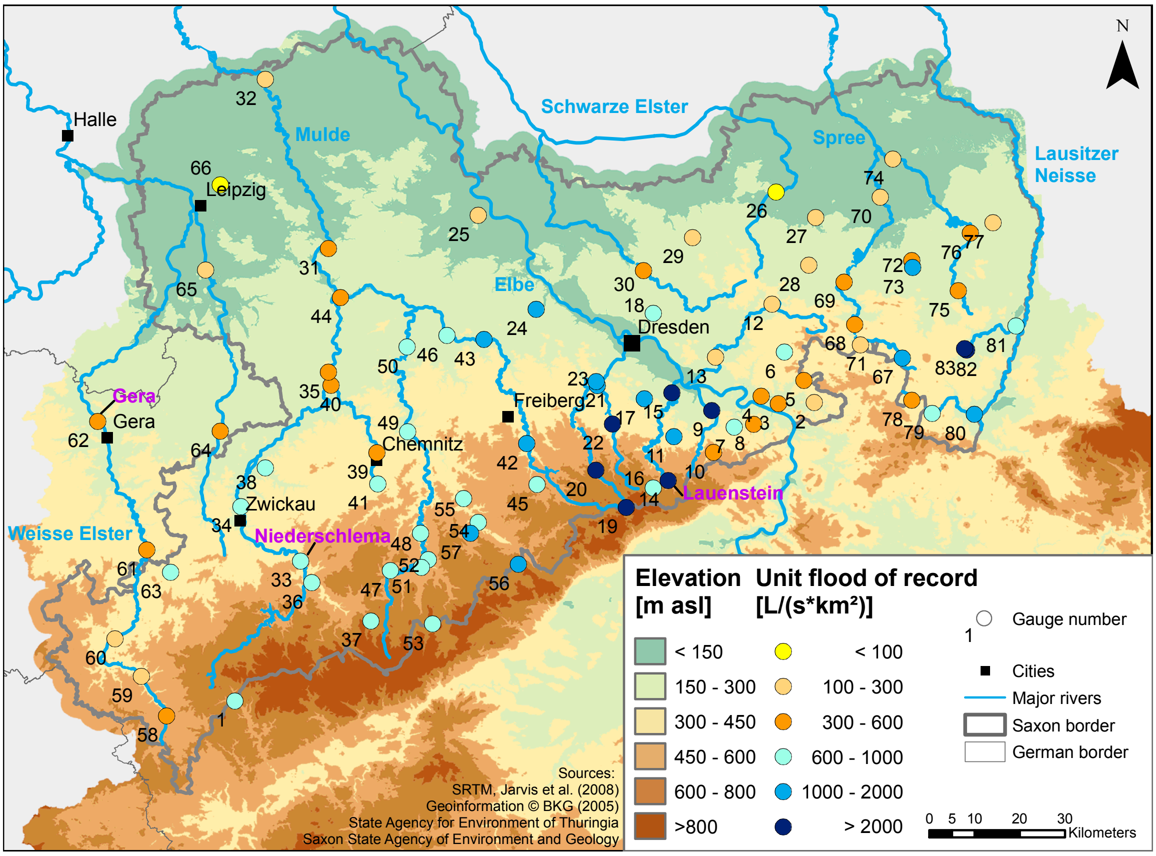

Fig. 4.1: Study region (Saxony/Germany) and selected discharge gauges coloured by their unit floods of record (modified from Guse et al., 2009). The three gauges which were used in the application (see Sect. 4.5) are named in purple.

4.3 Envelope curves

We used upper bound discharges derived from empirical envelope curves (ECs) and large flood quantiles provided by probabilistic regional envelope curves (PRECs). Both methods are briefly introduced. Envelope curves bound the observed floods of record of regional sites. Therefore, the floods of record Q_{FOR} are normalised by their catchment size A and then related to A in a double-logarithmic plot. Envelope curves are determined by their slope b and intercept a (Eq. (4.1), adapted from Castellarin et al., 2005).

$$\log\left(\frac{Q_{FOR}}{A}\right) = a + b * \log(A) \tag{4.1}$$

4.3.1 Empirical envelope curves

Three empirical envelope curves were constructed (Fig. 4.2). First, an envelope curve based on the Saxon floods of record only was derived. Second, the envelope curve for Germany EC_G from Stanescu (2002) was selected. Third, the European envelope curve EC_E of Herschy (2002) was used.

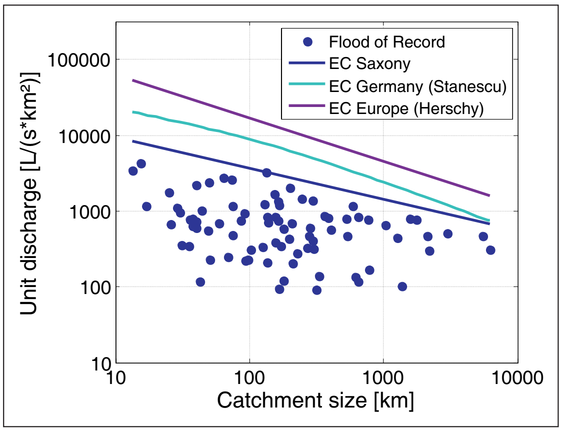

Fig. 4.2: Comparison of three different envelope curves. The floods of record of Saxon gauges are additionally shown.

In this study, an upper bound with an exceedance probability of zero for Saxony needs to be considered. The Saxon envelope curve was determined by the largest unit flood of record in Saxony. The floods of record of several gauges are close to this EC. Thus, it is inconsistent to assume that the Saxon envelope curve has an exceedance probability of zero with respect to T_{PREC} between 150 and 1500 years which were estimated by PRECs for this study region in Guse et al. (2010) (see Sect. 4.3.4). For a few gauging stations, the discharges provided from PRECs were close to or even larger than the Stanescu envelope curve for Germany. Since it was advisable to take an envelope curve which is certain to be the upper bound of Saxon flood discharges, we used the European envelope curve by Herschy (2002). This envelope curve is expected to be an upper bound which might not be exceeded in Saxony, since it is determined by significantly larger floods from the Mediterranean region. Stanescu (2002) and recently Gaume et al. (2009) compared ECs of European countries and determined the largest magnitude for Mediterranean countries. Stanescu (2002) concluded that larger floods are possible around the Mediterranean Sea than in Central European countries, owing to the warmer temperature and resulting larger humidity contained in the air masses. The Stanescu envelope curve was used only to investigate the sensitivity of the selection of the empirical envelope curve (see Sect. 4.4.3).

4.3.2 Probabilistic regional envelope curves

Probabilistic regional envelope curves (PRECs) (Castellarin et al., 2005; Castellarin, 2007) estimate an exceedance probability for a regional envelope curve (REC). PRECs can be derived for homogeneous regions as indicated in the index flood method (Dalrymple, 1960; Robson and Reed, 1999). In the case of regional homogeneity, the index flood (mean of the annual maxima series) is a function of the catchment size. The slope b of REC (Eq. (4.1)) is determined by a regression through all index flood values of the pooling group (Fig. 4.3). The intercept a is estimated by shifting the regression line up to the largest unit flood of record. Hence, the intercept a of REC is determined by the largest unit flood of record in the pooling group (Castellarin et al., 2005).

To estimate the exceedance probability of REC, the overall sample years of all regional annual maxima series (AMS) are reduced to the effective sample years of data. The intersite dependence among the AMS is examined by considering the reduction of the regional information content due to cross-correlated sites. Castellarin (2007) presented an empirical relationship for this case. The cross-correlation function of Tasker and Stedinger (1989) was used, which describes the decrease of the cross-correlation between the AMS with increasing distance between the catchment centroids. Because of the higher correlation among nested pairs of catchments, different parameter sets for nested and unnested pairs of catchments are used, as proposed by Guse et al. (2009), instead of the initial approach with one parameter set for all pairs of catchments.

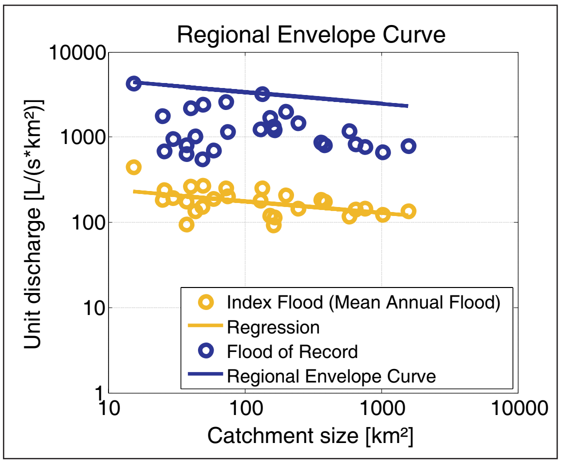

Fig. 4.3: Example of Regional Envelope Curve (REC) (from Guse et al., 2010).

The exceedance probability is calculated for the pair of the unit flood of record and its corresponding catchment size, which governs the REC (Castellarin, 2007). The PREC provides a discharge Q_{PREC} for each gauge of the pooling group with the same recurrence interval T_{PREC}.

4.3.3 Application of probabilistic regional envelope curves in Saxony

In previous studies, several PRECs were derived for Saxony (Guse et al., 2009, 2010). A major step in the PREC concept is the determination of the pooling group of sites. Guse et al. (2010) used cluster analysis and the Region of Influence (RoI) approach (Burn, 1990b) to construct several pooling groups using twenty candidate sets of two or three catchment descriptors. An own PREC was constructed for each pooling group, which fulfils the homogeneity criteria of the heterogeneity measure ($H_1 < 2$) of Hosking and Wallis (1993). Hence, the constitution of the homogeneous regions and thus PRECs differed depending on the grouping procedure.

The suitability of both pooling methods to derive PREC flood quantiles was assessed by comparing the PREC method with the index flood method. To this end, a leave-one-out jackknifing approach was used to calculate the PREC flood quantiles for ungauged conditions, denoted as $Q_{PREC-JK}(T_{PREC-JK})$ (Castellarin, 2007; Castellarin et al., 2007; Guse et al., 2010). The relative error between $Q_{PREC-JK}$ and $Q_{IF,}$ the estimated discharge for $T_{PREC-JK}$ with the index flood method, was estimated for each gauge of the pooling group. The comparison of the relative errors for cluster analysis and RoI showed that both pooling methods lead to similar performance (Guse et al., 2010). Therefore, PREC flood quantiles of both pooling methods were used. In this study, PREC flood quantiles with a relative error < 2 were used only. By doing so, PREC realisations that deviated strongly from the index flood method were not considered. This means that PREC flood quantiles of a site which were more than three times larger for ungauged conditions than the index flood estimates for the same T_{PREC} were excluded.

The number of PREC realisations varied among the gauges between 0 and 127. A site had a lower number of PREC flood quantiles when it belonged more often to heterogeneous regions due to the specific characteristics of this gauge. Of the 89 gauges available in the previous studies, only the 83 gauges with at least one PREC realisation were used for this study (see Fig. 4.1). In the previous study, T_{PREC} varied between 150 and 1500 years with a mean value of 650 years (Guse et al., 2009).

4.3.4 Comparison of empirical and probabilistic regional envelope curves

When comparing the traditional empirical envelope curves with the probabilistic regional envelope curves, one has to take note of the differences between the two approaches.

Several studies have illustrated the slope values of empirical envelope curves. On average, a slope of -0.5 is estimated with a variability between -0.2 and -0.7 (e.g., Herschy, 2002; Castellarin et al., 2005; Castellarin, 2007; Gaume et al., 2009). In our study, the slopes of the empirical envelope curves are close to -0.4. In contrast, the slope in the PREC approach has a lower negative value. Here, the slope b is about -0.2. This means that the effect of the catchment size is smaller in the PREC concept.

Since the intercept of the empirical envelope curve is larger than those of the PREC realisations in this study, it follows that the discharge of EC is larger than in the PREC concept. This result is understandable given that the EC has an exceedance probability of zero, while that of the PREC lies between $6.7*10^{-4}$ and $6.7*10^{-3}$ for this study region.

The PREC discharges should be lower than the upper bound discharge from EC in all cases. Hence, the consistency of PREC discharges was checked for all sites of each PREC realisation. Since the slopes of the PRECs are in the majority of the cases smaller than those of the ECs, PRECs approach the ECs with increasing catchment size. PREC discharges which were larger than the upper bound derived by the Stanescu envelope curve sites were removed. These cases were detected for sites with a large catchment size. It is assumed that the estimation of the empirical envelope curve was better than those of PREC in these cases with a large catchment size. In this way, consistency among both methods was ensured.

4.4 Methods

This study aims at inserting large flood quantiles and upper bound discharges as additional information into a distribution function to improve the flood quantile estimates for T >100 years. For this purpose, a distribution function is requested, into which large flood quantiles derived by PRECs, i.e. Q_{PREC} and corresponding T_{PREC}, as well as an upper bound discharge Q_{MAX}, provided by an empirical envelope curve, can be integrated. The method consists of two steps:

(1) Integration of the PREC flood quantiles into the observed flood series (Sect. 4.4.1)

(2) Application of a mixed bounded distribution function including PREC flood quantiles and an empirical envelope curve discharge as upper bound (Sect. 4.4.2)

Figure 4.4 gives an overview about our approach, including the most relevant variables. The core idea is an improvement of discharge estimates for a target recurrence interval T_t of 1000 years (orange line in Fig. 4.4). As additional information, PREC flood quantiles with recurrence intervals between 150 (lower value T_l) and 1500 (upper value T_u) years are used (dashed cyan lines) and combined with the observed flood series in a distribution function ($GEV_{sim\text{-}prec}$). As second additional information, an upper bound discharge (Q_{MAX}) (purple line) derived from an empirical envelope curve is integrated into a distribution function. The resulting mixed bounded distribution (GEV_{bound}) consists of two distribution functions, connected at the inflection point (T_X) (dashed magenta line) and approaching the upper bound (Q_{MAX}) asymptotically. The mixed distribution function is identical with $GEV_{sim\text{-}prec}$ up to the inflection point. From this point on, the bounded GEV is used.

4.4.1 Integration of PREC flood quantiles

In the first step, PREC flood quantiles were combined with the observed AMS. In a traditional regional flood frequency analysis, flood data from the site itself and from neighbouring sites are available. Since a PREC flood quantile comprises of a Q_{PREC} and its corresponding T_{PREC}, it was impossible to add a Q_{PREC} value directly to the AMS as one additional flood value. The additional information of the corresponding T_{PREC} needs to be considered to use the maximum information from PRECs. Hence, a novel method was developed.

The Generalised Extreme Value (GEV) distribution was fitted to the observed AMS of each gauge using L-moments (Hosking and Wallis, 1997), denoted as GEV_{obs}. The adequacy of the GEV for the flood series in this study was proven by L-moment ratio diagrams (see e.g., Vogel and Fennessey, 1993; Peel et al., 2001).

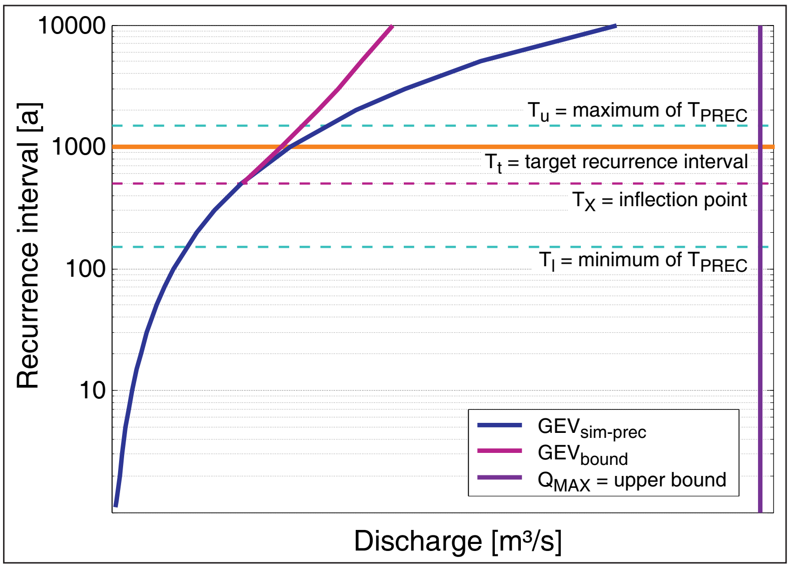

Fig. 4.4: Scheme of the proposed method including the most relevant variable names. The upper bound is illustrated in purple right of the legend. $GEV_{sim\text{-}prec}$ is the combined distribution function of the observed flood series and the PREC flood quantiles. GEV_{bound} is a bounded distribution function which includes PREC flood quantiles as well as an upper bound discharge.

The three at-site GEV_{obs} parameters (ξ, α, k) were used to generate synthetic flood series. For this, T_u random numbers between 0 and 1 (p_{sim}) were generated. T_u was selected, since it was the maximum of T_{PREC} for the study region. These p_{sim} values were inserted into the GEV (Eq. 4.2) resulting in T_u simulated discharge values, denoted as Q.

$$Q = \xi + \frac{\alpha}{k} * \left[1 - (-\ln(p_{sim}))^k\right] \text{ with } k \neq 0 \qquad (4.2)$$

Subsequently, the GEV was fitted to Q, denoted as GEV_{sim} with a new parameter set (ξ_{sim}, α_{sim}, k_{sim}).

To ensure consistency between GEV_{sim} and GEV_{obs}, the two should not differ considerably. For this, the flood quantiles for $T = T_u$ years of both GEV functions were compared. It was decided that the discharge estimates of both functions should not vary more than 1% for T_u. If $Q_{sim}(T_u)$ varied more than 1% from $Q_{obs}(T_u)$, the random selection of p_{sim} and the estimation of Q were repeated.

A second constraint was that there had to be nine or ten values, denoted as n_x, larger than $p_E = 0.9933\left(= 1 - \frac{1}{150}\right)$. This value was selected, because the T_{PREC} values were larger than 150 years ($T_l = 150$). It was therefore assumed that the PREC flood quantiles were representative for $T > T_l$ years. A binomial function showed that the largest probability was estimated when assuming that nine or ten floods with $T > T_l$ were expected to occur within T_u years. This constraint was considered to prevent an influence of a randomly selected number of PREC flood quantiles. Then, GEV_{sim} and GEV_{obs} were assumed as sufficiently similar for using the T_u simulated flood series instead of the shorter measured time series.

In a next step, PREC flood quantiles were integrated into the simulated flood series Q_{sim}. The random numbers p_{sim} were sorted in increasing order. Among $p_{sim,}$ the n_x values larger than p_E were removed from the simulated flood series Q_{sim} and replaced by n_x Q_{PREC} values.

This approach implicitly assumed that the observed flood series is appropriate up to T_l. However, the PREC discharges also influenced the combined function of observed and PREC discharges for T <T_l.

Since the previous studies provided more than n_x PREC flood quantiles for most of the gauges (see Sect. 3.3) (Guse et al., 2010), it was necessary to select n_x PREC flood quantiles among the PREC realisations of a given gauge. The n_x PREC flood quantiles were selected in a random process whereas the discharges were weighted according to their T_{PREC}. We considered the recurrence intervals using a binomial function B (Eq. 4.3). This approach was used to estimate the mean occurrence of a specific Q_{PREC} with a recurrence interval T_{PREC} within T_u years.

$$P(X = m) = B_{T_u;\frac{1}{T_{PREC}}}(X = m) \qquad (4.3)$$

with m=1,2,…,20

We checked m for one to twenty occurrences. Among these twenty results, we selected the m with the largest probability P_{max}, i.e. the maximum likelihood, denoted as m_{max}. The Q_{PREC} of this PREC realisation was assigned m_{max} times to a vector V_{PREC}. This implies that PREC discharges with a smaller T were assigned more often to V_{PREC}. In this way, the recurrence interval of the PREC realisations was evidently considered, since a PREC flood quantile with a smaller T_{PREC} was expected to occur more often than a PREC flood quantile with a larger one. This procedure was repeated for all PREC realisations of this gauge.

The n_x PREC values were then randomly selected without replacement from V_{PREC}. In order to adequately represent T_{PREC}, a specific Q_{PREC} could be selected as many times as it was included in V_{PREC}. The n_x discharges derived from PREC were assigned to the reduced simulated flood series of T_u - n_x values, so that the new flood series comprised T_u values again.

In the majority of cases, the length of V_{PREC} was larger than n_x, which required the random selection of PREC discharges. In the other cases, for sites with a lower number of PREC realisations in V_{PREC} than n_x, n_x values were removed from the simulated flood series as well. Then all values from V_{PREC} were added. In order to obtain T_u values again, the remaining discharges to T_u were selected randomly from the n_x discharges with T >T_l years.

The GEV was fitted to the new flood series, denoted as $GEV_{sim\text{-}prec}$, using L-moments. This approach allowed an integration of PREC flood quantiles in flood frequency estimations. Due to the random process, there might be differences in the magnitude of the selected PREC discharges, and therefore also in the final distribution function. Hence, we repeated the selection of Q_{PREC} one hundred times and estimated one hundred GEV parameter sets.

The GEV parameter sets which estimated the median discharge for T_t were used for the next steps. The corresponding GEV distribution was denoted as $GEV_{sim\text{-}prec\ 50}$. The influence of the PREC selection on the discharge estimates was expressed by showing the 5%- and 95%-quantiles of $GEV_{sim\text{-}prec}$ for T_t, denoted as $GEV_{sim\text{-}prec\ 05}$ and $GEV_{sim\text{-}prec\ 95}$, respectively. A comparison of $GEV_{sim\text{-}prec}$ with GEV_{sim} illustrated the effect of using PREC flood quantiles as additional information.

4.4.2 Mixed bounded distribution function

We used a mixed bounded distribution function which was developed in storm research (Hofherr et al., 2008). The use of this distribution function enables us to integrate an upper bound discharge as further additional information besides of the PREC flood quantiles.

In this mixed bounded distribution function, flood quantiles up to a recurrence interval threshold of T_X (inflection point) are estimated by an unbounded distribution function (here: $GEV_{sim\text{-}prec}$ with $k<0$), and quantiles above the inflection point T_X are estimated by a bounded distribution (here: GEV_{bound}). A higher T_X was used, as it would be representative for the observed flood series only. $GEV_{sim\text{-}prec}$ includes the PREC discharges which were representative for T between 150 and 1500 years and this additional information enables us to use the higher T_X. To adequately represent the PREC discharges, we selected an inflection point T_X =500 years. The sensitivity of this inflection point was analysed in Sect. 4.4.3.

GEV_{bound} has a positive shape parameter k and, hence, asymptotically approaches an upper bound. The three parameters of GEV_{bound} (ξ_{bound}, α_{bound}, k_{bound}) were determined in an optimisation process by three constraints using Eqs. (4.4) - (4.6). First, the upper bound Q_{MAX} which was provided by an empirical envelope curve was inserted into the GEV upper bound function (Eq. 4.4).

$$Q_{MAX} = \xi_{bound} + \frac{\alpha_{bound}}{k_{bound}} \tag{4.4}$$

Second, both GEV functions ($GEV_{sim\text{-}prec}$, GEV_{bound}) had to be identical at the inflection point to avoid inconsistencies. Therefore, both functions were equated at the inflection point (Eq. 4.5).

$$GEV_{sim-prec}(T = T_x) = GEV_{bound}(T = T_x) \tag{4.5}$$

The third constraint was that both GEV functions had the same slope at the inflection point. Therefore, their derivates were equated (Eq. 4.6).

$$GEV_{sim-prec}'(T = T_x) = GEV_{bound}'(T = T_x) \tag{4.6}$$

In the case of a successful optimisation, GEV_{bound} was fully defined, increasing monotonically.

The mixed bounded distribution function was not applied for sites with a positive k of $GEV_{sim\text{-}prec}$. In these cases, the $GEV_{sim\text{-}prec}$ was already bounded. The main advantage of a bounded distribution function is that it avoids an unlimited increase up to unrealistic discharge values, which was already prevented by the positive k values in these cases.

4.4.3 Sensitivity analysis

The effect of three choices in this method was investigated for a target recurrence interval T_t = 1000 years in a combined sensitivity analysis. The sensitivity of each choice was tested as follows:

1. the magnitude of the empirical envelope curve discharge: German EC (EC_G) (Stanescu, 2002) vs. European EC (EC_E) (Herschy, 2002),
2. the selection of PREC discharges: 5%- vs. 95%- of the $GEV_{sim\text{-}prec}$ estimates for T_t,
3. and the magnitude of the recurrence interval threshold (inflection point): T_X = 200 vs. 500 years.

For each choice, the four possible combinations of the two other choices were checked. The comparison of $Q_{bound}(T_t = 1000)$ between all possible combinations of these three choices allowed us to evaluate their effect on the discharge estimations of GEV_{bound} for T_t. The relative deviations are calculated for each choice (Eqs. 4.7 - 4.9). This procedure enabled us to determine the most sensitive choice of the discharge estimates for T_t.

$$E_{EC} = \frac{Q_{bound}(Q_{MAX} = EC_E) - Q_{bound}(Q_{MAX} = EC_G)}{Q_{bound}(Q_{MAX} = EC_G)} \tag{4.7}$$

$$E_{PREC} = \frac{Q_{bound}(GEV_{sim-prec,95}) - Q_{bound}(GEV_{sim-prec,5})}{Q_{bound}(GEV_{sim-prec,5})} \tag{4.8}$$

$$E_{T_X} = \frac{Q_{bound}(T_X = 500) - Q_{bound}(T_X = 200)}{Q_{bound}(T_X = 200)} \tag{4.9}$$

4.5 Results

4.5.1 Integration of PREC flood quantiles

Figure 4.5 illustrates exemplarily for the gauge Lauenstein (site 14 in Fig. 4.1) that GEV_{sim} agrees well with GEV_{obs} (orange and black lines in Fig. 4.5). The blue-coloured circles symbolise the PREC discharges which were selected for $GEV_{sim\text{-}prec\ 50}$. Most of the $Q_{PREC}(T_{PREC})$ are smaller than the $Q_{GEV}(T_{PREC})$. Hence, the integration of the PREC flood quantiles leads to a higher k (shape parameter of GEV) and a lower skewness of $GEV_{sim\text{-}prec}$ compared to GEV_{sim}. Therefore, $Q_{sim\text{-}prec}$ for a given T is smaller than Q_{sim}.

The PREC flood quantiles indicate that the skewness of the GEV might be too large when using the observed data only. The recurrence interval of the flood of record (flood discharge of 2002) might be larger than the at-site estimate. The effect of the flood of record on the estimation of large quantiles within the at-site flood frequency analysis seems to be too high. The smallest PREC discharge is identical with the flood of record of Lauenstein. This means that the intercept of this REC was determined by the at-site flood of record.

The shape parameter k of $GEV_{sim\text{-}prec}$ was positive for seven sites. Since they already approach an upper bound, even after integrating PREC discharges, the number of sites for which the mixed bounded distribution function was applied was reduced to 76.

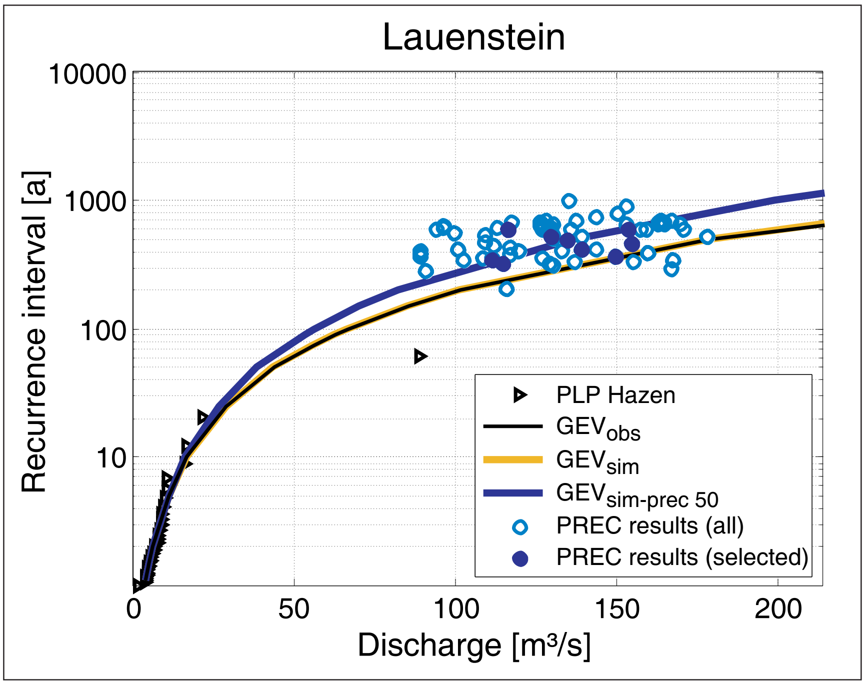

Fig. 4.5: Effect of integrating PREC flood quantiles into the at-site flood frequency analysis. GEV_{obs}, GEV_{sim} and $GEV_{sim\text{-}prec}$ are compared for the site Lauenstein. The observed flood series is illustrated as Hazen plotting position (PLP Hazen). The PREC flood quantiles which were selected for $GEV_{sim\text{-}prec\ 50}$ are coloured in blue.

4.5.2 Mixed bounded distribution function

$GEV_{sim\text{-}prec}$ was used to estimate the flood quantiles up to T_X = 500 years in the mixed bounded distribution approach. From T_X on, GEV_{bound} was used, which asymptotically approaches the upper bound discharge derived from the empirical envelope curve by Herschy (2002). Considering GEV_{sim} and GEV_{bound} for all gauges, three cases can be distinguished, which are shown in Fig 4.6a-c. The variability due to the selection of PREC flood quantiles is demonstrated by adding the 5%- and 95%-quantiles (cyan dashed line).

In the first case (gauge Lauenstein, Fig. 4.6a), GEV_{bound} estimates lower discharges than GEV_{sim} for all values of T. To give an example, GEV_{bound} estimates a discharge of 200 m³/s for T_t, whereas the GEV_{sim} discharge is about 300 m³/s. GEV_{sim} increases unlimitedly, whereas the gradient of GEV_{bound} decreases and approaches the upper bound.

Figure 4.6b shows an example (gauge Niederschlema, site 33 in Fig. 4.1) where several PREC discharges are larger than the GEV discharge estimates for the same recurrence interval. However, there are also various smaller PREC flood quantiles. On average, $Q_{PREC}(T_{PREC})$ is similar to $Q_{GEV}(T_{PREC})$, and therefore $Q_{sim\text{-}prec}$ is similar to Q_{sim}. The PREC flood quantiles support the GEV estimations, and the effect of the inclusion of PREC discharges is low.

In the third case, the PREC flood quantiles are larger than the GEV discharge estimates (gauge Gera in Fig. 4.6c, site 62 in Fig. 4.1). Here, Q_{bound} is about 1.5 times larger than Q_{sim}

for T_t. Despite the asymptotical approach towards the upper bound, Q_{bound} is larger than Q_{sim} even for T=10,000 years. There are gauges within the pooling groups of this site with significantly larger unit floods of record than those of Gera. The regional envelope curve has a considerably higher flood magnitude than the observed discharges. The PREC flood quantiles indicate that a flood larger than the current flood of record might occur.

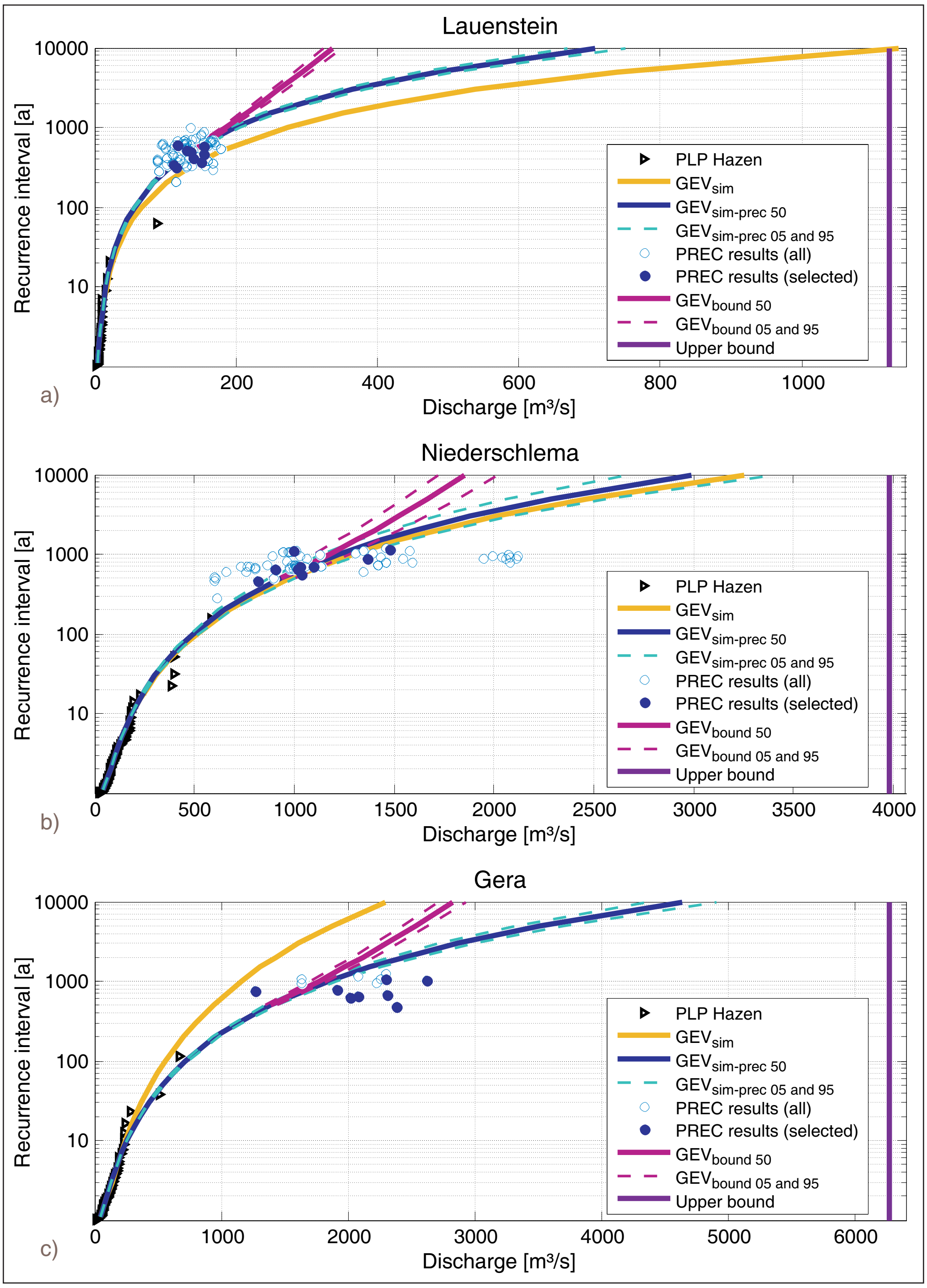

Fig. 4.6: The mixed bounded distribution function GEV_{bound} vs. the traditional GEV (GEV_{sim}) and the $GEV_{sim\text{-}prec}$ for the gauges (a) Lauenstein, (b) Niederschlema, (c) Gera. The blue-coloured PREC results show the selected PREC discharges which yielded a median discharge for the target recurrence interval of 1000 years among the hundred repetitions. The upper bound is illustrated in purple right of the legend.

4.5.3 Comparison of the three distribution functions

First, we compared GEV_{sim} and $GEV_{sim\text{-}prec}$. After that, we examined the differences between GEV_{sim} and GEV_{bound}. In both cases, discharge estimates for T_t were compared and we used the median of the hundred GEV estimations for $GEV_{sim\text{-}prec}$ and GEV_{bound}.

The comparison of GEV_{sim} and $GEV_{sim\text{-}prec\ 50}$ shows how strongly $GEV_{sim\text{-}prec\ 50}$ is affected by PREC flood quantiles. Figure 4.7 illustrates that the $GEV_{sim\text{-}prec\ 50}$ estimates larger discharges for almost all gauges. This result can be explained by the PREC flood quantiles. For the majority of the sites, the $Q_{PREC}(T_{PREC})$ values are larger than the corresponding $Q_{GEV}(T_{PREC})$ estimates. Hence, $GEV_{sim\text{-}prec\ 50}$ also estimates larger values than GEV_{sim} (see Gera, see Fig. 4.6c).

In a further step, Q_{sim} and $Q_{bound\ 50}$ are compared (Fig. 4.8). A positive relative deviation indicates that $Q_{bound\ 50}$ is larger than Q_{sim} despite the asymptotic behaviour towards the upper bound. The $Q_{bound\ 50}$ exceeds Q_{sim}, because $Q_{PREC}(T_{PREC})$ values are mostly larger in comparison to the corresponding $Q_{GEV}(T_{PREC})$ (see example of Gera (Fig. 4.6c)). This implies that the PREC discharges enormously affect the GEV and lead to larger discharges of $GEV_{bound\ 50}$ than GEV_{sim} for the same recurrence interval. Figure 4.8b shows that even for T=10,000 years a positive relative deviation is estimated for the half of the sites. Due to the asymptotic behaviour of $GEV_{bound\ 50}$, there are more sites with a negative relative deviation for T=10,000 than for T=1000 years.

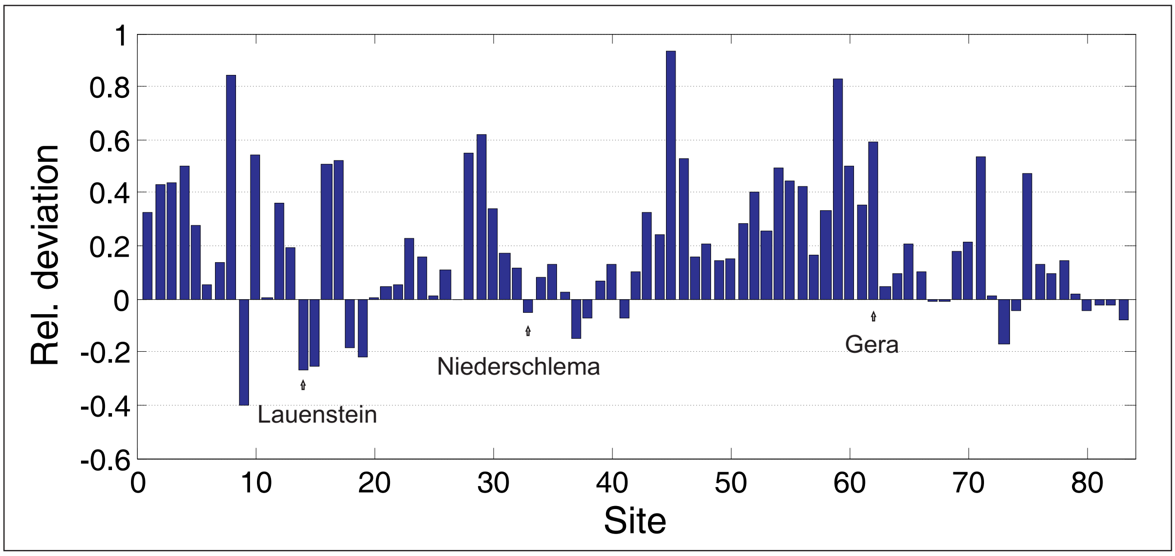

Fig. 4.7: Comparison of discharges estimated by GEV_{sim} and $GEV_{sim\text{-}prec\ 50}$ for the target recurrence interval of 1000 years for 83 gauges. The three sites shown in Fig. 4.6 are marked.

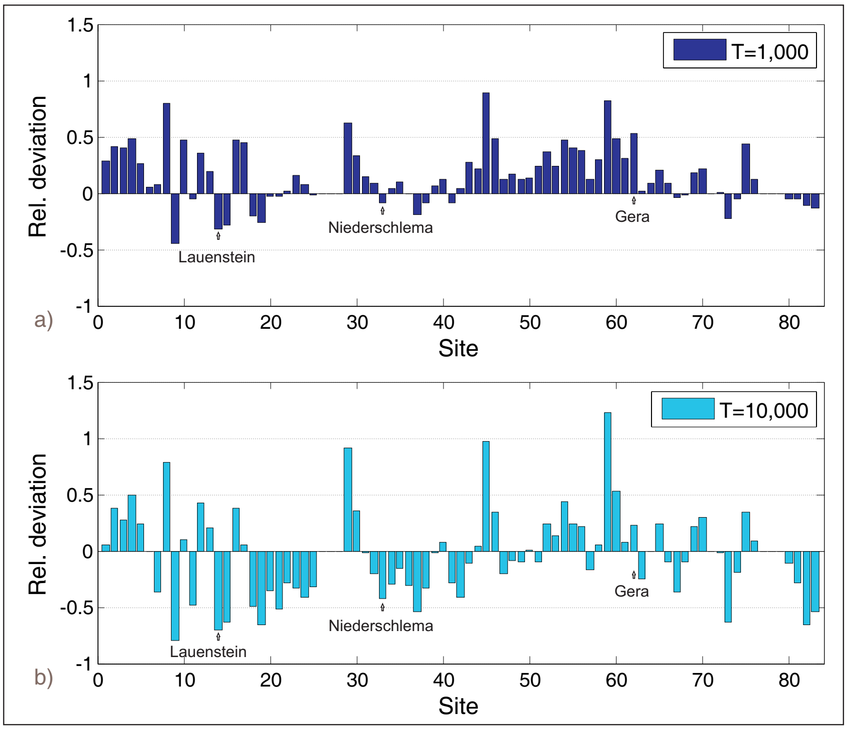

Fig. 4.8: Comparison of discharges estimated by GEV_{sim} and $GEV_{bound\ 50}$ for recurrence intervals of (a) 1000 and (b) 10,000 years. The three sites shown in Fig. 4.6 are marked. The seven sites with a positive k are not shown.

4.5.4 Sensitivity analysis

With a combined sensitivity analysis, the effect of the upper bound derived by the empirical envelope curve, the Q_{PREC}-selection and the inflection point is investigated. Figures 4.9a-c illustrate that the largest relative deviation is found when comparing the 5%- and 95%-quantiles of $GEV_{sim\text{-}prec}$ and emphasise that it is necessary to consider different PREC selections. This variation occurs due to the random selection of the PREC discharges.

The selection of the empirical envelope curve has the lowest relative deviation. There are only small differences in Fig. 4.9a. Its effect is slightly larger for T_X = 200. The smaller T_X, the smaller is the point at which GEV_{bound} asymptotically approaches to the upper bound and the stronger GEV_{bound} is influenced by the empirical envelope curve discharge.

The relative deviation due to the PREC selection is similar when varying the empirical envelope curve or the inflection point (Fig. 4.9b). Here, there is the inverse situation compared to the selection of the empirical envelope curve. The largest relative deviation is found for $T_X = 500$. This can be explained by the fact that, GEV_{bound} is affected from T_X on also by the asymptotic behaviour and not only by the selection of Q_{PREC}.

In Figure 4.9c, the largest deviation was estimated for the different T_X values when using the 95%-quantile of $GEV_{sim\text{-}prec}$. The $GEV_{sim\text{-}prec\ 95}$ is higher skewed than $GEV_{sim\text{-}prec\ 05}$, because of the inclusion of larger Q_{PREC} values. Thus, the difference between the two GEV_{bound} estimates with different T_X values is larger when using the 95%-quantile due to the higher skewness.

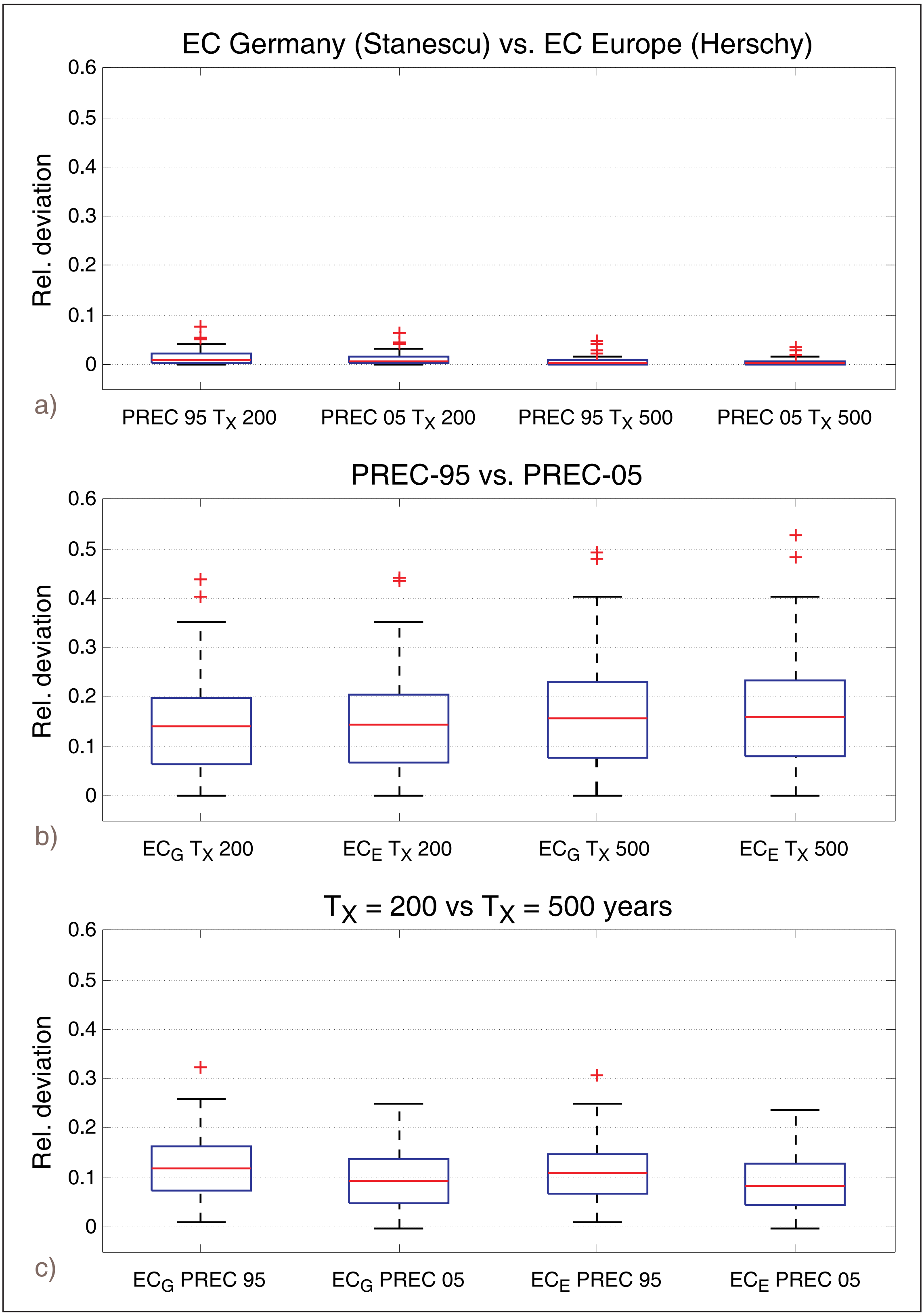

Fig. 4.9: Relative deviation between the quantile estimate of GEV_{bound} for T=1000 years when varying three choices. The boxplots show the results for the 76 sites which were used in the sensitivity analysis. (a) Empirical envelope curves (EC_G = Germany (Stanescu), EC_E = Europe (Herschy)), (b) PREC flood discharges (95-, 5-quantiles) and (c) inflection point (T_X).

The relative importance of the three choices is shown for all 76 gauges (Fig. 4.10). The gauges are ordered by the distance between their unit floods of record and E_{EC}. Figure 4.10 shows that the effect of the selection of the PREC flood discharges increases with larger distance to the REC, whereas the effect of the inflection point and of the empirical envelope curve decreases. This pattern can be explained when considering the three choices in detail.

The effect of the choice of the empirical envelope curve considerably influences the discharge estimates for T_t only for sites with a small distance to the largest unit flood of record, i.e. the sites which are close to the empirical envelope curve. The closer they are to the European one, the larger is the fraction of the empirical envelope curve selection.

The intercept of a REC is defined by the largest unit flood of record in the pooling group. The site which determines in all its PREC realisations the intercept of REC (Neundorf, site 9 in Fig. 4.1) has a relative deviation of zero related to the Q_{PREC} selection (site 3 in Fig. 4.10), because Q_{PREC} is always equal to the at-site flood of record. The smaller the at-site unit flood of record, the larger the distance to the largest unit flood of record of a pooling group could be within a REC. Because of that, the possible range of PREC discharges increases along with the distance between the at-site unit flood of record and the largest regional unit flood of record.

In addition, the effect of T_X is larger for sites with a high skewness. The larger the skewness, the larger are the differences between the discharge estimates for T = 200 vs. T = 500 years. Therefore, the influence of the choice of T_X also increases. Especially the sites with a large flood of record are characterised by a high skewness. Thus, the largest influence of the T_X selection is found for sites with floods of record close to EC. The fraction of the inflection point is highly correlated with the shape parameter k. The effect of the inflection point is negligible for sites with a small negative k, whereas its effect predominates when k is highly negative.

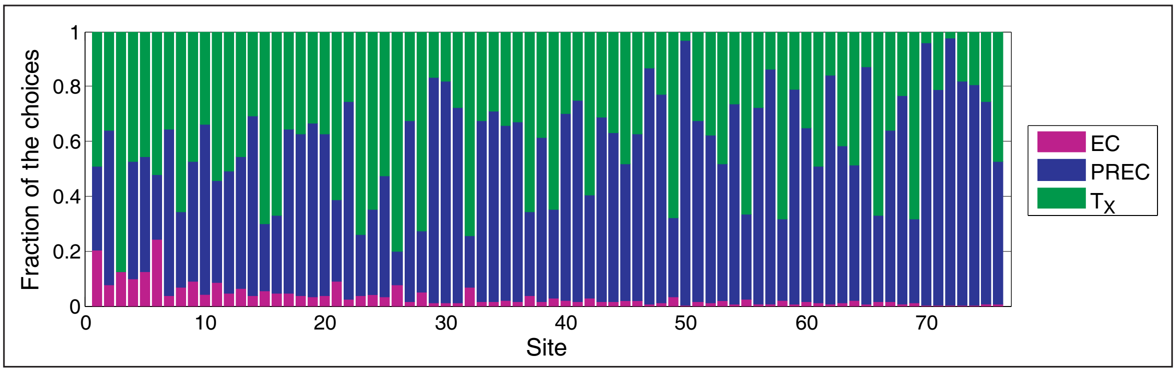

Fig. 4.10: Fraction of the three choices to the overall absolute relative deviation. The sites are ordered by the distance of the unit flood of record to the unit discharge of the European envelope curve. EC = selection of the empirical envelope curve (EC_G vs. EC_E); PREC = selection of PREC flood discharges (95- vs. 5-quantiles); T_X = selection of the inflection point (T_X = 200 vs. 500).

4.6 Discussion

A novel method to integrate additional regional information about upper tail behaviour into at-site flood frequency analyses was presented. This study aimed at improving the discharge estimates for large T. The core ideas were to combine PREC flood quantiles with traditional flood frequency approaches and to introduce a mixed bounded distribution function which considers large flood quantiles as well as an upper bound discharge. It is interesting to compare this method with the integration of historical events and to discuss the selection of PREC flood quantiles and the results of the sensitivity analysis.

There are some similarities between our method to integrate PREC flood quantiles and the use of historical floods as additional information in flood frequency studies. Historical floods are combined as non-systematic data with measured flood series. Generally, a threshold is fixed and the number of floods above this threshold in the historical period is determined (Stedinger and Cohn, 1986; Reis Jr. and Stedinger, 2005). The integration of historical information is based on the assumption that all extreme floods above the threshold are recorded because of the large amount of damages they have caused. However, in this approach discharge values are used only. The probabilities of the historic floods are unknown and are not considered (e.g., Martins and Stedinger, 2001). This is the largest difference to our method, which considers besides the discharge values also the recurrence interval of PRECs. Furthermore, whereas the use of historical data extends the time series, the integration of PREC flood quantiles is based on substituting the time period with spatial information.

Because of that, a different approach than for the integration of historic data was chosen, which enabled us to use the additional information in terms of T_{PREC} and to integrate several Q_{PREC} values. For this, we extended the flood series by using simulated flood series and replaced the simulated discharges above T_l by randomly selected Q_{PREC} values. The largest relative deviation between $GEV_{sim\text{-}prec}$, the flood series which includes the PREC discharges and GEV_{sim} which is based on the simulated flood series only, is calculated for sites with a large $Q_{PREC}(T_{PREC})$ in comparison to $Q_{GEV}(T_{PREC})$.

The selection of the PREC flood quantiles is the most sensitive step for T_t. As indicated, it was necessary to select PREC flood quantiles randomly, because more PREC realisations were provided from Guse et al. (2010) than are to be expected for $T > T_l$ in a T_u year flood series. The influence of the random process depends on two aspects. First, it is affected by the number of PREC realisations. The more PREC realisations, the more combinations of randomly selected PREC discharges are possible. Second, the results are influenced by the variation of the PREC flood quantiles in Q_{PREC} as well as in its corresponding T_{PREC}. Small differences between the PREC flood quantiles lead to low differences in $GEV_{sim\text{-}prec}$ independently of the number of PREC realisations.

As illustrated in Fig. 4.2, both empirical envelope curves differ strongly. However, the sensitivity analysis shows that the effect of the envelope curve selection on a discharge with T = 1000 years is smaller than those of the random selection of PREC discharges or of the inflection point. In this context, it is worth noting that we predefined a target recurrence interval of 1000 years. Since the envelope curve governs the asymptotical approach towards the upper bound, the influence of the envelope curve selection will be larger for increasing T.

4.7 Conclusion

A novel method to improve the quantile estimation for recurrence intervals larger than 100 years by using additional information was presented. Large flood quantiles were derived by probabilistic regional envelope curves (PREC). These PREC flood quantiles were combined with the measured flood series. A mixed bounded distribution function was presented which considers in addition to the PREC flood quantiles also an upper bound discharge derived by an empirical envelope curve. The mixed bounded distribution function avoids an increase up to unrealistic large discharges. Whereas the combination of PREC discharges and a simulated flood series based on at-site parameters was used for recurrence intervals of up to 500 years, a bounded distribution function was applied for larger T.

The main outcomes of this study are:

1) The use of the additional information of PREC flood quantiles and empirical envelope curves supports the estimation of large quantiles.
2) The effect of PREC flood quantiles on the quantile estimation is especially relevant when the PREC discharge varies largely from the at-site GEV estimate for the same recurrence interval.
3) The sensitivity of the flood quantile of 1000 years to the selection of empirical envelope curves providing the upper bound discharge on a flood quantile of 1000 years is smaller than the selection of PREC flood quantiles and of the inflection point between both functions of the mixed bounded distribution.

Acknowledgements

This work is part of the Center for Disaster Management and Risk Reduction Technology (CEDIM) (http://www.cedim.de), a joint venture between the Helmholtz Centre Potsdam - GFZ German Research Centre for Geosciences and the Karlsruhe Institute of Technology (KIT). We thank CEDIM and the GFZ for the financial support.

We thank the State Agency of Environment and Geology of the Free State of Saxony for the permission to use the discharge data. Furthermore we thank the State Agency for Environment of Thuringia for additional discharge data. We also thank the Federal Agency for Cartography and Geodesy of Germany (BKG) for the ATKIS-Basis-DLM and the digital elevation model for Saxony (BKG GeoDataCentre, 2005). The SRTM Digital Terrain Model was downloaded from (http://srtm.csi.cgiar.org/SELECTION/inputCoord.asp, 19.05.2008) (Jarvis et al., 2008).

CHAPTER 5

DISCUSSION AND CONCLUSION

5 Discussion and conclusion

5.1 Main achievements

The core aim of this thesis was an improvement of flood frequency estimates for large recurrence intervals. Therefore, empirical and probabilistic regional envelope curves were introduced into the flood frequency analysis. To resolve the core research question, six sub-questions were posed in the introduction. In the following, the main achievements are emphasised consecutively for each sub-question.

1. Appropriate pooling methods to construct PRECs

Two different pooling methods (cluster analysis, Region of Influence) were compared in chapter two. The performance of the two pooling methods for constructing pooling groups in order to derive PRECs was checked by a cross-validation assuming ungauged conditions for the site of interest and a comparison with the index flood method. It was concluded that the two pooling methods have a similar performance and that both are suitable to construct pooling groups for PRECs.

2. Effect of a modification of a pooling group on a PREC flood quantile

Both different candidate sets of catchment descriptors and different settings within the pooling methods affected the determination of the final pooling group. Several plausible settings within the PREC concept led to an ensemble of PREC flood quantiles which varied in discharge and in the recurrence interval. The PREC flood quantiles react considerably to a modification of the pooling group for both pooling methods. It was shown that the unit flood of record which governs the intercept of PREC has the largest effect on the determination of the PREC discharge. For an accurate estimation of a PREC flood quantile for a given site, it is therefore insufficient to construct only one pooling group.

3. Effect on the recurrence interval of PREC when refining the estimation method of the effective sample years of data

The differentiation into nested and unnested pairs of catchments is an objective criterion, and due to the river network system, nested catchments have a higher correlation than unnested ones. The proposed approach of different parameter sets for the cross-correlation function reduces the recurrence intervals considerably. This illustrates the importance of an accurate consideration of intersite correlation. The largest effect of this refinement of the intersite correlation approach was determined for pooling groups with a large amount of nested pairs of catchments and with a high number of total sample years of data. The selection of one single aspect isolated its influences on the flood quantile estimates and showed that an improvement of this aspect led to more plausible results. Hence, different parameter sets for nested and unnested pairs of catchments are recommended for the estimation of the effective sample years of data to improve the estimation of the recurrence interval.

4. Combination of PREC flood quantiles with measured flood series including the additional information of the PREC recurrence interval

PREC flood quantiles provide a discharge and its corresponding recurrence interval, whereas the traditional AMS includes discharge values only. Therefore, an appropriate method was developed which enables the combination of PREC flood quantiles with the AMS. Therefore, a long flood series was generated based on the at-site parameters of a GEV, and discharges above a certain threshold were replaced by PREC discharges. The recurrence interval was included in this approach by weighting PREC discharges according to their recurrence intervals. By doing so, the main merit of the PREC approach, i.e. the additional information of the recurrence interval, was included in flood frequency analyses. In this way, the combination of the PREC flood quantiles as a non-systematic data with the systematic AMS was realised.

5. Suitable distribution function to integrate PREC flood quantiles as well as an upper bound discharge

A mixed bounded distribution function was proposed which considers PREC flood quantiles and an upper bound discharge derived from an empirical envelope curve. The mixed bounded distribution function consists of two GEV functions which were equated at a predefined inflection point. First, the PREC flood quantiles were inserted into the GEV. Second, an upper bound discharge was included into a GEV with a positive shape parameter. Hence, the function asymptotically approaches the upper bound. In this way, two different types of additional regional information which are representative for the upper tail were combined with the measured flood series. In this way, the upper tail of the distribution function was supported by additional data.

6. Effect of the integration of PREC flood quantiles on the determination of a discharge with a recurrence interval of 1000 years

The influence of the integration of PREC flood quantiles into the mixed bounded distribution function depends on the difference between the PREC discharges and the GEV estimates. The largest differences are estimated when the PREC discharges are larger by a great amount than the GEV estimates for the same recurrence interval. The selection of PREC discharges was detected as the most sensitive choice for T= 1000 years in comparison to the empirical envelope curve and inflection point selection.

5.2 Discussion

5.2.1 Improving flood quantile estimates

The NRC (1988) suggested three principles to improve the quantile estimates of extreme floods. First, the short time period of at-site discharge values should be supported by additional flood data from neighbouring sites (flood regionalisation). Second, it is recommended to use more structural information. Third, NRC (1988) proposed a concentration on extreme floods to avoid a dominant effect of medium floods.

The first and the third principle were already included in the PREC method, which is an adequate regionalisation method and derives a PREC discharge with a large recurrence interval. Hence, additional information for the upper tail of a distribution function is provided. The mixed bounded distribution function contains PREC flood quantiles as well as an upper bound discharge from an empirical envelope curve. The introduction of both shifted the focus more on large floods than on the behaviour of the distribution function for medium-sized floods. The hydrologic behaviour of floods can vary for different recurrence intervals, since the dominant hydrologic processes can change from medium-sized to extreme floods as mentioned in Sect. 4.1.

It can be argued that also structural information is included, because we improved the estimation of the effective sample-years of data by separation into nested and unnested catchments. By doing this, the river network structure is better considered and the representation of the spatial correlation structure is improved which is especially relevant for local floods (see Sect. 3.5).

For the consideration of the particular conditions of the upper tail of a distribution function, it is therefore necessary to concentrate on extreme flood data and to include additional information which is representative for the target recurrence interval, which in this study was a recurrence interval of 1000 years. The estimation of this recurrence interval is improved by a better representation of the river network structure.

5.2.2 PRECs and the trade-off between the number of sites and the degree of heterogeneity

Probabilistic regional envelope curves are therefore a suitable regionalisation method. For an adequate representation of the upper tail of the distribution functions, it is worthwhile to estimate recurrence intervals as large as possible and hence to collect as many sites as possible. However, also the degree of heterogeneity increases with increasing number of sites which conflicts with the strict homogeneity criterion of the PREC concept.

The trade-off between the degree of heterogeneity and the number of sites within a pooling group is a long-lasting debate in flood regionalisation. By using different settings of the pooling methods (number of clusters, threshold of the Euclidean distance), the trade-off was taken into consideration (see Sect. 2.2.2). Both aspects and their interactions were investigated in chapters two and three and discussed as follows.

Figure 3.5 clearly shows that the effective sample years of data increase when more data is available. Hence, the target recurrence interval can also be increased. However, the additional gain of information decreases with increasing total sample years of data.

It was illustrated that relaxing of the homogeneity assumptions by using a higher threshold of the heterogeneity measure leads to a higher number of available pooling groups and therefore of PREC realisations (see Fig. 2.8). However, the performance of the PREC flood quantile estimates decreases with an increasing degree of heterogeneity (see Tab. 2.7).

The performance of PREC flood quantiles is especially affected by an upshift of the REC due to an inclusion of a site whose flood of record exceeds the REC. The high influence of the site which governs the intercept of PREC was emphasised in the Figs. 2.4 and 2.5. Hence, an inclusion of a site with a unit flood of record high above the flood of record of the site of interest lead to a higher relative error (see Fig. 2.7). The poorer performance of more heterogeneous regions can be explained with the increasing distance between the at-site and the largest unit flood of record in heterogeneous regions.

Considering the trade-off between the number of sites and the degree of heterogeneity, this study clearly shows that a relaxing of the homogeneity criterion leads to a decline in the PREC performance for both pooling methods. Therefore, a heterogeneity measure of two seems to be an appropriate threshold. Even in the case of a homogeneous pooling group, it might not be useful to increase the number of sites unlimitedly because of the small additional gain of information in larger pooling groups.

5.2.3 Sensitivity of PREC flood quantiles to the largest unit flood of record

In this thesis, first, the method of probabilistic regional envelope curves was thoroughly investigated and second, the PREC flood quantiles were integrated into the presented mixed bounded distribution function. Hence, the results and achievements of this first part affect the second part. All available PREC flood quantiles were used as additional information for a flood frequency analysis (chapter four) to express the variability in the PREC flood quantiles due to differently constituted pooling groups.

The magnitude of the PRECs was affected in particular by that PREC discharge which governs the PREC intercept. It was demonstrated that the inclusion of the site with the largest unit flood of record within the pooling group can significantly increase the PREC discharge (see Fig. 2.4) and despite of that the strict homogeneity criterion of the PREC concept can still be fulfilled. The high importance of the largest unit flood of record was already seen as a critical point when constructing empirical envelope curves (e.g. England Jr., 2006) and it is still for PRECs.

Two cases can be distinguished when considering the role of the largest unit flood of record within the PREC concept. First, the largest unit flood of record is identical in all or at least in the majority of the PREC realisations for a given site. In this case, this unit flood of record seems to be an adequate choice. Second, the sites which determine the largest unit flood of record differ within the PRECs. In the case of a large difference between the largest unit floods of record, there is a high variability of the PREC discharges for the given sites.

The particular role of that unit flood of record was considered in the integration process in chapter four. Due to the repetition of the PREC discharge selection for $GEV_{sim\text{-}prec}$ (see Sect. 4.4), the effect of a single PREC discharge is relatively low because of the whole ensemble of PREC realisations.

The presented mixed bounded distribution function considers the recurrence interval and also the variability of PREC flood quantiles. By doing so, the high sensitivity of the largest unit flood of record is reduced at least for those sites with several homogeneous pooling groups. Because of that, it is recommended to use several PREC realisations when constructing PRECs and using their results as additional information in a flood frequency analysis.

5.2.4 Comparisons of PREC with at-site and regional flood frequency analysis

The benefit of PREC as a flood regionalisation method can be detected by comparing the PREC results with other regionalisation methods. By doing so, it is possible to determine the particularity of the PREC approach. In chapter two, the PREC flood quantiles for ungauged conditions were compared with those of the index flood method. The integration of PREC flood quantiles into a distribution function in chapter four implicitly contains a comparison of the PREC flood quantiles with the at-site flood frequency analysis. The index flood discharges are not directly comparable with the GEV estimates because they represent the mean regional behaviour, whereas the GEV illustrates the at-site flood behaviour only. However, in these two cases, it was shown that in the majority of cases, PREC estimates larger discharges (see Figs. 2.6, 2.7 and 4.8). This can be explained by the fact that the magnitude of the PREC discharge is determined by the largest unit flood of record, and the more it is above the other floods of record of the pooling group, the more the PREC discharge is larger than other local or regional estimates for the same recurrence interval. In chapter two, it was demonstrated that the difference between the discharge estimates of PRECs and the index flood approach increases with increasing distance of the at-site unit flood of record to the unit PREC discharge (Fig. 2.7). This aspect emphasises the particularity of the largest unit flood of record. Hence, especially the largest unit flood of record needs to be considered accurately.

Probabilistic regional envelope curves estimate large flood quantiles, which can be used as additional spatial information in a flood frequency analysis as was illustrated in chapter four. By doing so, they contribute spatial information to the upper tail of a distribution function. It is recommended to use as much information as possible to reduce estimation uncertainty. Even when using all three types of information (spatial, temporal and causal), the upper tail benefits from adequate information. The introduction of PREC flood quantiles gives therefore precious information. Hence, it is proposed to use probabilistic regional envelope curves as an additional regionalisation method in a multi-pillar approach as suggested by Gutknecht et al. (2006) to improve the estimation of large flood quantiles.

In addition to PRECs, also empirical envelope curves were used to provide an upper bound discharge which avoids an unlimited increase of the distribution function. In this way, a second type of spatial information is included. Both types of envelope curve approaches, the empirical and the probabilistic ones, were used complementary. Whereas the flood quantiles between 150 and 1500 years were covered by PRECs, empirical envelope curves determinated the upper bound discharge. Hence, additional information for different parts of the distribution function was used. In chapter four, it was demonstrated that both envelope curve approaches benefit from a comparison and the inclusion of both improves the estimation of large flood quantiles.

5.2.5 Restrictions and limitations of the PREC concept

While investigating the PREC method, some limitations were detected, which were discussed as follows. In chapter two, it was shown that the number of homogeneous pooling groups varies highly within the study region, leading to a smaller number for the lowland sites. Even the Region of Influence approach, which considers explicitly the specific conditions of the site of interest, could not avoid that there are lowland sites with only a few or no homogeneous regions. The use of several candidate sets of catchment descriptors and the resulting enormous number of pooling groups leads to an increase of the number of sites with PREC realisations (see Sect. 2.4). However, it was not possible to derive PRECs for all sites, and, hence, additional information was not available for all sites.

It can be remarked that a small number of sites within the pooling group can be critical for the construction of a REC. It was observed that in rare cases, positive slopes were estimated for small pooling groups, which is against the idea of a decrease of the unit discharge with increasing catchment size. This is more related to the combination of sites within the pooling groups than a reflection of the real hydrologic situation. This limit can be solved in two ways. The minimum number of sites, which was set to four, can be increased to reduce the number of positive slopes, or a plausibility check of the REC slope can be included.

The spatial extent of a specific REC depends on the available catchment sizes within the pooling group, ranging from the smallest to the largest catchment size. This is a useful approach for pooling groups when the catchment sizes are log-normal distributed and/or include lots of gauging stations. However, in the specific situation of a small pooling group with a high range of catchment sizes, the discharge estimates might be too large for the sites with the largest catchment size because of the smaller slope in comparison to empirical envelope curves. In these cases, it might be appropriate to consider the distribution of catchment sizes when deriving RECs. It could then be useful to remove the site with the largest catchment size. Then the range of catchment sizes is reduced and more realistic discharges are estimated.

5.2.6 Is it possible to scale the PREC concept to whole Germany?

In this thesis, for the first time, PRECs were extensively investigated and several PRECs were derived for pooling groups in Saxony. Since only Saxon flood series were used, it is interesting to discuss which aspects can affect the scaling of PRECs to larger regions.

The PREC concept includes stricter assumptions than the empirical envelope curve approach. Empirical envelope curves were constructed for small regions as well as for the whole world. There is no limit of the spatial extension because of the simple approach. The method of PRECs is based on pooling groups which fulfil the homogeneity criterion of the index flood hypothesis and on a scaling of the index flood (mean annual flood) with the catchment size (see Sect. 2.2.4).

Because of the homogeneity requirements, PRECs could not be derived for a large amount of heterogeneous regions in this study. When scaling the PREC approach to larger regions, it can be expected that the heterogeneity among the sites increases. It was concluded that a heterogeneity measure of two is a reasonable choice to fulfil the homogeneity criterion. Because of that, it might be more difficult to construct homogeneous pooling groups for a larger region.

The second methodical assumption of PRECs is the relationship of the index flood with the catchment size. Hence, it is required to include a representative number of sites which allows an adequate estimation of the mean annual flood, which is required to determine the slope of the REC. This relationship has to be fulfilled for the whole pooling group. In this study, a representative number of gauges of all rivers was available and all parts of Saxony are adequately considered. For a larger region, a higher amount of sites is therefore required. However, the number of sites cannot be increased unlimitedly, because the heterogeneity of the pooling group will increase.

It was shown that a relaxing of the homogeneity criterion lead to a worse performance and this statement might be valid for a larger region. It is therefore expected that probabilistic regional envelope curves can be constructed for other regions in Germany with a comparable spatial extent and degree of heterogeneity.

However, it seems to be difficult to construct PRECs for a large region such as Germany based on all flood series. This is especially true when sites from another flood regime are included. In this case, it is possible to construct pooling groups and to derive PRECs separately for each flood regime.

5.2.7 Benefit for flood design estimations

For flood design, discharge estimates for recurrence intervals in the order of 1000 to 10,000 years are required. This thesis aimed at improving the flood quantile estimations for T = 1000 years, which was selected because of the recurrence intervals estimated by PRECs in this study. There is data available up to the maximum PREC recurrence interval (here: 1500 years). The estimation of larger flood quantiles was also affected by the introduction of PRECs. The use of a bounded distribution function which includes an upper bound discharge gives even more information and avoids the estimation of unrealistic large discharges. In this context, it is worth mentioning that for discharges with T = 10,000, no representative data was included. However, the area of the distribution function which is not supported by information is reduced.

It is also interesting to discuss a maximum recurrence interval which can be estimated by PRECs. It is clear that it is impossible to estimate an exact value, since the determination of the recurrence interval depends on data availability and regional conditions. However, figure 3.5 shows that the additional gain of information decreases with increasing total sample years of data. An approach to a limit at which the additional gain is close to zero can be assumed.

Hence, PRECs are an appropriate regionalisation method for flood quantiles in the order of 1000 years. However, it is not expected that PREC can be used to estimate discharges with recurrence intervals of 10,000 years.

5.3 Remarks on the CEDIM-project

The thesis was embedded in a synoptic view on three different natural hazards (earthquakes, storms, floods). There was an interaction among the three CEDIM research groups. Due to the underlying processes of the natural hazards, the hazard estimation of storms is closer to those of floods than for earthquakes (see Grünthal et al., 2006). A mixed bounded distribution function was used for the flood estimation which was also successfully applied in the winter storm estimation (see chapter four). The distribution function from the storm research was adjusted for the flood application. Storm frequency analyses are often characterised by a negative skewness (e.g. An and Pandey, 2007), which is also the case for the Saxon measurement stations. This implies that the GEV approaches an upper bound. For the storm research, the statistical upper bound from the GEV was replaced by an estimated one which was used within the mixed bounded distribution function (Hofherr et al., 2008). Therefore, there is no transition from a positive to a negative skewness as it is in the flood case at the inflection point. In contrast, flood series has mostly a positive skewness.

The mixed bounded distribution function in this study uses additional spatial information, which was not included in the storm application. The introduction of PREC flood quantiles gives more information and therefore a larger inflection point can be selected. The inflection point was set to a recurrence interval of twenty years in the initial storm study (Hofherr et al., 2008). Owing to the PREC flood quantiles, this value was increased to 500 years. Even without additional information, it seems appropriate to use an inflection point of one hundred years for the flood frequency analysis due to the flood data availability.

The results in chapter four show that discharge estimates of up to T = 1000 years are possible due to the additional spatial information. This was the requirement for the flood hazard estimation to allow flood risk estimates for this recurrence interval and a consistent comparison with the risk of earthquakes.

5.4 Further research

The improved discharge estimates for a given recurrence interval (e.g. T = 100, 200, 500 and 1000 years) can be used for flood risk analyses. Therefore, it is necessary to consider the whole flood risk chain and to include flood hazard estimation as well as the vulnerability of the inundated area (see Fig. 1.1). First, the improved at-site estimates by a flood frequency analysis need to be regionalised along the whole river system. The Topkriging-approach (Skoien et al., 2006) is therefore an adequate geostatistical regionalisation method which extended the traditional kriging approach by the inclusion of the river network system in the interpolation process. Second, the resulting inundation area has to be calculated. An inundation model for Germany is currently under development. As soon as this model is presented, it might be possible to use the regionalised improved flood quantile estimates in combination with the inundation model. Third, the inundated areas can be coupled with a damage estimation model such as FLEMOps (Thieken et al., 2008), which estimates the probable loss of a given recurrence interval (see Fig. 1.1).

Thus, the effect of the use of the mixed bounded distribution on the flood loss estimation can be shown by a comparison with the loss estimates based on the quantile estimates of a traditional distribution function. The final loss estimates, which include the flood hazard estimates based on the mixed bounded distribution function, can be compared with the loss estimates of storm and earthquakes.

5.5 Concluding remarks

This thesis improved the estimation of large flood quantiles. This was realised by using PREC flood quantiles and an upper bound discharge as additional spatial information. The main relevant details of the PREC concept were investigated and its principal influential aspects on the estimation of the PREC flood quantiles were emphasised. The variability of PREC flood quantiles due to a modification of the pooling group was estimated by a sensitivity analysis. The refinement of the intersite correlation approach leads to better estimates of the recurrence interval. The benefit of PRECs for flood frequency analyses was demonstrated by introducing PREC flood quantiles into the at-site flood frequency analysis. By doing so, and also under consideration of an upper bound discharges derived by an empirical envelope curve, flood quantile estimates for recurrence intervals of 1000 years were improved by the inclusion of these two different types of additional information for the upper tail of the distribution function.

References

Acreman, M.C., Sinclair, C.D., 1986. Classification of drainage basins according to their physical characteristics: an application for flood frequency analysis in Scotland. J. Hydrol. 84 (3), 365-380.

AG Boden, 1994. Bodenkundliche Kartieranleitung, Schweizerbart'sche Verlagsbuchhandlung, Hannover, 392 pp.

An, Y. and Pandey, M.D., 2007. The r largest order statistics model for extreme wind speed estimation. J. Wind Eng. Ind. Aerod., 95 (3), 165-182.

Apel, H., Thieken, A.H., Merz, B. and Blöschl, G., 2004. Flood risk assessment and associated uncertainty. Nat. Hazard Earth Sys., 4 (2), 295-308.

Benito, G., Lang, M., Barriendos, M., Llasat, M.C., Francés, F., Ouarda, T.B.M.J., Thorndycraft, V.R., Enzel, Y., A. B., Coeur, D., and Bobée, B., 2004. Use of Systematic, Palaeoflood and Historical Data for the Improvement of Flood Risk Estimation. Review of Scientific Methods, Nat. Hazards, 31 (3), 623-643.

Beurton, S., Thieken, A.H., 2009. Seasonality of floods in Germany. Hydrolog. Sci. J. 54 (1), 62-76.

BKG Geodatenzentrum (Federal Agency for Cartography and Geodesy -GeoDataCentre), 2005. Digital Landscape Model ATKIS Basis DLM, Frankfurt/ Main.

Boorman, D.B., Hollis, J.M., Lilly, A., 1995. Hydrology of Soil Types: A Hydrologically Based Classification of the Soils of the United Kingdom, IAHS Report 126. IAHS Press, Wallingford, UK.

Büchele, B., Kreibich, H., Kron, A., Thieken, A., Ihringer, J., Oberle, P., Merz, B. and Nestmann, F., 2006. Flood-risk mapping: contributions towards an enhanced assessment of extreme events and associated risks. Nat. Hazard Earth Sys., 6, 485-503.

Burn, D.H., 1989. Cluster analysis as applied to regional flood frequency analysis. J. Water Res. Pl.-ASCE 115 (5), 567-582.

Burn, D. H., 1990a. An appraisal of the 'region of influence' approach to flood frequency analysis, Hydrolog. Sci. J., 35 (2), 149-165.

Burn, D.H., 1990b. Evaluation of Regional Flood Frequency Analysis with a Region of Influence Approach. Water Resour. Res. 26 (10), 2257-2265.

Burn, D.H., 1997. Catchment similarity for regional flood frequency analysis using seasonality measures. J. Hydrol. 202 (1-4), 212-230.

Burn, D.H., Goel, N.K., 2000. The formation of groups for regional flood frequency analysis. Hydrolog. Sci. J. 45 (1), 97-112.

Castellarin, A., 2007. Probabilistic envelope curves for design flood estimation at ungauged sites. Water Resour. Res. 43 (4), W04406. doi:10.1029/2005WR004384.

Castellarin, A., Burn, D.H., Brath, A., 2001. Assessing the effectiveness of hydrological similarity measures for flood frequency analysis. J. Hydrol. 241 (3), 270-285.

Castellarin, A., Galeati, G., Brandimarte, L., Montanari, A., Brath, A., 2004. Regional flow-duration curves: reliability for ungauged basins. Adv. Water Resour. 27, 953-965.

Castellarin, A., Vogel, R.M., Matalas, N.C., 2005. Probabilistic behaviour of a regional envelope curve. Water Resour. Res. 41, W06018. doi: 10.1029/2004WR003042.

Castellarin, A., Vogel, R.M., Matalas, N.C., 2007. Multivariate probabilistic regional envelopes of extreme floods. J. Hydrol. 336 (3-4), 376-390.

Castellarin, A., Burn, D.H., Brath, A., 2008. Homogeneity testing: How homogeneous do heterogeneous cross-correlated regions seem? J. Hydrol. 360 (1-4), 67-76.

Chbab, E.H., Buiteveld, H., and Diermanse, F., 2006. Estimating exceedance frequencies of extreme river discharges using statistical methods and physically based approach, Österr. Wasser- und Abfallwirtschaft, 58 (3-4), 35-43.

Cohn, T.A., and Stedinger, J.R., 1987. Use of Historical Information in a Maximum Likelihood Framework, J. Hydrol., 96 (1-4), 215-233.

Condie, R., and Lee, K. A., 1982. Flood frequency analysis with historic information, J. Hydrol., 58 (1-2), 47-61.

Costa, J.E., 1987. A comparison of the largest rainfall-runoff floods in the United States with those of the People's Republic of China and the world. J. Hydrol., 96 (1-4), 101-115.

Crippen, J.R., Bue, C.D., 1977. Maximum Flood Flows in the Conterminous United States. Geological Survey Water-Supply Paper 1887, United States Printing Office, Washington DC.

Cunderlik, J.M., Burn, D.H., 2002. The use of flood regime information in regional flood frequency analysis. Hydrolog. Sci. J. 47 (1), 77-92.

Cunderlik, J.M., Burn, D.H., 2006. Switching the pooling similarity distances: Mahalanobis for Euclidean. Water Resour. Res. 42, W03409. doi:10.1029/2005WR004245.

Cunnane, C., 1978. Unbiased plotting positions: A review. Journal of Hydrology, 37 (3-44), 205-222.

Cunnane, C., 1988. Methods and Merits of Regional Flood Frequency Analysis, J. Hydrol., 100 (1-3), 269-290.

Dalrymple, T., 1960. Flood frequency analyses. US Geol. Surv. Water Supply Pap.: 1543-A, Reston, VA.

Durham, K., 2003. Treating the Risks in Cairns, Nat. Hazards, 30, 251-261.

El Adlouni, S., Bobée, B., and Ouarda, T.B.M.J., 2008. On the tails of extreme event distributions in hydrology, J. Hydrol., 355 (1-4), 16-33.

Eliasson, J., 1994. Statistical estimates of PMP values. Nordic Hydrology, 25 (4), 301-312.

Eng, K., Tasker, G.D., Milly, P.C.D., 2005. An analysis of Region of Influence methods for flood regionalization in the Gulf Atlantic Rolling Plains. J. Am. Water Resour. Assoc. 41 (1), 135-143.

Eng, K., Milly, P.C.D., Tasker, G.D., 2007. Flood Regionalization: A Hybrid Geographic and Predictor-Variable Region of Influence Regression Method. J. Hydraul. Eng.-ASCE 12 (6), 585-591.

England Jr., J.F., 2006. Frequency analysis and two-dimensional simulations of extreme floods on a large watershed, Dissertation, Department of Civil Engineering, Colorado State University.

England Jr., J.F., Jarrett, R.D., and Salas, J.D., 2003a. Data-based comparisons of moments estimators using historical and paleoflood data, J. Hydrol., 278 (1-4), 172-196.

England Jr., J.F., Salas, J.D., and Jarrett, R.D., 2003b. Comparisons of two moments-based estimators that utilize historical and paleoflood data for the log Pearson type III distribution, Water Resour. Res., 39 (9), 1243, doi:1210.1029/2002WR001791.

Enzel, Y., Ely, L.L., House, P.K., Baker, V.R., and Webb, R.H., 1993. Paleoflood evidence for a natural upper bound to flood magnitudes in the Colorado River basin, Water Resour. Res., 29 (7), 2287-2298.

Fernandes, W., and Naghettini, M., 2008. Integrated frequency analysis of extreme flood peaks and flood volumes using the regionalized quantiles of rainfall depths as auxiliary variables, J. Hydrol. Eng., 13 (3), 171-179.

Fernandes, W., Naghettini, M., and Loschi, R., 2010. A Bayesian approach for estimating extreme flood probabilities with upper-bounded distribution functions, Stoch. Env. Res. Risk A., doi:10.1007/s00477-010-0365-4, (in press).

Flemming, G., 2001. Angewandte Klimatologie von Sachsen. Tharandter Klimaprotokolle, Band 4, Technical University of Dresden, Tharandt.

Francés, F., 1998. Using the TCEV distribution function with systematic and non-systematic data in a regional flood frequency analysis, Stoch. Hydrol. Hydraul., 12 (4), 267-283.

Francés, F., and Botero, B.A., 2003. Probable maximum flood estimation using systematic and non-systematic information, in: Thorndycraft, V. R., Benito, G., Barriendos, M. and Llasat, M.C. (Eds.), Paleofloods, Historical Floods and Climatic Variability: Applications in Flood Risk Assessment (Proceedings of the PHEFRA workshop), Barcelona/ Spain, 223-229.

Gaál, L., Kyselý, J., Szolgay, J., 2008. Region of Influence approach to a frequency analysis of heavy precipitation in Slovakia. Hydrol. Earth Syst. Sci. 12 (3), 825-839.

Gaál, L. and Kyselý, J., 2009. Regional frequency analysis of heavy precipitation in the Czech Republic by improved region-of-influence method, Hydrol. Earth Syst. Sci. Discuss., 6, 273-317.

Gaume, E., Bain, V., Bernardara, P., Newinger, O., Barbuc, M., Bateman, A., Blaskovicova, L., Blöschl, G., Borga, M., Dumitrescu, A., Daliakopoulos, I., Garcia, J., Irimescu, A., Kohnová, S., Koutroulis, A., Marchi, L., Matreata, S., Medina, V., Precisco, E., Sempere-Torres, D., Stancalie, G., Szolgay, J., Tsanis, I., Velasco, D., and Viglione, A., 2009. A compilation of data on European flash floods, J. Hydrol., 367 (1-2), 70-78.

Gaume, E., Gaál, L., Viglione, A., Szolgay, J., Kohnová, S. and Blöschl, G., 2010. Bayesian MCMC approach to regional flood frequency analyses involving extraordinary flood events at ungauged sites. J. Hydrol., in press.

GREHYS, 1996a. Inter-comparison of regional flood frequency procedures for Canadian rivers, J. Hydrol., 186 (1-4), 85-103.

GREHYS, 1996b: Presentation and review of some methods for regional flood frequency analysis, J. Hydrol., 186 (1-4), 63-84.

Grünthal, G. and Wahlström, R., 2006. New Generation of Probabilistic Seismic Hazard Assessment for the Area Cologne/ Aachen Considering the Uncertainties of the Input Data, Nat. Hazards, 38, 159-176.

Grünthal, G., Thieken, A.H., Schwarz, J., Radtke, K.S., Smolka, A. and Merz, B., 2006. Comparative risk assessment for the city of Cologne, Germany - storms, floods, earthquakes. Nat. Hazards, 38 (1-2): 21-44.

Guillot, P., and Duband, D., 1967. La méthode du gradex pour le calcul de la probabilité des crues à partir des pluies, Floods and Their

Computation – Proceedings of the Leningrad Symposium, IASH Publication No. 84, 560-569.

Guse, B., Castellarin, A., Thieken, A.H., Merz, B., 2009. Effects of intersite dependence of nested catchment structures on probabilistic regional envelope curves. Hydrol. Earth Syst. Sci. 13 (9), 1699-1712, doi:10.5194/hess-13-1699-2009.

Guse, B., Thieken, A.H., Castellarin, A., and Merz, B., 2010. Deriving probabilistic regional envelope curves with different pooling groups, J. Hydrol., 380 (1-2), 14-26.

Gutknecht, D., Blöschl, G., Reszler, C., and Heindl, H., 2006. A "Multi-Pillar"-Approach to the Estimation of Low Probability Design Floods (in german), Österr. Wasser- und Abfallwirtschaft, 58 (3-4), 44-50.

Herschy, R., 2002. The world's maximum observed floods. Flow Meas. Instrum., 13, 231-235.

Hirsch, R.M., Helsel, D.R., Cohn, T.A., Gilroy, E.J., 1992. Statistical analysis of hydrological data. In: Maidment, D.A. (Ed.), Handbook of Hydrology. McGraw-Hill, New York, pp. 17.1 - 17.55.

Hofherr, T., Kottmeier, C., Heneka, P., and Ruck, B., 2008. Wintersturm Risiko Modell und Echtzeit-Schadensprognose, in: CEDIM – Center for Disaster Management and Risk Reduction Technology (Ed.), CEDIM Entwicklungsbericht Dezember 2008, 17-19 (unpublished work).

Hosking, J.R.M., and Wallis, J.R., 1986a. Paleoflood hydrology and flood frequency analysis, Water Resour. Res., 22 (4), 543-550.

Hosking, J.R.M., and Wallis, J.R., 1986b. The value of historical data in flood frequency analysis, Water Resour. Res., 22 (4), 1606-1612.

Hosking, J.R.M., Wallis, J.R., 1988. The effect of intersite dependence on regional flood frequency analysis. Water Resour. Res. 24 (4), 588-600.

Hosking, J.R.M., Wallis, J.R., 1993. Some statistics useful in regional frequency analysis. J. Hydrol. 29 (2), 271-281.

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

Houghton-Carr, H., 1999. Flood Estimation Handbook 4: Restatement and application of the Flood Studies Report rainfall-runoff method, Institute of Hydrology, Wallingford/UK

IKSE, 2004. Dokumentation des Hochwassers vom August 2002 im Einzugsgebiet der Elbe, IKSE (Internationale Kommission zum Schutz der Elbe), Magdeburg.

Jarvis, C.S., 1925. Flood Flow Characteristics. Trans. Am. Soc. Civil Eng. 88, 985-1032.

Jarvis, A., Reuter, H.I., Nelson, A., Guevara, E., 2008. Hole-filled SRTM for the globe Version 4, available from the CGIAR-CSI SRTM 90m Database. <http://srtm.csi.cgiar.org>.

Jensen, J., Mudersbach, C. and Koziar, C., 2004. Simulation von extremen Sturmflutwetterlagen in der Nordsee und deren statistische Analyse. In: G. Gönnert et al. (Eds.), Tagungsband Workshop: Klimaänderung und Küstenschutz. Universität Hamburg, pp. 159-170.

Kanda, J.A., 1981. A New Extreme Value Distribution With Lower and Upper Limits For Earthquake Motion and Wind Speeds, Theor. Appl., 31, 351-360.

Kjeldsen, T.R. and Rosbjerg, D., 2002. Comparison of regional index flood estimation procedures based on the extreme value type I distribution, Stoch. Env. Res. Risk A., 16 (5), 358-373.

Kjeldsen, T.R. and Jones, D. A., 2006. Prediction uncertainty in a median-based index flood method using L moments, Water Resour. Res., 42, W07414, doi:10.1029/2005WR004069.

Kroll, C.N. and Stedinger, J.R., 1998. Regional hydrologic analysis: Ordinary and generalized least squares revisted, Water Resour. Res., 34 (1), 121-128.

Kuczera, G., 1983. Effect of sampling uncertainty and spatial correlation on an empirical Bayes procedure for combining site and regional information, J. Hydrol., 65 (4), 373-398.

Lettenmaier, D.P., Wallis, J.R., Wood, E.F., 1987. Effect of regional heterogeneity on flood frequency estimation. Water Resour. Res. 23 (2), 313-323.

Madsen, H. and Rosbjerg, D., 1997a. Generalized least squares and empirical Bayes estimation in regional partial duration series index-flood modelling, Water Resour. Res., 33 (4), 771-781.

Madsen, H. and Rosbjerg, D., 1997b. The partial duration series method in regional index-flood modelling, Water Resour. Res., 33 (4), 737-746.

Martins, E.S., and Stedinger, J.R., 2001. Historical information in a generalized maximum likelihood framework with partial duration and annual maximum series, Water Resour. Res., 37 (10), 2559-2567.

Matalas, N.C., 1997. Stochastic hydrology in the context of climate change. Climatic change, 37, 89-101.

Matalas, N.C. and Langbein, W.B., 1962. Information content of the mean, J. Geophys. Res., 67 (9), 3441-3448.

Merz, B., 2006. Hochwasserrisiken - Grenzen und Möglichkeiten der Risikoabschätzung. E. Schweizerbart'sche Verlagsbuchhandlung (Nägele u. Obermiller), Stuttgart.

Merz, B. and Thieken, A., 2004. Flood risk analysis: concepts and challenges. Österr. Wasser- und Abfallwirtschaft, 56 (7/8): 27-34.

Merz, B., Thieken, A.H., 2005. Separating natural and epistemic uncertainty in flood frequency analysis. J. Hydrol. 309 (1-4), 114-132.

Merz, B., and Thieken, A.H., 2009. Flood risk curves and uncertainty bounds, Nat. Hazards, 51(3), 437-458.

Merz, B., Thieken, A.H. and Blöschl, G., 2002. Uncertainty analysis for Flood Risk Estimation. In: M. Spreafico and R. Weingartner (Eds.), International Conference on Flood Estimation, pp. 577-585.

Merz, B., Thieken, A.H. and Gocht, M., 2007. Flood Risk Mapping At The Local Scale: Concepts and Challenges. In: S. Begum, M.J.F. Stive and J.W. Hall (Eds.), Flood Risk Management in Europe - Innovation in Policy and Practice. Adv. Nat. Technol. Haz., Springer, Dordrecht/ Netherlands, pp. 231-253.

Merz, B., Kreibich, H. and Apel, H., 2008. Flood risk analysis: uncertainties and validation, Österr. Wasser und Abfallwirtschaft, 5-6, 89-94.

Merz, R., 2006. Regionalisierung von statistischen Hochwasserkenngrößen. In: Gutknecht, D. (Ed.), Methoden der hydrologischen Regionalisierung, Wiener Mitteilungen, Wasser-Abwasser-Gewässer, Band 197, Institut für Wasserbau und Ingenieurhydrologie, Technische Universität Wien, pp. 109-130.

Merz, R. and Blöschl, G., 2003. A process typology of regional floods, Water Resour. Res., 39 (12), 1340, doi:10.1029/2002WR001952.

Merz, R., Blöschl, G., 2004. Regionalisation of catchment model parameters. J. Hydrol. 287 (1-4), 95-123.

Merz, R., Blöschl, G., 2005. Flood frequency regionalisation - spatial proximity vs. catchment attributes. J. Hydrol. 302 (1-4), 283-306.

Merz, R., and Blöschl, G., 2008a. Flood frequency hydrology: 1. Temporal, spatial, and causal expansion of information, Water Resour. Res., 44, W08432, doi:08410.01029/02007WR006744.

Merz, R., and Blöschl, G., 2008b. Flood frequency hydrology: 2. Combining data evidence, Water Resour. Res., 44, W08433, doi:08410.01029/02007WR006745.

Merz, R., Blöschl, G., and Humer, G., 2008. National flood discharge mapping in Austria, Nat. Hazards, 46 (1), 53-72.

Mimikou, M., 1984. Envelope curves for extreme flood events in North-Western and Western Greece. J. Hydrol. 67 (1-4), 55-66.

Montgomery, D.C., Peck, E.A., Vining, G.G., 2001. Introduction to Linear Regression Analysis, third ed. John Wiley and Sons, Inc., New York, NY.

Moon, Y.-I., Lall, U., and Boswarth, K.: A comparison of tail probability estimators for flood frequency analysis, J. Hydrol., 151 (2-4), 343-363, 1993.

Mosley, M.P., 1981. Delimination of New Zealand hydrologic regions. J. Hydrol. 49 (1-2), 173-192.

Naghettini, M., Potter, K.W., and Illangasekare, T., 1996. Estimating the Upper Tail of Flood-Peak Frequency Distributions Using Hydrometeorological Information, Water Resour. Res., 32 (6), 1729-1740.

Nathan, R.J., McMahon, T.A., 1990. Identification of homogeneous regions for the purposes of regionalization. J. Hydrol. 121 (1-4), 217-238.

National Research Council (NRC), 1988. Estimating Probabilities of Extreme Floods, Methods and Recommended Research. National Academy Press, Washington D.C.

Ouarda, T.B.M.J., Girard, C., Cavadias, G., Bobée, B., 2001. Regional flood frequency estimation with canonical correlation analysis. J. Hydrol. 254 (1-4), 157-173.

Ouarda, T.B.M.J., Bâ, K.M., Diaz-Delgado, C., Cârsteanu, A., Chokmani, K., Gingras, H., Quentin, E., Trujillo, E., and Bobée, B., 2008. Intercomparison of regional flood frequency estimation methods at ungauged sites for a Mexican case study, J. Hydrol., 348 (1-2), 40-58.

Pandey, M.D., Van Gelder, A.J.M. and Vrijing, J.K., 2001. The estimation of extreme quantiles of wind velocity using L-moments in the peaks-over-threshold approach. Struct. Saf., 23 (2): 179-192.

Peel, M.C., Wang, Q. J., Vogel, R.M., McMahon, T.A., 2001. The utility of L-moment ratio diagrams for selecting a regional probability distribution. Hydrolog. Sci. J. 46 (1), 147-156.

Petrow, T., Thieken, A.H., Kreibich, H., Bahlburg, C.H., and Merz, B., 2006. Improvements on flood alleviation in Germany: Lessons learned from the Elbe flood in August 2002, Environ. Manage., 38 (5), 717-732.

Petrow, Th., Merz, B., Lindenschmidt, K.-E., Thieken, A.H., 2007. Aspects of seasonality and flood generating circulation patterns in a mountainous catchment in south-eastern Germany. Hydrol. Earth Syst. Sci. 11 (4), 1455-1468, doi:10.5194/hess-11-1455-2007.

Pitlick, J., 1994. Relation between peak flows, precipitation, and physiography for five mountainous regions in the western USA. J. Hydrol. 158 (3-4), 219-240.

Pohl, R., 2004. Historische Hochwasser aus dem Erzgebirge. Wasserbauliche Mitteilungen, Heft 28, Fakultät Bauingenieurwesen, Institut für Wasserbau und Technische Hydromechanik, Technische Universität Dresden, Dresden.

Rao, R. and Srinivas, V.V., 2006. Regionalization of watersheds by hybrid-cluster analysis, J. Hydrol., 318 (1-4), 37-56.

Reis Jr., D.S., Stedinger, J.R., and Martins, E.S., 2005. Bayesian generalized least squares regression with application to log Pearson type 3 regional skew estimation, Water Resour. Res., 41, W10419, doi:10.1029/2004WR003445.

Reis Jr., D.S., and Stedinger, J.R., 2005. Bayesian MCMC flood frequency analysis with historical information, J. Hydrol., 313 (1-2), 97-116.

Robson, A., Reed, D., 1999. Flood Estimation Handbook 3: Statistical Procedures of Flood Frequency Estimation. Institute of Hydrology, Wallingford/UK. 338 pp.

Rosbjerg, D., and Madsen, H., 1995. Uncertainty measures of regional flood frequency estimators, J. Hydrol., 167 (1-4), 209-224.

Rossi, F., Fiorentino, M., and Versace, P., 1984. Two component extreme value distribution for flood frequency analysis, Water Resour. Res., 20 (7), 847-856.

Schumann, A.H., 2004. Das hydrologische Risiko bei der Bemessung und der Bewirtschaftungsplanung von Talsperren, Wasserbauliche Mitteilungen, Heft 27, 33-46.

Schumann, A.H., 2005. Flood statistical assessment of the event from August 2002 in the Mulde river basin, based on seasonal statistics (in german), Hydrologie und Wasserbewirtschaftung, 49, 200-206.

Skoien, J.O., Merz, R., and Blöschl, G., 2006. Top-kriging - geostatistics on stream networks, Hydrol. Earth Syst. Sci., 10 (2), 277-287.

Stanescu, V.A., 2002. Outstanding floods in Europe; A regionalization and comparison, International Conference on Flood Estimation, Berne, Switzerland, 697-706.

Stedinger, J.R., 1983. Estimating a Regional Flood Frequency Distribution, Water Resour. Res., 19 (2), 503-510.

Stedinger, J.R., and Cohn, T.A., 1986. Flood frequency analysis with historical and paleoflood information, Water Resour. Res., 22 (5), 785-793.

Stedinger, J.R., Lu, L., 1995. Appraisal of regional and index flood quantile estimators. Stoch. Hydrol. Hydraul. 9 (1), 49-75.

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

Stedinger, J.R. and Tasker, G.D., 1986. Regional hydrologic analysis, 2, Model-error estimators, estimation of sigma and log-Pearson type 3 distributions, Water Resour. Res., 22 (10), 1487-1499.

Stedinger, J.R., Vogel, R.M., and Foufoula-Georgiou, E., 1993. Frequency Analysis of extreme events, in: Maidment, D. A. (Ed.), Handbook of Hydrology, McGraw-Hill, New-York, pp. 18.1 - 18.66.

Tasker, G.D., Stedinger, J.R., 1989. An operational GLS model for hydrologic regression. J. Hydrol. 111 (1-4), 361-375.

Thieken, A.H., Müller, M., Kreibich, H., and Merz, B., 2005. Flood damage and influencing factors: New insights from the August 2002 flood in Germany, Water Resour. Res., 41(12), W12430, doi:101029/102005WR004177.

Thieken, A.H., Kreibich, H., Müller, M., Merz, B., 2007. Coping with floods: preparedness, response and recovery of flood-affected residents in Germany in 2002, Hydrolog. Sci. J. 52 (5), 1016-1037.

Thieken, A.H., Olschewski, A., Kreibich, H., Kobsch, S. and Merz, B., 2008. Development and evaluation of FLEMOps - a new Flood Loss Estimation MOdel for the private section. In: Proverbs, D.; Brebbia, C.A.; Penning-Rowsell, E. (Eds.), Flood recovery innovation and response I. WIT Press, Southampton, UK, 315-324.

Troutman, B.M., Karlinger, M.R., 2003. Regional flood probabilities. Water Resour. Res. 39 (4), 1095. doi:10.1029/2001WR001140.

Uhlenbrook, S., Steinbrich, A., Tetzlaff, D., Leibundgut, C., 2000. Zusammenhang zwischen extremen Hochwassern und ihren Einflussgrößen. In: Arbeitskreis KLIWA (Eds.), KLIWA-Heft 1, KLIWA-Symposium - Klimaveränderung und Konsequenzen für die Wasserwirtschaft, Karlsruhe, pp. 187-203.

Ulbrich, U., Brücher, T., Fink, A.H., Leckebusch, G.C., Krüger, A., Pinto, J. G., 2003. The central European floods of August 2002: part 1 - Rainfall periods and flood development. Weather, 58, 371-377.

Viglione, A., 2008. The nsRFA Package, r package version 0.4-8.

Viglione, A., Laio, F., Claps, P., 2007. A comparison of homogeneity tests for regional frequency analysis. Water Resour. Res. 43, W03428, doi:10.1029/2006WR005095.

Vogel, R.M., Fennessey, N.M., 1993. L moment diagrams should replace product-moment diagrams. Water Resour. Res. 29 (6), 1745-1752.

Vogel, R.M., Zafirakou-Koulouris, A., and Matalas, N. C., 2001. Frequency of record-breaking floods in the United States, Water Resour. Res., 37 (6), 1723-1731.

Vogel, R.M., Matalas, N.C., England Jr., J.F., Castellarin, A., 2007. An assessment of exceedance probabilities of envelope curves. Water Resour. Res. 43 (7), W070403. doi:10.1029/2006WR005586.

Vogt, J.V., Soille, P., de Jager, A., Rimavičūité, E., Mehl, W., Foisneau, S., Bódis, K., Dusart, J., Paracchini, M.L., Haastrup, P., Bamps, C., 2007. A pan-European River and Catchment Database, European Commission, EUR 22920 EN – Joint Research Centre.

Vorogushyn, S., Merz, B., Lindenschmidt, K.-E. and Apel, H., 2010. A new methodology for flood hazard assessment considering dike breaches. Water Resour. Res.: in press.

Wiltshire, S.E., 1986. Identification of homogeneous regions for flood frequency analysis. J. Hydrol. 84 (3-4), 287-302.

Zrinji, Z., Burn, D.H., 1994. Flood frequency analysis for ungauged sites using a region of influence approach. J. Hydrol. 153 (1-4), 1-21.

Author's declaration

I prepared this dissertation without illegal assistance. The work is original except where indicated by special reference in the text and no part of the dissertation has been submitted for any other degree. This dissertation has not been presented to any other University for examination, neither in Germany nor in another country.

Björn Guse

Potsdam, July 2010