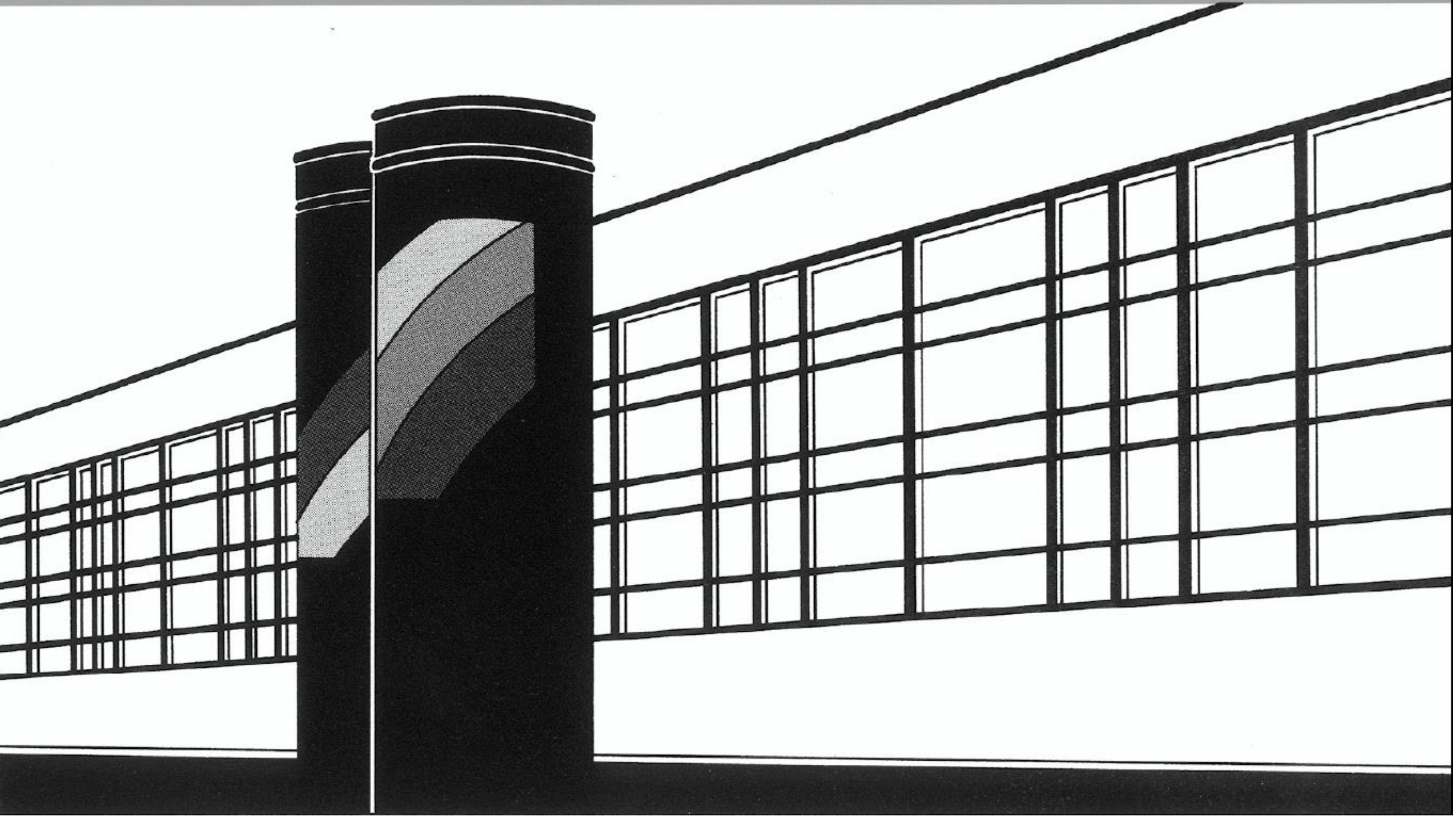
Universität Stuttgart

Institut für Wasser- und Umweltsystemmodellierung

Mitteilungen



Heft 304 Abbas El Hachem

Spatial Extent of Precipitation Extremes in Hydrology

Spatial Extent of Precipitation Extremes in Hydrology

Von der Fakultät Bau- und Umweltingenieurwissenschaften
der Universität Stuttgart zur Erlangung der Würde eines
Doktor-Ingenieurs (Dr.-Ing.) genehmigte Abhandlung

vorgelegt von
Abbas El Hachem
aus Akoura, Libanon

Hauptberichter:	Prof. Dr. rer. nat. Dr.-Ing. András Bárdossy
Mitberichter:	Prof. Dr.-Ing. Uwe Haberlandt

Tag der mündlichen Prüfung: 19. Dezember 2023

Institut für Wasser- und Umweltsystemmodellierung
der Universität Stuttgart
2023

Heft 304

Spatial Extent of Precipitation Extremes in Hydrology

von
Dr.-Ing.
Abbas El Hachem

Eigenverlag des Instituts für Wasser- und Umweltsystemmodellierung
der Universität Stuttgart

D93 Spatial Extent of Precipitation Extremes in Hydrology

Bibliografische Information der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über http://www.d-nb.de abrufbar

El Hachem, Abbas:
Spatial Extent of Precipitation Extremes in Hydrology, Universität Stuttgart. -Stuttgart:
Institut für Wasser- und Umweltsystemmodellierung, 2023

(Mitteilungen Institut für Wasser- und Umweltsystemmodellierung, Universität Stuttgart: H. 304)
Zugl.: Stuttgart, Univ., Diss., 2023
ISBN 978-3-910293-08-3
NE: Institut für Wasser- und Umweltsystemmodellierung <Stuttgart>: Mitteilungen

Herausgegeben 2023 vom Eigenverlag des Instituts für Wasser- und Umweltsystemmodellierung

Druck: P+K Solutions GmbH & Co. KG, Stuttgart

Acknowledgements

First, I would like to thank my supervisor, Prof. Dr. rer. nat. Dr.-Ing. András Bárdossy for his guidance and continuous support throughout my 5-year journey. He was always very patient, welcoming and offering solutions to the many challenges that came along the way. Prof. Bárdossy could always explain and quickly identify the wrong and correct results. After every meeting, I gained new motivation, clarity and strength to continue working on my research topic. I am very thankful for conducting my research under his supervision at the Department of Hydrology and Geohydrology (LHG) of the Institute for Modelling Hydraulic and Environmental Systems (IWS) at the University of Stuttgart.

Second, my gratitude goes to Prof. Dr.-Ing. Uwe Haberlandt from Leibniz Universität Hannover, who joined as my second supervisor. We had several meetings, in which he commented on the methods and results, and suggested future progress. After every meeting, I was able to learn, exchange, correct and improve my research progress.

Additionally, I am grateful for Dr. rer.nat. Jochen Seidel, who was not only the supervisor of my master's thesis, but continued to support me during my PhD. We discussed the methods and results and wrote together several project reports. Moreover, I am thankful to Astrid Lemp, who has accompanied and supported me over the years. Her friendliness and heartiness always offered a place of refuge and communication, even through bureaucratic hurdles.

I thank the Federal Ministry of Education and Research (BMBF) for offering me the chance and funding to work on the STEEP (Space-time statistics of extreme precipitation) subproject within the ClimXtreme research project. In addition, my acknowledgement goes to the International Doctoral Program Environment Water (ENWAT) at the University of Stuttgart and to the course director, Dr. -Ing. Gabriele M. Hartmann. By being part of ENWAT, I could present my results to a more general scientific community and receive great support from the beginning to the end of the PhD journey.

Much appreciation goes to my colleagues at LHG. We shared many discussions, coffee breaks and several Python scripts. Especially, I would like to mention Faizan Anwar, Dhiraj Raj Gyawali, Dr. -Ing Dirk Schlabing, and Masoud Mehrvand. As well as former colleagues Micha Eisele, Thomas Müller, Naibin Song and Claudia Teutsch. Further, I thank my colleague Golbarg Goshtasbpour from Leibniz Universität Hannover for engaging in several interactions within the STEEP project.

I received continuous encouragement and support from my parents in Lebanon and, lastly, from my wife and two sons, who joined in the course of my PhD. I am thankful to God and Jesus Christ for leading you into my life. Thank you for your support, understanding and for bringing so much joy into my life.

Contents

List of Figures

List of Tables

Abbreviation and acronym

Abbreviation and acronym	Description
ADDF	Area Depth Duration Frequency curve
ARF	Area Reduction Factor
CDF	Cumulative Distribution Function
CORDEX	Coordinared Regional Climate Downscaling Experiment
CR	Criteria Ratio
CV	Coefficient of Variation
DWD	Deutscher Wetter Dienst
DDF	Depth Duration Frequency curve
EDK	External Drift Kriging
ESD	Empirical Statistical Downscaling
EVA	Extreme Value Analysis
FFT	Fast Fourier Transformation
GC	Gravelius Compactness Coefficient
GCM	General Circulation Model
GEV	Generalized Extreme Value
KET	Kriging Error Term
KDE	Kernel Density Estimate
KS	Kolmogorov-Smirnov
IDF	Intensity Duration Frequency
LUBW	Landesanstalt für Umwelt Baden-Württemberg
OK	Ordinary Kriging
PWS	Personal Weather Station
Pcp	Precipitation
QC	Quality Control

QQ	Quantile Quantile
QPE	Quantative Precipitation Estimates
RADKLIM	Radar Climatology
RCM	Regional Climate Models
RCP	Representative Concentration Pathway
REA	Reanalysis
RMSE	Root Mean Square Error

Summary

Precipitation extremes are a space-time phenomenon that influences many engineering design decisions. The occurrence of precipitation extremes is, however, rare and with values that can deviate notably from "normal" observations. For design purposes, an estimate of areal rainfall depth for a corresponding return period is needed. Traditionally, point rainfall extreme value statistics are transferred to areal statistics using the concept of the area-reduction factor. The latter varies between 0 and 1 and is used to partition the area by point rainfall depth for the corresponding duration and return period. In this research, the use of this factor is investigated using high-resolution precipitation data.

The first section uses space-time statistical methods to evaluate the quality of recorded intense observations. To cope with the highly non-normally skewed rainfall distribution, a non-linear transformation is applied to normalize the data distribution. In the newly transformed space, the maximum values for each observation location and event duration are estimated using the surrounding stations and the computed spatial variogram. If the difference between the observed and estimated values exceeds the test's critical value, the value is marked as a possible outlier. To avoid singularities caused by convection, the same procedure is repeated for different aggregations. A major challenge is distinguishing outliers between single events and false measurements. Therefore, all flagged outliers are compared to the corresponding weather radar and discharge observations, and implausible data are removed. Using the corrected data the space-time dynamics of selected intense events is investigated. The aim is to visualize and estimate the dependence structure within a rainfall event, identify spatially occurring unusual events using the depth function, and clusters of variograms and step-functions. Different types of rainfall events were identified and accordingly classified.

Rain gauge networks, however, frequently lack the network density required to accurately capture the temporal and spatial extent of extreme events. An alternative is to use weather radar data, which provides spatially distributed rainfall fields. However, these observations are prone to errors. A copula-based conditional merging procedure is applied to minimize errors and artifacts in the weather radar observations. The aim is to combine in the rank space radar and station observations with high temporal resolution. Aspects such as geometrical anisotropy, horizontal wind displacement, advection, and External Drift Kriging are included. The quality of the final product is verified using a cross-validation procedure for several station data sets from two variant radar locations.

Area Depth Duration Frequency (ADDF) curves are a mathematical function relating the area of a location to the depth and frequency of a rainfall event for a certain temporal duration and return period. The calculation of the ADDF curves is not straightforward, as in contrast to point precipitation areal precipitation is not measured but must be estimated.

Although these curves could be directly calculated from the weather radar data, they were not yet fully investigated. Hence, for several locations, the ADDF curves are derived for small to large areas and for short to long durations. The result indicates the ability to derive these curves; however, crossings occur between the curves of the small and large areas. For long durations, the estimated rainfall depth for the same return period and duration is larger for the larger area than for the smaller one. This contradicts previous assumptions (and knowledge) regarding the behaviour of areal statistics. An investigation is conducted to understand why these crossings are occurring. Several reasons were identified such as the role of the sample size, the event type, and the choice of the area center. From the ADDF curves, the area reduction factors (ARF) are derived and their spatial-temporal behaviour is inspected. The derived values are compared to reference values showcasing that the latter underestimate areal extremes for short durations. Moreover, the ARF is seen to vary largely in space and might exceed the value of 1. The ARF exceeding the value of 1 can be attributed to the event type and sample size. For instance, the maximum within an area can occur anywhere independently of its location, with varying spatial frequency of occurrence.

Utilizing the information from the extracted blocks, the occurrence of rainfall extremes was analysed. Several research questions were considered: (1) How many extremes occur in a given area, independent of their location? (2) To what extent is their occurrence in space random, and to what extent is it a structured process? (3) How are the connected volumes behaving in space and time? (4) How does the areal extent relate to event duration, rainfall volume, and discharge volume? The first two research questions were examined for all of Germany, and the last two by analysing rainfall and run-off data in several small and medium size headwater catchments in southern and Western Germany. The results show that the occurrence of events in space is related to their spatial extent; regions exist where the frequency of occurrence of large spatially distributed events is greater than that of smaller ones. Also, interesting relationships between the spatial extent of an event, the event duration, and the event rainfall volumes are identified. For high discharge values, not only does the rainfall intensity matter but also the event duration and spatial distribution of rainfall within a catchment. Many discharge peaks are not necessarily caused by high-intensity events (hourly or daily maxima) but by the accumulation of rainfall cells in space and time.

To assess the impact of climate change, areal extreme value statistics from projected climate scenarios are derived. The EURO-CORDEX 11 data within the radar region of Hannover for historical and future periods are used. A recorrelation procedure of upscaled reference data enables correcting the model spatial dependence structure, which largely influences areal precipitation values. To account for any bias present in the marginal distribution function of future period data, a double quantile-quantile mapping approach is applied. This preserves the signal in future model projections while reducing the data discrepancy. Once the data are corrected a downscaling of the model to a point scale is performed. In a stochastic conditional simulation approach, random mixing is used. For each temporal duration, simulations conditioned on the corrected data and a scaled spatial model are generated. ADDF curves are then derived for future scenarios in a probabilistic manner and are associated with an uncertainty interval.

Zusammenfassung

Niederschlagsextreme sind eine raum-zeitliche Variante, die viele ingenieurtechnische Planungsentscheidungen beeinflusst. Das Auftreten von Niederschlagsextremen ist jedoch selten und mit Werten, die deutlich von normalen Beobachtungen abweichen können. Für Planungszwecke wird eine Schätzung der flächenbezogenen Niederschlagshöhe für eine entsprechende Dauer und Wiederkehrperiode benötigt. Traditionell werden Extremwertstatistiken für Punktniederschläge mit Hilfe des Konzepts des Flächenreduktionsfaktors in Flächenstatistiken übertragen. Letzterer variiert zwischen 0 und 1 und wird verwendet, um die punktuelle Niederschlagshöhe über das Gebiet für die entsprechende Dauer und Wiederkehrperiode zu verteilen. In dieser Studie wird die Verwendung dieses Faktors anhand von hochauflösenden Niederschlagsdaten untersucht.

Im ersten Teil dieser Forschungsarbeit werden statistische Raum-Zeit-Methoden verwendet, um die Qualität der aufgezeichneten intensiven Beobachtungen zu bewerten. Um der sehr ungleichmäßigen Niederschlagsverteilung gerecht zu werden, wird eine nicht lineare Transformation zur Normalisierung der Datenverteilung angewandt. Der optimale Transformationsparameter wird durch ein Optimierungsschema ermittelt, das stochastische Simulationen der gestutzten Standardnormalverteilung verwendet. In dem neu transformierten Raum werden die Maximalwerte für jeden Beobachtungsort und Ereignisdauer anhand der umliegenden Stationen und des berechneten räumlichen Variogramms geschätzt. Der geschätzte Wert am Zielort wird mit dem beobachteten Wert verglichen und, wenn die Differenz den kritischen Wert des Tests überschreitet, wird der Wert als möglicher Ausreißer markiert. Um durch Konvektion verursachte Singularitäten zu vermeiden, wird das gleiche Verfahren für verschiedene Aggregationen wiederholt. Eine große Herausforderung bei der Interpretation von Ausreißern ist die Unterscheidung zwischen Einzelereignissen und Fehlmessungen. Daher werden alle markierten Ausreißer mit den entsprechenden Wetterradar- und Abflussbeobachtungen verglichen und unplausible Daten entfernt.

In einem zweiten Schritt werden die korrigierten Daten verwendet, um die Raum-Zeit-Dynamik ausgewählter intensiver Ereignisse zu verstehen. Ziel ist es, die Abhängigkeitsstruktur innerhalb eines Niederschlagsereignisses zu visualisieren und abzuschätzen, räumlich auftretende ungewöhnliche Ereignisse anhand der Tiefenfunktion zu identifizieren und Variogramme und Stufenfunktionen zu clustern. Dies ermöglicht eine bessere Charakterisierung des zugrunde liegenden räumlichen und zeitlichen Verhaltens. Es wurden verschiedene Arten von Niederschlagsereignissen identifiziert und entsprechend klassifiziert.

Regenmessernetzen fehlt es jedoch häufig an der erforderlichen Netzdichte, um das zeitliche und räumliche Ausmaß von Extremereignissen genau zu erfassen. Eine Alternative

ist die Verwendung von Wetterradardaten, die räumlich verteilte Niederschlagsfelder liefern. Diese Beobachtungen sind jedoch anfällig für Fehler. Um Fehler und Artefakte in den Wetterradarbeobachtungen zu minimieren, wird ein Copula-basiertes Verfahren zur bedingten Zusammenführung angewendet. Ziel ist es, Radar- und Stationsbeobachtungen mit hoher zeitlicher Auflösung im Rangraum zu kombinieren. Aspekte wie geometrische Anisotropie, horizontale Windverschiebung, Advektion und External Drift Kriging werden berücksichtigt. Die Qualität des Endprodukts wird anhand eines Kreuzvalidierungsverfahrens für mehrere Stationsdatensätze von zwei unterschiedlichen Radarstandorten überprüft. Der Schwerpunkt liegt dabei auf der Fähigkeit, die jährlich beobachteten Maxima über die verschiedenen Aggregationen zu reproduzieren. Die fusionierten Daten dienen zusammen mit dem offiziellen deutschen Wetterradarprodukt als Grundlage für weitere Analysen.

Area Depth Duration Frequency (ADDF)-Kurven sind eine mathematische Funktion, die die Fläche eines Ortes mit der Tiefe und Häufigkeit eines Niederschlagsereignisses für eine bestimmte Dauer und Wiederkehrperiode in Beziehung setzt. Obwohl diese Kurven direkt aus den Wetterradardaten abgeleitet werden können, wurden sie bisher noch nicht vollständig untersucht. Daher werden die ADDF-Kurven für mehrere Standorte für kleine bis große Flächen und von kurzer bis langer Dauer abgeleitet. Das Ergebnis zeigt, dass diese Kurven abgeleitet werden können, allerdings gibt es Überschneidungen zwischen den Kurven der kleinen und großen Gebiete. Für lange Zeiträume ist die geschätzte Niederschlagshöhe für denselben Wiederkehrzeitraum und dieselbe Dauer für das größere Gebiet größer als für das kleinere. Dies steht im Widerspruch zu früheren Annahmen (und Erkenntnissen) über das Verhalten von Flächenstatistiken. Es wurde untersucht, warum es zu diesen Überschneidungen kommt. Es wurden mehrere Gründe identifiziert, wie z. B. die Rolle des Stichprobenumfangs, der Ereignistyp und die Wahl des Gebietszentrums. Bei einigen wenigen Ereignissen lagen die flächenbezogenen Niederschlagsmengen und die flächenbezogenen Höchstwerte nicht über dem Gebietszentrum, sondern konzentrierten sich auf eine bestimmte Region. Aus den ADDF-Kurven werden die Flächenreduktionsfaktoren (ARF) abgeleitet und ihr räumlich-zeitliches Verhalten untersucht. Die abgeleiteten Werte werden mit Referenzwerten verglichen, wobei sich zeigt, dass letztere die Flächenextreme für kurze Zeiträume unterschätzen. Außerdem zeigt sich, dass der ARF räumlich stark variiert und den Wert 1 überschreiten kann. Um dies zu verstehen, wurden aus den Wetterradarfeldern mehrere zusammenhängende Niederschlagsblöcke extrahiert, die einen Schwellenwert überschreiten. Für jedes zusammenhängende Gebiet werden die Verhältnisse zwischen dem Gebietsmittelwert, dem Gebietsmaximum und dem Gebietsmittelpunkt der Niederschlagsmenge berechnet. Diese zeigen, dass das Verhältnis zwischen dem Gebietsmittelwert und dem Gebietsmittelwert größer als 1 sein kann, das Verhältnis zwischen dem Gebietsmittelwert und dem Gebietsmaximum jedoch immer unter 1 liegt. Dies deutet darauf hin, dass ein ARF, der den Wert 1 überschreitet, auf den Ereignistyp und die Stichprobengröße zurückzuführen ist. Beispielsweise kann das Maximum innerhalb eines Gebiets unabhängig von seiner Lage überall auftreten, wobei die räumliche Häufigkeit des Auftretens variiert.

Anhand der Informationen aus den extrahierten Blöcken wurde das Auftreten von Niederschlagsextremen analysiert. Dabei wurden mehrere Forschungsfragen berücksichtigt: (1)

Wie viele Extreme treten in einem bestimmten Gebiet unabhängig von ihrem Standort auf? (2) Inwieweit ist ihr Auftreten im Raum zufällig und inwieweit ist es ein strukturierter Prozess? (3) Wie verhalten sich die zusammenhängenden Volumina in Raum und Zeit? (4) Wie verhält sich die flächenmäßige Ausdehnung zur Ereignisdauer, Niederschlagsmenge und Abflussmenge? Die ersten beiden Forschungsfragen wurden für ganz Deutschland untersucht, die letzten beiden durch die Analyse von Niederschlags- und Abflussdaten in mehreren kleinen und mittelgroßen Oberwassereinzugsgebieten in Süd- und Westdeutschland. Die Ergebnisse zeigen, dass das Auftreten von Ereignissen im Raum mit ihrer räumlichen Ausdehnung zusammenhängt; es gibt Regionen, in denen die Häufigkeit des Auftretens von großen räumlich verteilten Ereignissen größer ist als die von kleineren. Auch lassen sich interessante Zusammenhänge zwischen der räumlichen Ausdehnung eines Ereignisses, der Ereignisdauer und den Ereignisabflussmengen feststellen. Bei hohen Abflusswerten spielt nicht nur die Niederschlagsintensität eine Rolle, sondern auch die Ereignisdauer und die räumliche Verteilung der Niederschläge innerhalb eines Einzugsgebiets. Viele Abflussspitzen werden nicht unbedingt durch Ereignisse mit hoher Intensität (stündliche oder tägliche Maxima) verursacht, sondern durch die räumliche und zeitliche Akkumulation von Niederschlagszellen.

Um die Auswirkungen des Klimawandels zu bewerten, werden flächenhafte Extremwertstatistiken aus projizierten Klimaszenarien abgeleitet. Dazu werden die EURO-CORDEX 11-Daten in der Radarregion Hannover für den historischen und zukünftigen Zeitraum verwendet. Ein Rekorrelationsverfahren von hochskalierten Referenzdaten ermöglicht die Korrektur der räumlichen Abhängigkeitsstrukturen des Modells, die die flächenhaften Niederschlagswerte maßgeblich beeinflussen. Um etwaige Verzerrungen in der Randverteilungsfunktion zukünftiger Perioden zu berücksichtigen, wird ein Doppel-Quantil-Quantil-Mapping-Ansatz verwendet. Dadurch bleibt das Signal in den zukünftigen Modellprojektionen erhalten, während die Diskrepanz zwischen den Daten verringert wird. Sobald die Daten korrigiert sind, wird ein Downscaling des Modells auf die Punktskala durchgeführt. Für jede Dauer werden stochastische und bedingte Simulationen auf der Grundlage der korrigierten Daten und eines skalierten räumlichen Modells unter Verwendung von Random Mixing erstellt. Die ADDF-Kurven werden dann für zukünftige Szenarien auf probabilistische Weise abgeleitet und mit einem Unsicherheitsbereich versehen.

1 Introduction

1.1 The general problem

One of the main differences between handling floods and extreme rainfall events is the difference in their spatial occurrence. Generally speaking, flood events occur around river banks with a certain time delay after the occurrence of a rainfall event or snowmelt. Rainfall extremes, however, are not spatially bound and can be associated with varying spatial and temporal extents. For instance, a convective or frontal rainfall event with a short or long duration might lead to similar discharge values. Moreover, among the various climatic variables, rainfall presents a highly spatially and temporally variable process.

General practice in hydrology is to use local observations to estimate the spatial extent of rainfall events using different interpolation and simulation techniques. This often leads to an underestimation of the true maxima that could have occurred between the spatially distributed observations (*Teegavarapu and Chandramouli*, 2005). For example, the whole country of Germany is covered by almost 1000 rain gauges with a high temporal resolution. Such gauges usually have a catchment area of $200cm^2$ and the combination of the 1000 stations has a measuring area of $20m^2$. An observation area that is insignificant compared to the area of Germany of $357,386km^2$. The spatial distribution and density of daily and sub-daily stations can be seen in Fig. 1.1. For many locations, no available observations exist, and the values between the stations need to be estimated, leading to large uncertainty while reconstructing the ′true′ values. The station spatial distribution also displays a heterogeneous pattern, with some areas having a denser network density than others. Note that in many other countries, the available rain gauge network has an even smaller spatial density. Often, intense events occurred that were not correctly captured and were sometimes completely missed. This signifies the problem of trying to derive the correct temporal and spatial extent of extremes, especially using only point observations. Moreover, the high number of zeros in high-resolution rainfall data leads to several challenges. These occur when estimating and simulating the spatial and temporal dependence structures. In other words, an analysis of the hourly values shows that more than 80% of the positive observations occur in shorter than 5% of the considered period and, at the 5 minute temporal aggregation, around 95% or the values are equal to 0 millimeters. Hence, precipitation data are associated with a positively skewed distribution and belong to the zero-inflated variables (*Klemeš*, 2000).

Figure 1.2 shows an example of the high spatial variability within an intense rainfall event. Within a distance of 10 km, one station observed 60 mm, a second one observed 10 mm and a third one observed 0 mm. If only one of the stations was available, the observed rainfall depth would have been notably underestimated or overestimated. Estimating this location

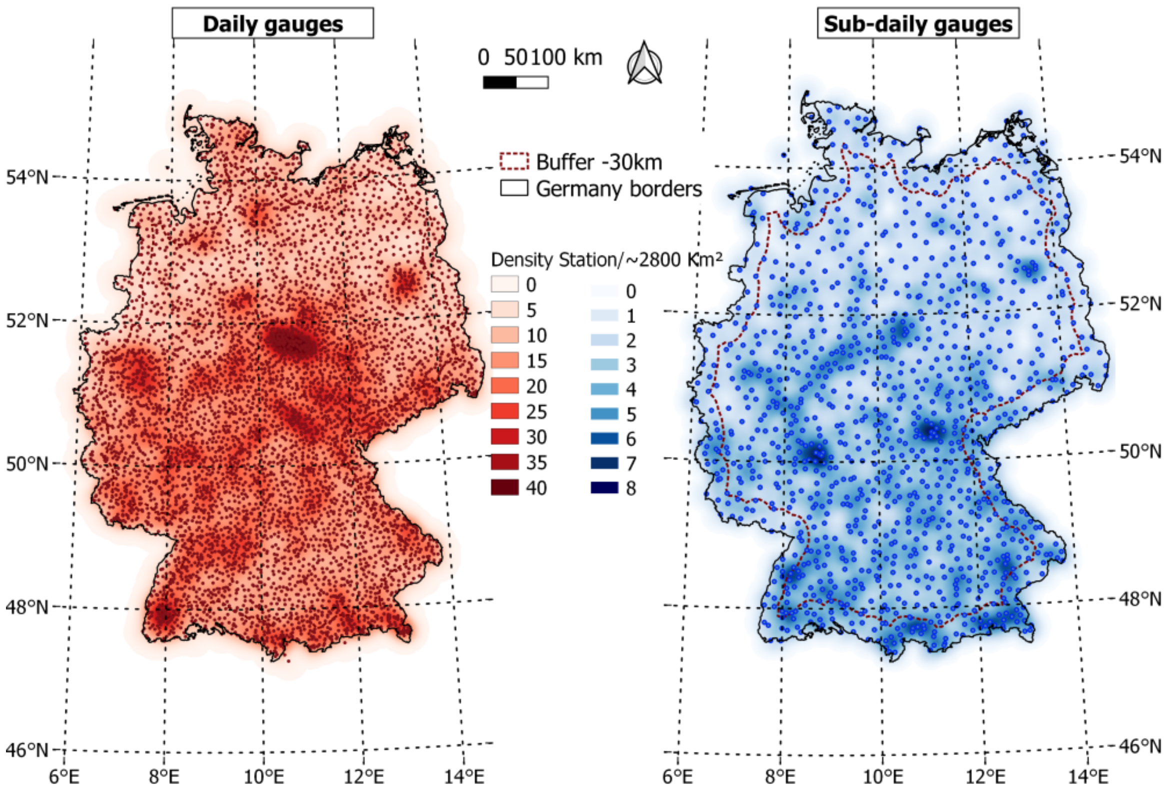


Figure 1.1: Spatial distribution and density of daily and sub-daily stations in Germany.

using the neighbouring observations gave a value of 6 mm, which is ten times smaller than the observed one. Another challenge is marked here: how reliable are the observations themselves? Additionally, the showcased event is associated with a directional dependent correlation, where rainfall is concentrated over a small portion of the region. Such events can be associated with convective rainfall and are even more challenging to model. If the intense precipitation observations were underestimated or missed, the later effects on urban hydrology could be of large magnitude.

Unusual events are non-traditional extremes that occur at a single location or simultaneously at several locations. An interesting aspect is the correct identification and description of unusual events not only in time or space but in a space-time aspect. Some tools are available to investigate this aspect. One of the methods is the data-depth function, such as Tukey's depth (*Tukey*, 1975). This measure is a non-parametric tool to identify unusual events within a dataset. In hydrology, data depth was used for identifying unusual events at single or several locations and for analyzing a single time series. Still, not enough research was undertaken concerning spatially distributed unusual events.

For understanding the spatial extent of extremes in hydrology, especially rainfall extremes, a combination of local information and other observation sources such as weather radar (or satellites) and numerical climate models is needed. Point observations are advantageous since they offer a ground truth, but they present drawbacks since they only cover

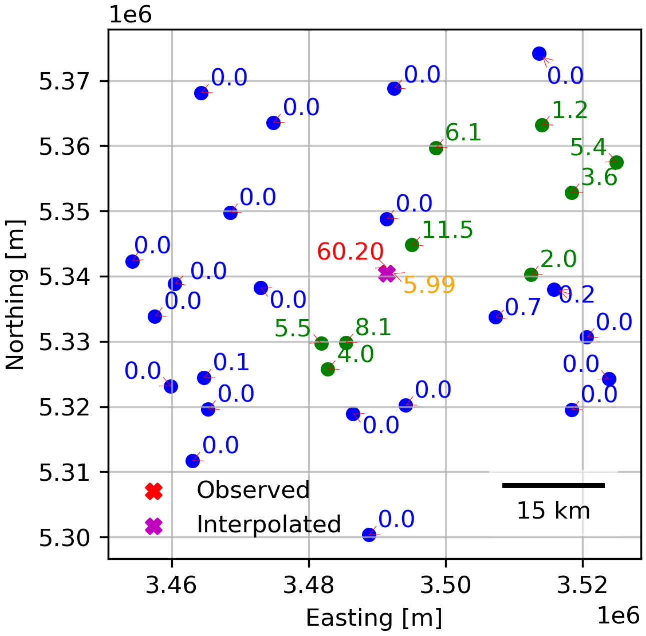


Figure 1.2: Example of spatial variability within an intense daily precipitation event. All gauges in blue recorded zero precipitation values and those in green had positive observations.

a small area and are sparsely distributed. Observations from weather radar, satellites, and numerical climate models might be inaccurate but offer valuable spatial information about the extent of rainfall events. Recently, precipitation estimates derived from weather radar, satellite data, and numerical weather models have become more available and reliable. Despite having their respective drawbacks, the derived rainfall field can be used to improve the estimation of precipitation and, of interest here, the spatial extent of extremes. Still, such data sources suffer from disadvantages and cannot replace gauge observations. For example, weather radar does not directly measure rainfall but rather the reflectivity of the radar signal. This does not depend on the rainfall intensity but on the size of the raindrops. The latter vary from event to event and vary within the event itself. Moreover, the radar antenna is usually measuring at high elevations and not on the ground. On the ground, rainfall is measured between time intervals, while with radar, rainfall (or reflectivity) is measured at a certain time point. Additional sources of errors are possible due to attenuation, clutter existence, variant Z-R relations, and many more. Satellites have a similar concept as radar, but at a coarser time interval. They provide a high spatial resolution image, but only for a single temporal instance, namely a snapshot. What happened in between the two images is 'unknown'. Numerical weather models are based on several physical equations that aim to describe the atmosphere and the interaction between its different parts. They usually provide possible values of several variables (temperature, pressure, wind, rainfall, etc.) over a certain region. Such models often cannot correctly capture the complex phenomena leading to rainfall occurrence over a certain small-scale area, especially for short periods. Despite being usually correct in the mean, they underestimate and overestimate the spatial and temporal variability of rainfall. Therefore, data from weather radars, satellites and numerical weather models need to be corrected and merged (or combined) with ground observations that give a 'true' representation of the measured rainfall at the corresponding locations.

The 'correct' estimation of areal precipitation volumes is still a challenging and much-needed task in hydrology. Deriving correct areal precipitation values depends on the accuracy of the available data and the corresponding spatial density and temporal resolution. Traditional methods were based on point observations derived from a rain gauge network. An underlying assumption was made, namely that the point observations registered the extreme events and their statistical distribution correctly. This assumption is often not fulfilled, as rainfall can differ greatly within nearby locations. To transfer the point observation and associated extreme value statistics to a catchment or area scale, an (Area Reduction Factor (ARF)) is used. Here another assumption is usually made: the extreme rainfall value is maximum at the point of observation and decreases with increasing area size. The ARF is a ratio that lies between zero and one and defines the partition of point design values into areal design values. For a given return period and duration, the statistical rainfall depth is transferred to the catchment or areal scale. *Bárdossy and Pegram* (2018) showed that even for a small area the ARF estimates can be inappropriate and lead to an underestimation of rainfall extremes. Most recently, a combination of point observations with weather radar and satellite data has been used to estimate areal precipitation volumes. The potential of weather radar data in contributing to directly deriving areal statistics is not yet fully investigated. These might offer an alternative to traditional ARF values, and their spatial and temporal variability can be inspected. Important aspects related to the variability of rainfall in space and time and the corresponding spatial distribution of local extremes and sub-regional deviations can now be investigated. This requires moving away from traditional approaches to analyzing extremes to alternative methods and data.

Spatially distributed rainfall fields such as those derived from weather radar data offer another possibility to look at rainfall events from a different perspective, specifically as connected 2D and 3D blocks. From each field, connected blocks exceeding a minimum threshold are identified, analysed, and classified in space and time. New insights, regarding the evolution of a rainfall event in terms of its area, duration, intensity, volume, and speed are investigated. Additionally, the later reaction in the catchments and questions such as how fast does a catchment react (?) which factors are relevant (?) and what is the relation between rainfall and discharge volumes (?) can be tackled. Moreover, areal statistics are space-time dependent, the shape of an object, its location, and the time of occurrence influence the underlying statistics. Comparing point to areal statistics reveals alternative insights into the frequency of occurrence and variability of areal extremes.

Climate change is defined as a change in the statistical properties of a climatic system. Regional climate models Regional Climate Models (RCM) use a change in the physical system (an increase in emissions) to derive possible scenarios for several climatic variables, such as precipitation. Despite having a finer resolution than global climate models (General Circulation Model (GCM)), RCM data need correction and adaptability to the region of interest. Several researchers noted the problems with directly using RCM data for hydrological investigations and offered possible correction methods. Not only are the values themselves important, but so is their spatial dependence. This affects the areal precipitation values and subsequent statistics. Downscaling is usually used to transfer the information from a coarse to a finer scale. To have many possible outcomes, the downscaling can be coupled with a simulation algorithm. One of them is Random Mixing, a stochastic conditional simulation method that preserves the information in spatial observations and results in several possible realizations. Hence, the downscaled fields are associated with the corresponding uncertainty and can be used for further analysis.

1.2 Objectives of the study

The focus of this study is to present the current results regarding investigating and understanding the spatial and temporal extent of intense rainfall events. The analysis is done using high-temporal resolution data such as point observation (station data) and spatial data (weather radar and numerical weather climate models). The main questions to answer are: (1) How can we move from point statistics to a space-time description of extremes? (2) How many extremes occur in a given area, independent of their location? (3) Is their occurrence in space and time: a random or structured process? (4) How does the area size of connected rainfall blocks relate to the volume they bring? (5) How do catchments react to rainfall events with large volumes? (6) How to estimate the spatial extent of extremes for past, present, and future times? For this, adaptive methods need to be established to help with understanding and characterizing the spatial extent of extremes in hydrology. The first step involves a thorough quality check of point observations. The data-depth function tool was tested in a space-time aspect to find the depth of unusual spatially distributed events and the simultaneous occurrence of extreme. From point observations, the spatial extent of extremes was evaluated using clustering of variogram and step functions. Afterward, to achieve a reliable spatially distributed data set, a method for merging station and radar data is developed. From the latter, Depth Duration Frequency curve (DDF) and Area Depth Duration Frequency curve (ADDF) curves along ARF ratios were derived. Their space-time variability and uncertainty were investigated. A new insight into rainfall as 2D and 3D connected blocks will be presented. The aim is to understand how areal extremes behave in space and time through a data-driven analysis. As for regional climate model data, a probabilistic upscaling and downscaling method will be implemented to upscale point observations to model scales and downscale them later to the finer scale. The main outcome of this analysis is the area-intensity-duration-frequency curves that will be calculated for present and future times.

1.3 Sections of this thesis

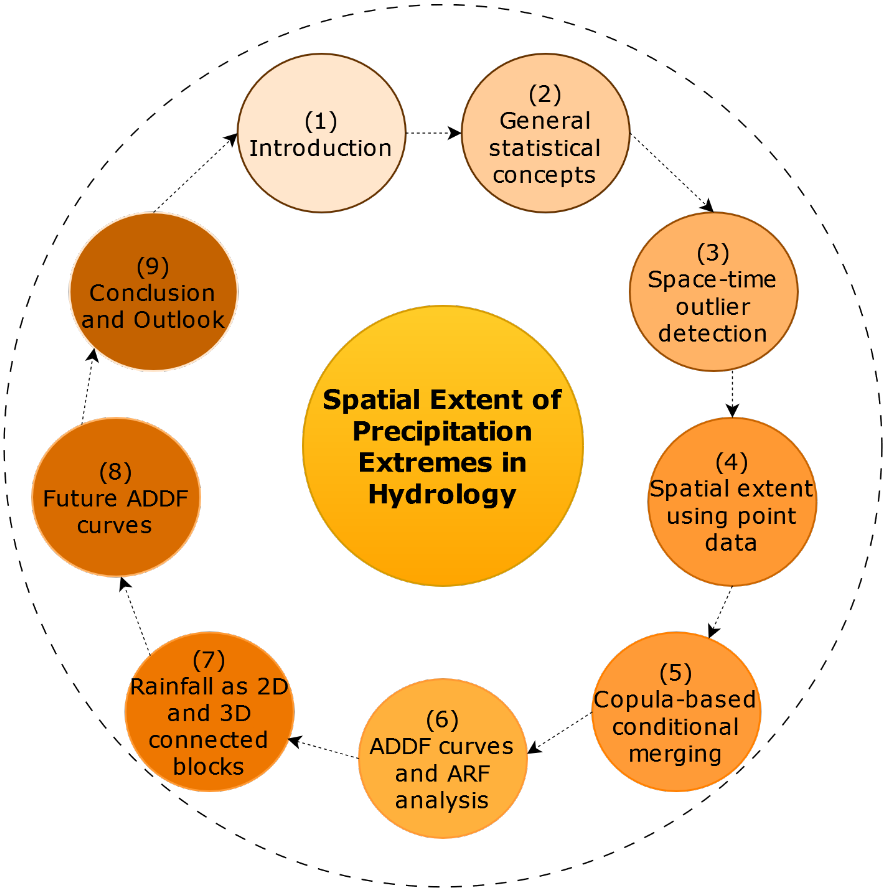


Figure 1.3: Overview of the thesis.

After the introduction, **Chapter 2:** Introduces several statistical measures to describe pairwise dependence. The Kolmogorov-Smirnov test, the K-mean clustering algorithm, and the definition of zero-inflated problems are briefly presented. These concepts and measures are used throughout the thesis.
Chapter 3: Involves a statistical space-time outlier detection method for intense rainfall gauge observations. To cope with the skewed rainfall distribution, a suitable transformation is implemented. For each observation set, a cross-validation scheme with a verification procedure to differentiate outliers between false measurements and single events is developed. The results are presented for the DWD rain gauge network in Germany for all available temporal aggregations. The validated data set serves as a foundation for further analysis.
Chapter 4: The filtered station data were used to inspect the spatial extent of extremes. Several approaches are presented, but the challenges of using point observations are also discussed. The focus was on analyzing event spatial dependence measures, variograms, and step-function clustering, detecting spatially distributed unusual events, and calculating the probability of missing the in-event rainfall maxima. The chapter ends with a discussion regarding the limits of using only point observations.

Chapter 5: A copula-based conditional merging algorithm to incorporate weather radar and station data with high temporal resolution is developed. The effects of anisotropy and the zeros on the variogram are investigated. The conditional merging is coupled with an algorithm to incorporate the horizontal wind displacement vector and anisotropy to derive the final merged product. The latter is cross-validated within two radar regions for several stations with a focus on rainfall extremes. The final resulting data were used along with the official German weather radar data to explore the spatial extent of extremes.

Chapter 6: Dives into design rainfall values by introducing depth duration frequency curves DDF and deriving area-DDF from the weather radar data. From the latter, traditional area reduction factors (ARF) are calculated and compared to reference values for the same duration and area size. The crossing between the ADDF curves is inspected and discussed. The chapter shows that the traditional use of ARF can be inappropriate for some durations and event types.

Chapter 7: The concept of connected rainfall blocks in 2D and 3D is used to compare point and areal statistics and relations between different area sizes. This offers an original approach to looking at areal extremes. The size of the 2D blocks, their location, frequency of occurrence, and areal mean are extracted from the RadKlim data. The movement of the 2D connected areas and their development in time are calculated, and the event volume is derived. The latter is related to the event duration, spatial extent, the amount falling in several headwater catchments, and the subsequent discharge reaction.

Chapter 8: Data from regional climate models are extracted for the region of Germany and used for calculating ADDF curves for future periods. Several data correction approaches were undertaken on the model scale using an upscaling of the reference data. The downscaling of the corrected data is done using a stochastic simulation approach (random mixing), leading to an ensemble of equally possible realizations of ADDF curves. These are then associated with an uncertainty band and are compared to the ADDF curves from the radar data.

Conclusion: A summary of the main findings in each chapter and the main outcomes of this research work are presented.

2 General statistical concepts

Several statistical approaches to describe the dependence and relationship between pairs of data or a cloud of data are discussed and presented in the following section. These measures are used throughout this work; hence, some theoretical background is given beforehand. The second part introduces the Kolmogorov-Smirnov (KS) test for comparing empirical distribution functions, the K-mean clustering algorithm, and the definition of zero-inflated problems.

2.1 Pearson correlation

The Pearson correlation coefficient was initially introduced in the thesis of August Bravais (*Bravais*, 1846) and was later established by Karl Pearson as a robust method to calculate the correlation between quantitatively measured variables (*Pearson*, 1901). Still today, it is one of the most widely used measures to identify the presence (or absence) and direction of a linear relationship between pairs of variables (A and B). Both pairs of data must have a similar size and no missing values. The range of the correlation coefficient is between -1 and 1, where -1 or 1 denotes the presence of a perfect linear relationship with either a downward or an upward slope, respectively. A value of -1 means that as A increases, B decreases. If the Pearson correlation coefficient has a value of 0, then no linear relationship between A and B exists. Though a nonlinear relationship can still be present. A downside of this correlation measure is, on the one hand, its sensitivity to outliers (or extreme values) and on the other hand, the assumption of the normality of the variables. Mathematically speaking, the Pearson correlation coefficient is the normalization of the covariance between the two variables by the product of their standard deviations (defined by equation 2.1). This eliminates the scale dependence problem of the covariance and indicates the presence or absence of a linear relationship.

$$r_{xy} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}} = \frac{cov(X,Y)}{\sigma_X \sigma_Y} \tag{2.1}$$

2.2 Spearman correlation

The Spearman correlation coefficient refers to the Pearson correlation calculated between the ranks of the values instead of the values themselves. It was first introduced by Charles Edward Spearman in his work 'The proof and measurement of association between two things'

(*Spearman, Charles Edward*, 1863-1945). Through this coefficient, the monotonic relationship between the two variables is examined. The value of the Spearman correlation coefficient is also between -1 and 1. The value of -1 means that the ranks of both variables are behaving oppositely. The value that has the highest rank in variable A corresponds to the one with the lowest rank in variable B. If the Spearman correlation coefficient has a value of 0 the absence of a monotonic relationship between the data can be established. One advantage of the rank correlation is its smaller sensitivity towards outliers. Moreover, it is not affected by linearly transforming the data, and no underlying assumption of the normality of the data exists and is, thus, a non-parametric coefficient. Equation 2.2 defines the Spearman correlation coefficient.

$$r_s = 1 - \frac{6\sum d_i^2}{n(n^2-1)} \tag{2.2}$$

Where:

r_s = Spearman's rank correlation coefficient
d_i = difference between the two ranks of each observation
n = number of observations

2.3 Indicator correlation

In the work of *Pearson* (1901) the correlation value between two Boolean vectors (0 and 1) was introduced as the phi (ϕ) coefficient. In this case, the Pearson and Spearman correlation estimated between two binary vectors will return the ϕ coefficient. The indicator correlation offers the possibility of calculating the ϕ coefficient for different quantiles or threshold values. Considering a quantile or a threshold value (α), the original values of A and B are transformed to binary values (0 and 1) by comparing them to the value of α (as in equation 2.3). The calculated correlation between the indicator series is called the indicator correlation. The latter can be calculated for classified non-quantitative data (*Řezanková and Everitt*, 2009). The indicator series is derived as follows:

$$I_{t,\alpha} = \begin{cases} 1 & \text{if A(t)} \geq \alpha \\ 0 & \text{otherwise} \end{cases} \tag{2.3}$$

The practical use of such a measure in hydrology is, for example, to compare the spatial dependence structure between one location and other locations. An example of using the indicator correlation to compare the spatial dependence structure between weather radar observation at gauge locations and gauge observation is seen in *Brommundt and Bárdossy* (2007). In the study done by *Bárdossy et al.* (2021), the indicator correlation was used as a measure of dependence to identify Personal Weather Station (PWS) with reliable observations. For this, the indicator correlation of the PWS was compared to a reference network with trustworthy data.

2.4 Cross-correlation

Cross-correlation is used to identify the degree of correlation over the whole distribution of two variables (two time series). The normality of the data is also assumed. The cross-correlation values range between -1 and 1, with a value of 1 indicating that the two data sets are identical. Equation 2.4 defines the cross-correlation coefficient.

$$r_{x,y} = \frac{\sum_{i=1}^{n}(h_x(t) - \bar{h}_x)(h_y(t) - \bar{h}_y)}{\sqrt{\sum_{i=1}^{n}(h_x(t) - \bar{h}_x)^2}\sqrt{\sum_{i=1}^{n}(h_y(t) - \bar{h}_y)^2}} \tag{2.4}$$

Where:

$r_{x,y}$ = cross-correlation between two locations (x, y)
$h_x(t)$ = time series at location x
$h_y(t)$ = time series at location y
n = number of observations

The cross-correlation can be calculated for indicator series and different thresholds (or quantiles), as in *Brommundt and Bárdossy* (2007). Additionally, the cross-correlation coefficient has been used to evaluate the spatial variability of rainfall between different locations and to derive a spatial correlation function (*Berndtsson*, 1987).

2.5 Coefficient of determination

Linear regression models are used to reproduce observed values using a simple function, whose parameters are derived from the data. To assess the agreement between the estimated and observed data, the coefficient of determination (R^2) can be used as a measure to quantify the quality of the applied model (defined by equation 2.5). It reflects the variation between the model input and output data and enables identifying the fraction of the observed variance that is explained (or not explained) by the simulated data (*Wright*, 1921). Usually it varies between 0 and 1, where 1 reflects a perfect match. Though R^2 can have negative values in worst-case scenarios.

$$R^2 = 1 - \frac{\sum_i (y_i - \bar{y})^2}{\sum_i (y_i - y_i^*)^2} = 1 - \frac{SS_{ref}}{SS_{tot}} \tag{2.5}$$

Where:

R^2 = coefficient of determination
y_i = observation dataset
y_i^* = simulated dataset
SS_{ref} = is the sum of squares of the residual errors
SS_{tot} = total sum of errors

2.6 Kolmogorov-Smirnov test

The goodness-of-fit test for a sample was initially introduced by *Kolmogorov* (1933). The Kolmogorov-Smirnov (KS) test for two samples was presented by *Smirnov* (1939). KS is a non-parametric test used to examine the relationship between two functions. It allows comparing the similarity of two distribution functions by measuring the separating distance between the two curves for every observation. The null hypothesis (H0) states that the two data originate from the same distribution function. The opposite hypothesis (H1) states that both distributions have different distribution functions. Usually, a maximum p-value of 5% is used to accept H0 or reject it. The KS test values lie between 0 and 1. Considering two cumulative distribution functions, $F(x)$ and $G(x)$, one can mathematically express the dominance of one distribution over the other by the first-order stochastic dominance test. If $F(x)$ lies completely to the right of $G(x)$, then $F(x)$ has first-order stochastic dominance over $G(x)$ (*Tanaka et al.*, 2012).

$$\int u(x)dF(x) \geq \int u(x)dG(x) \forall u \in \Re \tag{2.6}$$

The two-sided KS statistic is used to test the hypothesis that $F(x)$ and $G(x)$ are identical, and the null and alternative hypotheses H_0 and H_1, are formulated as follows:

$$H_0 : F(x) - G(x) = 0 \forall x \in \Re \tag{2.7}$$

$$H_1 : F(x) - G(x) \neq 0 \forall x \in \Re \tag{2.8}$$

As for the one-sided KS test, the stochastic dominance of one distribution over the other is tested:

$$H_0 : F(x) - G(x) \leq 0 \forall x \in \Re \tag{2.9}$$

$$H_1 : F(x) - G(x) \geq 0 \text{ for some x } \in \Re \tag{2.10}$$

The two-sided KS test KS_2 and the one-sided KS test KS_1 are expressed as:

$$KS_2 = \sqrt{\frac{nm}{N}} \max |F_{1,n}(X_i) - G_{2,m}(X_j)| \text{ for } 0 \leq i \leq N \tag{2.11}$$

$$KS_1 = \sqrt{\frac{nm}{N}} \max(F_{1,n}(X_i) - G_{2,m}(X_j)) \text{ for } 0 \leq i \leq N \tag{2.12}$$

Where:

n = sample size of $F(x)$
m = sample size of $G(x)$
N = n+m

Gnedenko and Korolyuk (1951) found a simpler combinatorial proof of the KS equations, a proof that was further elaborated by *Hodges Jr* (1958) and is implemented in the Python library Scipy (*Virtanen et al.*, 2020). The latter was used in this thesis.

2.7 K-Mean clustering

Diday and Simon (1976) defined clustering as a pattern recognition technique to identify objects with similarities or dissimilarities. One of the most common clustering approaches is the K-Mean algorithm. The aim of K-Mean clustering is to group data based on the separating distance between a centroid and the data points. The number of centroids is equal to the number of chosen clusters, and the optimal location of the centroid (cluster center) is found by an iteration procedure. The goal is to start with random locations for the centroids, calculate the distance between each point and centroid, assign each point to the group with the nearest distance, recalculate the centroid location as a mean of all points within the group, and repeat the procedure until the points belonging to each group do not change anymore. At the end of this process, points belonging to the correspondent group are identified and, hence, clustered. The idea was developed by *Steinhaus et al.* (1956). Different distance measures can be used; the most common one is the Euclidean distance.

2.8 Zero-inflated problems

Several environmental variables that have many zero observations cannot be well represented by standard distributions (for example, normal) and need special handling (*Heilbron*, 1994). The zero values could be due to correct observations, instrument malfunction, or a detection limit. A variable is considered zero-inflated if the number of available zero values is excessively large, which is the case for rainfall data. For the latter, the number of zeros is scale dependent and predominates the observations on the sub-hourly scale (see Table 2.1). For a 5-minute temporal resolution, the probability of having a zero value is around 94. The high number of zeros affects, for example, the variogram estimation, namely deriving the spatial correlation structure of the rainfall field and hence the interpolation results.

Temporal aggregation	5min	30min	60min	360min	720min	1440min
P_0 (probability Pcp=0)	0.94	0.89	0.87	0.72	0.63	0.49

Table 2.1: Average P_0 values derived from DWD stations for several temporal aggregations.

3 Space-time quality control of intense point observations

The following sections are based on a publication published as ”Technical Note: Space–time statistical quality control of extreme precipitation observations” (*El Hachem et al.*, 2022).

3.1 Literature review

Defining an outlier might be intuitive to many, but it has been stated alternatively done by several researchers. For instance, for *Barnett and Lewis* (1994) an observation is defined as an outlier if it represents an inconsistent behavior with regard to other data values. In the work of *Hawkins* (1980), an observation is seen as an outlier if it diverts substantially from other values. In other words, it might have been produced by an alternating mechanism. For *Iglewicz and Hoaglin* (1993) an outlier is defined as an observation that arouses suspicion to the analyst and as it might belong to a different data distribution. Generally speaking, there are two types of outliers: those associated with an error and those with a real observation. Several reasons can affect the measurement quality. For example, instrumental errors (e.g., use of the wrong instrument, equipment failure, inappropriate equipment operation) or/and human-related errors (false reading or documenting, or even derivation of observations). Errors can also happen if the measuring site is falsely chosen. For example, placing a rain gauge in the valley of a mountain or beneath a tree. Such measurements provide false representativeness of the underlying process.

Special care is needed for Hydroclimatological data, especially for precipitation data with unique occurrences in space and time. For instance, if an observation was not correctly captured, reconstructing it is hardly possible, especially for single (localized) rainfall events. However, information about precipitation extremes with reliable quality is predominant for many design purposes, such as flood prediction, extreme value statistics, and stationarity analysis, to name just a few. In order to provide reliable data, many Quality Control (QC) algorithms have been implemented by weather service agencies. As an example, a comprehensive QC algorithm for daily surface meteorological observations (temperature, precipitation, snowfall, and snow depth) was presented by (*Durre et al.*, 2010). The applied method for detecting false precipitation observations consisted of several steps. At first, a climatological outlier check is used to flag values exceeding a certain temperature-dependent threshold. Secondly, a spatial consistency check is used to compare the target value to neighboring ones. An observation is defined as an outlier if the difference compared to neighboring values exceeds a certain climatological percent rank threshold. Similarly, in the work done by *Hubbard et al.* (2005) for daily temperature and precipitation values, quality was checked by a four-step procedure. Using the long-term mean at a reference location, a value is flagged if it does not belong to the interval mean ± 3 standard deviation. A further step was to check how large the difference was between the observed and estimated values using a spatial regression technique. The method was tested for six stations, and the author concluded that the method had a success rate of 40%. A more recent study to identify erroneous hourly rain gauge observations was done using radar Quantative Precipitation Estimates (QPE) as support information *Qi et al.* (2016). Many QC algorithms use an interpolation method to estimate the reliability of a target observation using the surrounding locations. *Ingleby and Lorenc* (1993) defined this mathematically as the probability of an observation being suspicious. Other QC methods are available, but mainly for time series analysis and overlook the challenge of rainfall events' space-time dimension.

Rainfall measurements from rain gauges are typically taken at specific spatially distributed locations with defined time intervals, such as every minute, hour, or day. However, precipitation events are actually continuous processes in space and time. Furthermore, many precipitation observations have zero values, which can significantly impact the data distribution. For example, at the hourly scale, around 90% of observations may have zero values, and this percentage can be even higher, around 95%, for data with 5-minute temporal resolution. This can result in highly non-normally skewed distributions with heavy tails, which can affect the statistical analysis of the data. To address the issue of zero values, measures such as Box-Cox transformation with appropriate parameters can be applied to alleviate the effect of zeros, as suggested by (*Box and Cox*, 1964).

In the following section, a statistical space-time methodology is used to detect outliers in intense precipitation observations. Outliers are identified as observations that exhibit significant disagreement with their neighboring locations. To differentiate between correct and false observations, the detected outliers are validated using supporting information, such as discharge and radar measurements.

3.2 Data description

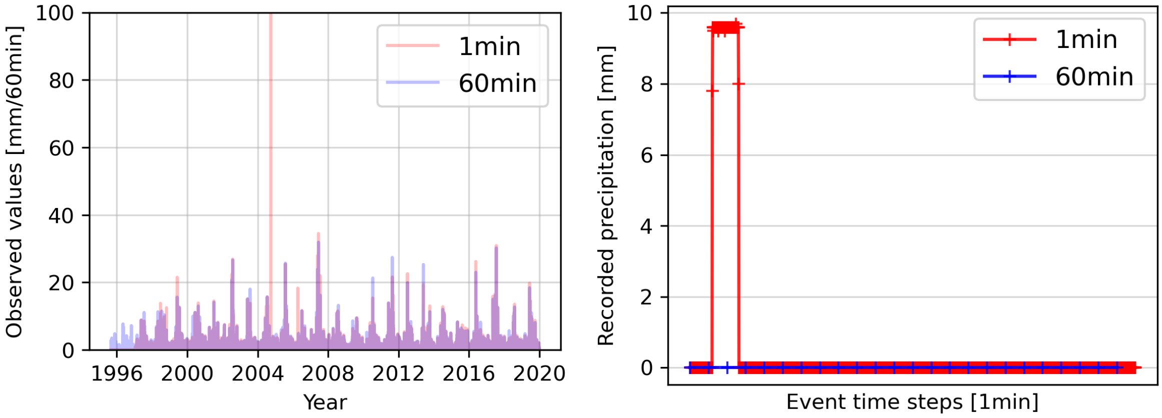
The investigation utilized the German-wide precipitation dataset from the Deutscher Wetter Dienst (DWD), which provides data on average annual rainfall in Germany, ranging from around 800 mm to up to 2100 mm in higher elevations of the Alps in the south. The DWD operates a network of approximately 1000 rain gauges across Germany, with varying temporal resolutions, including minute, hourly, sub-hourly, and daily data. Hourly and sub-hourly data have been available from the 1990s onwards, and the number of stations providing such data has been continuously increasing over time. However, the number of stations with daily observations has been decreasing as they are being replaced by automatic rain gauges. It's important to note that rain gauges located near the border, within a 30-kilometer inland buffer, were not included in the analysis. The use of the DWD precipitation dataset provides a comprehensive and reliable source of precipitation data for Germany, allowing for a detailed investigation of areal extremes and their spatial extent in the region.

In the 1990s, most DWD rain gauges were tipping buckets or drop counters. From 2000 onwards, these were replaced by weighing gauges (OTT Pluvio), and since 2017, these have been replaced by combined tipping buckets and weighing rain gauges (Lambrecht rain [e]). The precipitation data from the recent DWD (German Meteorological Service) observation network undergo multiple quality control steps to ensure data accuracy and reliability. The first step involves a quality control process at the automatic weather stations, where relatively wide thresholds are applied for tests such as completeness, thresholds, temporal consistency, and internal consistency. Based on these tests, a quality flag is assigned to the data, and it is then submitted to a database. Next, a tighter quality check is performed using the QualiMet software, which applies more stringent thresholds for tests related to completeness, climatological consistency, temporal consistency, spatial consistency, and internal consistency. Questionable values are manually checked and corrected, and the quality label is adjusted accordingly. A final quality check is conducted after all the data for a month are available, with a focus on aggregate values. DWD also has procedures in place for identifying and correcting or describing errors in historical data. Although the quality of historical data is generally considered to be reasonably good, there may still be some doubtful values, particularly for pre-1979 data, on the order of about 0.1-1%. It's important to note that the data may be influenced by non-climatic effects, such as changes in instrumentation or observation time. In most cases, the data are reported ″as observed,″ meaning no homogenization procedure was applied. Overall, DWD implements quality assurance measures to ensure the accuracy and reliability of precipitation data, but users should be aware of potential limitations and uncertainties associated with historical data and non-climatic effects.

As independent data for verification, radar-derived rainfall QPE and discharge observations from the state of Bavaria were used. The radar data used is the product RADOLAN-RW that is provided by the DWD in hourly and daily resolutions starting in 2005 (*DWD Climate Data Center (CDC)*, 2021b). These products have been gauge-adjusted with the observed hourly station data. The occurrence (or absence) of precipitation observations in the radar data over the target location is an indication of the quality of the observation. The discharge data were quality checked and provided by the environmental agency of Bavaria with hourly and daily resolutions for approximately 400 gauges within the region of Bavaria (*Bayerisches Landesamt für Umwelt*, 2022). Different headwater catchments were derived and selected to validate the results. A reaction (within a few hours) in the headwater catchment discharge is expected after the event occurs in the case of correct rainfall observations.

In Fig. 1.1 a map of the locations of daily and sub-daily rain gauges along their spatial density is shown. The map shows that the stations have a heterogeneous spatial distribution, where some areas have a higher observation density than others. The spatial density was calculated using a Kernel Density Estimate (KDE) with a Quartic shape and a radius of influence of 30 kilometres. Further information regarding KDE estimation can be found in *Yu et al.* (2015).

A quick analysis of the data shows unusual and unrealistic observations, especially in the 1-minute data. Fig. 3.1a shows an example of comparing minute to hourly data for a randomly selected station. One can see in Fig. 3.1b that in the hourly aggregated minutely data, physically unrealistic extreme values (above 400 mm) are present. When looking closely into the 1-minute data, a value of 9.6 [mm] was recorded for 86 continuous time steps, leading to such outliers in the data. The reason for such values is not specified but could be due to a technical error.



(a) Observed time series

(b) False observation values

Figure 3.1: Example of minutely and hourly data for the same station with an evident outlier in the minute data.

3.3 Method

3.3.1 Data transformation

As an initial step, a Box-Cox transformation as described by equation 3.13 was applied for every variable X and temporal aggregation t to reduce the effect of the skewed precipitation distribution (*Box and Cox*, 1964).

$$X_t^* = \begin{cases} \frac{(X_t^\lambda - 1)}{\lambda} & \text{if } \lambda \neq 0 \\ \log(X_t) & \text{if } \lambda = 0 \end{cases} \tag{3.13}$$

Where:

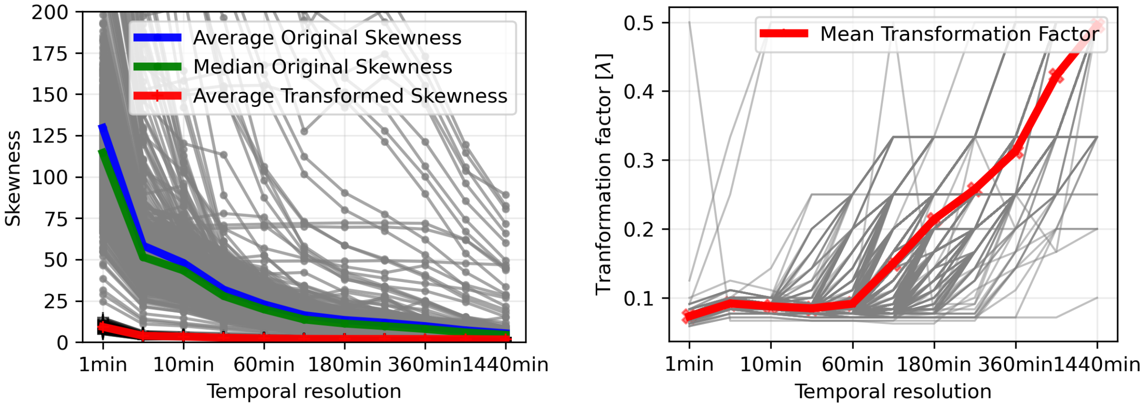
X^* = transformed precipitation data at location u and temporal aggregation t
X = original precipitation data at location u and temporal aggregation t
λ = transformation factor for temporal aggregation t

To find which transformation factor λ is most suitable, several simulated lower truncated standard normal distribution functions (sampling space bounded by $[-\infty < a = p_0,\ b = +\infty]$) were fitted to the original data (*Burkardt*, 2014). The probability of having a value above or below p_0 is then derived (p_0 probability of having 0 mm precipitation value).

From this probability (denoted p_{norm}) a new standard normal distribution is generated where ($x < p_{norm} = 0$, $x >= p_{norm} = x$). From this distribution, the skewness γ_{norm} is calculated. The goal now is to find which transformation factor minimizes the difference between γ_{norm} and the original data skewness. This was done for each station separately and for all aggregations (minutely to daily). Eventually, an average transformation factor from all stations (denoted hereafter as λ) was derived for each temporal aggregation. The results of this procedure can be seen in Table 3.1 and Figure 3.2.

Table 3.1: Average Box-Cox transformation factor (λ) per temporal aggregation.

Temporal duration [min]	**60**	**120**	**180**	**240**	**360**	**720**	**1440**
Transformation factor λ	0.097	0.155	0.219	0.262	0.318	0.427	0.499



(a) Skewness values

(b) Transformation factor

Figure 3.2: Left panel shows the skewness values along the mean and the median of all stations before and after transformation. The right panel shows the average transformation factor (λ) derived from all stations and for each temporal aggregation.

Once λ was calculated, the original precipitation data were transformed as in equation 3.13, and in the newly truncated normalized space, the following approach was implemented to find outliers in the precipitation data over several temporal resolutions.

3.3.2 Ordinary Kriging

Kriging was first introduced by the French mathematician Georges Matheron in his book 'Le Krigage universel' (*Matheron*, 1969). Ordinary Kriging (OK) is part of Geostatistics, which refers to multivariate statistics for neighboring values in space (*Cressie*, 1988). The main concept behind Geostatistics is the consideration of the data as spatially dependent random numbers with a variance that increases with increasing separation distance. As with other variance-dependent interpolation methods, OK aims to minimize the estimated variance at the unobserved location, providing the best linearly unbiased estimate.

The observed data at the corresponding locations are seen as a realization of the regionalized variable of the random space function. Since for every location u in the domain D there are many (infinite) random realizations $Z(u)$ describing each $Z(u)$ using its own distribution function $F_Z(u)$ is practically impossible. To simplify the problem, different hypotheses are considered. The first hypothesis, which is a central one in Geostatistics is stationarity. Simply put, the whole domain D is represented by a single distribution function regardless of the location of the points u in D. A further simplification is introduced with the second-order stationarity. For this, the expected value of the random function $E(Z(u))$ is constant over the domain, and the covariance of two random variables corresponding to two locations u_i and u_j depends only on the separating vector $h = u_i - u_j$ between the two points. This means that the covariance depends on the spatial configuration of points and not their exact values. The second-order stationarity hypothesis requires that a covariance function exists.

The previously mentioned hypotheses are followed by the following: a constant mean, a constant variance, and a spatial correlation that depends only on the distance lag vector h.

For any u in D:

$$E[Z(u)] = mE[(Z(u+h) - m)(Z(u) - m)] = C(h)C(h = 0) = Var[Z(u)] = const \tag{3.14}$$

The covariance function is a product between two terms, where each term is a deviation from the mean. The two variables are Z(u) and Z(u+h). In second-order stationarity, the correlation depends only on h and the covariance is a function of h, $C(h)$. The covariance function describes the dependence within the random space as a function of distance. Hence, the spatial correlation structure can be described by $C(h)$. Since for a separation distance of $h = 0$, the covariance is the same as the variance, the existence of a finite variance for D is required.

${\sigma_Z}^2 = C(h = 0)$ is finite.

This is often not the case in many natural processes (such as rainfall), where the variance increases with the distance. To solve this problem, the final hypothesis, known as the intrinsic hypothesis, was introduced. Same as the second-order stationarity, the expected value is constant all over the domain D, and the increment of the variance between two locations depends only on the separating vector h. The intrinsic hypothesis is a simplification of the second-order stationarity that is not constrained by the variance but by the variance of the increments.

The intrinsic random function is a simplification of second-order stationarity that does not need the variance. The assumption is that the variance of the increments Z(u+h)-Z(u) is constant.

For any u in D:

$$E[Z(u)] = m\frac{1}{2}Var[Z(u+h) - Z(h)] = \frac{1}{2}E[(Z(u+h) - Z(u))^2] = \gamma(h) \tag{3.15}$$

where: Z(u+h)-Z(h) is called the increment and $\gamma(h)$ the (semi-) variogram.

Since the mean is constant, the increments must have a mean of 0. The variance of the increment is denoted by the square root of the increments by taking the expected value. This is called the (semi-) variogram γ(h).

The (semi-) variogram is defined by:

$$2\gamma(h) = E[Z(u+h) - Z(u)]^2 \tag{3.16}$$

Where:

$Z(u+h)$ = observation value at location $u+h$
$Z(u)$ = observation value at location u
h = separation distance

The relation between the covariance function and the variogram is:

$$\gamma(h) = C(0) - C(h) \tag{3.17}$$

C(0) is the variance; in the intrinsic case, the variance does not exist; the variogram and the covariance are fully equivalent for second-order stationarity; however, for intrinsic, only the variogram is valid. That is why we focus on the variogram. The variogram reflects how the variance between the data increase with the separation distance between the data. C(h) starts at the variance of the points C(0), the covariance reflects the decrease of correlation with distance, and the variogram shows the increase of variability with distance.

Properties of the Variogram

For a given separating distance h between pairs of data, the variogram needs to have the following properties:

1. $\gamma(h=0)$ = 0 (always 0)
2. $\gamma(h) \geq 0$ (because it is a square, a variance)
3. $\gamma(h)$ = $\gamma(-h)$ (symmetrical, because it is a squared difference)
4. Variance of the increment is a function of h
5. Asymptotic behavior: there is a kind of limit of continuity
6. Nugget effect: for very small distances, there are differences
7. Anisotropy: $\gamma(h)$ may differ from one direction to the other

The experimental variogram is derived from the observed values and their spatial distribution. For each case, 30 observations were used. A theoretical variogram model was then fitted to the experimental one. Once the variogram was estimated, OK could be performed.

OK is a regionalization method to estimate an unknown value at a target location by solving a linear equation system by minimizing the estimation variance and maximizing the accuracy (no systematic error).

The estimation of the target location Z^* using the surrounding observations Z_i at the measurement locations n is defined by a linear estimation equation:

$$Z^* = \sum_{i=1}^{n} \lambda_i Zi \tag{3.18}$$

The Kriging estimation variance at the target location is formulated as:

$$\sigma^2(u) = -\sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i \lambda_j \gamma(ui - uj) + 2\sum_{i=1}^{n} \lambda_i \gamma(ui - u) \tag{3.19}$$

The weights λ_i are solved by guaranteeing the unbiased property of OK. Namely, the expected value of the estimation value should be equal to the expected value of the field Z. For this, the Lagrange multiplier μ is introduced, and the following linear system is solved:

$$\sum_{j=1}^{n} \lambda_j \gamma(u_i - u_j) + \mu = \gamma(ui - u) \qquad \forall i = 1, .., n \tag{3.20}$$

$$\sum_{i=1}^{n} \lambda_i = 1 \tag{3.21}$$

Note that γ refers to the (semi-) variogram which can also be replaced by the covariance function with the respected modifications. These two functions reflect the change in correlation as a function of the separating distance between the spatially distributed values. Note that for estimating the variogram, the cross-validated location is not used.

3.3.3 Outlier detection

Initially, a method was proposed for identifying outliers in groundwater quality data (*Bárdossy and Kundzewicz*, 1990). In this work, a similar approach was implemented for intense precipitation data and is extended by a verification of the results using additional data such as radar or discharge observations. For detecting suspicious values, the concept of jackknifing is used, a concept initially developed by *Quenouille* (1949, 1956). The concept is based on removing one (or each) value from the data and estimating it again. As the focus is on intense observations, the four largest annual values for every station are inspected. Each cross-validated value is estimated using the nearest 30 neighboring locations with valid observations. This is the minimal number of points needed for a reliable variogram estimation. The estimated value is only affected by the nearest points and their configuration. This is known as the shading effect in Kriging, i.e., the stations further away have smaller weights.

Since many possible faulty observations can only be detected at lower temporal resolutions, the procedure was applied over several temporal aggregations. For example, when looking at sub-daily and sub-hourly values, a single observation might not be unusual, but the accumulation of many values reveals suspicious sums. Furthermore, single events might occur on high temporal scales (e.g., hourly) and are not detected on lower aggregations (e.g., daily).

Each cross-validated value is estimated using the nearest 30 neighboring locations with valid observations. The spatial correlation structure is reflected by the variogram, which is derived in the rank-space domain and rescaled to the variance of the data. This allows for variogram calculation in a more robust manner (*Lebrenz and Bárdossy*, 2019). The target location is calculated by solving the Kriging equation, and the estimation variance is noted. For identifying unusual observations the ratio between the absolute value of the difference between The observed and estimated values and the estimation variance are calculated. This Criteria Ratio (CR) describes the relative agreement or disagreement between the observed value and the spatial surroundings for the corresponding time step. Larger CR values reflect high spatial-temporal disagreement, and low values denote greater agreement. Based on the CR value, different types of events can be identified, namely those occurring on a local scale with high CR values, and others on a regional scale with low CR values. Following *Bárdossy and Kundzewicz* (1990) a CR value of three is initially used to identify suspicious observations. The CR value is derived for every cross-validated event. Eventually, the CR value is related to all of the observed (interpolated) data, establishing the possibility of finding a suitable CR value for the identification of precipitation outliers.

$$CR_i(u) = \frac{|Z_i^*(u) - Z_i(u)|}{\sigma_i(u)} \tag{3.22}$$

Where:

$Z_i^*(u)$ = estimated value at location u and timestep i
$Z_i(u)$ = observed value at location u and timestep i
σ_i = kriging standard deviation at location u and timestep i

Since precipitation events occurring on a local scale might represent an actual small-scale event to validate the first or second case, the suspicious events are compared to the observed radar QPE or discharge values in the corresponding catchment. Despite having their own drawbacks, radar and discharge observations are used here as a qualitative decision-support tool.

The flowchart in Fig. 3.3 describes the implemented space-time precipitation outlier detection scheme.

3.3.4 Method testing using data corruption

For further testing of the reliability of the method, 20 stations without any detected outliers were selected, and their values (same events as before) were 'artificially' manipulated. The transformed observations of each target location were decreased and increased by several percentages (from 25 to 100 %) and the outlier detection method was tested. The results are presented in table 3.2. By decreasing the observed value until reaching a false zero observation, the method was able to identify on the hourly scale around 60 % and on the daily scale

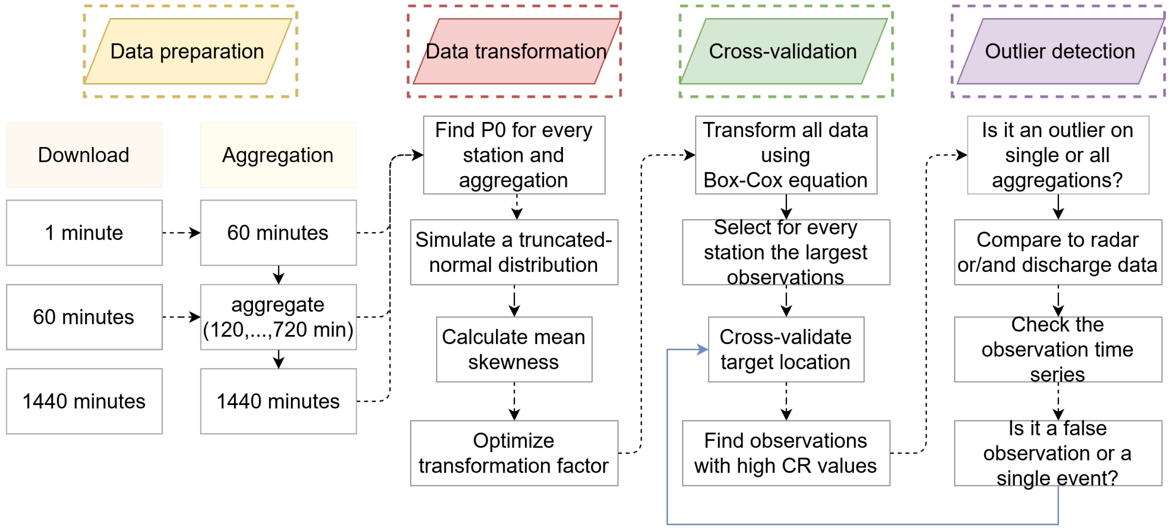


Figure 3.3: Flowchart summarizing the described method, starting with the data download procedure and ending with the identification of suspicious observations.

94 % of the cases as being outliers. On the other hand, by increasing the error value to up to 100 %, almost all values were detected on all temporal aggregations. This emphasizes the validity of the method, especially regarding identifying false high observations.

Table 3.2: Number of newly detected outliers after modifying by different percentages the cross-validated observations of 20 randomly selected stations.

	Temporal aggregation [min]	**60**	**120**	**180**	**240**	**360**	**720**	**1440**
	Number of events	**150**	**150**	**150**	**150**	**150**	**150**	**150**
	Minimum of the minima [mm]	5.12	5.16	5.17	5.13	5.26	5.6	5.17
	Average of all averages [mm]	11.05	12.67	14.41	14.41	16.85	19.8	24.03
	Maximum of the maxima [mm]	51.2	50.1	53.47	63.93	71.92	73.6	76.37
Percentage of error	**-100 [%] (false zero)**	88	115	102	125	100	124	141
	-50 [%]	10	29	38	41	33	46	65
	-25 [%]	2	3	13	9	8	8	4
	0 [%]	**0**	**0**	**0**	**0**	**0**	**0**	**0**
	+25 [%]	23	45	48	52	65	46	55
	+50 [%]	74	88	118	121	119	113	116
	+100 [%]	149	150	150	148	150	149	149

3.4 Results

3.4.1 Outliers vs. Single Events

The left and right panels of Fig. 3.4 represent the CR value versus the ratio between the interpolated and observed values. All values denoted in red have a CR value above 3. This Fig. allows for identifying the events that are of interest and relating the CR value to the interpolated and observed data. Note that the observed and interpolated values are in the original non-transformed space; only the CR values are calculated from the interpolation of the transformed values. The values in the plot having a ratio of interpolated to observed of 5 are values obtained when a neighboring station (or stations) simultaneously recorded an outlier (in this case, a false high observation). This leads to detecting a false outlier. This can be accounted for by running the method again after all neighbours have been checked.

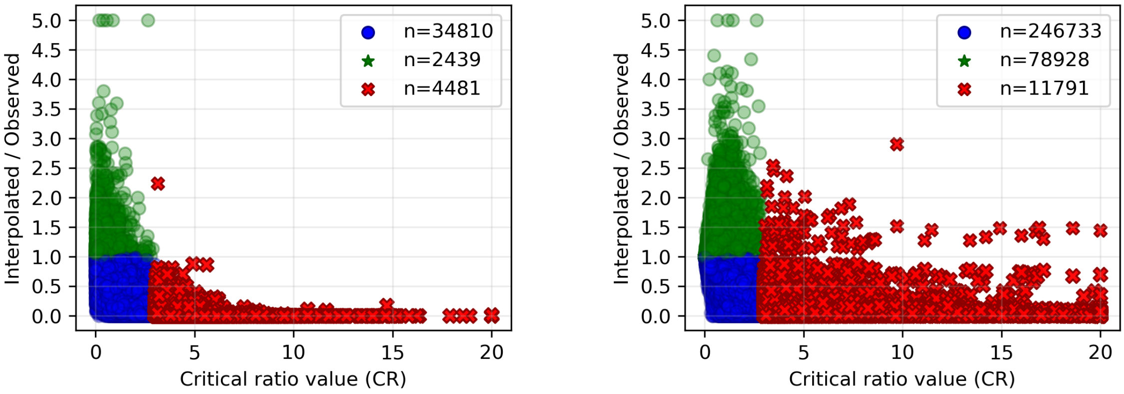


Figure 3.4: The left and right panels show for the minutely aggregated data and the daily data the CR values versus the ratio of interpolated and observed values, respectively.

In the left and right panels of Fig. 3.5 the Cumulative Distribution Function (CDF) from all investigated observations was calculated, and the location of the detected outliers is marked. These spread along the CDF, signifying that the approach is not focused only on high values but is also beneficial for identifying small values that highly differ from their neighboring observations. The left panel of Fig. 3.5 shows the results for the original hourly observations. The right panel shows those for the aggregated minute observations. By comparing the two, the quality control procedure of the DWD can be investigated. Spatial consistency is checked more intensively by the DWD for higher aggregated precipitation data (≥ 1 h) than for high temporal resolution data (e.g., 1 min). For example, in the hourly data, none of the largest sums (> 60 mm h) are detected as outliers, and only one observation is larger than > 80 mm h. In the hourly data based on the aggregated minute data, many values above 80 mm h exist and are mostly all detected as being suspicious. There are even unrealistic values with accumulated sums above 200 mm h, which can be caused by several faulty 'small' measurements or a few single large spikes in the data.

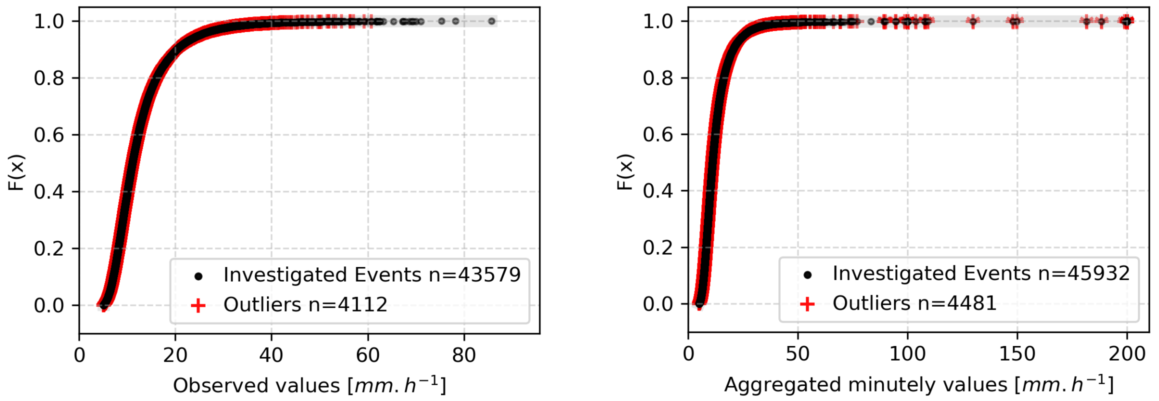


Figure 3.5: The left panel shows the CDF of all investigated hourly events, with the detected outliers marked in red. The right panel shows the CDF of all investigated hourly events, with the detected outliers marked in red. Note that an upper limit of 200 mm h^{-1} was set.

3.4.2 Selected case studies

The example in Fig. 3.6 shows the presence of unusual values in the minute data of the cross-validated station ($>$ 8 mm min^{-1}). In the right panel, the radar data for that hour do not show such a high-intensity event above the investigated location. Hence, the detected outlier is categorized as a false observation. In Fig. 3.7, the example shows a similar case in the minute data, but the radar image confirms the occurrence of the event. This observation is then marked as a single event.

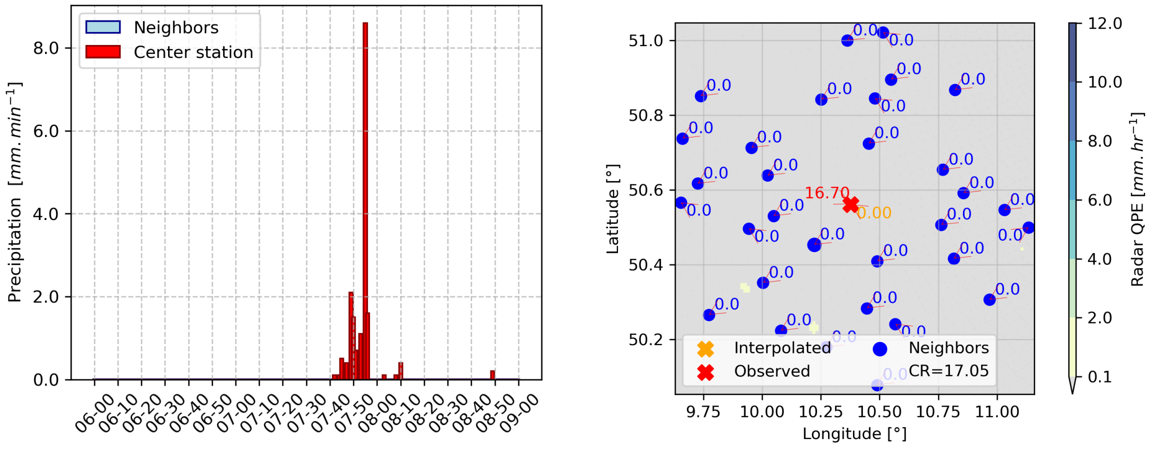


Figure 3.6: The left panel shows the time series of the target (in red) and neighboring observations (in blue). The right panel depicts the values at the neighboring stations (in blue), the observed value (in red), the estimated value (in orange), and the Radolan-RW QPE data for that hour.

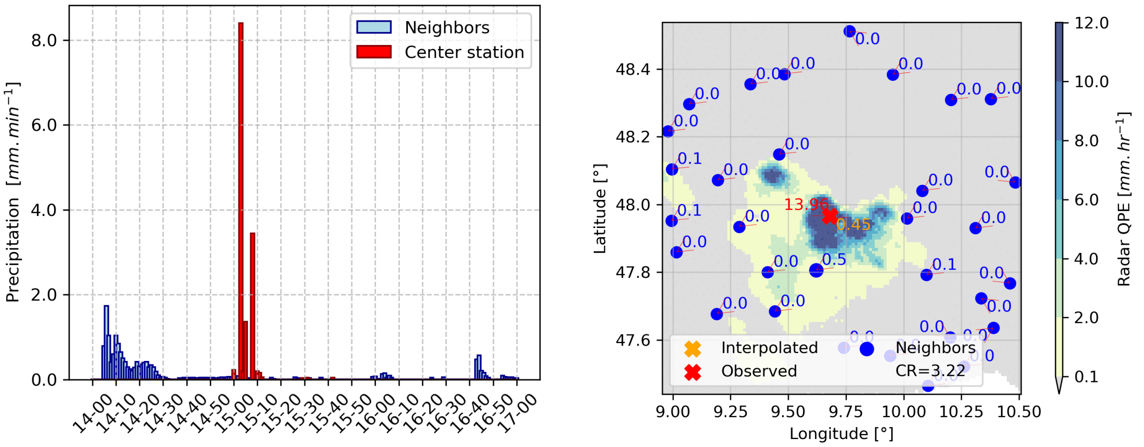


Figure 3.7: The left panel shows the time series of the target (in red) and neighboring observations (in blue). The right panel depicts the values at the neighboring stations (in blue), the observed value (in red), the estimated value (in orange), and the Radolan-RW QPE data for that hour.

Discharge data from small headwater catchments in the federal state of Bavaria with one (or many) rain gauge stations within the catchment were analysed. If a rain gauge observation was identified as suspicious, the discharge values for the next few hours following the event were checked. An example of this is shown for the upper Pegnitz catchment, which is located in the northern part of Bavaria. Panel (a) in Fig. 3.8 shows an hourly outlier observation that resulted in a reaction in the corresponding headwater catchment. On the other hand, panel (b) in Fig. 3.8 shows the opposite case, i.e., an hourly outlier that did not cause any reaction in the Pegnitz catchment.

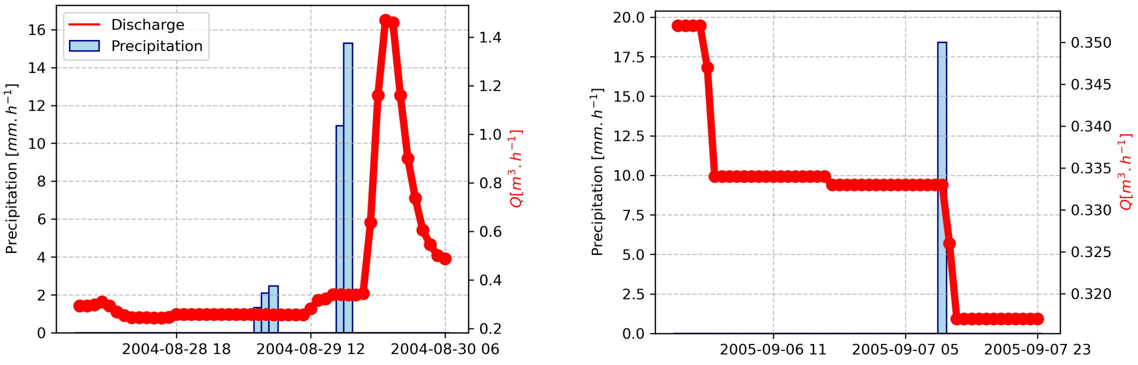


Figure 3.8: Two examples are shown for the observed discharge and precipitation data (+/- 1 day) for detected outliers with (a) a discharge increment and (b) without a discharge increase.

3.4.3 Results over all stations and aggregations

The method was applied over several temporal aggregations (hourly to daily), and events that were suspicious over single or several aggregations were identified. The result of this can be seen in Table 3.3. The diagonals show events that are common over the corresponding test and reference temporal aggregation. Some observations are only suspicious until a temporal aggregation is reached or exceeded, beyond which they are not detected anymore. The result of this can be seen in the values above and below the diagonals in Table 3.3. Another factor that affects the number of outliers is the number of active stations and the corresponding device quality. To investigate this, the number of active stations versus the number of detected outliers for the corresponding hour is visualized in Figure 3.9. The red curve in Fig. 3.9 represents the ratio of detected outliers to the number of active stations (for every hour), which is shown by the blue curve. As the number of active stations increases, the number of detected outliers decreases which is an indication that the quality of the observations is improving with time. In the right panel of Fig. 3.9 the effect of seasonality was inspected. The detected outliers were grouped by the month in which they occurred. The results show that in the summer period, the number of detected outliers is much larger than in the winter period. The most convectional rainfall processes occur in the summer period, leading to more small-scale single events. Finally, the percentage of outliers in the investigated events of every station for the hourly aggregated data is presented in the right panel of Fig. 3.10. The map does not present any clear structure related to elevation and topography. Moreover, the map shows that outliers can happen everywhere, meaning this is not a systematic problem.

Table 3.3: The diagonals show the number of unique days with identified outliers. The values above the diagonals reflect the number of different days between the reference and test aggregations. For example, according to the test, there are 358 days in the reference 60-minute aggregation that are not in the 120-minute test aggregation.

		Test aggregation						
		60 min	**120 min**	**180 min**	**240 min**	**360 min**	**720 min**	**1440 min**
Reference aggregation	**60 min**	**1581**	358	392	414	498	762	898
	120 min	218	**1441**	210	237	354	646	787
	180 min	344	302	**1533**	240	341	657	825
	240 min	437	400	311	**1604**	329	657	837
	360 min	539	535	430	347	**1622**	559	762
	720 min	771	795	714	643	527	**1590**	439
	1440 min	889	918	864	805	712	421	**1572**

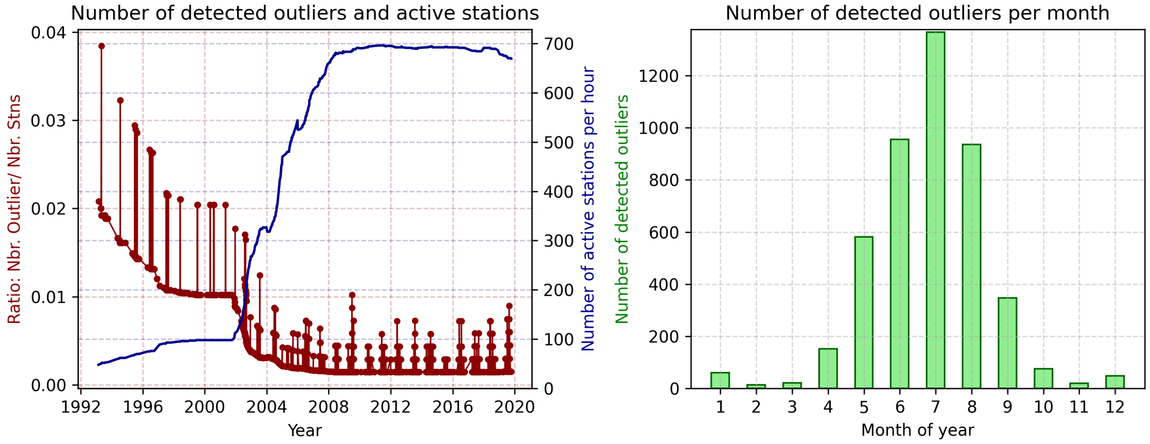


Figure 3.9: Left figure shows the number of hours with outliers within the investigated hourly events (aggregated from 1 minute) of all stations per year. The right figure shows the number of detected outliers within the investigated events for every month within the hourly data using a CR value of 3.

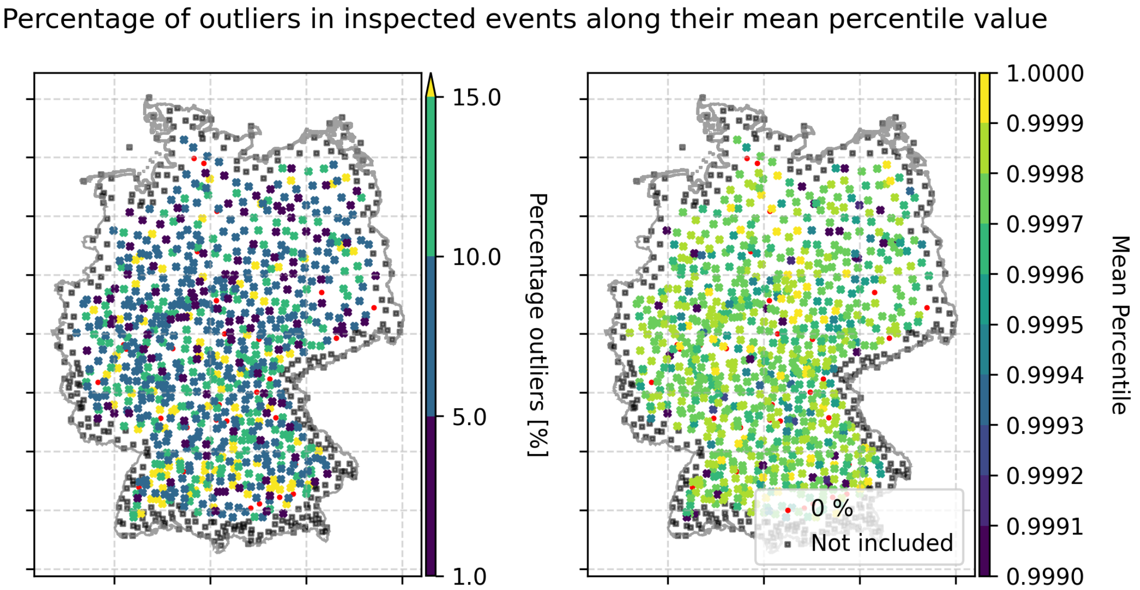


Figure 3.10: The left panel shows the percentage of outliers in the investigated events of every station on an hourly scale. The right panel shows the average percentile of the detected outliers.

3.5 Summary

The presence of outliers in the data affects the results of any analysis dramatically. In the event that these outliers are false observations, an incorrect evaluation and, consequently, design values are derived. On the other hand, if the outliers present a correct observation, disregarding them may lead to many losses and an underestimation of critical structures. Therefore, the previously presented method is essential to finding reliable outliers in intense point observation.

One key aspect when working with precipitation is the high spatial-temporal variability and the zero-inflated nature of the variable. To handle the first aspect, the procedure was applied over several temporal aggregations, which helps in handling convective-driven events. Such events are often correct but represent an outlier, in this case, a single event. To cope with the presence of the many zeros, a Box-Cox transformation with a suitable parameter was implemented. The procedure was applied to intense observations and revealed many outliers that are false measurements.

Furthermore, a verification step was undertaken, in which data were artificially corrupted in both directions and inspected. The method revealed success in detecting much of the corrupted data, especially as the value of the added error increased. Many events are detected as outliers, often because they are single events occurring in the warm season. Their occurrence or falsehood was validated by weather radar and/or discharge observations. Finally, the flagged data are gathered and kept aside for further inspections.

Still, there are some possible modifications. For example, on the sub-hourly scale, the effects of advection and correlation between subsequent time steps are more predominant and should be incorporated in the Kriging procedure. Another aspect is the directional dependency that can be seen in some events, where rainfall is occurring on a local scale or within a certain frontal system. This leads to detecting a false outlier due to the high level of disagreement between the target and neighboring locations. These cases could be further handled by including anisotropy in the interpolation method.

4 Spatial extent from point observations

The goal of this chapter is to present several methods that were developed to understand the spatial extent of extremes using point observations. The focus was laid on the behaviour of neighbouring locations within an intense event.

The following questions are related to capturing the spatial extent of extremes using point observations:

1. How well can the spatial extent of extremes be captured and described?
2. What are the limitations of using point observations?

4.1 Literature review

Investigating the spatial extent of rainfall extremes was traditionally done using point observations. Different statistical measures describing the correlation structure between spatially distributed data were used. One of the most straightforward approaches is to calculate the correlation values (Pearson, Spearman, or Indicator correlations) between pair-wise data and plot them against the separating distance. A typical behaviour shows a decreasing correlation with increasing distance. When done for event-based data, the procedure was seen to be reliable on daily and larger temporal aggregations, where a clear dependence between correlation and separating distance values was present (*Gunst*, 1995). However, for sub-daily and sub-hourly data, this dependence measure showed a more chaotic behaviour and was only relevant for small-scale regions (*Serinaldi*, 2008). One of the reasons for this is the time lag factor (advection) that plays a major role in sub-hourly scales. To cope with this challenge, *Burauskaite-Harju et al.* (2012) used the cosine similarity measure within a moving temporal window approach. By doing so, it was possible to inspect the spatial dependence of sub-daily events while accounting for the time lag of precipitation occurrence at subsequent stations (movement of the front). The presented method was seen as more successful for low-intensity events, but as the latter increased, the uncertainty of the results became larger, and the similarity measure dropped quickly with separating distance.

In the work done by *Touma et al.* (2018) the spatial extent of daily extreme precipitation events over the United States was investigated using daily station data with long records. The methodology was based on indicator semi-variograms that were derived from intense daily observations. The semi-variogram was used to model the dependence structure between neighbouring observations, from which the extent of the extreme events (defined as length scale) was derived. The analysis showed a clear relationship between the length scale and the corresponding region or season. A similar approach was presented by *Tan et al.* (2021) using information from semi-variogram models derived from extreme daily events on a global scale. In the work done by *Marzban and Sandgathe* (2009) the variogram is applied to assess the quality of gridded forecasts produced by numerical weather models. The aim was to identify if the models could reproduce the spatial structure of the observed fields derived from weather radar reflectivity fields. For simplicity, only isotropic variograms and one variogram per field were considered. These were derived by randomly sampling 40 samples from each gridded dataset. Their results showed the possibility of employing variograms as a tool to compare spatial fields, but with limitations. For instance, anisotropy was mentioned but not implemented, and by using all values (zeros as well), the variograms and subsequent results changed. Moreover, a single variogram was assumed to represent the entire field, reducing the spatial variation and leading to smoothing effects. Rainfall fields present high spatial heterogeneity, and describing them by the variogram is often not enough. As mentioned by *Diggle et al.* (2003), spatial fields derived from interpolation techniques such as kriging, where the spatial correlation structure is represented by the variogram, are only reliable if the focus is only on the mean and variance. Aspects such as asymmetry are not accounted for. In addition, in the work done by *Teegavarapu and Chandramouli* (2005) different advantages and disadvantages of several spatial interpolation methods used to estimate missing data at an observed location were discussed. They point out two scenarios where such interpolation techniques fail. The first case is where precipitation did not occur at the target location but occurred at all (or some) of the surrounding stations. In this case, the estimated value will be positive, whereas the observation was zero precipitation. The second case is the opposite, namely, precipitation only occurred at the target location and all surrounding stations recorded zero precipitation. For this situation, the estimated value will be zero.

Based on rank statistics, copulas were previously used to describe the dependence between spatially distributed data. An advantage of using copulas is the possibility of investigating the dependence by separating the distance between values of similar magnitudes, for example, between the high values. To account for the relationship between high and low values, the asymmetry function developed by *Bárdossy* (2006) could be used. Spatial copulas were implemented in several previous studies for geostatistical problems, for example, in *Bárdossy and Li* (2008) and *Guthke and Bárdossy* (2017). They are advantageous compared to kriging approaches, especially regarding the estimated error maps.

Still, one of the challenges is that at high temporal resolution, for example, on hourly and sub-hourly scales, the variability within a rainfall field can be substantial. On one hand, it cannot be correctly captured by sparse rain gauges, and on the other hand, it cannot be well replicated. Moreover, all interpolation schemes aim at minimizing the error term by solving a linear equation system. This leads to a smoothing effect and an underestimation of the true maxima and true variability.

In the following chapter, using the rain gauge observation, several events were selected across different temporal aggregations. For every event, the correlation between neighbouring observations was investigated, and the probability of a rainfall maxima within an event not occurring at the centre location was calculated. This was in accordance with analysing point maxima and the relation between point and areal maxima.

4.2 Correlation of neighbouring data vectors

The first step of the analysis was to derive a set of events to be investigated across several temporal aggregations. This was done by selecting for every temporal aggregation (5 minutes to daily) observations above a minimal threshold (as in Table 4.1) that occurred at a single or several locations simultaneously. Eventually, all selected events were reduced to a 5-minute resolution and analysed further on (also in chapter 5). These time intervals were referred to as *significant time steps*.

Table 4.1: Minimal threshold for every temporal aggregation.

T [min]	5	15	30	60	90	120	180	240	360	720	1440
P [mm]	1	2	3	6	9	12	15	20	25	35	50

To investigate how rainfall events behave in space, the spatial correlation between neighbouring stations was calculated. Different events were selected on a daily scale using aggregated sub-daily data based on significant time steps. For a certain location, all neighbouring locations with simultaneous observations were identified. The Pearson and Spearman correlation measures were calculated between the data vector (in this case, 24 observations) of the target and each neighbouring location. This was repeated for each location and several events. By applying this simple procedure, the spatial correlation between neighbouring locations for an intense event could be visualized. The result indicated that for some events, all neighbouring locations had a high correlation with the target (centre) location. However, for other events, only a few neighbouring locations had a high correlation. This was an indication of the presence of anisotropy in rainfall events. A simple definition of anisotropy can be stated as: direction-dependent correlation. Hence, if an interpolation technique were to be applied, anisotropy should be accounted for. Moreover, two types of events were identified: those that occur on a local scale (known as convective precipitation events, with short spatial continuity) and those that occur over a larger spatial domain (known as stratiform precipitation events, with long spatial continuity). The examples in Fig. 4.1 show the correlation values for two events. Panel (a) represents an event where there was high dependence between the data vector of the central location and most of the neighbouring locations. In this case, no anisotropy was noticed. Panel (b) is another event and reveals the presence of anisotropy (direction-dependent correlation). The high dependence between neighbouring and centre values is only limited to a few stations and follows a certain orientation. Similar results were acquired for the Spearman correlation.

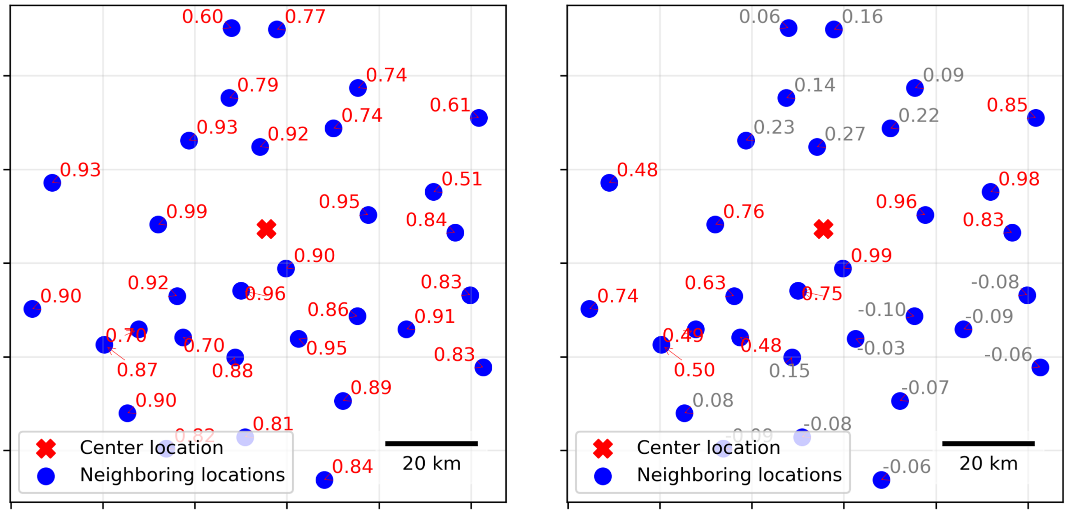


Figure 4.1: Pearson correlation values for two different events between the centre location and each neighbouring data vectors (24 hourly values). The gauges displayed in red are those with a correlation value above 0.5 with the centre location. The gauges in gray have a correlation below 0.5 with the centre location.

4.3 Correlation clouds

The spatial dependence structure between point gauges was investigated by deriving several correlation clouds. These were related to the separating distance and were derived by calculating the correlation between the time series at each location and all other locations. For example, in Fig. 4.2 the Spearman and Indicator correlations were calculated from the hourly rain gauge data in the radar area of Hannover for the period April-September. As the separating distance increases, the correlation values drop. The Spearman correlation in panel (a) shows the largest spatial continuity and the smallest scatter.

As described in Chapter 2 the latter was calculated between the ranks of the data and has a small sensitivity towards outliers. In panel (b), the Indicator correlation shows that after the separation distance of 50 km, the correlation value was negligible. Hence, on the hourly scale, the spatial extent between the high rainfall values (above the 99 percentile) is less than 50 kilometers. Such analysis of the correlation clouds offered insight into the spatial dependence structure between spatially distributed observation time series.

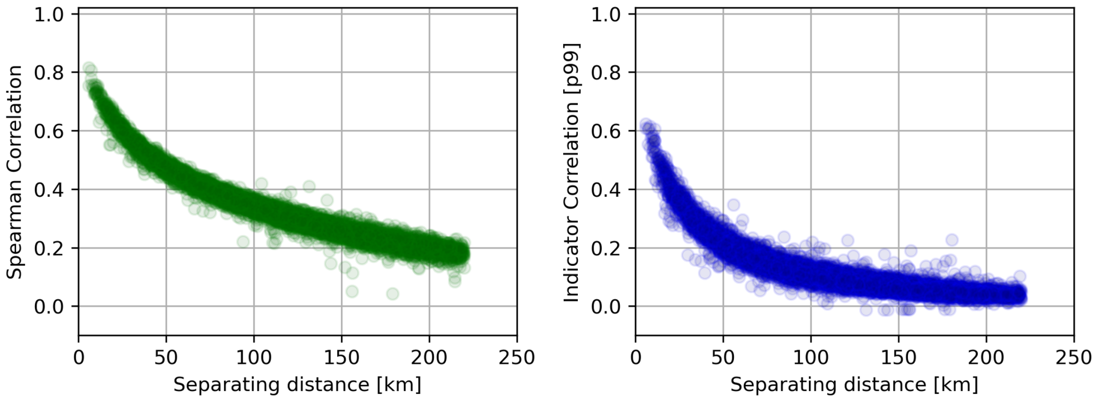


Figure 4.2: Spearman (panel a) and Indicator (panel b) correlation clouds calculated from pair-wise correlation between each and all other DWD rain gauges in the region of Hannover for the period April-September.

4.4 Local occurrence of intense events

After detecting the different types of rainfall events, the DWD sub-hourly station data set was divided into two equally spatially homogeneously distributed data sets. This was done by selecting for every station its nearest second neighbour. These datasets were denoted as Group A and Group B. For every station in Group A and for each event within the significant time steps, the neighbouring observations from Group B were identified. The following question arose: how often does it occur that a neighbouring station in group B has a higher observation than the centre location in group A? Answering this question was useful to quantify how often did the centre location A had the most intense value. An example of this procedure for one event and station can be seen in Fig. 1.1. The left panel shows that none of the neighbouring stations had a value larger than the centre one. In the right panel, 3 neighbouring stations had larger values than the centre one. In Fig. 4.4 the results for temporal aggregations from hourly to daily are shown.

Only the nearest neighbours (within 20 km distance) were considered, and the probability that one of the neighbours had a larger value than the centre one was calculated. For the hourly scale, the probability is around 14 % and for the daily scale, it is 39 %. This reflected the fact that intense observation occurring at a single location might not be the largest within the area of interest (in this case, around 1256 km^2).

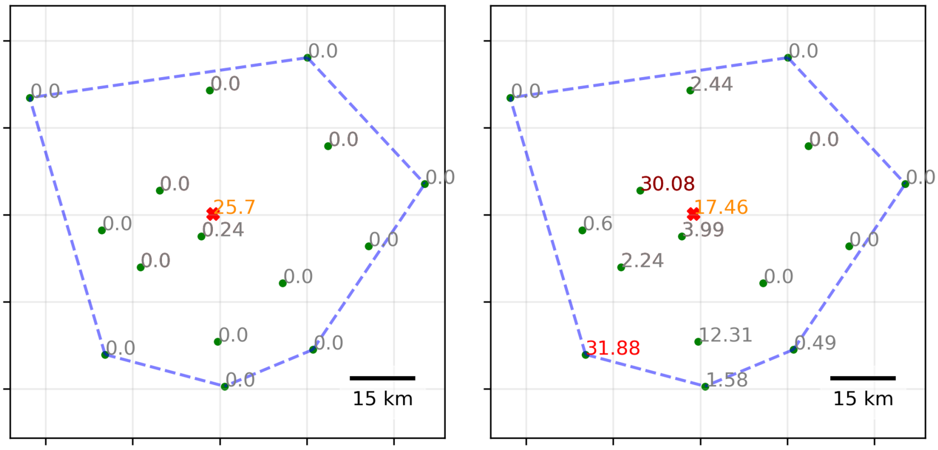


Figure 4.3: Left panel shows an hourly intense event observed at the center location, where most of the nearest neighbors recorded much lower values. The right panel shows for the same location a different hourly event where several nearest neighbours had a larger value.

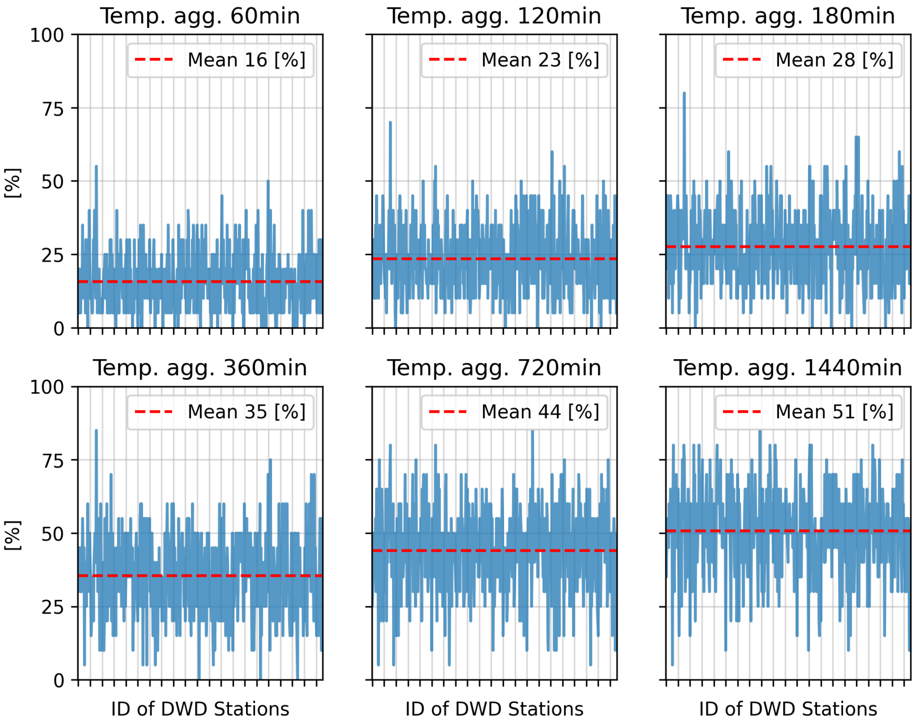


Figure 4.4: Average number of intense events where at least one of the nearest neighbours had a value higher than the center location. Each panel corresponds to a duration starting from hourly to daily. The x-axis represent the DWD stations and the y-axis is the average percentage. Note that the mean over all stations is displayed by the red-dashed line.

4.5 Clustering of step functions and variograms

Another aspect was to investigate how a rainfall event over a certain location evolved in space. To that end, for every event at a certain location, different circles (areas) with an increasing radius were constructed. For every area, the neighbouring observations enclosed by the circle were identified, and their mean value was calculated. The ratio between the mean of the neighbours and the central location was then calculated and assigned for this area. The procedure was repeated until reaching a radius of up to, for example, 200 kilometers. Eventually, this spatial step function was visualized, showing how the ratio between areal and centre rainfall within an intense event evolved in space. This was done for several events at a certain location, all other locations, and several temporal aggregations. For the event in Fig. 4.5 the calculated step-function shows an increase in the precipitation mean with increasing radius before starting to decrease as the radius (area) increases. By calculating these step functions, the assumption behind area-reduction factors, namely that with increasing area, the areal rainfall decreases, could be investigated. Indeed, for many events, the step function showed an increase with increasing area size. Moreover, from the cloud of all step functions, either for a certain location, a certain aggregation, or for all locations, several shapes of step functions were identified. A K-mean clustering algorithm was used to cluster the step functions into four groups. The number of representative clusters was derived using the elbow method (*Robert*, 1953). The implemented algorithm for step-function calculation was as follows:

1. For a given time resolution:
    a) For every station
        i. Construct the empirical distribution function.
        ii. Select all data above a certain percentile (ex: 99.9).
        iii. For a set of radius (for example, 50 km with a 2km increase).
        iv. Calculate for each event the normalized mean of the neighbours in the circle.
        v. Create a step function that describes the ratio of the areal mean to the centre location with an increase in the radius.
        vi. Cluster all Step functions for this station.

The same was done by calculating the empirical spatial variogram for each location and all selected events. The variogram is a function showing how dependence changes with distance. Depending on the spatial extent of the events, the variograms were different. Once repeated for all locations and aggregations, K-mean clustering was applied to the variogram clouds. Four types of variogram shapes were identified and used within the interpolation scheme for further analysis. The step functions and variograms were derived for all locations and several events and temporal aggregations. A K-mean clustering algorithm was used to cluster the derived functions. For example, for the daily aggregation, the results can be seen in the left and right panels of Fig. 4.6, respectively. To find the most suitable number of clusters, the elbow method was used. A clustering number of four was identified as being the most suitable. The clustered functions (step functions, variograms) were back-analyzed to find the corresponding type of event and spatial extent. To visualize the events, a spatial interpolation method, namely ordinary kriging, was applied to generate gridded rainfall values.

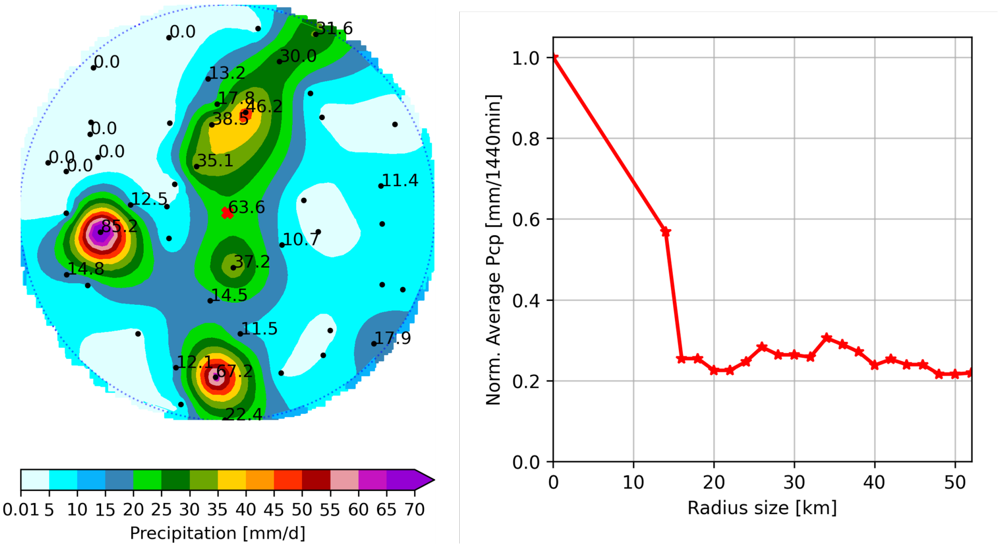
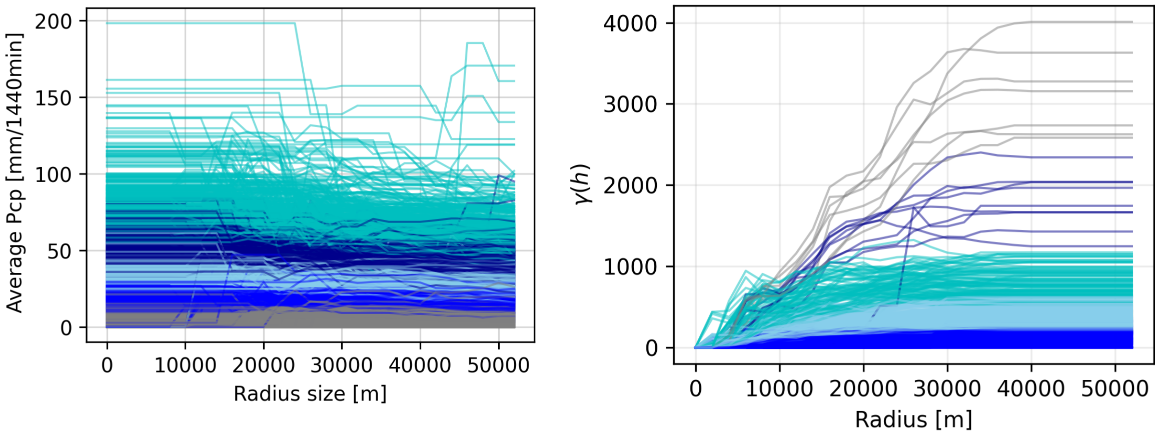


Figure 4.5: Panel (a) shows the interpolation map of the observed data over a 1 km grid. Panel (b) shows the derived normalized step function.



(a) Clustered step functions

(b) Clustered empirical variograms

Figure 4.6: Clustering of step functions and variograms for characterizing the spatial extent of intense daily events, derived from all stations.

4.6 Depth functions

In this paragraph, depth functions are briefly explained. A short theoretical background and some practical examples for deriving the depth of a data value are presented. This approach was used to identify unusual events in the precipitation station data set.

4.6.1 Statistical data-depth

Statistical data depth emerged around the year 1970. The main idea is to derive quantiles in a multivariate space for a multivariate observation defined by a random vector X distributed according to a probability distribution function P. In traditional univariate statistics, quantiles are derived via a total ordering of the data. A quantile of a certain point is calculated by finding which observations are located on its left or right. The central location is referred to as the median in a univariate space. The data depth concept attempts to generalize this in a multidimensional space. Namely, for any point x in d-dimensional space and any distribution P, assign a depth value for this point x with respect to the distribution P, where the depth value characterizes how centrally x is positioned with respect to the mass P (*Nagy et al.*, 2019).

$$D : \Re^d \times P(\Re^d) \to [0,1] : (x, P) \to D(x, P) \tag{4.23}$$

The example in panel (a) of Fig. 4.7 shows the calculated depth for a random sample corresponding to a normal distribution function. The points denoted in orange correspond to 10 percent of the data and are geometrically located in the center of the data space. A generalization of the median for this data set is done by considering the single observation that has the highest depth value with respect to this sample. The median is defined as the deepest point with respect to the given random sample. By starting from the center of the distribution and by moving outwardly in any direction, the depth decreases until the points on the boundaries are reached (points in red on Fig. 4.7). These points have the least depth (d = 1) with respect to this sample. The statistical data-depth calculation can be seen as a data- (or a distribution)-dependent ordering of points in a d-dimensional sample space; it is a non-parametric multivariate tool that requires no assumption on the distribution of the data.

To transfer the data depth calculation to space, the depth was calculated using common observations between neighboring locations. For every combination of locations, the depth of values for every time step with respect to other values was calculated. An example of the spatial locations can be seen in panel (b) of Fig.4.7. An example of the results is presented in the coming sections.

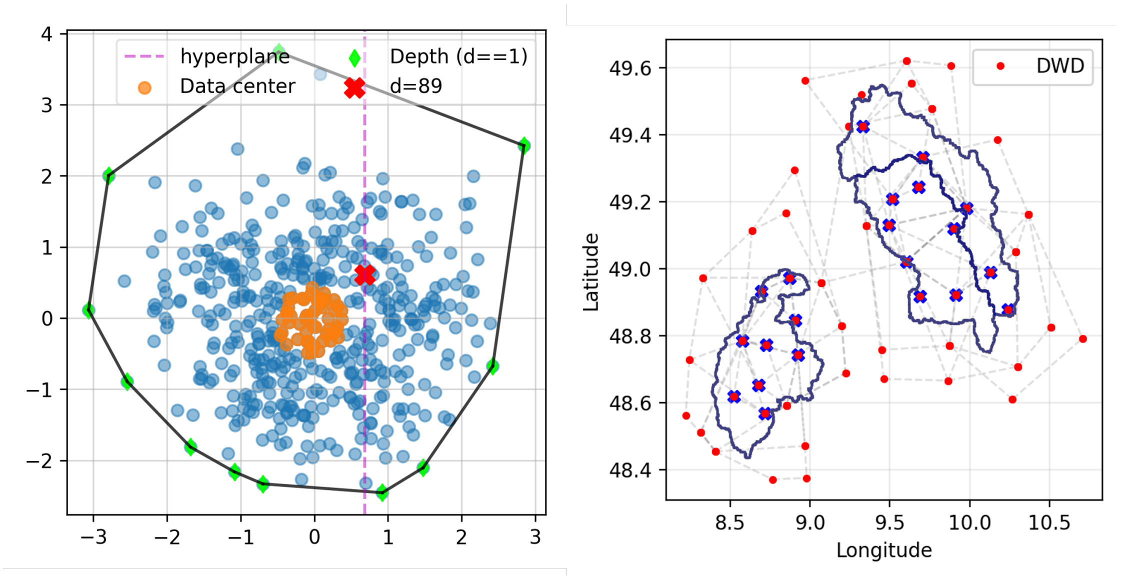


Figure 4.7: Example of depth calculation in a 2D space is seen in panel (a). Panel (b) shows a configuration of spatially clustered stations.

4.6.2 Data depth calculation

There are several functions to calculate the data depth, but the fundamental approach was introduced by Tukey in 1975 (*Tukey*, 1975). All methods aim to find the depth (or centrality) of a given point x with respect to the data space.

The data-depth function should fulfill the following properties:

1. Affine invariance: The depth value is independent of the coordinate system (or the choice of axes).
2. Maximality at the Centre: The maximum of the depth function is always centrally located.
3. Monotonicity is relative to the deepest point; the depth follows a monotonic decrease while moving in any direction away from the center point.
4. Vanishing at Infinity: If a point x reaches infinity, its corresponding depth value should reach zero.

In the following work, the Tukey data-depth function known as half-space depth was used. A depth of a point x with respect to a measure P is defined as the smallest probability of a closed half-space that contains x on the boundary. Mathematically defined in equation 4.24. Hence, the depth of a point x is defined by the minimum number of points lying on one side of a hyperplane passing through x. This is eventually calculated over all hyperplanes. Points located outside the convex hull have a depth of 0, and the depth increases as one moves towards the center.

$$D(x; P) = \inf P(H) : H \text{ is a closed halfspace with x} \in H \tag{4.24}$$

Statistical data-depth was used, for example, by *Yulizar and Bárdossy* (2020) to find unusual events occurring at a single location or several locations simultaneously. *Bárdossy and Singh* (2011) used the data-depth method to find regional parameter estimations for several catchments. In the work of *Bárdossy and Singh* (2008), data-depth was introduced as a tool to find robust parameter sets for hydrological modeling. Still, the data-depth function was not used for the spatial extent of rainfall. The problem of many dimensions often leads to difficulties in correctly estimating the depth. A possibility is to use a dimension reduction method such as principal component analysis.

The first analysis was conducted on neighboring pair-wise locations. The station data were divided into two data sets with equal length, either as past and present data or as an example for hourly data, the first half-hour and the second half-hour. The cross-depth of the first time series in the second and the second in the first were calculated. Fig. 4.8 shows the cross-depth of two neighboring stations. The highlighted points are those with low depth ($d < 10$) and are referred to as unusual events. The focus is on the simultaneous occurrence of unusual events, for example, the event in orange with a depth of 7. The right-hand side of Fig. 4.8 shows the time series of the investigated pair-wise stations (curves in red and blue) and neighboring stations. The event shown here is seen as unusual at both stations simultaneously.

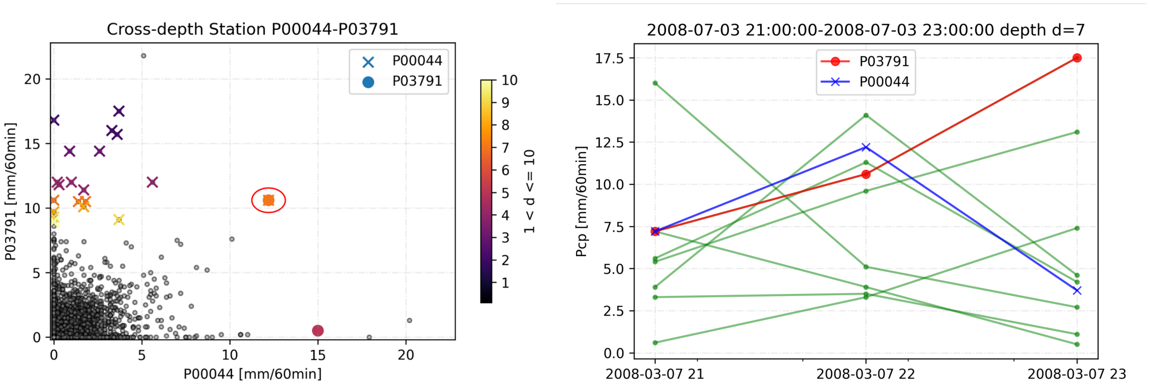


Figure 4.8: Panel (a) shows the cross-depth of two neighboring stations. The highlighted points are those with low depth ($1 < d < 10$) and are referred as unusual events. The time series for the point in red with a depth of 7 can be seen in panel (b).

4.6.3 Simultaneous or single occurrence of unusual events

In order to apply the depth function to find unusual observations over several locations, the following procedure was implemented:

1. Start by station A.
2. Select its nearest 8 neighbours.
3. Find common time steps and create a 9-dimensional data set.
4. Replace all zero values with a small random number.
5. Select all time steps where the maximum (the sum or the average) was above a minimal threshold (for example, 30 mm/day).
6. Calculate the cross-depth of the previous values.
7. Find events that have a low depth (for example equal to 2), namely unusual events.
8. Move to the first neighbor of A and repeat the procedure.
9. Save again all identified unusual events.
10. Repeat the procedure until all stations were mapped (namely all 8 neighbors of each neighbor).
11. Final result: a data frame with the time steps of the unusual events and values over all selected stations.
12. Test the procedure for a headwater catchment and check the discharge observations.

The method was tested for daily and hourly observations. The aim was to identify simultaneously occurring spatially unusual events. Unusual events over a cluster of neighbouring stations were first identified. An example can be seen in Table 4.2. One can distinguish two types of events, namely one over a portion of the stations (example, the first row) and one over all locations simultaneously (example, the third row).

Table 4.2: Identified unusual events within a cluster of 9 stations.

	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10
05/06/2011 14:00	0	0	0	0	0	0	4.5	36	38.7	5
30/05/2008 17:00	34.6	38.7	15.8	9.1	15.6	0.1	0	0	0.1	0
30/05/2008 18:00	13.1	3.6	2.7	3	6.3	7.7	23.2	23.4	12.4	25.3
30/06/2012 21:00	16.2	18.2	17.5	21	38.7	0	0	0	10	0
27/08/2010 00:00	20.9	6.1	6.3	14.9	17.4	25.8	16	23.6	19.3	13.8

Once the process was repeated for all configurations, a time series of unusual events was constructed and sorted by the sum over the stations of each time step. In Fig. 4.9 two daily events are showcased with very large daily sums. The first event in panel (a) presents a frontal-driven case, spreading continuously from the north-east to southwest of Germany with the highest values in the center of the front. The second event is a convective one, with rainfall cells scattered from north to south with dry spots and varying precipitation depths. Both events were detected as being unusual with low depth. The event in panel (a) occurred on the 6th of June, 1981, the latter was seen as the rainiest year in Germany and the event in panel (b) was part of the Elbe flood, which was seen as a century flood.

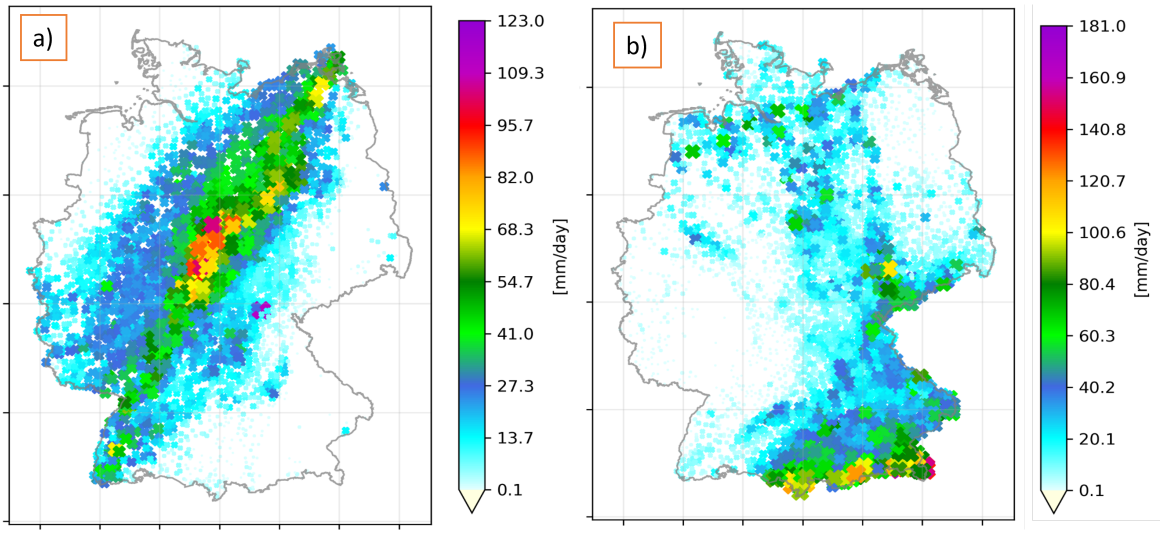


Figure 4.9: Panel (a) and (b) show two identified unusual events within the daily stations, showcasing a frontal and a convective event, respectively. The frontal precipitation event occurred on the 03-06-1981 and the convective event on 11-08-2002.

To be able to differentiate between the detected unusual events, discharge observations for three headwater catchments were used. Following the concept that the catchment serves the role of a rain gauge. The stations falling within and around the catchments were selected, and the previously described procedure was implemented for hourly and daily values. Fig. 4.10 shows a detected unusual event using the gauges available in the three headwater catchments. The precipitation values are presented in panel (a), and the reaction in the discharge of the three catchments is seen in panel (b). The event is seen as unusual because only two catchments were hit by a strong rainfall event, while the third one (Enz catchment) was not. Hence, from a spatial point of view, two neighbouring regions were hit asynchronously by a strong rainfall event. Note that in panel (b), the precipitation and discharge values are displayed for the event hour (at 18 p.m.) and the following 20 hours.

The advantage of applying such a procedure is in detecting spatially distributed unusual events that differ from traditional extremes. Although traditionally applied in the field of mathematics, the data-depth function is a useful tool for investigating and detecting unusual events that lie between the center and outer bounds of a data space.

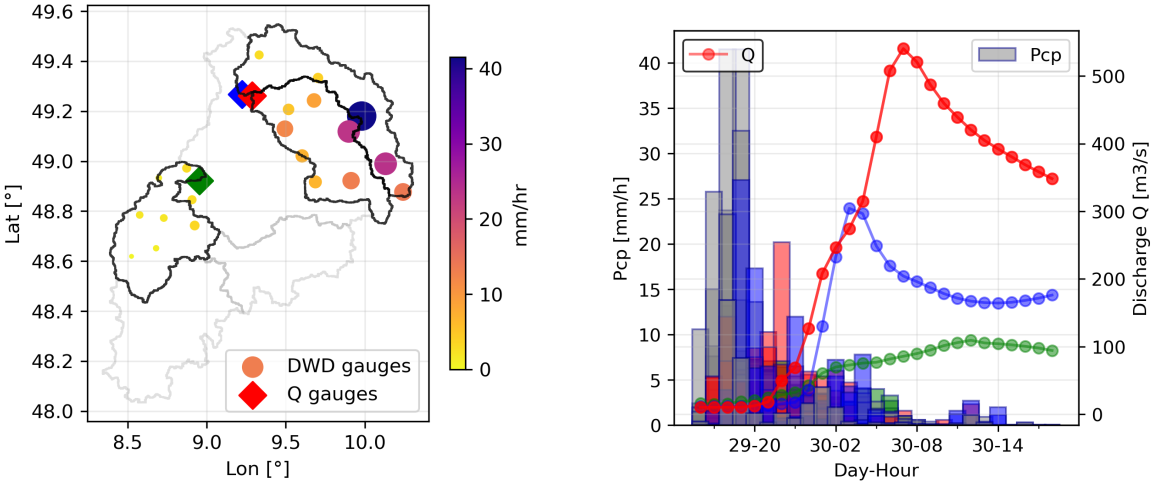


Figure 4.10: Panel (a) shows the recorded DWD values in three headwater catchments. Panel (b) displays the discharge values in the three catchments.

4.7 Limitations of using point observations

4.7.1 Detecting extremes in a given area, independent of their location

To highlight the challenge of catching extremes using only point observations, the following analysis was conducted: The area in question will be the whole country of Germany, and the DWD warning levels (presented in Table 4.3) defining heavy and severe rainfall thresholds were used as target values. For every station, the number of observations for every duration exceeding the three different warning levels was calculated. Eventually, the number of observed values was averaged over the number of stations with similar record periods. The first 3 rows of Table 4.4 show the results. One can see that, independently of the event location, with an increase in the number of available years, the probability that an event occurred increased. This indicates that, with longer time periods, the probability of an intense event occurring increases. However, the number of recorded values is relatively small, indicating that several events were missed by the gauges. Similarly, the procedure was repeated, but by randomly sampling a smaller number of stations. For example, from the total number of 1100 stations with hourly data, 300, 600, and 900 gauges were randomly selected and analyzed. The results are shown only for the stations with 15 years of data and presented in the lower 3 rows of Table 4.4. Not only does the availability of data in time matter, but so does their spatial distribution. One would expect that as the network density increases, the number of detected events increases. This is true when the network density changes from 300 to 600 stations, but as the number increases to 900, the average number of detected events decreases. This is because many locations do not capture any of the occurring events, which leads to a reduction in the average number of detected events. Though the total number of detected events is dependent on the network density and increases with the latter. The procedure was repeated using the weather radar data for the period 2000–2020. All pixels exceeding the three threshold levels were extracted and divided based on their time of occurrence (at least a 24-hour difference) in a way to capture independent events in time. The results show that within the weather radar data, many more events were detected, despite not being necessarily all correct (due to errors in the radar data). They indicate that relying only on point observations is not enough for the temporal-spatial analysis of rainfall extremes. A thorough description of missed extremes by the rain gauge networks can be seen in *Lengfeld et al.* (2020).

Table 4.3: Definition of warning levels based on the DWD.

		Warning level		
Duration [h]		1	2	3
1	Precipitation [mm]	15-25	25-40	> 40
6		20-35	35-60	> 60

Table 4.4: Average number of events exceeding the DWD warning levels derived from the rain gauges network depending on available years of observations.

		> 15	**> 25**	**> 40**	**-**	**> 20**	**> 35**	**> 60**
N years with data	5	3	0.5	0	-	4.2	0.3	0.15
	15	15	3	0.45	-	25	3	0.2
	25	25	6.5	1.5	-	37	5	0.3
N stations (15 years)	300	12.7	2.8	0.34	-	21.01	2.65	0.17
	600	15.7	3.4	0.5	-	29.1	3.63	0.22
	900	15.3	2.8	0.44	-	27	3.2	0.18

4.7.2 Deriving the spatial dependence structure

The assumptions behind describing the spatial dependence structure using the variogram are second-order stationary, a constant mean, and a variance that depends on the increments of separating distance. These assumptions lead to limits on the extent to which the variogram can be used as a tool to describe the spatial extent of extremes. For example, the presence of trends in the data should be removed beforehand; otherwise, the correct variogram calculation will be even more challenging.

Calculating the variogram from highly skewed data presents issues, and therefore, the asymmetry in the dependence structure cannot be represented. Moreover, the sparse observation network limits the possibility of correctly investigating and describing the anisotropy, whether geometrical or zonal, leading to the assumption of an isotropic variogram, which is rarely the case. Most, if not all, interpolation techniques lead to smoothing behavior and, hence, an underestimation of the true variability and maxima of the field. Moreover, as mentioned in chapter 2, precipitation belongs to the zero-inflated problems, making the variogram calculation extremely sensitive to the number of zero observations.

Another challenge is in the time domain. Precipitation is highly variable not only in space but also in time; therefore, the calculation of the variogram or the step functions for a single time step (for example, a certain hour) does not include the correlation between subsequent time steps, which is referred to as advection. It was shown in Fig. 4.1 that some events present a directional dependent correlation in time and space, an aspect that cannot be well described by the variogram or step functions.

To illustrate some of the problems previously mentioned, Fig. 4.11 shows some issues with calculating the covariance function $C(h)$ (or the variogram function ($\gamma(h) = 1 - C(h)$)) for describing the spatial dependence structure from a rainfall field. The latter was generated using a spectral simulation algorithm conditioned on a predefined spatial model defined by the covariance function (*Guthke*, 2013). Note that at high temporal resolution, rainfall fields are not Gaussian but have high asymmetric behaviour. Below were noted some remarks regarding Fig. 4.11.

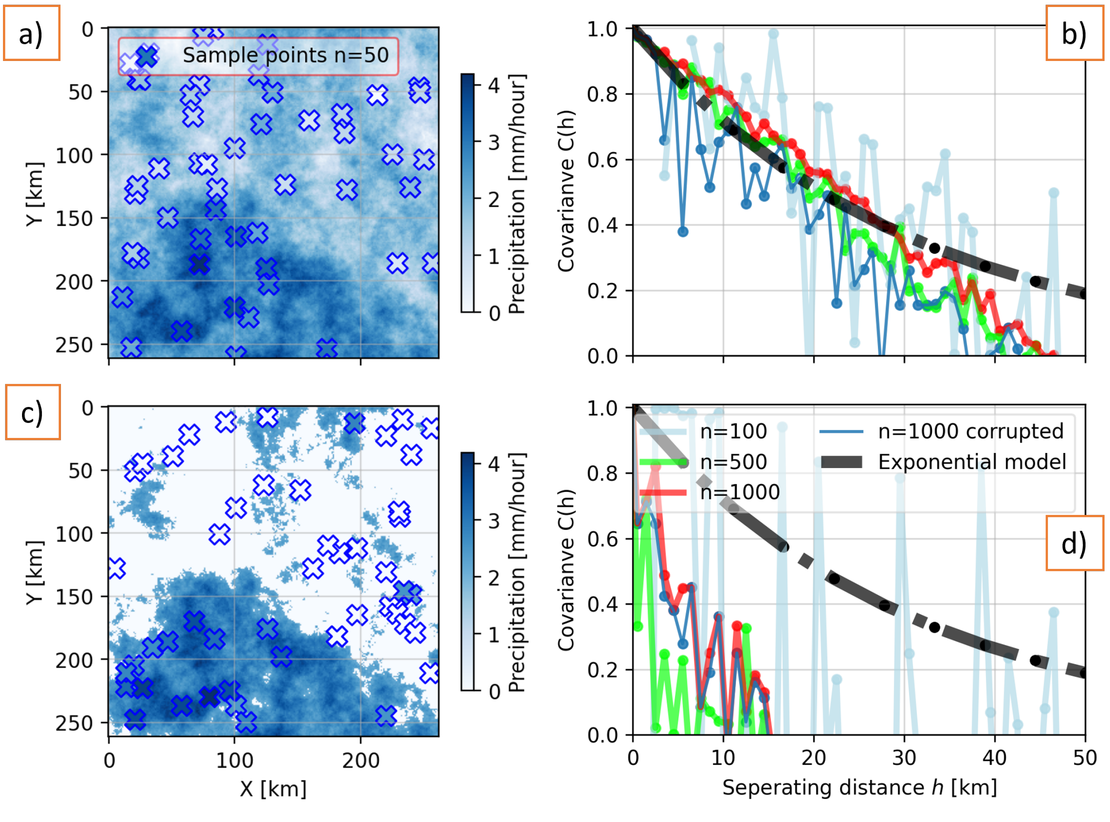


Figure 4.11: Panels (a) and (c) show the generated random fields using the same exponential covariance function shown in black in panel (b). The field in (c) has more zero values. In panels (b) and (d), the estimated covariance functions using the different sampled points from the two fields are shown.

Below some remarks regarding Fig. 4.11 are listed:

- Small sample size (50 points) leads to high uncertainty in the estimated model.
- Increasing the sample size (500, 1000 points) improves the model estimation.
- Including only one small outlier (2 larger than the maxima) in the data deteriorates the estimation despite a large sample size of 1000 points.
- Larger number of zero values leads to a bad covariance estimate and a smaller dependence range, regardless of the sample size.

One of the main challenges is that at high temporal resolution, precipitation fields present large space-time variability. The role of zero observations becomes predominant and hinders a 'correct' variogram estimation. Moreover, the assumptions behind Geostatitics such as stationarity and constant mean, are not well justified. The high spatial and temporal variability cannot be well captured by scarcely distributed point observations. Another aspect is the high asymmetry within a rainfall field; for example, there is a different dependence structure between the high and the low values. This variability within a rainfall cell, especially with a high gradient between the low and high values (a front), cannot be correctly represented by the measurement of rain gauges. In addition, interpolation cannot exceed the observed maxima, which leads to an underestimation of the true maxima that might have occurred between the values. An example of this is seen in Fig. 4.5, the center value of 63 mm is underestimated by the interpolation scheme by more than 50 %.

4.8 Summary

In the previous analysis, the station data were used to investigate the spatial extent of several rainfall events across different temporal aggregations. The first part showed that for many intense observations, the true maxima was located at a neighboring location. On average, over all aggregation and stations, for 20% of the events, neighboring values had higher observations than the center location. An indication that within a given area, the maxima and its location might be underestimated or misplaced. Moreover, a directional-dependent correlation structure present within rainfall events, especially on a local scale, was identified. Traditional approaches such as variogram calculations were seen as reliable, especially for low temporal aggregations, but had less reliability for sub-hourly and convective events. The use of step-functions helped in showing that as one goes further away from the center location (where the original intense value was), the average areal rainfall might exceed the center value. Hence, areal rainfall reduction in the traditional sense does not apply. Clustering of step functions and variograms was used to categorize events based on their spatial extent, especially for lower temporal aggregations and frontal-driven events. Another aspect was the use of the depth function to find simultaneously spatially occurring unusual events. The method showed success in identifying events on the local scale (a cluster of a few neighboring locations) and on the regional scale (over all neighboring positions). The results in this chapter support further understanding of the spatial dynamics of precipitation events and show some challenges with only using point observations.

5 Copula-based conditional merging of weather radar and point observations

In the following chapter, a copula-based conditional merging procedure is introduced to combine weather radar and rain gauge observations with high temporal resolution. The term copula is used here because the method is implemented in the rank space. A copula is a mathematical function used to derive and model the dependence between variables independent of their distribution functions. The first section describes the weather radar data, and then the theory behind external drift kriging and anisotropy is presented. A procedure previously developed by *Yan and Bárdossy* (2018) to incorporate the wind-displacement effect is used and complimented by a conditional merging algorithm of spatial and point observations (*Sinclair and Pegram*, 2005). The merging was applied for the data period 2001–2020 for the radar regions of Hannover and Türkheim. The final product was cross-validated with a focus on the extremes. Eventually, the final results showed the reliability of the derived product by being unbiased and applicable for further analysis.

5.1 Description of weather radar data

The radar data have great potential for looking at the spatial extent of rainfall events. They offer the possibility to investigate what was measured between point observations, as these might not correctly sample the rainfall minima and maxima and the event spatial distribution. In Germany, the DWD operates a weather radar network with 17 radars equipped with dual-polarization Doppler technology, covering the whole region with a high temporal resolution of 5 minutes. Doppler radars work by sending waves in both the vertical and horizontal directions, which are used to identify the reflecting object based on its shape.

Before using the radar data, the radar-derived QPE can have several errors. Several reasons are possible since the radar does not measure rainfall directly but rather measures the reflectivity of the radar signal. Large reflectivity values indicate that the beam is blocked and that most likely rainfall is failing. This does not depend only on the amount but also on the size of the raindrops. The latter, however, varies significantly between rainfall events. Depending on the end goal, different radar wavelengths might be used. Longer wavelengths are essential for qualitative information in convective and severe events, while shorter wavelengths are beneficial for a uniform representation of largely distributed precipitation events. The returned reflectivity values are used to calculate a reflectivity factor (Z), which is used to calculate the rainfall rate via the so-called Z-R relationships. These have a power function shape and vary based on the meteorological processes, the location, and the type of hydrometeors (rain, snow, hail, etc.). Several Z-R relationships were derived, and further research is pending. In reality, the Z-R function changes with every rainfall event and within the rainfall event itself. A common assumption is to consider the Z-R relation as an average value and apply it to the entire radar field.

Some of the most common error sources in the weather radar observations are, for example, overshooting, which happens when radar signals measure above the clouds (or rain cells). This happens at long distances from the radar location and leads to completely or partly missing the observed event. A common problem is attenuation. The radar signal is highly attenuated due to strong reflectivity (as in intense rainfall events). Everything behind the rainfall cell is then not measured since the signal is not strong enough. Also, since the upper atmosphere layers are colder than the lower ones, precipitation falling as snow can melt before reaching the ground. If the radar is measuring at the exact elevation where the water phase transformation is happening, a phenomenon known as the bright-band effect occurs. In other words, the radar signal is detecting the signal of the melting layer, leading to errors in the observations. Another common source of errors is the presence of static objects (buildings, windmills) or mobile objects (airplanes, birds) obstructing the radar signal. This phenomenon is known as the clutter effect and should be accounted for when deriving rainfall observations from radar reflectivity measures.

Besides these error sources, further aspects are to be considered. One of them is that the radar is measuring at a high altitude and at a certain time step. While on the ground the measurements are between time intervals, the radar rather gives an image for a certain time step showing the number of raindrops. The size of the raindrop is usually unknown and varies along the path before reaching the ground. Moreover, a displacement of the image exists; when the radar is measuring at high altitudes, the raindrop may be displaced by the wind and fall in a different location than indicated in the radar image. Another type of displacement is temporal displacement, in which the observed raindrops fall at a slightly different time than the observed ones, depending on the fall velocity.

The radar locations of Feldberg and Türkheim cover simultaneously the city of Reutlingen in southwest Germany. The city operates a network of rain gauges with high temporal (1 min) and spatial (6 km) resolution. A comparison between the two images of the two radars for a certain time step showed a similar spatial pattern but with different precipitation values. An example of this case can be found in *Yan and Bárdossy* (2018). To cope with such situations, weather services generally use the information sampled from the lowest level, but in some conditions, one can use the average, the maximum, the minimum, or a weighted combination of the overlapping data.

Finally, using several radar locations and rain gauges, a mosaic of gridded QPE values is produced. The end product is quality-checked by applying a bias correction using ground observations that are continuously updated. Correction factors are acquired from point observations and are applied to increase (or decrease) hourly or daily accumulations either on a local or field scale.

In this thesis, two datasets were used: the raw radar data for the locations of Hannover and Türkheim and the radar-based precipitation climatology-adjusted data (RadKlim) (*Winterrath et al.*, 2018a). Both have a temporal resolution of 5 minutes and cover the period

2000–2021. The RadKlim data were used as they are, and to reduce the errors in the raw data, several procedures were undertaken. These are based on several algorithms in the Python library wradlib (*Heistermann et al.*, 2013). First, the raw data are converted from polar to cartesian coordinate systems, and then dynamic and static clutter removal filters are applied (*Gabella and Notarpietro*, 2002). To account for the wetting of the Radome and the subsequent effect on the radar beam, the attenuation correction introduced by *Merceret and Ward* (2000) was applied. Finally, the enhanced three-part Z-R relation developed by the DWD was used to transfer reflectivity values to rainfall rates.

Still, when combining radar with ground observations, other effects such as wind displacement between the radar and ground observations, the spatial correlation structure of the precipitation field, and the advection vector should be considered. For this, an alternative conditional merging approach was developed based on the work of *Sinclair and Pegram* (2005) and *Yan and Bárdossy* (2018).

5.2 Methodology

The following sections describe the interpolation techniques used to estimate values at unknown locations. The first part briefly describes External Drift Kriging (EDK), which is an interpolation technique using external information as auxiliary data. In part two, conditional merging is introduced. Conditional merging grants the possibility of combining precipitation station data and weather radar observations into a combined product. To test the quality of the interpolation methods, cross-validation was used. This is described in the final part of this section.

5.2.1 External drift kriging

It is often the case that spatially distributed data are related to other variables. For example, precipitation processes have a direct relationship to topography. Such information can be included within the estimation method in the form of auxiliary information. Hence, a new family of kriging possibilities was derived that belongs to non-stationary kriging forms. Among them is External Drift Kriging (*Ahmed and De Marsily*, 1987). For non-stationary kriging methods, the intrinsic hypotheses are changed. The expected value is not a constant but a function of the location. Using an additional variable that is available at all grid nodes and is linearly correlated with the target variable, a slightly changed equation system as compared to Ordinary Kriging (presented in chapter 3) is used to estimate the values at unknown locations.

$$E[Z(u)|Y(u)] = a + bY(u) \tag{5.25}$$

where a and b are unknown constants. As for OK, for any a and b, the linear estimator should be unbiased:

$$Z^* = \sum_{i=1}^{n} \lambda_i Zi \tag{5.26}$$

The external drift kriging equation system is summarized by these three equations, given that μ_1 and μ_2 refer to the Lagrange multipliers. Same as in OK, γ refers to the variogram model. This remains similar.

$$\sum_i \lambda_j \gamma(u_i - u_j) + \mu_1 + \mu_2 Y(u_i) = \gamma(u_i - u) \tag{5.27}$$

$$\sum \lambda_j = 1 \tag{5.28}$$

$$\sum_j \lambda_j Y(u_j) = Y(u) \tag{5.29}$$

Here, weather radar rainfall values are used as drift values while performing the conditional merging of radar and station data. The radar data are available at the station locations and are distributed over a homogeneous grid with a 1 km spatial resolution.

5.2.2 Geometrical anisotropy

A simplification while working with the random function Z(u) is the assumption of isotropy. The spatial variability is independent of the direction and only depends on the separating distance h. This is often not the case since nature depicts anisotropic behavior. The spatial structure differs from one direction (angle) to another. Some examples of anisotropic processes in nature are groundwater or contaminants flowing in the subsurface, where the flow direction is influenced by the characteristics of the subsurface, and rainfall events where wind or elevation alter the spatial structure.

Anisotropy can be incorporated into the variogram calculation. The first option is to calculate the variogram not only for different separating distances but also for different directions. Hence, the directions of the major and minor axes are derived, and by a linear transformation (stretching, rotation) the anisotropic field is transformed into an isotropic one.

This type of anisotropy is known as geometric anisotropy and has, in a general form, two parameters: the stretching coefficient λ and the rotation angle ϕ.

$$x^* = \lambda(x \cos\phi + y \sin\phi)$$
$$y^* = -x \sin\phi + y \cos\phi$$

If for each direction, the still value is different or if the ranges do not fall on the ellipse, then the second type of anisotropy, known as *zonal anisotropy*, is present. One possibility to cope with this would be to divide the space into different zones based on intervals of angles. Each zone will have its own variogram that will be used for interpolating each zone. Another possibility is to fit a complex variogram model where individual terms represent different geometric anisotropies, which can be direction-dependent or not. An example of the derived variograms for different directions can be seen in Fig. 5.1. For every direction, the values falling within this orientation were selected, and the empirical variogram was calculated. The sill and range values of the variogram change with each selected direction. In an isotropic case, only the variogram in blue (an angle of 0 degrees) would have been derived. Moreover, an example of applying the conditional merging with and without the account of geometrical anisotropy can be seen in Fig. 5.2. From the radar field, the variogram in 2D was derived, and the direction of the main axis and minor axis along the scaling coefficient were derived. Panel (a) shows an isotropic case; panel (b) after accounting for anisotropy; and panel (c) is the difference map. The difference exceeds the values of 3 mm/5 min in both directions. Such differences influence the final result and the areal rainfall.

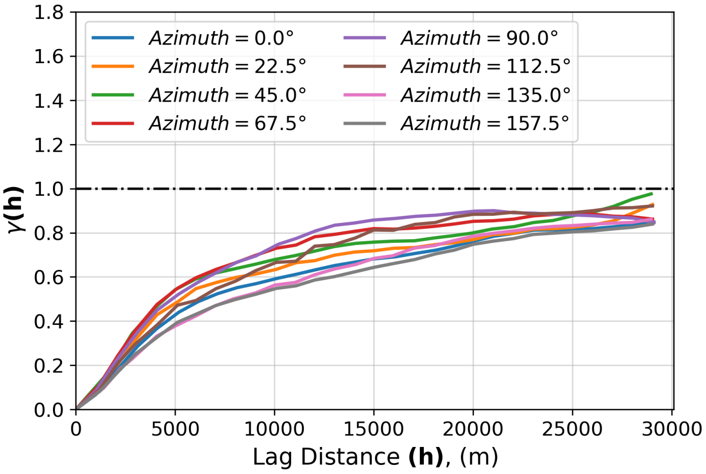


Figure 5.1: Example of geometrical and zonal anisotropy (where the empirical variogram and sill value change with direction).

To derive a suitable variogram describing the spatial dependence structure, the variogram was calculated from the combined radar and station observations. For this, the variogram was calculated for different directions (angles), and the directions of the main and minor axes along the stretching factor were calculated. Since precipitation is a variable that falls within the zero-inflated problem, the influence of the zeros on the variogram needed to be investigated. The following section covers that part.

5.2.3 Influence of the zeros on the variogram

On a 5-minute temporal resolution, around 95% of the precipitation values are zeros. The zeros have a large influence on the variogram: they lead to a larger sill, and larger range and affect the anisotropy direction and scaling factor. The large number of zeros influences the spatial dependence structure and can lead to a misrepresentation of the dependence of the extreme precipitations.

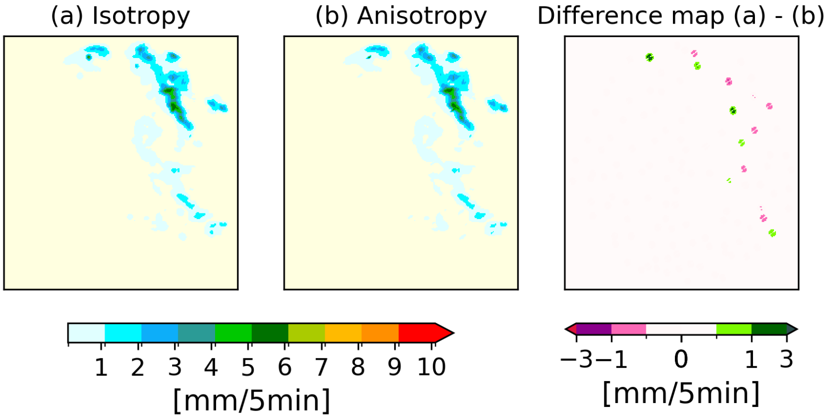


Figure 5.2: Example of a merged rainfall field using isotropic variogram (left figure) and anisotropic variogram (center figure). The right plot depicts the difference map in mm per 5 minutes.

To investigate this effect, four different approaches to derive the variogram were considered, and the effect on the Kriging Error Term (KET) defined in equation 5.30 was calculated.

$$KET(u,t) = \frac{Z^*(u,t) - Z(u,t)}{\sigma_{u,t}} \tag{5.30}$$

Where:

$Z^*(u)$ = estimated value at location u and time step t
$Z(u)$ = observation value at location u and time step t
σ_u = kriging standard deviation at location u and time step t

The error term values were calculated from the cross-validation results of 300 hourly events for 16 different DWD stations. To derive and fit a variogram to the empirical variogram map (variogram in 2D), the Fast Fourier Transformation (Fast Fourier Transformation (FFT)) was applied (*Marcotte*, 1996). Four different approaches for variogram calculation were considered; for each case, the variogram range, the sill, and the anisotropy (direction, scaling) were noted.

1. Original field: variogram derived from radar field with zeros
2. No zero field: variogram derived from radar field without zeros
3. Reduced field: variogram derived from the radar field with a layer of zeros
4. Indicator field: variogram derived from indicator radar field

Fig. 5.3 demonstrates an example of one hourly event. The lower panels show the estimated variogram map for each case. For each aforementioned case, the variogram parameters were calculated. For each case, the values of the range, direction of anisotropy and scaling fac tor changed. The results for this 5 minutes time step are displayed in Table 5.1.

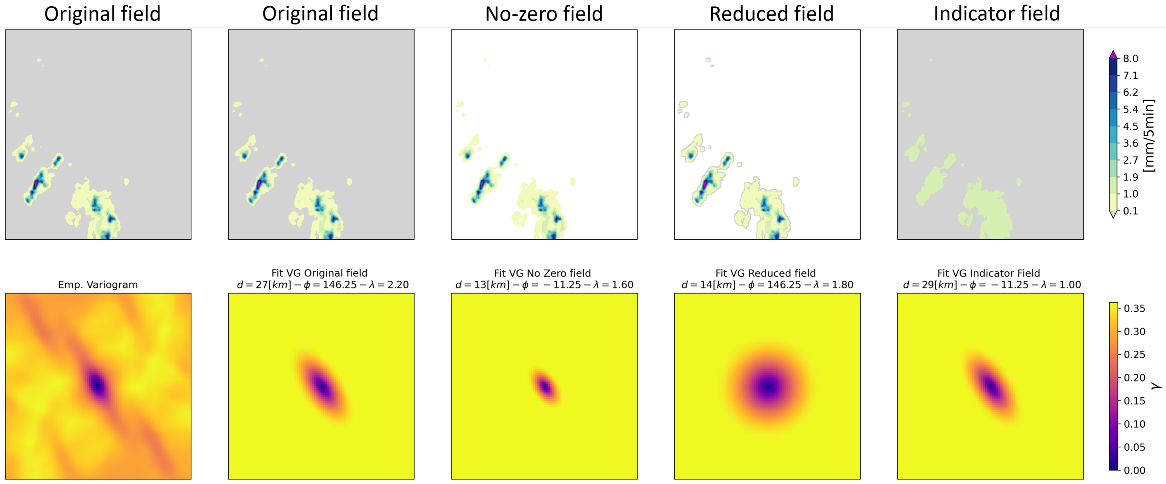


Figure 5.3: Example of calculated variogram from the original, the no zero, the reduced, and the indicator fields for a 5-minute time step. Each with correspondent fitted geometrical anisotropy parameters.

Table 5.1: Estimated geometrical anisotropy parameters from the fields in Fig. 5.3. Note that depending on the portion of zero values the parameters of the fitted dependence model varies.

	Original field	**No Zero field**	**Reduced field**	**Indicator field [0.1 mm]**
Variogram range [km]	27	13	14	29
Major axis direction ϕ [deg]	146.25	-11.25	146.25	-11.25
Scaling factor λ	2.20	1.60	1.80	1.00

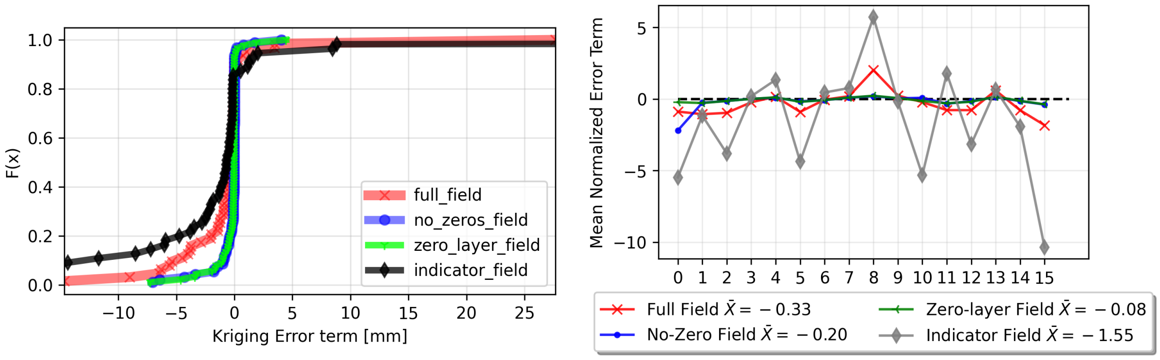
A summary of the cross-validation results can be seen in Fig. 5.4. The result shows that using the variogram with the zero layer has the least error compared to using the whole radar field, only the positive part of it, or an indicator field. The previous analysis showed that the reduced field, meaning the field containing the positive values enclosed by a small buffer of zero values, gave optimal results. Hence, when applying the merging, a suitable theoretical variogram without a nugget value was fitted to the empirical one, and the resulting model was used in the conditional merging.

5.2.4 Copula-based conditional merging

To combine the radar observations with the ground-point data, an adapted conditional merging approach was implemented. The main idea behind conditional merging is to have a final field that respects the point observations and incorporates the spatial distribution of the radar field (*Sinclair and Pegram*, 2005).

The different parts of the traditional conditional merging procedure are described below:

1. (a): The rainfall field is observed at discrete points by rain gauges.



(a) CDF of the KET (one location)

(b) Mean error term (per location)

Figure 5.4: Kriging error term depending on variogram calculations. Four rainfall fields are used: (1) with zeros, (2) without zeros, (3) only a layer of zeros and (4) with an indicator field.

2. (b): The rainfall field is also observed by radar on a regular, volume-integrated grid.
3. (c): Kriging of the rain gauge observations is used to obtain the best linear unbiased estimate of rainfall on the radar grid.
4. (d): The radar pixel values at the rain gauge locations are interpolated onto the radar grid using Kriging.
5. (e): At each grid point, the deviation C between the observed and interpolated radar values is computed.
6. (f): The field of deviations obtained from (e) is applied to the interpolated rainfall field obtained from Kriging the rain gauge observations.
7. (g) A rainfall field that follows the mean field of the rain gauge interpolation while preserving the mean field deviations and the spatial structure of the radar field is obtained.

Fig. 7.1 shows the radar area of Hannover along the location of the available DWD stations with sub-hourly resolution. The aim was to combine the radar observations with the ground data in order to gain the most from both observational data sets.

Despite correcting the radar observations for different possible errors (e.g., attenuation, beam blockage, clutter removal, etc.), the displacement due to the wind movement needs to be considered. The incorporation of the horizontal wind displacement vector was done similarly to the work of *Yan and Bárdossy* (2018). The reader is referred to the previous manuscript for a full description of the method. Below, a summary is provided.

To incorporate the effect of the horizontal wind displacement between the radar and ground observation, the radar field was shifted homogeneously for different horizontal and vertical displacement vectors. For each of the shifts, the rank correlation between the radar and ground observations was calculated and compared to the reference correlation (shift=0).

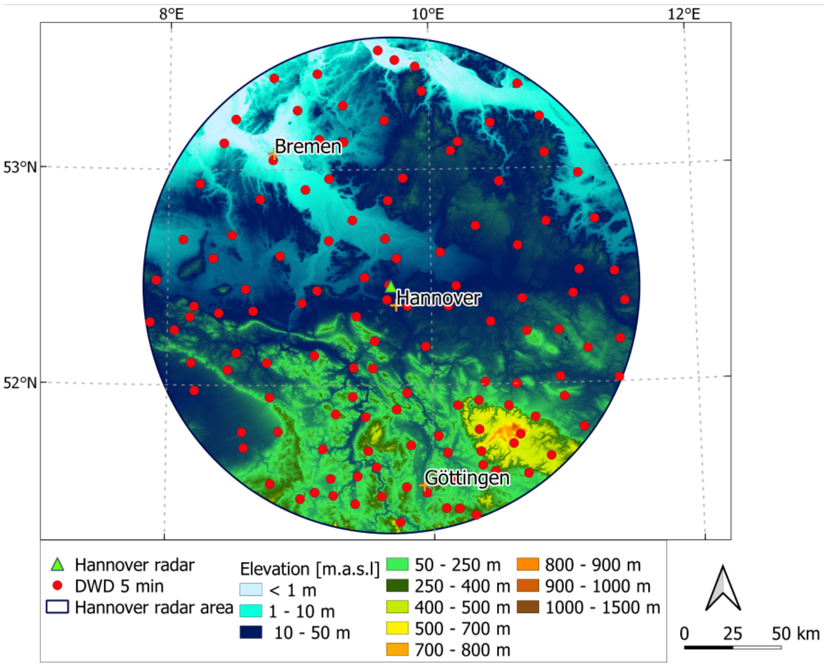


Figure 5.5: Digital elevation model of the radar coverage area of Hannover along the location of the DWD stations with sub-hourly resolutions.

From this, a rank correlation matrix that was later transformed into a rank probability matrix using a transformation function was calculated. Each shift that brought an increase in the rank correlation was then associated with the corresponding probability. A modification of the original work was done by incorporating the advection vector between subsequent observations. The procedure was applied for each 5-minute time step by considering one step before and one step afterwards. This granted greater continuity in time and space. The so-called expected field was then derived as a weighted product of the different fields that were transformed using the marginal distribution function. The latter was derived from the shifted radar observations at the ground locations. The expected field was used to derive the variogram map and the anisotropy parameters. These were incorporated in conditional merging. Hereby, EDK was applied as an interpolation technique, with the expected field as the drift term. This enabled the merging of the radar spatial structure and the station point observations and provided better results than using ordinary kriging.

A complete description of the copula-based merging workflow is given below and in Fig. 5.6. Note that the final product is a two-dimensional field incorporating the ground observations, the advection vector, the wind displacement information, the geometrical anisotropy, and the spatial distribution of the original radar field.

1. Preprocess the radar reflectivity observations (cutter removal, attenuation correction, etc.) and transform them to rainfall values using the Z-R relationship.
2. Read the station observations in the radar field with a temporal resolution of 5 minutes.

3. For a timestep i, consider the timesteps $i-1$ and $i+1$ to account for the advection vector.
4. Apply a Gaussian kernel smoothing of the radar field dependent on the variance of the field to reduce the 'noise' in the data (optional step).
5. Calculate the rank correlation between the radar observations at the station locations and the corresponding station values c_0.
6. Start to shift the radar field homogeneously horizontally i or vertically j or in both directions ij and transform the field to the rank space U_{ij}.
7. For every shift, calculate the rank correlation between the station data and the corresponding radar values.
8. Find all the shifts with a rank correlation greater than c_0.
9. Repeat this for timesteps $i-1$, i, and $i+1$ and construct an average rank correlation matrix.
10. If there is a shift in the average rank-correlation matrix that increases the correlation, consider that; otherwise, consider only the shifts for timestep i.
11. Construct the marginal distribution function of the shifted field using the radar observations at the station locations as defined in equation 5.31. Extend the maximum value using an exponential distribution.

$$F_{ij}(z) = \frac{u_k - u_{k-1}}{z_k - z_{k-1}}(z_k - z_{k-1}) + u_{k-1} \tag{5.31}$$

Where:

z = ground observation
z_k, z_{k-1}= nearest neighbors of $z_{k-1} < z < z_k$
u_k, u_{k-1}= quantiles corresponding to z_k, z_{k-1}

12. Transform the ranks of the radar field into precipitation observations using the corresponding marginal distribution function.
13. Assign for every transformed field a probability value P_{ij} dependent on the increase of correlation as compared to c_0, with emphasis on the shifts with a higher increase.
14. Construct the final radar field (the expected field Z), which is a weighted combination of the previously calculated fields using equation 5.32.

$$Z = \sum_i \sum_j P_{ij} F_{ij}^{-1}(U_{ij}) \tag{5.32}$$

Where:

U_{ij} = shifted radar field in the rank space
F_{ij}^{-1}= inverse of the marginal distribution function from equation 5.31
P_{ij} = value in the probability matrix for shift ij

15. Use a fast Fourier transformation to calculate the empirical variogram (variogram map) from the expected field and the station observations.
16. To account for geometrical anisotropy, find the direction and scaling factor of the major and minor axes and apply the corresponding coordinate transformation.
17. Use the station observations to interpolate the target grid using external drift kriging $EDK_{stations}$.
18. Use the expected field observations at the station location to interpolate the target grid using external drift kriging EDK_{radar}.
19. Calculate the difference between the expected field and EDK_{radar}.
20. Add to the $EDK_{stations}$ the calculated difference.
21. Back-transform everything to the original coordinate space.
22. This is the final merged radar-station field for this time step.
23. Repeat the procedure for the next time step.

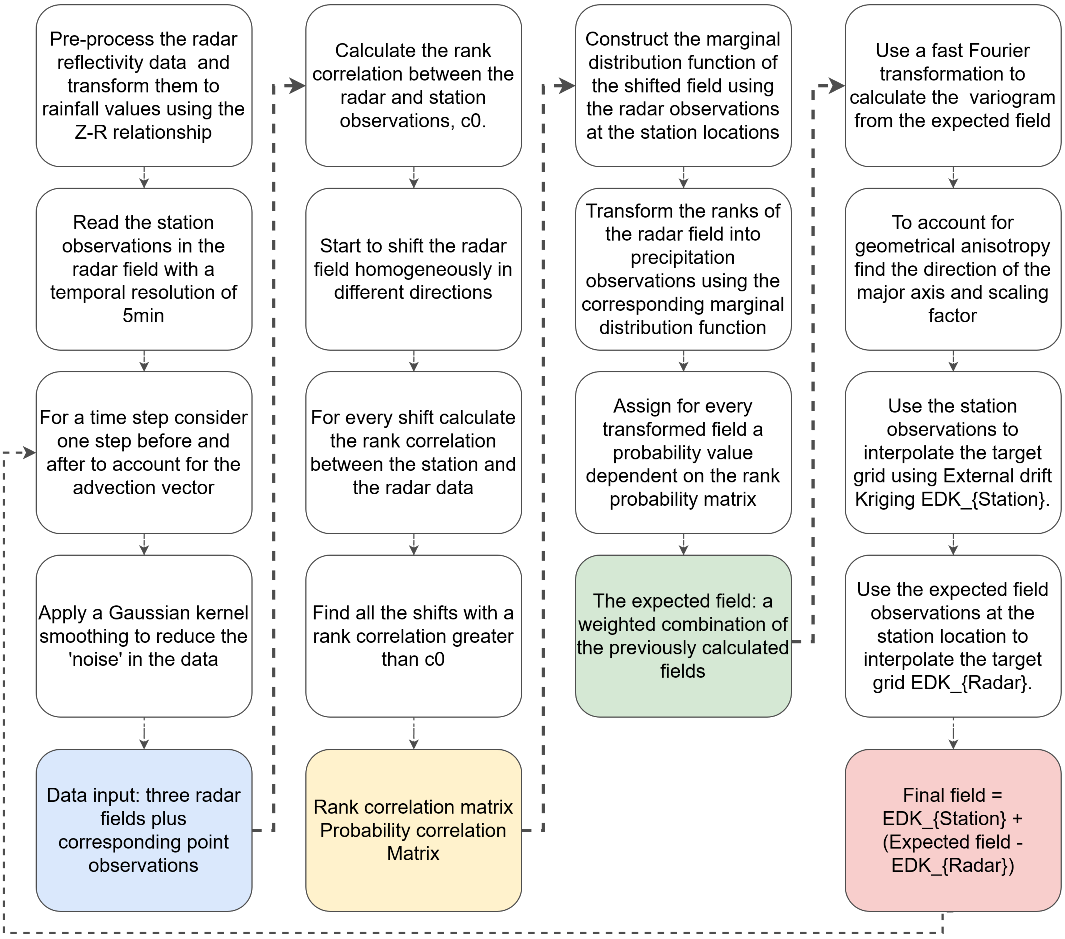


Figure 5.6: Flowchart summarizing the applied copula merging procedure.

5.2.5 Optical flow procedure

The weather radar observations are a snapshot of the observed time. This is not a representation of the continuous rainfall process happening continuously between the different time steps. If one tries to aggregate the images without considering the continuous flow in time and space of rainfall, the final result will present discontinuities. One might compare this with the scales of a fish, where each scale represents an image. The effect becomes more predominant with increasing temporal aggregations. Therefore, to account for the movement of rainfall, the optical flow method was used (*Ayzel et al.*, 2019).

The underlying assumption for solving the optical flow problem using the Lucas-Kanade method is that neighboring positions have similar movement behavior (*Lucas and Kanade*, 1981). The optical flow vector (u, v) is constant within a small neighborhood of the space window ω. This is described by equation 5.33 defined for each point (a, b) in the domain.

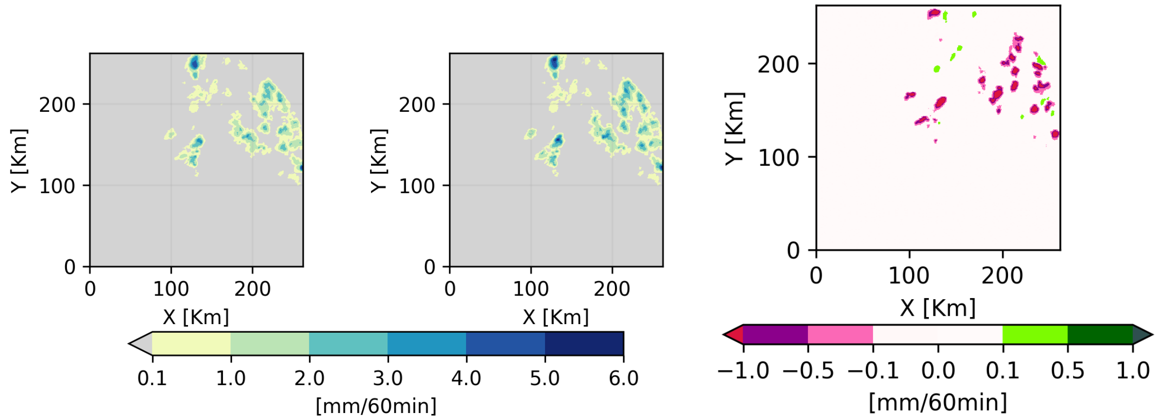
$$\frac{\partial I(a,b)}{\partial x}.u + \frac{\partial I(a,b)}{\partial y}.v + \frac{\partial I(a,b)}{\partial t} = 0 \tag{5.33}$$

Where:

(a, b) = location of point in domain ω
(u, v) = optical flow vector
$I(u, v)$ = intensity (function of directions x, y and time t)

This is the same for all points in ω leading to an equality system with a similar number of equations as the number of points. The advantage of this equation system is that all equations are linearly independent of each other. Since the derivative in the x and y directions is different for each point. The goal is to find the motion vector (u, v) for the window ω by solving the linear equation system. Two cases exist where the calculation is less reliable: the first is when neighbouring windows have high similarities (very small texture, namely all zero observations), and the second is when there is an edge between the neighbouring windows (high texture to low texture, a gradient between high and low values). However, when the texture is very rich (for example, within the rainfall field), the optical flow method is reliable.

In the case of aggregating sub-hourly radar images to higher aggregations, the advection vector (motion field) between consecutive images was accounted for using the motion field and interpolation method available in the Python radar now-casting library $pysteps$ (*Pulkkinen et al.*, 2019b). An example of the results for an hourly aggregated field (sum of twelve 5-minute images) with and without the optical flow method can be seen in Fig. 5.7. Note that the corrected image shows a better representation of the continuous rainfall process with fewer discontinuities (dry patches). The spatial continuity is important for areal rainfall analysis, which otherwise might be underestimated.



(a) Without and with correction

(b) Difference map

Figure 5.7: Example of an aggregated hourly image before (left panel) and after (right panel) applying the optical flow method. Panel (b) shows the difference map (without-with correction).

5.2.6 Cross-validation

Cross-validation is used to test if the assumptions made and the consequently estimated statistics are reliable. Observation values at certain locations (or all of them) are estimated using all neighboring data, disregarding the data at the target location. The estimated value is then compared to the observed one at the same location. The procedure is repeated for several observations at a certain location or/and different locations. To judge the results of the cross-validation procedure, the term in equation 5.34 was calculated. This normalized Criteria Ratio (CR) should theoretically have an average value of zero and a standard deviation of one. This indicates that the estimation is unbiased and has an appropriate variance value (assumption about the errors and variogram model).

$$CR_i(u) = \frac{|Z_i^*(u) - Z_i(u)|}{\sigma_i(u)} \tag{5.34}$$

Where:

$Z_i^*(u)$ = estimated value at location u and timestep i
$Z_i(u)$ = observed value at location u and timestep i
σ_i = kriging standard deviation at location u and timestep i

When doing a cross-validation procedure, one can decide to leave one station out and re-estimate all of its observations or to leave several stations out simultaneously. The latter is known as split-sampling and generally refers to splitting the observation data set into two parts one used for the data estimation and one for the validation (*Refaeilzadeh et al.*, 2009). Cross-validation was applied in this work to test the quality of the final merging product. The focus was on intense precipitation observations. Fifteen stations with the longest observation records were individually removed from the data set, and their 5-minute values were re-estimated using the previously described merging procedure. Eventually, the estimated and observed values were compared, and the data were used to derive depth-duration-frequency curves.

5.3 Results

5.3.1 The expected field

To derive the expected field, the method introduced by *Yan and Bárdossy* (2018) was used. The main idea behind calculating the expected field is to reduce the disagreement that might have occurred due to wind displacement. Therefore, the original radar field was shifted in the rank space by several kilometers (up to 4 km) in the horizontal and vertical directions. Shifts that increase the agreement in the rank space between the radar and ground observations were identified. Based on the difference between the original rank correlation and the shifted rank correlation, a probability value was assigned to the corresponding field. The procedure was repeated until all shift combinations had been tested, and only shifts with positive increases were considered. The resulting rank correlation matrix can be seen in the middle panel of Fig. 5.8. The shifts with a red value are those with a correlation above the reference correlation (in this case, 0.6). The maximal correlation of 0.7 is noted for the shift $i = -4$ and $j = -4$. The rank correlation matrix was transformed to a probability matrix using a transformation function denoted as $g(x)$, where x represents the rank correlation value. Here, the transformation function $g(x) = x^3$ was used, where $x = r_0 - r_{ij}$ with r_0 and r_{ij} being the reference and shifted field rank correlations. Although several other functions are possible, this one was chosen as it enables assigning the largest probability to the fields with the largest improvement in the rank correlation. Shifted fields with minor improvements in rank correlation were associated with low probabilities. Hence, fields with greater improvement were prioritized. Note that all probability values in the probability matrix must sum to one. Using equation 5.32 the expected field was derived as a weighted combination of all fields with positive probabilities.

5.3.2 The final field

Fig. 5.9 shows an example of a merged rainfall field incorporating the expected field and the station observations using EDK while accounting for the anisotropy direction and scaling factor. The left panel shows the original QPE field with a rank correlation of 0.57. The latter was calculated between the values of the rain gauges and the overlapping radar pixels. In the center panel, the derived expected field is shown with a correlation value of 0.66 with the ground observations. For this time step, the expected field has larger rainfall maxima, but compared to the original field, the proportion of small rainfall values decreased. The final field presented in the right figure incorporates the expected field and the ground observations.

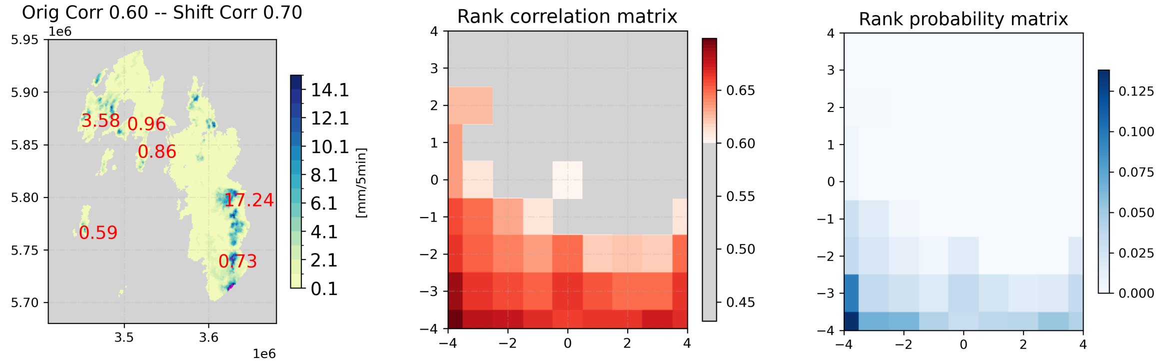


Figure 5.8: Original radar field with corresponding rain gauge values in the left plot. The associated rank correlation matrix and probability matrix are shown in the center and right plots, respectively.

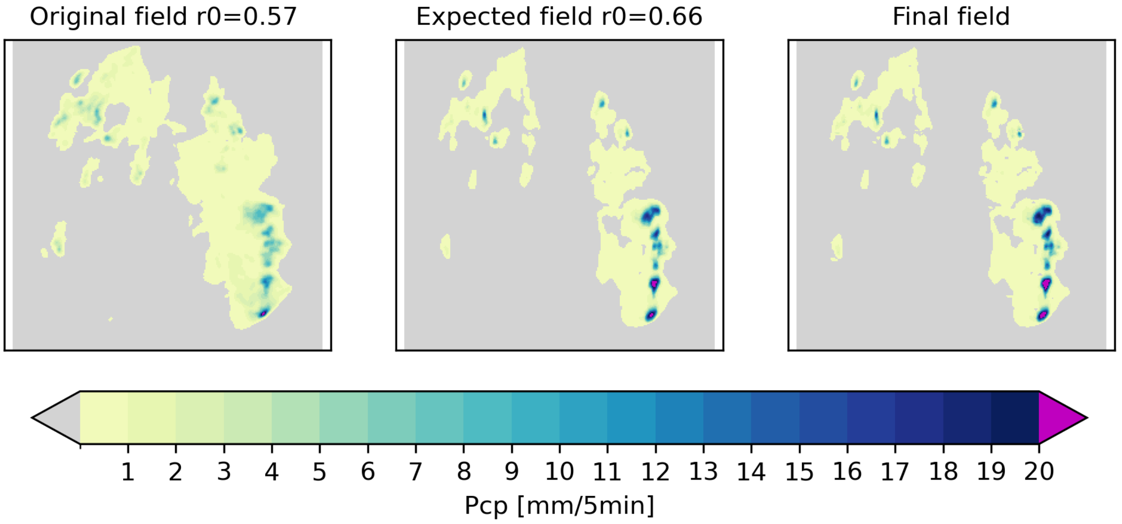
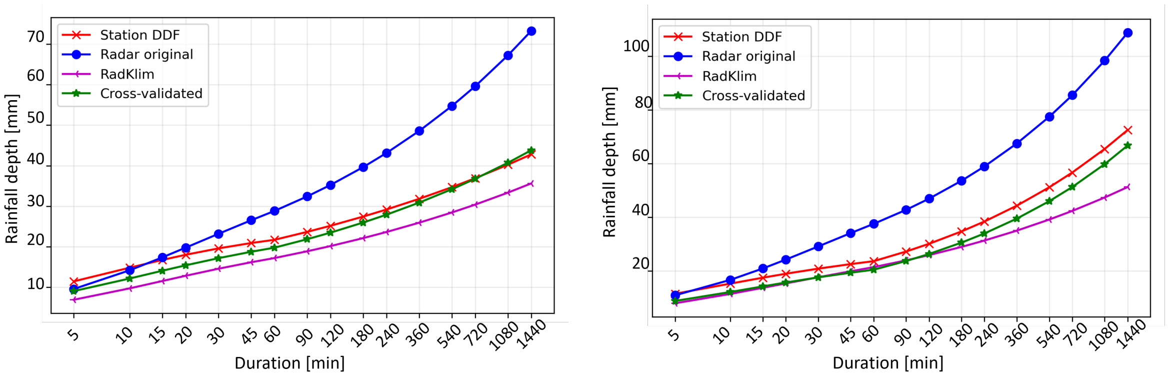


Figure 5.9: Left figure shows the original field with a rank correlation value of 0.57 with the ground observations. The center figure displays the expected field with a rank correlation of 0.66 with the ground observations. The right figure represents the final field incorporating the information from the expected field and the ground observations.

5.3.3 Cross-validation: DDF curves

To test the validity of the method, a cross-validation procedure was implemented. The 15 stations within each radar area of Hannover and Tuerkheim with the longest records, mainly from 2000 until 2019, were selected, and their data were re-estimated using the merged product via a leave-one-out procedure. This means that cross-validation was done thirty times. The results are compared to the observed values, and the focus was on the extreme values. To that end, the DDF curves were calculated, and different statistical measures such as the mean bias and the RMSE were calculated. In Fig. 5.10 the derived DDF curves for a 5-year return period for two locations are shown. The observed DDF curves were compared to the processed raw radar data, the station data, the RadKlim pixel data, and the cross-validated values. The processed raw radar DDF curve (in blue) shows an overestimation, while the Radklim data (curve in lilac) shows an underestimation of the DDF curves and hence the design rainfall values. The DDF from the merged product (in green) shows the largest agreement with the observed data (in red).



(a) DDF of cross-validated location 1

(b) DDF of cross-validated location 2

Figure 5.10: Calculated DDF curves for two cross-validation stations using the observed station data (red curve), the processed raw weather radar data (blue curve), the merged estimated point value (green curve), and RadKlim data (purple curve).

Moreover, the quality of the merged product was further investigated by inspecting the error between the quantiles of the DDF curves for the cross-validated and reference rain gauge data. The bias value was calculated between the gauge DDF and the DDF curves derived from each other product. The results are presented for the fifteen locations in Hannover in Fig. 5.11 for every duration. Panel (a) uses the processed raw radar data showing the largest bias and panel (b) depicts the copula-based merging with the smallest bias. Note that in RadKlim information from the cross-validated station were incorporated into the data, hence the comparison might not be fair. Still, the copula-based merging outperformed the RadKlim results (not shown here). Similar results were seen for other metrics, such as root mean square error, where the merged products outperformed all other products. The cross-validation results reveal that the error term for large durations (e.g., daily data) is still high, and the merged data show an increase in the general error term and bias across all durations. The DDF curves obtained from the processed raw radar data also showed an overestimation for the long durations. This overestimation may be attributed to remaining errors in the processed radar fields. Additionally, the error terms were found to be proportional to the return period, which could be attributed to the relatively short observation period used in the analysis. To mitigate the errors, a maximum return period of 5-10 years were used in this study.

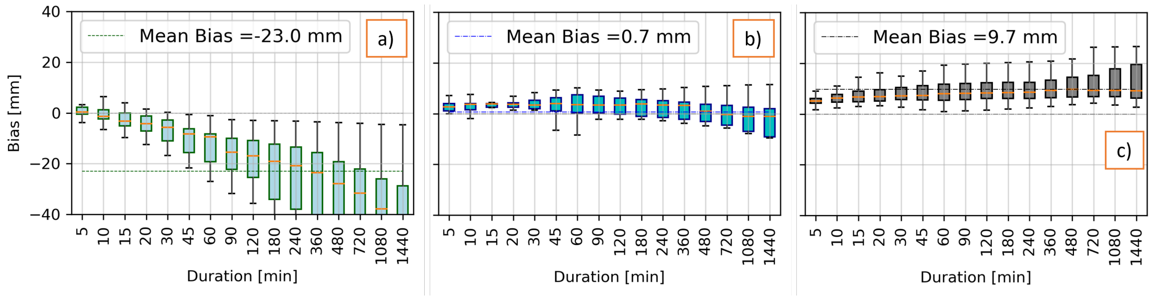


Figure 5.11: Panel (a) shows the results from the raw processed radar data and panel (b) from the copula merging data. Panel (c) the bias from the Radklim data. The bias is calculated between the quantiles of the DDF curves compared to the observed data over the different durations and locations.

5.4 Summary

The presented method has been successful in deriving an unbiased and statistically validated combined spatial product by integrating information from radar fields and gauge observations. The procedure was applied directly to high-temporal resolution data using processed raw radar reflectivity observations spanning over 20 years. The quality of the final product was validated using a cross-validation procedure, focusing on extremes, with data sets from two different radar locations. Modifications to this method could include using the original reflectivity field instead of quantitative precipitation estimates to derive the expected field, which may further reduce uncertainties. Moreover, two assumptions were made that may not always hold. First, the gauge observations are representative of the entire field. This might not be adequate in cases where a rainfall event was partly or completely missed by the gauges. The second assumption is that the field is being homogeneously shifted. In cases of a heterogeneous wind field over the domain, the results might not be accurate. Also, the final product may still have deficiencies due to persistent beam blockages in certain regions (such as in the southwest region of Hannover and Feldberg), as inherited from the processed raw radar observations. These have also been observed in the official DWD weather radar products. Despite these limitations, the merged product was used for investigating the spatial extent of extremes and for deriving design values. Acknowledging these limitations and uncertainties in the data is important when using the merged product for applications such as extreme value analysis and design purposes. Further research and improvements in data processing techniques may be necessary to mitigate errors and uncertainties associated with longer durations and larger return periods and to address persistent issues such as beam blockages. However, advantages of using RadKlim is the country wide availability and the use of overlapping several radar coverage areas.

6 From point to areal precipitation with the area reduction factor.

6.1 Theoretical background regarding DDF, ADDF curves, and ARF values

6.1.1 DDF curves - Ombrian relationship

To estimate rainfall depth for design values, a statistical analysis of rainfall maxima derived from long records is required. In general, design values are associated with the corresponding duration and return period. The latter varies between 1 and 100 years, depending on the intended application. The rainfall depth (or height) depends on the duration, and for the same duration, it must increase with a larger return period. The previous concept forms the basis of the statistical analysis of heavy rainfall. Since rainfall observations are generally more abundant and easier to acquire than discharge values (especially for flood events), the calculated data are used to derive the hydrograph for the corresponding duration and return period. The relation between rainfall depth, duration, and return period is commonly known as the Depth-Duration-Frequency (DDF) or Intensity-Duration-Frequency (IDF) curve or Ombrian curve (*Koutsoyiannis et al.*, 1998).

The idea behind the DDF curve is to derive a mathematical expression relating the average rainfall intensity (i) occurring over a timescale (d) for a predefined return period (T) (*Koutsoyiannis and Papalexiou*, 2017). DDF curves are used to estimate the probability of non-exceedance of a certain rainfall amount for a certain duration. These can be derived from a frequency analysis of the observed station data for different durations. For acquiring reliable values, a minimal 30-year observation period is recommended. However, for a 10-year return period, an observation record of 20 years is seen as suitable. Otherwise, high uncertainty in the design values exists. Rainfall depth for a short duration can be adequately derived only from the summer observation period. This is because high-intensity precipitation occurs mostly in the warm months and is generally driven by convective events. For longer durations, however, the winter records must be included. Rainfall maxima are extracted from the observation time series for the different durations either by considering the yearly maxima (annual series) or the values exceeding a minimal threshold (partial series) (*DWA-A*, 2012). Standard practice is to fit to the empirically calculated DDF curve a theoretical extreme value distribution function (e.g., Gumbel Type 1), from which one can derive the possible rainfall depth (or intensity) for a certain return period and timescale.

The estimation of the DDF curves is done using the following steps:

1. For each station, aggregate the data to the required durations.
2. Consider the values for each year separately.
3. Rank the observed data while assigning equal ranks to equal values.
4. Calculate for each observation its probability of exceeding p.
5. Compute the corresponding return period $T = \frac{1}{p}$.
6. Transform the observed value into intensities by dividing by the considered duration.
7. Find the average values from all years for each return period and duration.
8. This is the empirical DDF curve.
9. Fit a Generalized Extreme Value distribution to empirical DDF.
10. Use the fitted function for design purposes.

After calculating the empirical values, the traditional procedure consists of fitting a Generalized Extreme Value (GEV) distribution (e.g., Gumbel Type I) to the annual maxima or an exponential distribution to the partial series. The parameters of the distribution function are estimated by the mean of the method of moments, the maximum likelihood estimate, or the L-moments method for each duration separately. The reasoning behind fitting a distribution function to the sampled annual or partial maxima is that these represent only one realization of the possible rainfall values for the corresponding duration. The theoretical distribution is then used to derive design values for the desired duration and return period. The CDF of the GEV distribution is defined in equation 6.35:

$$F(x, \mu, \sigma, \theta) = e^{-[1+\theta\frac{x-\mu}{\sigma}]^{\frac{-1}{\theta}}} \tag{6.35}$$

Where:

μ = location parameter (location of distribution peak)
σ = scale parameter (distribution spread)
θ = shape parameter

The location parameter μ corresponds to the precipitation amount that is exceeded on average once per year. The scale parameter σ determines how strongly the precipitation amount increases with the return period (i.e. corresponds with the slope of the GEV distribution in a Gumbel probability plot). Finally, the shape parameter θ determines the degree of deflection (upward or downward) in a Gumbel probability plot. The GEV distributions are divided into three families (types), the Gumbel, Fréchet and Weibull distributions. The first two are of interest in hydrological design and have been widely used for modelling extremes. However, the last type is more adequate for variables with a light-tailed distribution and not for heavy-tailed distributions, such as rainfall extremes. The Gumbel distribution function is applied in the following form:

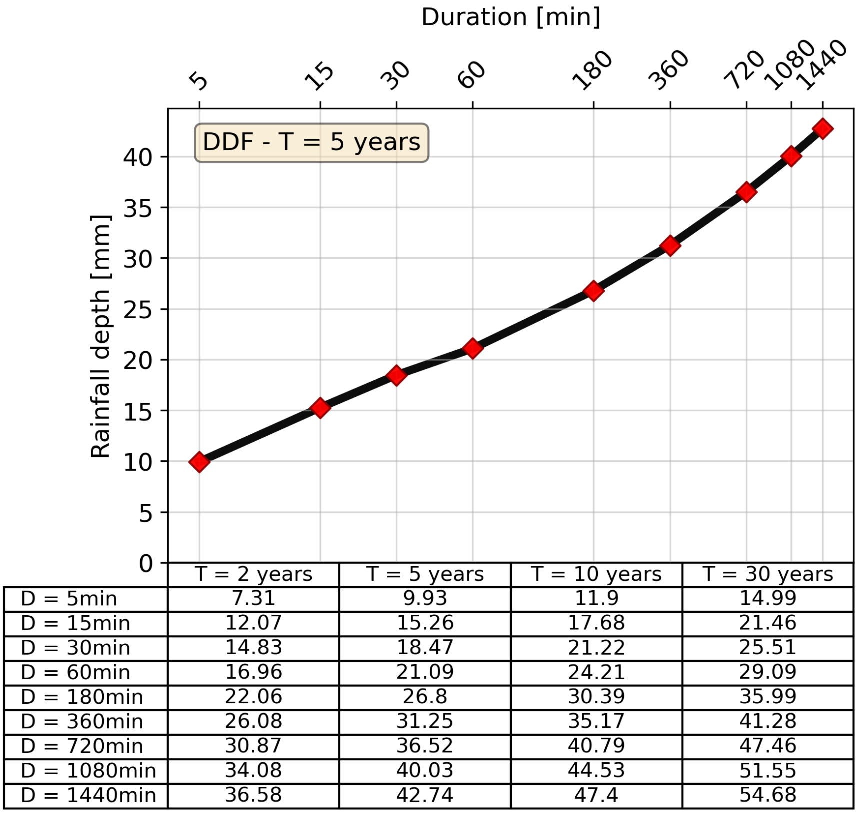
$$h_N(T_n) = u_j + w_j(-ln\ ln\frac{T_n}{T_n - 1}) \tag{6.36}$$

Where:

h_N = rainfall depth in [mm]
T_n = return period of the annual maxima in years [a]
u_j, w_j = parameters of the distribution function

An example of the derived theoretical DDF curve can be seen in Fig. 6.1. The curve represents the DDF for DWD rain gauge data for a return period of 5 years. The x-axis refers to the duration (on the log scale) and the y-axis to the rainfall depth in millimeters. The table below the figure shows the values for selected durations and return periods. Depending on the hydraulic structure to be designed, different rainfall values for the corresponding return period and duration are required.

For example, urban drainage networks are designed to account for high-intensity and short-duration rainfall events with a relatively small return period between 2 and 5 years. However, large drainage areas, such as river bridges, are designed to account for events with long duration (long concentration time) and large return periods (between 50 and 100 years).



	T = 2 years	T = 5 years	T = 10 years	T = 30 years
D = 5min	7.31	9.93	11.9	14.99
D = 15min	12.07	15.26	17.68	21.46
D = 30min	14.83	18.47	21.22	25.51
D = 60min	16.96	21.09	24.21	29.09
D = 180min	22.06	26.8	30.39	35.99
D = 360min	26.08	31.25	35.17	41.28
D = 720min	30.87	36.52	40.79	47.46
D = 1080min	34.08	40.03	44.53	51.55
D = 1440min	36.58	42.74	47.4	54.68

Figure 6.1: Example of calculated fitted theoretical DDF curve from one station data.

According to *Koutsoyiannis* (2004) the GEV type two (also known as the Fréchet distribution) represents the distribution of the annual maxima better than the Gumbel distribution. The Gumbel distribution has been widely used in practical applications, for example, by the

German Association for Water, Wastewater, and Waste (DWA) recommendations for calculating design storm values (*DWA* (2006), *DWA-A* (2012)). However, the problem lies in the small sample of extremes that disguise the true behavior of the maxima. Typically, samples of maxima with 30 to 50 years of observations tend to follow the GEV Type I distribution. Recent work by *Koutsoyiannis and Iliopoulou* (2022) advanced further the concept of Ombrian curves to theoretically founded Ombrian models. This allows a stochastic generation of Ombrian curves while accounting for the properties of the parent process. These are, for example, the first and second moments (mean and variance), the intermittency of the rainfall process (wet and dry behavior) and conserving the properties of the maxima. In this thesis, however, the *DWA-A* (2012) approach was used to estimate the DDF curves from an annual series of 20 years (2000–2020).

DDF curves are derived from point observations and represent a single location. For hydrological modeling and catchment scale analysis, areal precipitation values are required. For this, the same procedure for calculating DDF curves should be done using areal rainfall products over the catchment area. Unfortunately, this information is still not correctly available for long observation periods (unless the catchment is completely covered by rain gauges). Common practice is to use ARF to transfer the point data to the area of the catchment, resulting in spatially distributed rainfall extreme values. By which the design values for the corresponding area, return period, and duration are derived.

6.1.2 Area reduction factor - Literature review

Design rainfall values are necessary for several applications. Often, the concept of Area-Reduction-Factor (ARF) is used to transfer the point-derived design values to an area or catchment scale. Initially, ARF was introduced to calculate area-depth rainfall within a study conducted by the US Weather Bureau in the year 1957 (*Bureau*, 1957). Many studies have investigated several methods and properties to derive ARF values for a target region. A comprehensive methodological review can be found in *Svensson and Jones* (2010). The main assumption behind ARF is that the maximum areal rainfall average value is smaller than the point maximum value for the same duration and return period. In other terms, the ARF is used to relate point to areal rainfall values. In principle, the ARF for a basin with an area A for a duration D and return period T is calculated as the ratio between the area-average rainfall intensity and the point rainfall intensity for the same duration and return period (*Langousis*, 2005).

Several factors influence the ARF calculation. Since the spatial structure of rainfall depends on the season, it has an influence on the resulting ARF values. For example, convective events occurring in the summer periods (April to September) have smaller spatial dependencies compared to frontal events happening in the winter (October to March). The rapid decay in the spatial correlation within convective events results in a smaller ARF. Hence, the seasonal effect is that in the summer period, the ARF is smaller than in the winter period (*Allen and DeGaetano*, 2005). Other factors affecting precipitation formation and hence the derived ARF values are topographic variables such as elevation, slope, orientation, and exposure. These were found to be linearly related to precipitation sums (*Basist et al.*, 1994).

ARF was shown to present spatial heterogeneity depending on the target region and altitude (topography), but is less influenced by the catchment characteristics. In addition, the ARF values are expected to decrease with increasing catchment area and increase with the storm duration (*Veneziano and Langousis*, 2005).
Traditionally, areal rainfall values were calculated from point observations by using, for example, the Thiessen polygon method, the arithmetic mean, or the inverse distance-weighted approach. A common disadvantage of these methods is the lack of consideration of the spatial dependence structure between the point data. Moreover, many extremes might not have been correctly captured by the rain gauges, especially at high elevations, where the network density is usually smaller. For reliable ARF calculation, a long time series is required; this increases the probability of capturing the extremes values over that location.

In the work of *Asquith and Famiglietti* (2000) three different rain gauge networks at the same locations were used and yielded different ARF results due to different measurement instruments. This points out the fact that combining several data sources will influence the derived ARF values. Moreover, different ARF calculation methods yield different results.
Two main methods predominate in the ARF calculation. The first requires long data sets, empirical models (probabilistic approach), and a fixed area. While the second is analytically based. The first is bound to the area but is advantageous because it leads to statistically valid ARF values. The second method is independent of the area but delivers only empirically calculated ARF. In this context, both methods were applied. The ARF for scaling areal-to-point rainfall and vice-versa is defined as follows:

$$ARF(A,d,T) = \frac{R_a(A,d,T)}{R_p(d,T)} \tag{6.37}$$

Where:

$ARF(A,d,T)$	=	ARF for area size A and duration d and return period T
$R_a(A,d,T)$	=	areal rainfall depth (or intensity) for duration d and return period T
$R_p(d,T)$	=	point rainfall depth (or intensity) for duration d and return period T

Traditional use of ARF:

Given a certain catchment with an example size of 1000 km^2 and assuming a DWD rain gauge exists within the catchment, the following steps are needed to transfer the point to catchment scale statistics:

1. Calculate from the DWD rain gauge the DDF curve for different return periods.
2. Calculate the catchment concentration time (for example, 10 hours) (time needed for the farthest drop of water to reach the outlet, function of catchment shape).
3. Depending on the design structure, a return period (T) is selected (for example, T = 5 years).
4. If the catchment has a 10-hour concentration time, find for this duration the estimated rainfall depth (P) using the DDF curve (for example, P = 35 mm).

5. Find the corresponding ARF value depending on catchment size, duration, and return period (for example, ARF = 0.8).
6. Calculate the catchment scale rainfall depth (P_c) by multiplying the estimated DWD rainfall depth by the ARF ($P_c = 35 * 0.8 = 28mm$).
7. P_c is then the design rainfall value to use.

Recently, weather radar data were used to investigate the validity and reliability of ARF values. *Thorndahl et al.* (2019) used a 15-year weather radar data set covering the region of Denmark to derive storm-centered ARF values for short durations and small-scale areas for the design of urban hydrological systems. The radar-based ARF values were compared to several rain gauge area-fixed derived ARF values. It was stated that the radar-derived and storm-centered ARF are smaller than the area-fixed and rain gauge based values.

In *Kim et al.* (2019) radar data were used to extract different storms over the region of South Korea that were used to estimate the ARF from a storm-centered approach. The authors show that the ARF estimates are not only dependent on the area and duration but also on the storm's inner variability. The latter was expressed by the coefficient of variation (CV) and was seen to be strongly related to the derived ARF. As the coefficient of variation increased (hence, more variability), the ARF of the corresponding storm decreased. The authors point to the fact that traditional methods for estimating ARF should be coupled with rainfall spatial variability, made more visible with the use of weather radar data. Moreover, the authors noted a difference between the ARF for circular and elliptical-shaped objects. The ARF for a circular object was seen on average to be 20% smaller than that of an elliptical storm object for the same area size and duration.

6.1.3 Area-Depth-Duration-Frequency curves

Traditional DDF curves are derived from point observations. As seen in chapter 4 these cannot always correctly represent the spatial extent of rainfall. Area-DDF curves are derived similarly to DDF curves, but instead of using point observations, an areal rainfall field is used. For a given area, rainfall observations are used to derive the DDF curve. This is the DDF curve for that specific area, known as the Area-DDF curve. This can be done for different area sizes.

Bennett et al. (2016) first suggested the use of interpolated rainfall data for the direct estimation of the statistics of areal extremes by introducing the IDFA (intensity-duration-frequency-area) curves. The need for ARF to convert point to spatial rainfall could then be eliminated. Radar-derived precipitation data can be used to calculate the ADDF (or AIDF) curves. However, the quality of the radar QPE data and the relatively short observation time period highly influence the results (*Haberlandt and Berndt*, 2016). *Marra and Morin* (2015) used radar QPE to derive IDF curves for the region of Israel and compared the results to nearby rain gauges. Despite efforts to reduce errors in the radar QPE, the final results showed an overestimation of radar-derived IDF curves. An effect that increased with larger durations. Another study done by *Ghebreyesus and Sharif* (2021) used the radar QPE data over the state of Texas to derive state-wide IDF curves. The latter, showed mostly an underestimation of short-duration maxima. The goal of using the radar QPE data is to be able to derive spatially and temporally reliable IDF curves, eliminating the need for area reduction factors (*Ghebreyesus and Sharif*, 2021). In *Bárdossy and Pegram* (2018) the radar data for selected extreme events were used to investigate the cumulative distribution function of point, areal and within an area precipitation extremes. It was shown that the areal precipitation value can exceed the point extreme value and that within an area, the local occurrence of extremes can be problematic. This already shows that the use of the area reduction factor might not always be applicable.

The raw processed radar data, the merged processed radar, and the RadKlim data were used to derive ADDF curves for the radar areas of Hannover and Türkheim for different area sizes. The ADDF curves are used to find the estimated rainfall depth for a given duration over a given area. One would expect that as the area increases, the maximum average rainfall depth is expected to decrease. This is to be tested.

6.2 Results

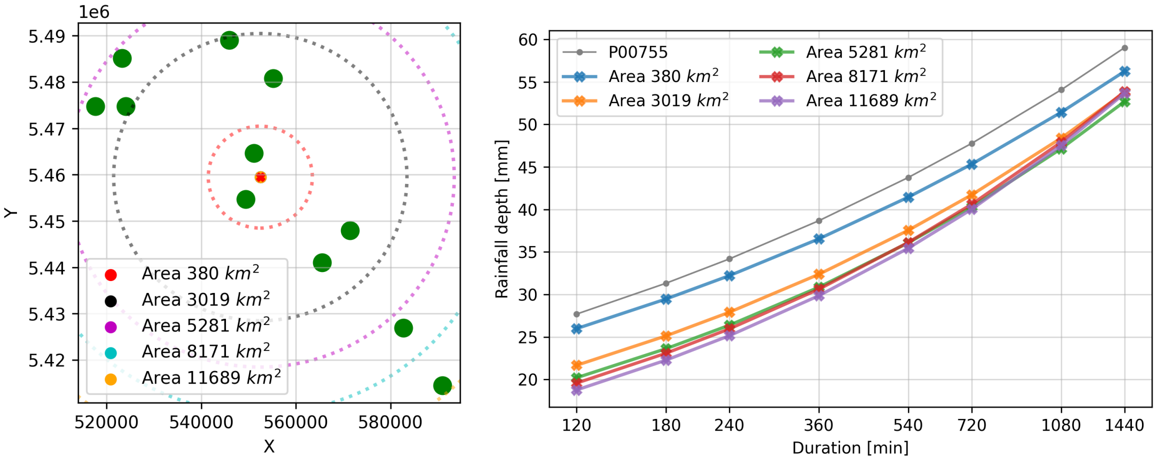
6.2.1 ADDF from station data

Note: The crossings of the ADDF curves were first identified by Golbarg Goshtasbpour *Goshtsasbpour et al.* (2022) during collaborative work within the ClimXtreme project.

The DWD point observations were used as an initial step to go from point DDF to areal DDF curves. For this purpose, several circular areas around a center station were constructed; as the area size increased, stations within the new area were identified, and the average of all included stations per time step was calculated. From the averaged time series, the DDF was calculated and assigned to the corresponding area. The procedure was repeated for several area sizes and center locations. The point DDF was then compared to the area DDF. An example is shown in Fig. 6.2. The curves show unusual behavior; with increasing area size and temporal duration, the ADDF curve of area size 11689 km^2 crosses the one from smaller areas but not the DDF curve. This indicates that larger areas have a higher probability of capturing the event maxima failing within that area. Smaller areas and point data can miss the event maxima, and hence the yearly maxima could be underestimated.

6.2.2 ADDF from merged and Radklim data

The merged weather radar and station data developed in Chapter 5 were used to derive ADDF curves for the radar regions of Hannover and Türkheim. Several randomly selected points were chosen as center of the areas. These were constructed as squares and circles with sizes ranging from 4 km^2 to 1024 km^2. For every area size, the pixels falling within the area were identified, and for every time step, the precipitation average over all pixels was calculated. From this new time series, the DDF was calculated for different durations



(a) Spatial locations

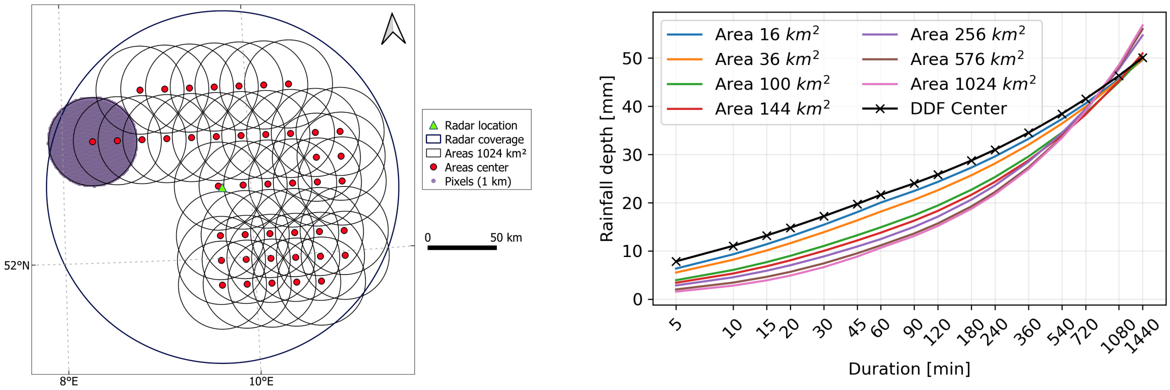
(b) ADDF for T=10 years

Figure 6.2: Panel (a) shows the location of the created areas around the center point. Panel (b) is an example of the calculated DDF and ADDF curves for the denoted areas in (a). The DDF curve was calculated from the station data. The different curves represent the ADDF values for the corresponding areas.

(5 minutes to daily) and assigned to the corresponding area size. This approach is beneficial and allows statistical evaluation of the areal precipitation time series as a point-time series. The ARF was then calculated as the ratio between the point-DDF and the area-DDF values for the corresponding duration and return period.

In Fig. 6.3 the results for one areal configuration are presented. The black curve refers to the point DDF derived from the area center. The different colored curves represent the ADDF curves for the corresponding area sizes. For a short duration, the point DDF exceeds all areal DDF curves. The ADDF curves follow a decreasing gradient with increasing area size. However, starting from a duration of around 18 hours, the ADDF curve crosses with the point DDF curve. Beyond this duration, the ADDF exceeds the DDF values. This is an interesting result, contradicting the ARF concept. Larger areas have higher rainfall depth values than point rainfall values. This behavior was observed in almost all considered locations. Note that the south-west region of the radar coverage area was not considered due to the presence of a beam blockage that hindered the correct sampling of QPE values.

After repeating the previous procedure for several randomly selected locations within the radar area of Hannover, the average ARF for every duration and area size could be calculated. The results are displayed in Fig. 6.4. Panel (a) shows for each duration, the relation between the area size (x-axis) and the ARF values (y-axis). For example, for the 5-minute duration, the reduction increases with increasing area size. In other words, for the same duration, as the area size increases, the areal extreme decreases as compared to the point extreme. For small areas (4 km^2), the areal mean is quasi-equal to the point value, however, for large areas (1024 km^2) the reduction is the largest, with an average ARF value of 0.45 for the 5 minute duration. For the daily duration, the ARF is rather independent of the area size, with an average value of 0.95.

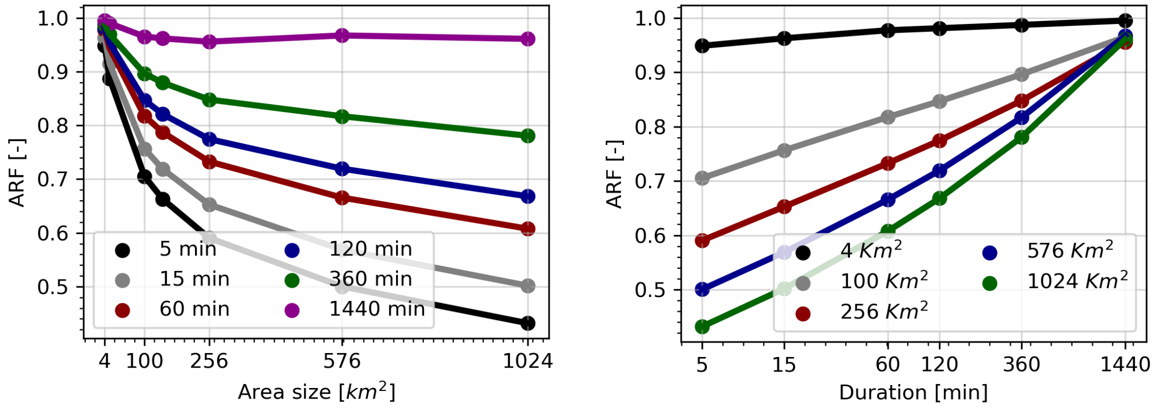


(a) Spatial locations A=1024km^2

(b) ADDF for T=10 years

Figure 6.3: Panel (a) shows the location of the selected largest areas within the radar area of Hannover along their center point (denoted by the point in red). Panel (b) is an example of the calculated DDF and ADDF curves for the denoted areas in (a).

In the right panel of Fig. 6.4, the relation between the duration (x-axis) and the ARF (y-axis) is presented. For a constant duration, for example, 15 minutes, the reduction is proportional to the area size. Specifically, the larger the area, the larger the reduction. For the duration of 18 hours (1080minutes), a crossing between the curves of the larger areas (1024 km^2) and the curve of the smaller areas (100 km^2) was noted. This contradicts the traditional assumptions behind the ARF, namely that larger areas have smaller ARF values. A discussion about this is presented later in the discussion section.



(a) ARF with area

(b) ARF with duration

Figure 6.4: Example of average ARF derived from several randomly selected locations within the radar area of Hannover. Panel (a) shows the ARF values (y-axis) as a function of the area size (x-axis) for different durations. The different curves represent the ARF for the specific duration as a function of the area. In panel (b) the ARF is displayed as a function of the duration (x-axis). The different curves represent the ARF for selected area sizes.

6.2.3 ARF from ADDF curves and connected areas

ARF shows high uncertainty even for neighboring locations; a regional average ARF value leads to an underestimation or overestimation of the areal extremes for design values. This can be seen in the left panel of Fig. 6.5. From 51 ADDF locations, the ARF value was calculated and is presented for the area size of 1024 km^2. The average ARF value is presented in red, while each gray curve presents the estimated ARF for that specific location. One can see that the spatial variation is not included using a single ARF value, and the ARF is not stationary in space but presents larger variations over all durations. Moreover, there are regions and durations with ARF values exceeding the value of 1. This indicates that the point sampling underestimates the true maxima, especially for long durations. Depending on the spatial extent of rainfall events, the areal average can exceed the point value. Other (but few) studies have already noted an ARF value exceeding the pre-assumed upper bound of 1. For example, *Catchlove and Ball* (2003) and *Bennett et al.* (2016) identified ARF values exceeding 1 in their works. Especially when deriving regionalized IDF or ARF values, the underlying regionalization scheme underestimates the rainfall extremes and spatial variability (as mentioned in chapter 4).

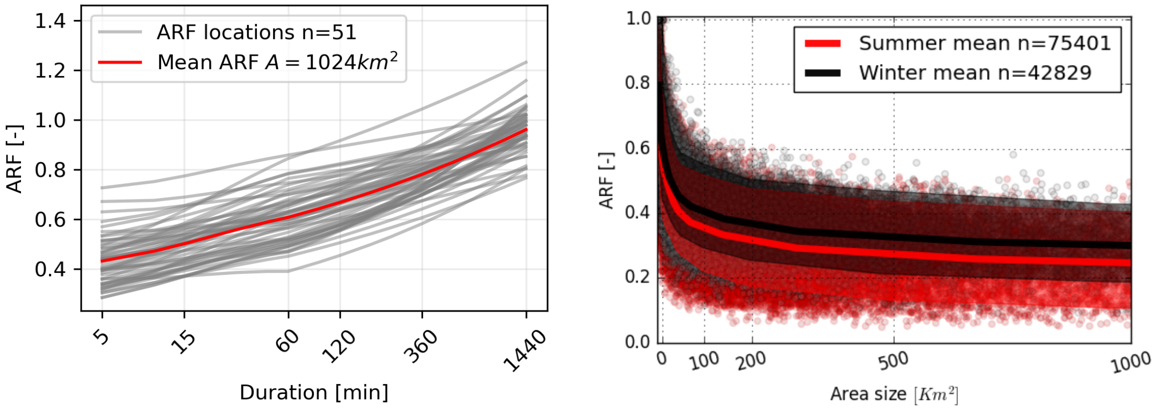


Figure 6.5: Panel (a) depicts the ARF for several areal locations using the ADDF curves. Panel (b) shows empirically calculated ARF values for a duration of 5 minutes.

ARF from storm-centered approach:

Another aspect was to investigate the storm-based ARF values. For this, connected rainfall blocks were extracted from the radar fields (see chapter 7). The ARF was calculated as the ratio between the mean and maximal within-area precipitation value and was assigned to the corresponding area size. This is an event-based approach with varying results depending on convective and stratiform rain formations and is not suitable for statistical analysis. However, the derived values were used as a case study to determine the reliability of the statistically derived ARF values. The results show large variation; for a similar area size, several ARF values were obtained. In practice, however, only the average of these values is considered, which can lead to an underestimation in many cases. The right panel of Fig. 6.5 shows average empirically calculated ARF values from connected rainfall blocks exceeding a size of 4 km^2 within the Neckar catchment for a duration of 5 minutes. The values show the variation in the ARF values. For similar area sizes, the values vary and spread greatly around the mean and show a clear distinction between the summer (values in red) and winter (values in black) seasons. The average ARF value (0.42) in panel (a) for a duration of 5 minutes and area size of 1024 km^2 is larger than the estimated one in panel (b) (winter 0.35 and summer 0.25). In other terms, the statically derived ARF delivers larger areal rainfall values than the empirically derived values. A similar result was noticed by *Thorndahl et al.* (2019), where the radar storm-centered ARF were smaller than the fixed area ARF values. However, these are average values in both cases.

6.2.4 Comparing derived ARF to ARF used in practice

In the work done by *Verworn* (2008) areal reduction factors for area sizes ranging between 25 km^2 and 1000 km^2 were determined using a rain gauge network with high spatial and temporal resolution. The large number of rain gauges covering several regions meant that the obtained ARF values were adequate to be transferred to other locations. In this work, these values were considered as reference values and compared to the derived ones using the ADDF curves. The result is displayed in Table. 6.1, the columns show the area size and the rows the ARF values. For every duration, the results were compared. For lower temporal aggregations (hourly to daily), the difference between the calculated and reference ARF values was in the same range. The noticeable difference was, however, that for sub-hourly aggregations (first row) and large areas, the reference values show lower ARF values. This leads to an underestimation of the areal extremes. For the area of 1024 km^2 the reference value shows an ARF value of around 0.36, while the calculated average ARF is 0.5. If the point extreme value for the same duration was, for example, 50 mm, the resulting areal values will be 18 mm and 25 mm, respectively. Hence, an underestimation of the areal extreme. The reason for the difference is mostly due to the presence of convective-driven events that are mostly better detected with the weather radar than with the station data. Though the true ARF value cannot be exactly pointed out, this analysis showed the uncertainty in the derived and used ARF values, especially for the sub-hourly durations.

Table 6.1: The reference and calculated ARF values for different durations and area sizes are shown in the red-titled and black-titled columns, respectively.

		Area size [km^2]							
		10	16	**50**	100	**200**	256	**1000**	1024
Dur. $[min]$	**15**	0.91	0.91	0.74	0.76	0.56	0.65	0.36	0.50
	60	0.95	0.94	0.84	0.82	0.70	0.73	0.56	0.61
	360	0.98	0.97	0.95	0.90	0.88	0.85	0.80	0.78
	1440	0.99	0.99	0.98	0.97	0.94	0.96	0.90	0.96

6.3 ARF with catchment size, form, and orientation

Change of ARF with catchment size and form:

Gravelius (1914) introduced the Gravelius Compactness Coefficient (GC) to describe the catchment shape defined in equation 6.38 (*Sassolas-Serrayet et al.*, 2018).Several catchments in the Germany were selected and their GC value calculated. Panel (a) of Fig. 6.6 shows the different catchments along their shape parameter. Higher GC values indicate a more longitudinal (and non-uniform) form (catchments in orange and red). Smaller GC values refer to more homogeneously spatially formed catchments (locations in blue). Using the RadKlim data, for each catchment and for the period 2001–2020, the precipitation time series for the catchment centroid pixel and for the areal average precipitation value were extracted. For the centroid pixel, the DDF curve was constructed, and from the areal average, the ADDF curve was derived. The ARF was then calculated as the ratio between the DDF and ADDF curves. The aim was to identify if there is a relationship between the catchment size and form and the ARF value. Panel (b) in Fig. 6.6 depicts the hourly and daily duration of the ARF value as a function of the catchment area size for a 10-year return period. For the hourly duration (dots in blue), the ARF decreases with increasing area size and shows variability for small-scale catchments. However, as seen in panel (b) of Fig. 6.7 neighboring catchments show similar ARF values. On a daily scale, some catchments have ARF exceeding the value of 1. This is most likely due to undersampling of the extremes at the centroid location. Another possibility would have been to derive the DDF from the pixel with the largest mean. For the daily duration, the ARF seems to be less influenced by the area size.

In panel (c) of Fig. 6.7, neighboring catchments tend to have similar ARF values with minor exceptions. For instance, in the Neckar catchment, the Jagst (in blue) and the Enz (in orange) sub-catchments show different behavior than the other sub-catchments. However, the difference is the largest for the Jagst catchment, with a much smaller ARF on a daily scale. The subplot in panel (c) of Fig. 6.6 shows the effect of the form (or shape) parameter on the ARF. The catchments with an area size exceeding 1000 km^2 were selected, and the relationship between the ARF and GC was plotted. The size and color of the dots are proportional to the area. This explains the different behavior of the Jagst catchment with the largest GC value of 3.1, associated with a small ARF. Moreover, catchments with similar sizes and smaller GC values have larger ARF values. Hence, as mentioned by *Veneziano and Langousis* (2005) the shape of a catchment affects the ARF, and longitudinal catchments are expected to have a smaller ARF value compared to regularly shaped catchments of similar size. Additional factors such as topography and elevation affect the areal rainfall statistics and the derived ARF values (*Basist et al.*, 1994).

$$GC = \frac{P}{2\sqrt{\pi A}} \tag{6.38}$$

Where:

GC = Gravelius compactness coefficient
A = Area of catchment
P = Perimeter of catchment

Change of ARF with catchment orientation:

To determine if the ARF is constant or dependent on the underlying shape and its orientation, the following analysis was conducted: Several object forms with the same area size (around 1024 km^2) were created around the same centre point. For each form, the ADDF curve was calculated and the ARF was derived as the ratio between the ADDF and centre pixel DDF curves. The created forms were the circle used previously for deriving the ADDF curve and a horizontal rectangle (width = 20 km and length = 51.2 km). The horizontal rectangle (east-west) was then rotated vertically (south-north) and in the directions $\pm$ 45 °. Eventually, five forms were created with equal size and centre point.

The ARF values for the different shapes and durations corresponding to the return period of 10 years can be seen in Table. 6.2. The columns to the right display the average, minimal, and maximal ARF values for each duration. Since the DDF curve was constant, the difference in the ARF values was related to the average areal rainfall within each form. The minimum and maximum columns show the range of possible ARF values with a 20 % deviation. Moreover, the maximum ARF was associated with the shape of a circle, while the minimal was associated with the rotated rectangles. This indicates that, for the same area size, different ARF values resulted from different geometrical forms, with the maximum associated with a circular form and the minimum with a longitudinal shape.

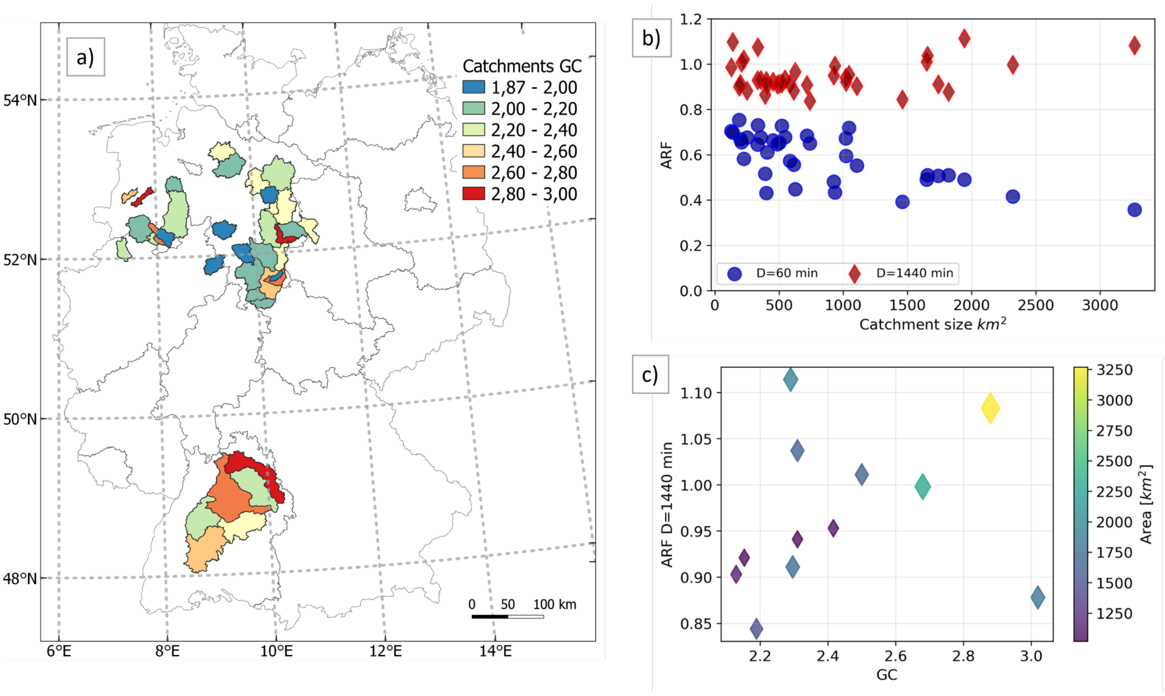


Figure 6.6: Panel (a) shows several selected catchments with the GC coefficient. Panel (b) shows the ARF values for the hourly and daily durations as a function of area size. Panel (c) shows for the daily duration the GC coefficient for the large catchments as a function of the ARF.

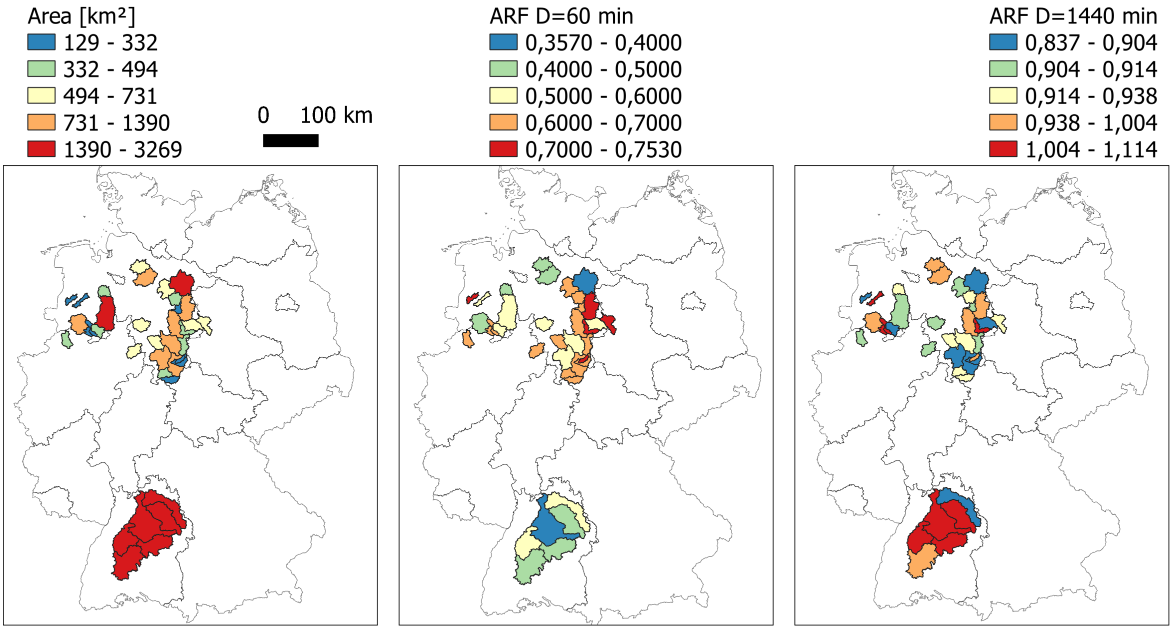


Figure 6.7: Panel (a) shows several selected catchments with their area size. Panels (b) and (c) show the ARF values for the hourly and daily durations, respectively.

Table 6.2: ARF values for several durations and for the same area size but for different shapes (seen in panel (a) of Fig. 6.8). The average, minimal and maximal ARF values were calculated for every duration over the different shapes.

		Area Reduction Factor (ARF) - Area sizes of 1 and 1024 km^2							
		Circle	**Horizontal**	**Vertical**	**Shifted +45°**	**Shifted -45°**	**Average**	**Min**	**Max**
Duration [min]	**5**	0.197	0.175	0.174	0.156	0.182	0.176	0.156	0.197
	15	0.278	0.249	0.247	0.233	0.255	0.252	0.233	0.278
	30	0.353	0.32	0.315	0.306	0.32	0.323	0.306	0.353
	60	0.437	0.4	0.392	0.393	0.388	0.402	0.388	0.437
	120	0.53	0.489	0.49	0.489	0.479	0.495	0.479	0.53
	360	0.723	0.68	0.698	0.695	0.672	0.693	0.672	0.723
	720	0.881	0.84	0.875	0.869	0.834	0.859	0.834	0.881
	1080	0.99	0.953	0.998	0.991	0.947	0.975	0.947	0.998
	1440	1.076	1.043	1.096	1.088	1.037	1.068	1.037	1.096

Considering the vertical and the ± 45 ° rotated rectangles, the ARF was the largest for the vertical orientation. For a short duration (below 60 minutes), the +45 ° rotated rectangle has a larger ARF than the -45 ° rotated rectangle. Beyond the hourly duration, the ARF becomes larger for the -45 ° rotated rectangle. This indicates that, depending on the duration, orientation plays an important role.

Panel (a) of Fig. 6.8 shows the accumulated sum of a 4-hour event that moved from South-East to North-East in July 2019. For every 5-minute time step, the ratio between the accumulated areal rainfall in every shape (the black polygons in panel (a)) and the accumulated centre pixel rainfall value was calculated. The results are seen in panel (b) of Fig. 6.8. As the event spread, the ratio (namely, the ARF) neared the value of 1. However, the development of each curve was different depending on the shape and orientation. The vertical shape reaches the ARF of 1 very quickly and then fluctuates around that value. This is because the vertical shape was oriented in the same direction in which the rainfall field started and developed. Hence the order of the curves: after the vertical shape, the - 45° rotated rectangle, then the circle, then the horizontal shape, and finally the +45 ° rotated rectangle. The rainfall field reached the latter towards the end of the event. This analysis indicated that the ARF depended on the storm orientation (or anisotropy) compared to the catchment orientation. If the moving storm was perfectly oriented with the catchment orientation, the ARF would reach a value of 1 very quickly. However, if the storm orientation was perpendicular to the catchment, the ARF would gradually increase along the spatial-temporal development of the event.

In *Veneziano and Langousis* (2005) the effect of advection on ARF values was investigated. The authors stated that depending on the catchment shape the effect of advection differed. For instance, for highly elongated catchments, not only does the magnitude of advection matter but also the direction. Whereas in more circularly shaped basins, only the magnitude of advection and not the direction were relevant. This supports the results displayed in Table 6.2, where for the same area size and centre DDF curve, the ADDF varied depending on the object form (geometry) and orientation. The largest ADDF (or ARF ratio) was seen for the circle shape, while the smallest was seen for the elongated and oriented shapes (associated with large GC values).

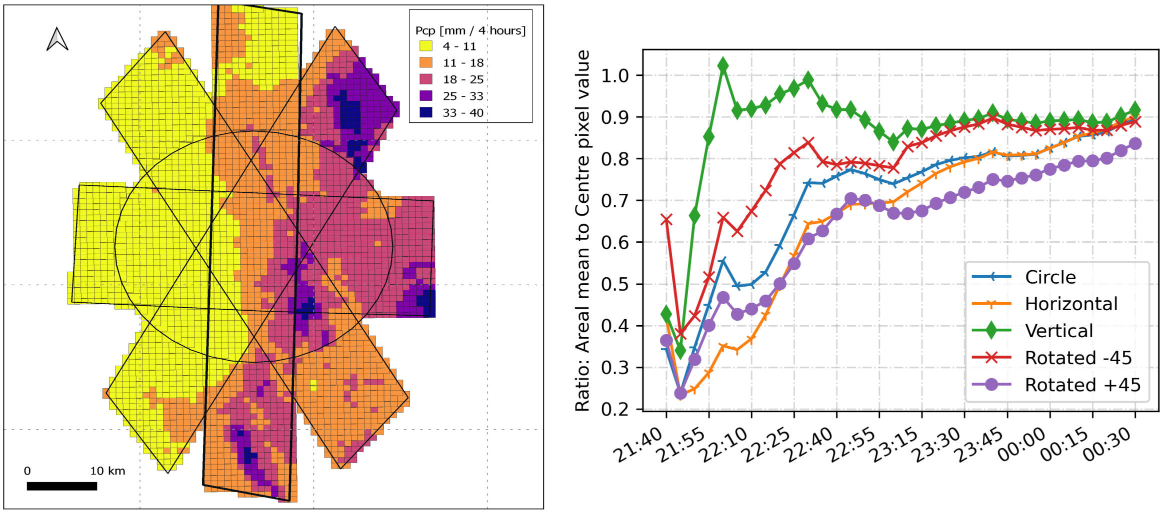


Figure 6.8: Panel (a) shows the accumulated values for a single 4-hour event. Panel (b) displays the ratio between the areal average and centre pixel rainfall values for each shape and time step for the event in Panel (a).

6.4 Discussion

Uncertainty in DDF curves: To illustrate the degree of uncertainty present in the estimation of the DDF curves, point-time series from Radklim were used. The DDF curves were derived for all pixels within an area size of 16 km^2. Assuming that the centre pixel represented the location of a rain gauge, the DDF of all other pixels was compared accordingly. Fig. 6.9 shows the result and hence the degree of uncertainty in representing the correct extremes even for such a small area. The DDF curves show a high degree of variability and, hence, uncertainty. This has the following consequences: if the rain gauge was placed at the location of the minimum or maximum DDF curves, the subsequent design values would vary substantially. Even calculating the ARF value for such an area showed a large deviation depending on the choice of the 'point' or pixel DDF curve. In the left side of Fig. 6.9 the red curve represents the center point DDF curve, while the blue shows the ADDF for the corresponding area. The plot on the right side shows the location of the selected pixels in the orange square within the radar area of Hannover (pixels in blue).

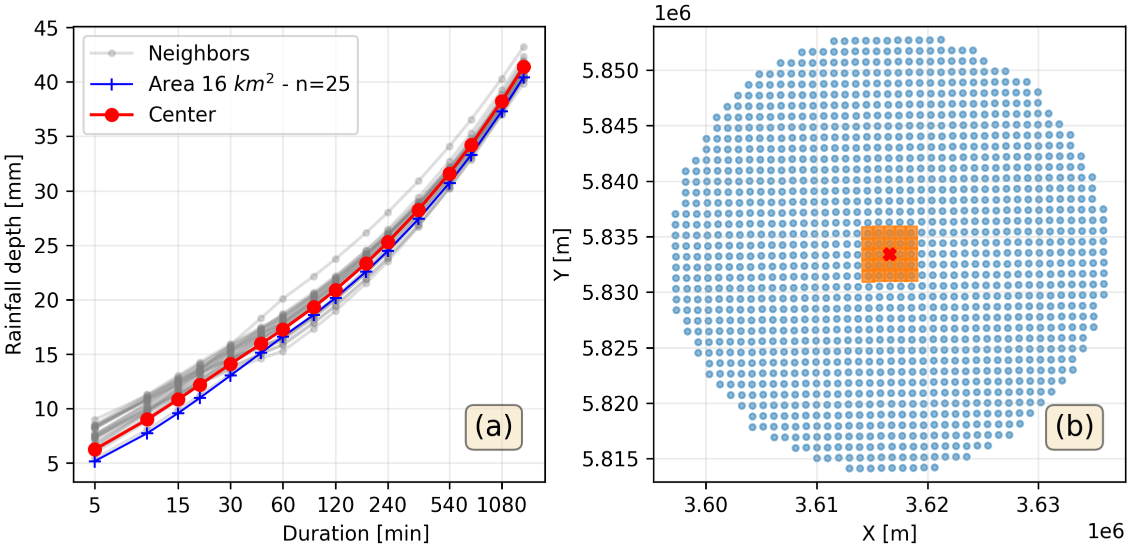


Figure 6.9: Example of calculated DDF curves for all 25 pixels within an area of 16 km^2.

Uncertainty in ADDF curves: One of the main findings with ADDF curves was the presence of crossings between the curves. Large areas have values exceeding the point values for higher aggregations. Hence, the ARF value is not relevant in this case. As seen in the results, crossing the ADDF curves, first between the point and area curves, and then between the ADDF of different area sizes, limits the current usability of the ADDF curves.

For the crossings between point DDF and the ADDF curve, the reason is an underestimation of the maxima by the point value, especially for a large duration. One example of this is an event from the year 2003, which can be seen in Fig. 6.10. The left panel shows the event time series, and the right panel shows the event spatial distribution. The event maxima (curve in blue) was missed by the center value (curve in red). The accumulated sum of the areal average (black curve) exceeds the point value, which results in crossings of the DDF and

ADDF curves. The spatial distribution of the event shows a concentration of the values on the left and right sides of the area centre (represented by the x pixel). The event was seen to be associated with anisotropy and uneven spatial distribution.

In addition, depending on the length of the available data sample, crossings might not occur. For example, if the event in 2003 did not exist, the crossings would not exist. Moreover, in the left panel of Fig. 6.9 the choice of the center point and the subsequent effect on the derived DDF curve can be seen. Usually, the center point is considered the point representative of the area or catchment, however, choosing a different point than the center will lead in some cases to a better sampling of the extremes within the area and hence no crossings between the DDF and ADDF curves.

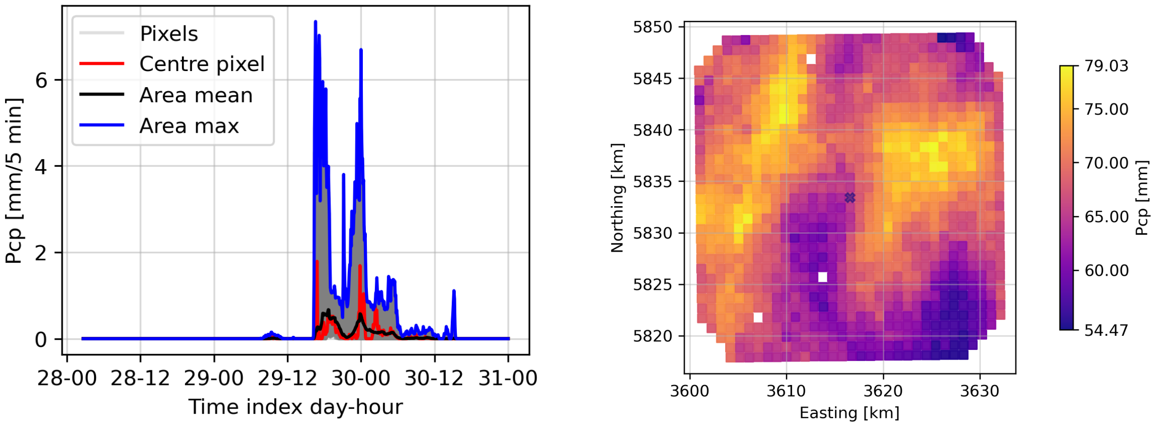
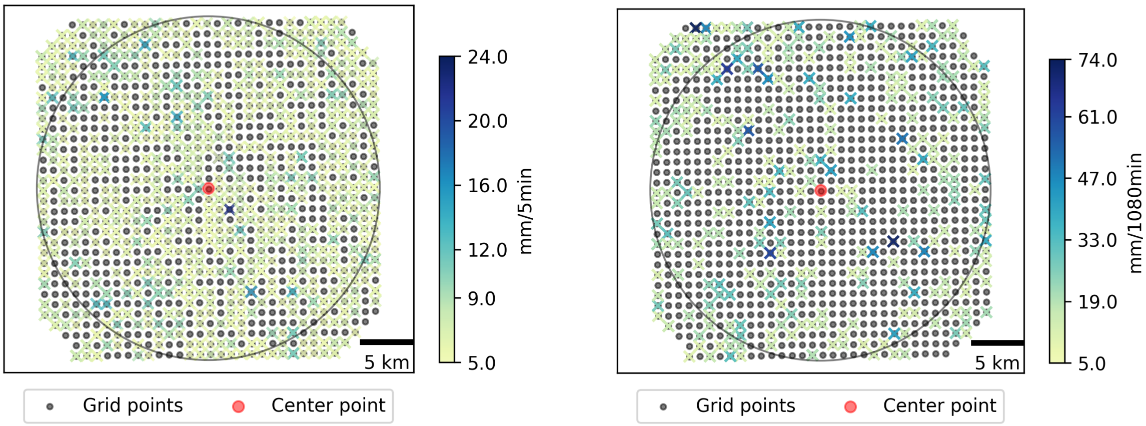


Figure 6.10: Panel (a) depicts the time series of the event in 2003. Panel (b) shows the spatial distribution of the precipitation sum over the event. The centre pixel is marked by **X**.

Where is the maxima within the area (?): Fig. 6.11 shows the location of the maximum pixel value within the area of the ADDF curve. The center point is shown in red. For every time step between the years 2001 and 2021 with data exceeding the minimal threshold of 5 mm per 5 minutes and 10 mm per 18 hours, the location of the maximum was marked. Both panels demonstrate that the occurrence of the maxima is independent of its location. This explains why sampling the DDF from any point value might lead to missing or catching the event maxima. For example, the maxima occurred over some pixels more frequently than over others, but interestingly, some locations never recorded any maxima (grid points in gray).

Uncertainty in ARF values: The assumption that a single ARF value can represent a region was not confirmed. ARF was seen to be highly variable in space and time. Even within a single radar area with similar duration and area size, the ARF values vary significantly. Moreover, a test case study was done using the derived ARF values. From the connected areas, the values of the areal maximum, areal mean, and area size are known. Hence, the empirical ARF value for that area could be calculated. Using the values derived from all connected areas with similar sizes and for a similar duration, an average ARF was obtained. When using the average ARF value (which is the case in practice), the uncertainty space was



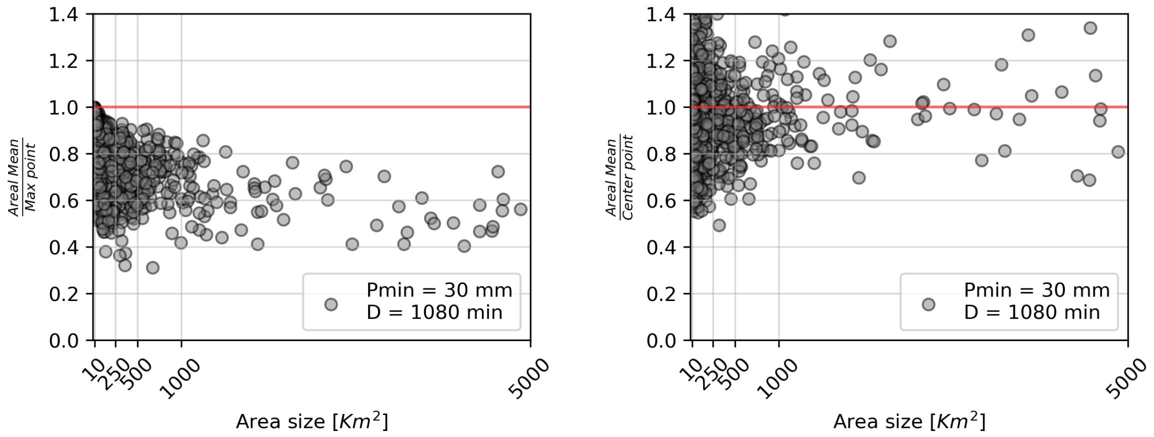
(a) Location of maximum (d=5 minutes) (b) Location of maximum (d=18 hours)

Figure 6.11: Panel (a) and (b) depict the location of the maximum values ($>$ 5 mm per 5 min; $>$ 10 mm per 1080 min) within the area of 1024 km^2 independent of their time of occurrence for the durations of 5 minutes and 18 hours respectively.

quite high. For example, a connected area with a size of 500 km^2 was considered, for which the mean ARF was around 0.35 for the 5-minute duration. As seen in the right panel of Fig. 6.5 the calculated ARF varies between 0.1 and 0.6.

A further investigation was done to showcase that, independently of the location of the maximum, the in-area maxima exceed the areal mean. However, if the location of the representative point was confined to a certain fixed location, for example, the center of the area, the areal average might exceed the point value. Fig. 6.12 presents the previous two cases, respectively. Panel (a) depicts the ARF calculated as the ratio of areal average to maximum precipitation values from 2D connected areas exceeding the $>$ 30 mm per 1080 min threshold for the year 2021 within the Neckar catchment. Panel (b) shows the ratio between the areal mean and area maximal value, for the same period and location. Panel (a) shows that the area maximum is always larger than the mean. Panel (b) shows that the areal mean might exceed the area center rainfall values in some cases. Hence, the assumption that the areal average is always smaller than the in-area maximum was confirmed by panel (a). However, the assumption that the center point (or any other single point) is representative of the in-area maximum precipitation value is wrong, as seen in panel (b).

Dependence of ARF on return period: for several selected catchments (in total 37) with various area sizes and geographical locations (in the North and South of Germany) the ARF was derived as the ratio between the catchment specific ADDF curve and the DDF curve of the centroid pixel. For that purpose, the RadKlim data for the period 2005–2020 were used. The hourly and daily durations were selected as test cases. The ARF values for the return period of 10 years were chosen as reference values. The ARF for the return periods of 30, 50, and 100 years were compared against the reference values. The results show that the ARF is quasi-independent of the return period. Between the different return periods, the values vary slightly, with differences of 5 %. As the return period increases, the ARF decreases. Similar conclusions were found in the literature review, for example, in *Svensson and Jones*

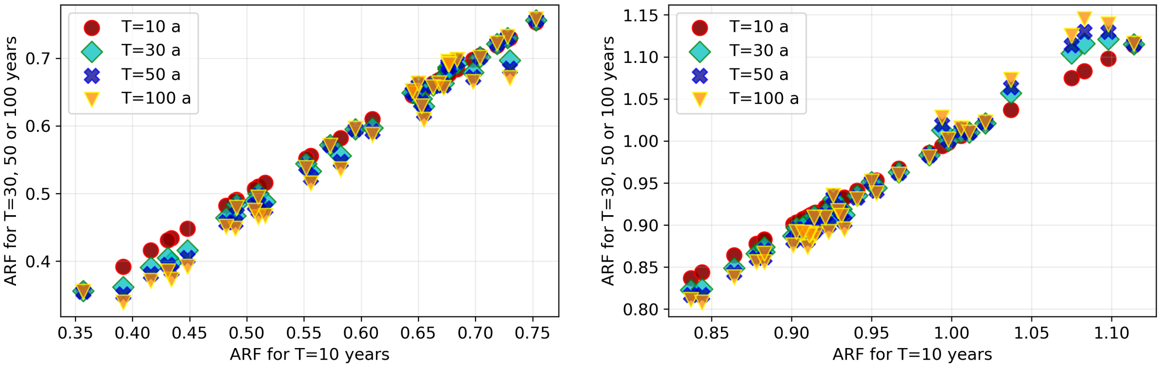


(a) Ratio areal to maximum values

(b) Ratio areal to center values

Figure 6.12: Panel (a) depicts the ARF calculated as the ratio of areal average to maximum precipitation values from 2D connected (> 30 mm per 1080 min). Panel (b) shows the ratio between the areal mean and area maximal value.

(2010). The larger decrease in the hourly ARF values was associated with the greater role convection plays in hourly heavy rainfall values. For catchments where the ARF is larger than one, namely where the centroid is not representative of the point maxima, the ARF increases with the return period. This behavior, however, was associated with sampling problems of the point DDF curve (as seen in previous sections). Moreover, the decrease in the ARF was noted to be larger for the hourly duration (panel (a) of Fig. 6.13) compared to the daily duration (panel (b) of Fig. 6.13). Note that the bias in the RadKlim data was largely canceled out as all results were derived from the same database.



(a) ARF with return period - D=60 min

(b) ARF with return period - D=1440 min

Figure 6.13: Panel (a) and (b) depict the ARF values for 37 catchments and for 4 different return periods for the hourly and daily durations, respectively.

6.5 Summary

Extreme precipitation values associated with a given duration and return period are essential for the design of different engineering systems. In this chapter, traditional point gauges were used to calculate the DDF curves, and from the radar data, the ADDF curves were derived. The ADDF curves refer to design curves that can be used directly for the respective area size. These would have been traditionally acquired using the area reduction factor. From the ADDF curves, area reduction values were derived and compared to reference values. The result highlighted that the ADDF-based ARF values are larger for short durations, which implies higher areal precipitation depth. Moreover, the ARF was derived from different storms extracted from the radar data. The average ARF for every area size and duration was calculated. If the average ARF value is used, large uncertainty exists in estimating the true storm maxima. Large uncertainty was also identified in the DDF curves within a small proximity. Neighboring locations could have varying DDF curves, depending on the sampled extremes. Thus, a long sampling period is essential to ensure that a location has observed the maximum of the maxima for each duration. The analysis undertaken showed that applying the ARF might not always be reliable and can lead to an underestimation of the areal precipitation values. In addition, the shape, orientation and geometry of a catchment or area affects the ARF value. A factor with important aspects for urban areas.

7 Areal extremes from a different perspective: rainfall as connected 2D and 3D blocks

In the following chapter, the spatial extent of precipitation events was inspected using the weather radar data made available for the last 20 years with 5 min temporal and 1 km spatial resolutions. Connected rainfall areas were extracted from every radar field. Corresponding statistics regarding area size and areal mean rainfall, area size and areal max rainfall, and spatial density distribution were calculated. For selected events, the relationship between rainfall volume and discharge volume was established. The study region was Germany, with headwater catchments in the Neckar catchment.

7.1 Literature review

Precipitation extremes are traditionally assessed using single-point observations. Corresponding extreme value statistics are applied to these point observations and aggregated for different duration levels. However, the consequences of heavy precipitation, such as flooding, are related to water volume. Hence, the spatial aspect should not be disregarded. However, most studies with observed series neglect the spatial extent of precipitation events. The focus is on viewing precipitation as a spatial phenomenon without using purely point statistics.

Weather radar data have advantageous aspects when looking at the spatial extent of rainfall events. They offer the possibility to investigate what was measured between point observations and offer greater variability than using interpolation methods. In *Lengfeld et al.* (2019) the spatial extent of hourly and daily data was derived for the period 2001–2016 from the German weather radar data (Radklim). The goal of the study was to find regions with highly correlated behaviour of rainfall extremes. This was done by calculating the rank correlation between the positive values of each grid cell (smoothed to 5km grid size) and all other cells with simultaneously positive values. The distance from which the correlation drops was used to reflect the spatial extent between the reference and neighbouring locations. As in *Touma et al.* (2018), the spatial extent of the daily data was seen to be strongly dependent on topography, and for the hourly data, the extent was on average four times smaller than for the daily data and affected local regions. In *Lengfeld et al.* (2018) rainfall events exceeding a certain threshold or return period were extracted from the radar data for the region of Germany. The events were extracted as contiguous rainfall objects over several durations. The events were looked at in time to ensure that no event was repeated for different durations. The events were then classified based on their time of occurrence, size, and location. Eventually, the authors provided a catalog of extreme rainfall events in Germany. In *Matte et al.* (2022) the spatial extent of extremes was derived from EURO-CORDEX climate model data with an 11 km spatial and a daily temporal resolution. One of the research questions was if the spatial extent of extremes, especially in terms of size, was changing with climate change. The authors investigated and showed that the intensity of larger areal events was getting bigger with the increase in area size.

Based on the concept of trading space for time, the probability of the occurrence of an event with a 5-year return period over a certain location was proportional to the area size. *Bárdossy and Pegram* (2018) demonstrated this by using a high-density network of rain gauges associated with weather radar data. This shows that the occurrence of an extreme event is not only a time-dependent problem but a spatial one as well. A previous study done by *Doneaud et al.* (1984) established a power-law relation between the storm area, duration, and corresponding volume. *Atlas et al.* (1990) analysed the behaviour of convective storms and expressed it in terms of their duration (or lifetime), their areal extent and their volume. *Ayat et al.* (2022) used sub-hourly radar data within an object detection and tracking approach to investigate the presence of a trend in the intensification and occurrence of sub-hourly extremes. Traditionally, investigating the relationship between rainfall and runoff was done using interpolated gauge observations. More recently, weather radar data were used as input for runoff simulations. In both cases, however, the focus was on high-intensity events. An aspect not often considered is that the accumulation of low-intensity rainfall over a certain catchment for a duration long enough could lead to a flood situation. By accumulation, not only are temporal aggregations meant, for example, as hourly or daily maxima, but the accumulation of events individually, defined by the start and end of an event. This differs for each event and offers the possibility of investigating events with varying durations, intensities, and spatial distributions. The focus was usually on hourly or daily maxima; however, a data-based analysis might lead to identifying events with low intensity and long duration, leading to high discharge values.

Why connected areas or volume? Point observations offer a perspective on surface precipitation. Namely, what reached the ground at the gauge location. Numerical climate models offer spatially distributed fields but are the output of meteorological products. They provide precipitation from an atmospheric perspective. The weather radar data offers the opportunity to look at spatially distributed rainfall fields with high resolution in space and time. This is advantageous as it allows for covering the gap between rain gauges, satellites, and climate model data. More specifically, the high spatial and temporal resolution can be used to track the rainfall objects in space and time and apply an object-oriented analysis. Interestingly, the location of the maximum rainfall value within an event can be tracked, and the total event volume can be evaluated. This approach is not feasible using any other observation data. In *Langousis* (2005) the effect of advection velocity and rainfall intensity on subsequent discharge was underlined. The largest effect was associated with slow-moving objects with a velocity of around 10 km/h. Such information can be derived from the connected volumes and related further to the areal extent of extremes and discharge statistics.

In the following chapter, an approach to investigating the spatial extent of extremes by considering rainfall as spatially (2D) and temporally (3D) connected blocks is presented. The focus was not only on singular rainfall blocks but also on the statistical relation between different area sizes (and corresponding areal mean and areal maxima) and their spatial density (frequency of occurrence). Moreover, from the connected areas, the ratio between the areal mean and areal maxima values was derived. This is related to storm-based area-reduction factors (see chapter 6). After introducing the data and the used method, the results were mapped and discussed for the study area. An example of 3D rainfall blocks for sub-catchments in the Neckar catchments is shown, and a relation between spatial extent, rainfall volume, and event duration is presented.

7.2 Method

7.2.1 Data and study area

In this study, the RadKlim data set made available by the DWD (*Winterrath et al.*, 2018b) were used. The data have been climatically and gauge adjusted. The product is derived from the combination of all the weather radar data operated by the DWD. A complete description of the RadKlim data quality can be found in (*Kreklow et al.*, 2020). For processing the data from binary into yearly netCDF files, the software package *radolan_to_netcdf* was used (*Chwala and Polz*, 2021). The data have a 5 minutes temporal resolution and are gridded on a Cartesian grid with the dimensions of 900*1100 kilometers.

The catchments Enz, Jagst, and Kocher are part of the Neckar catchment, located in the southwest of Germany. They have a drainage area of 1.655, 1.820, and 1.943 km^2, respectively. The upper right panel of Fig. 7.1 shows the location of the three catchments along their gauging stations. The catchments are located near two mountainous regions, namely the Black Forest on the west side and the Swabian Alps on the east side. The elevation ranges from approximately 200 to 1000 meters above sea level. The three headwater catchments are used to compare the precipitation volume to the discharge volume. The discharge data were made available by the state environmental agency (LUBW) for the period 2010–2018 with an hourly resolution (*LUBW Landesanstalt für Umwelt*, 2019). The Ahr catchment, shown in the lower right panel of Fig. 7.1 was hit in July 2021 by a severe flooding event that caused large-scale damages. This event was taken as an example case study.

7.2.2 Definition of connected areas and volumes

A connected area was defined by neighboring connected pixels with values exceeding a predefined minimal threshold.

1. For time step t, identify all pixels U in the rainfall field exceeding the minimal threshold P_{min}.
2. Find which pixels in U are connected with one another.

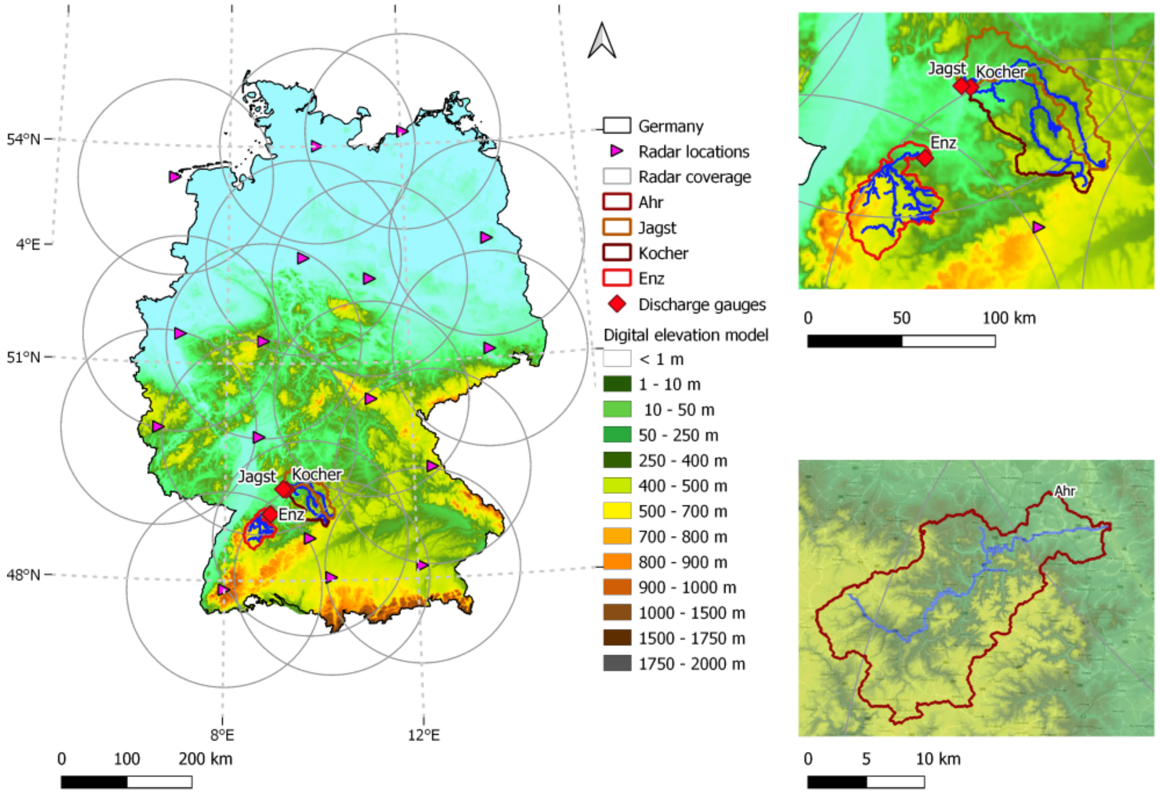


Figure 7.1: Study region with the locations of weather radar coverage areas.

3. Given a pixel $Z(x_0, y_0)$ at locations x_0, y_0 with a precipitation value $Z >= P_{min}$

$$
\begin{aligned}
&Z(x_0, y_0) \text{ connects } Z(x_{0+i}, y_{0+j}) \text{ if} \\
&Z(x_{0+i}, y_{0+j}) >= P_{min} \\
&\text{for i, j in [-1, 0, 1]}
\end{aligned}
\tag{7.39}
$$

4. Repeat the procedure for the next pixel $Z(x_1, y_1)$ until $Z(x_u, y_u)$ is not connected to any other pixels
5. This is a single **2D connected area** defined as $Z(x, y)$ with the number of pixels $i = 1, \ldots, u$ and the coordinate vectors x and y
6. Repeat the procedure until all spatially connected blocks in U are identified.
7. Save for every block the coordinates of the pixels, the area size, the average, and the maximum precipitation values.
8. Move to next time step $t + 1$
9. Extract all connected blocks in $t + 1$
10. Find which connected blocks in $t + 1$ are connected to the blocks in t

11. For a block $Z(t, x, y)$ find if $Z(t+1, x_1, y_1)$ connects $Z(t, x, y)$:

$$\begin{aligned}&\text{for } K(t, x_i, y_i) \text{ in } Z(t, x, y)\\ &\ \text{for } \ T(t+1, x_{1,j}, y_{1,j}) \text{ in } Z(t+1, x_1, y_1)\\ &\ \text{if } \ x_{1,j}, y_{1,j} == \ x_i, y_i\\ &\text{end iteration}\end{aligned} \tag{7.40}$$

12. The two blocks $Z(t, x, y)$ and $Z(t+1, x_1, y_1)$ form a **connected block in 3D**, namely a connected volume
13. Continue the procedure using the newly formed block and find all connections in other time steps.
14. Extract parameters such as locations, duration, volume, and speed of travel for each connected volume.

The method aims at extracting and investigating spatially and temporally connected precipitation areas from weather radar data. It is classified as an object-oriented approach. The gridded data are available with 5 minutes temporal resolution and 1 km spatial resolution. Along with using the 5-minute data, the fields were also aggregated to hourly and daily resolutions. The radar observations are a snapshot of the observed time. This is not a representation of the continuous rainfall process happening between successive time steps. If one tries to aggregate the images without considering the continuous flow in time and space of rainfall the final result will present discontinuities. The effect becomes more predominant with increasing aggregations. To account for this, the advection correction presented by *Anagnostou and Krajewski* (1999) was applied. To find the advection vector between successive fields, the Lukas-Kanade optical flow method (described in chapter 5 was used (*Lucas et al.*, 1981)). The procedure was adapted from the pysteps Python framework (*Pulkkinen et al.*, 2019a).

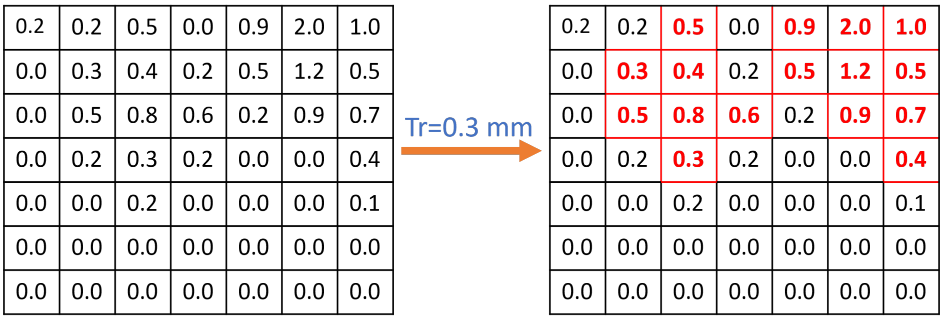
7.2.3 Extracting connected areas

Connected rainfall areas were derived from every radar field. Depending on the temporal aggregation, a different threshold was used. The chosen thresholds can be found in Table 7.1. To extract the pixels that are connected within a certain time step, the radar field was scanned, and all pixels with precipitation values above the given threshold were marked. Then an iteration through each marked pixel and its 8 nearest neighbours was done. If any other neighbour was previously marked, the two pixels were now connected. The procedure was repeated until each object with connected pixels was extracted. For every connected object, the following information was gathered:

- Time of occurrence.
- The area size (km^2).

- The average areal precipitation value (mm).
- The maximum precipitation value (mm).
- The location of the pixels (longitude, latitude).

An illustration of the principle of extracting one connected area from part of a radar field can be seen in Fig. 7.2. In this example, a single 2D-connected area was identified. Information regarding the time of occurrence, area size, areal mean and maximum precipitation, and the grid coordinates were stored. Note that for a larger threshold (for example, 0.6 mm), two connected areas would have been identified.



0.2	0.2	0.5	0.0	0.9	2.0	1.0
0.0	0.3	0.4	0.2	0.5	1.2	0.5
0.0	0.5	0.8	0.6	0.2	0.9	0.7
0.0	0.2	0.3	0.2	0.0	0.0	0.4
0.0	0.0	0.2	0.0	0.0	0.0	0.1
0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0

0.2	0.2	**0.5**	0.0	**0.9**	**2.0**	**1.0**
0.0	**0.3**	**0.4**	0.2	**0.5**	**1.2**	**0.5**
0.0	**0.5**	**0.8**	**0.6**	0.2	**0.9**	**0.7**
0.0	0.2	**0.3**	0.2	0.0	0.0	**0.4**
0.0	0.0	0.2	0.0	0.0	0.0	0.1
0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0

Figure 7.2: Example of connected area extraction. Using the radar QPE values and minimal precipitation threshold (in this case, 0.3 mm/5 min), all cells above and below the threshold are identified. The cells that exceed the threshold (shown in red) are checked for spatial continuity (or connectivity).

Table 7.1: Minimal precipitation threshold per temporal aggregation for identifying connected blocks.

Duration [min]	5	5	60	120	360	1440
Pmin Germany [mm]	0	5	5	10	15	20

To focus on the extremes, a minimal threshold of 5 mm was chosen for the whole study area. While for the Neckar sub-catchments (Jagst, Kocher, and Enz) and the Ahr catchment, a threshold of 0.1 mm was chosen, namely, rain or no rain. This is needed for investigating the connected areas in 3D.

7.2.4 Extracting connected volumes

Many extreme precipitation events are not necessarily driven by high intensities (hourly or daily maxima) but are caused by the accumulation of rainfall cells in space and time. For example, if a connected block remained relatively constant over a certain region, even if with a small intensity, the accumulated volume might cause certain damage. On the other hand, if an event with a strong intensity had a large speed (not concentrated over a certain location) and/or was spatially spread, the subsequent volumes might not be problematic.

Therefore, the connected areas were inspected in a 3-dimensional context, namely space and time. The procedure was simple: once the connected areas in the 5-minute data were identified in space, if intersecting pixels between consecutive 2D blocks (in subsequent time steps) existed, the connected areas would form a 3D connected rainfall block (until no more intersections were present). With this new approach, connected blocks were inspected. as the accumulation of individual areas, defined by start and end times. This differed for each event and offered the possibility to investigate cases with varying durations, intensities, and spatial distributions. The following information describing each block could be extracted:

1. The temporal duration between the first occurrence and spatial separation (min).
2. The accumulated area, the union of all connected areas (km^2).
3. The corresponding volume (each pixel rainfall multiplied by pixel area) (km^3).
4. The average extent, the average distance between the centre point of the first block and the nodes of the convex hull (km).
5. The average speed, calculated between centre points of consecutive blocks (km/h).
6. The spatial density, calculated as the spatial frequency of the locations of the pixels.
7. The form of the blocks defined as the ratio between the convex hull perimeter and the square root of the area.

The accumulated volume over a 3D block could be derived as in equation 7.41.

$$V = \int_T \int_A R \, da \, dt \tag{7.41}$$

Where:

V = accumulated precipitation volume km^3
R = precipitation estimate at each pixel u in area A for time step t
T = Block duration

An example of this can be seen in Fig. 7.3. In panel (a), a connected 2D area is extracted from the radar image for time step T1. In panel (b), two connected areas are identified for T2 (the red and yellow blocks), and if any of them intersects with the area in panel (a), the union is considered a new connected block (the two red areas in panel (b)). The same is seen in panel (c), where new objects intersect with those in panel (b). The procedure is continued until no more blocks intersect. In panel (d), two connected blocks are identified along the trajectory of the centroids (V1, V2, and V3). The one in yellow has a duration of 10 minutes, while the red one is for 15 minutes. Note that for each block, all information from 2D areas was extracted and expanded with time as a third dimension.

The approach was done for the volumes occurring within and around several headwater catchments in the southwest of Germany during the period between 2010 and 2018. The period was chosen accordingly with the availability of the high-resolution discharge data.

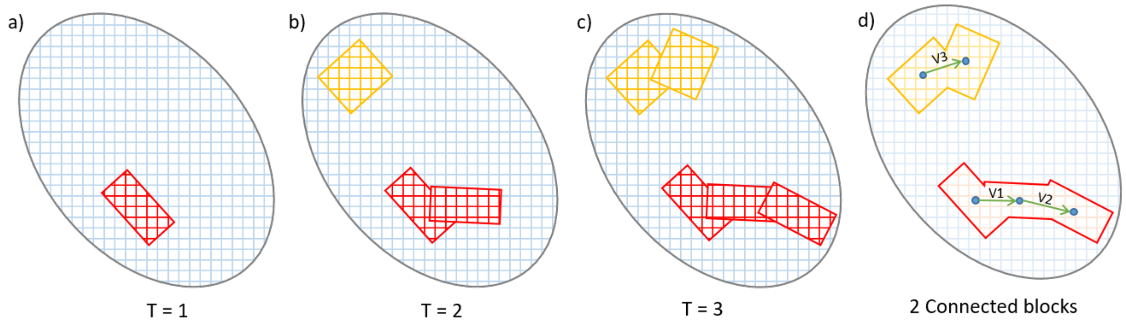


Figure 7.3: Example of 3D connected area extraction within a 2-dimensional field. Panel (a) shows one extracted connected 2D area for time step $t = 1$. Panel (b) shows for time step $t = 2$ newly identified 2 connected areas from which one is connected with the area in panel (a). Panel (c) shows the results for time step $t = 3$ and panel (d) for time step $t = 4$. In panel (d) 2 connected volumes (or 3D blocks) were extracted along their spatial and temporal characteristics.

These were used for event verification. A sort of rainfall-runoff scheme was established by using statistics of the connected 3D blocks and discharge measurements. The main questions to answer are listed below:

1. When do high rainfall volumes correspond to high discharge values?
2. What is the relation between the 3D blocks and the subsequent discharge behaviour, represented by the hydrograph?
3. How fast does the catchment react, defined by the slope of the rising limb?
4. When does the peak occur?
5. How is the spatial and temporal distribution of rainfall influencing the catchment reaction?
6. Which part of the 3D block is failing above the catchment, and what is the ratio between the volumes above the catchment and the total event volume?
7. How much is the discharge volume compared to the 3D rainfall volume?
8. Can and when is the discharge volume exceeding the rainfall volume?

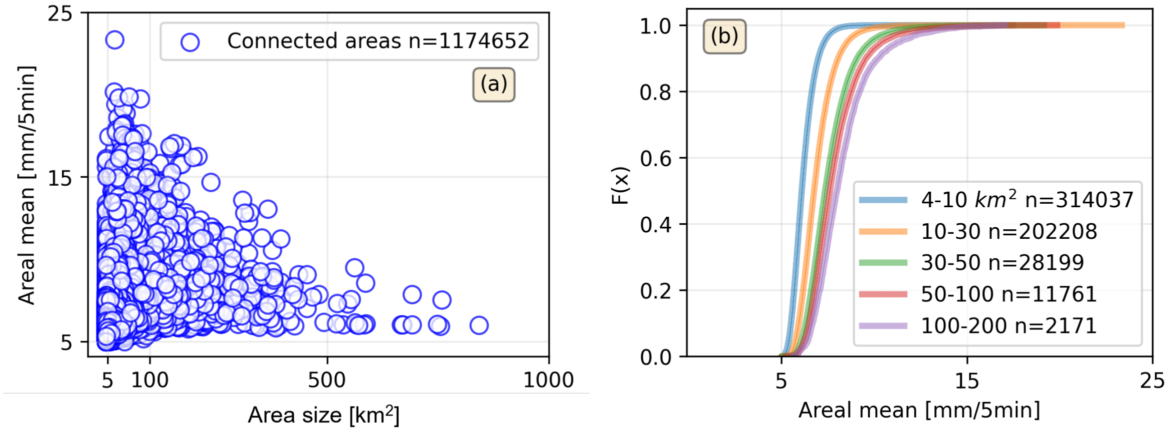
To distinguish between snow and liquid precipitation events, the temperature hourly data for the surrounding gauges were used as additional information. If the average temperature was below the freezing value during the duration of the 3D block, the event was attributed to a snow-driven event.

7.3 Results

7.3.1 Connected areas in 2D

The results of categorizing 2D connected areas based on their size and corresponding areal mean precipitation are shown in Fig. 7.5. The data have been acquired using data from Germany at a 5-minute resolution with a minimum threshold of 5 mm. Panel (a) presents a scatter plot depicting the relationship between area size and areal mean precipitation. Generally, as the area size increases, the areal mean precipitation tends to increase. However, the maximum value was observed for an area with an approximate size of 30 km^2. This might be due to a convective-driven event with high small-scale intensities. However, there is significant variability in the values, with multiple averages possible for the same area size.

In panel (b) of Fig. 7.5, the 2D objects are categorized based on their size into five categories: 4-10 km^2, 10-30 km^2, 30-50 km^2, 50-100 km^2, and 100-200 km^2. The CDF for each category was constructed. As the area size increases, the corresponding areal mean precipitation also increases. The reasoning behind this is the following: since each area contains only pixels exceeding the threshold, larger area sizes tend to have a higher frequency of pixels with larger precipitation values, leading to higher areal precipitation. Namely, the connected area size and the probability of having larger values are proportionally increasing. This indicates that the occurrence of areal extremes in space and time is not a complete random process but rather a structured one.



(a) Area size vs area mean

(b) CDF per area size

Figure 7.4: Panel (a) shows a scatter plot between the area size and the area mean. Panel (b) shows the CDF of the area mean dependent on area size. The threshold used for extracting the connected areas was 5 mm for the duration of 5 minutes. The sample size is shown in both panels.

In panel (a) of Fig. 7.5, the timeline of occurrence of 2D connected areas is presented. The area size in km^2 is shown on the y-axis. The green values represent the summer period (April to September), while the blue ones represent the winter period (October to March).

The result shows that the largest 2D areas on the 5-minute scale occur during the summer period, with the maximum area size being below 800 km^2. In winter, the size of the 2D areas is generally smaller compared to summer, with some exceptions, such as in January 2018. The red dashed lines represent the end of each calendar year. However, the figure will vary depending on the temporal aggregation and minimum threshold used for analysis. For example, on the hourly or daily scale, the largest areas occur most predominantly in winter and not in summer.

To visualize the extracted areas in space, the coordinates of the 2D objects were stacked and binned in a 2-dimensional histogram. The spatial density of the occurrence of a certain area size (and intensity) over a region was calculated and is shown in panel (b) of Fig. 7.5. The south of Germany appears to have a higher spatial density compared to the north. The frequency of the occurrence of events seems to be partly related to topography, but not completely. Some artifacts within the radar data that were not correctly handled are also visible, such as beam blockage in the southeast and southwest radar locations of Feldberg and Hannover, respectively. Moreover, the density tends to increase towards the outward edges, which may be caused by oversampling at the edges and errors during the conversion of the data from the Polar to the Cartesian system.

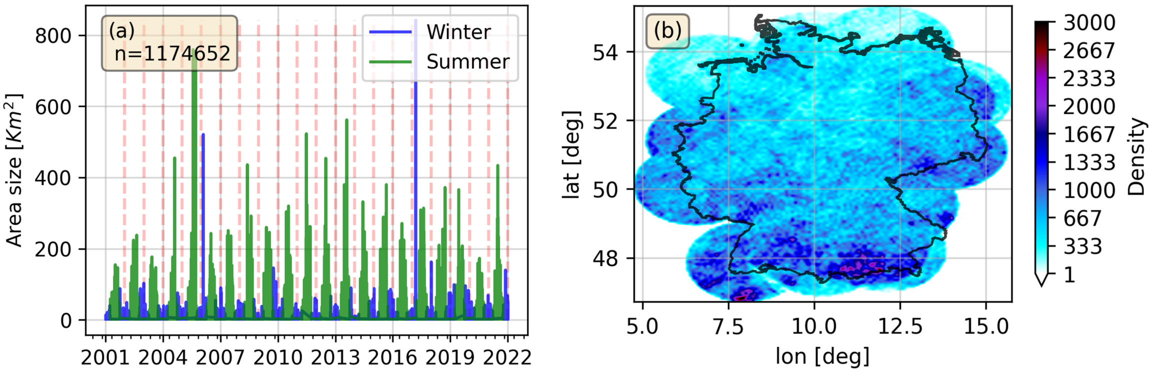


Figure 7.5: Panel (a) shows the area size with time for 5 min data, threshold of 5 mm for the winter and summer seasons. Panel (b) shows their spatial density calculated as a 2-dimensional histogram.

Fig. 7.6 displays the changes in the frequency of occurrence of a 2D block with a given areal extent in the two different seasons. Here, hourly data and a minimum threshold of 5 mm were considered. Panel (a) shows the spatial density for 2D blocks with a size between 30-300 km^2 during the summer period, while panel (b) shows the spatial density for the same area size during the winter season. In winter, the frequency of occurrence is highly dependent on the topography, with higher frequencies for larger elevations. On the other hand, in summer, the results are less dependent on topography and are driven by a north-to-south gradient. Note that the results will change depending on the temporal aggregation, areal extent, and minimal threshold used for extracting the 2D areas.

The extracted areas from the data were further divided based on their size, and the probability of their occurrence at a certain location was calculated. Fig. 7.7 shows a comparison

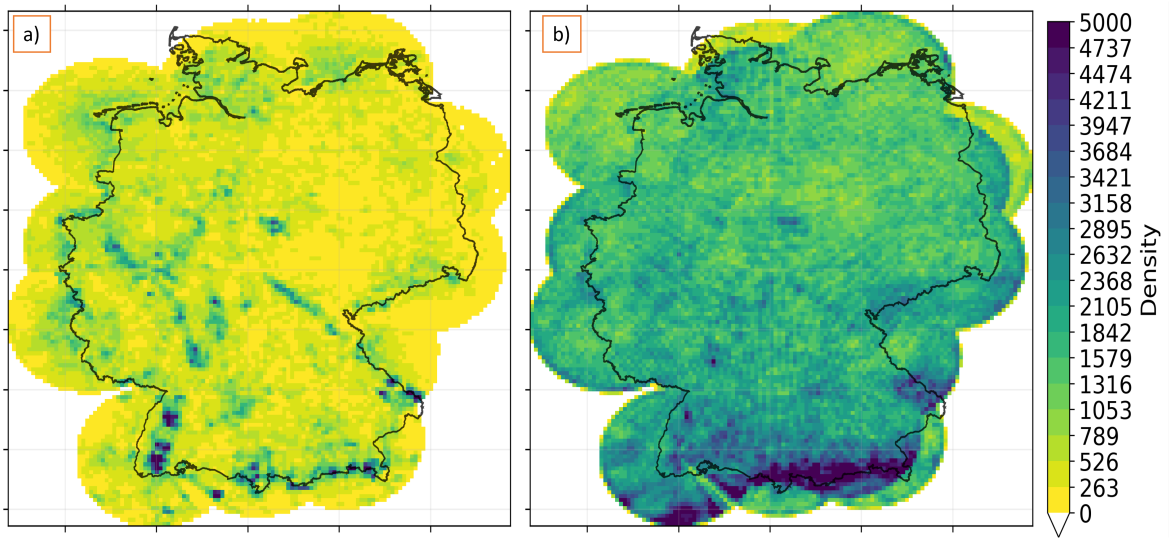


Figure 7.6: Spatial density for hourly 2D areas for size between 30 and 300 km^2 exceeding a threshold of 5 mm for the years 2001-2021 depending on the season. Left panel for the winter months and the right panel for the summer months.

between the spatial density of hourly 2D data. Panel (a) for an area size of 4-50 km^2 and for 30-300 km^2 in panel (b). The ratio of the two maps can be calculated by dividing the spatial density map of larger areas by that of smaller areas. The results show that in most regions, the ratio is greater than one, indicating that larger areas are more frequent than smaller ones. The non-general and space-time-dependent behavior regarding the frequency of occurrence of larger areas compared to smaller ones can be seen in panels (a) and (b) of Fig. 7.7 and panel (b) of Fig. 7.4.

7.3.2 Connected volumes confined to the catchments

The extracted areas from the data can be used to calculate the accumulated volume over a given event with a 5-minute temporal resolution. The total volume of precipitation within a 3D block (an event) was calculated by summing the rainfall values corresponding to each pixel in the 2D block. This provided an idea of the total amount of precipitation that can fall over a given area during the duration of the block. The travel speed of successive 2D areas was calculated as the ratio between the distance between block centers and travel time. The final travel speed of the 3D block was averaged from individual travel speeds. This indicated the average travel speed of the 3D block. Note here the difference between speed and velocity: speed is directionless, while velocity is associated with a direction (a vector quantity). The final block area was derived as the difference between the area of the accumulated 2D blocks and the area of the successive block (as seen in Fig. 7.3). The event volume was compared to the event area, duration, and travel speed, providing insights into the spatial characteristics of precipitation events.

In the first step of the analysis, the identification of 3D blocks was done only within each headwater catchment (Jagst, Kocher, Enz and Ahr). For this, a mask was applied over the

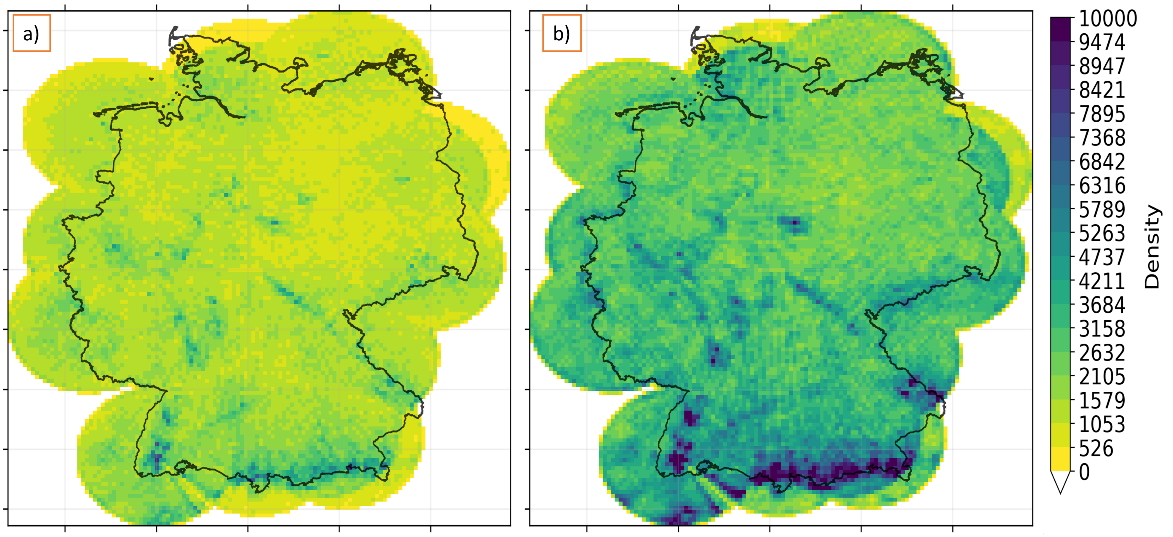


Figure 7.7: Panels (a) and (b) show the spatial density for hourly 2D areas for size between 4-50 km^2 and 30 and 300 km^2 exceeding a threshold of 5 mm, respectively. The spatial density was calculated as an accumulation of all extracted 2D areas for the years 2005-2020.

catchment area, and all 2D connected areas and subsequently 3D blocks for the period 2008–2019 were extracted. An example of a 3D block from the Jagst catchment is shown in Fig. 7.8. Various information were extracted, including the total area of 2034 km^2 (which covers the entire catchment), the duration of 29 hours, the volume of 0,0104 km^3, and the average travel speed of 3,5 km/h. Note that the event started at around 1 p.m. and ended at 6 p.m. the next day. Panel (a) of Fig. 7.8 displays the calculated spatial density as a 2-dimensional histogram of the accumulated locations. The latter can be weighted by precipitation values, allowing for the identification of areas with the highest precipitation intensity. In the case of the Jagst catchment, the spatial density is observed to be highest around the downstream region and diminishes upstream. The upper right panel displays the reaction in the discharge, showing a fast-reacting catchment. The vertical red lines correspond to the start and end of the event, and the vertical bars display the average catchment areal rainfall. The discharge values are displayed by the blue dotted curve and show a quick increase in the discharge values. The purple curve refers to the discharge values after subtracting the base flow and was used to calculate the discharge volume, noted as Qv=0,043 km^3. This accounts for 40 % of the rainfall block volume. The lower panel showcases the discharge time series from 2008 to 2019, with the blue bar corresponding to the event in question. Noticeably, the event corresponds to the maximum observed discharge for the Jagst catchment.

Fig. 7.9 displays all results for the Jagst catchment for the period 2008–2019. The x-axis represents the block volumes, and each sub-figure (panels a–c) shows a different variable on the y-axis. Panel (a) shows the duration. There appears to be a quasi-linear relationship between event duration and volume. Higher volumes are associated with longer event durations. However, events with different durations can result in similar volumes, indicating that intensity (associated with the 2D connected areas) plays an important role in the overall volume of an event. Panel (b) displays the event area on the y-axis, with the largest area being for the catchment itself, which is around 2000 km^2. Interestingly, for the same area size, different volumes are possible, ranging from low to high values. Besides the spatial areal extent, intensity and duration influence the overall volume of precipitation. Panel (c) depicts the volume as a function of the mean travel speed. Clearly, larger volumes are associated with smaller speeds. Overall, the analysis revealed that there are interesting relationships between event volume, duration, area, speed, and spatial density of the events in the Jagst and other catchments.

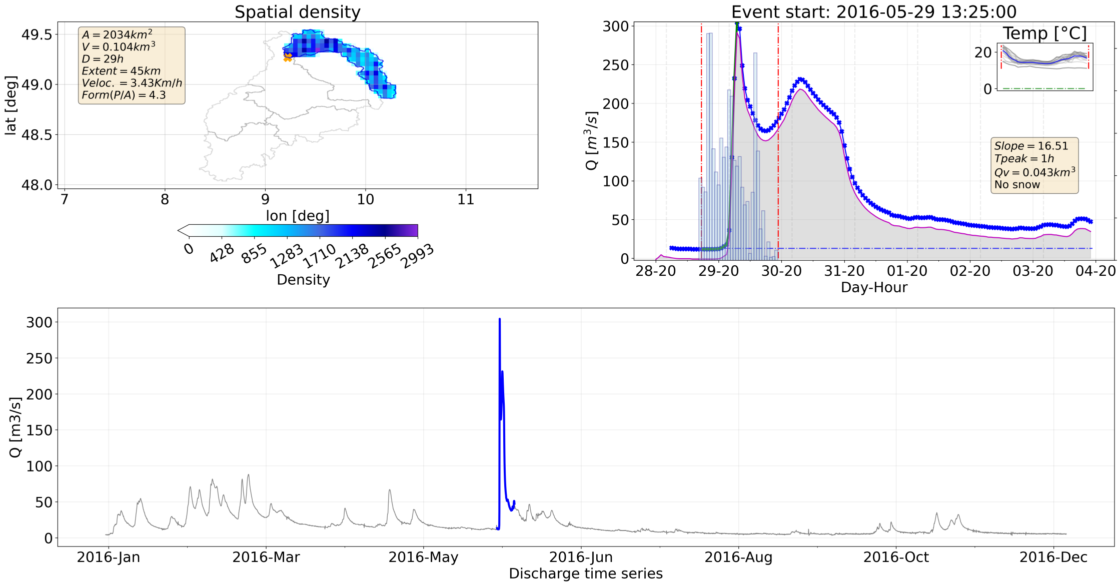


Figure 7.8: Example of a 3D extracted block within the Jagst catchment with the corresponding discharge reaction and timeline values.

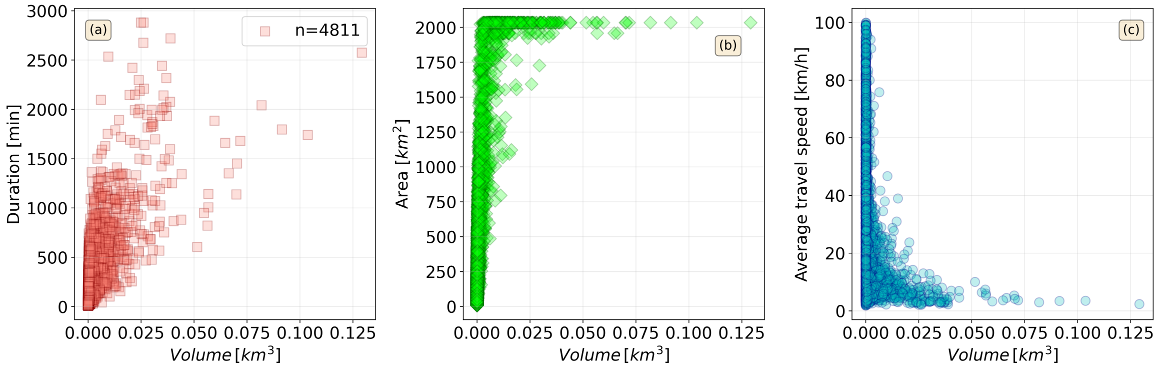


Figure 7.9: Example of the relationships between the precipitation volume delimited to the Jagst catchment with the area (a), the duration (b) and the travel speed (c). The extracted volumes are presented for the whole period.

In the lower panel of 7.10 the relation between the in-catchment precipitation volume and resulting discharge volume in the Jagst catchment is shown. Note that the size of the values in the scatter plot is proportional to the precipitation volume. Moreover, the points shown in black refer to events where the discharge volume exceeded the precipitation volume. Some of these events are associated with below-zero temperatures and might be due to snowfall. In the upper two figures, the location of these events in comparison to the average surrounding temperature is displayed. For those events, the radar data were most likely erroneous, which led to an underestimation of the precipitation volume. In the upper left and right panels, the average catchment temperature corresponding to event rainfall volume and discharge volume are shown, respectively. The temperature ranges between -5 and 25 degrees Celsius. The largest rainfall volumes occur at a temperature of 15 deg C except for some events with an average temperature of 5 $\circ$C. The largest discharge volumes are associated with temperatures between 10 and 15 deg C.

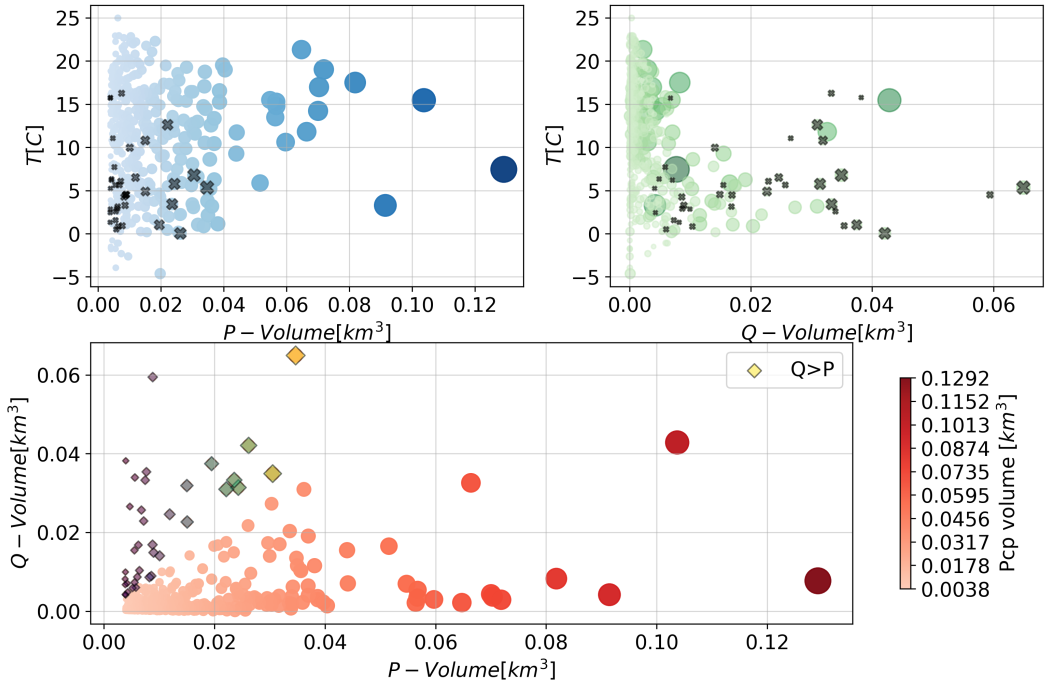


Figure 7.10: Example of the relationships between the top 500 volumes of the 3D blocks in the Jagst catchment against the discharge volume. The upper left and right panels show the relation between rainfall and discharge volume to temperature, respectively.

In Fig. 7.11 the time to the discharge peak is shown as a function of the event duration. The time to the Q peak is calculated in hours. For example, in Fig. 7.8, the discharge peak is reached after one hour. This indicates that this was a very fast-reacting case with a large volume, leading to a high discharge value. Most likely, a flash-flood type of event. To calculate the time to the peak, the start and end times of the rainfall event are identified, the time corresponding to half of the time is considered, and the duration to the maximum discharge value is determined. The value corresponding to this event can be seen in Fig. 7.11, the light green circle with a duration of 29 hours and a time to the peak of 1 hour. In general, the time to the peak discharge is spread around the 24-hour mark. Such information is beneficial for rainfall-runoff modeling. Note that the size of the circles and the colours refer to the in-catchment precipitation volume.

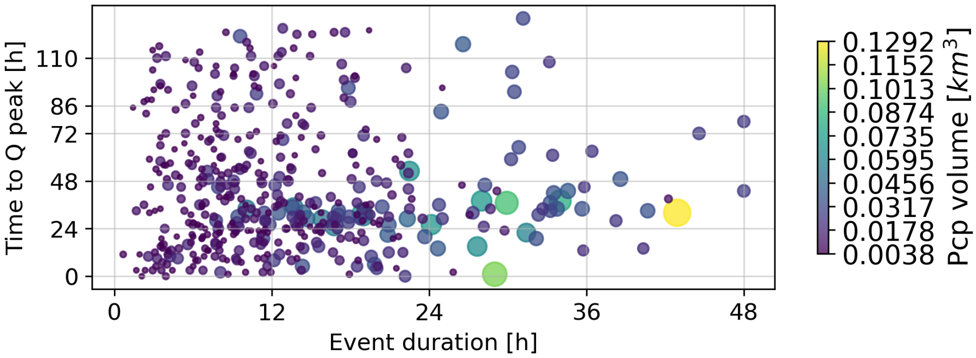


Figure 7.11: Scatter plot showing the relation between event duration and time of discharge peak for the largest 500 precipitation volumes.

On July 14 and 15, 2021, a severe flooding event occurred within the Ahr catchment. The 2D blocks and subsequent event volume were extracted and analysed using RadKlim data. Fig. 7.12 shows the spatial density, weighted by the precipitation depth at each pixel and corresponding time step. The figure shows that the event had a large spatial concentration over the catchment, especially the upper part. In Table. 7.2 the statistics regarding the event duration, area size, in-catchment event volume, and travel speed are presented. The blocks moving over the catchments had a very small traveling speed but a long duration, hence the large volume and subsequent flooding. Although the radar data were seen to be underestimated due to attenuation, these results are insightful for event investigation.

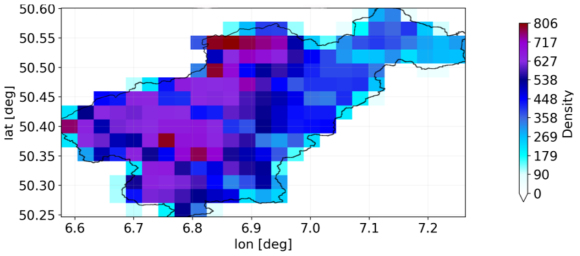


Figure 7.12: Pcp weighted spatial density.

Total duration [$hours$]	16.5
Total area [km^2]	989
Total volume [km^3]	0.0896
Travel speed [km/h]	3.5

Table 7.2: Event volume characteristics.

7.3.3 Connected volumes from total event tracking

In the second part, a total event-based approach was used to extract connected volumes from the space-time-connected 2D blocks. An example of a 3D block with a total duration of 72 hours is shown in Fig. 7.13. The 3D block consists of 3572 separate but connected 2D areas, with a total area of around 2 million km^2. The block properties, including the average speed of 10 km/h and the integrated volume of 4.563 km^3, are noted in the upper left part of Fig. 7.13.

The event spatial density was calculated and shows the accumulation of the 2D blocks in space. The maximum density is observed over the north part of the Neckar catchment, particularly over the headwater catchments of Jagst and Kocher, as depicted in the left panel of Fig. 7.13. The right panel displays the observed discharge series, with the green bar showing the Kocher discharge values. In the lower panel, the reaction in the three headwater catchments (Jagst, Kocher, and Enz) is shown. The red dashed lines indicate the start and end of the rainfall event. The blue bar presents the average areal rainfall over the catchments, with corresponding units shown on the second y-axis. For each discharge gauge, the corresponding base flow was subtracted, and the resulting discharge hydrograph is displayed for each catchment separately. The largest peak is observed for the Kocher catchment in green, followed by the Jagst catchment in orange, and eventually the Enz catchment in blue. To calculate the discharge volume, the integral of each hydrograph over time was calculated, and the information regarding the discharge volume in each catchment is shown in the text box in the upper right part of the lower figure. For example, the volume in the Jagst catchment (denoted as Q_{Jagst}) is 0,048 km^3. Furthermore, the volume of the 3D block that fell in each catchment separately was identified and shown in the right part of the same figure. The ratio of the discharge volume to the rainfall volume was calculated. It is noted that for this event, the average temperature was below zero degrees, which means that the "observed" rainfall volume might be underestimated. This serves as an example of one extracted 3D block.

Table 7.3: Information associated with the connected volume in Fig. 7.13.

Event information		**Catchment precipitation**		**Catchment discharge**	
Area [km^2]	205295	Volume in Enz [km^3]	0.034	Volume in Enz [km^3]	0.025
Volume [km^3]	4.56	Volume in Jagst [km^3]	0.069	Volume in Jagst [km^3]	0.045
Duration [hours]	72	Volume in Kocher [km^3]	0.067	Volume in Kocher [km^3]	0.068
Speed [km/h]	9.5				
Number of 2D blocks	3572				

Fig. 7.14 displays the relationship between connected 3D blocks' properties. Panel (a) shows the relationship between event volume and areal extent. As event volume increases, the areal extent also tends to increase proportionally. However, for the same areal extent, the event volume can vary widely, ranging from 2 to 8 km^3, indicating variability in rainfall intensity within individual blocks. Panel (b) presents the relationship between event volume and total duration. It suggests that event duration may not always be associated with large volumes. Many events with shorter durations can have larger volumes, and for the same duration (e.g., 50 hours), the volume can vary significantly.

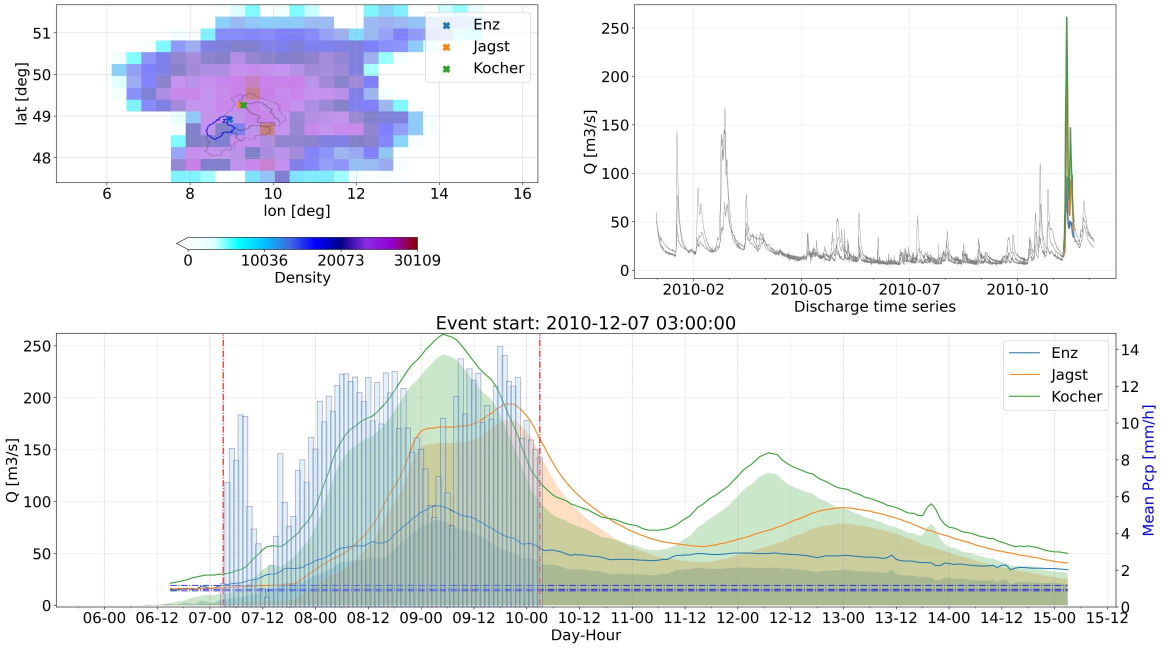


Figure 7.13: Example of a connected volume with a duration of 72 hours causing simultaneous high discharge in the 3 headwater catchments. Upper left panel displays the spatial extent along the spatial density. The upper right panel displays the discharge time series for the year 2010 with the green curve corresponding to the Kocher catchment. In the lower panel the time series of precipitation and discharge volumes for the selected event are shown.

Panel (c) illustrates the relationship between event volume and average travel speed. It shows that larger volumes are generally associated with slower-moving events, with an average travel speed of approximately 10 km/h. However, smaller volumes, including those up to around 0.01 km^3, can occur at different travel speeds, ranging from 10 to 40 km/h. The results provide insights into the areal extent of precipitation extremes and highlight the variability in rainfall volumes within connected 3D blocks, showing that rainfall intensity can vary even within individual blocks. The results also indicate that event duration may not always be a reliable indicator of event volume, and that travel speed play an important role in determining the volume of extreme precipitation events over a certain area or region.

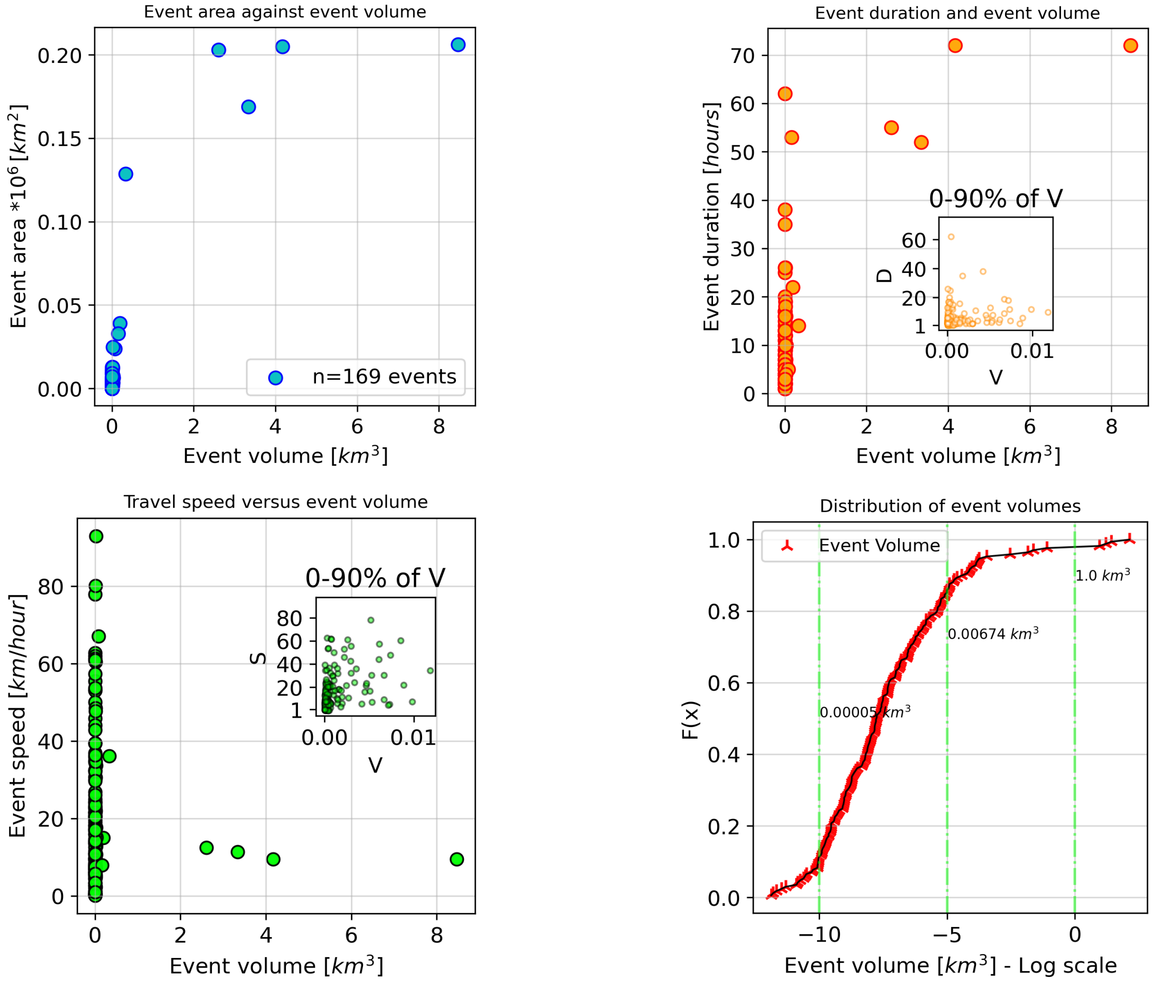


Figure 7.14: The relation of event volume with areal extent, event duration, and average travel speed are presented in panels (a), (b), and (c), respectively. Panel (d) shows the CDF of the volumes of the 3D block.

The distribution of 3D block volumes can be seen in panel (d) of Fig. 7.14, using a log-scale to display the volumes, with dashed vertical lines indicating volumes in the original data space. The CDF shows that most event volumes are spread along the curve, with only a few values exceeding 1 km^3. In addition, the portion of event rainfall volume that falls exclusively in each catchment was determined and related to the total event volume and corresponding travel speed.

Panels (a) and (b) of Fig. 7.15 depict this relationship. Panel (a) shows that although only a portion of the total event volume falls in each catchment, high discharge volumes occur. Panel (b) illustrates that 3D blocks with low travel speeds (around 10 km/h) are associated with the highest discharge volumes. Highlighting the importance of travel speed in determining the magnitude of extreme precipitation events. This suggests that the relationship between event volume, in-catchment volume, and travel speed plays a direct role in determining the discharge volumes of extreme precipitation events.

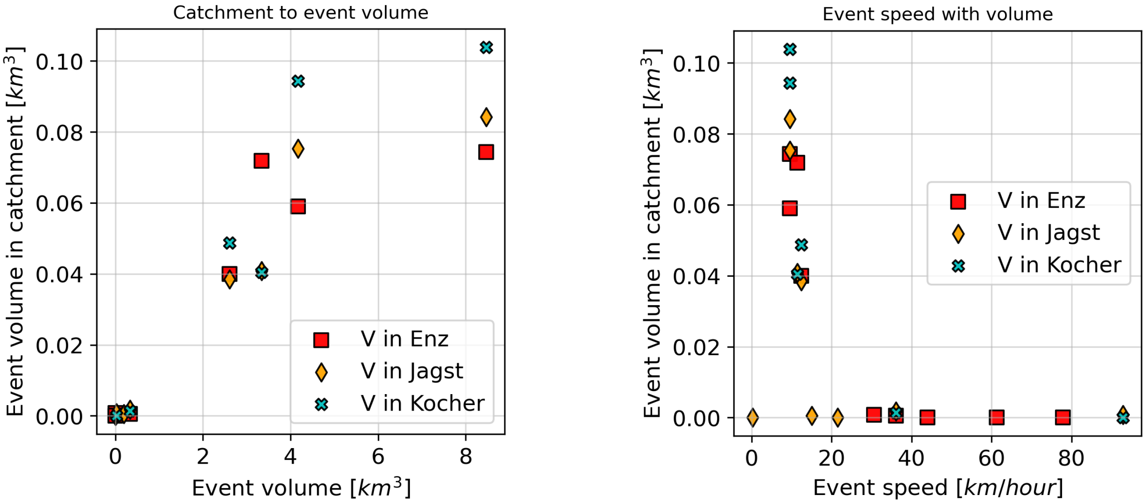


Figure 7.15: Scatter plots of the event to in-catchment volume in panel (a) and event travel speed to in-volume catchment in panel (b). The three headwater catchments Enz, Jagst and Kocher are represented by the red, orange and blue dots, respectively.

7.4 Summary

The previous chapter introduced a new approach to analyzing areal extremes and their spatial extent. Two important aspects were considered: the frequency of occurrence of a certain areal event with a given size and mean areal precipitation over a target region, and the influence of factors such as topography and seasonality on the occurrence and extent of spatial extremes. The results can be valuable when choosing a new observation location or inspecting the areal extremes behavior in a certain region or city. The analysis involved comparing the frequency of occurrence of larger area sizes to smaller ones instead of traditional point-to-areal statistics. The behavior of connected areas with the same size and areal mean differed significantly between summer and winter, as well as across different geographical locations. The comparison between different categories of areal extent, duration, and areal mean revealed that, for a certain areal extent and duration, the probability of a larger area occurring was greater than that of a smaller area. Which can play a role in the use of area reduction factors. The analysis of 3D connected areas provided information on block volumes, the footprint of a block, event travel speed, and event duration. Other aspects were also investigated, such as the change of intensity with advection, the change of the CDF of each 2D block, and the distribution of the final 3D block. Moreover, the ratio of each 2D area to the total block area was calculated. If a large volume occurred over a persistent location, the resulting discharge would likely be high. However, if the same volume occurred over a larger spatial extent with a higher travel speed, the resulting discharge would likely be smaller. This highlights the importance of considering not only the volume of rainfall but also its spatial extent, duration, and travel speed in understanding the resulting discharge and potential impacts of extreme precipitation events.

8 Spatial extent from regional climate model data

The final section of this thesis investigated the changing statistical properties of areal extremes with climate change. Data from regional climate models for the European continent were used with historical and future projections of increasing emission levels. Before being able to use the climate data, the data had to be inspected and any underestimation or overestimation of the dependence structure and/or bias corrected. Once done, the values had to be downscaled from the model spatial scale to a finer one. The aim here was to use a stochastic downscaling method with many (infinitely) equally probable outcomes. Hence, associating the results with the corresponding uncertainty interval.

8.1 Literature review

The output of climate models can be used to investigate the change in specific statistical properties (namely the mean and variance but also extreme value analysis) of different atmospheric quantities over a long period of time. A significant deviation over an extended period of these statistical properties is attributed to climate change. Such deviations can be induced by natural activities (internal climate variability) or human-related activities (such as emissions of greenhouse gases and urbanization). Climate change is different from climate variability and is pronounced by a series of continuous anomalies (the frequent occurrence of low-frequency attributed events). Performing suitable statistical tests is required to distinguish the first from the latter (*Benestad*, 2008). Climate scenarios offer projections about possible changes in the climate system by changing the boundary conditions (for example, greenhouse emissions). Altering the emission scenario and, hence, the model forcing data affects the physical reactions within the climate system. In the case of EURO-CORDEX data, the governing general circulation model (GCM) is driven by a set of emission scenarios known as RCP (Representative Concentration Pathways). Several climate projections defined as an ensemble of projections can be derived and further evaluated by slightly varying the experimental setup (change of initial conditions or model).

In general, GCMs are applied to model the past (forced by observations) and project the future (forced by emissions scenarios) changes in climate systems over global scales. They are based on numerically solving mathematical representations of the physical processes and their interactions (such as conservation of energy, mass, and momentum, the thermodynamic equations such as the Gas law, and relating processes on the land to the atmosphere, and vice versa) (*Hennemuth et al.*, 2017). These calculations are very computationally demanding, especially the discretization of the partial differential equation systems on a continuous temporal scale with a suitable spatial resolution. GCM provides output variables at a relatively coarse space-time resolution, usually between 100 and 500 km^2 and for 6-hour intervals. However, transferring and employing the global results to local and regional climate analysis requires a finer space-time resolution, which downscaling can achieve. The latter can be divided into two categories, empirical-statistical downscaling (ESD) and dynamical downscaling via regional climate models (RCM). ESD exploits the statistical nonlinear relations between small and large-scale information on climate variables. RCM utilizes the GCM output as lateral boundary conditions and, coupled with parameterization schemes to account for local aspects (such as topography), higher-resolution climate data are acquired.

The EURO-CORDEX data are derived from an RCM-based downscaling approach. Several RCM ensembles are available, depending on the driving GCM model and applied parameterization schemes. Differences in the parameterization of the models arise, for example, when representing convective processes, the chosen microphysics and land use schemes (*Kotlarski et al.*, 2014). Despite providing practical information regarding climate systems, GMC and RCM data cannot completely and correctly simulate all relevant spatial and temporal processes. Their outputs should be well interpreted and suitable studies applied (*Ford et al.*, 2016).

Representative concentration pathways (RCP) provide information about possible future climate scenarios. Different RCP outcomes are available depending on the amount of greenhouse gas released. The change in the emission concentration is integrated within the GCM calculations and converted to carbon dioxide (CO_2) equivalents. Two main scenarios present the best (RCP2.5) and worst (RCP8.5) situations. RCP2.5 presents a fast reduction of the produced emissions after the years 2020-2025 while RCP8.5 represents a passive case (zero reduction, a continuous rate of emissions) as compared to historical situations. In between exist two scenarios, RCP4.5 and RCP6.0, which are likely to occur and display less of a catastrophic scenario. An increase in the amount of greenhouse gases implies an increase in the global temperature and hence an alteration of the climatic system (temperature, precipitation, and sea water level) (*Pachauri et al.*, 2014).

A common challenge in creating realizations with climate models is correctly representing the spatial variability of the underlying variable. An evaluation of several RCM hourly data done by *Kotlarski et al.* (2014) showed that, even on the seasonal and regional scales, precipitation values present a bias of $\pm$ 40 $\%$ with a tendency to be overestimated. *Berg et al.* (2019) investigated deriving summer depth-duration-frequency (DDF) statistics from hourly EURO-CORDEX 0.11 $\deg$ data for several European countries (including Germany). Reference national DDF curves were used for comparison. Several RCMs were chosen, and for the long duration, the quality was seen as reasonable, but for the short duration, the models had a bad representation of the hourly extremes. The rainfall depth corresponding to a 10-year return period was severely underestimated by the RCMs output. In the work of *Meredith et al.* (2021), the precipitation diurnal cycle for present and future periods from the EURO-CORDEX 0.11 $\deg$ was examined. Most models have timing errors regarding the occurrence of the maximum hourly precipitation intensities. All models have a peak occurring several hours before the one in the observations. The timing errors were aggravated by increasing precipitation thresholds.

Essential questions to be investigated in this work are:

1. Knowing that extremes look different depending on scales, to what extent can the climate model produce extremes correctly?
2. How can the spatial and temporal dependence structures of climate models be successfully corrected?
3. How will the statistics of extremes change with the climate projections?

To answer these questions, the following scheme was proposed: The first step consisted of upscaling the point to the model scale data. This enabled the derivation of a reference for temporal and spatial dependence structures. A recorrealtion procedure was implemented to correct the model spatial dependence structure to match the reference one. In the second step, the bias in the magnitudes of future projections and any subsequent bias in the marginal distribution function were corrected using a double-QQ transformation. The corrected data on the model scale were then employed within random mixing, a conditional simulation-based method to derive spatial fields and areal statistics for future temporal projections on a finer spatial scale. The downscaling and analysis of extreme value statistics were showcased in the third section of this manuscript. The results were discussed, and a conclusion involving key messages finalized this section.

8.1.1 Reference data

The German Weather Service (DWD) operates a rain gauge network of around 1000 rain gauges across Germany, with varying temporal resolutions (minutely to daily). In the weather radar region of Hannover, 127 DWD stations exist with hourly data (*DWD Climate Data Center (CDC)*, 2021a). The data for these stations were acquired between 2000 and 2020. The second data set consisted of the hourly weather radar data (RADOLAN) for the period 2005-2020 also made available by the DWD (*DWD Climate Data Center (CDC)*, 2021b). The third data set was the EURO-CORDEX data. These have been provided within the Coordinated Regional Downscaling Experiment (CORDEX) for the European continent with two horizontal simulation domains of 50 km (EUR-44) and 12,5 km (EUR-11). The simulation output consists of several data sets with hourly resolution representing different atmospheric and surface-near variables (from which is precipitation) (*Jacob et al.*, 2014). Although several RCM outputs are available, the REMO model is especially advantageous for precipitation analysis on the hourly scale since it accounts for advection on the local scale (*Jacob*, 2001). In this work, the MPI-M-MPI-ESM-LR-GERICS-REMO2015-v1 was used. It was developed by Max-Planck-Institute für Meteorologie (MPI-M) and the Climate Service Center Germany (GERICS). The data were made available by the ClimXtreme Central Evaluation System framework (*Kadow et al.*, 2021). Although other RCM models or an ensemble of RCMs could have been used, the methodology remains the same.

Fig. 8.1 presents the location of the EURO-CORDEX 11 ° center grid points in Germany and in the radar area of Hannover (black circle). To avoid edge effects in the interpolated and simulated fields only model points falling within a 10-kilometer inward buffer (red circle) of the radar coverage boundary were selected. The DWD rain gauges are visualized as orange triangles. The red points in the lower red map represent the center of the radar pixels. The weather radar grid with a spatial resolution of 1 kilometer constitutes the interpolation and simulation grid. The orange box is a 12.5*12.5 km polygon presenting one EURO-CORDEX block. In total, 273 blocks were available within the selected area.

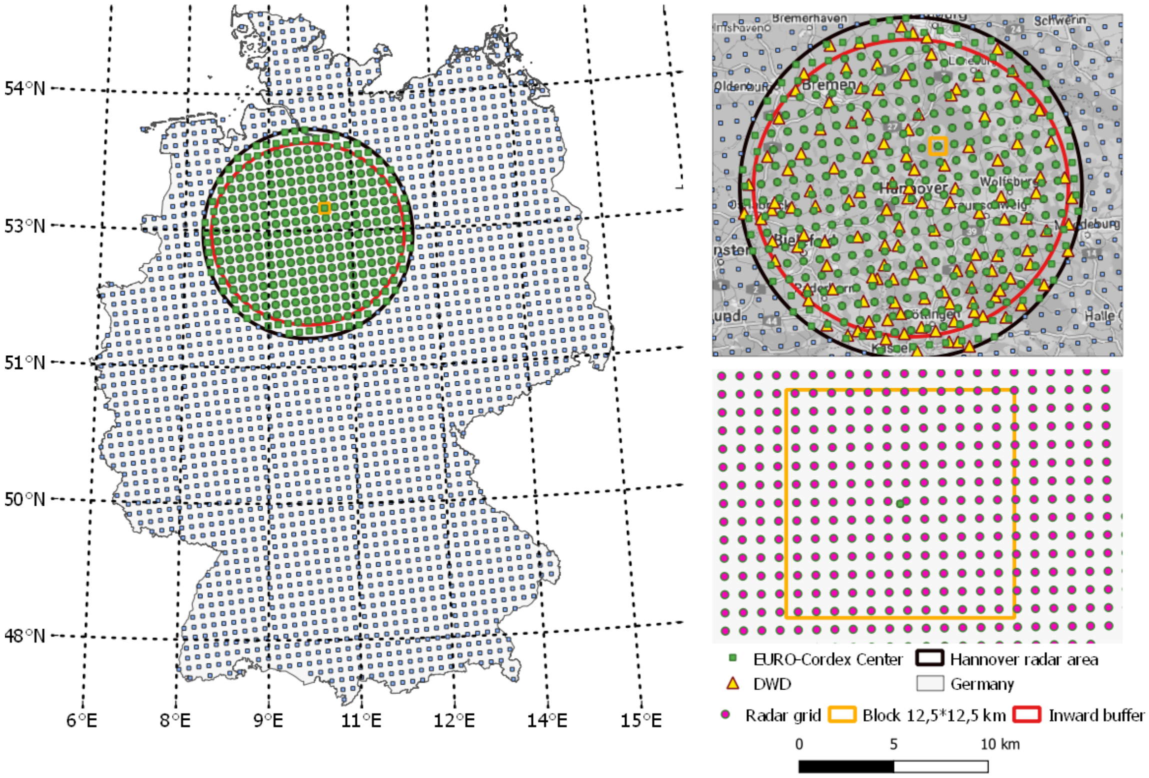


Figure 8.1: Map of Germany with the EURO-CODEX 11 ° grid center locations. The study area is defined by the weather radar area of Hannover along the DWD rain gauge data and the radar grid.

The goal was to find the reliability/usability of the RCM model data for areal precipitation analysis, with a special focus on extremes. Since the model data consist of block averages, it is not suitable to compare them to rain gauge (point) data. Hence, the first reference data is based on interpolated fields using the DWD station hourly data. The interpolated fields on the 1 km radar grid were aggregated spatially to match the EURO-CORDEX 11° grid. This reference data set will be denoted hereafter as DWD_{interp}. The second data set is the spatially averaged radar data for the period 2005-2020. The aggregated data match the model grid and are denoted hereafter as Radar_{avg}. Note that in all cases, the arithmetic mean of the 1 km pixels falling within each model cell was calculated and assigned to the corresponding model pixel.

8.2 Methodology

The procedure for analysing, correcting, and downscaling the RCM data was divided into three parts and is presented in flowchart 8.2.

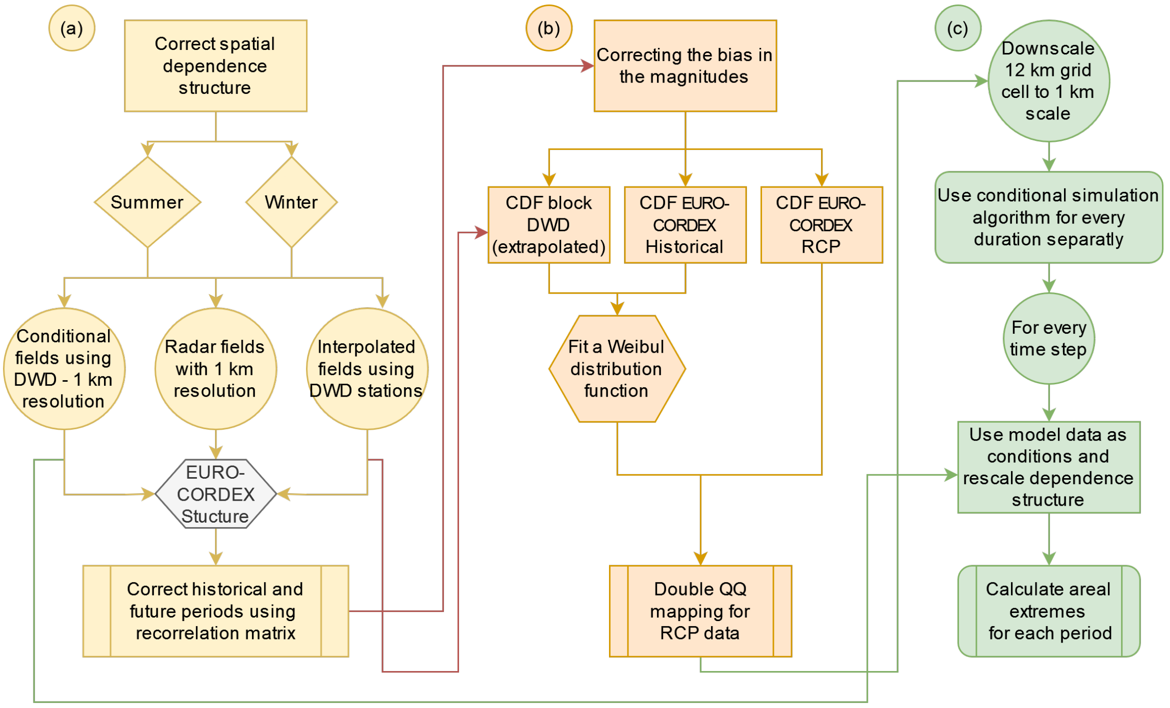


Figure 8.2: Flowchart describing the methodology for correcting spatial and temporal structures of EURO-CORDEX 11 ° data.

First, the spatial dependence structure of the model needed to be corrected according to a reference-based structure. The latter was derived from calculating pixel-wise correlations (cross-correlation or rank correlation) from the reference data (for example, DWD$_{interp}$). The correction was based on a recorrelation procedure that is described in the coming section. The dependence structure greatly affects the areal extremes (*Bárdossy and Pegram*, 2012). This part refers to the yellow section of the flowchart. Once the dependence structure was rectified, any remaining bias in the marginal distribution function of future climate projections was to be handled. For this, a double-QQ transformation involving information from the reference data and the model historical data was applied (*Bárdossy and Pegram*, 2011). This section represents the red part of the flowchart. Afterwards, the final corrected data were downscaled using a stochastic simulation algorithm (random mixing, (*Hörning et al.*, 2015)) to a finer spatial resolution of 1 km. The final fields were eventually used for the analysis of the spatial extent of extremes via calculating the extreme value (EV) statistics such as DDF and ADDF curves. The DDF curves were calculated pixel-wise (1 km scale) and area-wise (up to 1024 km^2). The possible impact of climate change on the statistics of areal extremes was to be investigated.

8.2.1 Correction of dependence structure

The spatial dependence of precipitation plays a major role in the distribution of areal rainfall and the corresponding extremes. An often neglected problem is the capability of the climate model data to replicate the observed dependence structure. Previous researchers have already discussed and presented solutions to the problem. An example of this can be found in (*Bárdossy and Pegram*, 2012). A similar procedure was here implemented, however, for hourly data and 12.5 km spatial resolution. The reference and model data were divided between the summer (April-September) and winter (October-March) seasons, and for each period, the dependence structure was derived. To cope with the zero values, a mixed-type distribution was used, defined by a censored Gaussian copula. All values greater than zero were transformed into the Gaussian space using a standard normal distribution. Values equal to zero were replaced by $P(0)/2$. The transformation was applied for each block separately and is defined by equation 8.44.

$$F_i(z) = P(Z_i(t) < z | Z_i(t) > 0) \tag{8.42}$$

$$P_{i0} = P(Z_i(t) = 0) \tag{8.43}$$

$$W_j(t) = \begin{cases} \phi_1^{-1}(F_j(Z(t)(1 - P_{i0}) + P_{i0})) & \text{if } Z_i(t) > 0 \\ \phi_1^{-1}(\frac{P_{i0}}{2}) & \text{if } Z_i(t) = 0 \end{cases} \tag{8.44}$$

Where:

$F_i(z)$ = distribution function at block (i) for positive precipitation
P_{i0} = p0 for block (i)
ϕ_1 = standard normal distribution function N(0,1)

The reasoning behind using the rank correlation was that the ranks were later involved in the double-QQ transformation. In Fig.8.3 and Fig.8.4 the pair-wise rank correlation values for each data set individually plotted against the separating distance between each and all other blocks. For the EURO-CORDEX data, the historical and future periods presented similar behavior and only the historical data are displayed. Note that each correlation matrix has a shape of (273, 273).

Fig.8.3 shows the correlation structure derived from the DWD rain gauges located within the study area and DWD$_{interp}$. The correlation of the point data shows less spatial continuity and would have been inappropriate for correcting and comparing to the correlation of the blocks.

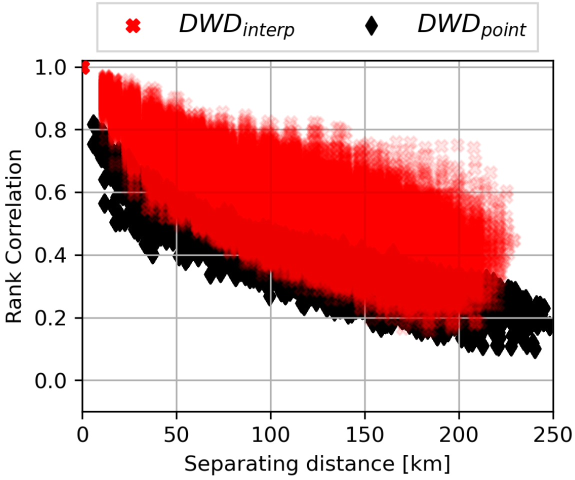


Figure 8.3: The rank correlation values between the time series of the DWD rain gauges against the separating distance are shown in black dots. The red dots display the rank correlation between the time series of the DWD averaged block values. Both are shown for the winter period.

The DWD$_{point}$ correlation values are lower than those of the DWD$_{interp}$ data. This difference can be justified theoretically - assuming stationarity of the spatial dependence - as the interblock variability leads to the reduction of the variance. In fact, the covariance between two grid cells V_i and V_j can be written as a function of the covariance function of the point values $C(x, y)$:

$$\text{Cov}(V_i, V_j) = \frac{1}{|V_i|}\frac{1}{|V_j|}\int_{V_i}\int_{V_j} C(x, y)\ dx\ dy \tag{8.45}$$

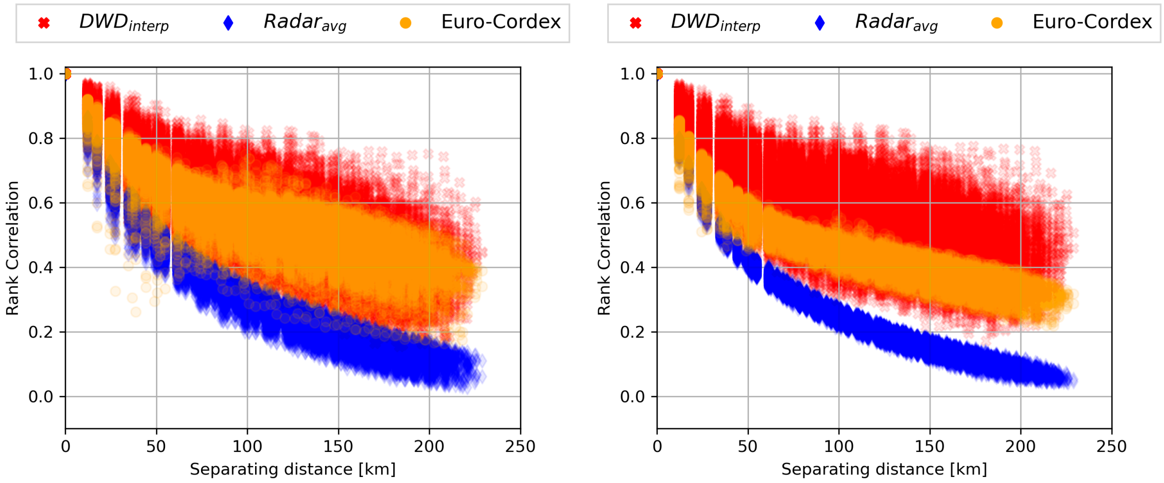
While for the variance of the grid cell V:

$$\text{Var}(V) = \text{Cov}(V, V) = \frac{1}{|V||V|}\int_{V}\int_{V} C(x, y)\ dx\ dy \tag{8.46}$$

Both the covariance and the variance decrease. The decrease of the covariance is less than or equal to that of the variance, thus the correlation increases. Thus due to scale difference, using the DWD$_{point}$ correlation structure as a reference for the correction of the EURO-CORDEX spatial structure would be incorrect.

In panels (a) and (b) of Fig.8.4, the red dots represent the DWD$_{interp}$, the blue dots the Radar$_{avg}$ and the orange dots the Euro-Cordex data. Panel (a) for the winter period and panel (b) for the summer period. The results for DWD$_{interp}$ show a typical behavior of decreasing correlation with increasing separating distance associated with a large scatter. In both cases, the correlation structures of Radar$_{avg}$ present a smaller scatter and fall below those of DWD$_{interp}$ and Euro-Cordex. In other words, Radar$_{avg}$ shows less spatial continuity (a quick drop of correlation) and larger variability between the blocks. Compared to

DWD$_{interp}$ the EURO-CORDEX results show an underestimation of the dependence structure, especially in the summer period. The aim of the method was to adjust the model dependence structure to match the reference one over all temporal aggregations.



(a) Period: October-March

(b) Period: April-September

Figure 8.4: Panels (a) and (b) show the calculated block pair-wise rank correlation values from interpolated fields (red points), radar fields (blue points), and EURO-CORDEX fields (orange dots). The x-axis refers to the separating distance between the blocks.

The recorrelation scheme was based on linear methods applied to the Gaussian-transformed data. Following the same notation as in *Bárdossy and Pegram* (2012), the reference transformed data were denoted by **W** and their correlation matrix by **C**. The transformed EURO-CORDEX data were denoted by **Y**. The aim was to recorrelate **Y** to new data **V** with the same correlation as **W**. The values in **V** were then back-transformed to the original data space at each location individually. The main assumption behind this technique was that the pair-wise correlation matrix was independent of the areal mean values. An assumption that was found to be valid by the researchers. The matrix **C** was transformed into the matrix **S** by calculating the square root of applying single value decomposition (SVD) to **C**. The correlation matrix **R** of **Y** was transformed to **T** by calculating the inverse square root of applying SVD to **R**. The final matrix for decorrelating and recorrelating **Y** to **V** was defined by **F=TS** and **V=YF**.

The recorrelation matrix **F** derived from the historical observations could be used for the projected RCP scenarios. The method was in-depth described in *Bárdossy and Pegram* (2012) with main modifications in this work regarding the temporal and spatial resolution of the RCM data.

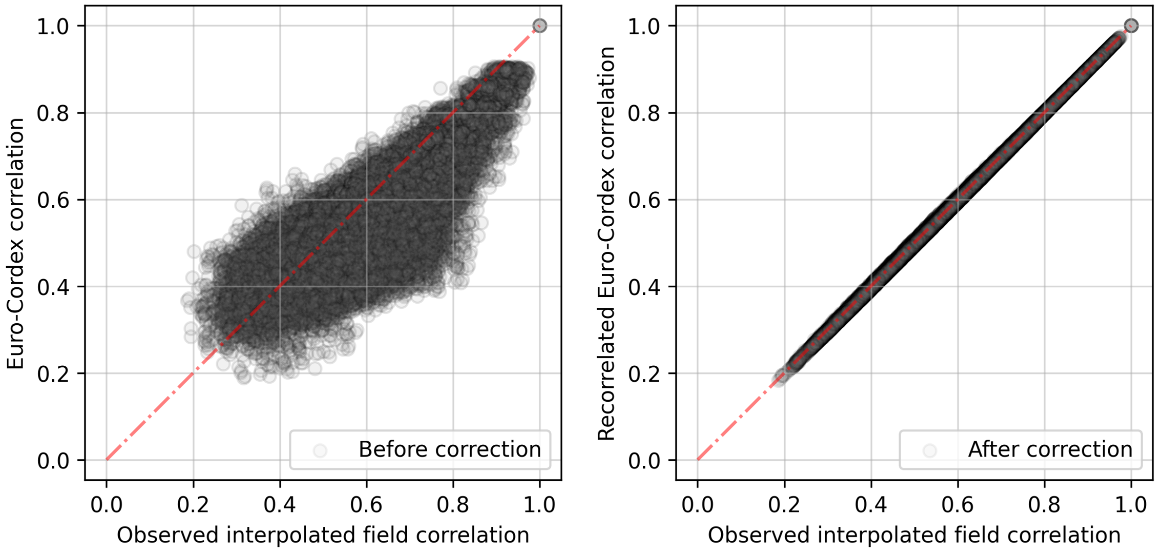
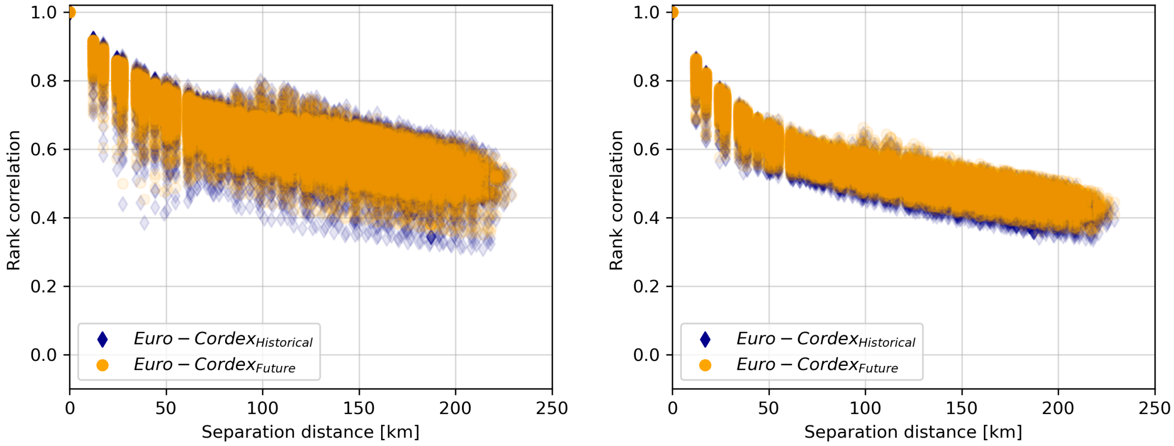


Figure 8.5: Recorrelation of the EURO-CORDEX data using the interpolated fields derived dependence structure for the summer season. Left panel before recorrelation and right panel after recorrelation.

Fig.8.5 shows the recorrelation procedure using the DWD$_{interp}$ as reference data. Panel (a) displays the cross-correlation of DWD$_{interp}$ and EURO-CORDEX data. In panel (b), the final results of the recorrelation are displayed and show a complete success in relocating the RCM data for the historical period. The same recorrelation matrix was applied to the projected RCP data. Eventually, using the inverse of equation 8.44 the recorrelated data in matrix **V** were back-transformed to the original data space. The recorrelation matrix **F** could be applied directly to correct the RCP data because the correlation structure was assumed to be constant. That means that the spatial dependence structure from the historical period and the reference data were considered stable despite climate change.

Fig. 8.6 shows the pair-wise block rank correlation values for the EURO-CORDEX historical and future data. Panel (a) shows the results for the winter period, and panel (b) for the summer period. For the latter, the correlation values seem to be highly similar, with few differences. However, for the winter period, the historical data show a quicker drop in the correlation values than the future data. For instance, for a separation distance of 150 kilometers, the future data indicate a minimal correlation of 0.42, while the historical data indicate a correlation of 0.38. In addition, both datasets appear to have an unusual drop in correlation for short separation distances (around 50 km) and an unusual increase in correlation for longer separation distances (between 100 and 120 km). This behavior might be caused by erroneous data at one single block location.



(a) Correlation period October-March.

(b) Correlation period April-September.

Figure 8.6: Rank correlation dependence structure for historical and future EURO-CORDEX data for winter (a) and summer (b) periods.

To demonstrate the effect of the recorrelation on the model data, the results for one location are shown in Fig.8.7. Panel (a) shows how recorrelated values scatter around the x=y axis (shown in red), indicating no clear under- or over-correction of the data. Note that the size of the dots is proportional to the precipitation values before correction. The upper tail of the distribution function before and after recorrelation can be seen in panel (c) of Fig.8.7. The effect of the recorrelation on the values is minimal; however, the spatial dependence structure has been adjusted to fit the reference one. Other measures, such as the indicator correlation (p99) and Cross correlation measures (described in chapter 2) were also tested. The recorelation procedure was successful in improving the dependence structure, but a complete match was not possible. Moreover, if the correction was done based on any of the correlation measures (for example, Pearson correlation), other dependence measures (such as Indicator and Spearman correlation) showed an improvement. Hence, although the correction was applied to one measure, the procedure improved other properties simultaneously.

The recorrelated data on the hourly scale were aggregated into higher temporal aggregations (such as 2, 6, 18, and 24 hours). The correlation structure was calculated before and after the recorrelation and was compared to the reference data. Although the correction was done on an hourly scale, the procedure was found to be successful in improving the spatial dependence structure over other temporal aggregations. An example of the daily aggregation can be seen in Fig. 8.8. The correction was the most successful on the hourly scale and the least successful on the daily scale. Though the corrected daily data (panel (b)) show much better agreement than the non-corrected data (panel (a)), which largely underestimates the dependence structure. Although a slight bias was still present, it was not seen as necessary to apply the recorrelation procedure for each temporal aggregation separately.

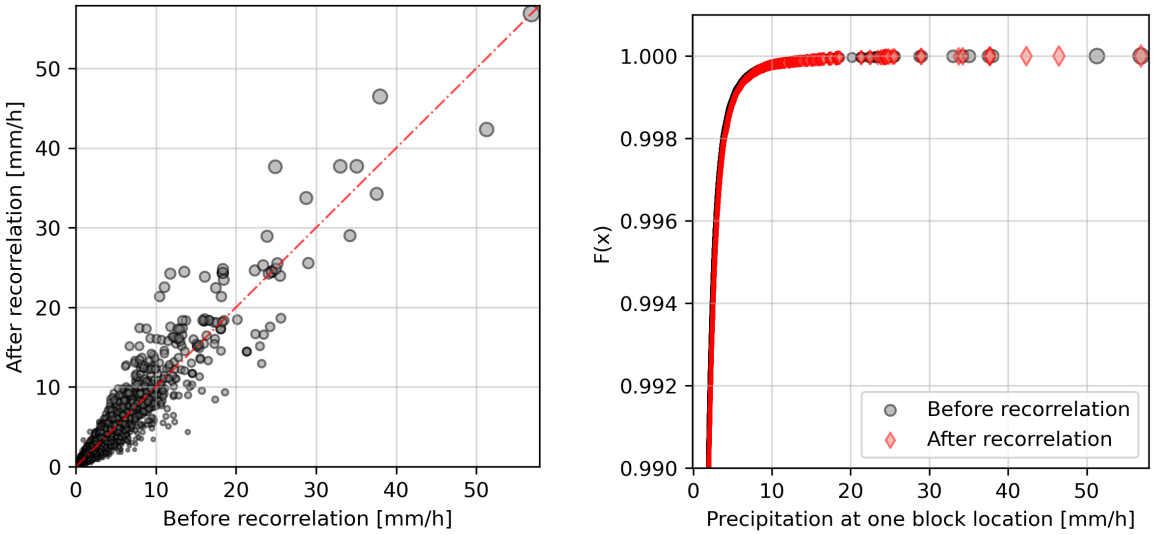
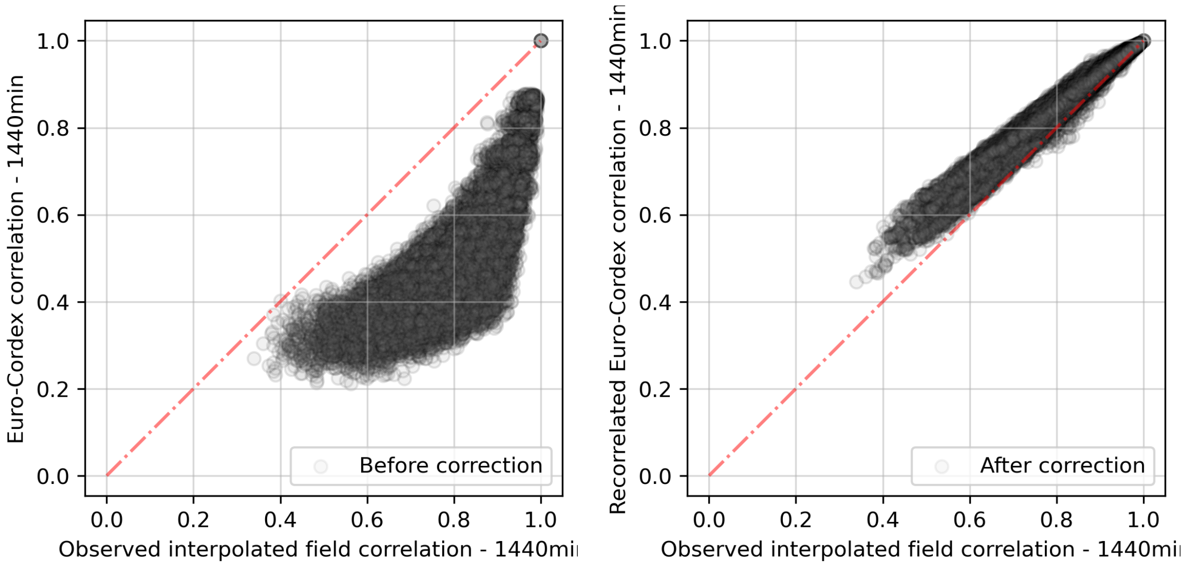


Figure 8.7: Panel (a) shows a scatter plot of hourly rainfall values at a single block from RCP8.5. Panel (b) shows the upper 1% tail of the CDF before and after the recorrelation procedure.



(a) Before correction

(b) After correction

Figure 8.8: Rank correlation dependence structure for daily historical EURO-CORDEX data before (a) and after (b) the recorrelation procedure. Both panels showcase the data for the period April-September.

Note that this complete part refers to the yellow block in the flowchart 8.2. After correcting the dependence structure, the next step was to proceed with correcting the marginal distribution function of the model data.

8.2.2 Double quantile-quantile mapping

In statistics, quantile-quantile plots (QQ plots) are used to compare two distributions and find if they both belong to the same distribution function. Often, a test distribution is compared to a theoretical one. The comparison is based on the quantiles. Namely, a scatter plot between the quantiles of both data values is constructed. If the data have the same distribution function, the QQ plot will be defined by a linear function (y=x). To derive the QQ plot, the data of the two distribution functions are sorted and their quantiles are calculated. In this section, a QQ transformation was used to correct the distribution of the projected EURO-CORDEX data.

For this, the CDF of the observed data at location X (for example, DWD$_{interp}$) and the distribution function of the model results for the present and future periods for location X were used. A Q-Q transformation was applied to correct the bias in the projected model data while preserving the ranks of the values. An example of this was shown by *Bárdossy and Pegram* (2011) where the distribution function of regional climate models (RCM) was corrected using a double QQ transformation as defined by equation 8.47.

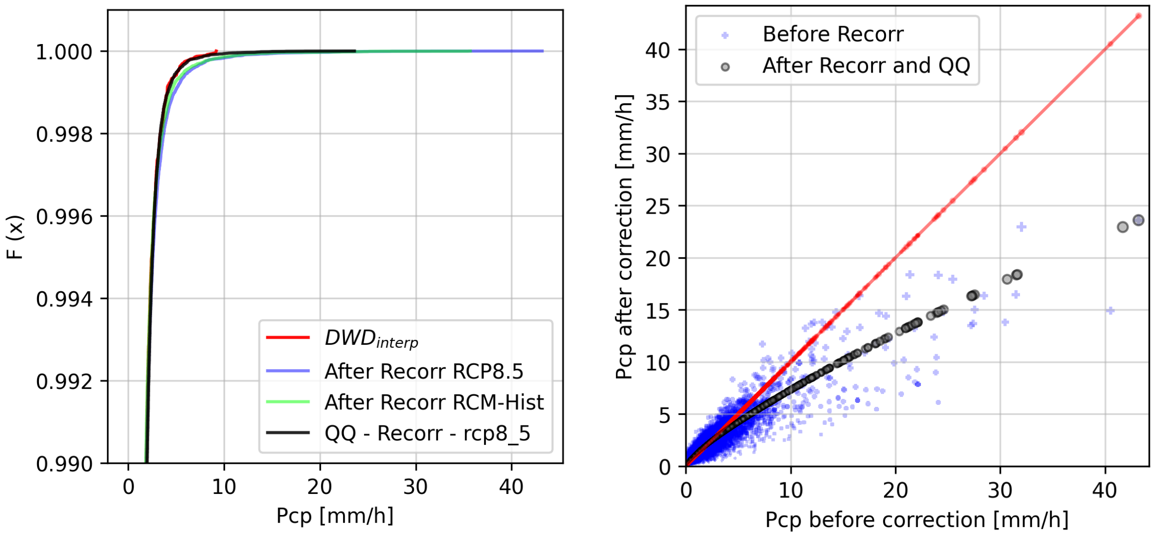
$$Z(x,t) = F_o^{-1}(F_R(Z_R(x,t),x),x) \tag{8.47}$$

Where:

x = target location
t = time step (hour)
$Z(x,t)$ = corrected precipitation value
F_o^{-1} = inverse of the fitted CDF to the reference data
F_R= CDF of the RCM data
$Z_R(x,t)$ = precipitation simulated by the RCM

A similar approach to *Bárdossy and Pegram* (2011) was used to correct the distribution function and temporal structure of climate model future projections. Using the CDF of the spatially aggregated observations, the CDF of the recorrelated historical data, and the CDF of the future time periods were corrected. For the observation and historical data, a Weibull distribution function with suitable parameters was fitted using the maximum likelihood method (*Singh*, 1987). Note that in this case, the maximum of the fitted function was not limited to the observed maxima. This allows for correcting the RCP future data while allowing for extreme values slightly exceeding the current observed one. These might be underestimated due to observation errors or under-sampling of the spatial block maxima. In panel (a) of Fig.8.9 the x-axis refers to precipitation in millimeters and the y-axis to the cumulative probabilities. The red curve is the distribution function of the reference data. The green curve is the RCM distribution for the present period, and the curve in blue is the RCM curve for a future scenario. For a precipitation value in the future curve, the corresponding present value was found, and for the same quantile level, the observed value from the observation curve was found and re-assigned to the future value. This QQ transformation reduced the bias in the model data while preserving the signal in the RCP data.

An example of this is presented in panels (a) and (b) of Fig.8.9 for one location. The light blue curve refers to the RCP CDF before the recorrelation procedure, the Lilac curve for the recorrelated RCP data, and the red curve for the reference data for the observation period. The black curve shows the QQ corrected RCP distribution function using the previously described procedure. For every precipitation value in the future data, the corresponding value in the historical data was found, and for that, the corresponding value for the same quantile level in the observation data was assigned as the future value. To incorporate the uncertainty in the CDF of the blocks, simulated fields were used. Hence, an ensemble of corrected data can be acquired. The correction using any of the reference data seemed to not strongly influence the final results. The simulation of DWD$_{rdm}$ indicates that the maxima of a 12.5*12.5 km block could exceed those aggregated from DWD$_{interp}$ and Radar$_{avg}$; hence, the fitted Weibull distribution was not bound to the observed DWD$_{interp}$ or Radar$_{avg}$ maxima for the corresponding location. The procedure was repeated for all values and for both data sets, namely RCP2.5 and RCP8.5. Note that this section refers to the orange block in the flowchart 8.2.



(a) Upper 1% tail of the CDF before after QQ

(b) Values before and after correction.

Figure 8.9: CDF and scatter values for the recorrelated historical and RCP data before and after the QQ correction for one example location.

8.2.3 Random mixing theory

Most interpolation techniques can only deliver smoothed fields. For example, in all Kriging applications, the equation system is solved by minimizing the estimation variance. Hence, the interpolated fields have less variability than the original ones. This can be demonstrated by calculating the variogram from the original and interpolated values. For a similar separating distance, the variance of observed data is greater than that of the interpolated values. In hydrology, many processes (for example, rainfall and catchment reactions) are driven by the variability of the variable in question. To incorporate variability in the estimation method, a simulation could be undertaken. Two major types of simulations exist, conditional and unconditional simulations. Random mixing is a conditional simulation method that allows stochastic generation of realizations fulfilling several conditions (*Hörning et al.*, 2015). The method is extended from the work of *Hu* (2000) regarding the gradual deformation of Gaussian fields.

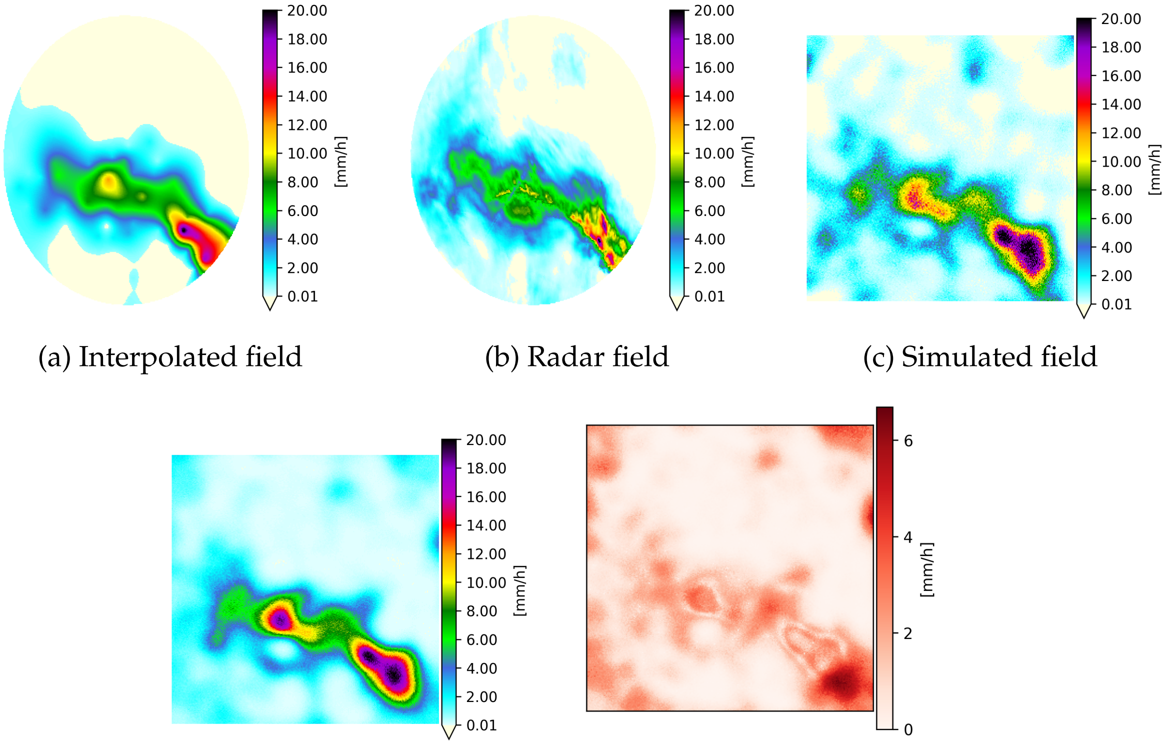
For achieving the goal of having realistic realizations, the following criteria must be fulfilled:

1. Match the measured data (conditioned on observations).
2. Have the same spatial dependence structure (represented by a variogram or spatial copula function).
3. The realizations should have a similar range as the observed values (no extreme values).

Traditional kriging methods are constrained by the mean and variance of the observations. The random mixing methods allow capturing the spatial dependence structure by a spatial copula function. A copula is a mathematical function used to derive and model the dependence between variables independent of their distribution functions. If a spatial copula function is used, the asymmetry of the spatial field can be better accounted for (asymmetry can be seen as the skewness measure of univariate data). Another advantage of using random mixing is the possibility to incorporate several conditional observations, which can be integrated as a linear or a non-linear equation. Simulation methods offer the possibility of generating several possible realistic realizations of a certain event. Through this, the mean uncertainty field can be derived. This method offers a probabilistic simulation technique, which is advantageous compared to the traditional interpolation method.

In this section, random mixing was used to upscale and downscale the station observations to the climate model scale and vice versa. Based on the DWD station hourly data for the period 2005-2020 several conditional fields were simulated for every hour with positive precipitation. The simulated fields were conditioned on the DWD point observations, the corresponding marginal distribution, and the fitted Gaussian copula model representing the spatial dependence structure. The fields were then spatially aggregated to the model scale. Hence the upscaling part.

Note that in this step, the fitted spatial model was saved and used later for the downscaling part. Once the model data were corrected using the recorrelation and double-QQ procedures, simulations conditioned on the model areal averages were generated on the finer spatial grid of 1 km. Hence, the downscaling part. An example of a randomly selected realization can be seen in Fig. 8.10. The realization in panel (c) is compared to the results of ordinary kriging (panel (a)) and the observed weather radar image (panel(b)). The advantage of this method as compared to normal interpolation techniques is that it allows a probabilistic upscaling or downscaling of the point data to the model scale. The associated uncertainty field from the 20 simulations can be seen in panel (e). Note that the average simulated field in panel (d) is highly similar to the interpolated field in panel (a).



(a) Interpolated field (b) Radar field (c) Simulated field

(d) Average of 20 simulations (e) Uncertainty field

Figure 8.10: Example of a randomly selected hourly rainfall field using the DWD rain gauges for interpolation (panel (a)), simulation (panel (c)), and the radar observed field (panel (b)). Panel (d) shows the average simulated field along with the associated uncertainty field in Panel (e).

A limitation of using random mixing is that it generates spatial fields that are not temporally correlated. In other words, the simulated fields for successive time steps were not temporally correlated, and effects such as advection were not considered. For example, the correlation structure shown in Fig. **??** could not be correctly calculated from the upscaled fields. The correlation values were calculated between pair-wise time series; hence, the autocorrelation of the data played a major role. The latter was destroyed by random mixing, as the focus is on the spatial aspect. This would require a modification to the algorithm to include advection and conserve the autocorrelation within a time series.

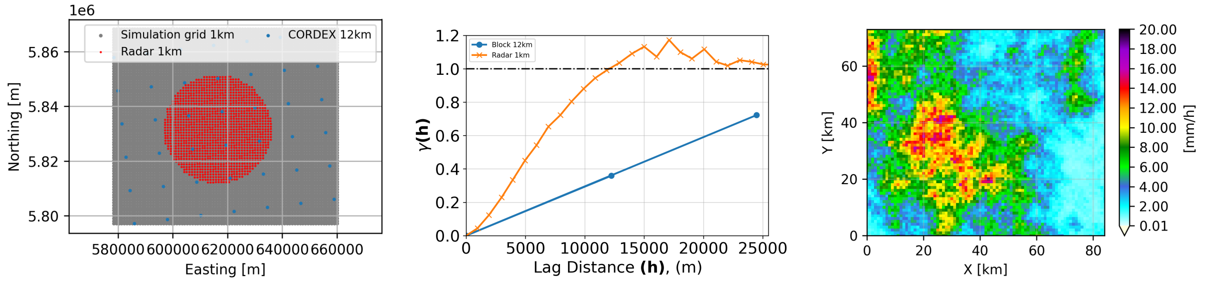
8.2.4 Downscaling model to point scale

After correcting the spatial dependence structure and the bias in the CDF of the future RCP data, conditional realizations at a finer spatial scale of 1 km were generated using random mixing. Note that this section corresponds to the green block in the flowchart shown in Fig. 8.2. This process incorporated the climate signal from the projections and provided a dataset that could be used to derive future DDF curves for the study area. To account for the uncertainty acquired by the simulation approach, 50 realizations for every time step were performed. Namely, 50 downscaled and equally probable time series for every pixel in the simulation domain were generated. While performing the downscaling, the spatial model calculated on the larger scale needed to be adjusted for the lower scale. To adjust the spatial correlation models from the climate model data to the 1 km scale, the variance and range of the block variogram needed to be rescaled to the point variogram. An example of this is shown in panel (b) of Fig.8.11. The block variogram derived from the average radar data ($Radar_{12km}$) shows a smaller variance and is smoother compared to the empirical pixel-based variogram ($Radar_{1km}$).

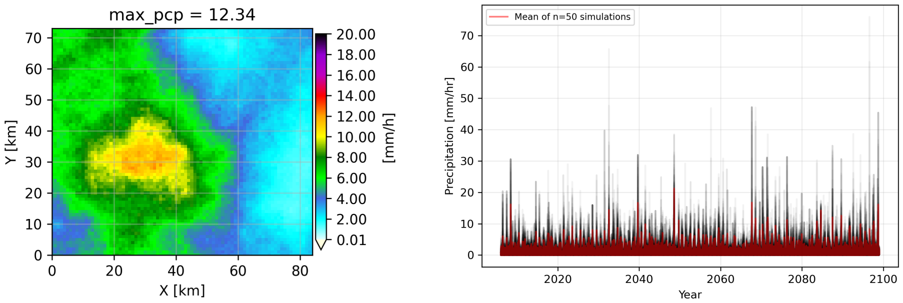
A re-scaling of the variogram parameters was performed. From the upscaled fields, the variograms were saved and clustered using a K-mean clustering approach, similar to the procedure in chapter 4. Depending on the day of the year, the fitted variogram from the upscaled fields was used to rescale the calculated variogram from the EURO-CORDEX data. This was essential; otherwise, the downscaled fields will be smoother than the actual 1 km data. The step-by-step procedure for downscaling the climate data for a given Area-DDF location using random mixing is described below:

1. Create a buffer enclosing the ADDF largest area (1024 km^2).
2. Find all EURO-CORDEX blocks falling within the buffer (conditional values).
3. Find all radar pixels falling within the buffer (simulation domain).
4. For every hour in the projected RCP8.5 data, read the corrected values.
5. Fit a non-parametric marginal distribution using KDE with a Gaussian kernel.
6. Optimize the kernel width and build CDF and INVCDF.
7. Transform the observations to standard normal space using the fitted CDF.
8. Fit a Gaussian copula model for spatial dependence.
9. Fit an exponential or spherical covariance model.
10. Select the model with the largest likelihood.
11. Correct the model to match the reference point model.
12. Run Random Mixing for conditional simulations (50 simulations).
13. Back-transform to original data space using INVCDF.
14. Repeat for next time steps with positive Pcp.

An example of the simulation domain for one ADDF location is shown in panel (a) of Fig.8.11. The domain shown in red has a total area of 1024 km^2. Within this domain, the areas of 16 and 576 km^2 were considered for deriving the ADDF curves. These were derived as DDF curves, but by using a time series of the average of all pixels within each area instead of a pixel time series. For example, for the area size of 1024 km^2, for every time step, the average of the pixels in red was calculated. Afterwards, using the same procedure for DDF curves (described in chapter 6), the ADDF curve for this area was derived.



(a) Simulation domain (b) VG scaling (c) Simulated field



(d) Average for one time step (e) Average pixel wise

Figure 8.11: Panel (a) simulation domain for one ADDF area (shown in red) and panel (b) example of one simulated hourly field. Panel (c) shows one simulation for one time step. Panel (d) depicts the average field of 50 simulations for the same time step, and panel (e) depicts the average for one pixel over the simulation period.

In panel (a) of Fig.8.11, the gray pixel represents the complete simulation domain, and the blue dots are the center of the EURO-CORDEX blocks. These served as conditional values for the downscaling. In panel (b), an example of the calculated variogram using the block (12 km) and pixel (1km) values. Note the difference in the range and sill of the variograms.

For downscaling, the block variogram was rescaled to the pixel-based one. In panel (c), a single realization based on the recorrelated and double-QQ-corrected RCP8.5 data for the simulation domain is presented. Panel (d) of Fig.8.11 displays the average simulated field derived from 50 realizations for a one-time step. The average field is highly similar to the interpolated one (not shown here), with a maximum precipitation value of 12.32 mm per hour. Panel (e) depicts the simulated hourly time series for one pixel over the whole period. In the ensemble of the simulations for one pixel, the variability of the values can be better visualized.

8.3 Results

8.3.1 DDF curves for future scenarios

The downscaled fields were used to calculate the ADDF and DDF curves for the selected region or pixel in the study area. The aim was to derive areal extremes for future periods, especially from the RCP8.5 data. For every time step, 50 simulations were generated over the simulation domain for each duration separately. The selected durations range from hourly to daily resolutions. This is essential; otherwise, the fields will not be space-time continuous. Random mixing generates realizations that are all equally possible but are limited to spatial constraints and are not temporally connected (advection is not included). For example, on the hourly scale, each realization for each time step will most likely differ from the realization of the next time step, despite being equally probable and statistically correct. Aggregating these fields will result in a false representation of areal rainfall. A simple but computationally intensive solution was to aggregate the hourly corrected data for each required duration and run the simulations again. A different possible solution would have been to use the generated fields for time step i as unconditional fields (instead of random fields) for time step $i + 1$. This would have required modifying the simulation algorithm to incorporate time as a third dimension. However, since the focus is on areal statistics, especially yearly maxima, and not on event reconstruction, the first solution was seen as adequate enough for this scope. Moreover, an alternative stochastic simulation approach could have been tested. For example, *Papalexiou et al.* (2021) presents a framework for simulating space-time rainfall fields with characteristics such as velocity field, advection and anisotropy.

Once the simulations were terminated, the DDF and subsequent ADDF curves were calculated from the generated time series. For comparison purposes, the simulations were done using the raw and corrected RCP8.5 data. Fig.8.12 shows an example of the DDF values for the centre pixel in the ADDF location for the return period of 5 years and 2 different periods. Panel (a) for the period 2005-2005 and panel (b) for the period 2065-2099. Each box plot for every duration consists of 50 simulations. The DDF values for the different durations from the raw and corrected RCP8.5 data are displayed in the blue and green boxes, respectively. The blue crosses and the green dots present the outliers in the raw and corrected RCP8.5 data, respectively. As reference data, the DDF curve from RADOLAN data for the period 2005-2020 is displayed in the red dashed line (and crosses). These fall within the bounds of the simulations. Compared to the RADOLAN data, the raw RCP8.5 values for both periods show an overestimation of the maxima over all durations. The corrected values, however, fall within the range of the RADOLAN data and show a slight underestimation for the period 2005-2025. Additionally, the box plots of the raw data indicate a larger range compared to the corrected data. For example, for the duration of 720 minutes (12 hours), the estimated rainfall depth from raw data varies between 40 and 100 mm with a mean value of 60 mm. In the corrected data, however, the range is between 25 and 40 mm. This indicates that the raw data have a larger spread and variability, limiting the use of a reliable uncertainty interval.

An increase in the expected rainfall depth from the first to the second period was clearly noticeable in both data sets. Though with different magnitudes. The raw data showed for the hourly duration almost a doubling of the values (from 24 to 39 mm per hour). In the corrected data, however, the increase was from 13 to 19 mm per hour. Both data sets indicated that the RCP8.5 projected scenario is associated with a larger rainfall depth over all durations.

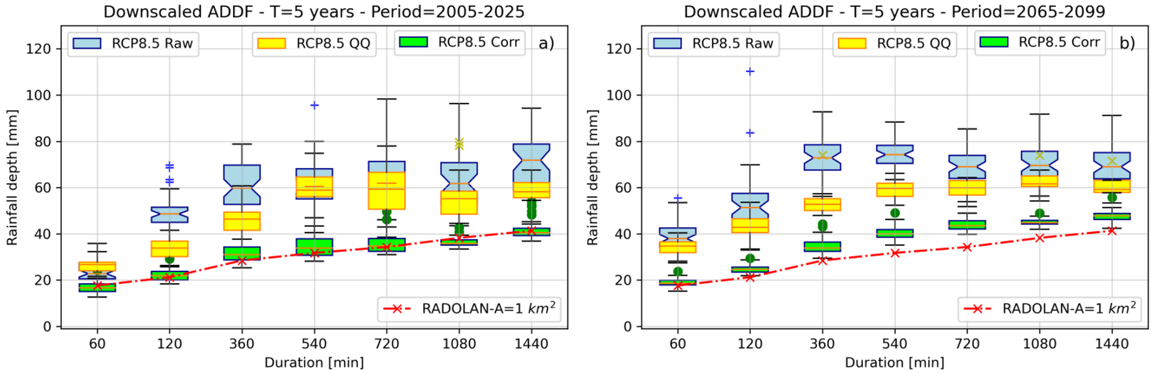


Figure 8.12: Derived ADDF curves (A=1 km^2) from RCP8.5 data before and after data correction for the ADDF center pixel for two different periods and a return period of 5 years. In panel (a) for the period 2005-2025 and panel (b) for the period 2065-2099. For every duration (x-axis), 50 simulations were generated and summarized in the boxplots. In both panels, the blue boxes refer to the raw RCP8.5 data, the yellow boxes to the double-QQ corrected data (without recorrelation), and the green boxes to the recorrelated and double-QQ corrected RCP8.5 data. The rainfall depth values derived from the RADOLAN data for the period 2005-2020 are displayed by the red crosses (or red curve).

8.3.2 ADDF curves for future scenarios

To investigate the effect of the recorrelation on the areal extremes, the ADDF curves were derived from the double-QQ corrected data only. In other words, the CDF of RCP8.5 data was corrected using the double-QQ procedure and used for downscaling to the finer spatial scale. From the downscaled fields, the ADDF curves were derived and compared to those derived with the recorrelated and double-QQ-corrected fields. The results of this approach are shown in the orange boxes in panels (a) and (b) of Fig. 8.13. Although the bounds (the scatter) of the simulations are reduced by the double-QQ correction, an overestimation of the ADDF curves was persistent over all area sizes, durations, and temporal periods. This supports the need for the recorrelation procedure to ensure a better representation of the spatial dependence correction and subsequent areal extremes.

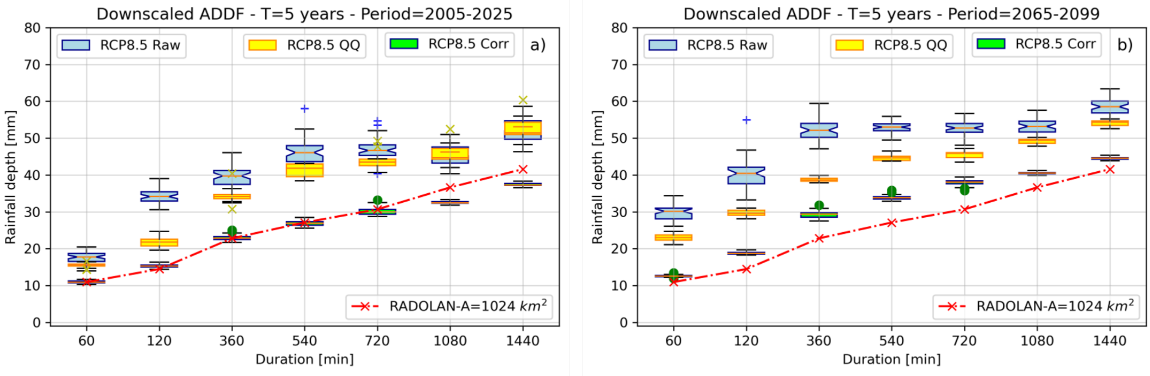


Figure 8.13: Estimated rainfall depth and ADDF curve from RCP8.5 data before and after data correction for the ADDF area of 1024 km^2 for a return period of 5 years. For every duration (x-axis), 50 simulations were generated and summarized in the boxplots. Panel (a) shows the results for the period 2005-2025. In panel (b), the ADDF curve for the period 2065-2099 is displayed. In both panels, the blue boxes refer to the raw RCP8.5 data, the yellow boxes to the double-QQ corrected data (without recorrelation), and the green boxes to the recorrelated and double-QQ corrected RCP8.5 data. The rainfall depth values derived from the RADOLAN data for the period 2005-2020 are displayed by the red crosses (or red curve).

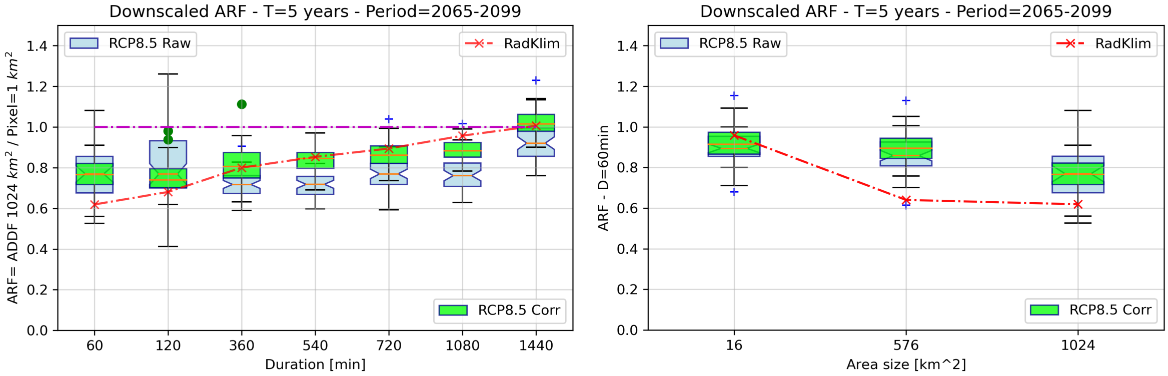
Panel (a) shows the hourly values for the ADDF curve for an area size of A=1024 km^2 and a return period of 5 years divided by the chosen climatic periods. The results for the areal statistics are similar to those for pixel-based analysis. Firstly, compared to the RadKlim-derived values, raw RCP8.5 data show a large increase in the areal maxima. From the first period (2005-2025) to the second period (2025-2045) the average hourly areal rainfall depth increased from 18 mm to 32 mm. Almost a doubling of the values. The period 2045-2065 shows a slight increase in the mean and the period 2065-2099 a slight decrease. The downscaled and corrected data indicate an increase in the rainfall maxima, but with a smaller magnitude. Between the first and last periods, the hourly rainfall depth increases on average from 10 to 15 mm. The values for the areal statistics are in compliance with the radar-derived results. For the raw data, the uncertainty interval (minimum and maximum) of the box plots is quite large; the corrected data show, however, a smaller uncertainty interval over all durations.

Panel (b) of Fig.8.13 shows the Area-DDF (ADDF) curves for the area size of 1024 km^2 for the period 2065-2099 and a return period of 5 years. The radar ADDF curve values are added to the plot (red curve) for comparison purposes. The raw RCP8.5 data are displayed in blue, and the corrected one is in green. Similarly to Fig.8.12, each box represents 50 simulations. The purpose of this analysis is to showcase if the pixel (point) and areal extremes behave in a similar manner. First, looking at the values derived from the raw data, an overestimation of the reference radar values is noted. The RCP8.5 corrected data presents a similar range as the weather radar data. For durations below 12 hours, the RCP8.5 values are larger than the radar values, however, the opposite case occurs for durations beyond 12 hours (720 min).

A possible explanation for this is that in the period 2065-2099, the probability of the occurrence of a rainfall event exceeding the radar values is minimal. To test this, the ADDF curves for the period 2045-2065 were calculated and show that for the same duration of 18 hours (1080 min), the corrected RCP8.5 data exceeds the radar values. Namely, within this period, the areal maxima for this duration were larger than in the later period. Interestingly, the uncertainty interval for larger areas is less than for smaller areas. In other words, the pixel DDF curve shows the largest variations, while the ADDF curve for A=1024 km^2 shows the smallest. Note that the point DDF are all derived from the same pixel located in the centre of the corresponding area. This indicates the risk associated with these values.

8.3.3 ARF values for future scenarios

A final point to mention was that the area reduction factor (ARF) values could be calculated as the ratio of pixel to area DDF (or the ratio between large and smaller ADDF curves). An example of this can be seen in panels (a) and (b) of Fig.8.14. Panel (a) shows the ARF calculated as the ratio between the estimated rainfall depth of 1024 km^2 and 1 km^2 area sizes. The x-axis denotes the duration (from hourly to daily) and the y-axis the ARF (usually between 0 and 1). As the duration increases, the ARF increases. In other words, for large durations, the areal maxima for small and large areas are in a similar range. However, for the hourly duration, the RadKlim values show an ARF of 0.62 (areal maxima is 62% of pixel maxima). Interestingly, the raw and corrected values are almost identical in the mean and have comparable ranges. The ARF on the hourly duration has a mean value of 0.75 (75 %) and indicates an increase in the areal precipitation of 20 % for the period 2065-2099. An increase in the areal mean was also seen for the 2-hour duration. The raw RCP8.5 shows a simulation with an hourly ARF value greater than one. For the other durations (between hourly and daily), the corrected data show a larger ARF (similar to RaKlim data) as compared to the raw data. In panel (b) of Fig.8.14, the duration is kept constant (hourly values), but the area size is changing. The x-axis shows three area sizes (16, 576, and 1024 km^2) and on the y-axis, the ARF is calculated as the ratio between each areal rainfall and centre pixel values. For example, the first RadKlim value of 0,96 for the area size of 16 km^2 indicates that the areal rainfall depth is 96% of the pixel (1 km^2) estimated value. As expected, for larger areas the ratio decreases. Between the raw and corrected RCP8.5 data, the difference is minimal. However, between the RADOLAN ARF and RCP8.5 ARF, the difference is notable, especially for larger areas, for which the RCP data show larger ARF values. The correction of the RCP data seems to keep the relationship between point and areal extremes preserved. Note that these results are for the period of 2065-2099 with a return period of 5 years. The results for other periods are similar.



(a) ARF for the period 2065-2099

(b) ARF function of area size

Figure 8.14: Derived ARF value for RCP8.5 before and after correction. The results are for the area sizes of 1 and 1024 km^2. The ARF values derived from the RADOLAN data for the period 2005-2020 are displayed by the red dots in both panels. Panel (a) depicts the ARF for different durations and panel (b) shows the ARF for different area sizes

8.4 Summary

Investigating the changes in the statistical properties of climatic variables, such as rainfall, due to a changing climate is essential for coping and preparing for future periods. Regional climate models provide useful information about climatic data for historical and future scenarios. These have been generated based on increasing emission levels, thereby changing the physical, energetic, and thermodynamic balance between the atmospheric components. The outcome of RCM data is in general too coarse for local analysis, and often spatial or temporal downscaling is applied. For reasonably downscaled values, the original RCM data should be inspected and eventually corrected. Here, two major aspects were crucial. The first is related to the spatial dependence structure. This influenced the distribution of areal rainfall and, as a false structure, altered any subsequent results. The second was related to the presence of a bias in the data for future scenarios. A bias in any direction (overestimation or underestimation) of the marginal distribution function affected the temporal structure and the quality of the data. If the model distribution differed largely from the observed one, it would be as if it were not a realization of the same process. Hence, the first and second parts of this work were related to correcting the spatial dependence structure and marginal function, respectively. Both of these steps were undertaken on the same spatial and temporal scale as the model data. This required upscaling the reference data to the model scale. The corrected data could be further integrated into a downscaling scheme. In this part, a probabilistic scheme involving conditional simulations using random mixing was applied. For each time step and duration, several realizations on the 1-km scale were generated and used for analyzing areal extremes. The final results show realistic DDF and ADDF curves that can be used for design purposes.

9 Conclusions

In this thesis, the spatial extent of precipitation extremes in Hydrology was investigated. The main goal of this work was to derive new methods for characterizing areal extremes and gain additional insights into their statistical properties. New aspects regarding ADDF curves, copula merging of weather and gauge data, statistics of event volumes, and stochastic downscaling of climate weather data were introduced. Eventually, this work showed that transferring points to areal statistics must involve a space-time approach using spatially distributed data.

After the introduction (first chapter), several statistical measures were described in the second chapter. Typical dependence measures and notions used throughout the work were presented. An outlier detection method was developed and implemented for the largest yearly observations from the rain gauge data. The detected outliers were verified by additional information, and false observations were disregarded further on. The procedure was presented in the third chapter. This was an essential step in acquiring reliable observation data for evaluating the rainfall maxima. The corrected rain gauge data were used to derive the spatial extent of yearly maxima over different temporal aggregations.

Several approaches, such as variogram and step-function clustering, event-based correlation values, and detection of unusual events by the depth-function tool, were elaborated on in the fourth chapter. The focus was on simultaneously occurring unusual events. The results revealed that effects such as anisotropy and advection can play an important role in a rainfall event. Moreover, the probability of neighbouring observations having larger values than local maxima was calculated. Showcasing that on the hourly scale, for around 25% of the events, the maxima was at one of the nearest neighbours. This indicated that the point observations missed a large number of intense values. Moreover, the observed point values might have underestimated the rainfall maxima within an event, depending on their location. This was especially noted on the hourly and sub-hourly scales.

To that end, the fifth chapter introduced a new copula-based merging of point and radar data. Several aspects, such as high-temporal analysis, anisotropy, advection and the wind displacement vector were incorporated. The method was tested for two radar regions and validated with a cross-validation of the DDF for 30 rain gauges. The result showed an unbiased and reliable product suitable for further use. Despite many efforts to obtain an error 'free' product, automating the process showed that, in some cases, the fields were not reliable. For example, if no careful care is taken when deriving the variogram model, false parameters will introduce errors in the final field, especially for low-intensity rainfall fields. However, the overall performance of the merged data, especially with regard to extreme value statistics, was acceptable. The merged product as well as RadKlim were used to investigate space-time statistics for areal extremes.

Chapter six introduced the derivation of ADDF curves and associated area reduction factors for several locations within the radar area of Hannover. The derived ADDF curves can be theoretically used for deriving design values and replacing the traditional use of ARF. In that way, for a given area size (or catchment size), the rainfall depth associated with the corresponding duration and return period can be derived. The derived ARF values were compared to reference ones, showcasing that the latter underestimate the average ARF for a short duration. The variability in the ARF values was found to be very large. Using an average ARF value for a duration or area size can lead to a large underestimation or overestimation. In addition, the analysis of the ADDF curves showed the presence of crossings of the curves between point and area DDF curves and small to large areas. An investigation of the crossings revealed that for long durations, events can occur where the areal mean exceeds the centre areal value. Such cases lead to crossings of the curves. However, if a longer observation period were present, the sampled areal values would change and the crossingswouldl disappear. The crossings are also influenced by the fact that, given an area, the location of the maximum can occur anywhere. Hence, considering the centre point for the point DDF might lead to an underestimation of the 'true' point extremes. The ARF values were linked to the catchment size, form, and orientation, and a spatial mapping of ARF values for several catchments was undertaken.

The weather radar data were further exploited, and a new way to look at extremes was developed: rainfall as 2-dimensional and 3-dimensional connected blocks. This was the content of chapter seven. A large data set of connected 2D areas was acquired from the 20-year RadKlim data. The connected areas were classified, and the frequency of occurrence of the corresponding area size with an areal mean over a region was derived. Interestingly, there are regions that experience larger blocks than smaller ones for the same area size and season. A strong dependence in winter on topography was found, and in summer, a north-to-south gradient for frequency of occurrence was found. Moreover, the largest areal mean and areal maxima were found to be associated with small-scale events. The analysis was continued into 3D blocks with a focus on three headwater catchments in the southwest of Germany. The event volume, areal extent, duration, in-catchment volume, and corresponding discharge reaction were statistically related. Not only are strong events relevant, but slow-moving rainfall cells could lead to a high discharge peak.

The final chapter used the EURO-CORDEX 11 hourly data for historical and future periods to derive the ADDF curves. An upscaling of the 1-km reference data (interpolation based on DWD gauges, aggregated radar data, and simulation based on the DWD gauges) to the model scale (12.5 km) served as the basis of model data quality analysis. First, the spatial dependence structure of the model was compared to the reference one, and a recorrelation procedure was implemented. This ensured that the model has a similar spatial dependence structure essential for areal extremes. Secondly, the double-QQ transformation of the marginal distribution function allowed conserving the signal in the model projections while reducing the bias. Both procedures were implemented for every season separately and showed success in improving the model data quality. The final step was a downscaling of the corrected model data using a stochastic conditional simulation method to derive the ADDF curves for future periods. To ensure that advection and correct spatial and temporal aggregation of the spatial field were achieved, the simulation was done for each duration separately. The final derived curves showed better agreement than those using the original data and fell within the range of the weather radar curves. Moreover, for future climatic periods, the ADDF curves showed an increase in the estimated rainfall depth.

In conclusion, this work presented new methods and results concerning the space-time statistics of rainfall maxima. The focus was laid on the spatial extent of areal extremes and on transferring points to areal statistics. Below are important remarks:

1. Analysing the spatial extent of extremes only with rain gauges was beneficial but not sufficient to observe the space-time dynamics of rainfall events.
2. Merging weather radar and rain gauge data in the rank space enhanced the performance of the results.
3. ADDF curves can be derived from weather radar data; however, longer records are needed to avoid problems with small samples.
4. ARF was found to be larger than one for events with long durations, a strong advection vector, or/and under sampling of the pixel (point) extremes.
5. For several catchments, the ARF was derived, and the values were linked to the catchment size, form, and orientation.
6. Areal extremes were not only space-time-dependent but also changed with the object size, duration, time of occurrence, and geographical location.
7. High discharge peaks were not necessarily caused by hourly or daily rainfall maxima but by an accumulation of slow-moving rainfall cells.
8. Correcting the dependence structure and marginal distribution function of regional climate model data provided a foundation for reliable further analysis.
9. Deriving ADDF for future (and present) periods was associated with an uncertainty measure to showcase the possible error term.

Future work can build on the methods developed and provided in this thesis. The results are transferable to other locations and can be applied to other data sets. The developed methods could be used to update state-of-the-art information regarding area reduction factors and transferring point to areal extremes. Moreover, investigating the volume of rainfall blocks along their space-time development and spatial characteristics could be further improved and incorporated into a rainfall-runoff model.

Bibliography

Ahmed, S., and G. De Marsily, Comparison of geostatistical methods for estimating transmissivity using data on transmissivity and specific capacity, *Water resources research*, *23*(9), 1717–1737, 1987.

Allen, R. J., and A. T. DeGaetano, Areal reduction factors for two eastern united states regions with high rain-gauge density, *Journal of Hydrologic Engineering*, *10*(4), 327–335, 2005.

Anagnostou, E. N., and W. F. Krajewski, Real-time radar rainfall estimation. part i: Algorithm formulation, *Journal of Atmospheric and Oceanic Technology*, *16*(2), 189–197, 1999.

Asquith, W., and J. Famiglietti, Precipitation areal-reduction factor estimation using an annual-maxima centered approach, *Journal of Hydrology*, *230*(1-2), 55–69, 2000.

Atlas, D., D. Rosenfeld, and D. A. Short, The estimation of convective rainfall by area integrals: 1. the theoretical and empirical basis, *Journal of Geophysical Research: Atmospheres*, *95*(D3), 2153–2160, 1990.

Ayat, H., J. P. Evans, S. C. Sherwood, and J. Soderholm, Intensification of subhourly heavy rainfall, *Science*, *378*(6620), 655–659, 2022.

Ayzel, G., M. Heistermann, and T. Winterrath, Optical flow models as an open benchmark for radar-based precipitation nowcasting (rainymotion v0.1), *Geoscientific Model Development*, *12*(4), 1387–1402, doi:10.5194/gmd-12-1387-2019, 2019.

Bárdossy, A., Copula-based geostatistical models for groundwater quality parameters, *Water Resources Research*, *42*(11), 2006.

Bárdossy, A., and Z. Kundzewicz, Geostatistical methods for detection of outliers in groundwater quality spatial fields, *Journal of Hydrology*, *115*(1-4), 343–359, doi:10.1016/0022-1694(90)90213-H, 1990.

Bárdossy, A., and J. Li, Geostatistical interpolation using copulas, *Water resources research*, *44*(7), 2008.

Bárdossy, A., and G. Pegram, Downscaling precipitation using regional climate models and circulation patterns toward hydrology, *Water Resources Research*, *47*(4), doi:https://doi.org/10.1029/2010WR009689, 2011.

Bárdossy, A., and G. Pegram, Multiscale spatial recorrelation of rcm precipitation to produce unbiased climate change scenarios over large areas and small, *Water Resources Research*, *48*(9), 2012.

Bárdossy, A., and G. Pegram, Interpolation of precipitation under topographic influence at different time scales, *Water Resources Research*, *49*(8), 4545–4565, 2013.

Bárdossy, A., and G. Pegram, Intensity–duration–frequency curves exploiting neighbouring extreme precipitation data, *Hydrological Sciences Journal*, *63*(11), 1593–1604, doi:10.1080/02626667.2018.1524987, 2018.

Bárdossy, A., and S. Singh, Robust estimation of hydrological model parameters, *Hydrology and earth system sciences*, *12*(6), 1273–1283, 2008.

Bárdossy, A., and S. K. Singh, Regionalization of hydrological model parameters using data depth, *Hydrology Research*, *42*(5), 356–371, doi:10.2166/nh.2011.031, 2011.

Bárdossy, A., J. Seidel, and A. El Hachem, The use of personal weather station observations to improve precipitation estimation and interpolation, *Hydrology and Earth System Sciences*, *25*(2), 583–601, doi:10.5194/hess-25-583-2021, 2021.

Barnett, V., and T. Lewis, *Outliers in statistical data*, 608 pp., John Wiley and Sons Location, Hoboken, NJ, 1994.

Basist, A., G. D. Bell, and V. Meentemeyer, Statistical relationships between topography and precipitation patterns, *Journal of Climate*, *7*(9), 1305 – 1315, doi:10.1175/1520-0442(1994)007⟨1305:SRBTAP⟩2.0.CO;2, 1994.

Bayerisches Landesamt für Umwelt, Wasserstand und abfluss, 2022.

Benestad, R. E., A simple test for changes in statistical distributions, *Eos, Transactions American Geophysical Union*, *89*(41), 389–390, 2008.

Bennett, B., M. Lambert, M. Thyer, B. C. Bates, and M. Leonard, Estimating extreme spatial rainfall intensities, *Journal of Hydrologic Engineering*, *21*(3), 04015,074, 2016.

Berg, P., O. B. Christensen, K. Klehmet, G. Lenderink, J. Olsson, C. Teichmann, and W. Yang, Summertime precipitation extremes in a euro-cordex 0.11 deg ensemble at an hourly resolution, *Natural Hazards and Earth System Sciences*, *19*(4), 957–971, 2019.

Berndtsson, R., On the use of cross-correlation analysis in studies of patterns of rainfall variability, *Journal of Hydrology*, *93*, 113–134, doi:10.1016/0022-1694(87)90198-3, 1987.

Box, G. E., and D. R. Cox, An analysis of transformations, *Journal of the Royal Statistical Society: Series B (Methodological)*, *26*(2), 211–243, 1964.

Bravais, A., Mathematical analysis on the probabilities of the situation errors of a point, *Memories presented by various scientists. The Royal Academy of Sciences of the Institute de France*, *9*, 255–332, 1846.

Brommundt, J., and A. Bárdossy, Spatial correlation of radar and gauge precipitation data in high temporal resolution, *Advances in Geosciences*, *10*, 103–109, 2007.

Burauskaite-Harju, A., A. Grimvall, C. Achberger, A. Walther, and D. Chen, Characterising and visualizing spatio-temporal patterns in hourly precipitation records, *Theoretical and Applied Climatology*, *109*(3), 333–343, 2012.

Bureau, W., *Rainfall intensity-frequency regime*, US Department of Commerce, Weather Bureau, 1957.

Burkardt, J., The truncated normal distribution, *Department of Scientific Computing Website, Florida State University*, *1*, 35, 2014.

Catchlove, R. H., and J. E. Ball, A hydroinformatic approach to the development of areal reduction factors, in *28th International Hydrology and Water Resources Symposium: About Water; Symposium Proceedings*, pp. 1–9, Institution of Engineers, Australia Barton, ACT, 2003.

Chwala, C., and J. Polz, cchwala/radolan_to_netcdf, doi:10.5281/zenodo.4452204, 2021.

Cressie, N., Spatial prediction and ordinary kriging, *Mathematical geology*, *20*(4), 405–421, 1988.

Diday, E., and J. Simon, Clustering analysis, in *Digital pattern recognition*, pp. 47–94, Springer, 1976.

Diggle, P. J., P. J. Ribeiro, and O. F. Christensen, An introduction to model-based geostatistics, in *Spatial statistics and computational methods*, pp. 43–86, Springer, 2003.

Doneaud, A., S. Ionescu-Niscov, D. L. Priegnitz, and P. L. Smith, The area-time integral as an indicator for convective rain volumes, *Journal of Applied Meteorology and Climatology*, *23*(4), 555 – 561, doi:10.1175/1520-0450(1984)023⟨0555:TATIAA⟩2.0.CO;2, 1984.

Durre, I., M. J. Menne, B. E. Gleason, T. G. Houston, and R. S. Vose, Comprehensive automated quality assurance of daily surface observations, *Journal of Applied Meteorology and Climatology*, *49*(8), 1615–1633, 2010.

DWA, A., 118: Hydraulische bemessung und nachweis von entwässerungssystemen, *DWA-Regelwerk, Arbeitsblatt A*, *118*, 2006.

DWA-A, ., Arbeitsblatt DWA-A 531: Starkregen in Abhängigkeit von Wiederkehrzeit und Dauer, 2012.

DWD Climate Data Center (CDC), Historical daily station observations (temperature, pressure, precipitation,sunshine duration, etc.) for Germany, 2021a.

DWD Climate Data Center (CDC), Historical hourly RADOLAN grids of precipitation depth (GIS-readable), version 2.5, 2021b.

El Hachem, A., J. Seidel, F. Imbery, T. Junghänel, and A. Bárdossy, Technical note: Space–time statistical quality control of extreme precipitation observations, *Hydrology and Earth System Sciences*, *26*(23), 6137–6146, doi:10.5194/hess-26-6137-2022, 2022.

Ford, J. D., L. Cameron, J. Rubis, M. Maillet, D. Nakashima, A. C. Willox, and T. Pearce, Including indigenous knowledge and experience in ipcc assessment reports, *Nature Climate Change*, *6*(4), 349–353, 2016.

Gabella, M., and R. Notarpietro, Ground clutter characterization and elimination in mountainous terrain, in *Proceedings of ERAD*, vol. 305, 2002.

Ghebreyesus, D. T., and H. O. Sharif, Development and assessment of high-resolution radar-based precipitation intensity-duration-curve (idf) curves for the state of texas, *Remote Sensing*, *13*(15), 2890, doi:10.3390/rs13152890, 2021.

Gnedenko, B., and V. Korolyuk, On the maximum discrepancy between two empirical distributions, in *Dokl. Acad. Nauk., SSSR*, vol. 4, pp. 525–528, 1951.

Goshtsasbpour, G., U. Haberlandt, A. E. Hachem, J. Seidel, and A. Bárdossy, Statistical analysis of space-times dynamics of extreme precipitation, 2022.

Gravelius, H., Morphometry of drainage bassins, *Quaternary International, v. Goschen Ve*, 1914.

Gunst, R. F., Estimating spatial correlations from spatial-temporal meteorological data, *Journal of Climate*, *8*(10), 2454–2470, 1995.

Guthke, P., *Non-multi-Gaussian spatial structures: process-driven natural genesis, manifestation, modeling approaches, and influences on dependent processes*, 2013.

Guthke, P., and A. Bárdossy, On the link between natural emergence and manifestation of a fundamental non-gaussian geostatistical property: asymmetry, *Spatial Statistics*, *20*, 1–29, 2017.

Haberlandt, U., and C. Berndt, Development of rainfall area-intensity-duration-frequency curves from weather radar data, in *AGU Fall Meeting Abstracts*, vol. 2016, pp. H13J–1557, 2016.

Hawkins, D. M., *Identification of outliers*, 188 pp., Springer Dordrecht, doi:10.1007/978-94-015-3994-4, 1980.

Heilbron, D. C., Zero-altered and other regression models for count data with added zeros, *Biometrical Journal*, *36*(5), 531–547, 1994.

Heistermann, M., S. Jacobi, and T. Pfaff, An open source library for processing weather radar data (wradlib), *Hydrology and Earth System Sciences*, *17*(2), 863–871, 2013.

Hennemuth, T. I., et al., Guidance for euro-cordex climate projections data use, *Version1. 0-2017.08. Retrieved on*, *6*(2019), 118, 2017.

Hodges Jr, J., The significance probability of the smirnov two-sample test, *Arkiv för matematik*, *3*(5), 469–486, 1958.

Holleman, I., *Bias adjustment of radar-based 3-hour precipitation accumulations*, Citeseer, 2006.

Hörning, S., A. Bardossy, and S. Tyson, Simulation of conditional spatial fields using random mixing, in *Petroleum Geostatistics 2015*, pp. cp–456, European Association of Geoscientists & Engineers, 2015.

Hu, L. Y., Gradual deformation and iterative calibration of gaussian-related stochastic models, *Mathematical Geology*, *32*(1), 87–108, 2000.

Hubbard, K., S. Goddard, W. Sorensen, N. Wells, and T. Osugi, Performance of quality assurance procedures for an applied climate information system, *Journal of Atmospheric and Oceanic Technology*, *22*(1), 105–112, 2005.

Iglewicz, B., and D. Hoaglin, The asqc basic references in quality control: statistical techniques, in *How to detect and handle outliers*, vol. 16, edited by E. Mykytka, pp. 1–87, ASQC Quality Press Milwaukee, WI, 1993.

Ingleby, N. B., and A. C. Lorenc, Bayesian quality control using multivariate normal distributions, *Quarterly Journal of the Royal Meteorological Society*, *119*(513), 1195–1225, 1993.

Jacob, D., A note to the simulation of the annual and inter-annual variability of the water budget over the baltic sea drainage basin, *Meteorology and Atmospheric Physics*, *77*(1), 61–73, 2001.

Jacob, D., et al., Euro-cordex: new high-resolution climate change projections for european impact research, *Regional environmental change*, *14*, 563–578, 2014.

Kadow, C., et al., Introduction to freva–a free evaluation system framework for earth system modeling, *Journal of Open Research Software*, *9*(1), 2021.

Kim, J., J. Lee, D. Kim, and B. Kang, The role of rainfall spatial variability in estimating areal reduction factors, *Journal of Hydrology*, *568*, 416–426, 2019.

Klemeš, V., Tall tales about tails of hydrological distributions. i, *Journal of Hydrologic Engineering*, *5*(3), 227–231, 2000.

Kolmogorov, A. N., Foundations of the theory of probability, 2nd english edition, *NY: Chelsea Publishing Co*, 1933.

Kotlarski, S., et al., Regional climate modeling on european scales: a joint standard evaluation of the euro-cordex rcm ensemble, *Geoscientific Model Development*, *7*(4), 1297–1333, doi:10.5194/gmd-7-1297-2014, 2014.

Koutsoyiannis, D., Statistics of extremes and estimation of extreme rainfall: Ii. empirical investigation of long rainfall records/statistiques de valeurs extrêmes et estimation de précipitations extrêmes: Ii. recherche empirique sur de longues séries de précipitations, *Hydrological Sciences Journal*, *49*(4), 2004.

Koutsoyiannis, D., and T. Iliopoulou, Ombrian curves advanced to stochastic modeling of rainfall intensity, *Rainfall*, pp. 261–284, 2022.

Koutsoyiannis, D., and S. M. Papalexiou, Extreme rainfall: Global perspective, *Handbook of Applied Hydrology; McGraw-Hill: New York, NY, USA*, pp. 74–1, 2017.

Koutsoyiannis, D., D. Kozonis, and A. Manetas, A mathematical framework for studying rainfall intensity-duration-frequency relationships, *Journal of hydrology*, *206*(1-2), 118–135, 1998.

Kreklow, J., B. Tetzlaff, B. Burkhard, and G. Kuhnt, Radar-based precipitation climatology in germany-developments, uncertainties and potentials, *Atmosphere*, *11*, doi:10.3390/atmos11020217, 2020.

Langousis, A., The areal reduction factor (arf): a multifractal analysis, Ph.D. thesis, Massachusetts Institute of Technology, 2005.

Lebrenz, H., and A. Bárdossy, Geostatistical interpolation by quantile kriging, *Hydrology and Earth System Sciences*, *23*(3), 1633–1648, 2019.

Lengfeld, K., T. Winterrath, T. Junghänel, and A. Becker, Characteristic spatial extent of rain events in germany from a radar-based precipitation climatology, 2018.

Lengfeld, K., T. Winterrath, T. Junghänel, and A. Becker, Characteristic spatial extent of rain events in germany from a radar-based precipitation climatology, in *Rainfall Monitoring, Modelling and Forecasting in Urban Environment. UrbanRain18: 11th International Workshop on Precipitation in Urban Areas. Conference Proceedings*, pp. 61–66, ETH Zurich, Institute of Environmental Engineering, 2019.

Lengfeld, K., P.-E. Kirstetter, H. J. Fowler, J. Yu, A. Becker, Z. Flamig, and J. Gourley, Use of radar data for characterizing extreme precipitation at fine scales and short durations, *Environmental Research Letters*, *15*(8), 085,003, 2020.

LUBW Landesanstalt für Umwelt, M. u. N. B.-W., Abfluss-bw, 2019.

Lucas, B., and T. Kanade, An iterative image registration technique with an application to stereo vision (ijcai), 1981.

Lucas, B. D., T. Kanade, et al., *An iterative image registration technique with an application to stereo vision*, vol. 81, Vancouver, 1981.

Marcotte, D., Fast variogram computation with fft, *Computers & Geosciences*, *22*(10), 1175–1186, 1996.

Marra, F., and E. Morin, Use of radar qpe for the derivation of intensity–duration–frequency curves in a range of climatic regimes, *Journal of Hydrology*, *531*, 427–440, doi:https://doi.org/10.1016/j.jhydrol.2015.08.064, hydrologic Applications of Weather Radar, 2015.

Marzban, C., and S. Sandgathe, Verification with variograms, *Weather and Forecasting*, *24*(4), 1102 – 1120, doi:10.1175/2009WAF2222122.1, 2009.

Matheron, G., *Le krigeage universel*, vol. 1, École nationale supérieure des mines de Paris Paris, 1969.

Matte, D., J. H. Christensen, and T. Ozturk, Spatial extent of precipitation events: when big is getting bigger, *Climate Dynamics*, *58*(5), 1861–1875, 2022.

Merceret, F. J., and J. G. Ward, Attenuation of weather radar signals due to wetting of the radome by rainwater or incomplete filling of the beam volume, *Tech. rep.*, 2000.

Meredith, E. P., U. Ulbrich, H. W. Rust, and H. Truhetz, Present and future diurnal hourly precipitation in 0.11 euro-cordex models and at convection-permitting resolution, *Environmental Research Communications*, *3*(5), 055,002, 2021.

Nagy, S., C. Schütt, and E. M. Werner, Halfspace depth and floating body, *Statistics Surveys*, *13*, 52–118, 2019.

Pachauri, R. K., et al., *Climate change 2014: synthesis report. Contribution of Working Groups I, II and III to the fifth assessment report of the Intergovernmental Panel on Climate Change*, Ipcc, 2014.

Papalexiou, S. M., F. Serinaldi, and E. Porcu, Advancing space-time simulation of random fields: From storms to cyclones and beyond, *Water Resources Research*, *57*(8), e2020WR029,466, 2021.

Pearson, K., Liii. on lines and planes of closest fit to systems of points in space, *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, *2*(11), 559–572, 1901.

Pulkkinen, S., D. Nerini, A. A. Pérez Hortal, C. Velasco-Forero, A. Seed, U. Germann, and L. Foresti, Pysteps: an open-source python library for probabilistic precipitation nowcasting (v1. 0), *Geoscientific Model Development*, *12*(10), 4185–4219, 2019a.

Pulkkinen, S., D. Nerini, A. A. Pérez Hortal, C. Velasco-Forero, A. Seed, U. Germann, and L. Foresti, Pysteps: an open-source python library for probabilistic precipitation nowcasting (v1.0), *Geoscientific Model Development*, *12*(10), 4185–4219, doi:10.5194/gmd-12-4185-2019, 2019b.

Qi, Y., S. Martinaitis, J. Zhang, and S. Cocks, A real-time automated quality control of hourly rain gauge data based on multiple sensors in mrms system, *Journal of Hydrometeorology*, *17*(6), 1675–1691, 2016.

Quenouille, M. H., Approximate tests of correlation in time-series, *Journal of the Royal Statistical Society: Series B (Methodological)*, *11*(1), 68–84, 1949.

Quenouille, M. H., Notes on bias in estimation, *Biometrika*, *43*(3/4), 353–360, 1956.

Refaeilzadeh, P., L. Tang, and H. Liu, Cross-validation., *Encyclopedia of database systems*, *5*, 532–538, 2009.

Řezanková, H., and B. Everitt, Cluster analysis and categorical data, *Statistika*, *89*(3), 216–232, 2009.

Robert, L., Thorndike."who belongs in the family?", *Psychometrika*, *18*(4), 267–76, 1953.

Sassolas-Serrayet, T., R. Cattin, and M. Ferry, The shape of watersheds, *Nature communications*, *9*(1), 3791, 2018.

Serinaldi, F., Analysis of inter-gauge dependence by kendall's τk, upper tail dependence coefficient, and 2-copulas with application to rainfall fields, *Stochastic Environmental Research and Risk Assessment*, *22*(6), 671–688, 2008.

Sinclair, S., and G. Pegram, Combining radar and rain gauge rainfall estimates using conditional merging, *Atmospheric Science Letters*, *6*(1), 19–22, 2005.

Singh, V. P., On application of the weibull distribution in hydrology, *Water Resources Management*, *1*, 33–43, 1987.

Smirnov, N. V., On the estimation of the discrepancy between empirical curves of distribution for two independent samples, *Bull. Math. Univ. Moscou*, *2*(2), 3–14, 1939.

Spearman, Charles Edward, The proof and measurement of association between two things, 1863-1945.

Steinhaus, H., et al., Sur la division des corps matériels en parties, *Bull. Acad. Polon. Sci*, *1*(804), 801, 1956.

Svensson, C., and D. A. Jones, Review of methods for deriving areal reduction factors, *Journal of Flood Risk Management*, *3*(3), 232–245, 2010.

Tan, X., X. Wu, and B. Liu, Global changes in the spatial extents of precipitation extremes, *Environmental Research Letters*, *16*(5), 054,017, 2021.

Tanaka, A., et al., Firm productivity and foreign direct investment in the services sector: A firm-level analysis using japanese data, *Tech. rep.*, 2012.

Teegavarapu, R. S., and V. Chandramouli, Improved weighting methods, deterministic and stochastic data-driven models for estimation of missing precipitation records, *Journal of hydrology*, *312*(1-4), 191–206, 2005.

Thorndahl, S., J. E. Nielsen, and M. R. Rasmussen, Estimation of storm-centred areal reduction factors from radar rainfall for design in urban hydrology, *Water*, *11*(6), 1120, 2019.

Touma, D., A. M. Michalak, D. L. Swain, and N. S. Diffenbaugh, Characterizing the spatial scales of extreme daily precipitation in the united states, *Journal of Climate*, *31*(19), 8023–8037, 2018.

Tukey, J. W., Mathematics and the picturing of data, in *Proceedings of the International Congress of Mathematicians, Vancouver, 1975*, vol. 2, pp. 523–531, 1975.

Veneziano, D., and A. Langousis, The areal reduction factor: A multifractal analysis, *Water Resources Research*, *41*(7), 2005.

Verworn, H.-R., Flächenabhängige abminderung statistischer regenwerte, *Korrespondenz Wasserwirtschaft*, *1*(9), 493–498, 2008.

Virtanen, P., et al., Scipy 1.0: fundamental algorithms for scientific computing in python, *Nature methods*, *17*(3), 261–272, 2020.

Winterrath, T., C. Brendel, M. Hafer, T. Junghänel, A. Klameth, K. Lengfeld, E. Walawender, E. Weigl, and A. Becker, Radklim version 2017.002: Reprocessed quasi gauge-adjusted radar data, 5-minute precipitation sums (yw), *Sci. Tech. Data*, 2018a.

Winterrath, T., C. Brendel, M. Hafer, T. Junghänel, A. Klameth, K. Lengfeld, E. Walawender, E. Weigl, and A. Becker, Radklim version 2017.002: Reprocessed quasi gauge-adjusted radar data, 5-minute precipitation sums (yw), *Sci. Tech. Data*, 2018b.

Wright, S., Correlation and causation. joumal of, *Agricultural Research*, 1921.

Yan, J., and A. Bárdossy, *Nonlinear Estimation of Short Time Precipitation Using Weather Radar and Surface Observations*, Mitteilungen / Institut für Wasser- und Umweltsystemmodellierung, Universität Stuttgart, Eigenverlag des Instituts für Wasser- und Umweltsystemmodellierung der Universität Stuttgart, 2018.

Yu, W., T. Ai, and S. Shao, The analysis and delimitation of central business district using network kernel density estimation, *Journal of Transport Geography*, *45*, 32–47, 2015.

Yulizar, Y., and A. Bárdossy, Study of changes in the multivariate precipitation series, *Modeling Earth Systems and Environment*, *6*(2), 811–820, 2020.

Institut für Wasser- und
Umweltsystemmodellierung
Universität Stuttgart

Pfaffenwaldring 61
70569 Stuttgart (Vaihingen)
Telefon (0711) 685 - 60156
Telefax (0711) 685 - 51073
E-Mail: iws@iws.uni-stuttgart.de
http://www.iws.uni-stuttgart.de

Direktoren
Prof. Dr.-Ing. Rainer Helmig
Prof. Dr.-Ing. Wolfgang Nowak
Prof. Dr.-Ing. Silke Wieprecht

Emeriti
Prof. Dr.-Ing. habil. Dr.-Ing. E.h. Jürgen Giesecke
Prof. Dr.h.c. Dr.-Ing. E.h. Helmut Kobus, PhD

Lehrstuhl für Wasserbau und Wassermengenwirtschaft
Leiterin: Prof. Dr.-Ing. Silke Wieprecht
Stellv.: Dr.-Ing. Kristina Terheiden
Versuchsanstalt für Wasserbau
Leiter: Stefan Haun, PhD

Lehrstuhl für Hydromechanik und Hydrosystemmodellierung
Leiter: Prof. Dr.-Ing. Rainer Helmig
Stellv.: apl. Prof. Dr.-Ing. Holger Class

Lehrstuhl für Stochastische Simulation und Sicherheitsforschung für Hydrosysteme
Leiter: Prof. Dr.-Ing. Wolfgang Nowak
Stellv.: apl. Prof. Dr.-Ing. Sergey Oladyshkin
Hydrogeophysik der Vadosen Zone
(mit Forschungszentrum Jülich)
Leiter: Prof. Dr. J.A. Sander Huisman

***VEGAS,* Versuchseinrichtung zur Grundwasser- und Altlastensanierung**
Leiter: Dr.-Ing. Simon Kleinknecht
PD Dr.-Ing. Claus Haslauer

Verzeichnis der Mitteilungshefte

1 Röhnisch, Arthur: *Die Bemühungen um eine Wasserbauliche Versuchsanstalt an der Technischen Hochschule Stuttgart*, und Fattah Abouleid, Abdel: *Beitrag zur Berechnung einer in lockeren Sand gerammten, zweifach verankerten Spundwand*, 1963

2 Marotz, Günter: *Beitrag zur Frage der Standfestigkeit von dichten Asphaltbelägen im Großwasserbau*, 1964

3 Gurr, Siegfried: *Beitrag zur Berechnung zusammengesetzter ebener Flächentrag-werke unter besonderer Berücksichtigung ebener Stauwände, mit Hilfe von Rand-wert- und Lastwertmatrizen*, 1965

4 Plica, Peter: *Ein Beitrag zur Anwendung von Schalenkonstruktionen im Stahlwasserbau*, und Petrikat, Kurt: *Möglichkeiten und Grenzen des wasserbaulichen Versuchswesens*, 1966

5 Plate, Erich: *Beitrag zur Bestimmung der Windgeschwindigkeitsverteilung in der durch eine Wand gestörten bodennahen Luftschicht*, und Röhnisch, Arthur; Marotz, Günter: *Neue Baustoffe und Bauausführungen für den Schutz der Böschungen und der Sohle von Kanälen, Flüssen und Häfen; Gestehungskosten und jeweilige Vorteile*, sowie Unny, T.E.: *Schwingungsuntersuchungen am Kegelstrahlschieber*, 1967

6 Seiler, Erich: *Die Ermittlung des Anlagenwertes der bundeseigenen Binnenschiffahrtsstraßen und Talsperren und des Anteils der Binnenschiffahrt an diesem Wert*, 1967

7 *Sonderheft anläßlich des 65. Geburtstages von Prof. Arthur Röhnisch mit Beiträgen von* Benk, Dieter; Breitling, J.; Gurr, Siegfried; Haberhauer, Robert; Honekamp, Hermann; Kuz, Klaus Dieter; Marotz, Günter; Mayer-Vorfelder, Hans-Jörg; Miller, Rudolf; Plate, Erich J.; Radomski, Helge; Schwarz, Helmut; Vollmer, Ernst; Wildenhahn, Eberhard; 1967

8 Jumikis, Alfred: *Beitrag zur experimentellen Untersuchung des Wassernachschubs in einem gefrierenden Boden und die Beurteilung der Ergebnisse*, 1968

9 Marotz, Günter: *Technische Grundlagen einer Wasserspeicherung im natürlichen Untergrund*, 1968

10 Radomski, Helge: *Untersuchungen über den Einfluß der Querschnittsform wellenförmiger Spundwände auf die statischen und rammtechnischen Eigenschaften*, 1968

11 Schwarz, Helmut: *Die Grenztragfähigkeit des Baugrundes bei Einwirkung vertikal gezogener Ankerplatten als zweidimensionales Bruchproblem*, 1969

12 Erbel, Klaus: *Ein Beitrag zur Untersuchung der Metamorphose von Mittelgebirgsschneedecken unter besonderer Berücksichtigung eines Verfahrens zur Bestimmung der thermischen Schneequalität*, 1969

13 Westhaus, Karl-Heinz: *Der Strukturwandel in der Binnenschiffahrt und sein Einfluß auf den Ausbau der Binnenschiffskanäle*, 1969

14 Mayer-Vorfelder, Hans-Jörg: *Ein Beitrag zur Berechnung des Erdwiderstandes unter Ansatz der logarithmischen Spirale als Gleitflächenfunktion*, 1970

15 Schulz, Manfred: *Berechnung des räumlichen Erddruckes auf die Wandung kreiszylindrischer Körper*, 1970

16 Mobasseri, Manoutschehr: *Die Rippenstützmauer. Konstruktion und Grenzen ihrer Standsicherheit*, 1970

17 Benk, Dieter: *Ein Beitrag zum Betrieb und zur Bemessung von Hochwasserrückhaltebecken*, 1970

18 Gàl, Attila: *Bestimmung der mitschwingenden Wassermasse bei überströmten Fischbauchklappen mit kreiszylindrischem Staublech*, 1971, <u>*vergriffen*</u>

19 Kuz, Klaus Dieter: *Ein Beitrag zur Frage des Einsetzens von Kavitationserscheinungen in einer Düsenströmung bei Berücksichtigung der im Wasser gelösten Gase*, 1971, <u>*vergriffen*</u>

20 Schaak, Hartmut: *Verteilleitungen von Wasserkraftanlagen*, 1971

21 *Sonderheft zur Eröffnung der neuen Versuchsanstalt des Instituts für Wasserbau der Universität Stuttgart mit Beiträgen von* Brombach, Hansjörg; Dirksen, Wolfram; Gàl, Attila; Gerlach, Reinhard; Giesecke, Jürgen; Holthoff, Franz-Josef; Kuz, Klaus Dieter; Marotz, Günter; Minor, Hans-Erwin; Petrikat, Kurt; Röhnisch, Arthur; Rueff, Helge; Schwarz, Helmut; Vollmer, Ernst; Wildenhahn, Eberhard; 1972

22 Wang, Chung-su: *Ein Beitrag zur Berechnung der Schwingungen an Kegelstrahlschiebern*, 1972

23 Mayer-Vorfelder, Hans-Jörg: *Erdwiderstandsbeiwerte nach dem Ohde-Variationsverfahren*, 1972

24 Minor, Hans-Erwin: *Beitrag zur Bestimmung der Schwingungsanfachungsfunktionen überströmter Stauklappen*, 1972, <u>*vergriffen*</u>

25 Brombach, Hansjörg: *Untersuchung strömungsmechanischer Elemente (Fluidik) und die Möglichkeit der Anwendung von Wirbelkammerelementen im Wasserbau*, 1972, <u>*vergriffen*</u>

26 Wildenhahn, Eberhard: *Beitrag zur Berechnung von Horizontalfilterbrunnen*, 1972

27 Steinlein, Helmut: *Die Eliminierung der Schwebstoffe aus Flußwasser zum Zweck der unterirdischen Wasserspeicherung, gezeigt am Beispiel der Iller*, 1972

28 Holthoff, Franz Josef: *Die Überwindung großer Hubhöhen in der Binnenschiffahrt durch Schwimmerhebewerke*, 1973

29 Röder, Karl: *Einwirkungen aus Baugrundbewegungen auf trog- und kastenförmige Konstruktionen des Wasser- und Tunnelbaues*, 1973

30 Kretschmer, Heinz: *Die Bemessung von Bogenstaumauern in Abhängigkeit von der Talform*, 1973

31 Honekamp, Hermann: *Beitrag zur Berechnung der Montage von Unterwasserpipelines*, 1973

32 Giesecke, Jürgen: *Die Wirbelkammertriode als neuartiges Steuerorgan im Wasserbau*, und Brombach, Hansjörg: *Entwicklung, Bauformen, Wirkungsweise und Steuereigenschaften von Wirbelkammerverstärkern*, 1974

33 Rueff, Helge: *Untersuchung der schwingungserregenden Kräfte an zwei hintereinander angeordneten Tiefschützen unter besonderer Berücksichtigung von Kavitation*, 1974

34 Röhnisch, Arthur: *Einpreßversuche mit Zementmörtel für Spannbeton - Vergleich der Ergebnisse von Modellversuchen mit Ausführungen in Hüllwellrohren*, 1975

35 *Sonderheft anläßlich des 65. Geburtstages von Prof. Dr.-Ing. Kurt Petrikat mit Beiträgen von*: Brombach, Hansjörg; Erbel, Klaus; Flinspach, Dieter; Fischer jr., Richard; Gàl, Attila; Gerlach, Reinhard; Giesecke, Jürgen; Haberhauer, Robert; Hafner Edzard; Hausenblas, Bernhard; Horlacher, Hans-Burkhard; Hutarew, Andreas; Knoll, Manfred; Krummet, Ralph; Marotz, Günter; Merkle, Theodor; Miller, Christoph; Minor, Hans-Erwin; Neumayer, Hans; Rao, Syamala; Rath, Paul; Rueff, Helge; Ruppert, Jürgen; Schwarz, Wolfgang; Topal-Gökceli, Mehmet; Vollmer, Ernst; Wang, Chung-su; Weber, Hans-Georg; 1975

36 Berger, Jochum: *Beitrag zur Berechnung des Spannungszustandes in rotationssymmetrisch belasteten Kugelschalen veränderlicher Wandstärke unter Gas- und Flüssigkeitsdruck durch Integration schwach singulärer Differentialgleichungen*, 1975

37 Dirksen, Wolfram: *Berechnung instationärer Abflußvorgänge in gestauten Gerinnen mittels Differenzenverfahren und die Anwendung auf Hochwasserrückhaltebecken*, 1976

38 Horlacher, Hans-Burkhard: *Berechnung instationärer Temperatur- und Wärmespannungsfelder in langen mehrschichtigen Hohlzylindern*, 1976

39 Hafner, Edzard: *Untersuchung der hydrodynamischen Kräfte auf Baukörper im Tiefwasserbereich des Meeres*, 1977, ISBN 3-921694-39-6

40 Ruppert, Jürgen: *Über den Axialwirbelkammerverstärker für den Einsatz im Wasserbau*, 1977, ISBN 3-921694-40-X

41 Hutarew, Andreas: *Beitrag zur Beeinflußbarkeit des Sauerstoffgehalts in Fließgewässern an Abstürzen und Wehren*, 1977, ISBN 3-921694-41-8, vergriffen

42 Miller, Christoph: *Ein Beitrag zur Bestimmung der schwingungserregenden Kräfte an unterströmten Wehren*, 1977, ISBN 3-921694-42-6

43 Schwarz, Wolfgang: *Druckstoßberechnung unter Berücksichtigung der Radial- und Längsverschiebungen der Rohrwandung*, 1978, ISBN 3-921694-43-4

44 Kinzelbach, Wolfgang: *Numerische Untersuchungen über den optimalen Einsatz variabler Kühlsysteme einer Kraftwerkskette am Beispiel Oberrhein*, 1978, ISBN 3-921694-44-2

45 Barczewski, Baldur: *Neue Meßmethoden für Wasser-Luftgemische und deren Anwendung auf zweiphasige Auftriebsstrahlen*, 1979, ISBN 3-921694-45-0

46 Neumayer, Hans: *Untersuchung der Strömungsvorgänge in radialen Wirbelkammerverstärkern*, 1979, ISBN 3-921694-46-9

47 Elalfy, Youssef-Elhassan: *Untersuchung der Strömungsvorgänge in Wirbelkammerdioden und -drosseln*, 1979, ISBN 3-921694-47-7

48 Brombach, Hansjörg: *Automatisierung der Bewirtschaftung von Wasserspeichern*, 1981, ISBN 3-921694-48-5

49 Geldner, Peter: *Deterministische und stochastische Methoden zur Bestimmung der Selbstdichtung von Gewässern*, 1981, ISBN 3-921694-49-3, vergriffen

50 Mehlhorn, Hans: *Temperaturveränderungen im Grundwasser durch Brauchwassereinleitungen*, 1982, ISBN 3-921694-50-7, vergriffen

51 Hafner, Edzard: *Rohrleitungen und Behälter im Meer*, 1983, ISBN 3-921694-51-5

52 Rinnert, Bernd: *Hydrodynamische Dispersion in porösen Medien: Einfluß von Dichteunterschieden auf die Vertikalvermischung in horizontaler Strömung*, 1983, ISBN 3-921694-52-3, vergriffen

53 Lindner, Wulf: *Steuerung von Grundwasserentnahmen unter Einhaltung ökologischer Kriterien*, 1983, ISBN 3-921694-53-1, vergriffen

54 Herr, Michael; Herzer, Jörg; Kinzelbach, Wolfgang; Kobus, Helmut; Rinnert, Bernd: *Methoden zur rechnerischen Erfassung und hydraulischen Sanierung von Grundwasserkontaminationen*, 1983, ISBN 3-921694-54-X

55 Schmitt, Paul: *Wege zur Automatisierung der Niederschlagsermittlung*, 1984, ISBN 3-921694-55-8, vergriffen

56 Müller, Peter: *Transport und selektive Sedimentation von Schwebstoffen bei gestautem Abfluß*, 1985, ISBN 3-921694-56-6

57 El-Qawasmeh, Fuad: *Möglichkeiten und Grenzen der Tropfbewässerung unter besonderer Berücksichtigung der Verstopfungsanfälligkeit der Tropfelemente*, 1985, ISBN 3-921694-57-4, vergriffen

58 Kirchenbaur, Klaus: *Mikroprozessorgesteuerte Erfassung instationärer Druckfelder am Beispiel seegangsbelasteter Baukörper*, 1985, ISBN 3-921694-58-2

59 Kobus, Helmut (Hrsg.): *Modellierung des großräumigen Wärme- und Schadstofftransports im Grundwasser*, Tätigkeitsbericht 1984/85 (DFG-Forschergruppe an den Universitäten Hohenheim, Karlsruhe und Stuttgart), 1985, ISBN 3-921694-59-0, vergriffen

60 Spitz, Karlheinz: *Dispersion in porösen Medien: Einfluß von Inhomogenitäten und Dichteunterschieden*, 1985, ISBN 3-921694-60-4, vergriffen

61 Kobus, Helmut: *An Introduction to Air-Water Flows in Hydraulics*, 1985, ISBN 3-921694-61-2

62 Kaleris, Vassilios: *Erfassung des Austausches von Oberflächen- und Grundwasser in horizontalebenen Grundwassermodellen*, 1986, ISBN 3-921694-62-0

63 Herr, Michael: *Grundlagen der hydraulischen Sanierung verunreinigter Porengrundwasserleiter*, 1987, ISBN 3-921694-63-9

64 Marx, Walter: *Berechnung von Temperatur und Spannung in Massenbeton infolge Hydratation*, 1987, ISBN 3-921694-64-7

65 Koschitzky, Hans-Peter: *Dimensionierungskonzept für Sohlbelüfter in Schußrinnen zur Vermeidung von Kavitationsschäden*, 1987, ISBN 3-921694-65-5

66 Kobus, Helmut (Hrsg.): *Modellierung des großräumigen Wärme- und Schadstofftransports im Grundwasser*, Tätigkeitsbericht 1986/87 (DFG-Forschergruppe an den Universitäten Hohenheim, Karlsruhe und Stuttgart) 1987, ISBN 3-921694-66-3

67 Söll, Thomas: *Berechnungsverfahren zur Abschätzung anthropogener Temperaturanomalien im Grundwasser*, 1988, ISBN 3-921694-67-1

68 Dittrich, Andreas; Westrich, Bernd: *Bodenseeufererosion, Bestandsaufnahme und Bewertung*, 1988, ISBN 3-921694-68-X, vergriffen

69 Huwe, Bernd; van der Ploeg, Rienk R.: *Modelle zur Simulation des Stickstoffhaushaltes von Standorten mit unterschiedlicher landwirtschaftlicher Nutzung*, 1988, ISBN 3-921694-69-8, vergriffen

70 Stephan, Karl: *Integration elliptischer Funktionen*, 1988, ISBN 3-921694-70-1

71 Kobus, Helmut; Zilliox, Lothaire (Hrsg.): *Nitratbelastung des Grundwassers, Auswirkungen der Landwirtschaft auf die Grundwasser- und Rohwasserbeschaffenheit und Maßnahmen zum Schutz des Grundwassers*. Vorträge des deutsch-französischen Kolloquiums am 6. Oktober 1988, Universitäten Stuttgart und Louis Pasteur Strasbourg (Vorträge in deutsch oder französisch, Kurzfassungen zweisprachig), 1988, ISBN 3-921694-71-X

72 Soyeaux, Renald: *Unterströmung von Stauanlagen auf klüftigem Untergrund unter Berücksichtigung laminarer und turbulenter Fließzustände*,1991, ISBN 3-921694-72-8

73 Kohane, Roberto: *Berechnungsmethoden für Hochwasserabfluß in Fließgewässern mit übeströmten Vorländern*, 1991, ISBN 3-921694-73-6

74 Hassinger, Reinhard: *Beitrag zur Hydraulik und Bemessung von Blocksteinrampen in flexibler Bauweise*, 1991, ISBN 3-921694-74-4, *vergriffen*

75 Schäfer, Gerhard: *Einfluß von Schichtenstrukturen und lokalen Einlagerungen auf die Längsdispersion in Porengrundwasserleitern*, 1991, ISBN 3-921694-75-2

76 Giesecke, Jürgen: *Vorträge, Wasserwirtschaft in stark besiedelten Regionen; Umweltforschung mit Schwerpunkt Wasserwirtschaft*, 1991, ISBN 3-921694-76-0

77 Huwe, Bernd: *Deterministische und stochastische Ansätze zur Modellierung des Stickstoffhaushalts landwirtschaftlich genutzter Flächen auf unterschiedlichem Skalenniveau*, 1992, ISBN 3-921694-77-9, *vergriffen*

78 Rommel, Michael: *Verwendung von Kluftdaten zur realitätsnahen Generierung von Kluftnetzen mit anschließender laminar-turbulenter Strömungsberechnung*, 1993, ISBN 3-92 1694-78-7

79 Marschall, Paul: *Die Ermittlung lokaler Stofffrachten im Grundwasser mit Hilfe von Einbohrloch-Meßverfahren*, 1993, ISBN 3-921694-79-5, *vergriffen*

80 Ptak, Thomas: *Stofftransport in heterogenen Porenaquiferen: Felduntersuchungen und stochastische Modellierung*, 1993, ISBN 3-921694-80-9, *vergriffen*

81 Haakh, Frieder: *Transientes Strömungsverhalten in Wirbelkammern,* 1993, ISBN 3-921694-81-7

82 Kobus, Helmut; Cirpka, Olaf; Barczewski, Baldur; Koschitzky, Hans-Peter: *Versuchseinrichtung zur Grundwasser- und Altlastensanierung VEGAS, Konzeption und Programmrahmen*, 1993, ISBN 3-921694-82-5

83 Zang, Weidong: *Optimaler Echtzeit-Betrieb eines Speichers mit aktueller Abflußregenerierung*, 1994, ISBN 3-921694-83-3, *vergriffen*

84 Franke, Hans-Jörg: *Stochastische Modellierung eines flächenhaften Stoffeintrages und Transports in Grundwasser am Beispiel der Pflanzenschutzmittelproblematik*, 1995, ISBN 3-921694-84-1

85 Lang, Ulrich: *Simulation regionaler Strömungs- und Transportvorgänge in Karstaquiferen mit Hilfe des Doppelkontinuum-Ansatzes: Methodenentwicklung und Parameteridenti-fikation*, 1995, ISBN 3-921694-85-X, *vergriffen*

86 Helmig, Rainer: *Einführung in die Numerischen Methoden der Hydromechanik*, 1996, ISBN 3-921694-86-8, *vergriffen*

87 Cirpka, Olaf: *CONTRACT: A Numerical Tool for Contaminant Transport and Chemical Transformations - Theory and Program Documentation -*, 1996, ISBN 3-921694-87-6

88 Haberlandt, Uwe: *Stochastische Synthese und Regionalisierung des Niederschlages für Schmutzfrachtberechnungen*, 1996, ISBN 3-921694-88-4

89 Croisé, Jean: *Extraktion von flüchtigen Chemikalien aus natürlichen Lockergesteinen mittels erzwungener Luftströmung*, 1996, ISBN 3-921694-89-2, *vergriffen*

90 Jorde, Klaus: *Ökologisch begründete, dynamische Mindestwasserregelungen bei Ausleitungskraftwerken*, 1997, ISBN 3-921694-90-6, *vergriffen*

91 Helmig, Rainer: *Gekoppelte Strömungs- und Transportprozesse im Untergrund - Ein Beitrag zur Hydrosystemmodellierung-*, 1998, ISBN 3-921694-91-4, *vergriffen*

92 Emmert, Martin: *Numerische Modellierung nichtisothermer Gas-Wasser Systeme in porösen Medien,* 1997, ISBN 3-921694-92-2

93 Kern, Ulrich: *Transport von Schweb- und Schadstoffen in staugeregelten Fließgewässern am Beispiel des Neckars*, 1997, ISBN 3-921694-93-0, *vergriffen*

94 Förster, Georg: *Druckstoßdämpfung durch große Luftblasen in Hochpunkten von Rohrleitungen 1997*, ISBN 3-921694-94-9

95 Cirpka, Olaf: *Numerische Methoden zur Simulation des reaktiven Mehrkomponententransports im Grundwasser*, 1997, ISBN 3-921694-95-7, vergriffen

96 Färber, Arne: *Wärmetransport in der ungesättigten Bodenzone: Entwicklung einer thermischen In-situ-Sanierungstechnologie*, 1997, ISBN 3-921694-96-5

97 Betz, Christoph: *Wasserdampfdestillation von Schadstoffen im porösen Medium: Entwicklung einer thermischen In-situ-Sanierungstechnologie*, 1998, SBN 3-921694-97-3

98 Xu, Yichun: *Numerical Modeling of Suspended Sediment Transport in Rivers*, 1998, ISBN 3-921694-98-1, vergriffen

99 Wüst, Wolfgang: *Geochemische Untersuchungen zur Sanierung CKW-kontaminierter Aquifere mit Fe(0)-Reaktionswänden,* 2000, ISBN 3-933761-02-2

100 Sheta, Hussam: *Simulation von Mehrphasenvorgängen in porösen Medien unter Einbeziehung von Hysterese-Effekten*, 2000, ISBN 3-933761-03-4

101 Ayros, Edwin: *Regionalisierung extremer Abflüsse auf der Grundlage statistischer Verfahren*, 2000, ISBN 3-933761-04-2, vergriffen

102 Huber, Ralf: *Compositional Multiphase Flow and Transport in Heterogeneous Porous Media*, 2000, ISBN 3-933761-05-0

103 Braun, Christopherus: *Ein Upscaling-Verfahren für Mehrphasenströmungen in porösen Medien*, 2000, ISBN 3-933761-06-9

104 Hofmann, Bernd: *Entwicklung eines rechnergestützten Managementsystems zur Beurteilung von Grundwasserschadensfällen*, 2000, ISBN 3-933761-07-7

105 Class, Holger: *Theorie und numerische Modellierung nichtisothermer Mehrphasenprozesse in NAPL-kontaminierten porösen Medien,* 2001, ISBN 3-933761-08-5

106 Schmidt, Reinhard: *Wasserdampf- und Heißluftinjektion zur thermischen Sanierung kontaminierter Standorte*, 2001, ISBN 3-933761-09-3

107 Josef, Reinhold: *Schadstoffextraktion mit hydraulischen Sanierungsverfahren unter Anwendung von grenzflächenaktiven Stoffen*, 2001, ISBN 3-933761-10-7

108 Schneider, Matthias: *Habitat- und Abflussmodellierung für Fließgewässer mit unscharfen Berechnungsansätzen*, 2001, ISBN 3-933761-11-5

109 Rathgeb, Andreas: *Hydrodynamische Bemessungsgrundlagen für Lockerdeckwerke an überströmbaren Erddämmen*, 2001, ISBN 3-933761-12-3

110 Lang, Stefan: *Parallele numerische Simulation instätionärer Probleme mit adaptiven Methoden auf unstrukturierten Gittern*, 2001, ISBN 3-933761-13-1

111 Appt, Jochen; Stumpp Simone: *Die Bodensee-Messkampagne 2001, IWS/CWR Lake Constance Measurement Program 2001*, 2002, ISBN 3-933761-14-X

112 Heimerl, Stephan: *Systematische Beurteilung von Wasserkraftprojekten,* 2002,
ISBN 3-933761-15-8, vergriffen

113 Iqbal, Amin: *On the Management and Salinity Control of Drip Irrigation,* 2002,
ISBN 3-933761-16-6

114 Silberhorn-Hemminger, Annette: *Modellierung von Kluftaquifersystemen: Geostatistische Analyse und deterministisch-stochastische Kluftgenerierung,* 2002, ISBN 3-933761-17-4

115 Winkler, Angela: *Prozesse des Wärme- und Stofftransports bei der In-situ-Sanierung mit festen Wärmequellen,* 2003, ISBN 3-933761-18-2

116 Marx, Walter: *Wasserkraft, Bewässerung, Umwelt - Planungs- und Bewertungsschwerpunkte der Wasserbewirtschaftung,* 2003, ISBN 3-933761-19-0

117 Hinkelmann, Reinhard: *Efficient Numerical Methods and Information-Processing Techniques in Environment Water,* 2003, ISBN 3-933761-20-4

118 Samaniego-Eguiguren, Luis Eduardo: *Hydrological Consequences of Land Use / Land Cover and Climatic Changes in Mesoscale Catchments*, 2003, ISBN 3-933761-21-2

119 Neunhäuserer, Lina: *Diskretisierungsansätze zur Modellierung von Strömungs- und Transportprozessen in geklüftet-porösen Medien*, 2003, ISBN 3-933761-22-0

120 Paul, Maren: *Simulation of Two-Phase Flow in Heterogeneous Poros Media with Adaptive Methods*, 2003, ISBN 3-933761-23-9

121 Ehret, Uwe: *Rainfall and Flood Nowcasting in Small Catchments using Weather Radar*, 2003, ISBN 3-933761-24-7

122 Haag, Ingo: *Der Sauerstoffhaushalt staugeregelter Flüsse am Beispiel des Neckars - Analysen, Experimente, Simulationen -,* 2003, ISBN 3-933761-25-5

123 Appt, Jochen: *Analysis of Basin-Scale Internal Waves in Upper Lake Constance,* 2003, ISBN 3-933761-26-3

124 Hrsg.: Schrenk, Volker; Batereau, Katrin; Barczewski, Baldur; Weber, Karolin und Koschitzky, Hans-Peter: *Symposium Ressource Fläche und VEGAS - Statuskolloquium 2003, 30. September und 1. Oktober 2003,* 2003, ISBN 3-933761-27-1

125 Omar Khalil Ouda: *Optimisation of Agricultural Water Use: A Decision Support System for the Gaza Strip,* 2003, ISBN 3-933761-28-0

126 Batereau, Katrin: *Sensorbasierte Bodenluftmessung zur Vor-Ort-Erkundung von Schadensherden im Untergrund*, 2004, ISBN 3-933761-29-8

127 Witt, Oliver: *Erosionsstabilität von Gewässersedimenten mit Auswirkung auf den Stofftransport bei Hochwasser am Beispiel ausgewählter Stauhaltungen des Oberrheins*, 2004, ISBN 3-933761-30-1

128 Jakobs, Hartmut: *Simulation nicht-isothermer Gas-Wasser-Prozesse in komplexen Kluft-Matrix-Systemen*, 2004, ISBN 3-933761-31-X

129 Li, Chen-Chien: *Deterministisch-stochastisches Berechnungskonzept zur Beurteilung der Auswirkungen erosiver Hochwasserereignisse in Flussstauhaltungen,* 2004, ISBN 3-933761-32-8

130 Reichenberger, Volker; Helmig, Rainer; Jakobs, Hartmut; Bastian, Peter; Niessner, Jennifer: *Complex Gas-Water Processes in Discrete Fracture-Matrix Systems: Up-scaling, Mass-Conservative Discretization and Efficient Multilevel Solution,* 2004, ISBN 3-933761-33-6

131 Hrsg.: Barczewski, Baldur; Koschitzky, Hans-Peter; Weber, Karolin; Wege, Ralf: *VEGAS - Statuskolloquium 2004,* Tagungsband zur Veranstaltung am 05. Oktober 2004 an der Universität Stuttgart, Campus Stuttgart-Vaihingen, 2004, ISBN 3-933761-34-4

132 Asie, Kemal Jabir: *Finite Volume Models for Multiphase Multicomponent Flow through Porous Media*. 2005, ISBN 3-933761-35-2

133 Jacoub, George: *Development of a 2-D Numerical Module for Particulate Contaminant Transport in Flood Retention Reservoirs and Impounded Rivers,* 2004, ISBN 3-933761-36-0

134 Nowak, Wolfgang: *Geostatistical Methods for the Identification of Flow and Transport Parameters in the Subsurface,* 2005, ISBN 3-933761-37-9

135 Süß, Mia: *Analysis of the influence of structures and boundaries on flow and transport processes in fractured porous media*, 2005, ISBN 3-933761-38-7

136 Jose, Surabhin Chackiath: *Experimental Investigations on Longitudinal Dispersive Mixing in Heterogeneous Aquifers,* 2005, ISBN: 3-933761-39-5

137 Filiz, Fulya: *Linking Large-Scale Meteorological Conditions to Floods in Mesoscale Catchments,* 2005, ISBN 3-933761-40-9

138 Qin, Minghao: *Wirklichkeitsnahe und recheneffiziente Ermittlung von Temperatur und Spannungen bei großen RCC-Staumauern*, 2005, ISBN 3-933761-41-7

139 Kobayashi, Kenichiro: *Optimization Methods for Multiphase Systems in the Subsurface - Application to Methane Migration in Coal Mining Areas*, 2005, ISBN 3-933761-42-5

140 Rahman, Md. Arifur: *Experimental Investigations on Transverse Dispersive Mixing in Heterogeneous Porous Media,* 2005, ISBN 3-933761-43-3

141 Schrenk, Volker: *Ökobilanzen zur Bewertung von Altlastensanierungsmaßnahmen,* 2005, ISBN 3-933761-44-1

142 Hundecha, Hirpa Yeshewatesfa: *Regionalization of Parameters of a Conceptual Rainfall-Runoff Model*, 2005, ISBN: 3-933761-45-X

143 Wege, Ralf: *Untersuchungs- und Überwachungsmethoden für die Beurteilung natürlicher Selbstreinigungsprozesse im Grundwasser,* 2005, ISBN 3-933761-46-8

144 Breiting, Thomas: *Techniken und Methoden der Hydroinformatik - Modellierung von komplexen Hydrosystemen im Untergrund*, 2006, ISBN 3-933761-47-6

145 Hrsg.: Braun, Jürgen; Koschitzky, Hans-Peter; Müller, Martin: *Ressource Untergrund: 10 Jahre VEGAS: Forschung und Technologieentwicklung zum Schutz von Grundwasser und Boden,* Tagungsband zur Veranstaltung am 28. und 29. September 2005 an der Universität Stuttgart, Campus Stuttgart-Vaihingen, 2005, ISBN 3-933761-48-4

146 Rojanschi, Vlad: *Abflusskonzentration in mesoskaligen Einzugsgebieten unter Berücksichtigung des Sickerraumes,* 2006, ISBN 3-933761-49-2

147 Winkler, Nina Simone: *Optimierung der Steuerung von Hochwasserrückhaltebeckensystemen,* 2006, ISBN 3-933761-50-6

148 Wolf, Jens: *Räumlich differenzierte Modellierung der Grundwasserströmung alluvialer Aquifere für mesoskalige Einzugsgebiete*, 2006, ISBN: 3-933761-51-4

149 Kohler, Beate: *Externe Effekte der Laufwasserkraftnutzung*, 2006, ISBN 3-933761-52-2

150 Hrsg.: Braun, Jürgen; Koschitzky, Hans-Peter; Stuhrmann, Matthias: *VEGAS-Statuskolloquium 2006,* Tagungsband zur Veranstaltung am 28. September 2006 an der Universität Stuttgart, Campus Stuttgart-Vaihingen, 2006, ISBN 3-933761-53-0

151 Niessner, Jennifer: *Multi-Scale Modeling of Multi-Phase - Multi-Component Processes in Heterogeneous Porous Media*, 2006, ISBN 3-933761-54-9

152 Fischer, Markus: *Beanspruchung eingeerdeter Rohrleitungen infolge Austrocknung bindiger Böden,* 2006, ISBN 3-933761-55-7

153 Schneck, Alexander: *Optimierung der Grundwasserbewirtschaftung unter Berücksichtigung der Belange der Wasserversorgung, der Landwirtschaft und des Naturschutzes*, 2006, ISBN 3-933761-56-5

154 Das, Tapash: *The Impact of Spatial Variability of Precipitation on the Predictive Uncertainty of Hydrological Models,* 2006, ISBN 3-33761-57-3

155 Bielinski, Andreas: *Numerical Simulation of* CO_2 *sequestration in geological formations*, 2007, ISBN 3-933761-58-1

156 Mödinger, Jens: *Entwicklung eines Bewertungs- und Entscheidungsunterstützungssystems für eine nachhaltige regionale Grundwasserbewirtschaftung*, 2006, ISBN 3-933761-60-3

157 Manthey, Sabine: *Two-phase flow processes with dynamic effects in porous media - parameter estimation and simulation,* 2007, ISBN 3-933761-61-1

158 Pozos Estrada, Oscar: *Investigation on the Effects of Entrained Air in Pipelines,* 2007, ISBN 3-933761-62-X

159 Ochs, Steffen Oliver: *Steam injection into saturated porous media – process analysis including experimental and numerical investigations*, 2007, ISBN 3-933761-63-8

160 Marx, Andreas: *Einsatz gekoppelter Modelle und Wetterradar zur Abschätzung von Niederschlagsintensitäten und zur Abflussvorhersage, 2007,* ISBN 3-933761-64-6

161 Hartmann, Gabriele Maria: *Investigation of Evapotranspiration Concepts in Hydrological Modelling for Climate Change Impact Assessment*, 2007, ISBN 3-933761-65-4

162 Kebede Gurmessa, Tesfaye: *Numerical Investigation on Flow and Transport Characteristics to Improve Long-Term Simulation of Reservoir Sedimentation*, 2007, ISBN 3-933761-66-2

163 Trifković, Aleksandar: *Multi-objective and Risk-based Modelling Methodology for Planning, Design and Operation of Water Supply Systems*, 2007, ISBN 3-933761-67-0

164 Götzinger, Jens: *Distributed Conceptual Hydrological Modelling - Simulation of Climate, Land Use Change Impact and Uncertainty Analysis*, 2007, ISBN 3-933761-68-9

165 Hrsg.: Braun, Jürgen; Koschitzky, Hans-Peter; Stuhrmann, Matthias: *VEGAS – Kolloquium 2007,* Tagungsband zur Veranstaltung am 26. September 2007 an der Universität Stuttgart, Campus Stuttgart-Vaihingen, 2007, ISBN 3-933761-69-7

166 Freeman, Beau: *Modernization Criteria Assessment for Water Resources Planning; Klamath Irrigation Project, U.S.*, 2008, ISBN 3-933761-70-0

167 Dreher, Thomas: *Selektive Sedimentation von Feinstschwebstoffen in Wechselwirkung mit wandnahen turbulenten Strömungsbedingungen*, 2008, ISBN 3-933761-71-9

168 Yang, Wei: *Discrete-Continuous Downscaling Model for Generating Daily Precipitation Time Series*, 2008, ISBN 3-933761-72-7

169 Kopecki, Ianina: *Calculational Approach to FST-Hemispheres for Multiparametrical Benthos Habitat Modelling,* 2008*,* ISBN 3-933761-73-5

170 Brommundt, Jürgen: *Stochastische Generierung räumlich zusammenhängender Niederschlagszeitreihen*, 2008, ISBN 3-933761-74-3

171 Papafotiou, *Alexandros: Numerical Investigations of the Role of Hysteresis in Heterogeneous Two-Phase Flow Systems*, 2008, ISBN 3-933761-75-1

172 He, Yi: *Application of a Non-Parametric Classification Scheme to Catchment Hydrology*, 2008, ISBN 978-3-933761-76-7

173 Wagner, Sven: *Water Balance in a Poorly Gauged Basin in West Africa Using Atmospheric Modelling and Remote Sensing Information*, 2008, ISBN 978-3-933761-77-4

174 Hrsg.: Braun, Jürgen; Koschitzky, Hans-Peter; Stuhrmann, Matthias; Schrenk, Volker: *VEGAS-Kolloquium 2008 Ressource Fläche III*, Tagungsband zur Veranstaltung am 01. Oktober 2008 an der Universität Stuttgart, Campus Stuttgart-Vaihingen, 2008, ISBN 978-3-933761-78-1

175 Patil, Sachin: *Regionalization of an Event Based Nash Cascade Model for Flood Predictions in Ungauged Basins*, 2008, ISBN 978-3-933761-79-8

176 Assteerawatt, Anongnart: *Flow and Transport Modelling of Fractured Aquifers based on a Geostatistical Approach,* 2008, ISBN 978-3-933761-80-4

177 Karnahl, Joachim Alexander: *2D numerische Modellierung von multifraktionalem Schwebstoff- und Schadstofftransport in Flüssen*, 2008, ISBN 978-3-933761-81-1

178 Hiester, Uwe: *Technologieentwicklung zur In-situ-Sanierung der ungesättigten Bodenzone mit festen Wärmequellen,* 2009, ISBN 978-3-933761-82-8

179 Laux, Patrick: *Statistical Modeling of Precipitation for Agricultural Planning in the Volta Basin of West Africa,* 2009, ISBN 978-3-933761-83-5

180 Ehsan, Saqib: *Evaluation of Life Safety Risks Related to Severe Flooding,* 2009, ISBN 978-3-933761-84-2

181 Prohaska, Sandra: *Development and Application of a 1D Multi-Strip Fine Sediment Transport Model for Regulated Rivers,* 2009, ISBN 978-3-933761-85-9

182 Kopp, Andreas: *Evaluation of* CO_2 *Injection Processes in Geological Formations for Site Screening,* 2009, ISBN 978-3-933761-86-6

183 Ebigbo, Anozie: *Modelling of biofilm growth and its influence on* CO_2 *and water (two-phase) flow in porous media,* 2009, ISBN 978-3-933761-87-3

184 Freiboth, Sandra: *A phenomenological model for the numerical simulation of multiphase multicomponent processes considering structural alterations of porous media*, 2009, ISBN 978-3-933761-88-0

185 Zöllner, Frank: *Implementierung und Anwendung netzfreier Methoden im Konstruktiven Wasserbau und in der Hydromechanik*, 2009, ISBN 978-3-933761-89-7

186 Vasin, Milos: *Influence of the soil structure and property contrast on flow and transport in the unsaturated zone*, 2010, ISBN 978-3-933761-90-3

187 Li, Jing: *Application of Copulas as a New Geostatistical Tool*, 2010, ISBN 978-3-933761-91-0

188 AghaKouchak, Amir: *Simulation of Remotely Sensed Rainfall Fields Using Copulas*, 2010, ISBN 978-3-933761-92-7

189 Thapa, Pawan Kumar: *Physically-based spatially distributed rainfall runoff modelling for soil erosion estimation*, 2010, ISBN 978-3-933761-93-4

190 Wurms, Sven: *Numerische Modellierung der Sedimentationsprozesse in Retentionsanlagen zur Steuerung von Stoffströmen bei extremen Hochwasserabflussereignissen*, 2011, ISBN 978-3-933761-94-1

191 Merkel, Uwe: *Unsicherheitsanalyse hydraulischer Einwirkungen auf Hochwasserschutzdeiche und Steigerung der Leistungsfähigkeit durch adaptive Strömungsmodellierung*, 2011, ISBN 978-3-933761-95-8

192 Fritz, Jochen: *A Decoupled Model for Compositional Non-Isothermal Multiphase Flow in Porous Media and Multiphysics Approaches for Two-Phase Flow*, 2010, ISBN 978-3-933761-96-5

193 Weber, Karolin (Hrsg.): *12. Treffen junger WissenschaftlerInnen an Wasserbauinstituten*, 2010, ISBN 978-3-933761-97-2

194 Bliefernicht, Jan-Geert: *Probability Forecasts of Daily Areal Precipitation for Small River Basins,* 2011, ISBN 978-3-933761-98-9

195 Hrsg.: Koschitzky, Hans-Peter; Braun, Jürgen: *VEGAS-Kolloquium 2010 In-situ-Sanierung - Stand und Entwicklung Nano und ISCO -,* Tagungsband zur Veranstaltung am 07. Oktober 2010 an der Universität Stuttgart, Campus Stuttgart-Vaihingen, 2010, ISBN 978-3-933761-99-6

196 Gafurov, Abror: *Water Balance Modeling Using Remote Sensing Information - Focus on Central Asia,* 2010, ISBN 978-3-942036-00-9

197 Mackenberg, Sylvia: *Die Quellstärke in der Sickerwasserprognose: Möglichkeiten und Grenzen von Labor- und Freilanduntersuchungen,* 2010, ISBN 978-3-942036-01-6

198 Singh, Shailesh Kumar: *Robust Parameter Estimation in Gauged and Ungauged Basins,* 2010, ISBN 978-3-942036-02-3

199 Doğan, Mehmet Onur: *Coupling of porous media flow with pipe flow,* 2011, ISBN 978-3-942036-03-0

200 Liu, Min: *Study of Topographic Effects on Hydrological Patterns and the Implication on Hydrological Modeling and Data Interpolation,* 2011, ISBN 978-3-942036-04-7

201 Geleta, Habtamu Itefa: *Watershed Sediment Yield Modeling for Data Scarce Areas,* 2011, ISBN 978-3-942036-05-4

202 Franke, Jörg: *Einfluss der Überwachung auf die Versagenswahrscheinlichkeit von Staustufen,* 2011, ISBN 978-3-942036-06-1

203 Bakimchandra, Oinam: *Integrated Fuzzy-GIS approach for assessing regional soil erosion risks,* 2011, ISBN 978-3-942036-07-8

204 Alam, Muhammad Mahboob: *Statistical Downscaling of Extremes of Precipitation in Mesoscale Catchments from Different RCMs and Their Effects on Local Hydrology,* 2011, ISBN 978-3-942036-08-5

205 Hrsg.: Koschitzky, Hans-Peter; Braun, Jürgen: *VEGAS-Kolloquium 2011 Flache Geothermie - Perspektiven und Risiken,* Tagungsband zur Veranstaltung am 06. Oktober 2011 an der Universität Stuttgart, Campus Stuttgart-Vaihingen, 2011, ISBN 978-3-933761-09-2

206 Haslauer, Claus: *Analysis of Real-World Spatial Dependence of Subsurface Hydraulic Properties Using Copulas with a Focus on Solute Transport Behaviour,* 2011, ISBN 978-3-942036-10-8

207 Dung, Nguyen Viet: *Multi-objective automatic calibration of hydrodynamic models – development of the concept and an application in the Mekong Delta,* 2011, ISBN 978-3-942036-11-5

208 Hung, Nguyen Nghia: *Sediment dynamics in the floodplain of the Mekong Delta, Vietnam,* 2011, ISBN 978-3-942036-12-2

209 Kuhlmann, Anna: *Influence of soil structure and root water uptake on flow in the unsaturated zone,* 2012, ISBN 978-3-942036-13-9

210 Tuhtan, Jeffrey Andrew: *Including the Second Law Inequality in Aquatic Ecodynamics: A Modeling Approach for Alpine Rivers Impacted by Hydropeaking,* 2012, ISBN 978-3-942036-14-6

211 Tolossa, Habtamu: *Sediment Transport Computation Using a Data-Driven Adaptive Neuro-Fuzzy Modelling Approach,* 2012, ISBN 978-3-942036-15-3

212 Tatomir, Alexandru-Bodgan: *From Discrete to Continuum Concepts of Flow in Fractured Porous Media,* 2012, ISBN 978-3-942036-16-0

213 Erbertseder, Karin: *A Multi-Scale Model for Describing Cancer-Therapeutic Transport in the Human Lung,* 2012, ISBN 978-3-942036-17-7

214 Noack, Markus: *Modelling Approach for Interstitial Sediment Dynamics and Reproduction of Gravel Spawning Fish,* 2012, ISBN 978-3-942036-18-4

215 De Boer, Cjestmir Volkert: *Transport of Nano Sized Zero Valent Iron Colloids during Injection into the Subsurface*, 2012, ISBN 978-3-942036-19-1

216 Pfaff, Thomas: *Processing and Analysis of Weather Radar Data for Use in Hydrology*, 2013, ISBN 978-3-942036-20-7

217 Lebrenz, Hans-Henning: *Addressing the Input Uncertainty for Hydrological Modeling by a New Geostatistical Method*, 2013, ISBN 978-3-942036-21-4

218 Darcis, Melanie Yvonne: *Coupling Models of Different Complexity for the Simulation of* CO_2 *Storage in Deep Saline Aquifers*, 2013, ISBN 978-3-942036-22-1

219 Beck, Ferdinand: *Generation of Spatially Correlated Synthetic Rainfall Time Series in High Temporal Resolution - A Data Driven Approach*, 2013, ISBN 978-3-942036-23-8

220 Guthke, Philipp: *Non-multi-Gaussian spatial structures: Process-driven natural genesis, manifestation, modeling approaches, and influences on dependent processes*, 2013, ISBN 978-3-942036-24-5

221 Walter, Lena: *Uncertainty studies and risk assessment for* CO_2 *storage in geological formations*, 2013, ISBN 978-3-942036-25-2

222 Wolff, Markus: *Multi-scale modeling of two-phase flow in porous media including capillary pressure effects,* 2013, ISBN 978-3-942036-26-9

223 Mosthaf, Klaus Roland: *Modeling and analysis of coupled porous-medium and free flow with application to evaporation processes,* 2014, ISBN 978-3-942036-27-6

224 Leube, Philipp Christoph: *Methods for Physically-Based Model Reduction in Time: Analysis, Comparison of Methods and Application*, 2013, ISBN 978-3-942036-28-3

225 Rodríguez Fernández, Jhan Ignacio: *High Order Interactions among environmental variables: Diagnostics and initial steps towards modeling*, 2013, ISBN 978-3-942036-29-0

226 Eder, Maria Magdalena: *Climate Sensitivity of a Large Lake*, 2013, ISBN 978-3-942036-30-6

227 Greiner, Philipp: *Alkoholinjektion zur In-situ-Sanierung von CKW Schadensherden in Grundwasserleitern: Charakterisierung der relevanten Prozesse auf unterschiedlichen Skalen*, 2014, ISBN 978-3-942036-31-3

228 Lauser, Andreas: *Theory and Numerical Applications of Compositional Multi-Phase Flow in Porous Media*, 2014, ISBN 978-3-942036-32-0

229 Enzenhöfer, Rainer: *Risk Quantification and Management in Water Production and Supply Systems*, 2014, ISBN 978-3-942036-33-7

230 Faigle, Benjamin: *Adaptive modelling of compositional multi-phase flow with capillary pressure*, 2014, ISBN 978-3-942036-34-4

231 Oladyshkin, Sergey: *Efficient modeling of environmental systems in the face of complexity and uncertainty*, 2014, ISBN 978-3-942036-35-1

232 Sugimoto, Takayuki: *Copula based Stochastic Analysis of Discharge Time Series*, 2014, ISBN 978-3-942036-36-8

233 Koch, Jonas: *Simulation, Identification and Characterization of Contaminant Source Architectures in the Subsurface*, 2014, ISBN 978-3-942036-37-5

234 Zhang, Jin: *Investigations on Urban River Regulation and Ecological Rehabilitation Measures, Case of Shenzhen in China*, 2014, ISBN 978-3-942036-38-2

235 Siebel, Rüdiger: *Experimentelle Untersuchungen zur hydrodynamischen Belastung und Standsicherheit von Deckwerken an überströmbaren Erddämmen,* 2014, ISBN 978-3-942036-39-9

236 Baber, Katherina: *Coupling free flow and flow in porous media in biological and technical applications: From a simple to a complex interface description,* 2014, ISBN 978-3-942036-40-5

237 Nuske, Klaus Philipp: *Beyond Local Equilibrium — Relaxing local equilibrium assumptions in multiphase flow in porous media,* 2014, ISBN 978-3-942036-41-2

238 Geiges, Andreas: *Efficient concepts for optimal experimental design in nonlinear environmental systems*, 2014, ISBN 978-3-942036-42-9

239 Schwenck, Nicolas: *An XFEM-Based Model for Fluid Flow in Fractured Porous Media*, 2014, ISBN 978-3-942036-43-6

240 Chamorro Chávez, Alejandro: *Stochastic and hydrological modelling for climate change prediction in the Lima region, Peru*, 2015, ISBN 978-3-942036-44-3

241 Yulizar: *Investigation of Changes in Hydro-Meteorological Time Series Using a Depth-Based Approach*, 2015, ISBN 978-3-942036-45-0

242 Kretschmer, Nicole: *Impacts of the existing water allocation scheme on the Limarí watershed – Chile, an integrative approach*, 2015, ISBN 978-3-942036-46-7

243 Kramer, Matthias: *Luftbedarf von Freistrahlturbinen im Gegendruckbetrieb*, 2015, ISBN 978-3-942036-47-4

244 Hommel, Johannes: *Modeling biogeochemical and mass transport processes in the subsurface: Investigation of microbially induced calcite precipitation*, 2016, ISBN 978-3-942036-48-1

245 Germer, Kai: *Wasserinfiltration in die ungesättigte Zone eines makroporösen Hanges und deren Einfluss auf die Hangstabilität*, 2016, ISBN 978-3-942036-49-8

246 Hörning, Sebastian: *Process-oriented modeling of spatial random fields using copulas*, 2016, ISBN 978-3-942036-50-4

247 Jambhekar, Vishal: *Numerical modeling and analysis of evaporative salinization in a coupled free-flow porous-media system,* 2016, ISBN 978-3-942036-51-1

248 Huang, Yingchun: *Study on the spatial and temporal transferability of conceptual hydrological models*, 2016, ISBN 978-3-942036-52-8

249 Kleinknecht, Simon Matthias: *Migration and retention of a heavy NAPL vapor and remediation of the unsaturated zone*, 2016, ISBN 978-3-942036-53-5

250 Kwakye, Stephen Oppong: *Study on the effects of climate change on the hydrology of the West African sub-region*, 2016, ISBN 978-3-942036-54-2

251 Kissinger, Alexander: *Basin-Scale Site Screening and Investigation of Possible Impacts of CO_2 Storage on Subsurface Hydrosystems*, 2016, ISBN 978-3-942036-55-9

252 Müller, Thomas: *Generation of a Realistic Temporal Structure of Synthetic Precipitation Time Series for Sewer Applications*, 2017, ISBN 978-3-942036-56-6

253 Grüninger, Christoph: *Numerical Coupling of Navier-Stokes and Darcy Flow for Soil-Water Evaporation*, 2017, ISBN 978-3-942036-57-3

254 Suroso: *Asymmetric Dependence Based Spatial Copula Models: Empirical Investigations and Consequences on Precipitation Fields*, 2017, ISBN 978-3-942036-58-0

255 Müller, Thomas; Mosthaf, Tobias; Gunzenhauser, Sarah; Seidel, Jochen; Bárdossy, András: *Grundlagenbericht Niederschlags-Simulator (NiedSim3)*, 2017, ISBN 978-3-942036-59-7

256 Mosthaf, Tobias: *New Concepts for Regionalizing Temporal Distributions of Precipitation and for its Application in Spatial Rainfall Simulation,* 2017, ISBN 978-3-942036-60-3

257 Fenrich, Eva Katrin: *Entwicklung eines ökologisch-ökonomischen Vernetzungsmodells für Wasserkraftanlagen und Mehrzweckspeicher,* 2018, ISBN 978-3-942036-61-0

258 Schmidt, Holger: *Microbial stabilization of lotic fine sediments*, 2018, ISBN 978-3-942036-62-7

259 Fetzer, Thomas: *Coupled Free and Porous-Medium Flow Processes Affected by Turbulence and Roughness–Models, Concepts and Analysis*, 2018, ISBN 978-3-942036-63-4

260 Schröder, Hans Christoph: *Large-scale High Head Pico Hydropower Potential Assessment*, 2018, ISBN 978-3-942036-64-1

261 Bode, Felix: *Early-Warning Monitoring Systems for Improved Drinking Water Resource Protection*, 2018, ISBN 978-3-942036-65-8

262 Gebler, Tobias: *Statistische Auswertung von simulierten Talsperrenüberwachungsdaten zur Identifikation von Schadensprozessen an Gewichtsstaumauern*, 2018, ISBN 978-3-942036-66-5

263 Harten, Matthias von: *Analyse des Zuppinger-Wasserrades – Hydraulische Optimierungen unter Berücksichtigung ökologischer Aspekte*, 2018, ISBN 978-3-942036-67-2

264 Yan, Jieru: *Nonlinear estimation of short time precipitation using weather radar and surface observations*, 2018, ISBN 978-3-942036-68-9

265 Beck, Martin: *Conceptual approaches for the analysis of coupled hydraulic and geomechanical processes*, 2019, ISBN 978-3-942036-69-6

266 Haas, Jannik: *Optimal planning of hydropower and energy storage technologies for fully renewable power systems*, 2019, ISBN 978-3-942036-70-2

267 Schneider, Martin: *Nonlinear Finite Volume Schemes for Complex Flow Processes and Challenging Grids*, 2019, ISBN 978-3-942036-71-9

268 Most, Sebastian Christopher: *Analysis and Simulation of Anomalous Transport in Porous Media*, 2019, ISBN 978-3-942036-72-6

269 Buchta, Rocco: *Entwicklung eines Ziel- und Bewertungssystems zur Schaffung nachhaltiger naturnaher Strukturen in großen sandgeprägten Flüssen des norddeutschen Tieflandes*, 2019, ISBN 978-3-942036-73-3

270 Thom, Moritz: *Towards a Better Understanding of the Biostabilization Mechanisms of Sediment Beds*, 2019, ISBN 978-3-942036-74-0

271 Stolz, Daniel: *Die Nullspannungstemperatur in Gewichtsstaumauern unter Berücksichtigung der Festigkeitsentwicklung des Betons*, 2019, ISBN 978-3-942036-75-7

272 Rodriguez Pretelin, Abelardo: *Integrating transient flow conditions into groundwater well protection*, 2020, ISBN: 978-3-942036-76-4

273 Weishaupt, Kilian: *Model Concepts for Coupling Free Flow with Porous Medium Flow at the Pore-Network Scale: From Single-Phase Flow to Compositional Non-Isothermal Two-Phase Flow*, 2020, ISBN: 978-3-942036-77-1

274 Koch, Timo: *Mixed-dimension models for flow and transport processes in porous media with embedded tubular network systems*, 2020, ISBN: 978-3-942036-78-8

275 Gläser, Dennis: *Discrete fracture modeling of multi-phase flow and deformation in fractured poroelastic media*, 2020, ISBN: 978-3-942036-79-5

276 Seitz, Lydia: *Development of new methods to apply a multi-parameter approach – A first step towards the determination of colmation*, 2020, ISBN: 978-3-942036-80-1

277 Ebrahim Bakhshipour, Amin: *Optimizing hybrid decentralized systems for sustainable urban drainage infrastructures planning*, 2021, ISBN: 978-3-942036-81-8

278 Seitz, Gabriele: *Modeling Fixed-Bed Reactors for Thermochemical Heat Storage with the Reaction System* $CaO/Ca(OH)_2$, 2021, ISBN: 978-3-942036-82-5

279 Emmert, Simon: *Developing and Calibrating a Numerical Model for Microbially Enhanced Coal-Bed Methane Production*, 2021, ISBN: 978-3-942036-83-2

280 Heck, Katharina Klara: Modelling and analysis of multicomponent transport at the interface between free- and porous-medium flow - influenced by radiation and roughness, 2021, ISBN: 978-3-942036-84-9

281 Ackermann, Sina: *A multi-scale approach for drop/porous-medium interaction*, 2021, ISBN: 978-3-942036-85-6

282 Beckers, Felix: *Investigations on Functional Relationships between Cohesive Sediment Erosion and Sediment Characteristics*, 2021, ISBN: 978-3-942036-86-3

283 Schlabing, Dirk: *Generating Weather for Climate Impact Assessment on Lakes*, 2021, ISBN: 978-3-942036-87-0

284 Becker, Beatrix: *Efficient multiscale multiphysics models accounting for reversible flow at various subsurface energy storage sites*, 2021, ISBN: 978-3-942036-88-7

285 Reuschen, Sebastian: *Bayesian Inversion and Model Selection of Heterogeneities in Geo-statistical Subsurface Modeling*, 2021, ISBN: 978-3-942036-89-4

286 Michalkowski, Cynthia: *Modeling water transport at the interface between porous GDL and gas distributor of a PEM fuel cell cathode*, 2022, ISBN: 978-3-942036-90-0

287 Koca, Kaan: *Advanced experimental methods for investigating flow-biofilm-sediment interactions*, 2022, ISBN: 978-3-942036-91-7

288 Modiri, Ehsan: *Clustering simultaneous occurrences of extreme floods in the Neckar catchment*, 2022, ISBN: 978-3-942036-92-4

289 Mayar, Mohammad Assem: *High-resolution spatio-temporal measurements of the colmation phenomenon under laboratory conditions*, 2022, ISBN: 978-3-942036-93-1

290 Schäfer Rodrigues Silva, Aline: *Quantifying and Visualizing Model Similarities for Multi-Model Methods*, 2022, ISBN: 978-3-942036-94-8

291 Moreno Leiva, Simón: *Optimal planning of water and renewable energy systems for copper production processes with sector coupling and demand flexibility*, 2022, ISBN 978-3-942036-95-5

292 Schönau, Steffen: *Modellierung von Bodenerosion und Sedimentaustrag bei Hochwasserereignissen am Beispiel des Einzugsgebiets der Rems*, 2022, ISBN 978-3-942036-96-2

293 Glatz, Kumiko: *Upscaling of Nanoparticle Transport in Porous Media,* 2022, ISBN 978-3-942036-97-9

294 Pavía Santolamazza, Daniela: *Event-based flood estimation using a random forest algorithm for the regionalization in small catchments*, 2022, ISBN 978-3-942036-98-6

295 Haun, Stefan: *Advanced Methods for a Sustainable Sediment Management of Reservoirs*, 2022, ISBN 978-3-942036-99-3

296 Herma, Felix: *Data Processing and Model Choice for Flood Prediction*, 2022, ISBN 978-3-910293-00-7

297 Weinhardt, Felix: *Porosity and permeability alterations in processes of biomineralization in porous media - microfluidic investigations and their interpretation*, 2022, ISBN 978-3-910293-01-4

298 Sadid, Najibullah: *Bedload Transport Estimation in Mountainous Intermittent Rivers and Streams*, 2023, ISBN 978-3-910293-02-1

299 Mohammadi, Farid: *A Surrogate-Assisted Bayesian Framework for Uncertainty-Aware Validation Benchmarks*, 2023, ISBN 978-3-910293-03-8

300 Praditia, Timothy: *Physics-informed Neural Networks for Learning Dynamic, Distributed and Uncertain Systems*, 2023, ISBN 978-3-910293-04-5

301 Gyawali, Dhiraj Raj: *Development and parameter estimation of conceptual snow-melt models using MODIS snow-cover distribution*, 2023, ISBN 978-3-910293-05-2

302 Görtz, Jan: *Coupled modeling approach for physico-chemical processes during the deterioration of cement-based structures*, 2023, ISBN 978-3-910293-06-9

303 Veyskarami, Maziar: *Coupled free-flow–porous media flow processes including drop formation*, 2023, ISBN 978-3-910293-07-6

304 El Hachem, Abbas: *Spatial Extent of Precipitation Extremes in Hydrology*, 2023, ISBN 978-3-910293-08-3

Die Mitteilungshefte ab der Nr. 134 (Jg. 2005) stehen als pdf-Datei über die Homepage des Instituts: www.iws.uni-stuttgart.de zur Verfügung.