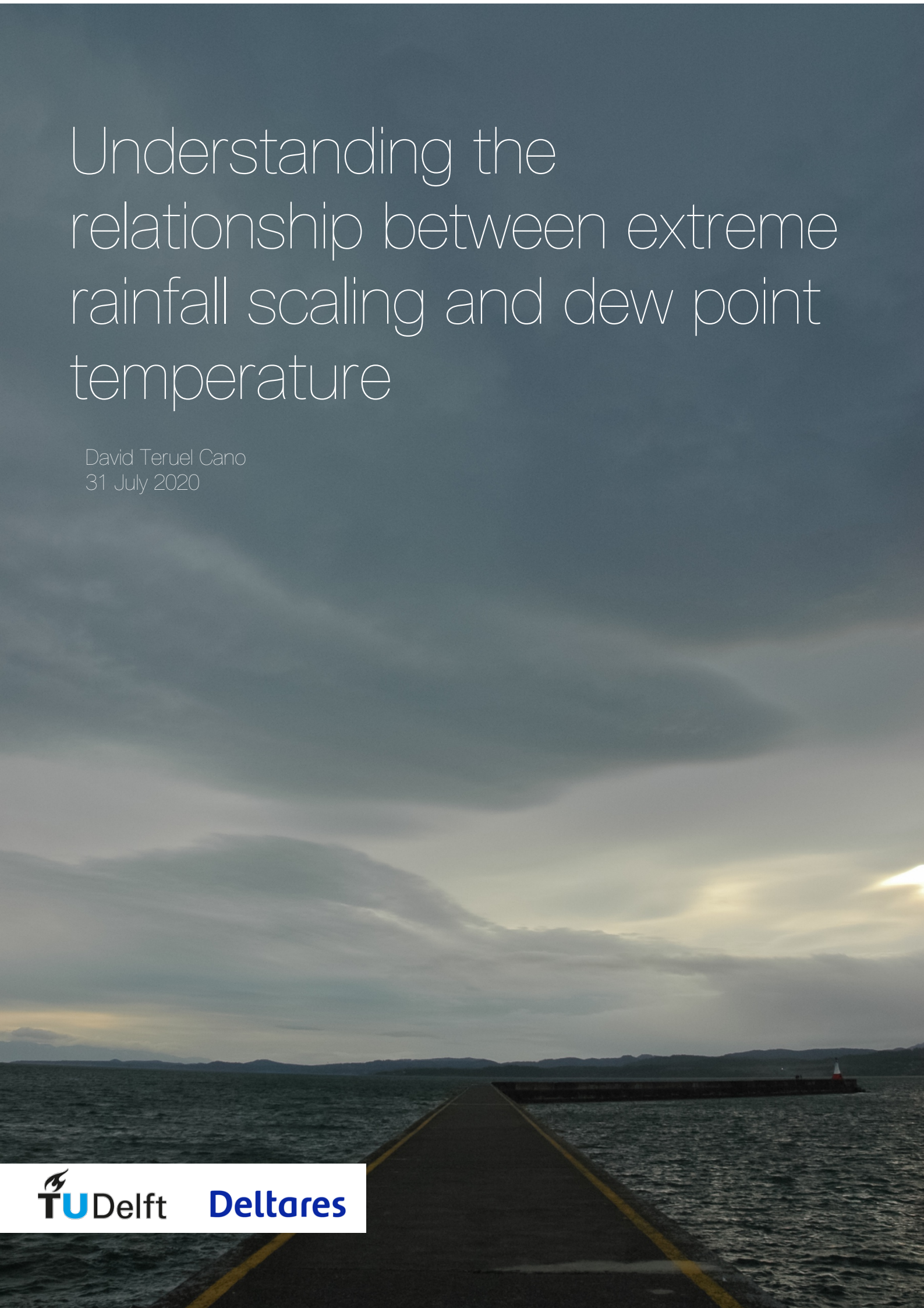
Understanding the relationship between extreme rainfall scaling and dew point temperature

David Teruel Cano
31 July 2020



TU Delft Deltares

Understanding the relationship between extreme rainfall scaling and dewpoint temperature

by

David Teruel Cano

to obtain the degree of Master of Science at the Delft University of Technology,

Student number:	4582969	
Project duration:	December 2019 – July 2020	
Thesis committee:	Dr. ir. M.C. ten Veldhuis,	TU Delft, chair
	Dr. M. Schleiss,	TU Delft
	Dr. ir R.J. van der Ent,	TU Delft
	Ir. R.J. Dahm,	Deltares
	Dr. F.S.W Sperna Weiland	Deltares

Cover photo: Shutterstock royalty-free stock photo ID: 715073116.

An electronic version of this thesis is available at http://repository.tudelft.nl/.

TU Delft

Preface

After 8 months of intense work, this thesis is the final step as a student of the Master in Civil Engineering, in the specialisation of Hydrology and Water Management. The Netherlands has not only been my home for the past 2 years, but it was also where I moved 4 years ago to finish my Bachelor in Civil Engineering, as an exchange student. These years in Delft have been a continuous learning process, and I am grateful for all the people I have encountered during the process.

Firstly, I would like to thank the ”La Caixa” Scholarship for believing in my project, which made it possible for me to continue my studies at TU Delft. Secondly, I would like to thank all the members of my committee for their constant support over these past months, for their availability and eagerness to discuss about this thesis. In the past months of isolation because of Corona, they have always been there, taking care not only about the project but about my well-being.

I would like to express my gratitude to the Chair of my committee, Marie-Claire ten Veldhuis for her constant support throughout the process and for helping me structure my thoughts. I would to thank Marc Schleiss for always coming up with interesting ideas for the project and for helping me during the design of this research. I would like to thank Ruud van der Ent for being available during the thesis. His input and comments on the different stages of the project have helped shape the contents of this work.

I would also like to thank Deltares for giving me the opportunity to work on this topic and for being my home for the past months. In particular, I would like to thank my supervisors at the company, Frederiek Sperna and Ruben Dahm. Thanks to them, I have been able to dive into topics I thought I would never explore, such as meteorology and climate physics. I also want to thank them for the freedom and support I have received. The insight of all the members of my panel has contributed greatly to the quality of this work.

There are a group of people that have accompanied me during this past year. I would like to thank the ”Simonsstraat family”, my flatmates, who have followed with me all the ups and downs during the development of this thesis. Also, I would like to give a special thanks to the interns group at Deltares, with whom I have shared stories and experiences over the virtual coffee breaks during lock down.

Finally, I would like to express my gratitude to my parents and my sister. Although I have been living abroad for 4 years, I still feel constantly supported and guided in every decision I made.

David Teruel Cano July 2020.

Abstract

Due to global warming, temperatures are expected to rise and, with it, the increase in moisture holding capacity of the atmosphere. Changes in precipitation extremes with temperature are governed by the Clausius-Clapeyron relationship (CC), which states that precipitation increases 7% per degree of warming. While global models and observations are consistent with CC, research at hourly observations have showed deviations in the CC relationship, showing steeper increase up to 14% per degree (super-CC) and <7% per degree for larger temperatures (sub-CC). Studies have shown that scaling rates can be influenced by several factors, such as the length of observations, rainfall percentiles or the timescale of rainfall, among others. Apart from methodological limitations, research on scaling rates is not readily available in all parts of the world, with scarce studies carried out in the African continent.

In this context, there is a need for better understanding the relationship between precipitation and temperature as well as the implications for future extremes. To address these challenges, this research investigated 3 main topics. Firstly, existing methodologies were compared and analysed for 353 rain gauges in Germany. Using as extreme rainfall indicators wet percentiles (80th-99th quantile), 3 methods were tested and statistically compared to estimate scaling rates with dew point temperature (T_{dew}): (i) binning methods, (ii) quantile regression and (iii) change-point models at an hourly timescale. Secondly, the presence of super-CC was evaluated for storm events instead of fixed hourly intervals. For that, data from Germany at 10-min resolution was used. The last step involved the analysis of scaling rates for a dense network of newly available rain stations in 4 countries in sub-Saharan Africa: Ghana, Kenya, Tanzania and Uganda.

The results from measuring the scaling of extreme precipitation with T_{dew} shows that two regimes arise for Germany: for T_{dew} below 13ºC- 15ºC the scaling rates are close to the CC rate across all quantiles. For warmer T_{dew} of 15ºC or more, scaling rates above the CC are observed, with increasing slope for the highest quantiles. The study of storm events shows that, for mean intensity and peak rainfall intensity, there is a strong sensitivity to T_{dew}, with higher T_{dew} associated with higher rainfall. The analysis of storm duration indicates that shorter events are more frequent for larger temperatures than for shorter ones. The study in sub-Saharan countries shows consistent negative scaling rates with dew point temperature. In addition, for large surface air temperatures (SAT), there is a decrease in relative humidity, suggesting that there is a general lack of moisture availability at the higher SAT ranges. To better understand the physical processes that influence these results, the use of alternative predictors that describe the atmospheric moisture are proposed.

In a future warming world, current literature indicates that the greatest rainfall increases will be more visible for short-duration storms rather than increases in annual mean precipitation. As such, within this project, a novel analysis on scaling rates for storm properties shows the potential of these approach to obtain detailed information on the storm properties and their relationship with temperature, which could be used in storm simulations and projects characterising current or future climate.

Summary

In the context of global warming, temperatures are likely to rise and, since warmer air has the potential to hold more water vapour, extreme weather events are also expected to intensify. The rate of increase of precipitation with temperature is described by the Clausius-Clapeyron (CC) relationship, which states that the water vapour holding capacity of the air increases 7% for every degree increase. While this has been observed globally at a daily time-scale, an empirical study carried out in 2008 in the Netherlands showed that, for sub-daily rainfall, the scaling rates obtained were up to 2 times the CC rate (super-CC). The causes behind super-CC are yet unknown, although there are 2 hypothesis that are being studied, one arguing that super-CC has a physical origin related to the properties of the convective regime and the other which states that it is a statistical effect originated by a transition between rainfall types, with convective precipitation dominating at high temperatures and large-scale precipitation dominating for colder temperatures.

The study of scaling rates has been carried out in many locations around the world, which have shown scaling rates larger than CC are observed at sub-daily resolution. Common approaches to obtain the scaling rate have been binning precipitation data in bins, from which an exponential regression is fitted to derive the scaling rate. However, these methods can be biased depending on how many observations each bin has. To address this issue, unbiased methods such as quantile regression have also emerged. Recent research argue that observations do not show a constant CC over the whole temperature range (both T_{dew} and SAT). To address this issue, change-point models have been proposed. In this context, the influence that the methodology has on the scaling rates is large and not fully understood, limiting the applicability and certainty of the scaling rates. In addition, while most studies have looked at pairs of Precip-T_{dew} at fixed time intervals, it has been shown that this method might not fully describe rainfall variability.

In order to tackle this limitations, and to address the influence that existing methodologies have in obtaining scaling rates, the aim of this project is to study what are the dependencies between dew point temperature and precipitation through a comparative analysis between existing methods, and the study of the scaling rate against several storm properties. To do so, 2 climatological regions are used, Germany, where 10-min observations over 25 years are used, and 4 countries in sub-Saharan Africa (Ghana, Kenya, Tanzania and Uganda), where a 5-min high resolution dataset with 3 years is used.

After critically evaluating the sensitivity that binning methods, quantile regression and change-point models have on scaling rates, the results from the application to a 25-year dataset in Germany shows that while binning methods can be biased, using bins separated with equal length of observations result in comparable scaling rates to the quantile regression method. From the models with no change-point, quantile regression gives the smallest confidence interval, with rates ranging from 1.01%/ºC to 1.88%/ºC. This is explained by the fact that all observations are used when computing the scaling rate, as opposed to binning methods where the fit depends on the number of observations per bin. The comparison with change-point models show that, out of 353 stations studied, 80% have a better model performance with a change-point model. From those selected stations, most change-point occurs within 8ºC and 15ºC, with a transition from CC to super-CC scaling rates. This occurs because of a change in the distribution of rainfall across temperatures. For lower dew point temperatures, the extreme rainfall observations are similarly constant whereas, for larger dew point temperatures, larger extreme precipitation observations occur when increasing the temperatures.

The analysis based on storm events characteristics (peak rainfall intensity, mean intensity during the event, total rainfall and duration) demonstrates that peak rainfall and mean intensity are the variables with the strongest Precip-T_{dew} relationship, reaching larger scaling rates than the CC. A negative relationship between duration-dew point temperature suggests that shorter duration events are predominant for larger temperatures, as opposed to longer events which take place in colder temperatures. The study shows that conditioning storm events to duration influences the scaling rates. While for total

rainfall depth, the scaling rates are almost constant around 4-5%/ºC, when the rainfall depth is conditioned to duration, shorter events show scaling rates slightly larger than the CC rate, while long events show flatter behaviour.

The spatial representation of scaling rates both at fixed time intervals and based on storm events in Germany show that the largest results are observed on North-East and center of the country, and the lowest values in the north-west. These patterns indicate that there are other effects apart from temperature (dynamic contribution) that might be influencing the precipitation in the region.

With a focus on the sub-Saharan Africa, the main findings show predominant negative scaling rates. Since for some stations the data is scarce, there are cases where initial bins are undersampled, which, when using binning methods results in steep positive scaling rates. In the cases where the initial bins are well sampled, most stations show negative slopes larger than -22%/ºC. Given the high resolution of the dataset and the short period of observations, it is suggested that further research look at the storm events characteristics instead of fixed time intervals.

In summary, the results from applying several methods to estimate scaling rates indicate that there are two precipitation regimes across dew point temperature, one in lower T_{dew} that shows scaling rates between 5-7%/ºC and one for larger T_{dew} with scaling rate exceeding the CC. This suggests that, although the uncertainty associated to change-point predictions is still large, they capture in a better way rainfall distribution across T_{dew}. The findings from the scaling rates of storm events show that storm properties are sensitive to T_{dew}, and should be taken into account in present and future extreme rainfall analysis. In particular, the differences in scaling rates observed for varying duration could be used when defining storms in stochastic storm generators. With this project, a better knowledge on how different storm properties increase with dew point temperature for short and long events provides room for better characterising not only individual storms, but the impact of different storm combinations.

Contents

List of Figures

List of Tables

1 Introduction

1.1. Background

It is widely agreed that, due to global warming, temperatures around the world will rise. As warmer air has the potential to hold more water vapour, precipitation extremes are also expected to increase (Trenberth et al., 2003). The expected changes in precipitation extremes are generally based on the Clausius-Clapeyron relationship (CC), which states that for every degree increase in temperature, there is a 7%/ºC more moisture holding capacity in the atmosphere (Allen & Ingram, 2002).

Over the last decade, research has focused on studying the dependence of precipitation with temperature and moisture content, scaling it with dew point temperature for a set of geographies: USA (Mishra et al., 2012), Australia (Wasko et al., 2018), Netherlands (Lenderink & Van Meijgaard, 2008), Germany (Berg et al., 2013) or Switzerland (Molnar et al., 2015), among many others. For example in the Netherlands, it was shown that, for dew point temperatures larger than 12º, one-hour precipitation extremes increased twice as fast with rising temperatures as expected from the CC relation (referred to as super-CC) (Lenderink & Van Meijgaard, 2008). In a warming climate context, it is thus vital to understand the relationship between extreme rainfall and temperature.

Apart from regional and country-level analysis, Ali et al. (2018) studied the scaling relationship on a global scale, taking daily precipitation observations and showed that, for 75 % of the stations, the relationship between dew point temperature and precipitation was mostly consistent with the CC, only reaching super-CC in some instances.

The general conclusions from these studies indicate that, overall, the relationship between temperature and extreme rainfall is not always consistent with the scaling rate of 7%/ºC as described by the CC relation. In fact, for a range of temperatures between 12ºC and 22ºC, it has been shown to scale up to 2 times more than CC (Figure 1.1). In a study carried out in Australia, Hardwick Jones et al. (2010) also showed that, for temperatures above 24ºC, there was a decrease in extreme rainfall intensity, associated with a drop in the relative humidity. Berg et al. (2009a) also found that a decline in moisture availability for high temperatures resulted in a negative CC slope.

In general, the physical explanation behind the presence of super-CC is still under discussion, although it is thought to occur due to changes in the dynamics of the atmosphere and the convective cloud, such as the size of convective clouds (O'Gorman & Schneider, 2009). There are several factors that have been seen to influence the scaling rate inferred from temperature and precipitation observations, such as (1) the type of regression model applied (Westra et al., 2014; Van de Vyver et al., 2019), (2) spatial scale of the analysis (Hardwick Jones et al., 2010; Pumo et al., 2019), (3) the storm characteristics (Berg et al., 2013; Molnar et al., 2015; Wasko et al., 2015; Lochbihler et al., 2019) or (4) the influence of difference climate regions (Ali et al., 2018; Westra et al., 2014).

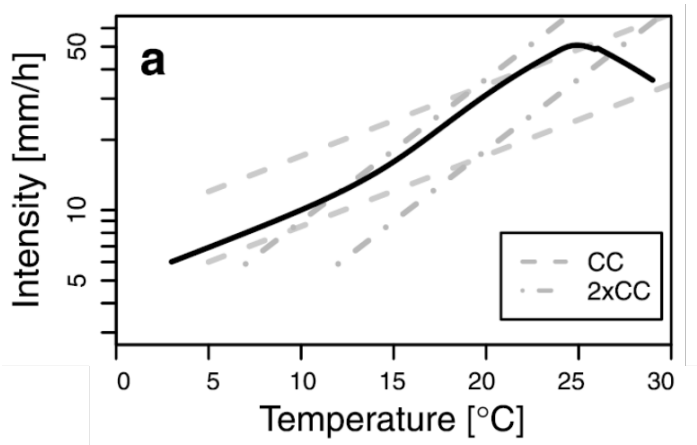


Figure 1.1: Conceptual diagram of observed relationship between temperature and extreme rainfall (Westra et al., 2014).

1.2. Problem statement

Although multiple observational and modelling studies have found scaling rates that deviate from the CC rate, there are only 2 or 3 studies that have investigated the use of change-point models (Pumo et al., 2019; Van de Vyver et al., 2019; Dahm et al., 2019). In addition, these studies has only looked at individual locations for different countries, lacking an in-depth analysis a large set of stations within the same country. From a statistical perspective, questions arise about (i) the validity of change-point models versus constant slope models, and (ii) the uncertainty of the estimated change-point temperature and slope estimates.

When estimating scaling rates, there are several factors that can influence the result. These range from the method used, the input data and predictors, the timescale of analysis or the indicators used to define extreme rainfall. One of the key challenges is the influence of observations at different timescales. Existing research looking at the behavior of precipitation intensity from the instantaneous to the daily resolution has shown that using fixed time intervals might not provide the best representation of rainfall variability. (Haerter et al., 2010). This shows that, while efforts are now in studying the scaling rate at an hourly resolution, sub-hourly analysis will still be needed.

With a focus on the coverage of scaling relationship studies in the world, studies in different climatological regions have shown that there are substantial differences in scaling rates between different locations (Westra et al., 2014). However, data scarcity at a sub-daily timescale difficult the development of projects in certain locations. This is specially relevant in the African continent.

1.3. Research Objectives

The main aim of this thesis is to understand what are the underlying factors influencing differences in the T_{dew}-Precip scaling rate, by studying its sensitivity to hourly observations and several storm event characteristics. To address this objective, this research focuses on 2 climatological regions: Germany and West and East African countries. The thesis is organised in 3 main areas of work: Part 1 aims to explore the potential and limitations of existing methods for estimating scaling rates. Part 2 will evaluate the relationship of dew point temperatures with different storm event properties. For these sections, the analysis will be focused on Germany. In parallel, this project will put into perspective the results of Part 1 and 2 by analysing the scaling rates on four West and East African countries (Figure 1.2).

Part 1. Understanding the influence that existing methods have on estimating scaling rates: In the first part of the research, hourly precipitation observations are studied by developing 2 types of models, exponential regression and piecewise regression models. For the first type, a binning approach and quantile regression are tested. For the second type, a change-point model based on piecewise linear regression and piecewise linear quantile regression are developed. The aim of this section is to:

- Compare the suitability and sensitivities of using exponential regression models and piecewise regression models for estimating precipitation scaling rates.

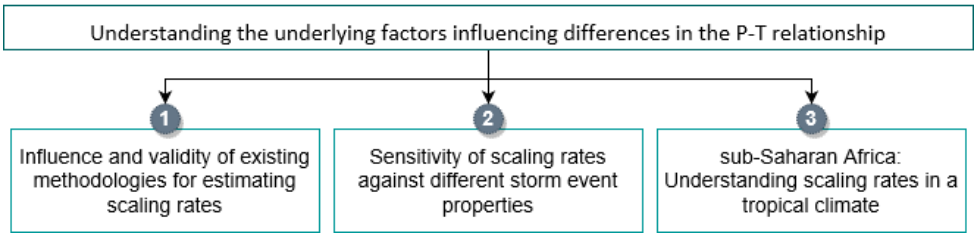


Figure 1.2: Research objectives.

- Study the presence of super-CC for hourly extreme precipitation intensity (from the 80th to 99th Quantile)
- Evaluate the success and validity of change-point models to detect super-CC scaling.

Part 2. Variability of scaling rates against different storm event properties: While the first part looks at a fixed time interval (1-hour observations), in this one the analysis is based on storm events. For each storm, several storm properties are retrieved, such as duration of event, peak intensity, mean intensity and total rainfall depth. The aims are to:

- Test the influence that storm properties have on scaling rates
- Evaluate the presence of super-CC rates when storm properties are related to dew point temperature
- Investigate the significance of adding storm characteristics as covariates in regression models for T_{dew}-P scaling rate.
- Investigate a methodology for predicting changes in storm event characteristics based on scaling properties.

Part 3. Scaling rates on different climatological regions: West and East Africa: In parallel to Part 1 and 2, this last section will study the scaling rates for a set of observations from 4 countries in sub-Saharan Africa (Ghana, Kenya, Uganda and Tanzania), using data at hourly resolution. The aims in this part are:

- Evaluate existing hypothesis on expected scaling rates in sub-Saharan Africa
- Test the usefulness and limitations of binning approaches and quantile regression for estimating scaling rates in the tropics.

1.4. Research questions

The main question that this thesis aims to answer is: What is the influence that different storm event characteristics have on the scaling rate? Can this be used to predict extreme rainfall characteristics with changes in temperature? In order to answer this question, the project first addresses several sub-questions organised in the main Parts covered in this thesis:

Part 1: On the comparison of scaling rate methods at a fixed hourly precipitation interval:

- What differences exist between methods to estimate scaling rates? How does the choice of fitted model influence the scaling results?
- What is the validity of the estimated scaling rates and detected change-points?
- What model considerations should be taken into account when applying the different methods?

Part 2: On the scaling rates observed by storm event properties:

- What are the scaling rates associated with different storm properties?
- Are super-CC rates also present when looking at storm event properties?

Part 3: On the scaling rates in sub-Saharan Africa:

- What are the observed scaling rates? How does the result compare to existing research in the region?
- What are the limitations in quantifying the scaling rates with short length of observations?

1.5. Contribution of the research

While the scaling of precipitation with temperature and dew point temperature has been identified in a number of studies, there are still many uncertainties that need to be addressed. New modelling approaches based on changing-point methods have been developed to better capture the complexity of extreme rainfall behavior (Van de Vyver et al., 2019). However, further research is needed not only to evaluate the influence that different methodological approaches have on the obtained scaling rates but also to validate the significance of super-CC behavior across different locations (Westra et al., 2014). The first Part of this thesis aims to address this issue.

In current research, scaling rates have been calculated based on combinations of precipitation and temperature defined at varying temporal scales, at a fixed time interval or based on storm properties. So far, most studies focus on hourly and daily analysis of precipitation. However, research has shown that a fixed time interval might not always provide an accurate representation of rainfall variability (Haerter et al., 2010). Therefore, in the second Part of this thesis, I aim to look at different storm properties and check whether, in those conditions, super-CC would also be observed. While scaling rate analysis based on storm properties has been covered less in research, the outcomes could have greater applications than fixed-interval analysis, as it would allow to extrapolate event characteristics with changing temperatures. This would be specially relevant for finding new uses of scaling rates for managing extreme weather, for example in urban areas.

Lastly, while global analysis on the scaling of extreme precipitation have already been carried out (Ali et al., 2018), an analysis for the African continent at a sub-daily resolution is still not available. As such, this thesis has the opportunity to study, for the first time, the scaling relationship in sub-Saharan Africa. Using high-resolution data from an ongoing project that is currently installing weather stations across Africa, the results from this study could help raise awareness about the opportunities that these stations could provide for climate research.

1.6. Thesis outline

The thesis is organised in several chapters:

- Theoretical background: An initial description of the topic and the current methodologies used to estimate the scaling rate are presented. In addition, an overview of the scaling rates at different locations and temporal resolutions is shown, together with the hypothesis for the underlying causes of large scaling rates. Lastly, the climate characteristics of the study areas are briefly introduced.
- Methodology: This section provides information about the different methods that have been applied in to estimate the scaling rate, as well as the statistical test and metrics used for comparison between models. For the storm event analysis, a description of the algorithm created to retrieve the storm events is provided as well as the details on the storm event analysis.
- Part 1: Analysis of scaling rates methodologies: This section covers the results from the analysis of different models to observations for Germany. In detail, exponential and quantile regression models are compared with change-point models.
- Part 2: Analysis of storm events: In this section, the results from estimating scaling rates based on storm properties is presented. This is done with observations of 25 years from Germany.
- Part 3: Scaling rates in sub-Saharan Africa: An introduction to the first results in Africa are presented. In particular, the results from applying the exponential regression models to several stations of the countries.
- Discussion: The discussion is presented by addressing the research objectives of the thesis.

- Conclusions: This is the last section of the thesis and provides a summary on the key findings while also referring to some limitations during the development of the project and recommendations for future research.
- Appendix: At the end of this document, additional figures are presented to support the discussion and analysis of the different parts in the thesis.

2

Theoretical Background

2.1. The Clausius-Clapeyron relation

The capacity of the atmosphere to hold water is established by the Clausius-Clapeyron (CC) equation (Westra et al., 2014), which relates the water-holding capacity of a gas and its temperature. Under the assumptions that there is a constant relative humidity, a relationship can be established that the moisture-holding capacity of the atmosphere increases at about 7%/ºC at 0ºC and around 6%/ºC at 24ºC (Trenberth et al., 2003).

Global studies have shown that, over higher latitudes, extreme daily precipitation increases with surface air temperature at a similar rate, while in the tropics it decreases for high temperatures (Utsumi et al., 2011). However, it has been seen that other variables can affect the scaling rates, for example, changes in storm duration (Molnar et al., 2015), the type of extreme precipitation (convective or stratiform) (Berg et al., 2013), or seasonality (Berg et al., 2009a). In addition, many studies have also tried to estimate what is the actual thermodynamic contribution (based on temperature changes) and the dynamic contribution (connected to atmospheric circulation) to the scaling values (Oueslati et al., 2019).

2.1.1. Super Clausius-Clapeyron

An increase in precipitation intensity of twice as the CC scaling (super-CC) was first shown by Lenderink and Van Meijgaard (2008). The research showed that the 99th percentile of precipitation increased with surface air temperature at a rate up to 14%/ºC for temperatures larger than 12ºC (Figure 2.1).

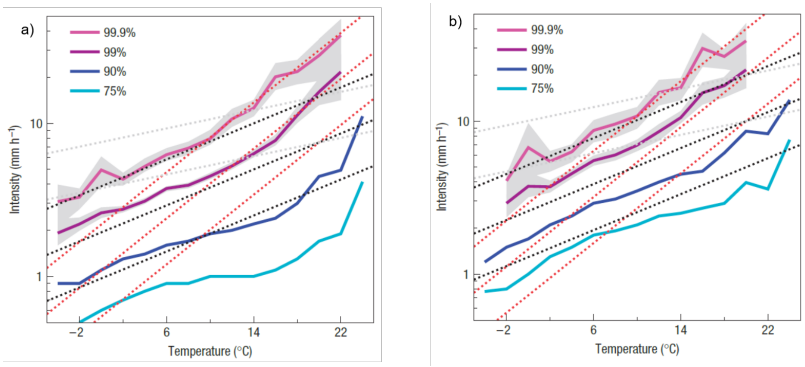


Figure 2.1: Observed 1h precipitation intensity with SAT (a). Observed maximum 1h precipitation intensity (b). Grey lines are the 90% confidence intervals estimated by the bootstrap. Dotted black lines represent the 7%/ºC and red lights the 14%/ºC (Lenderink & Van Meijgaard, 2008)

.

This increase of the scaling rate beyond 7%/ºC supported the evidence that dynamic processes also influence the relationship between precipitation and temperature (O'Gorman & Schneider, 2009). In

addition, it showed that the scaling rate might not be constant for the whole range of temperatures.

The underlying causes of super-scaling are yet to be fully understood. There are two main arguments among the research community. The first one, proposed by Berg et al. (2009b) suggests that super-CC occurs due to statistical behaviour of data. In detail, they argue that the transition from CC to super-CC occurs because: (1) there is a very large timescale difference between large scale and convective precipitation and (2) convective precipitation dominates at high temperatures whereas large scale precipitation dominates for cool temperatures. Their approach is that the unexpected rise of CC occurs because of a change in rain type. Figure 2.2 shows the conceptual description of the proposed argument.

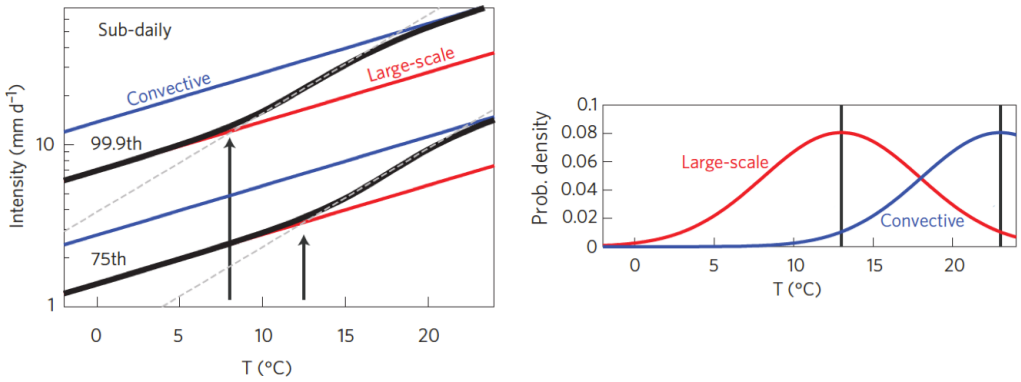


Figure 2.2: Conceptual model on super-CC causes. Left: 99.9th (upper curves) and 75th (lower curves) sub-daily precipitation intensity percentile of large-scale (red), convective (blue) and total precipitation (black), and double Clausius–Clapeyron increase (dotted grey). Right: Weighting functions for large-scale (red) and convective (blue) precipitation as function of temperature (Berg et al., 2009b).

In 2013, an analysis that used synoptic codes to separate stratiform and convective precipitation showed that the super-CC increase occurred in the transition region between the stratiform (low temperatures) and convective precipitation (high temperatures) (Berg & Haerter, 2013).

The second argument about super-CC is proposed by Lenderink and van Meijgaard (2009), who argue that super-CC has a physical origin related to the properties of the convective regime. In their view, as temperature rises, there is an increase latent heat resulting in convective clouds with stronger updrafts and dynamics.

2.2. Scaling rates at a global scale

In 2011, a study with almost 9,000 stations revealed the global distribution of scaling rates for daily observations under the 99th wet percentile (Utsumi et al., 2011). The results showed 3 types of slope patterns (Figure 2.3). For the higher latitudes (above 55ºN), a monotonic increase of daily precipitation was found with increasing surface air temperatures. In mid-latitudes (between 20º-55ºN and S), the relationship had a peak-like structure and, for the tropics (between 20º-20ºN), extreme daily precipitation decreased monotonically (Utsumi et al., 2011). For warmer climates like the tropics, the decrease in slope occurred between SAT of 22-24ºC.

Local studies in the tropics at both daily and hourly resolution showed similar results to those obtained by Utsumi (Wasko et al., 2015). There is not a unique explanation to why it occurs, although several hypothesis have been formulated. The following sections discusses some of them.

2.2.1. Scaling rates in the tropics

In a study in the French Mediterranean region, Drobinski et al. (2016) suggested that negative scaling could occur because of arid surface conditions; In detail, they concluded that SAT was not a good proxy for temperature because, in dry conditions, the clouds and precipitation form at a higher level and, thus, have much lower temperature than the SAT. Berg et al. (2009b) and Hardwick Jones et al. (2010) argued that the reduction in precipitation with high temperature was due to a limitation on moisture availability. Hardwick analysed in Australia, the relative humidity relationship with SAT and showed that there was

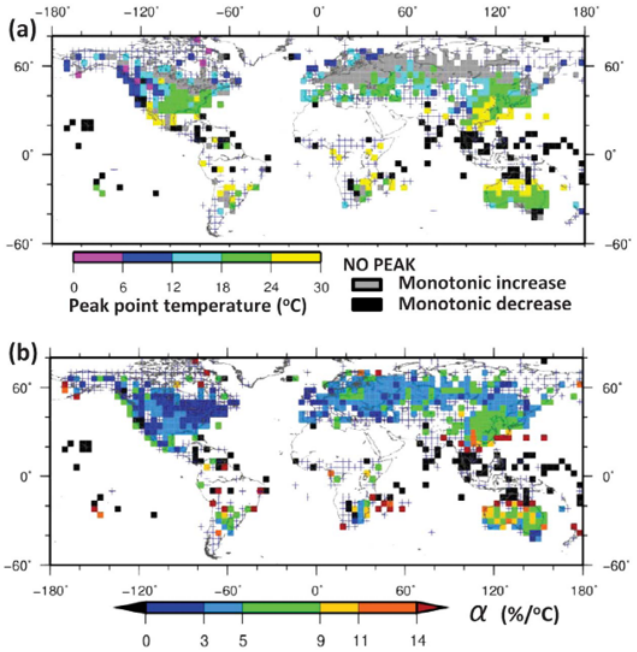


Figure 2.3: Global distribution of (a) peak point temperatures and (b) for $P99_d$. The colored grid boxes have a with 5% significance level. Grid boxes with no significance are indicated by crosses (Utsumi et al., 2011).

a decline in moisture availability at large temperatures, which matched the inflexion point of the scaling rates. For temperatures larger than 24ºC, there was a reduction of moisture availability. This outcome indicated that not only is important to consider the capacity of the atmosphere to hold moisture, but also how much moisture is present Hardwick Jones et al. (2010).

Another explanation for negative scaling was proposed by Roderick et al. (2019) after investigating the relationship between SAT and integrated water vapour (IWV). In a recent research they showed that, while for smaller SAT there was an increase in IWV, for larger SAT the IWV did not increase but actually decrease at a specific temperature. They argued that the reason behind the negative scaling scaling was because, for warmer regions, there is less surface moisture available for evaporation.

The whole process was explained by the limitations of temperature caused by evaporation under warm climates. As described by Roderick et al. (2019), if there is surface moisture available and SAT increases, the SAT will remain constant and higher evaporation will occur. Therefore, when there is insufficient surface moisture, the SAT increase because there is no moisture available, which, in turn, leads smaller extreme precipitation (because of reduced IWV). In conclusion, there is consensus that arid surface conditions can be the cause associated to negative scaling (Drobinski et al., 2016).

This suggested hypothesis on negative scaling rates pose the question to what extent and in which context does the use of dew point temperature prove useful in the tropics or whether alternative predictors will be needed in the tropics.

2.2.2. Dew point versus surface air temperature

In order to overcome the limitation of relative humidity on the larger temperature ranges, several studies that have compared the benefit of estimating scaling rates based on dew point temperature instead of surface air temperature (Ali et al., 2018; Van de Vyver et al., 2019; Bui et al., 2019; Hardwick Jones et al., 2010; W. Zhang et al., 2019). Dew-point temperature is defined as ”the temperature to which an air parcel must be cooled (at constant pressure) to reach saturation” (Lenderink et al., 2011).

The authors of previous studies all agree that dew-point temperature is a better predictor of extreme precipitation than air temperature. For some locations in the tropics, estimating the scaling rate with dew-point temperature has been specially relevant, showing a strong positive scaling relationship using dew point (in contrast to using SAT) (Wasko et al., 2018). In a study by Lenderink in 2011, both SAT and dew point temperature were used on the same precipitation data. The results suggested dew point

temperature is a better predictor than temperature. When SAT was used, there was a decay at 24ºC whereas for dew point temperature, there was no decay observed. In a similar study, the comparison between SAT and dew point temperature was also made by Panthou et al. (2014) for Canada and showed that, while for inland regions, using dew point temperature strongly smoothed the peak-like structure obtained by SAT, this was not the case for coastal regions, where the results of the scaling rate between SAT and dew point were almost the same (humidity does not limit the increase of extreme rainfall with temperature). This means that, for inland locations, humidity was a limiting factor for these regions, but not for coastal locations.

2.3. Results from climate models

To compare the results from analysis with historic observations, most studies make use of Global Climate Models (GCM) and Regional Climate Models (RCM) (Lee et al., 2019). Given the challenge to obtain sub-daily observations from all regions of the world, studies look at how reliable can projections be made with climate models. At a daily time-scale, evidence from high-resolution climate models have shown an increase in the intensity similar to the CC relation (Westra et al., 2014). However, at a sub-daily level, GCM are not able to explicitly solve convective storms. Instead, implicit parametrization are developed (Lenderink & Van Meijgaard, 2008). In order to better estimate projections of extreme rainfall, model resolutions in the order of kilometers are required, although the computational cost is high (Westra et al., 2014).

2.4. Existing methodology approaches

The accuracy in representing the scaling relationship can be susceptible to the type of regression model used (Wasko & Sharma, 2014). Since this topic was first studied, there have been several approaches taken to evaluate and study the relationship between extreme rainfall and temperature. A short overview of the most widely used methods and their tradeoffs is presented in this section.

2.4.1. Binning Approach

One of the first methods used to study the rainfall-temperature scaling phenomena is the binning technique. This method, applied for the first time by Lenderink and Van Meijgaard (2008) in a 99-year record dataset, consists in grouping hourly or daily rainfall observations in temperature bins, for which percentiles are computed. The precipitation percentile is log-transformed and a regression with temperature is done per bin (Wasko et al., 2015).

Over the years, there has been a lot of variation in the method used for binning temperature, such as (i) the predictors used, (ii) the temporal resolution of precipitation, (iii) the percentiles of interest and (iv) the number of bins or spacing between them. An analysis from Wasko and Sharma (2014) highlighted some limitations of binning precipitation with temperature after making a comparison between 4 variations of binning methods. The results showed that, with varying lengths of observations, the methods that kept giving similar scaling rates while increasing the nº of observations were (a) linear regression on percentiles estimated from bins with equal sample number, followed by (b) linear regression on percentiles estimated from 2ºC equal distance bins and (c) linear regression on percentiles estimated from 2ºC equal distance bins with a 1ºC overlap; Overall, the performance of having equal number of observations per bin was higher compared to the other binning methods because, for the 2º bin method, there were less number of observations under the upper-range temperatures (Hardwick Jones et al., 2010; Wasko & Sharma, 2014). Based on the results by Wasko and Sharma (2014), 2 binning methods will be analysed in this study, one with equal number of samples per bin and a one with bins of 2ºC distance.

2.4.2. Quantile regression

In the same research by Wasko and Sharma (2014), they demonstrated that, when the scaling rate is constant against a temperature range, the use of quantile regression can give better results compared to binning methods. (Van de Vyver et al., 2019). For a given predictor variable, quantile regression techniques models the conditional quantiles of a response variable. The main difference from standard linear regression is that quantile regression only models the conditional mean of the response variable (Koenker, 2005). The comparison between binning and quantile approaches for different length of ob-

servations showed that the regression estimators for quantile are asymptotically unbiased (R. Koenker & Bassett Jr, 1978). An extensive description of the quantile regression method is presented in the Methodology section.

From the abovementioned models, there is still the question about how much detail can they capture. While several papers have observed super-CC, mostly at a (sub)daily rainfall level, they have, in most cases, considered linear regression methods. This is not always the best approach because, as Lenderink and Van Meijgaard (2008) showed, super-CC does not occur over the whole temperature spectrum, but can transition from CC to (sub) and (super) CC and return back to CC for large temperatures. As such, it is probably relevant to include models with change-points.

2.4.3. Piecewise regression models

To account for these scaling changes, only few studies have incorporated transition-points models. For the Netherlands, Dahm et al. (2019) considered a fixed transition approach between CC and super-CC based on a specific break-point temperature that a previous research found. While this approach produced a better model compared to simple quantile regression, it did not allow for its extrapolation to other locations (since different regions can experience a transition from CC to super-CC at different temperature range).

In addition, existing changing-point approaches have mostly focused on one station, leaving the question of whether change-points would be observed at multiple stations in the same country. Recent methods are currently being applied to study an automated way to obtain the transition between CC and super-CC, namely piecewise linear regression (Van de Vyver et al., 2019). The main idea behind this method is to not only determine what scaling is found but also the temperature range where the transition occurs. However, this approach has only been tested in Western Europe countries, leaving the question to whether such methods will also be applicable at different locations with different scaling behaviours (like in the tropics). In addition, this study only looked at the transition between CC and super-CC, leaving out of the analysis the behaviour of the scaling rate for large temperature observations. Lately, new research has emerged (Pumo et al., 2019) which has looked at piecewise regression models using the binning approach. Overall, there are still many uncertainties in these models and it remains unsure whether change-point occurs for the whole country, for some stations or under specific characteristics. Within this research, this aspect will be addressed.

2.5. Influence of storm events characteristics

Existing research looking at the behavior of precipitation intensity from the instantaneous to the daily resolution has shown that using fixed time intervals might not provide the best representation of rainfall variability (Haerter et al., 2010). The implications of considering only fixed intervals can have consequences in the estimated scaling. Wasko et al. (2015) observed that in tropical areas, only looking at historical relationship between precipitation and temperature resulted in contradicting negative scaling relationships. While research shows that this can be due to moisture availability (Hardwick Jones et al., 2010), they proved it was due to the applied method. In their study, they incorporated storm duration as a covariate and demonstrated positive scaling rates in locations where negative scaling was previously found.

However, studying the scaling rates for storm event properties is still a novelty in the field, only studied by few (Molnar et al., 2015; Panthou et al., 2014; Wasko et al., 2015). Molnar et al. (2015) looked at 4 storm properties (rain duration, peak rainfall, mean intensity and total rainfall depth) and showed that the strongest scaling relations to air temperature come from mean and peak event intensity, with scaling rates larger than the CC relationship. In addition, the study of mean and peak intensities for increasing duration also showed that the scaling rates of hourly observations were smaller than the sub-hourly timescales.

Conditioned on storm duration, Haerter et al. (2010); Molnar et al. (2015) have shown that as temperature increases, event duration decreases. Also related to the event duration, a study on peak intensity of short and long events showed that on the hourly and sub-hourly scale, long rainfall events increase slower with temperature than peak event intensities in short rainfall events (Schroeer & Kirchengast, 2018).

One of the hypothesis behind the causes of super-CC is that it occurs due to a transition from different rainfall types (from stratiform to convective) (Berg et al., 2009b). To address this issue, there have been several publications investigating the scaling rate of different types of rain, each one using different proxies for inferring convective or stratiform storms. In Molnar et al. (2015), storm types were obtained by checking the presence of lightning (in order to identify convective events). Berg et al. (2013) grouped the stations based on their location, creating 2 groups, one inland and another coastal and, using synoptic data, they classified precipitation measurements according to the cloud types. The outcomes of studying different storm types show that the scaling rates obtained using mean event intensity or peak intensity when storm types are mixed can result in an amplification of the scaling rates (Berg et al., 2013; Molnar et al., 2015).

Recently, new storm properties have been studied, looking not just on a temporal resolution but spatially. In 2019, a study investigated the spatial extent of convective cells and concluded that some other factors that can affect the scaling rates are: source area of moisture, the cloud size, or the degree of mesoscale organization (Lochbihler et al., 2019).

2.6. Other factors influencing scaling rates

2.6.1. Influence of observation length

While the analysis of scaling rates has been carried out mostly at a gauge level, some studies have investigated whether the scaling rate can differ when stations are clustered. Pumo et al. (2019) studied the scaling relationship between single gauges and pooled gauges for 10 min sub-hourly precipitation and observed that, while similar results between single and pooled stations were observed, the amount of data justifying the scaling was very different between methods. For example, looking at a dry Period (summer season), a decrease in the scaling was observed at high temperatures. While in the pooled method there were many precipitation measurements that described the behaviour at high temperature range, when the single station was analysed, the fit at high temperature was only given by 2 bin points (and thus had to be statistically neglected). This shows that, depending on the quantity of data, pooling stations can reduce the uncertainty from the results observed at a single gauge analysis.

A similar study by Molnar et al. (2015) divided Switzerland in 4 regions based on orography characteristics (Plateau, Pre-Alps, Alps and Tessin), and showed that the scaling rates for each region had similar value, therefore allowing for clustering of gauges.

2.6.2. Influence of different climate regions

While most studies have addressed the scaling relationship on a local level and individual stations, the question on how the trends in sub-daily rainfall events will evolve in the near future still remains. Overall, on a continental scale, an increase of short-duration events (minutes to hours) is expected. (Westra et al., 2014). There are, however, some differences between regions: Europe is expected to experience an increase in subdaily rainfall extremes, although it has been seen to vary between regions, seasons and duration. For North America and Central America, it is expected to have wetter storms occurring less frequently. Lastly for Asia, it has been seen that while short-duration events are expected to increase, eastern China rainfall intensity seems to decrease by late summer (Westra et al., 2014).

One of the latest studies from Ali et al. (2018) compared the scaling relationship between climates and not continents, and grouping the daily precipitation scaling on the 5 Koppen-Geiger climate classification types (tropical, dry, temperate, continental and polar). Overall, in the paper it was shown that the tropic and continental climates obtained higher median scaling as compared to the dry and polar climates. In addition, the research concluded that, when looking at quantile regression with seasonality, the polar and tropical climates had the largest percentage of stations with super-CC compared to the other regions. For the purpose of this project, such level of detail on the quantile regression approach is not considered.

It is worth to mention that these comparison between continents are the result of projects with differences in the nature of the data and the methodology used. Unfortunately, there are still 4 large regions where there is lack of data available to study extreme rainfall trends: South America, Central America, Africa and Middle East and Southeast Asia (Westra et al., 2014).

2.7. Study Area

For this project 2 climatological regions are used for estimating the scaling rates, in Europe and West and East Africa. In particular, 5 countries are chosen: Germany, Ghana, Tanzania, Kenya and Uganda. In the following section, a brief description of the climate characteristics in the area is given, which will facilitate the interpretation of results and the possible explanation about the spatial distribution of scaling rate.

2.7.1. Germany

Germany is a country located in Central and Western Europe. In western Germany, Atlantic winter cyclones are most relevant for heavy precipitation (Hofstätter et al., 2018). While winter storms are characterised for having large spatial dimension, convective storms occur mostly during the summer due to thermally driven vertical air motions with smaller dimensions (Mudelsee, 2020). Based on the Köppen-Geiger climate classification, Germany has 2 main climate regions, the east characterised for having a cold climate and the west characterised for having a temperate climate with warm summers.

2.7.2. Kenya

Kenya is located in the East of Africa, reaching in the south-east with Indian Ocean. Kenya's climate varies from tropical along the coast, and temperate inland to arid in the north and northeast parts of the country (Figure 2.4), according to the Köppen-Geiger classification (Beck et al., 2018).

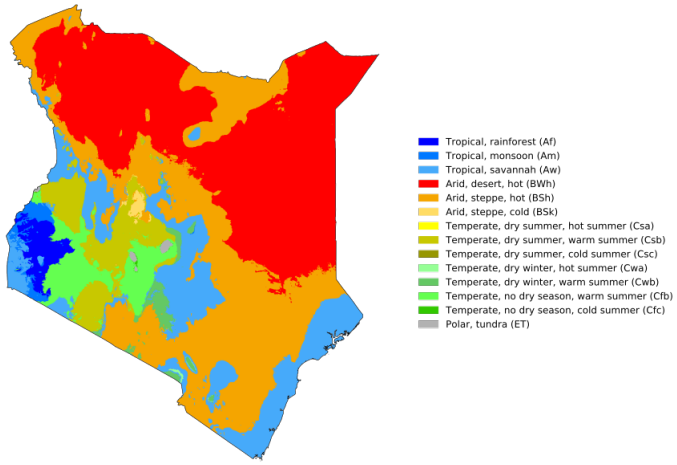


Figure 2.4: Köppen-Geiger climate classification in Kenya between 1980-2016 (Beck et al., 2018).

There are two rainy seasons that can be considered, separated by two dry seasons JJAS (June–September) and JF (January–February) (Mumo et al., 2019). Even though there is a large dry season, research shows that considerable rainfall can occur in the west of the country, characterised by a tropical climate.

2.7.3. Ghana

The Republic of Ghana is a country located along the Gulf of Guinea, only a few degrees north from the Equator. It reaches, to the south, with the Atlantic Ocean. It has a sub-tropical warm and humid climate, and there are six major ecological zones (Figure 2.5), established based on their soil, vegetation and climate (Asare-Nuamah & Botchway, 2019).

With a focus on the rainfall, there are two differentiated zones, a bimodal and mono-modal. The bi-modal zones have a major season that starts from March to July and a minor season from September to October. The ecological zones that it comprises are the Southern-Part of Ghana, namely Rain forest, Deciduous forest, Coastal Savanah and the Central part (Transition zone). The mono-modal season occurs in the Northern Part, Guinea and Sudan Savanah with the wet season during July and September (Asare-Nuamah & Botchway, 2019).

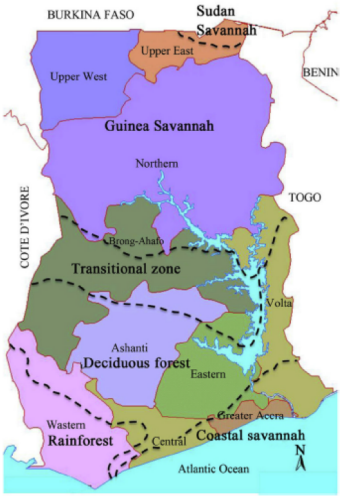


Figure 2.5: Ecological zones in Ghana as described by Asare-Nuamah and Botchway (2019)

2.7.4. Tanzania

Tanzania is located in East Africa, limiting with 2 of the other countries studied in this project, Uganda to the north and Kenya to the northeast. It is in the region of African Great Lakes. According to WFP (2013) and Suleiman (2018), there are two predominant rainfall regimes in which Tanzania can be divided. As Figure 2.6 shows, the bimodal regime occurs in the North, the Northeastern Highlands, the Victoria Basin, Zanzibar Islands and part of the Coastal Belt. For those regions, there are two rainy seasons, ranging from March to May (long rain season) and from October to December (short rain season). The unimodal regime covers the Central, Southern and Western parts of Tanzania, with a rain season that spans from November/December to April (Suleiman, 2018).

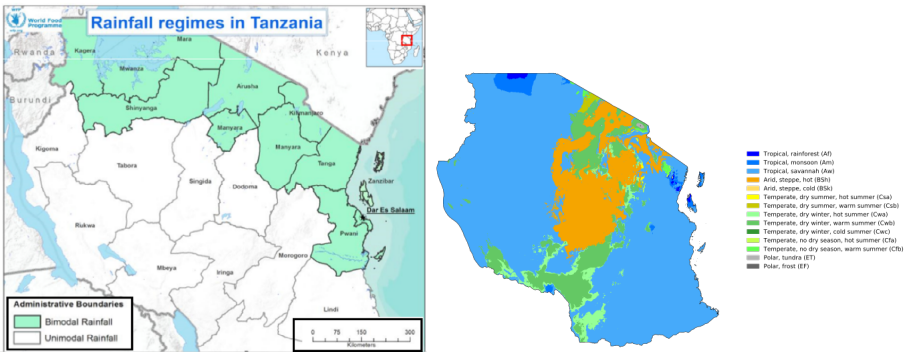


Figure 2.6: Rainfall regime in Tanzania (left) (WFP, 2013). Köppen-Geiger climate classification map between 1980-2016 (right) (Beck et al., 2018).

2.7.5. Uganda

The Republic of Uganda is a country located in the Central-East of Africa. Although it is a landlocked country, there is one of the world's biggest lakes, Lake Victoria, located in the South. Figure 2.7 shows the climate as described by (Beck et al., 2018). Concerning rainfall, there is an unimodal rainfall regime in the northern parts of the country, between June and August (Kisembe et al., 2019).

In this project, there are no stations located in the North. Easter Uganda is located in the north of Lake Victoria and is characterised for having a bi-modal rainfall seasons: between March-July and between September-November (Nimusiima et al., 2019). Near the equator, there is a bi-modal regime with the major season in March–May and a shorter rainfall season in September–November (Kisembe et al., 2019).

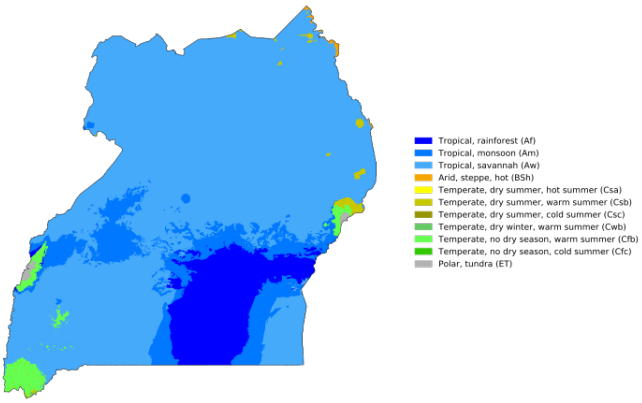


Figure 2.7: Köppen-Geiger climate classification in Uganda between 1980-2016 (Beck et al., 2018)

3

Methodology

3.1. Input data

For this project, precipitation, temperature and humidity data was obtained at different temporal resolutions for Germany, Ghana, Tanzania, Uganda and Kenya. For Part 1 and Part 3, which entails the study and development of different methods to estimate scaling rate, an analysis based on hourly interval was used. For the storm properties analysis (Part 2), precipitation at varying temporal scales from sub-hourly to hourly were used. Figure 3.1 shows the location of the stations in each country, as well as the digital elevation model, obtained from the Global Multi-Resolution Terrain Elevation Data 2010 at 30-arc-second spatial resolution (Danielson & Gesch, 2011).

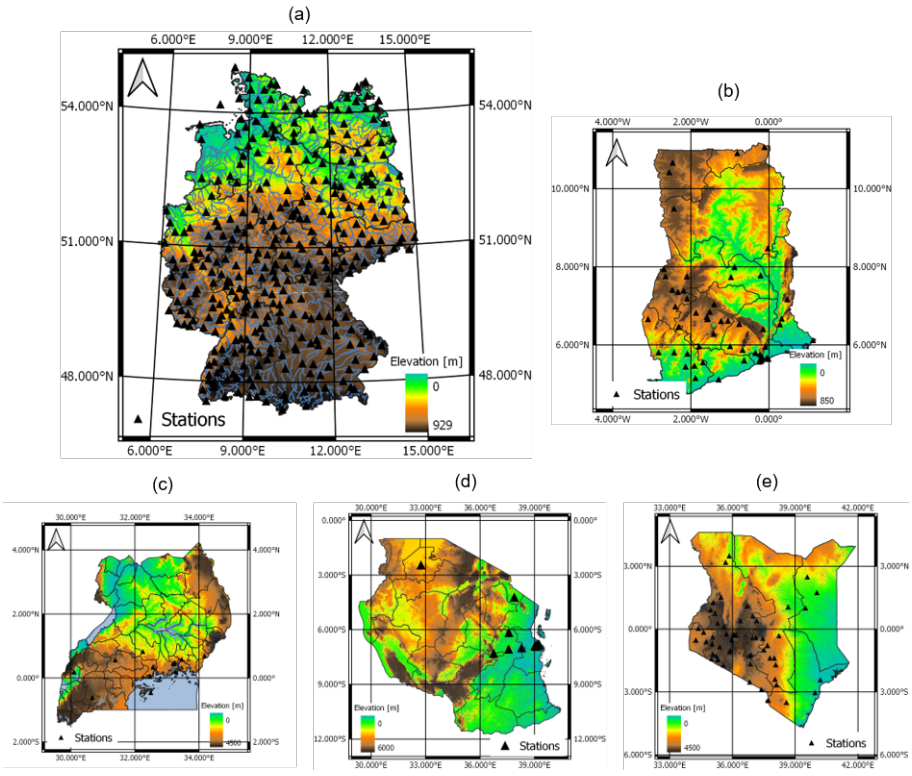


Figure 3.1: Study areas with elevation and location of stations considered in the research. (a) Germany, (b) Ghana, (c) Uganda, (d) Tanzania and (e) Kenya.

Over Germany, quality-controlled rain-gauge data was obtained by Germany's national meteorological service (Deutscher Wetterdienst, DWD), at 10-min and 1-h resolution, covering a period from 1995 until

2019 (DWD, 2018). In total, 515 stations were initially retrieved from the DWD.

The data from the Sub-Saharan Africa countries was obtained from the Trans African Hydro Meteorological Observatory (TAHMO). The TAHMO is an initiative that aims to develop at least 1 station every $1000Km2$ throughout sub-Saharan Africa (van de Giesen et al., 2014). Currently, the project has over 500 stations distributed around the continent. The data available can be up to 5-min resolution. For the development of this project, 10-minute and 1-h resolution data was used. Only the countries with the largest amount of stations and observation length from the TAHMO network were chosen, in this case: Ghana, Kenya, Tanzania and Uganda. For the same locations in sub-Saharan Africa, an additional dataset was included, the Global Summary of the Day (GSOD). This dataset is derived from The Integrated Surface Hourly (ISH) dataset and can be accessed from the National Oceanic and Atmospheric Administration, Department of Commerce (https://catalog.data.gov/dataset/global-surface-summary-of-the-day-gsod). The dataset contains daily precipitation observation and temperature for a set of stations around the world. In Table 3.1, the total number of gauges considered for this project are detailed per country.

Table 3.1: Number of stations used in the quality analysis

	Number of stations								
	Ghana		Tanzania		Uganda		Kenya		Germany
	GSOD	TAHMO	GSOD	TAHMO	GSOD	TAHMO	GSOD	TAHMO	DWD
Available stations	21	99	21	28	15	56	37	136	515

3.2. Quality Analysis

The quality analysis consisted in 4 steps, summarised in Figure 3.2. For some steps, different considerations were taken between Germany and Africa, due to the fact that each dataset had its own characteristics.

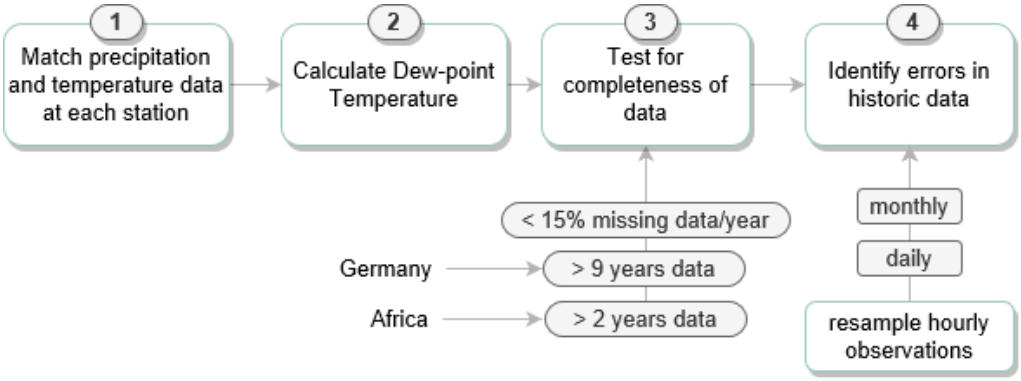


Figure 3.2: Flowchart of the quality analysis procedure

Overall, the steps that were taken were:

1. Compilation of data from available stations: Initially, precipitation and temperature data were combined, given that, for Germany, the temperature and precipitation datasets were in different files. Since not all stations with precipitation records had available temperature datasets, only the stations that had information for both temperature and precipitation were considered. For the sub-Saharan dataset, this step was not considered as datasets already combined precipitation and temperature measurements.

2. Calculating dew point Temperature: For Germany, most of the stations included this variable. However, in some cases, dew point data was missing. Since the relative humidity and temperature was provided, the dew point temperature was calculated both for Germany and African countries. It was estimated following the Magnus formula (Lawrence, 2005), with $A = 17.625, B = 243.04, T$ in (^oC) and RH in %. :

$$T_d = \frac{B \cdot [\ln(\frac{RH}{100}) + \frac{A \cdot t}{B+t}]}{A - \ln(\frac{RH}{100}) - \frac{A \cdot t}{B+t}} \tag{3.1}$$

The results from applying this equation were compared, for Germany, with the available dew point temperatures and they were shown to match.

3. Completeness Test: The third step consisted in filtering out the stations with large missing data on a certain year. Based on existing research, the threshold was set to 15%. For Germany and the GSOD dataset from sub-Saharan countries, years with more than 15% of missing data and stations with less than 9 years of data available were excluded. For the TAHMO dataset, the approach taken was different. Since the maximum amount of years available was smaller (between 3-4 years), it was decided to exclude those stations that had less than 2 years of data, keeping the 15% requirement.

4. Identification of errors in historic data: Once the final selection of stations was done, extreme rainfall values were analysed at different temporal timescales. For that, hourly precipitation data was resampled to daily, monthly and annual precipitation, for which total precipitation values were extracted. These values were compared to existing literature. The most extreme values across countries were checked individually with historic maximum precipitation in the region to ensure that real extreme rainfall was not classified as error data.

On the Results section, an outline of the final stations considered in the project is presented.

3.3. Definition of extreme rainfall indicators

Extreme rainfall can be defined by many different indicators at varying temporal resolutions. The common definition of extremes is based on thresholds, mainly those events that exceed the 90th, 95th and 99th percentile of a cumulative distribution functoin (Westra et al., 2014). For flood hydrology, other indicators such as the annual maxima or the number of days where total rainfall exceeds 10 mm or 20 mm are also common proxies (Westra et al., 2014).

In order to be consistent with existing research in the field and to allow for a comparison of results with previous research, it is decided to work mainly with wet percentiles at hourly and sub-hourly resolution. In addition, they are chosen because working with wet percentiles allows for further insight on the changes of extreme precipitation without the influence of the ”occurrence of events” (Schär et al., 2016).

However, in evaluating the impact of rainfall to society and when addressing future projections of extreme events, using all hours, dry and wet is important (Lenderink & Van Meijgaard, 2008; Ban et al., 2015; Schär et al., 2016). To evaluate the differences between relative percentiles (wet observations) and absolute percentiles (wet and dry) several scenarios are tested:

- Analysis of wet percentiles at 80th, 90th, 95th, and 99th
- Comparison of wet percentiles with all-observations percentiles of the same intensity
- Analysis of 99th percentile for the case of all-observations:

It is important to highlight that, to compare events from a relative percentile against absolute percentiles, the events need to have the same frequency (or return period). For that, some adjustments need to be made, in particular recalculate what would be the absolute percentile to obtain similar extreme events as the relative percentile. Being $q95_{wet}$ the 95th quantile of wet observations, the equivalent percentile for all-day observations can be defined by equation 3.2 (Schär et al., 2016) as:

$$q_{all} = \frac{f_w \cdot q95_{wet}}{100} - (1 - f_w) \tag{3.2}$$

where q_{all} is the absolute percentile that will lead to similar events as the $q95_{wet}$, and f_w is the fraction of wet events. Since the analysis is done with hourly observations, the wet fraction per station is calculated computing the number of wet hours divided by total hours in each dew point temperature bin (Lenderink & van Meijgaard, 2009).

3.4. Part 1: Comparison of regression models

3.4.1. Influence of the length of observations

In existing research, the study of scaling rates is most of the times limited to the number of observations available.In order to check the influence of that in estimating scaling rates, data from Germany was used. While the total number of stations available was 515, not all of them had observations for 25 years. Since the aim was to study how different length of observation influence the scaling rates, only the stations having the maximum length of observations were considered.

Figure 3.3 describes the methodology used for studying this section. For each station, the total length of observations was divided in 6 subsets of 3, 6, 9, 15, 20 and 25 years each. For each station, the scaling rates were obtained under 2 cases:

1. Comparison of subsets with fixed years: For this case, each subset had specific years starting from the most recent observations and going backwards up to 1995 (i.e the 3 year period comprised 2017-2019 and 6 years 2014-2019).

2. Comparison of subsets with random years: To perform the analysis based on random years, for each station and period, the scaling rate was calculated 25 times, each one with a random selection of years. As opposed to before, the selection of years was not consecutive. The only condition set was that the same year could not be repeated on a specific iteration. Given the computational cost of this part, it was decided to run 25 iterations.

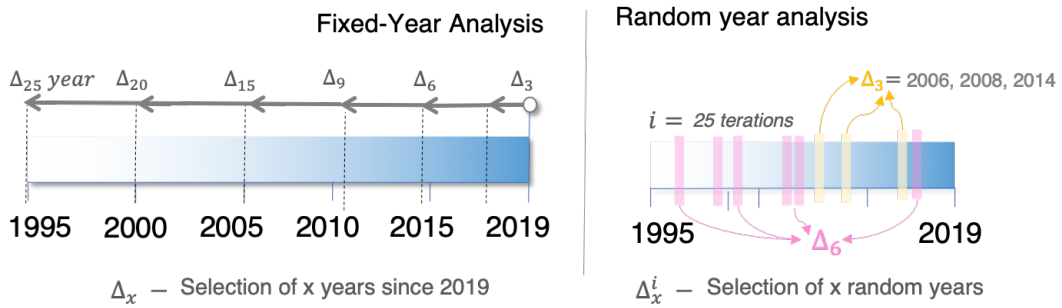


Figure 3.3: Schematisation of the methodology for studying the influence from the length of observations

The analysis based on fixed years was done to have an example to compare with the combination of random years. Also, it was done to check what would be the differences in scaling when the same years are considered. If only the last 3 years of data would be considered (fixed approach), it would not be possible to generalise the differences of scaling rates. For example, over a 25 year period, there are a lot of 3-year combinations, each of which it could lead to higher or lower scaling rate than the period 2017-2019. With an analysis based on random years, the variability of rainfall within different subsets of years is better captured and, thus, conclusions can be drawn from the analysis.

3.4.2. Selection of models

The relationship between extreme rainfall and dew point temperature is estimated by developing 2 types of models: Exponential Regression and Piecewise Regression (Figure 3.4).

Within the exponential regression, 3 models are developed: Binning approach with 2º bins separation (M1), Binning approach with equal length of observations per bin (M2) and Quantile regression (M3). For the piecewise regression, 2 scenarios are considered, Piecewise linear regression for both Binning methods (M4-M5) and a Piecewise Linear Quantile Regression with 1 change-point (M6). The following section contains the characteristics of each model.

3.4.3. Binning Approach

This method consists of grouping precipitation observations in temperature bins. There are several alternatives on how the bins are defined, either fix the number of bins, the distance between them or the number of observations per bin.

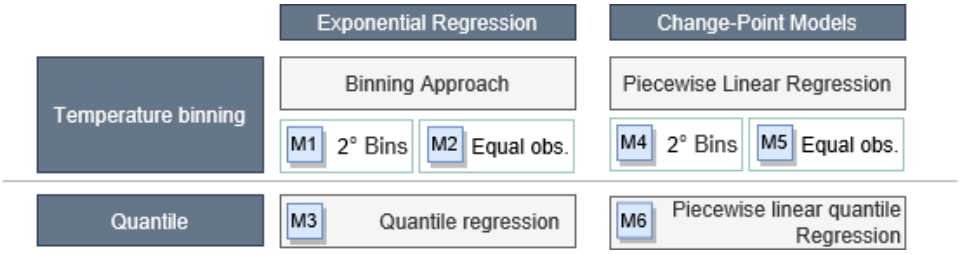


Figure 3.4: Types of models developed and compared in this project

For this project, 2 binning models were developed, one with a fixed distance of 2 degrees between bins and another one with equal number of observations per bin. In both cases, only non-zero precipitation observations of at least 0.1mm/h were considered, with mean dew point temperature for each bin.

The precipitation-temperature pairs for each hour were placed in bins according to temperature. For the method with 2º bins, those are created based on the minimum and maximum temperature from the dataset, whereas for the method with equal number of observations, the number of bins is set so that it matches the same number of bins as the 2º bin approach. Within each bin, the 80th, 90th, 95th and 99th quantile of precipitation is computed. Then, for each percentile, an exponential regression is applied (Eq. 3.3) by fitting a least-square linear regression to the logarithm of precipitation (Wasko et al., 2018).

$$P_2 = P_1(1+\beta)^{\Delta T} \tag{3.3}$$

In Eq. 3.3, P_2 is the precipitation of the last bin, P_1 the precipitation of the initial bin, β the sensitivity of precipitation per degree Celsius temperature change and ΔT the difference of temperature between rain measurements.

For this project, the change in precipitation per increase of temperature is calculated by taking into consideration the regression line between the initial bin (P_1) and the peak point temperature as the last precipitation (P_2), following the approach by Ali et al. (2018). The peak point temperature is defined as the bin where the maximum precipitation is obtained. As Ali et al. (2018); Wasko and Sharma (2014) points out, bins in the extreme temperatures are shown to have a small number of observations. Therefore, fitting to the peak point temperature avoided fitting a regression model with bins of only a few observations (that can have unrealistic precipitation percentiles).

The scaling rate $(\Delta P_q(\%)/^o\mathrm{C}$ for a given quantile (q) is estimated using the exponential transformation of the slope coefficient from the linear regression (β) given by:

$$\Delta P_q(\%)/^o\mathrm{C} = 100 \cdot (10^{\beta} - 1) \tag{3.4}$$

For the binning method where the bins are defined based on equal number of observations, the selection of the total number of bins was made taking into consideration the total number of wet observations and the minimum number of measurements each bin should have to reduce bias. The minimum number of points per bin was determined looking at existing methodologies (Dahm et al., 2019; Schleiss, 2018), that set that the number should be connected to the quantile of study (q) with the following relation:

$$f = \frac{100}{100-\tau} \tag{3.5}$$

Given that 99 is the maximum quantile in this research, a condition was set to have at least f=100 measurements per bins to consider. To determine the final number of bins, the station with the smallest number of wet observations in Germany (n=8131) was taken as a reference. Given that Germany has precipitation observations for an average range of dew point temperature equal to 27, 12 bins were considered. Therefore, the total number of observations per bin would be around 677, satisfying the condition of Equation 3.5. This decision was also made to allow comparison with the previous binning method (that had around 13 bins), with the only difference being the position of the first temperature bin.

The fit the linear model of the binning method the ”lm” function in the statistical programming language R was used (R. Koenker et al., 2012).

3.4.4. Quantile regression

The method of quantile regression estimates the scaling rate by modelling the conditional quantile of a response variable (R. Koenker & Bassett Jr, 1978). Standard linear regression models only consider the conditional mean of the response variable (Wasko & Sharma, 2014).

Considering a pair of data points, (P_i, T_i) for $i = 1, ..., n$, the scaling of high quantiles of $log(P)$ with hourly dew point temperature can be expressed as a linear model:

$$log(P_i) = \alpha + \beta T_i \tag{3.6}$$

For a given set of data points (P_i, T_i), the parameters α and β are obtained by solving the minimization problem (Koenker, 2005):

$$QR\ model: \quad \hat{S} = min_{(\alpha,\beta)} \sum_{i=1}^{n} \rho_\tau(\log(P_i) - QR_\tau(T_i)) \tag{3.7}$$

where $\rho_q(u) = u(q - I_{u<0})$ is the loss function that will be minimized. In quantile regression, instead of minimizing the sum of squared errors, as it would be done in linear regression, the absolute-deviation of the errors is minimized, with a penalty for under prediction (q) or over-prediction $(1 - q)$ based on the quantiles (Wasko & Sharma, 2014).

The quantile regression was carried out using the ”quantreg” package in the statistical programming language R (R. Koenker et al., 2012).

3.4.5. Piecewise linear regression model

For the Binning approach models [M1-M3], a linear regression with 1-change point was estimated. Several papers have set a range of temperatures where break-point can be expected (Dahm et al., 2019; Van de Vyver et al., 2019). In this project, the model was defined to automatically detect whether a change-point temperature (T_c) would be found, without setting any boundary on the range of temperatures where such change-point should be checked.

The piecewise linear regression model was carried out using the R package ”segmented” in R (Muggeo, 2015). In the specifications, it was chosen to look for 1 change-point, and carry out a bootstrap of 1000 replicates.

The following equation shows the change-point model considering 1 transition:

$$PR(T_i) = \begin{cases} \alpha_1 + \beta_1 T_i & for\ \ T \leq T_c \\ \alpha_2 + \beta_2 T_i & for\ \ T > T_c \end{cases} \tag{3.8}$$

where β_1 is the scaling rate corresponding to the CC, T_c the changing-point temperature, and β_2 the one corresponding to the super-CC (>7ºC).

To ensure continuity of the model at change-point, there were 4 unknown, which were connected with the relation: $\alpha_2 = \alpha_1 + (\beta_1 - \beta_2)T_c$.

In Part 1 of this study, 1 change-point was considered, excluding the study of change-point on the very high temperatures. This was decided because research shows that, for the highest temperatures, decay occurs due to statistical artifact of small samples Boessenkool et al. (2017).

3.4.6. Piecewise linear quantile regression

The Piecewise linear quantile regression estimates simultaneously the scaling slope and a change-point. The methodology followed in this project is based on Tomal and Ciborowski (2017). The procedure consists of:

1. An initial piecewise linear regression model is calculated using the R package ”segmented”. The model looks for a breakpoint temperature. If a change-point is found, the change-point and the obtained parameters of both linear regressions are stored. It is important to highlight that this part

aims to obtain initial parameters that will be used as input for the piecewise quantile regression. Up until this step, the whole dataset is used, without any restriction based on quantile

2. The second step uses the initial parameters obtained from the linear regression model as input for a quantile regression model defined in the R package "quantreg" where, for each quantile, the final parameters are estimated.

For a Piecewise linear quantile regression, the solution of the minimization problem is defined by:

$$PR\ model: \quad \hat{S} = min_{(\alpha_1,\beta_1,\beta_2,T_c)} \sum_{i=1}^{n} \rho_\tau(\log(P_i) - PR_\tau(T_i)) \tag{3.9}$$

In the piecewise linear quantile regression, there is no automatic way to estimate the parameters through the "segmented" package. To estimate the change-point, Van de Vyver et al. (2019) calculated the scaling through quantile regression for a range of T_c values, and chose the value where the minimization of the problem was minimum. Dahm et al. (2019) did the study for a location in The Netherlands where the expected change-point had already been studied in a previous research.

In this research, one of the aims of this first part is to evaluate of a change-point is detect within the middle range temperatures. To avoid the obtaining change-point caused by the jump of data at the beginning of the temperature range (due to undersampling of very small events), the analysis was set to look if a change-point would occur at dew point temperatures around 8 and 17ºC.

3.5. Comparison and validity of regression models

In total there are 6 models that were compared. For that, several statistics and information criteria methods were chosen:

3.5.1. Confidence Interval

To study the uncertainty related to the parameter estimates, the confidence intervals for each regression models are generated through the bootstrapping method, which re-samples data to estimate the sampling distribution of the maximum likelihood parameter estimates, without requiring the sampling distribution to be known. This methodology has been tested in several studies (Lenderink & Van Meijgaard, 2008, 2010; Berg et al., 2013; X. Zhang et al., 2013; Lenderink et al., 2014; Lenderink & Attema, 2015; X. Zhang et al., 2017; Pumo et al., 2019; Van de Vyver et al., 2019; Li et al., 2019). For this project the 95%-confidence intervals are computed, from 1000 bootstraps. In the case of sub-Saharan Africa, where the number of observations is smaller, 200 bootstraps are calculated.

3.5.2. Binning approach comparison

To compare between the Binning Approach methods, the goodness-of-fit of the model (R^2) was used. However, given that each binning methods was defined by different number of bins, different alternative indicators were also considered. A visual comparison between both methods was also chosen, in order to compare not only the number of bins and the shape of the regression, but also the underlying percentile values that were taken as a reference for measuring the quantile. This can help to understand what improvements could be proposed for such models.

3.5.3. Quantile and piecewise regression comparison

In this project, the main interest lies in understanding how strong is the presence of a change-point shifting from CC to super-CC at the mid-temperature range. To evaluate that, some change-points obtained from the different models have to be excluded, mostly those associated with change-point detected at the extremes. For that, a list of criteria are defined:

- Criteria 1: Change point at extreme temperatures: It is possible that, due to data characteristics, some models identify a changepoint around 0ºC or in the upper temperature extreme. Looking in detail to the reasons behind that, it is seen that the change-point is associated to the statistical effects, due to a jump from very small observations to a larger sample. For that, the change-points between 0-4 degrees were be excluded from the change-point (after validating whether

those identified change-point occur due to statistical effects). As described in the Background, there are several changes of slope that can occur, generally from CC to super-CC, but, for larger temperatures even a decay with negative slope has been identified. The extreme temperature change-points were excluded from the analysis because they are mostly associated to under-sampling of data. The first criteria filtered those stations that have a change in slope sign, which could occur in 2 situations: (i) at low temperatures where there is a transition from negative to positive slope or (ii) at large temperatures there is a transition from positive to negative slope.

- Criteria 2: Performance of the change-point model against the models without change-point: For quantile-based models, there is no R^2 available to compute. In addition, as it has been pointed out in this chapter, R-squared is sensitive to the inclusion of additional variables in the model, which does not make it the best metric to compare change-point models. Therefore, it is decided to use unbiased estimate of model performance, such as Bayesian Information Criterion (BIC).

The BIC statistic can be defined as -2 times multiplied by the log of the maximized value of the likelihood function of the model ($\hat{\mathcal{L}}$) plus the logarithm of the number of data points (n) multiplied by a penalty for the number of parameters fit (p) (Equ. 3.10). On of the motivations for using this metric is because it has already been tested in a previous research (Van de Vyver et al., 2019). In addition, it incorporates a penalty for the number of parameters, which the goodness-of-fit statistic does not have.

$$BIC = -2\log\hat{\mathcal{L}} + \ln(n)\, p \tag{3.10}$$

From this method, it is considered that the model with the lower BIC is the preferred. This project will look at the number of stations where a change-point model is preferred and, after accepting or rejecting the result based on the acceptance criterion, the results will be analysed.

3.6. Part 2: Analysis of storm event properties

3.6.1. Defining a storm event

A storm event is defined based on continuous measurement of 10-minutes rain with allowed interrupting intervals (intermittency). In existing research, the duration of no-rain taken between storm events can range between 2 and 6 hours. This allows 2 storms to be considered independent (Molnar et al., 2015). For this project, storm events were considered when there was at least 3 hours of no-rain in-between wet observations. A wet measurement was considered when the rain measurement was > 0.1 mm/10-min. This was chosen because the minimum measurement provided by the rain gauges in Germany was 0.1 mm/10-min.

In addition to filtering what was considered wet or dry measurement, 2 more filters were applied to exclude storm events with very short duration. Since the data available was given at 10-minute resolution, it was decided to exclude all storm events that did not last for at least 30 minutes. That way, the storm events could be represented by at least 4 rain observations, and be more accurately described as opposed to very short events that for which only 1 or 2 observations at 10-min was available.

3.6.2. Algorithm to retrieve storm events

In order to extract the storm events, and the associated storm characteristics needed, an algorithm is developed using the programming language R (R Core Team, 2017). The flowchart in Figure 3.5 describes the steps taken to extract storm events

3.6.3. Storm event properties

There have been several studies looking at storm event properties, each one using different predictors and precipitation variables. For this project, 4 variables have been selected, with the aim to get the main characteristics from storms: Mean intensity during the event, the duration, the total rainfall depth and the maximum 10-min rainfall.

To look at the relationship between each of these variables with temperature, dew-point observations are used. In order to make sure that the temporal resolution of each storm variable and temperature match, it is decided to use 2 dew point temperatures. For mean intensity, total rainfall depth and

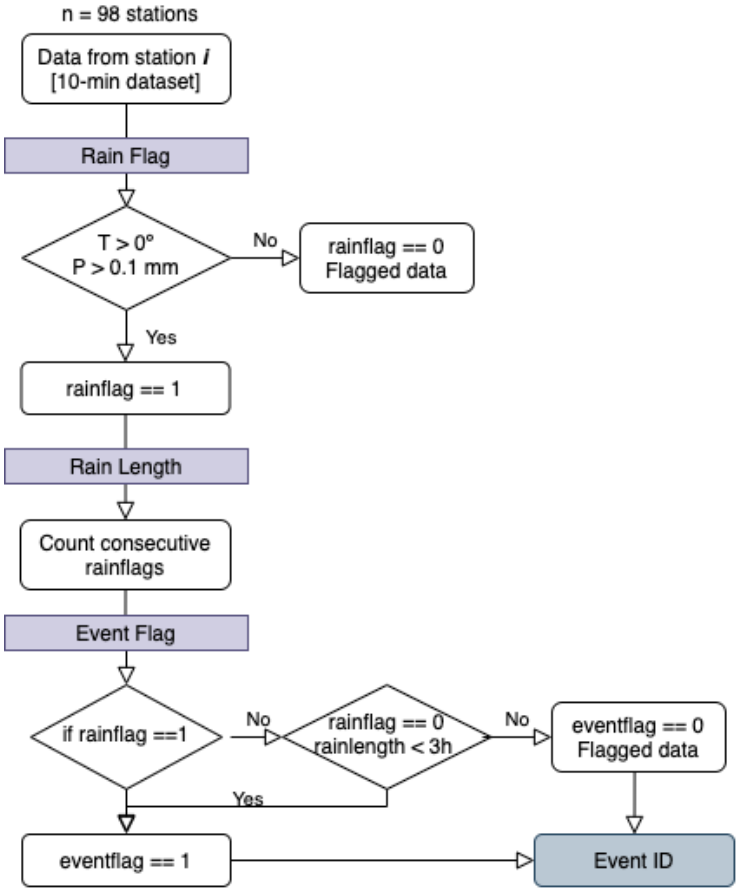


Figure 3.5: Flowchart with the algorithm developed for retrieving storm events.

duration, the mean dew point temperature during the event is used. In the case of maximum 10-min precipitation, which is given in mm/10-min, the temperature is provided as the maximum dew point temperature during the event.

Out of the 4 variables of analysis, the maximum 10-min rainfall is the only one where precipitation is fitted with dew-point at a 10-min measurement. Given that these temperature can change more than the mean dew point during the event, it is important to check its sensitivity for a relationship with precipitation. To put this into practice, 4 temperatures will be extracted from each storm event: the maximum temperature during the event (baseline used for the main analysis), the temperature at the start of the event, the maximum temperature when the precipitation peak is reached and the temperature 4 hours before the event started. The last temperature observation is chosen because, on previous research by (Lenderink et al., 2011; Dahm et al., 2019), they found the best relationship between hourly precipitation intensity and dew-point temperature occurred when the dew point temperature values 4h before the observed precipitation were considered.

Through goodness-of-fit, each predictor will be compared to establish what temperature of a storm best matches with the maximum 10-min rainfall.

3.6.4. Storm event properties conditioned to duration

One of the main hypothesis on why super-CC occurs is, as presented in the Background, because there is a shift from large stratiform precipitation at low temperatures to a short but extreme convective events at large temperatures. Given that all storms in a region will be available for a period of 25 years, this hypothesis will be tested by evaluating the proportion of different duration events at each temperature bin. To determine the lengths of the duration classes, 2 considerations are made, one based on existing research and also according to the histogram of rainfall duration of Germany dataset.

In a research by Panthou et al. (2014), rainfall events in Canada are split in 3 groups, short, intermediate

and long duration events, for which their proportion at each temperature bin is evaluated. While the duration of the event does not directly refer to convective or stratiform events, it can be a proxy, since convective storms are normally of short duration and stratiform rainfall can last hours. 3 classes were chosen instead of 2 because there are many rainfalls that are neither convective or stratiform, but rather a mixed combination. Figure 3.6 (left), shows the distribution of rainfall duration commonly observed for Germany. This, together with existing research (Panthou et al., 2014) was used to define duration classes: A short event was considered from 0 - 2 hours, an intermediate duration between 2 - 10 hours and a long event if it lasted more than 10 hours. To be able to study in-depth the outcomes of constraining the scaling analysis on duration, a subset of 6 stations was selected, which covered both coastal and inland stations, from all parts of Germany and with varying scaling rates. The idea behind is to have a group that can represent potential differences in rain distribution. Figure 3.6 shows the stations that were chosen:

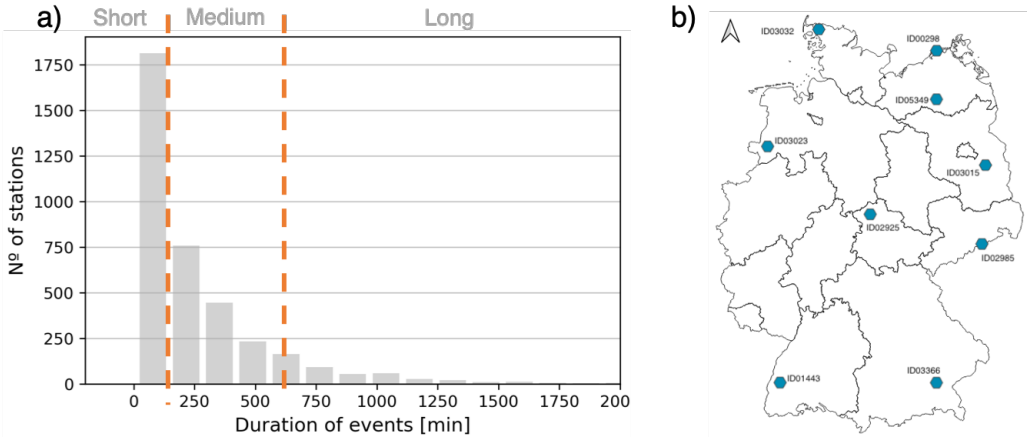


Figure 3.6: Histogram of rainfall duration for a common station in Germany (a). Selection of 9 stations where the storm event analysis will be carried in detail. Stations chosen based on different geographical regions, inland and coastal and with varying scaling rates (b).

3.7. Part 3: Scaling analysis in sub-Saharan Africa

The methodology to estimate the scaling rates in sub-Saharan Africa followed similar steps than in Part 1. For sub-Saharan African countries, the 3 linear regression models tested in Part 1 were implemented: Binning approach with 2º bins, binning with equal number of observations and quantile regression. While the method is the same than with Germany, some aspects were changed to fit the data characteristics of the TAHMO stations.

First, the range of dew point temperature for which precipitation occurs was obtained for all the stations in each country. The average per country was used, together with the number of observations per station to establish the approximate number of bins that would be required to have enough observations in each bin of the binning method with equal number of observations.

To begin with, the study first looked at the scaling rate of the 4 stations with the largest number of observations. After the initial results were compared, the methodology was refined and applied to all the remaining stations. Specific modifications that were carried out are described in the Results section. For the study of scaling rates in all stations, only the quantile regression was tested.

In addition to obtaining the scaling rates, and to evaluate the underlying causes of the scaling results, for all stations, the relationship between hourly relative humidity and surface air temperature was extracted. This was done to evaluate whether there is a decay in the relative humidity for larger temperatures as previous research has shown in other locations in the tropics. Lastly, the same steps taken for the TAHMO dataset were applied to the daily observation stations from the GSOD. For those stations,

4

Part 1: Analysis of scaling rates methodologies

4.1. Quality analysis and final selection of stations

In this chapter, the results from the quality analysis described in the Methodology section are presented.

For Germany, from an initial 515 number of stations available, 29 were excluded because they had less than 9 years of data. In addition, 6 stations had to be excluded because there were large gaps of data. Therefore, 478 stations were considered in the project, ranging between 9 to 25 years of data. For the 4 sub-Saharan Africa countries, the results from the quality analysis for both the hourly observations of the TAHMO (Table 4.1) are presented:

Table 4.1: Final number of TAHMO stations considered in the project

Number of stations with an estimated change-point				
	Uganda	Kenya	Tanzania	Ghana
Total n^o available stations	56	136	28	99
Stations with no measurements	9	8	3	4
N^o stations with < 15% missing data	13	91	20	59
N^o stations with >= than 2 years of data	13	71	15	59
Excluded stations due to errors in the datasets	-	7	1	1
Final stations considered	**12**	**56**	**14**	**58**

4.2. Influence of length of observations on scaling rate

Existing research has shown that super-CC is specially relevant for sub-daily observations (Lenderink & Van Meijgaard, 2008). Normally, having a large quality-controlled dataset of hourly observations is a challenge (Westra et al., 2014). Therefore, it is vital to know what is the sensitivity of the scaling rate to various lengths of observations. Before proceeding to compare the scaling results obtained from different methods, this section presents how much can the scaling rates change when an increasing amount of years of data is provided. This is done for the stations in Germany and the comparison is made based on results for applying a quantile regression. This model was chosen for this analysis because it is unbiased by the number of observations available (Wasko & Sharma, 2014).

4.2.1. Scaling differences with fixed selection of years

The results from calculating the scaling rates with fixed years are shown in Figure 4.1. For all the quantiles analysed, as more years of observations are included, the range of estimated scaling rates reduces. This is specially visible in the 99th quantile, where for 3 years, 50% of the stations have scaling rate between 9-12%/ºC, opposed to the 25 years where it goes from 7.7-9%/ºC.

When evaluating the 80th quantile, one can see that the scaling rates between 3 years and 25 years

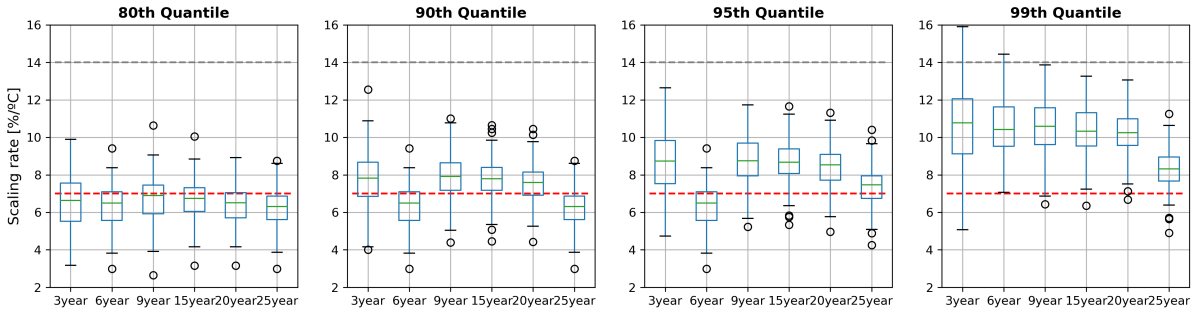


Figure 4.1: Boxplot of scaling rates obtained for quantiles 80th-99th through the quantile regression with fixed subsets of data for 3, 6, 9, 15, 20 and 25 years.

are more similar than comparing them for to 95th and 99th quantiles. This could indicates that the impact of shorter length of observations is more acute on extreme rainfall events (which are also the more relevant in climate studies). One possible explanation of this is because less extreme events occur more often whereas extreme events are more an exception, which means that there are less number of observations compared to smaller intensity rainfall.

For higher quantiles, Figure 4.1 reveals that, for fixed years, shorter years of data lead to higher quantiles compared to the 25 years. This occurs in all cases except the 6-year subset at the 95th quantile where smaller results are observed. When looking at the 80th and 99th quantile, the results of scaling rate with 6 years do not deviate from the rest of subsets. The fact that it only occurs for the 90th and 95th percentile might be because of the specific rainfall characteristics of the years analysed. To evaluate this hypothesis, the same analysis is done but changing the 6 years (which in the fixed case is between 2014-2019), with different 6 years.

The result of the new boxplot for the 95th quantile (Figure 4.2 shows that, no conclusions on the actual changes of scaling rates can be extracted from this comparison, since they are dependent on the years considered. In order to infer a statistical conclusion on the influences of the scaling rate and the actual order of magnitude of differences between 3 years and 25 years, it is necessary to add randomness to the analysis.

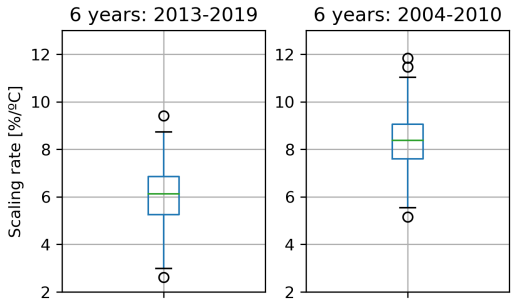


Figure 4.2: Boxplot of scaling rate at the 95th quantile for different subsets of 6 years

4.2.2. Scaling differences with random selection of years

In this section, the scaling rates presented are the result of iterating, for 25 times, the years that each subset uses for estimating the scaling rate.

In Figure 4.3, the results for the 80th, 95th, and 99th quantile show that, when random years are considered, the range of scaling rates is way larger than in the previous section. In particular, it can be seen that, as one can expect, for the 3 years data the range of scaling rates is the largest from all the years. The more years of data are included, the shorter the range of scaling rates are. This is similar to what the fixed-years analysis showed. The range of values is larger because different sets of years are considered. For example, in the previous section, the comparison between 3 and 6 years means that the years compared are between 2017-2019 and 2014-2019 respectively. Since they have half of

the years in common, the differences are smaller than in the random analysis.

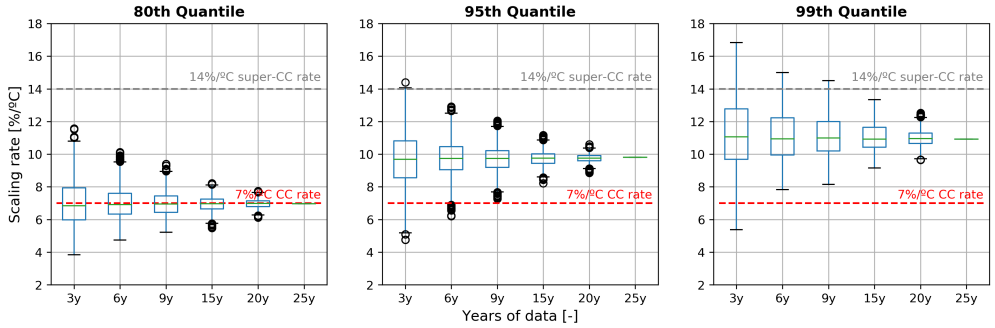


Figure 4.3: Boxplot of scaling rates obtained for quantiles 80th-99th through the quantile regression with random subsets of data for 3, 6, 9, 15, 20 and 25 years.

In the following figure, the boxplot of the difference between the scaling rate at 25 years versus the ones obtained by the subsets is presented (it is worth to note that this is done for each station and iteration, to make sure that the difference calculated corresponds to the same stations). With a focus on the quantiles, the boxplot show that for higher quantiles, the range of scaling rate differences is larger. For instance in the case of 6 years of data, the 80th quantile has an interquarlile between -0.8 and 0.5 %/ºC whereas for the 99th quantile it is between -1.5 and 1%/ºC.

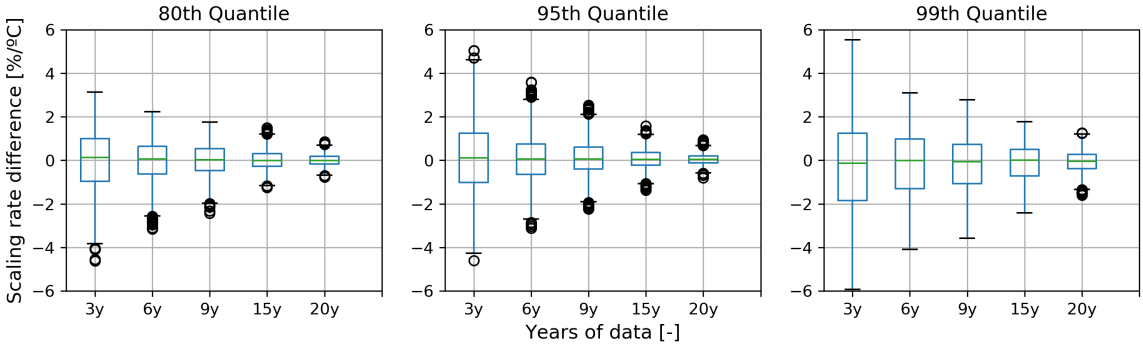


Figure 4.4: Boxplot of scaling rate differences between 25 years and different subsets (3,6,9,15 and 20).

Comparing the influence of the number of years, results show that with more years of data, the scaling rate difference reduces. In the most scarce scenario (3 years of data), half of the stations show a scaling difference between -1 to 1 %/ºC (80th quantile) up until -2 to 1.7 %/ºC (99th quantile). For some of the iterations computed, some scaling rates differed with 25 years up until -4/3.7 (80th quantile) and until -6 and 6 (99th quantile).

Exploring the results for the 99th extreme rainfall percentile (the quantile with the largest differences), it is shown that up until 9 years of data, the scaling rates of most of the stations differ with the 25 year data 0.8 - -1%/ºC. In addition, although in the boxplot is not visually evident, in general, for the largest quantiles the scaling rate difference is has more negative than positive results. This means that smaller datasets tend to overestimate the scaling rates compared to larger samples.

In an attempt to visualise the reason behind larger scaling rates with small observations, Figure 4.5 shows, for the 95th quantile, the observations available at different subsets of years for a random station. As it can be see, for a log precipitation of 0.50 mm/hour, the number of observations remains constant within dew point temperature However, the higher the quantile, the less distributed the precipitation are across temperatures. This with the fact that the largest rainfall events are observed in the highest dew point temperature might explain why for smaller datasets larger scaling rates are obtained.

Following the results in this section, it was decided to limit the number of stations form which the models would be applied to. In order to reduce as much as possible the influence of a small length of observations on the overall scaling rates, it was decided to exclude all the stations that had less than

15 years of data. In total, there are 353 stations (out of the initial 478) with at least 15 years, which is the number that will be considered from now on.

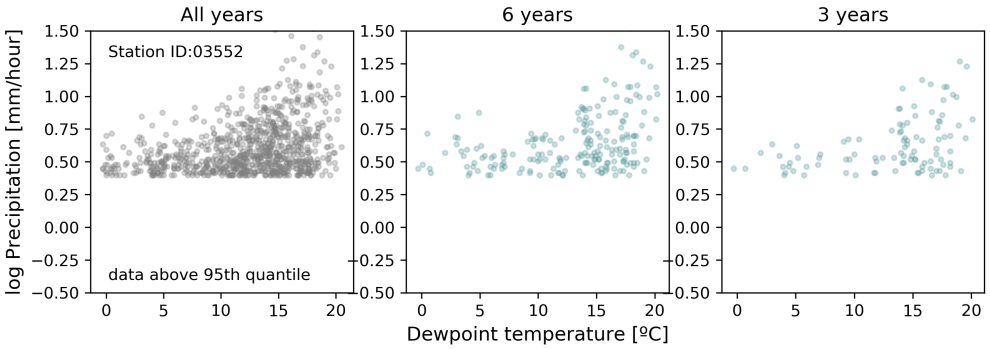


Figure 4.5: Rainfall observations above the 95th quantile for different subsets of data for Station ID03552.

4.3. Elevation influence on the scaling rate

Existing research has shown that the dynamic contribution of orography can play a role in precipitation extreme scaling (Drobinski et al., 2016). A study by Schroeer and Kirchengast (2018) in Austria showed that temperature sensitivities in the mountainous western region are lower than in the eastern lowlands. One way to test whether this hypothesis holds for Germany is to look at where the stations are located. While in Schroeer and Kirchengast (2018) regions with strong orographic differents are already given, the approach taken in this project is to test whether looking at the elevation of the stations alone can provide accurate representation of orographic properties.

With the results obtained from the quantile regression analysis and before comparing the results with other methods, it was checked whether there was a correlation between the scaling rate and the elevation. Figure 4.6 shows the scaling rates with increasing height, ranging from 0 meter to 1400 for all the quantiles. The results from this analysis shows that there is no apparent correlation between these variables, at least not by only looking at the elevation of the stations. This means that more details on the orographic characteristics might be needed, such as the slope of a mountain where a station is located, the proximity to the sea or even the surrounding characteristics (urban or rural area).

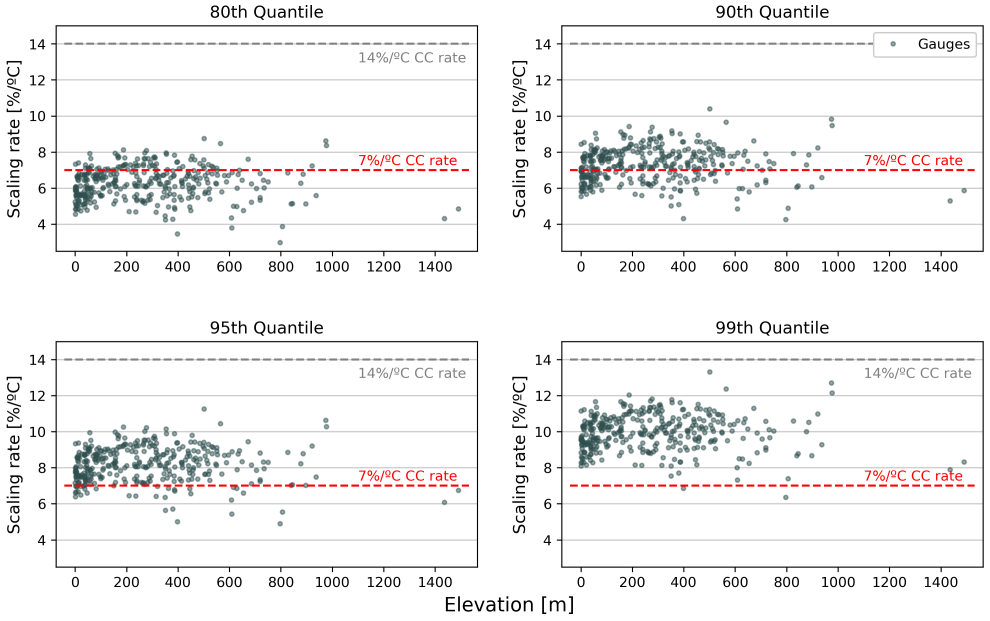


Figure 4.6: Scaling rate obtained relative to the elevation

4.4. Exponential Regression Models

The following section presents the results from 3 models, two of which are based on different temperature binning techniques (2º bin separation and equal samples per bin) and one based on quantile regression. As described in the Methodology, the exponential regression models are characterised for computing the linear regression using the log of precipitation.

4.4.1. Binning Temperature

The binning approach is used to analyse the 80th, 90th, 95th, and 99th precipitation quantiles through 2 scenarios: One which defined temperature bins are given every 2 degrees and one based on 12 bins with equal length of observations (details on the Methodology).

Table 4.2 shows the results for the Binning Approach with a 2 bin separation. The results are based on the stations that scored at least an $R^2 = 0.70$ (341 stations). The R^2 was set to 0.7 to filter out the worst estimates and set a minimum threshold. Figure 4.7 shows that, for all the quantiles, the interquartile range has always values larger than the 7%/ºC. In addition, the range of values of scaling rates is large, with the 80th quantile showing scaling rates between 5-15%/ºC.

Table 4.2: Median scaling rate per Quantile for the Binning approach with 2º bins

	Quantile			
	80th	**90th**	**95th**	**99th**
Median Scaling rate	9.5	10.9	11.8	14.0

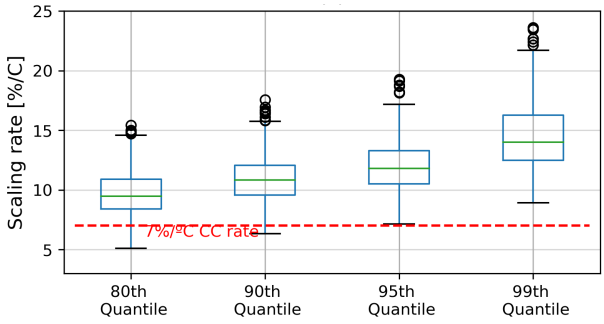


Figure 4.7: Boxplot of scaling rate results applying the binning approach with 2º bin separation

Checking the distribution of scaling rates, it is observed that, for the larger quantiles such as the 99th, there are some very steep scaling slopes that reach above 20-23%/ºC Figure 4.7. In order to better understand the underlying cause of these scaling rates, Figure 4.8 shows 3 illustrative examples from stations that recorded scaling values larger than the median.

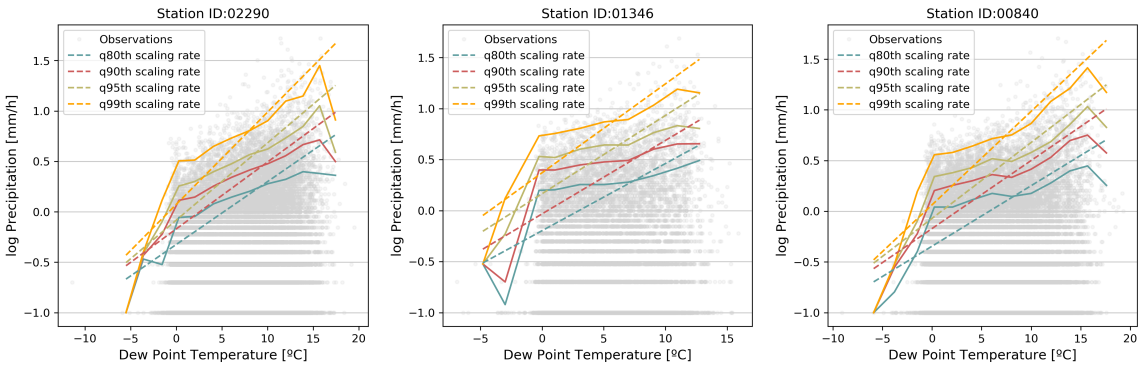


Figure 4.8: Application of the Binning approach for 3 stations with large scaling rates. The dashed line represent the estimated scaling rates with the 2 degree binning method and the solid lines the quantiles of binned data.

An inspection on the initial temperature bins showed that, for some stations, there were very few observations available. For example the stations shown in Figure 4.8, have in the first bins 4, 15, 13 observations respectively. This initial bins with low number of observations are not representative of

the overall precipitation range observed for the station (as it can be seen, that happens around 0ºC dew point temperatures).

However, as the model includes these bin in the analysis, it causes the slope of the linear regression model to increase, reaching higher scaling rates than expected. As mentioned in the description of the models, negative surface air temperature observations were excluded from the analysis, to avoid including snow into the analysis. Excluding those temperatures does not mean the dew point temperatures will be all above 0ºC, which could occur in winter.

To account for the small number of observations at the initial temperature bins, an improved version of this model was proposed and developed, based on only selecting as the initial bins, the one that has at least 60 measurements in it. This number was chosen to be close to 100 (which, a mentioned in the methodology is the minimum number that we would need to have a consistent analysis of the 99th percentile), but at the same time not restrict the selection of the bin by setting up a very large threshold. By keeping it at 60 measurements, the model can capture the progression towards more observations for higher temperatures while not excluding information at lower temperatures. The main goal is to make sure that the regression model accurately uses an initial bin that contains enough observation points. Results from the corrected binning approach are presented in the following Table, as well as in Figure 4.9:

Table 4.3: Median scaling rate per quantile for the corrected Binning approach with 2º bins

	Quantile			
	80th	**90th**	**95th**	**99th**
Median Scaling rate	7.7	8.9	9.6	13.4

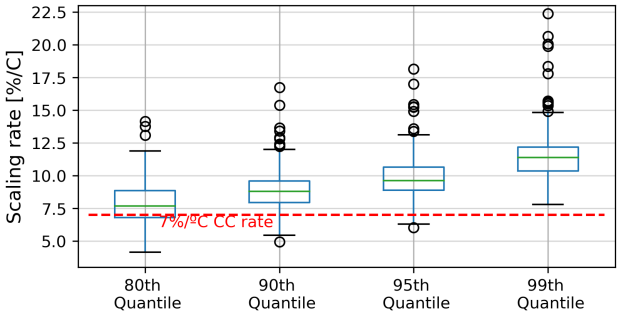


Figure 4.9: Boxplot of scaling rate results applying the corrected binning approach with 2º bins

With this model, the median scaling rate results for the 80th quantile reaches a scaling slightly higher that the Clausius-Clapeyron 7.66%/ºC. Compared to the initial binning approach, the median scaling is reduced for each quantile at least 2%/ºC. Another aspect to compare is that the interquartile range of this model is reduced than in the first case, meaning that, from the first and third quartile (25%-75%), the scaling rates values do not differ as mutch as those in the binning approach with 2º bin separation.

In addition, the overall performance of the 2 models is investigated using the goodness-of-fit, which also reveals that the corrected model yields better results, with its R^2 all above 0.920, whilst for the original binning model the R^2 ranges between 0.857-0.878 (Table 4.4).

Table 4.4: Goodness of fit comparison between original and corrected model

Quantiles	2^o bins R-squared	Corrected 2^o bins R-squared
80th	0.857	0.920
90th	0.871	0.925
95th	0.878	0.929
99th	0.863	0.925

Given the better performance and accuracy of the scaling rate estimation, it was decided to continue the analysis using the corrected model.

In the following table, an overview of the median scaling rates estimated from the Binning Approach with equal number of observations is presented, including as well the goodness-of-fit and the extreme scaling rates estimated. For this model, 12 bins of equal observations were assigned, with a total median number of wet observations per bin of 1292.

Table 4.5: Results of the Binning Approach with equal number of observations

	Median scaling rate [%/ºC]	R-squared	Maximum scaling [conf. interval]	Minimum scaling [conf. interval]
80th	6.34	0.936	9.00 [8.24-9.77]	3.04 [2.19-3.89]
90th	7.40	0.934	10.30 [9.60-11.02]	4.33 [3.48-5.18]
95th	8.38	0.926	11.54 [10.67-12.42]	5.02 [3.89-6.16]
99th	10.24	0.901	13.28 [11.73-14.94]	6.78 [5.34-8.24]

It can be seen that the median scaling rates are slightly smaller than the previous binning approach. For the 80th quantile, a sub-CC of 6.34%/ºC is observed. For the 99th quantile, the median scaling rate is 10.24%/ºC. The results from the other quantiles show values larger than the CC relationship (Figure 4.10).

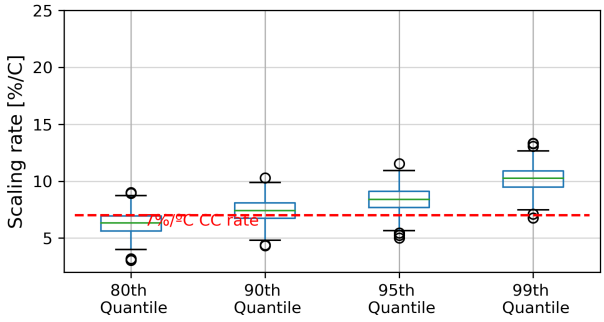


Figure 4.10: Boxplot of scaling rate results applying the binning approach with equal length of observations

Checking the extreme scaling results, it is observed that, for the 99th quantile, there is one station with a scaling rate reaching almost two times the Clausius-Clapeyron relation (α=13.28%/ºC). The goodness of fit of the model for the 80th and 90th quantile is slightly higher than the corrected 2ºC binning approach. In Figure 4.11, the specific scaling rates of 2 stations that have the minimum and maximum scaling rates are shown, in order to visually inspect whether the differences are associated with statistical effects of the method or due to other causes.

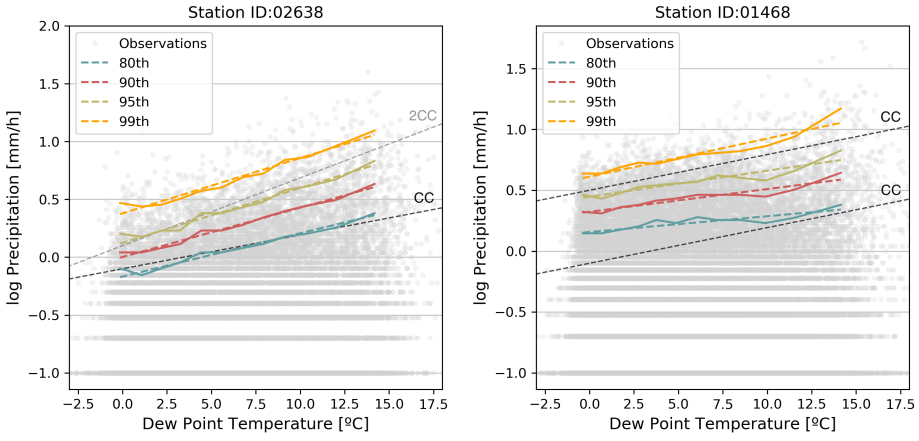


Figure 4.11: Comparison of 2 stations where the maximum scaling are observed for the equal number of observations (left) and minimum scaling rate (right)

As it can be seen, the station where small scaling rates are observed (ID01468), the number of observations between dew point temperature 0º and 5ºC dew point temperature is a lot larger than in

the case of station ID:02638. This means that the precipitation distribution between stations vary, with station ID01468 having more extreme rainfall at lower temperatures than for station ID02638. Overall, this shows that in Germany there might be regions with large differences in precipitation distribution, which translates to large variation in scaling rates across stations.

4.4.2. Quantile regression

The scaling rates obtained through the quantile regression are shown in Figure 4.12. Scaling were considered statistically significant at the 95% level. A t-test is carried out to test whether the slope of the regression line is different from zero under the null hypothesis that the slope is zero. This analysis is done following existing procedures for checking quantile regression significance (Wasko & Sharma, 2014). From the 353 stations, all showed to be significant.

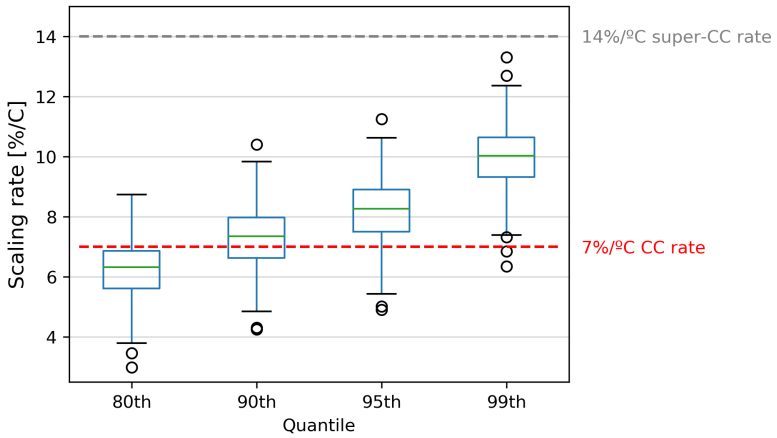


Figure 4.12: Boxplot of scaling rates (α) obtained using quantile regression.

Overall, the analysis of results shows that, for the 80th quantile, 2/3 of the stations have a scaling rate smaller than the CC. For the 90th quantile, the number of stations below CC is reduced to 1/3 of the stations. In addition, for the higher quantiles, scaling rates vary from 8 to 12, with none of the stations reaching 2 times the Clausius Clapeyron relation. Checking the confidence interval of the results showed that they were larger for higher quantiles. This was expected and could be attributed to the fact that there are less data observations for the higher quantiles.

Similarly to what was shown in the Binning Approach section, Figure 4.13 shows 3 illustrative examples of the estimated scaling rates with the quantile regression. In the figure, the binning approach with 2 bins is also included for reference.

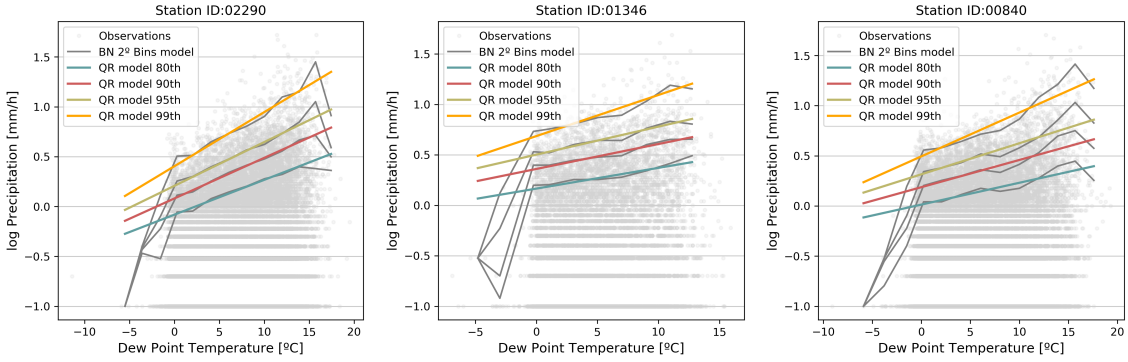


Figure 4.13: Application of the Quantile regression (coloured lines) for 3 stations with large scaling rates. Grey solid line represents the quantile of binned data for the 2 degrees binning approach.

It can be seen that when using the quantile regression, the extreme values in the early bins do not have any influence in the overall scaling, as compared to Figure 4.8. The fact that the quantile uses all the observations to obtain the regression slope instead of just the results in each bin like the binning

approach results in a better adjustment to the bulk of precipitation observations, being less affected to outliers.

4.4.3. Comparison of regression models

In this section results from the Binning Approach and Quantile regression with no change-point are presented. Figure 4.14 shows the scaling rates estimated from each model. In general, the binning approach with 2ºC bin separation has the largest scaling rates. As it was proved in the individual analysis of results, it was caused because of the small number of observations in the initial bins. Once this was corrected for the binning approach with 2º bins, it can be seen that the estimated rates are reduced and closer values to the other methods were obtained. Table 4.6 shows the median results between each method.

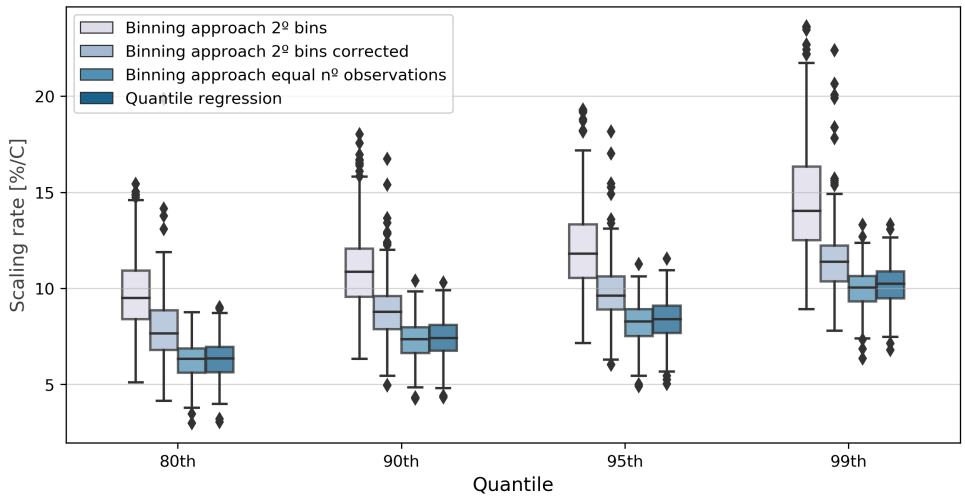


Figure 4.14: Boxplot of scaling rates estimated with the 3 exponential regression approaches. The corrected binning approach is also provided.

The values of the corrected binning approach still show larger rates when compared with the equal number of observations models. Checking the initial bins of such method shows that, while the corrected binning method excludes initial bins when not enough observations are found, it can still occur that a bin has a small number of observations but still passes the minimum threshold (for example the bins that are closer to main bulk of data). It is thought that this is why, although the corrected binning approach works better than the original binning approach, scaling results were still larger than the other methods.

Table 4.6: Comparison of median scaling rates between exponential regression models.

	Median scaling rate [$\%/^{o}C$]		
	Binning Approach		Quantile regression
	2^{o} bins (corrected 2^{o} bins)	Equal nº of observations	
80th	9.49 (7.66)	6.34	6.31
90th	10.97 (8.78)	7.40	7.35
95th	11.82 (9.63)	8.38	8.26
99th	14.04 (11.38)	10.24	10.03

When comparing the scaling rates of the 2 types of binning models, it can be seen that using equal number of observations per bin results in smoother scaling behaviour. This is probably because it groups, in one bin, a larger number of observations compared to the 2ºC bins. In fact, when comparing the number of bins between one method and the other it shows that while the average number of observations per bin in the 2º method is very large in the middle of the temperature range, in the extreme bins, the number of observations is around 200 measurements, compared to the equal observation method which has, for some stations around 600 points.

On the other side, it is shown that binning approach using equal observations and quantile regression

show similar results, around 6.3-7.5%/ºC for 80th and 90th quantile and values around 10%/ºC for the largest quantiles.

Looking at the spread of scaling rates, the interquartile range (IQR) groups the values between the first and third quartile. From all models, the quantile regression is the one with the smallest IQR, with values from 1.25%/ºC to 1.32%/ºC for increasing quantiles. While both quantile regression and the binning approach with equal number of observations have similar IQR, the binning approach with 2º bins has larger values between 2.06%/ºC and 1.87%/ºC for increasing quantiles. Overall, the scaling rate half of the stations in the binning approach at 2 bins range between 2 degrees from the median value, as opposed to the quantile regression where the range reduces to 1.3%/ºC.

Not only the scaling value is important to consider, but the confidence interval of the estimate. For simplification, the confidence intervals for the quantile 80th and 99th are compared between 3 models. Table 4.7 shows the mean of the differences between the upper bound and the lower bound of each scaling rate. As it can be seen, the model that yields the smallest confidence interval is the Quantile Regression with mean confidence interval of 1.01%/ºC for the 80th quantile, followed by the Binning Approach with Equal length of observations and lastly the Binning Approach with 2º bin separation.

Table 4.7: Average confidence interval difference between Binning Approaches and Quantile Regression.

	Mean confidence interval difference between for the 80th and 99th quantile		
	BN Corrected 2 degrees	**BN Equal Observations**	**Quantile Regression**
Quantile 80th	3.31	1.94	1.01
Quantile 99th	4.55	3.88	1.88

4.5. Piecewise regression model comparison

In this section, the results from the Binning Approach with 1 change-point and the Piecewise linear quantile regression are presented and compared. In addition, for the piecewise linear quantile regression, an in-depth study on the cases of change-point detected, the range of temperatures and confidence interval is presented.

For all the models, the first step consists in identifying the stations from which a change-point is estimated. As shown in Table 4.8, depending on the model used, different number of stations found a change-point.

Table 4.8: Number of stations for which the model estimated a change-point.

	Quantile			
Model	80th	90th	95th	99th
BN normal	238	215	220	215
BN corrected	224	243	253	267
BN equal observation	320	340	351	353
Quantile regression	350	343	344	329

An initial inspection on the estimated change-point temperature (Figure 4.15) shows that, for the 3 models (the original Binning method is excluded), the range of temperatures where transition occurs goes from -5ºC dew point temperature up to 20ºC.

For the piecewise linear models (based on Binning Approach methods), most of the change-points are concentrated between 10 and 15 degrees. For example, at the 95th quantile, the BN with equal observations has 186 stations within 10-15 ºC. Piecewise Quantile regression is the model that detects the maximum number of change-points, almost all the stations of study. From those, the largest range of temperatures is between 10-15ºC. On average, there are 80 stations identified with a change-point in that range between quantiles.

From this analysis, it is concluded that while the majority of change-points are estimated between 10-15 degrees, there is a large distribution of change-points between models, suggesting that the estimated temperature can be sensitive to the characteristics of the model behind its calculation.

For simplification, the comparison between piecewise and non-piecewise models presented in this report refer only to the Quantile Regression and Piecewise quantile regression models, given that

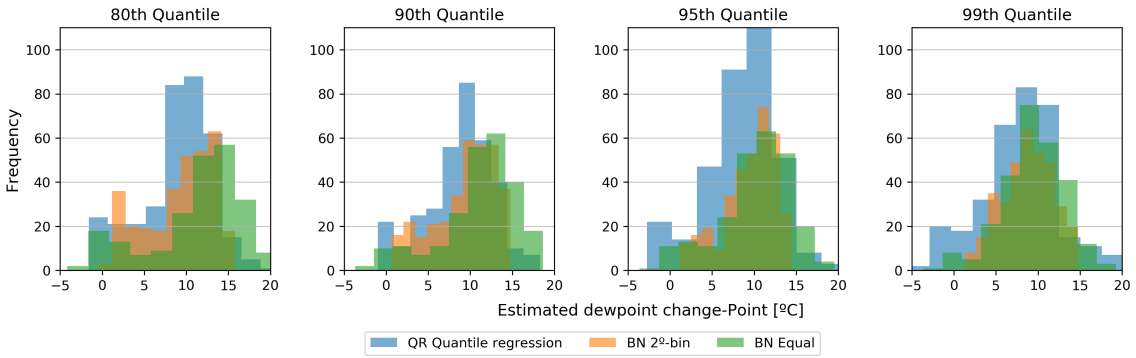


Figure 4.15: Distribution of estimated change-point temperature for the binning methods with a changepoint and the piecewise quantile regression.

the quantile regression was the model with the smallest confidence interval and best performance compared to binning approaches.

To better understand what does the change-point really represent for the stations, a visual inspection on all gauges is carried out, for which change-points are categorized based on similar patterns. Figure 4.16 summarises the results with examples for each case.

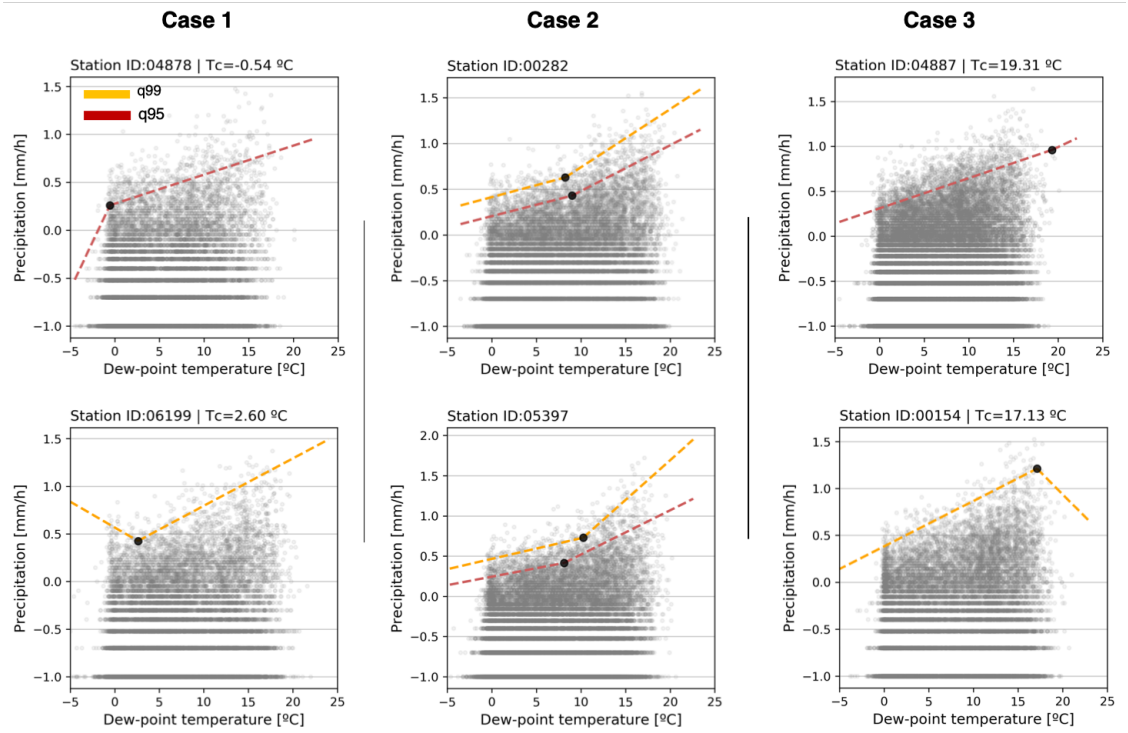


Figure 4.16: Categorisation of 3 types of change-point temperatures for the 95th and 99th quantiles.

There are 3 main groups of estimated change-points, mostly organised according to the dew point temperature that they occur.

- Case 1 Initial temperature bins: For these stations, the change-point occurs for temperatures between - and Y and occur because the of the number of very small rainfall events. Although wet events are considered if hourly rainfall is P>0.1 mm/hour, some stations have a very large number of observations at 0.2mm/h or more, which causes the model to select that shift from very small observations towards larger precipitation as the change-point. In this cases, it is common to find changes in slope sign shifting from negative to positive slope. In these cases, the beta 2 values are normally very similar to those scaling rates obtained by quantile regression, since

they are based on almost the whole range of observations.

- Case 2 Middle temperature bins: For this group of stations, the predominant temperatures are between 9º and 12ºC. In general, the scaling rates at lower temperatures follow smaller values than the CC rate (around 5%/ºC) whereas for higher temperatures in most cases reach scaling rates near the super-CC and higher.

- Case 3: High temperature bins: These group of stations have change-point around 14-15ºC. Similar to Case 1, this group has the scaling rate of lower temperatures closer to the Quantile regression, since the change-point occurs at the very end of the temperature range. In addition, the slopes after the scaling rate are above the super-CC, reaching very steep slopes of even twice the super-CC. The reason behind these steep behaviour is mainly because the change-point in this case mostly occurs in the last temperature bin of analysis. In this cases, some of the sloped after the change-point are negative, which is associated to a decay due to data undersampling.

Since cases 1 and 3 are generally estimated because of undersampling at the extreme temperature bins, the values of interest that will be taken for further analysis are the ones where the change-point occurs in the middle range. A comparison between the change-points at 95th and 99th percentile show that, in the majority of cases, higher quantile results in a change-point in lower temperature bins.

4.5.1. Acceptance criteria for change-point temperature selection

In accordance to the criteria described above, and as described in the Methodology, the results from the change-point are filtered based on the 2 criteria: (i) exclusion of change-point at the extremes of temperature range and (ii) exclusion of models performing worse than quantile regression based on BIC test.

The first step for filtering the stations is to flag the stations with change-point temperature lower than 4ºC as well as the negative scaling rate slopes. The 4º threshold is selected to identify stations that have enough observations at the very low range of temperatures (and thus avoid a change-point caused by undersampling). The second step is to look at the results from the BIC test. The stations where the change-point model is preferred are those that have a lower BIC (Van de Vyver et al., 2019). Table 4.9 shows the total number of accepted stations following the application of the acceptance criteria. Overall, there are between a 70-80% of stations that satisfy these conditions for every quantile.

Table 4.9: Nº of stations that meet the acceptance criteria.

	Stations that pass Criteria 1			
Criteria 1	**80th**	**90th**	**95th**	**99th**
	266/353	276/343	287/344	236/329

Overall, it can be seen that 2/3 of the stations have a better performance from the model with a change-point compared to the quantile regression. Figure 4.17 shows the histogram of change-point temperature from stations that fulfil the both criteria.

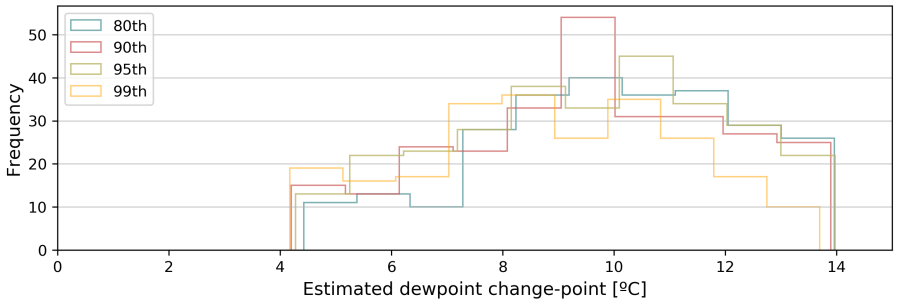


Figure 4.17: Distribution of Estimated dew point change-points for the stations with better BIC performance compared to Quantile Regression.

4.5.2. Inference results for the change-point model

In this section, the parameter estimates obtained from the change-point model are presented. Figure 4.18 shows the overall results of the scaling rates for the range of β before and after the change-point. In general, the boxplot shows that the uncertainty in β_2 is larger than in β_1, with an interquartile range going between 10-15%/ºC for the 80th quantile opposed to a 2%/ºC range for the β_1. There are some outliers identified within the stations analysed, that reach scaling rates for β_2 between 30 to 40%/ºC. Running an analysis to those stations show that, for the maximum slopes of beta 2, the associated change-point occurs at the very last range of temperatures, for example in temperatures larger than 15ºC. Therefore, a pattern observed is that steeper scaling rates (or scaling rates considered as outliers) occur for change-point temperatures of 15ºC or more, in stations where the upper limit of the range of temperatures is close to the change-point.

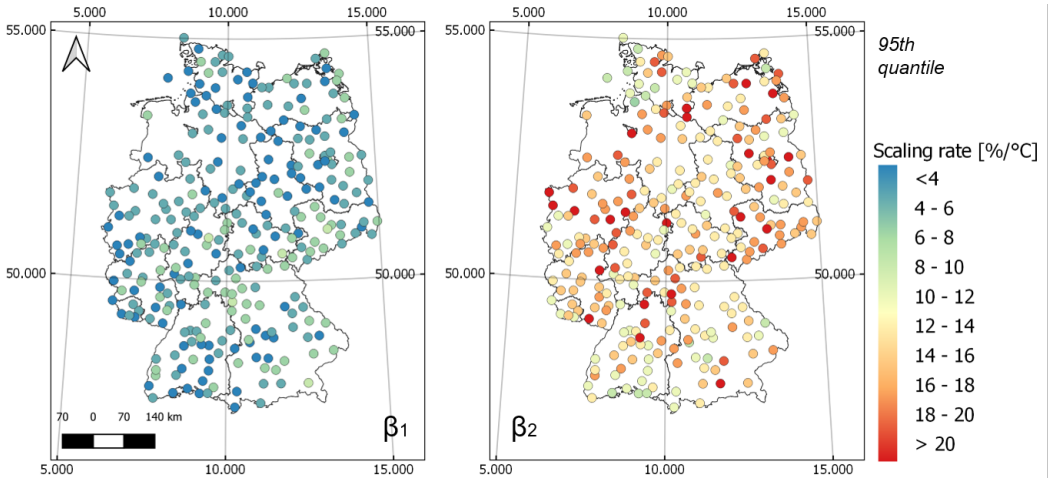


Figure 4.18: Boxplot of scaling rates for β_1 (lower T_{dew} values) and β_2 (higher T_{dew} values).

Table 4.10 shows in more detail the mean values, for the 95th quantile, of the parameters obtained through the piecewise linear quantile regression model, together with the confidence interval which were obtained through the bootstrap method.

Table 4.10: Beta parameters and Change-point estimates for the Piecewise linear quantile regression model.

Inference results	Mean	90% Confidence Interval (mean upper and lower band)
Beta 1	4.65 %/ºC	3.29 - 6.04
Beta 2	14.51 %/ºC	11.64 - 18
Change-point	9.77 ºC	8.30 - 11.06

As it was previously highlighted, the confidence interval of β_1 is smaller than β_2 , with a scaling range of 3%/ºC opposed to the almost 10%/ºc observed for β_2. The results from β_1 and the confidence interval for extreme rainfall show smaller values compared to the 7%/ºC CC rates. For β_2, values similar to the super-CC are observed, with confidence interval between 11.54 - 21.24%/ºC. It is important to notice that larger confidence interval are reported when results are observed station by station. This is because the values presented on the Table are obtained doing the upper and lower band, for easier representation.

This puts into perspective the validity of 1-change point methods. While, overall, the estimation of the change-point temperatures occurs over a range that can be considered correct, there are some outliers with scaling rates 10%/ºC larger than the super-CC scaling rates. In those cases, a quantile model would be preferred, since large scaling rate deviations have been seen to occur due to change-point closer to the upper limit of temperatures in a station.

4.6. Spatial Analysis of results

4.6.1. Distribution of scaling rates

This section investigates the sensitivity of hourly precipitation to dew point temperatures across Germany. The estimated scaling rates used here are those measured by the quantile regression model.

Figure 4.19 shows that, for the center part of Germany, a scaling rate larger than the Clausius-Clapeyron is observed already at the 80th Quantile. The smallest values are located on the North-West of the country, where it ranges between 3-4%/ºCC. The lowland region in the North-West is mainly influenced by the North Sea, that carries a uniform moist air resulting in mild summers and mild winters (Turner & Berentsen, 2020). This might explain why the scaling rates are lower than in other parts of Germany. In the North-Eastern direction, the scaling rates are higher.

While annual precipitation in that area is smaller than in the west, the average temperatures for the warmest and coldest months are more extreme (Turner & Berentsen, 2020), possibly leading to more extreme rainfall events than in the West. The largest scaling rates (around 10-12%/ºC for the 99th quantile) are observed in the Center of Germany. The geography in Central Germany is based on low mountains and hills (Turner & Berentsen, 2020). The range of scaling in this region might be associated to the slopes of the mountains. For the slopes facing the west, large precipitation amounts are observed, due to air masses coming from the sea. On the other side, for the east-facing slopes, the total precipitation is very low (Turner & Berentsen, 2020).

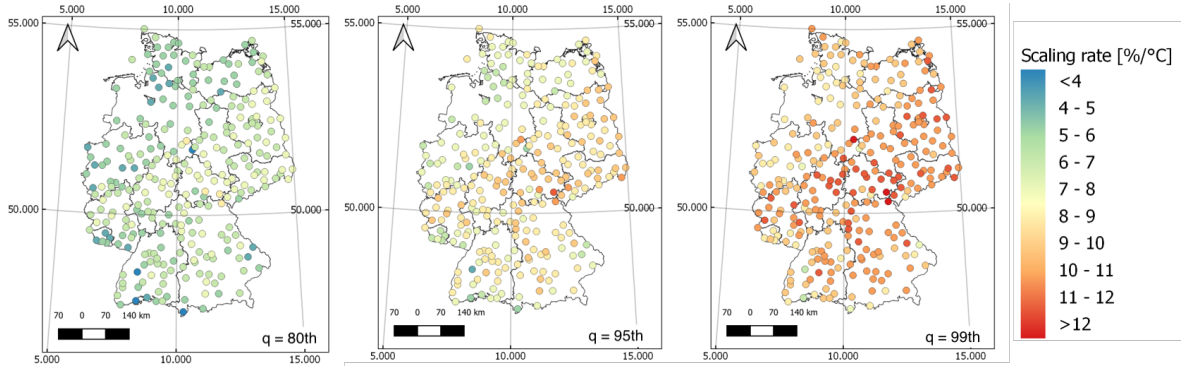


Figure 4.19: Scaling rates distribution across Germany estimated for the 80th, 95th and 99th Quantile.

4.6.2. 1 change-point temperature distribution

In this section, the change-point temperatures that were obtained through the application of the Piecewise linear quantile regression are shown for the 95th quantile. Figure 4.20 shows only the stations that passed the acceptance criteria described in the previous section.

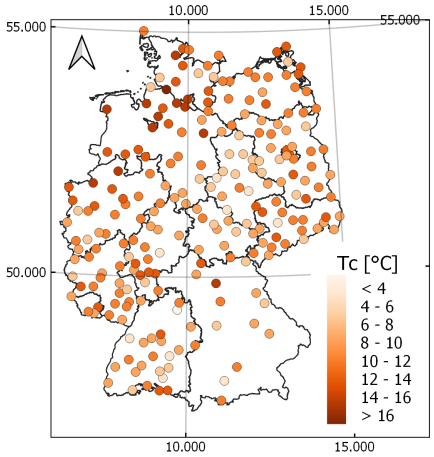


Figure 4.20: Distribution of the estimated change-point dew point temperature for the 95th quantile

The map shows an heterogeneous distribution of change-point temperatures across Germany. However, it can be seen that in the North-west, the change of slope occurs at higher temperature range, whereas in the centre of the country values of change-point are around 9-12 degrees. In general, If results from the distribution of change-point temperature are compared to the scaling rates obtained by the quantile regression, it can be observed that, in the North-west, where the scaling rates where the lowest, the change-point is located in the upper range of temperatures and, in the center, where

the scaling was largest, the change-point occurs at middle range temperatures. This is a reflection of the distribution of precipitation per bins. For the North-west case, quantile regression results in smaller scaling which means that the extreme rainfall occur only at the very end of dew point temperatures, with most temperature bins have a CC or smaller scaling rate. For the center stations, the fact that quantiles obtained were above CC, reaching for larger quantiles a scaling rate of 9-10%/ºC indicates that the extreme rainfall already starts for smaller dew point temperatures and, thus, the change-point occurs at smaller temperatures.

4.6.3. Scaling rates for different climate regimes

Given the high variability of scaling rates shown in the previous sections, it was decided to study whether there is a relationship between scaling rates and climate regimes. This is done to check whether patterns or clusters of similar scaling estimates can be found. There are several approaches available to characterize precipitation regimes, mostly looking at seasonal and climatological properties. For this project, the precipitation regimes from Lu et al. (2020) are used, which were obtained by using satellite-based cloud-top temperature (CTT) time series. There are 3 regions established according to Lu et al. (2020): The North-East, North-West and South (Figure 4.21), each one containing 50, 163 and 140 stations.

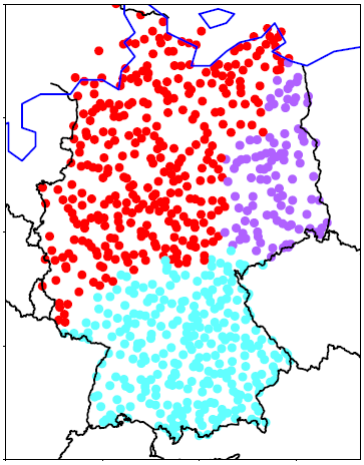


Figure 4.21: Climate regimes as defined by Lu et al. (2020)

Grouping the scaling rates between the 3 precipitation regimes (Figure 4.22 identified the North-East as the one with the largest scaling values, around 7%/ºC for the 80 quantile and up to 11%/ºC for the 99th quantile. Compared to the North-West and South, which, both had similar rates for the 4 quantiles, the North-East region only differs about 1%/ºC. This is not a significant difference considering as well that the given scaling rates have normally confidence intervals larger than 1%/ºC. In addition, the number of stations that fell into the Noth-East category were only 1/3 of the ones from the other regions.

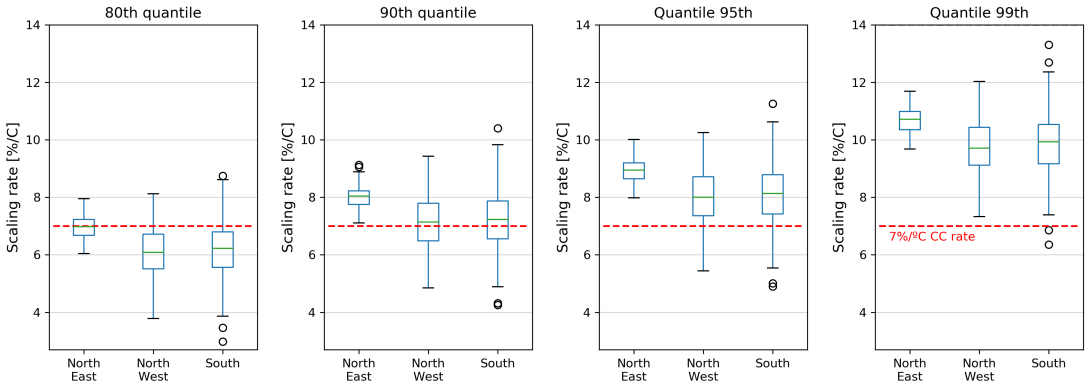


Figure 4.22: Boxplot of scaling rates for the 3 climate regions for Quantiles 80th-99th

4.7. Comparison of wet vs dry percentiles

In the Background section, the importance of comparing the results from relative percentiles and absolute percentiles was addressed. Mainly, it is important because absolute percentiles incorporate the occurrence of events, which can be specially relevant for studies of the impact of extreme rainfall, as well as climate change studies.

In Figure 4.23, the spatial distribution of the scaling rates are shown for 3 cases, the quantile based on relative percentiles (Fig.a), the absolute frequency for same intensity events (Fig.b) and the results from applying the absolute 99th quatile (Fig.c). Comparing the 95wet percentile with the Absolute percentile (wet and dry) accounting for the wet fraction, shows that, between both cases, the scaling rate are very similar. This means that there is almost none sensitivity to whether dew point temperature are used with all hours or wet hours. This has some interesting outcomes. (Lenderink et al., 2011) showed that, when comparing the same analysis as done in this research but with SAT instead of dew point temperature, using absolute percentiles resulted in a poor scaling relationship. Therefore, only wet percentiles reproduced robust scaling similar to the CC rate.

Having obtained similar results of scaling rates for both percentiles shows that it is not relevant what type of percentile is used in Germany. Indirectly, this also shows, as also Lenderink et al. (2011) points out, that it is atmospheric moisture (represented by dew point temperature) what determines the increase of precipitation with temperature.

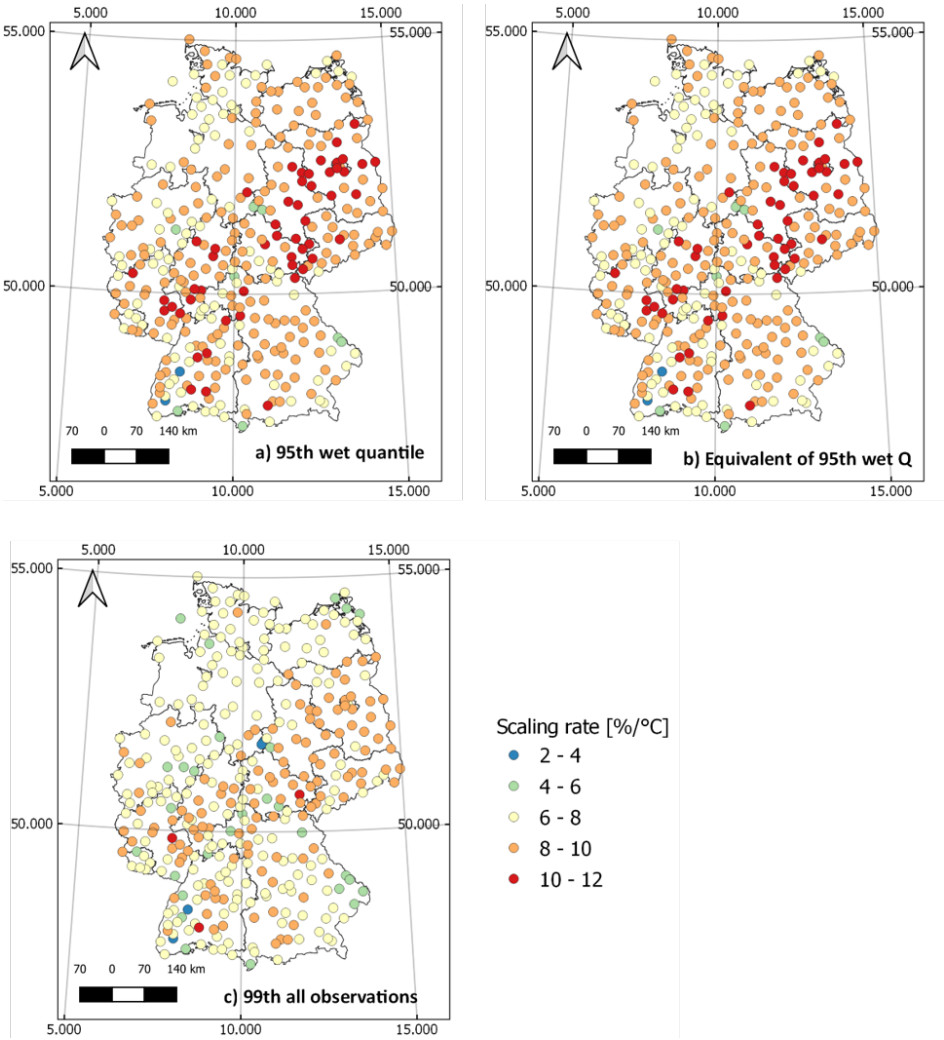


Figure 4.23: Spatial distribution of scaling rates in 3 scenarios. a) 95th quantile of wet observations, b) the equivalent of the 95th wet quantile but for all observations (wet and dry) and c) the 99th quantile for all observations

The results from the 99th quantile are shown to evaluate and put into perspective what does a 95th

wet quantile mean in absolute percentile. Given that the results of scaling rates are lower than those obtained for the 95th wet quantile, it means that in the study based on wet observations, the absolute percentiles of the 95th wet quantile that are considered are above the 99th percentile.

4.8. Summary of results

In the following section, the outcomes of this first block are grouped according to the objectives and research questions of this thesis.

Obj 1: Compare the suitability of using exponential regression models and piecewise regression models for estimating precipitation scaling rates.

Some of the learning lessons from this analysis are:

- As existing research points out (Wasko & Sharma, 2014), binning temperature methods to obtain scaling rates can be biased, and influenced by the length of observations. This research can support this statement, after showing that applying 2 binning methods, with varied number of observations per bin, can substantially alter the overall scaling rate.

  Within this project, it has been shown that the bias is mostly due to the behaviour of the extreme bins. For the 2ºC binning method, the small amount of observations at the initial bins resulted in an overestimation of scaling rates exceeding the Clausius-Clapeyron rate already for not extreme rainfall quantiles (i.e 80th).

- Using binning approaches of equal observations shows accurate results of scaling rates, with only 1 %/ºC in the confidence interval (except for the 99th quantile). While the estimated scaling is accurate, the range of scaling rate results across Germany is large (i.e for the 95th quantile the minimum scaling rate is 5.02%/ºC compared to the maximum which is 11.54%/ºC). After visually investigating the cases of two stations with very different scaling rates, it is shown that the cause of such difference is not because of the methodology applied but because of climatological differences in Germany. The stations with lower scaling rate have, in general, larger number observations in low dew point temperature bins as opposed to stations with steeper scaling rate.

- When A corrected version of the binning approach with, restricting the initial number of bins, showed a reduction in the scaling rates, obtaining closer results to the method using equal number of observations of bins.

- From the 2 binning approaches, the one using equal number of observations per bin has a higher goodness-of-fit as opposed to the 2º bin method.

- The estimation of scaling rates through the Quantile Regression shows that, for the 80th quantile, the rates are close to the expected Clausius-Clapeyron relationship, around 6-7 %/ºC. With increasing quantiles, the rates also increase to values larger than 7%/ºC. For the 99th quantile, the median scaling rate is around 10%/ºC. None of the stations reach the double clausius-clapeyron relationship.

- Overall, the largest scaling rates are observed in the center of the country, from the North-east towards the center. In contrast, although the distribution of scaling rates is heterogeneous in the country, the North-west region shows a cluster of stations with lower scaling rate.

- An analysis on the relationship between station elevation and scaling shows no correlation between both variables. In a research by Schroeer and Kirchengast (2018), it was shown that mountainous regions in Austria had lower temperature sensitivities than in the eastern lowlands. Although there were also elevation differences there, it was suggested that, rather than orography, what really influenced the scaling is the different weather patterns that can be associated in specific area, and not just the elevation of where the station is located.

- From the analysis of 3 models and following existing research, a quantile regression model is considered to be less biased. In addition, it is the model that has the smallest confidence interval, thus resulting in more accurate estimations of the scaling rate. Another benefit of quantile regres-

sion is that it can help comparing the behaviour of the scaling rates against different predictors. For the study of different storm variables (Part 2), quantile regression is thus the best approach.

Obj 2. Evaluate the success and validity of change-point models to detect super-CC scaling

- The outcomes of applying a piecewise regression model shows that there is a large range of temperatures where a change of the scaling rate is estimated. From all the tested models, the Piecewise linear quantile regression estimated a change point in all the 344/353 stations, while the other methods found a significant change-point for around 100 stations.
- There are 3 main cases of change-point detected, (i) at the lowest temperature bin, (ii) in the mid-range of temperatures and (iii) closer to the maximum temperature range. Both cases in the extreme range of temperatures are based on statistic artifacts of the data (abrupt changes of data or reduction in number of observations). For the project, only the changes in the mid-range temperatures are studied.
- Filtering the change-point based on previous condition and on BIC test), the number of stations where a change-point model is identified is 280.
- The identification of around 280 stations with a change-point re-affirms the results from previous research that suggest that the precipitation does not have a constant relationship with temperature. There are, however, large uncertainties in the range of temperatures where it might happen and the underlying assumptions behind it.
- The result from filtering the stations show that there are some outliers, with very large scaling rates. rates. An analysis on those locations show that the stations with scaling rates above 20%/ºC, the detected change-point is always larger than 15ºC, and with the upper temperature range very close. Considering that, on the extreme temperature range, the number of observations is smaller, this could explain why the scaling rates are larger than the other stations.
- Concerning the scaling rates observed from filtered change-point model, it can be seen that the median scaling rate before the change-point follows values slightly lower than the CC relation (~6%/ºC), whereas the slope after the break-point reaches values larger than the 2CC (~6%/ºC).
- For Germany, the increase from CC to super-CC varies across quantiles. For the 80th quantile, the median transition point from CC to super-CC occurs at 11.51ºC, reducing to 10.80ºC, 10.37ºC and 9.10ºC respectively for increasing quantiles to 90th, 95th, and 99th.
- The analysis of the change-point results show that, the β_1 values obtained from the piecewise linear quantile regression model and their confidence interval are compatible with the 7%/ºC CC rate. For the β_2 results, the results show larger range of scaling rates, but with the interquartile range between 11-15%/ºC. These results agree with the estimation of super-CC scaling rates for several locations (Van de Vyver et al., 2019; Lenderink & Van Meijgaard, 2008). However, it is important to highlight that there is a large confidence interval for β_2 ranging between 11-21%/ºC.
- Comparing the scaling rates of β_1 with the ones obtained through quantile regression, it can be seen that the later model has larger range of slopes for increasing quantiles (and median values larger then the CC), whereas the β_1 results stay within a smaller range of results (and closer to the CC rate). This means that, when using quantile regression approaches, for lower temperature ranges the scaling rates might be overestimated, and for larger temperatures underestimated.

5

Part 2: Analysis of storm events

The following chapter presents the results of 98 stations from which the storm events were retrieved. All the stations have records for 25 years of data. In this chapter, dew point temperature is defined as temperature unless stated otherwise.

5.1. Distribution of storm events

In Figure 5.1 (left), the total number of storm events obtained per station is presented. This excludes those events with a duration smaller than 10 minutes, given that the observations are provided in a 10-minute time interval.

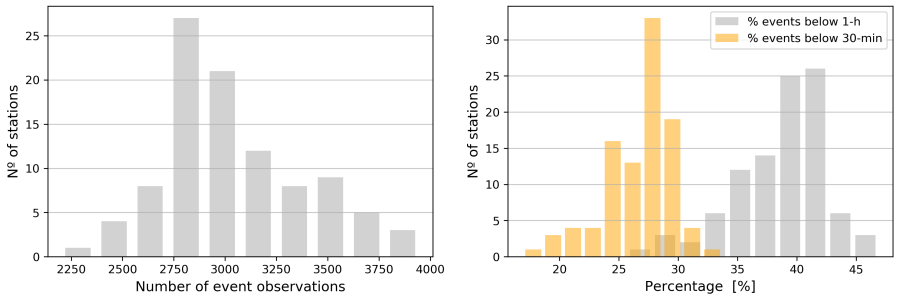


Figure 5.1: Left: Histogram on number of stations vs number of storm events. Right: % of events below 1-h and 30-min.

Overall, the range of storm events in Germany for a 25 year period go between 2250 and 4000 with most stations having 3000 events. For more than 65 stations, the duration of events of 1-hour or less represents the 30-40 % of all events. In addition, for 55 stations, more than 25% of their storm events last for 30 minutes or less.

This shows that, for Germany, the proportion of short events from the total dataset is considerably large having between 20% and 40% of sub-hourly storm events.

5.2. Range of values from storm properties

Before studying the scaling rates results, an overview on the range of values obtained for each of the 4 storm properties (duration, total rainfall depth, max 10-min rainfall and mean intensity) is presented. This is done by compiling the storm events from all stations together and plotting in as a boxplot. As Figure 5.2 shows, the range of values for the variable duration is between 10 minutes and up to 7000 minutes (almost a 5 days event). The majority of storm events last up to 690 minutes (around 11 hours events). The very long events observed as outliers occur only 1 or 2 times for each station.

Looking at maximum 10-min precipitation, most storm events have a maximum 10-min rainfall between 0.25-1 mm/10-min. However, there are values observed that can reach up until 40 mm/10-min.

Similarly, the mean intensity (mm/10min) results show that most storm events are between 0.1-0.3 mm/10-min with values reaching up to 14mm/10-min rainfall. Lastly, looking at the total rainfall depth, the interquartile range goes from 1.5 to 5.8 mm of rain. While most storms will have total rainfall events around these values, the results indicate that total rainfall of a storm event can reach up to 400 mm.

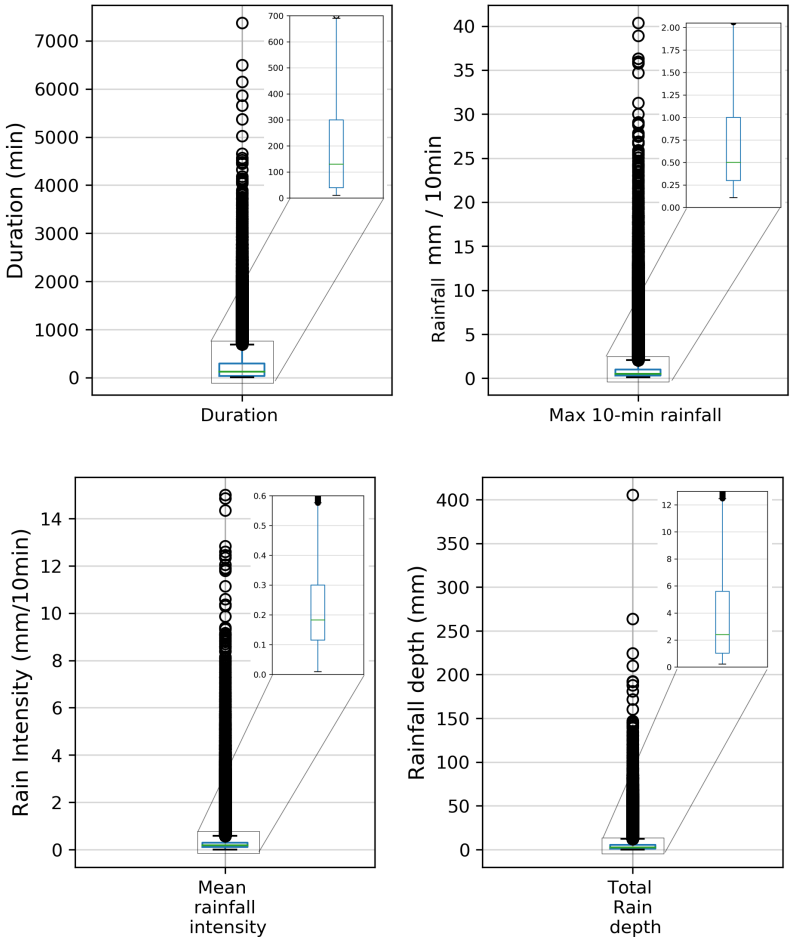


Figure 5.2: Range of storm properties values observed across all stations.

5.3. Scaling rates of storm properties

To evaluate the scaling rate of storm event properties, the quantile regression method is applied to 98 stations on 4 storm variables: storm duration, peak rainfall at 10-min, mean intensity during the event and total rainfall. In Figure 5.3, the results from the analysis are shown as a boxplot with the 4 quantiles of interest.

From all the studied storm properties, the storm duration is the only one that shows a negative scaling rate slope. For the 80th quantile, this negative relationship between duration and dew point temperature has a median value of -1.15%/ºC. The fitted slope for the 99th quantile has a larger interquartile range, reaching to lower scaling rates (up to -4%/ºC).

For the total rainfall depth, the overall scaling rates stay almost constant with increasing quantiles, with only a slight increase from a median of 5.05%/ºC (80th Quantile) to 5.50%/ºC for the 99th Quantile.

As it can be seen on Figure 5.3, larger differences of scaling rates between quantiles can be observed for the mean rainfall intensity (median values ranging from 6.56-10.32 %/ºC) and maximum 10-min rainfall (median scaling rate between 10.25-12.89 %/ºC). In particular, the maximum 10-min rainfall scaling rates are in all cases larger than the CC relationship, reaching for the largest quantile values surpassing the super-CC scaling rate.

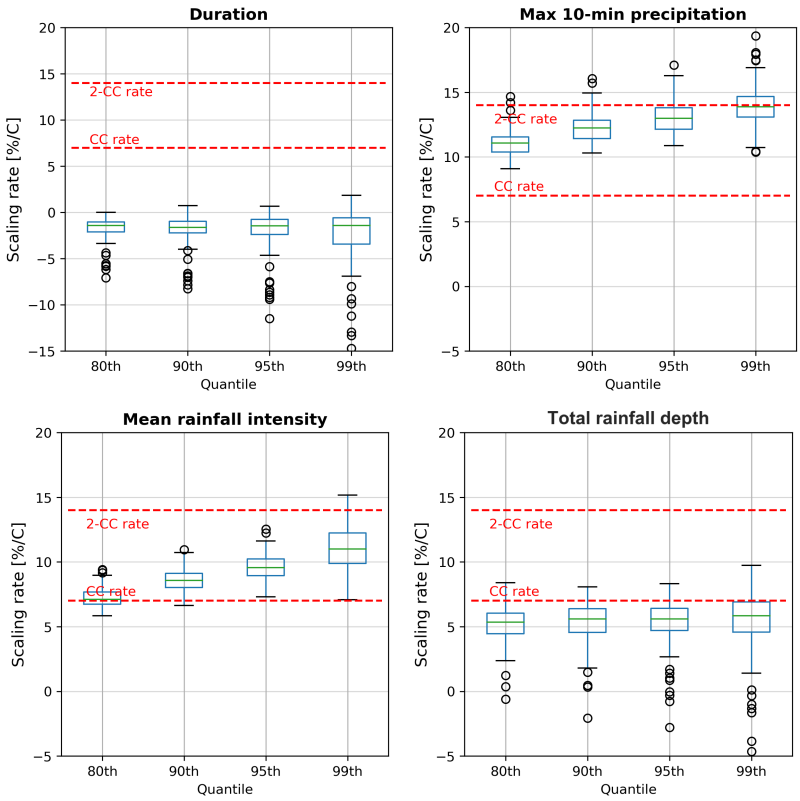


Figure 5.3: Boxplot of scaling rates results [%/ºC] for quantiles 80th-99th on storm duration, maximum 10-min intensity, mean rainfall intensity and total rainfall depth. Computed for 98 stations with the quantile regression method. Red dashed lines represent the Clausius-clapeyron scaling rate (7%/ºC) and super-CC (14%/ºC).

In Table 5.1, the median values and confidence interval of the scaling rates are presented for the 95th and 99th quantile. Overall, the confidence interval of for all variables ranges around 1.1 %/ºC (95th Quantile) and 1.8 %/ºC %/ºC (99th Quantile).

Table 5.1: Median scaling slopes fitted to the 95th and 99th percentile with confidence interval obtained from bootstrapping method.

95th Quantile	Median Scaling rate	Confidence Interval
Duration	-1.96	-2.84 to -0.89
Max 10-min precipitation	13,16	12.10 to 14.30
Mean Intensity	9.69	8.66 to 10.85
Rainfall depth	5.21	4.28 to 6.20

99th	Median Scaling rate	Confidence Interval
Duration	-2.09	-2.89 to -0.56
Max 10-min precipitation	13.89	12.04 to 15.81
Mean Intensity	11.07	9.34 to 12.79
Rainfall depth	5.50	3.98 to 7.35

5.4. Maximum 10-min event rainfall analysis

5.4.1. Sensitivity analysis on temperature predictor

Given that the largest scaling rates are observed for peak rainfall intensity, in this section a more in-depth analysis is provided on the influence of different dew-point temperatures when obtaining the scaling relationship between peak intensity and temperature. The comparison is based on 4 studied dew point temperatures: (i) 4 hours before the storm event, (ii) at the start of the event, (iii) the maximum dew point temperature during the event and (iv) the dew point temperature when the maximum precipitation is reached.

Figure 5.4 shows the result of the goodness-of-fit taking the mean between the 98 stations.

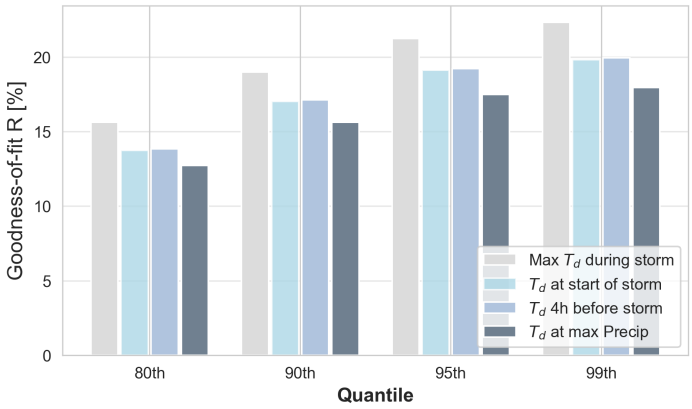


Figure 5.4: Goodness-of-fit criterion R. Mean R obtained from calculating the scaling rate of maximum 10-min temperature with 4 predictors of dew point temperature.

In general, it can be seen that the best temperature predictor with the maximum 10-min rainfall is the maximum dew-point temperature during the event, with a mean goodness-of-fit that ranges between 16.38% (80th quantile) up to 22.51% (99th quantile). The temperature that has the worst goodness-of-fit is the maximum temperature when the maximum precipitation is reached, which as a range of R from 13.17% to 17.94%. Both the temperature at the beginning of the event and 4 hours before have very similar goodness-of-fit results, as well as similar temperature values when individual stations are evaluated.

A comparison between the maximum dew point temperature during the event and the temperature when the precipitation peak occurs is done for several random stations. In Figure 5.5, an example from a random station is presented for support. In general, the temperatures when it is taken at the maximum precipitation is either equal or smaller to the maximum temperature during the event. For shorter events, the temperatures are normally equal. For larger events, the temperatures are generally lower. Translating this information into the scaling rate, it is seen that when temperature matches the maximum precipitation, the scaling rates are generally smaller than the case where the maximum dew point temperature is taken.

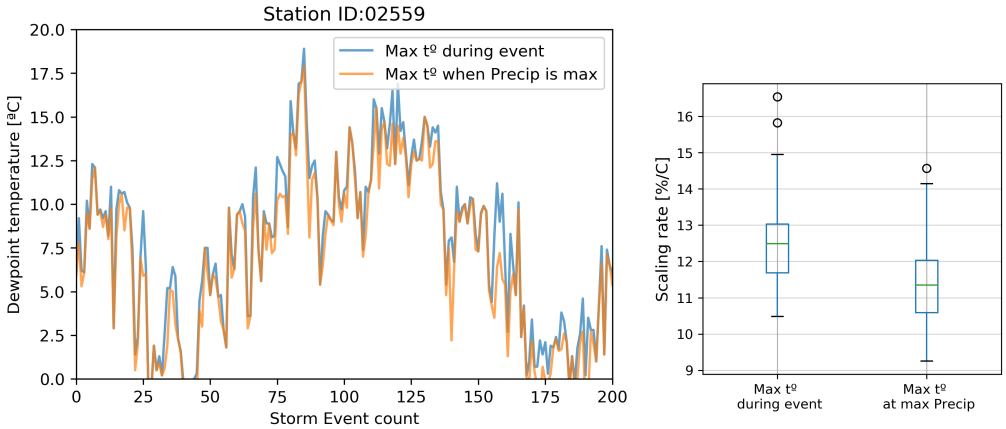


Figure 5.5: Left: Comparison of highhest and lowest dew point temperature predictor for the first 200 storm events of Station ID02559. Right: Boxplot of scaling rates from the temperature predictors.

From this analysis it is observed that, opposed to what Lenderink et al. (2011); Dahm et al. (2019) showed in their analysis, using a dew point temperature predictor 4 hours prior to the storm event has lower goodness-of-fit. The reasons behind the different results are presented in the discussion section.

5.4.2. Peak rainfall at varying temporal resolution

The results from this analysis show that, comparing peak rainfall intensity with mean rainfall during the event, the scaling rates observed are higher. In order to study the actual influence that the temporal resolution of rain can have on the scaling rate, the methodology followed to obtain the scaling rates for the maximum 10-min rainfall is repeated with a 30-min and 1-hour dataset. Figure 5.6 shows the results of varying time resolutions for the 95th and 99th quantile.

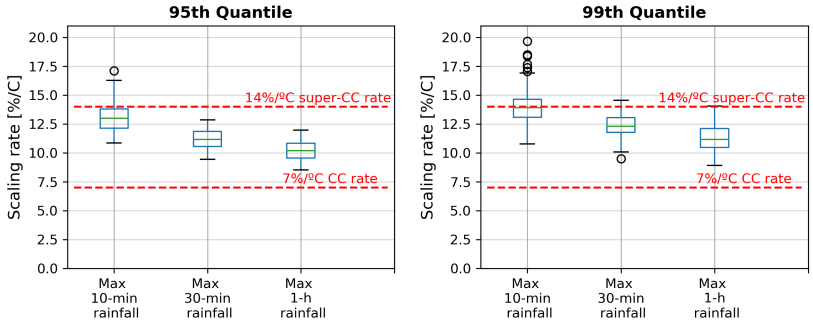


Figure 5.6: Boxpot of scaling rate of peak rainfall intensity at varying temporal resolution: 10 minute, 30 minute and 1 hour rainfall. Computed for 98 stations through the quantile regression model.

Looking specifically at the 95th quantile, Table 5.2 shows that changes in time resolution from 10-min resolution, 30-min and 1-hour resolution, leads to a reduction of the scaling rate of around 1 to 2%/ºC. This is also observed for the 99th quantile.

Table 5.2: Median scaling slopes of maximum precipitation at 3 temporal resolutions, fitted to the 95th percentile with confidence interval obtained from bootstrapping method.

Peak Intensity at 95th Quantile	Scaling rate [%/ºC]	Confidence Interval
10-min	13.06	11.99 - 14.15
30-min	11.19	10.30 - 12.08
1-h	10.18	9.30 - 11.06

To evaluate the existing hypothesis that suggests that the causes are related to an increase of dry periods within an event, when averaging to larger quantiles, the intermittency is evaluated for each peak rainfall. To do so, the intermittency of the all the peaks at 30-min and 1-hour were obtained for each station. Then, the average intermittency from all the peak intensity was obtained for each station. The results in Figure 5.7 show the mean percentage of intermittency in the 30-min and 1-hour peaks and it clearly shows that, for 1-hour measurements, the maximum peak has, on average, 57% of intermittency, and the 30-min rainfall around 41%.

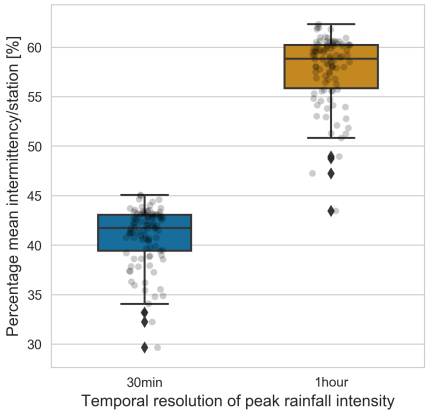


Figure 5.7: Mean of the intermittency percentage for 30-min and 1-hour peak rainfall [%].

5.5. Storm duration relationship with dew-point temperature

Studying the relationship between event duration with temperature is key, as research has shown that short and long events can scale very differently (Berg et al., 2009b). In this section, a more in-depth analysis on the influence of duration is proposed breaking down the analysis on the maximum 10-min rainfall into duration classes of events.

As it is shown on the Methodology, the analysis of the event duration is done for 9 stations distributed across Germany covering different regions and scaling rates. In order to evaluate the hypothesis made by Berg et al. (2009b) that different rain types can be found at different temperature range, 3 rain duration classes are proposed: Short events lasting a maximum of 2 hours, middle duration events that last between 2 and 10 hours and long events that last more than 10 hours.

While this method does not determine whether an event is more convective, stratiform or mixed, working in terms of event duration can help match the processes involved in rainfall generation.

5.5.1. Max 10-min peak rainfall

Taking as a reference the relationship between maximum 10-min rainfall during the event and max dew point temperature, Figure 5.8 shows per temperature bin, the proportion of the the 3 different event classes.

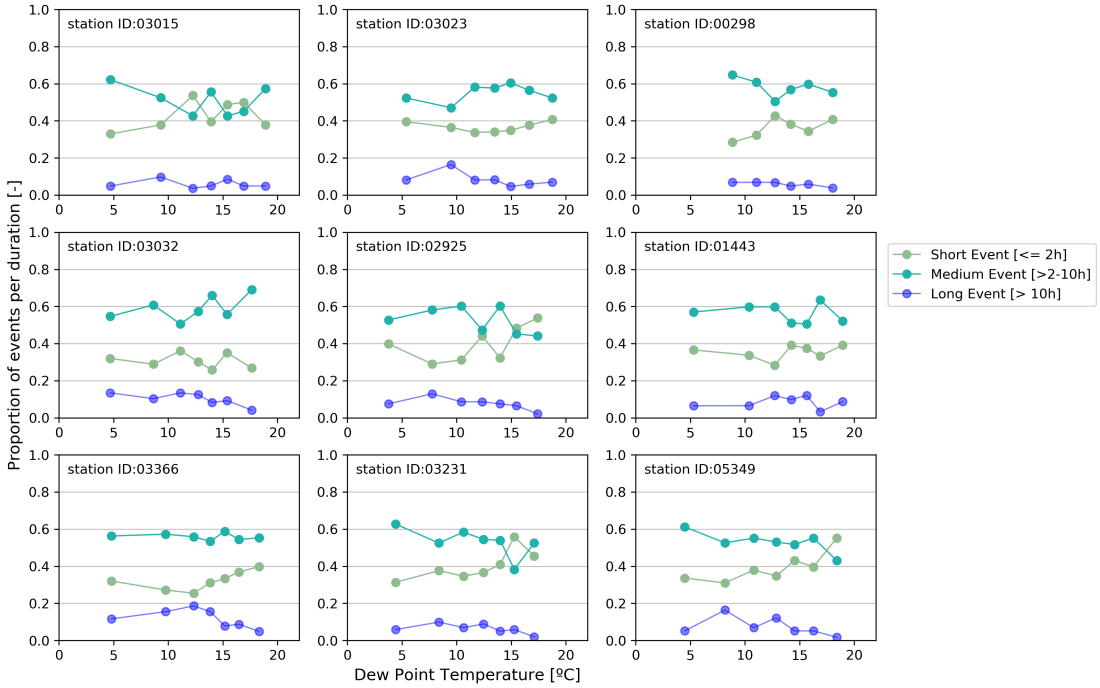


Figure 5.8: Proportion of small, medium and long duration events as a function of dew-point temperature based on events above the 95th percentile. Evaluated on 9 stations across Germany based on max 10min precipitation - max dew point temperature.

The results show that, overall, the most frequent events are normally the events between 2 - 10 hours (for small temperatures they represent around 60% of the events), followed by short duration events (around 40% of the events at small temperatures) and long events which represent the smallest group. When looking at the changes over temperature, in almost all the stations analysed, the long duration events reach a peak around 10 degrees, after which the number of events reduce almost to 0 for temperatures larger than 17ºC. The proportion of short events across temperatures changes between stations. In most cases, the percentage of cases per temperature is stable at 40%, although for larger temperatures, there are several stations where the proportion increases. This can be seen in stations ID02925, ID03366, ID03231 and ID05349. Lastly, medium duration events do not have large changes with increasing temperatures. Around 10-15 degrees, some stations present increases in the number of events while other show decrease.

The results obtained vary when compared to a similar research carried out by Panthou et al. (2014) that analysed different storm duration for different regions in Canada. In those cases, not individual stations were considered, but regions with similar climate characteristics. In addition to the proportion of events at different duration based on maximum 10-min peak rainfall, the scaling rates obtained for the peak rainfall were visually checked conditioned to the event duration. The motivation of this approach is to evaluate whether the peak of short events and long events have scaling rate similar to those obtained for the events combined.

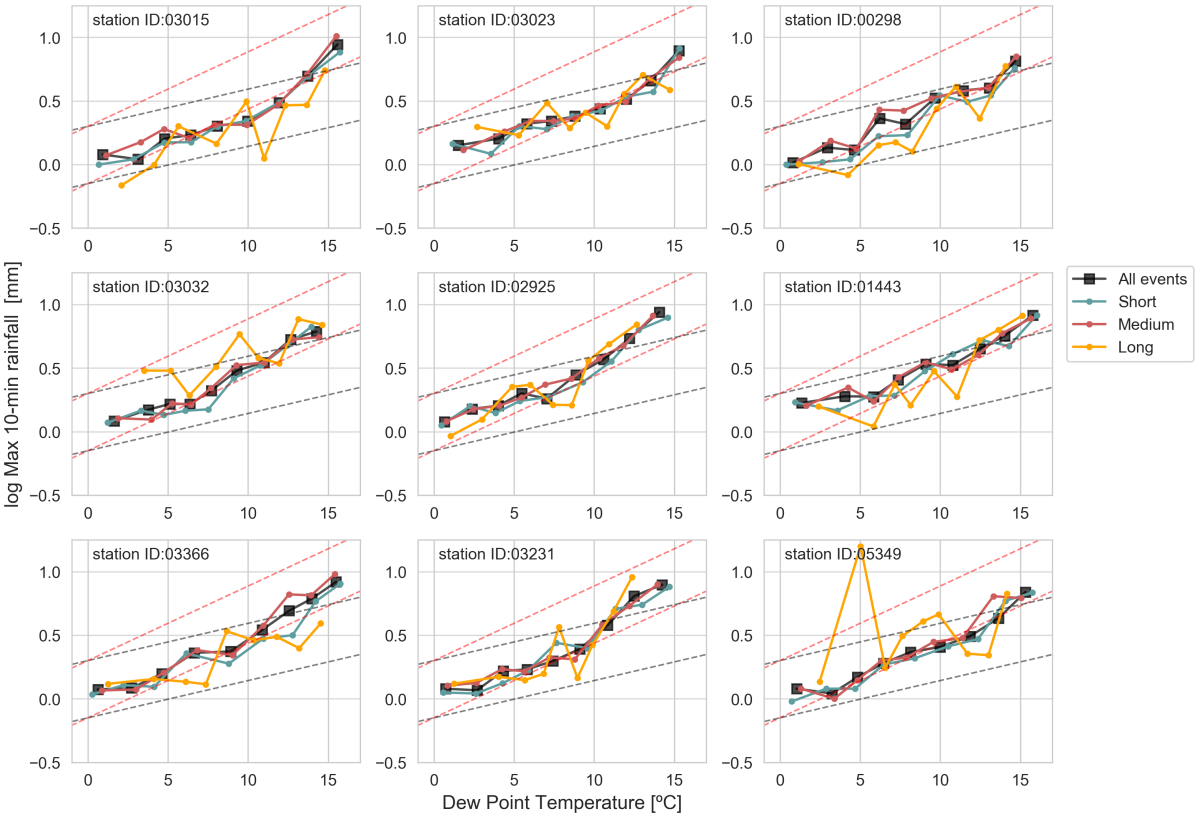


Figure 5.9: Relationship between the 95th quantile of maximum 10-min rainfall with maximum dew point temperature during the event. The relationship between precipitation and temperature is presented separate for short, medium and long events. Grey dashed lines represent the 7%/ºC CC scaling rate and red lines represent the 14%/ºC super-CC scaling rate.

As Figure 5.9 shows, the short duration events have a similar scaling rate to the combination of all events. Similarly, medium duration events scale like short events. On the other side, the long event relationship does not show a relationship as clear as the other duration. For example, stations ID3015, ID3032, ID03366, ID5349 have a flatter slope when compared to the other relationships. However, it can also be seen that some stations ID02925 or ID03231 have a similar slope than the other cases. Therefore, while some stations can show that longer events have smaller scaling rate than short events, the statement can't be made for all locations in Germany.

A feature that is important to highlight is that, by representing the scaling rates for a set of temperatures, the figure shows that, around 10ºC, some stations have a steeper increase in the relationship between dew point temperature and precipitation, larger than the super-CC (dashed line). This result could not be seen when plotting the overall scaling rates as a boxplot (in Figure 5.3). In case a change-point analysis would be proposed, it can be seen that some stations would show an increase in scaling rate for maximum dew point temperatures larger than 10ºC.

5.6. Total rainfall depth relationship with dew point temperature

Another variable of interest to study conditioned on storm duration is the total rainfall depth. This is specially relevant because the total rainfall depth can be very different between long events and short events. Following a similar approach than in the previous section, Figure 5.10 shows the results for the 95th Quantile, separating with short, medium and long the relationsip between mean dew point temperature during the event and total rainfall depth.

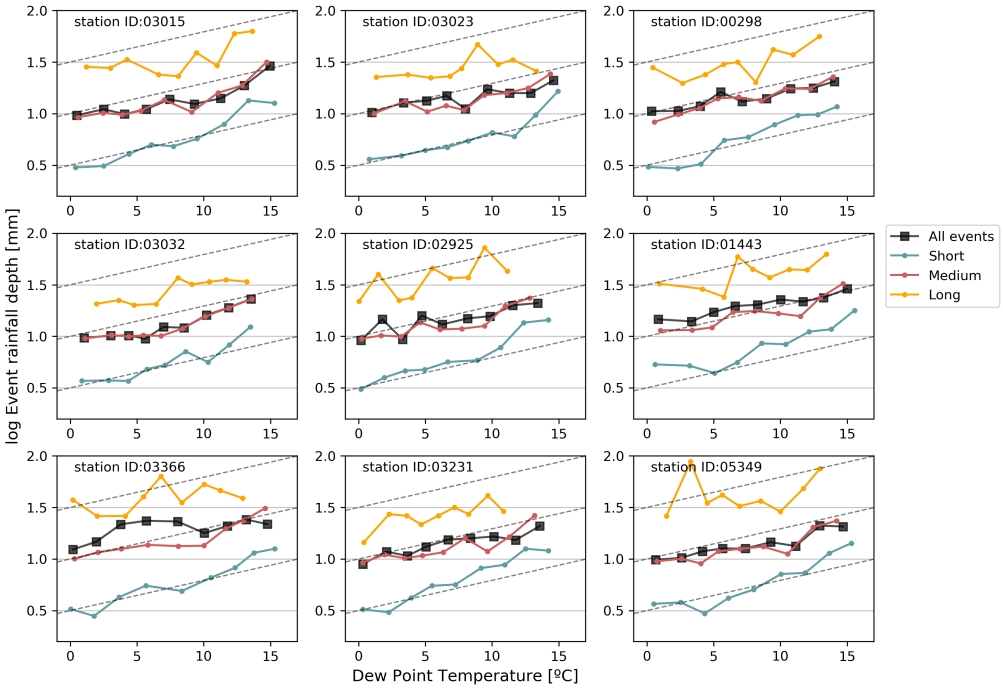


Figure 5.10: Relationship between the 95th quantile of total rainfall depth with mean dew point temperature during the event. Total rainfall depth is relationships are presented for the total events, separated between short, medium and long events. Dashed lines represent the 7%/ºC CC scaling rate.

One important outcome of this analysis is the fact that most of the stations show that short events have a larger scaling rate than the CC (dashed line). In stations ID01443, ID02925, ID03015, ID03231, ID03032 and ID03023, the slope is not always constant with the CC, but can transition between CC to larger scaling rates, for instance ID03023, ID02925 that have a transition for temperatures around 10º 15º dew point. One the opposite side, when looking at long events, while for some stations the rainfall depth and temperature scale similar to the CC, stations like ID1443 and ID3032 have less dependence and a smaller scaling rate.

5.7. Spatial distribution of scaling rates

To conclude the storm event analysis, this chapter presents the spatial distribution of the scaling rates specifically for the 10-min peak rainfall and the mean rainfall intensity.

As Figure 5.11 shows, the North-East region has, for both peak and mean intensity, larger scaling rate results than in the North-West and South. For the 10-min peak rainfall results, the North-East shows scaling rates between 13 and 14%/ºC, which also are the highest when looking at the mean rainfall intensity (between 9.57 to 10.58%/ºC).

These patterns reflect similar results to those observed during the analysis using hourly observations. Having a similar outcome also for storm characteristics analysis indicates that the temperature alone is not influencing the precipitation intensity, but that there are other factors, such as topography or local surface-atmosphere feedbacks that could be playing a role (Ban et al., 2015).

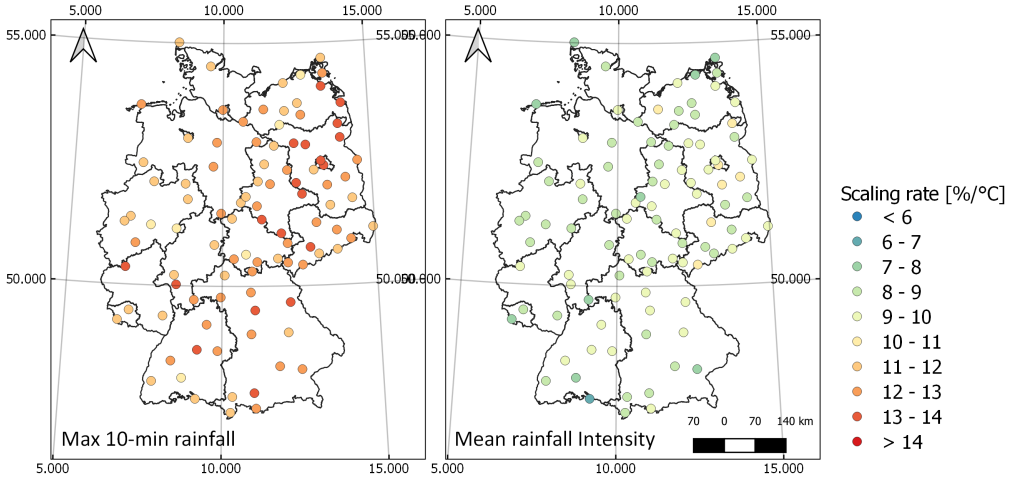


Figure 5.11: Spatial distribution of storm event characteristics for the 95th Quantile. Left: 10-min peak rainfall results. Right: Mean rainfall intensity.

5.8. Summary of results

Obj 1. Test the influence that storm properties have on scaling

Results from looking at the relationship between storm properties and dew point temperature indicate that the strongest relationship is for the mean intensity and the maximum-10 min rainfall. For both variables, the scaling rates are the highest among the other storm properties. Taking as a reference the quantile 95th, the mean rainfall intensity has a median value above CC, of 9.69%/ºC and middle range of scaling rates between 9-10.2%/ºC. The maximum 10-min rainfall has the strongest relationship with dew point temperature with the 95th quantile being 13.16%/ºC and with a interquartile range between 12-14%/ºC.

The study of the scaling rate between storm duration with mean dew point shows a negative scaling relationship. The results coincide with previous research carried out in Canada and Switzerland (Molnar et al., 2015; Panthou et al., 2014). For the 80th quantile, the median rate is -1.4%/ºC and stays almost constant with increasing quantile. The interquartile range is larger for increasing quantiles, reaching for the 99th quantile a range between -0.5%/ºC to -4%/ºC. In addition, for larger quantiles, some outliers point towards negative scaling rates up to -15%/ºC. Overall, the negative relationship indicate that shorter duration storms take place during warmer days. Taking into consideration the total rainfall depth, scaling rates for all 4 quantiles of study show similar scaling rate at 5.5%/ºC, with the interquartile range approximately 1%/ºC. There are no evident changes when increasing quantile, which indicate as well the relationship between rainfall depth and dew point temperature is weaker compared to other variables.

Obj 2. Evaluate the presence of super-CC rates when storm properties are related to dew point temperature

The 2 variables that have showed increases in scaling rates larger than the CC are mean rainfall intensity maximum 10-min rainfall intensity. From these 2, maximum 10-min rainfall show the largest scaling rates across quantiles, with also the smallest interquartile range of all the properties analysed. Already for the 80th quantile, the scaling rates are

When the relationship between peak intensity and dew point temperature is measured at different temporal resolution (10-min, 30-min and 1 hour), this research shows that the temporal resolution has a strong effect on the peak intensity variable, with greater scaling rates observed for higher resolution timescales (10-min). The results obtained (shown in this project for the 95th rainfall quantile) point out that the scaling rates at 10-min are on average 1%/ºC larger than compared to the 30-min analysis and 2%/ºC with the 1-hour case. Overall, the range of scaling rates for 50% of the stations goes from 12.5%/ºC-14%/ºC for the 10-min scenario, 10.5%/ºC-12%/ºC for the 30-min scenario and 9%/ºC-11%/ºC for the 1-hour scenario.

This behaviour, which is consistent with existing research supports the idea that pairs of hourly observations with dew point temperature might still not fully represent rainfall variability. The causes behind this increase in scaling rate might be related to increases of the intermittency for events measured at timescales of 1 hour or more (compared to those at the 10-min).

Obj 3. Investigate the significance of adding storm characteristics as covariates in regression models for dew point temperature-precipitation scaling rates

As this research pointed out in the Background section, mixing different types of storm events can result in unrealistic scaling rates (Berg et al., 2013). To address this issue, in this project the variables *mean intensity*, *maximum 10-min rainfall* and *total rainfall depth* are analysed conditioned to the event duration. With the aim to investigate if the scaling rates of the 4 variable studied change when events conditioned to duration are considered, the project first looked at the proportion of 3 event classes at each temperature bin (short, medium and long events). From the analysis to 9 stations that cover different geographical regions in Germany, it is shown that long events are the less frequent, with only contributing 10% of the total events. While there is a small increase of long events in dew point temperatures close to 10ªC, for larger temperatures there is slight decay of long events.

This is consistent with results from (Panthou et al., 2014) that also showed how long events were less predominant at higher temperatures. For short and medium events, the research shows that in most cases, medium events (between 2-10 hours) dominate with an average proportion at each temperature bin of 60%. The remaining 40% is associated to short events. In most cases, out of the 9 stations studied, short events show a slight increase from 40% to 60% of total events with increasing temperatures around 10-15ºC. This is at the expense of a reduction in medium and long events. Given that the analysis is presented for 9 stations, it is not possible to extend this as a general behaviour for the whole Germany. However, having chosen stations that cover different regions and different scaling rates, the research seems to align with current literature that indicates a larger proportion of short events at higher temperature ranges. In this case, this outcome supports the outcomes of the scaling rates between duration and dew point temperature, which showed that shorter duration events are located at larger dew point temperature ranges.

Obj 4. Investigate a methodology for predicting changes in storm event characteristics based on scaling properties and properties

In existing research, the study of scaling rates mostly focuses on understanding the physics behind the scaling rates relationship. There have only a few studies looking into applications for flood mitigation and adaptation. Westra et al. (2014) tried to bridge the gap between flood hydrology and atmospheric science by providing a list on the set of methods available for hydrologist to measure flooding. However, there was no connection on how the scaling rate could be incorporated to mitigate flood risk. In this section, the outcomes of this part of the research are used to motivate potential applications in the field of hydrology, mostly related to flood alleviation.

The level of detail provided by the storm characteristics analysed in this project highlights the importance of studying storm properties to better understand potential impacts of short but extreme rainfall. From literature, a summary on the main areas of work that studies in the field propose is presented:

- Studying the impact of climate change on the scaling behaviour of extreme short-duration precipitation (Van de Vyver et al., 2019).
- Applying the changepoint models and the storm analysis to characterise extremes in specific regions (Schroeer & Kirchengast, 2018).
- Trend analysis of daily and sub-daily precipitation aimed at studying how the scaling rate will change in the context of a warming climate (X. Zhang et al., 2013).
- Simulations of extreme precipitation events for current and warmer climate (Van de Vyver et al., 2019)

In this study, the focus is on simulating extreme precipitation events. It is widely agreed that the type, or in this case the duration of storms, can influence the scaling rates. However, it is complex to know,

for each storm, whether characteristics will be similar to convective or stratiform type. This is because, while some events during summer might be fully convective, there is still a spectrum of mixed rain types that can occur. For that, the use of storm properties becomes an opportunity to focus on the characteristics of the storm rather than the type.

Several studies have pointed out that there is no evidence of increased flood magnitudes with increasing precipitation extremes (Sharma et al., 2018). Some of the reasons proposed by Sharma et al. (2018) are:

- Shift to more frequent, higher intensity but shorter convective storms.
- Increase of temperature which result in decrease snowpack and earlier snowmelt.
- Increasing temperatures resulting in increased periods of drought.

Taking these arguments into account, the applications of scaling rate analysis should aim at planning for abrupt changes in hydrological extremes not assuming that with climate change there will be a constant increase in precipitation, but rather understand the dynamics and interaction between long and short events that, although less frequent, can have a larger impact.

With this in mind, what should be the steps to design a stochastic weather generator that can help cope with extreme precipitation events? To address this question, the generic structure of an stochastic weather generator developed by (Singer et al., 2018) is taken as a reference.

There are several steps involved in defining a stochastic weather generator which have been combined with the outcomes of this study:

1. Data input: The first step involves the definition of probability density functions provided for the variables of interest. In particular, PDFs on intensity-duration, storm duration, peak rainfall or total rainfall would be used, as well as more generic indicators such as monthly rainfall, seasonal rainfall or annual rainfall.
2. Check seasonal characteristics of storms: As this study has shown, different event duration can have different scaling rates. Therefore, it is key to know for the study area, the distribution of storm duration across the year. For example, in Germany, shorter storms associated to convective events are found more in summer than in winter. The aim of this step is to try to match as good as possible storm characteristics associated to a specific times-scale (month or season).
3. Threshold selection: Once the proportion of storms is known, the simulator can set a thresholds based on: a) The climate change scenario of interest: Current climate, climate-changed scenario based on RCP, etc... b) Total monthly, seasonal or yearly precipitation reached (based on historical observations from the PDF). In this step, the aim is to select the period of time and the total rainfall from which storms will be simulated. Relating it to Germany, this step would determine whether the storm simulator will generate storms for the summer or winter period (since that will have an influence on the storm characteristics that will be simulated).
4. Set the boundary conditions of the watershed. As it is mentioned above, climate change is already affecting the characteristics of a watershed, from the soil moisture to presence of snow,... Setting up the watershed characteristics will be needed for 2 reasons: to test the simulated storms and to change the characteristics of the watershed conditions (drier period, wetter soil moisture,...) and check the impact that the same simulated storm can have on an environment with different characteristics.
5. The next step would require to combine storm events based on their duration and characteristics. For that, the information provided in step 2 is comes to play. This is the section where this research would mostly address. From the study of the relationship between storm event and temperature, information on how much is the increase of precipitation per temperature based on storm duration can be provided.

   In this study, it has been shown that properties of short events do not necessarily scale on the same rate as long term. With this information, when simulating storm events, the combination of how short and long events scale and evolve over time could provide a better insight on the flood impact. To check what would be the impact that different storm duration could have when

occurring at similar times, incorporating a parameter like the interarrival time between storms would provide estimates on the flood impact of combining storms.

To conclude, the goal of a stochastic weather generator is to simulate observed historical rainfall characteristics for small watersheds. Thanks to the analysis proposed in this project, stochastic weather generators could now input more information on historic storm characteristics, differentiating not only on the rate of increase of several variables based on temperature but also on how events with different duration relate differently with temperature.

6

Part 3: Scaling rates in sub-Saharan Africa

In this last part of the project, the aim is to evaluate what are the scaling rates observed in 4 countries of sub-Saharan Africa using hourly precipitation observations. The following sections describes the data used and the results from applying a Binning Approach and Quantile Regression Models. Existing research has already shown that scaling rates vary across different regions in the world. In this section, a similar methodology as followed for Germany is tested to evaluate the extension of current methods for complex climates like the tropics and to investigate the performance of dew point temperature as a predictor.

6.1. Data characteristics

To start with, and following the methodology defined for the binning approach, in order to develop the temperature binning method, it is necessary to know the range of temperatures for which precipitation occurs. Figure 6.1 shows the temperature density function of each of the stations available. This is done only considering wet observations and is meant to give an indication on the potential number of bins that the model will have.

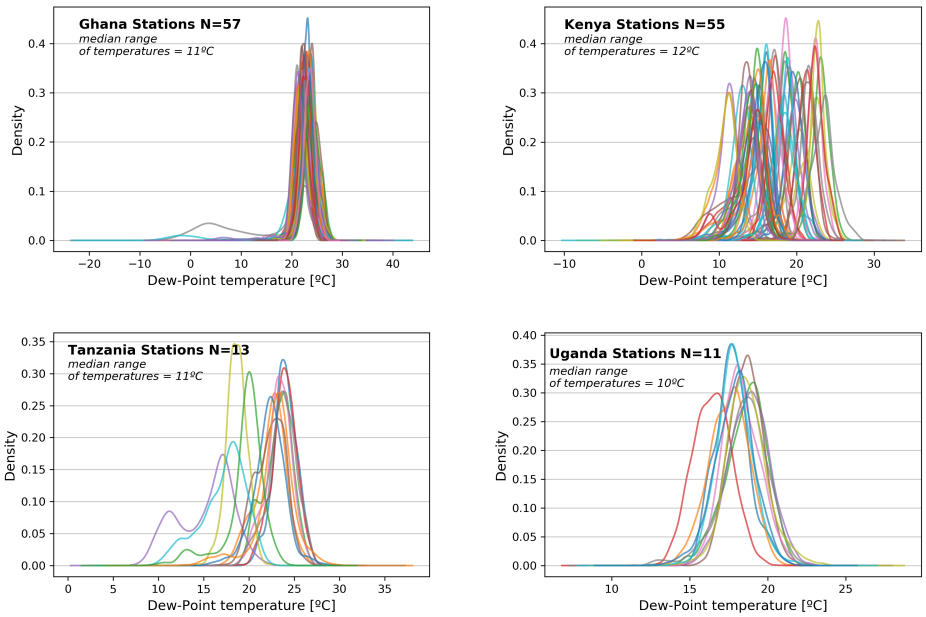


Figure 6.1: Density function of dew point temperature for each station, based on wet observations.

Overall, Ghana is the country with the most homogeneous range of dew point temperature between stations. For almost all the stations, precipitation occurs around dew point temperatures between 20-26ºC. Similarly, precipitation in Uganda occurs on the range of 16-26ºC.

The largest difference occur in Kenya, where the range of dew point temperatures can span for a larger set of temperatures. Apart from 2 peaks of data at 10 degrees shown by 2 stations, most of the stations show a small number of observations below that range, which, during the application of the binning method could result in temperature bins of very few observations.

With a focus on the extreme dew point temperature values observed, Figure 6.2 shows how Ghana, Uganda and Tanzania have a narrow window of dew pointtemperatures while Kenya has a wider range. Overall, the median range of temperatures with precipitation observations varies between 10ºC, [Uganda], 11ºC [Ghana and Tanzania] and 12ºC [Kenya]

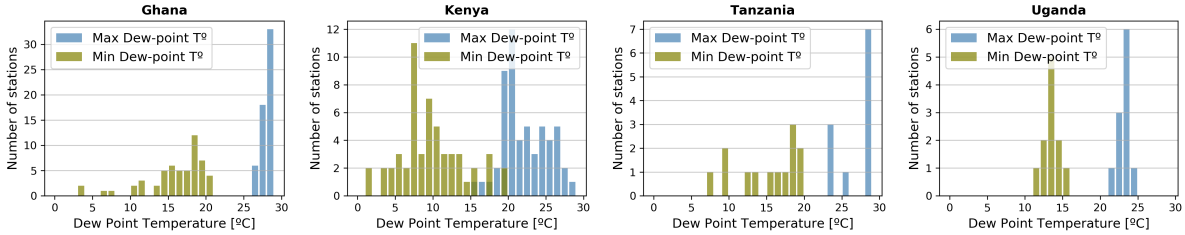


Figure 6.2: Distribution of maximum and minimum dew point temperature data for sub-Saharan African countries.

Another variable to consider before defining the binning approach is the length of observations. Given that the TAHMO project is still recent, there is not a large record of observations. As such, Figure 6.3 shows that the frequency in the number of wet observations available is quite low for extreme rainfall analysis. In general, the number of observations is around 6000, except for Uganda where the value drops to 4000.

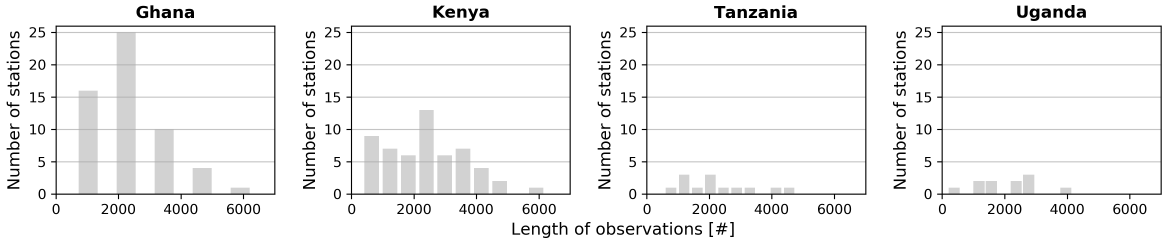


Figure 6.3: Frequency of wet observation for the sub-Saharan countries.

Given the limitation in the amount of data, it was decided to start first by estimating the scaling of the 4 stations with the largest number of observations (1 station per country). Table 6.1 shows the selected stations, with the total number of observations and the range of temperatures with precipitation data.

Table 6.1: Description of the initial stations to test

Country [station name]	Number of observations	Range of temperatures with precipitation data
Ghana [TA00011]	6099	14 - 27 ºC
Kenya [TA00140]	6205	4º - 14ºC
Tanzania [TA00493]	4767	16 - 28 ºC
Uganda [TA00214]	4155	13 - 23ºC

6.2. Binning Approach

6.2.1. Bins separated every 2ºC

The results from applying the binning approach (Figure 6.4) show that, grouping precipitation in 2ºC bins results in very different scaling rate results between the 4 stations of analysis. There are 2 main patterns observed: In one case, similar to what was observed for Germany, the initial temperature bins have a very small number of precipitation observations (Figure 6.4, station TA00493 and TA00140).

For those cases, there is a mismatch between the initial bin and the rest of the bins, which affects the outcome of the scaling rate estimate. This can clearly be seen when looking at the number of observations (grey points) around each bin. For example on station TA00140, the first 2 bins have only values in the order of 0.03 mm/1h, which are higher than the minimum threshold set but still very small.

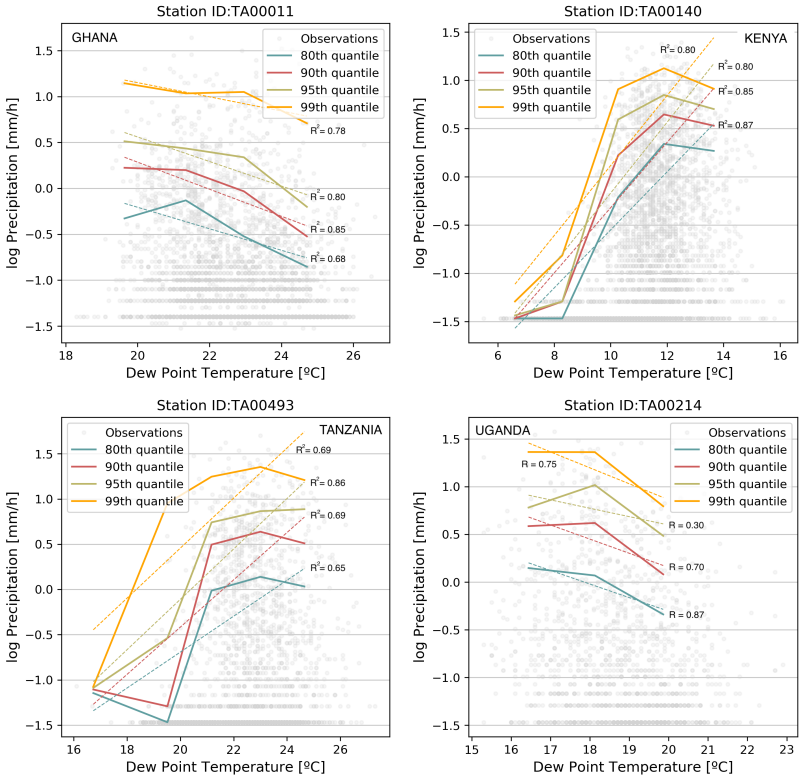


Figure 6.4: Results from scaling hourly precipitation with dew point temperature for 4 stations through the Binning Approach (every 2ºC bins). Solid lines are percentiles computed for each bin from raw data. Dashed line represents the fitted regression model.

For the stations TA00011 and TA00214, this situation does not occur, and the first bins are already located where most observations occur, with the temperatures of the first bin around 17-20ºC . In the Binning Approach method, a 2º degrees separation is used to fit the regression. For stations TA00011 and TA00214, there are 3 bins whereas for the other 2 stations, there are 5 bins.

As a result of these statistical effect in the first bins, the scaling rates observed vary greatly, as shown in Table 6.2.

Table 6.2: Scaling rate results using a Binning approach every 2^oC for different percentiles (0.80, 0.90, 0.95, 0.99)

Country [station name]	Scaling rates [$\%/^oC$]			
	80th	90th	95th	99th
Ghana [TA00011]	-23.61	-28.75	-18.31	-16.29
Kenya [TA00140]	99.92	118.01	132.84	130.56
Tanzania [TA00493]	57.95	82.50	89.87	89.01
Uganda [TA00214]	-27.95	-28.95	-26.47	-31.92

The first limitation that Figure 6.4 shows is that, given the small range of temperatures where rainfall occurs, applying a binning approach leads to a very small number of bins (between 3-5). In addition, out of these bins, it has also been seen that if there is a lot of very low-intensity rain, the first bin might cause bias in the estimation of the scaling slope. In the cases where the initial bins alter the overall scaling, it has already been seen that the results observed have positive scaling rates higher than would be expected if the first bins would be excluded.

Comparing these results with Germany, one of the approaches used to correct the influence of biased bins due to small number of observations was to not consider them for the regression model. In this

case, given the small range of temperatures, such method is not recommended because it would reduce the number of bins that would be used to computing the scaling rate, fitting a regression with less than 4 points (for example on station TA00140).

Instead, the approach that is taken is to test the binning approach with equal number of observations per bin, to see whether a similar behaviour occurs and to investigate if the observed scaling rates would be the same.

6.2.2. Bins with equal number of observations

For this model, some changes on the number of bins compared to Germany were made. For that, the number of observations is taken into account (for the 4 stations of analysis between 4000-6000 wet observations), and the range of temperatures (10 degrees). With this, it is chosen to have 10 fixed bins, which will give at least 150 measurements per bin. Results of the scaling rates can be seen on Table 6.3. Compared to the binning approach every 2ºC bins, the scaling rates estimated have a more steep behaviour, both for the stations with negative slope (TA00011 and TA00214) and positive slope (TA00140 and TA00493). The stations that had a decreasing scaling rate have in this model reach steeper negative scaling of around -33%/ºC and-35%/ºC for varying quantiles and for the stations where a positive scaling rate was observed in the previous model, now steeper positive scaling rates are seen, with the Uganda station TA00214 giving results of 137%/ºC scaling rate.

Table 6.3: Scaling rate results using a Binning approach with equal length of observations for different percentiles (0.80, 0.90, 0.95, 0.99)

Country [station name]	Scaling rates [$\%/^{o}C$]			
	80th	90th	95th	99th
Ghana [TA00011]	-35.697246	-36.817485	-34.057108	-37.550681
Kenya [TA00140]	137.353592	132.272046	131.293814	126.134327
Tanzania [TA00493]	77.493935	83.621949	55.656047	12.727954
Uganda [TA00214]	-33.910815	-32.519431	-27.206641	-8.159791

Looking at the percentiles of rain computed for each bin from raw data, (Figure 6.5) and comparing them with the results from the previous binning approach, it shows that the trend of the scaling rate is the same, but now with larger number of bins. For example, taking station TA00011 as a reference, it can be seen that while in the binning approach with 2ºC bin separation, there is a straight decreasing line, for the model with equal number of observations there are more peaks within the same range of dew point temperatures (zig-zag behaviour).

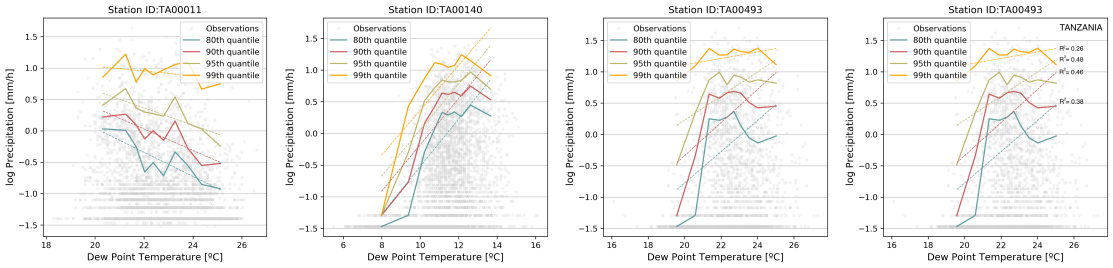


Figure 6.5: Results from scaling hourly precipitation with dew point temperature for 4 stations through the Binning Approach (equal number of observations). Solid lines are percentiles computed for each bin from raw data. Dashed line represents the fitted regression model

6.3. Quantile regression

Lastly, the results from applying the quantile regression are presented in Table 6.4. The results vary per station but, like the previous models, the stations for Ghana and Kenya have negative scaling rate and the stations for Tanzania and Uganda show a very steep positive slope.

The results of these scaling rates are shown in comparison with the Equal Observation results in Figure 6.6. As it can be seen, the main difference between the results is that, on the small temperature ranges, the scaling rates based on quantile show a less pronounced slope compared with the equal

Table 6.4: Scaling rate results using the Quantile Regression for different percentiles (0.80, 0.90, 0.95, 0.99)

Country [station name]	Scaling rates [$\%/^{o}C$]			
	80th	90th	95th	99th
Ghana [TA00011]	-35.39	-30.57	-24.57	-7.39
Kenya [TA00140]	116.75	130.44	115.78	56.12
Tanzania [TA00493]	46.43	36.13	27.49	7.72
Uganda [TA00214]	-31.64	-36.15	-36.64	-39.64

observations method. This is because for equal observations the line is fitted based on bins at all temperature range whether the quantile bases the overall slope based on all measurements.

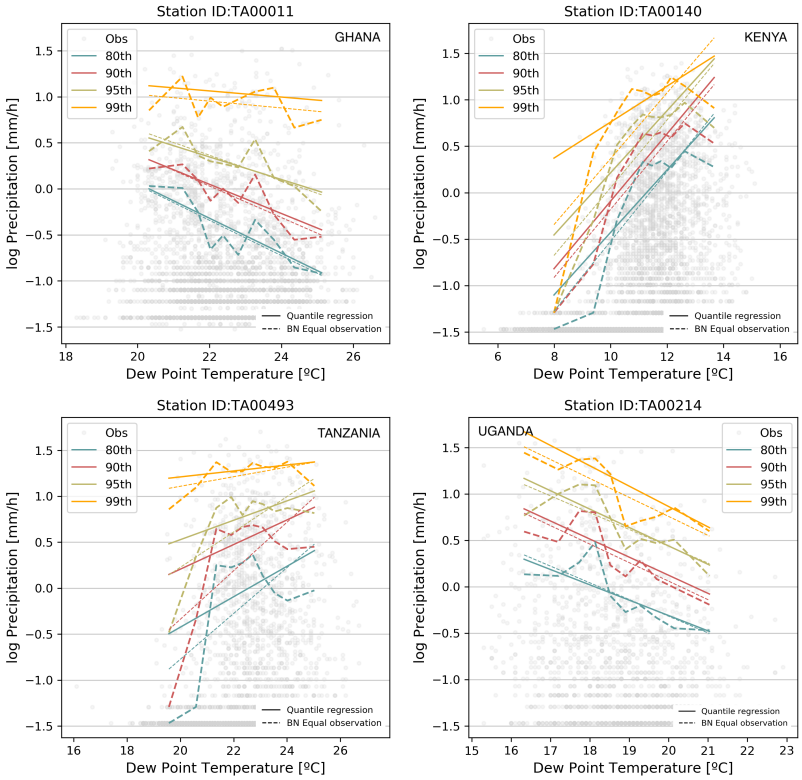


Figure 6.6: Results from scaling hourly precipitation with dew point temperature for 4 stations through the Quantile Regression. Dashed lines are percentiles computed for each bin from raw data. The straight dashed line represents the fitted regression model for Equal number of observations and the bold line represents the Quantile regression

While quantile approaches do not classify per bins the precipitation dataset, the positive results observed for Kenya highlight one interesting characteristic of this method which was not seen for Germany. In the case of the Kenya station, the range of dew point temperatures with very few measurements ranges between 6 and 10ºC. For Germany, the initial bins with few observations only occurred for 1ºC. While in the Germany cases the small number of measurements in first bins did not influenced the overall scaling rate obtained by the quantile regression, results in sub-Saharan Africa show that if there is a large range of initial temperatures without enough observations, the quantile regression will consider those points.

In order to obtain a scaling rate that was representative of the large bulk of rain observations (in the case of the Kenya station between 10 and 14ºC dew point temperature), an improvement of the quantile regression approach was proposed. The method applied consisted of:

1. Divide the temperature in bins and calculate the number of observations per bin.
2. For the initial bin, set a threshold of minimum observations. Since the sub-Saharan dataset has lower observations than Germany, the threshold was not set to 60 observations but to 20. The value was chosen after looking at the average number of observations at the initial bins of all stations.

3. Apply the quantile regression from the minimum dew point temperature that passed the threshold of the minimum number of observations.

As it will be shown in the following sections, applying this methodology allowed for the quantile regression model to estimate the scaling rate for the large bulk of data and exclude the initial bins where only a couple of observations were observed.

6.4. Analysis of results

For all the stations of the TAHMO dataset that passed the quality test, the quantile regression was applied. In the following section, a more in-depth analysis on the extended results as well as the study of the potential causes is explored. More examples on the scaling rates per country and quantile are given in Appendix B.

In Figure 6.7, the complete results from the quantile regression is presented. For most stations, the scaling rate obtained have negative slope. As expected, the countries where more stations are studied (Ghana and Kenya) have a larger range of scaling rates. In Kenya, for the 95th quantile, the interquartile range is between 14 and -20%/ºC and for Ghana between -10 to -20%/ºC. For countries with less stations studied (Uganda and Tanzania) the scaling rates vary less, but show similar results, between -17 to -30%/ºC and -3 to -20%/ºC respectively.

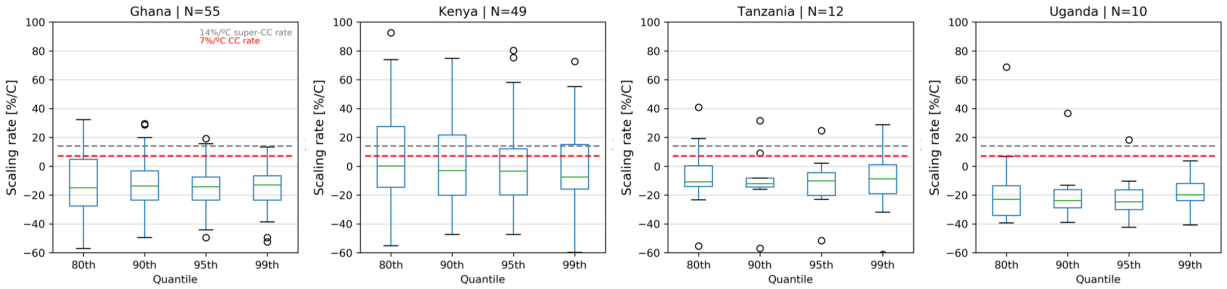


Figure 6.7: Boxplot of scaling rates for quantiles 80th, 90th, 95th and 99th computed with the quantile regression for the TAHMO dataset.

From the majority of stations, it is clear that a negative scaling is found. In particular, Kenya is the exception, having the larger number of stations with positive scaling. A visual representation of the scaling rates for the stations in Kenya reveals that the cases in which positive or negative scaling rate are obtained is strongly dependent on the range of dew point temperatures where precipitation occurs. To check how that compares with the other countries, the range of dew point temperatures where rainfall occurs was retrieved for all locations (see Table 8.1 in Appendix B). The results show that, when the range of temperatures is between 11-17ºC, the scaling rates are positive (mostly in Uganda and Kenya), whereas for 18ºC up until 26ºC the scaling rates are negative. This finding might suggest that the negative scaling is a feature caused at the very large temperature, which would go in line with the outcomes by Hardwick Jones et al. (2010).

However, the range of temperatures obtained with the TAHMO datasets could be influenced by the fact that the length of observations is small (3 years). To check that the range of temperatures where precipitation occurs is well sampled, a comparison with an analysis of stations from the GSOD (Daily observations) is carried out (some examples of the GSOD provided in the Appendix B, Figure 8.7). For those stations, daily measurements are available for at least 10 years. Exploring the range of temperatures where precipitation occurs and the overall scaling rates, there are 2 main outcomes:

- With more years of analysis, the range of temperatures of precipitation does not change compared to the hourly observations of the TAHMO dataset. For Ghana and Tanzania it is between 22-26 degrees dew point, for Kenya between 11-17ºC, and for Uganda between 15-21ºC approximately.
- For the Daily observations analysis, there is a tendency of stations with temperature below 17ºC to show a positive scaling rate and for temperatures larger than 23ªC, to have a negative slope.

For daily observations, several stations show a peak-like structure (see Figure 8.7 in Appendix B.) with an increase of precipitation with temperature until a peak at around 23ºC dew point temperature is

reached. This feature is not shown in hourly observations. The actual reasons behind these differences are not explored in detail but some hypothesis are presented: It could be because larger length of observations better capture the precipitation-temperature relation (as there is a distribution of rainfall for at least 10 years instead of only 3). The difference could also be related to the fact that one dataset is based on hourly observations while the other has daily measurements. Averaging rainfall at a daily resolution has been seen not to fully capture describe rainfall patterns with temperature. In addition, one possible explanation could be that the comparison is based on different stations, with different equipment and subject to different quality tests.

From the analysis in sub-Saharan Africa one can ask what can be the explanations for having consistent negative scaling rates. Some hypothesis and previous research behind negative scaling rates is already mentioned in the Literature Review. In this section, some of them are explored for the cases in sub-Saharan Africa. In Hardwick Jones et al. (2010), they argue that one possible explanation for why the scaling rates observed are negative is because of a reduction in the relative humidity at higher surface air temperatures. For that, they looked at the relationship between daily relative humidity and temperature, which showed that their decay in the scaling rate occurred at the same temperature where relative humidity started to decay. In a similar way, for each station, the relationship between relative humidity and surface air temperature for each hour is calculated for the TAHMO dataset, as shown in Appendix B (Figure 8.5). Results are presented for some stations but already show that, indeed, for very large SAT there is a reduction of relative humidity. Interestingly, in those plots, it can be seen that for smaller SAT, in some cases there is constant horizontal range of values with large relative humidity, up until a point where the relative humidity drops. This could explain why, for example in Kenya, stations with smaller range of dew point temperatures resulted in positive scaling rates, as opposed to larger temperatures which resulted in negative ones.

In past research in The Netherlands and Canada, using dew point temperature has resulted in a better predictor compared to surface air temperature. However, as Panthou et al. (2014) pointed out, it is not always the case, since, for example in coastal areas of Canada, the use of dew point did not prevent the decay at large temperatures (meaning that relative humidity was not the reason why the decay occurred). For the outcomes of the TAHMO stations, the relationship between relative humidity and SAT shows similar behaviour to previous research: is is within a constant range for smaller temperature and decays for larger temperatures. The fact that when using dew point temperature, the majority of stations show negative scaling rate in sub-Saharan Africa could mean that, similarly to what Panthou2014 found for coastal areas, relative humidity alone could not explain why the decay occurs.

To try to infer potential explanations about the causes of differences in scaling rates, the spatial distribution for the 90th and 95th quantile are presented in the Appendix B (Figure 8.6). From the spatial distribution, it is difficult to determine a special pattern in scaling rate, both because the scaling rates are very different between the stations and because the distribution of stations in each country is not homogeneous (some regions are very well represented and other do not have many stations).

Overall, from the comparison between quantile regression and binning approach with equal number of observations, it can be concluded that the method itself is not the cause of obtaining large negative scaling rates (since similar results are obtained between methods). Undoubtedly, the fact that the data available is only for a period of 2-3 years has an impact, even more considering that the regions are characteristed for having long dry periods. However, the outcomes match existing research that suggest negative scaling rates and, while not specific conclusions can be made on whether the scaling rates found are correct, the study of the range of temperatures, the influence of the relative humidity provide the grounds for further analysis.

6.5. Summary of results

The discussion of the initial results from different exponential regression models for Ghana, Kenya, Uganda and Tanzania is presented. Through the application of a binning approach to Africa, it is shown that for the studied locations, precipitation occurs in a smaller range of temperatures (~18ºC-26ºC). This has implications when applying the binning method, resulting in only 4-5 bins. This leads to an large changes of precipitation percentiles across bins.

Plotting the results of the binning approach for several stations shows that, for the range of dew point

temperatures between 20-25ºC, there is a zig-zag behaviour. The scaling rates estimated for the 4 stations with the larges number of observations shows very extreme results. In general, there are 2 patterns observed:

1. Similarly to the stations in Germany, a small number of observations in the initial bins can lead to very high and unrealistic scaling rate.
2. On the other side, if that initial influence of bins does not occur, or is taken out, the scaling rates shows a very steep negative scaling rate.

Similarly to the case of Germany, the African results show that, from the 3 regression models, the scaling rates follow a similar pattern. Those stations that have a negative scaling rate for one model, have a similar behaviour for the other 2 models.

7

Discussion

In this chapter the main results from this research are presented Firstly, a critical analysis on the impact that different models have on estimating scaling rates is presented. Secondly, the outcomes and applications of studying scaling based on storm properties is presented. Thirdly, the validity of using dew point temperature for warm climates is evaluated through the hourly results from sub-Saharan countries. Lastly, the limitations of this research are also proposed.

7.1. Discussion on existing methodologies for estimating scaling rates

An initial analysis on the influence of length of observations indicates that studies have an inherent uncertainty already based on how many years of data are available. Most studies looking at scaling rates have used daily or hourly data for a period of approximately 10 years or more (Hardwick Jones et al., 2010; Ali & Mishra, 2017; Panthou et al., 2014). From a randomised analysis at multiple lengths, it is shown that, for high quantiles (wet 99th), which are normally of interest in this type of studies, 9 years of data can have differences in scaling rates between 0.7 and -1%/ºC compared to larger datasets (25 years), in some cases reaching more than 2 and -2%/ºC.

The implications of this are that, when comparing analysis between stations with different length, like most projects do, a difference of 1 or 2%/ºC can be due to data limitations instead of other factors. Therefore, there is a trade-off between data quantity and consistency. From this project, it is recommended to work with stations that have less than 5 years difference between them, since my results show that only some outliers reach a scaling rate difference of -1/1%/ºC.

With a focus on the methods, recent studies have showed that relationship between precipitation and temperature might not be constant across the temperatures (Westra et al., 2014). There is only one research that has looked at change-point detection in Germany (Van de Vyver et al., 2019). In their study they showed that, in Berlin, the change point detected was 13.6º, 8.4ºC and 7.5º for the quantiles 0.90, 0.95 and 0.99 respectively. In my research, such analysis is extended to locations which offers a better knowledge on the spread and variability of change-point detection.

Through the analysis of scaling rates in Germany, the results from the quantile regression have shown that the Clausius-Clapeyron relationship is exceeded, for some stations, for the 90th quantile onwards. In fact, for the 99th quantile all stations with CC above 7%/ºC. These results coincide with (Van de Vyver et al., 2019), which for the quantile analysis the slopes for the 0.90, 0.95 and 0.95 for Berlin reach 9.5, 10.1 and 10.4%/ºC respectively (Figure 7.1).

There are between 236-287 stations where a change-point is determined (depending on the quantile). Compared to other projects, change-point analysis have only been carried out at individual stations from various countries (Lenderink & Van Meijgaard, 2008; Van de Vyver et al., 2019; Pumo et al., 2019). Through my research, I have proved that the precipitation-temperature does not follow a constant value

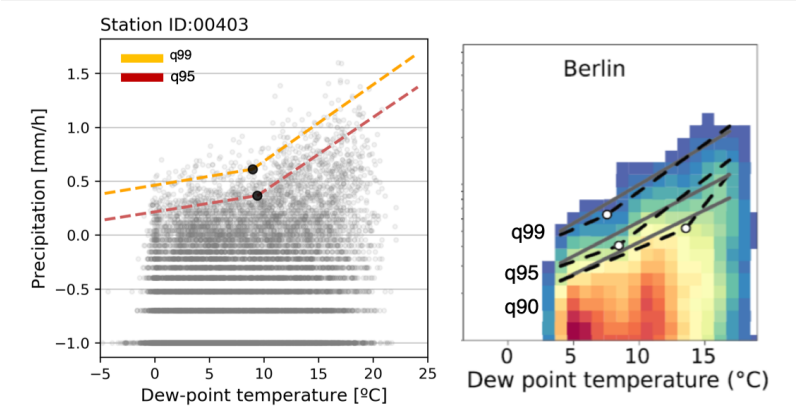


Figure 7.1: Comparison of results for a station in Berlin between this research (left) and the one by (Van de Vyver et al., 2019) (right).

across temperatures for most of locations in Germany, stressing out that it is not a feature of individual places, but rather a generalised pattern.

There are also 2 main findings concerning the change-point analysis that are interesting to discuss: the confidence interval of results and the comparison of change-point between quantiles and locations. From my study, the smallest confidence intervals of the parameters is β_1 . This matches research with similar methodologies but in different locations. For β_2 and T_c, the confidence interval is much larger. A suggestion to reduce confidence interval would be to group 2 or 3 nearby stations with similar results. This would be possible since, from the spatial analysis provided in this project, it shows similar scaling rates for nearby stations. Specifically for Germany, the project by (Van de Vyver et al., 2019) grouped 4 stations to determine the change-point. That allowed to have larger total years of data.

There is a contrast when comparing the results from scaling rates from quantile regression and change-point model. In the first model, an increase in quantile leads to increases in the scaling rates, reaching median values of 10%/ºC for the 99th quantile. This is a constant through all the stations and, as previously shown, for higher quantiles it results to scaling rates between 9-11%/ºC. The results from the piecesise analysis suggest that focusing on a quantile regression might overestimate the scaling rates for the lower temperatures.

In Figure 7.2 a schematisation is made on the distribution of rain across temperatures. This is a patterns seen for all stations in Germany. The distribution of rain across bins has 2 main regimes, when looking at the extreme percentiles. Between 80th and 99th quantile, before a dew point temperature, there is an almost constant precipitation with increasing temperatures, up until a change-point where with increasing temperatures the precipitation increases.

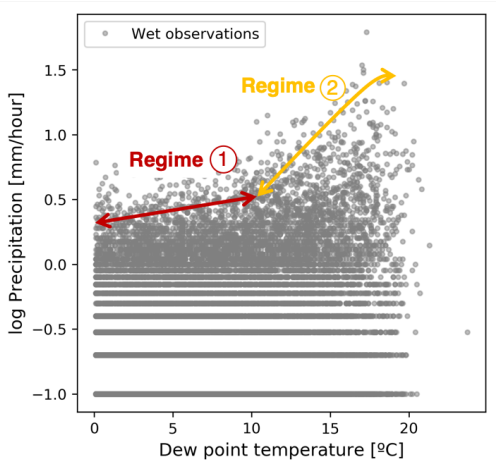


Figure 7.2: Schematic of the dew point temperature-precipitation relationship for Germany

Another way to this can be shown is by looking at the range of β_1 values for increasing quantiles. In all quantiles, the range of scaling rates go between 5 and 7%/ºC. However, when β_2 increases when more extreme quantiles are used. This indicates that it is the slope at the largest temperature range what increases with higher quantiles, and not the low temperature range.

7.2. Discussion on sub-Saharan Scaling rates

In this project, the study of the scaling rates of 4 sub-Saharan countries with the quantile regression method has shown that the presence of negative scaling rates is a constant feature in most stations. In existing research, it has been proposed to use dew point temperature instead of SAT because, in research done in several locations the model corrected the decay observed at larger temperatures.

In this project, the aim was to evaluate if incorporating relative humidity (by using dew point temperature) would result in a correction of these negative scaling rates. Even though the TAHMO hourly observations are only available for $\tilde{3}$ years, both these datasets and the GSOD daily observations show similar negative results. As opposed to existing research, using dew point temperature still results in negative scaling rate which indicates that the explanation of why the decay occurs might not be related to the reduction in relative humidity. The differences between previous studies and mine is that so far, the correction done using dew point has only been tested mostly in regions with different climate than the tropics (The Netherlands and Canada).

To find a better explanation there are several things that can be done. In general, there are 2 main factors that in the tropics are important to consider when looking at the relationship between extreme rainfall and temperature: the moisture storage capacity of the atmosphere and the moisture saturation. In a warm climate, the atmosphere has a higher moisture holding capacity (Berg et al., 2009a). Therefore, it is not as easily saturated as colder climates, like for example Germany. In previous studies in the tropics, Hardwick Jones et al. (2010) pointed out that not only capacity is important, but the amount of moisture available is also key. As this research has also shown, the relationship between SAT and relative humidity for sub-Saharan Africa indicates that for higher temperatures, the relative humidity is reduced. This means that the moisture availability for larger temperatures is small in the first place. The results from the relative humidity relationship with SAT, one could see that there were differences between stations, it would be interesting to consider what the relationship would be between stations close to water bodies. In general, oceans can contribute up to 85% of water to the atmosphere (Bigg et al., 2003; Hardwick Jones et al., 2010).

Panthou et al. (2014) evaluated in inland and coastal stations whether the use of predictors that incorporated the relative humidity (dew point temperature) eliminated the decay found for larger temperatures. Their results showed that there are indeed differences and, while for near-ocean stations humidity did not explain or corrected the decay from CC it did smooth the results for inland stations. Such a differentiation for the stations in sub-Saharan Africa could give more information on the root causes of the negative scaling.

On another topic, in this research the range of dew point temperature where precipitation occurs, in Kenya, showed that, whereas for large temperature scaling rates are normally negative, for smaller temperatures the scaling is positive. This reinforces the idea that what drives negative scaling occurs at the very large range of temperatures.

First, one could catalog the stations between inland and coastal and compare in more detail the differences between P-SAT and P-T_{dew} in order to see if there is an influence of relative humidity into the deviations of scaling rates.

The results obtained in this project suggest that maybe, the reasons behind negative scaling might be linked to other factors. One direction is towards the hypothesis of Drobinski et al. (2016), who stated that negative scaling is a result of arid surface conditions. He argued that surface air temperature or dew point temperature alone do not work as a good proxy for relating to extreme rainfall. In the process of finding better predictors, Roderick et al. (2019) presented the results from the relationship with Integrated water vapour (IWV). This variable is obtained through satellite data by combining moisture at different pressure levels in the vertical air column. As it is discussed already in the Literature Review and Block 3, the use of IWV showed an increasing relationship with SAT. One potential recommen-

dation for future research would be to explore the relationship between IWV and extreme precipitation for the sub-Saharan Africa stations. If it was a better predictor, this could be used as a as opposed to dew point or SAT temperature. Only by finding better predictors would it be possible to move towards studying potential implications in a warming climate.

To conclude, there is one aspect that I believe is important to highlight when addressing extreme rainfall in a warm climate. Although it has not been fully exploited for the sub-Saharan Africa dataset, the analysis of wet observations with all observations carried out in Germany highlight that in climates where there is a big fluctuation of wet fraction over the year, or where the wet fraction is smaller (like in the tropics), it is crucial for impact-based assessments to include the analysis with wet and dry observations. In future research where the study of scaling rates is directed towards taking decisions of flooding projects, it would be necessary to include wet and dry observations analysis.

7.3. Discussion on storm event properties

An important feature for Europe and, in particular for Germany, is that different types of storms are characteristic for different seasons (Berg et al., 2009a). For example in Germany, convective storms are predominant for the summer and longer stratiform precipitation are more common in winter. This means that while scaling analysis with fixed intervals can give an insight on the how precipitation changes with temperature, the applications for impact-based assessment or flooding projects need to be related to the storm characteristics. I believe that this makes the analysis of storm events as a point of inflexion in the application of scaling rates.

From the analysis, results indicate that for Germany, shorter duration events occur mostly at larger dew point temperature range. In addition, an analysis of 10-min, 30-min and 1-h peak intensity has shown that hourly observations have smaller scaling rate because of intermittency, meaning that, in real life, the peak intensity could increase higher with dew point temperature. This matches previous research by (Schleiss, 2018) that showed for the US, that intermittency can mask the real scaling rate.

These two outcomes indicate that thanks to storm-based analysis, it is now possible to have an order of magnitude of how sensitive different duration storms are to changes in dew point temperature. This opens a wide range of opportunities in simulating storms for seasons with predominant duration types or for specific applications like in urban drainage.

One of the main objectives with the study based on storms was to test if scaling rates larger than the CC were also observed for storm properties. The results for Germany have shown that mean and peak rainfall intensity show higher scaling rates with increasing quantiles. This means that, for peak rainfall, the highest values are consistently found at higher dew point temperatures. However, through this analysis all storm events are considered (1 hour events versus 10 hours), which makes it difficult to draw conclusions. For that, the scaling rates were conditioned to 3 duration, short medium and high. In the case of maximum rainfall intensity, results show that regardless of the duration of the storm, there is a steep increase of peak intensity at higher dew point temperatures.

On the other side, in the analysis about total rainfall depth, no major sensitivity was found with dew point temperature. When the storm events were classified according to the duration, results showed that short events show a steeper increase of total rainfall depth with temperature. This suggests that, for short events, higher temperatures normally lead to larger total rainfall (more accumulation of water). On the other side, for long and medium events, no major sensitivity with dew point temperature was found, indicating that the total rainfall depth is not really influence by the dew point temperature.

7.4. Limitations of current research

One of the main challenges during this project has been the validation of the scaling rates. It is a challenge not only for this project, but for anyone who estimates scaling rates based on observations. This limitation is, partly, the reason why I decided to research the first block, in which I have looked and compared different method for estimating the scaling rate. Thanks to my project, people in the field can now have a better idea on the uncertainties behind each method, and the potential issues that might arise when applying them at different locations or with different variables.

One limitation of my research lies in the applicability of the results. As it shown in all the blocks of

analysis, there are several sources of uncertainty: From the data input, to the method used to obtain the scaling rate. With these uncertainties combined, it is difficult to address what the real scaling rates are, or what conditions result in stronger relationships between temperature and precipitation. To address this issue, this project has focused on a very fine analysis at a sub-hourly level. However, not many locations have access to such high resolution data, which raises the question again on what uses can be proposed to incorporate scaling rates.

With a focus on the sub-Saharan analysis, the main limitation has been the small length of observations available. While the dataset was of very high resolution (5-min observations), scaling rate analysis mostly depend on long observation analysis. This constraint together with the fact that sub-Saharan African research in this field is scarce, has made it difficult to interpret the results obtained. The lack of observations was addressed by including daily observations that span for 10 years (GSOD dataset). However, since observations are provided at a daily level, the total number of observations did not really increase significantly compared tot he TAHMO dataset.

8

Conclusions

The relationship between precipitation and temperature is known to be affected by several factors, from the percentiles used, whether the study is done for all year or between seasons, the time scale, the region or the type of models implemented. In this study, the sensitivity of extreme rainfall to dew point temperature has been studied both at a fixed hourly time interval and based on storm events, for over 300 stations in Germany and several stations in Ghana, Kenya, Tanzania and Uganda. While it is unfeasible to study all the influencing factors all at once, with this project, those related to the methods behind have been addressed. In this section, the main research question will be addressed, and the key findings will be presented, summarizing the main results for each of the three parts of this thesis. After that, a set recommendations and suggestions of future work are proposed.

The main question that this thesis aimed to address was what is the influence that different storm event characteristics have on the scaling rate, and whether that could be used to predict extreme rainfall characteristics with changes in temperature?

In an analysis of the scaling rates of storm properties such as peak rainfall, mean rainfall intensity, total rainfall depth and duration, it can be concluded that some storm properties, in particular peak rainfall and mean rainfall intensity, are particularly sensitive to increases in dew point temperature (Figure 8.1). In fact, for peak rainfall, the scaling rates already on the 80th percentile exceed the CC rate. For the largest percentiles, the scaling rates exceed even the super-CC (> 14%/ºC).

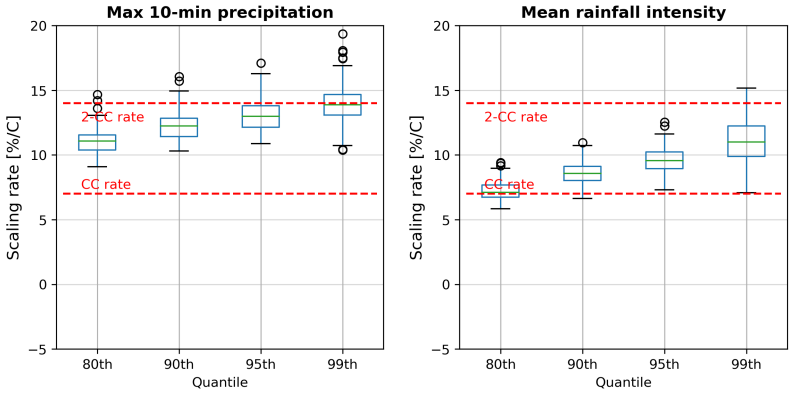


Figure 8.1: Boxplot of scaling rate results for the variables with largest sensitivity to dew point temperature.

The implications of the results from the storm analysis are large. To begin with, the application of the results for the peak rainfall intensity could be beneficial, for example, in the field of urban drainage, or in the quantification of pluvial flooding in cities, where peak rainfall is an important variable. In addition, one key outcome of the analysis based on storms, is that conditioning the study of scaling rates to the duration can result in differences in scaling patterns. This was mostly evident for the total

rainfall depth property. When all events were considered, constant scaling rates around 4-5%/ºC were observed for increasing quantiles. However, when the analysis was done by grouping events in storm duration classes (short, medium, long), results showed an increase in scaling rates up to CC rates for short events, with a transition to higher rates around 10ºC. This means that, when larger dew point temperatures occur, higher total rainfall depths are observed for short duration storms, whereas, for long storms, the influence of temperature is less. In climates where seasons are characterised by different types of storms, the outcomes of the analysis indicate that in a warming climate, it could occur that total precipitation in seasons where long events occur most would not change as much as seasons where shorter events are predominant (under the hypothesis that the occurrence of the events would be the same as in present conditions).

There are, however, several limitations on the applications of studies based on storm events, due to the access to high-resolution observations. From an analysis of peaks at different time resolutions, scaling rates showed that at a sub-hourly resolution (10-min and 30-min times), the scaling rates observed can be up to 1%/ºC higher than when compared to hourly observations (Figure 8.2). In 1-hour observations, there was a 60% of mean intermittency which shows that, already at an hourly observation, results are smoothed and affected by short periods of no rain.

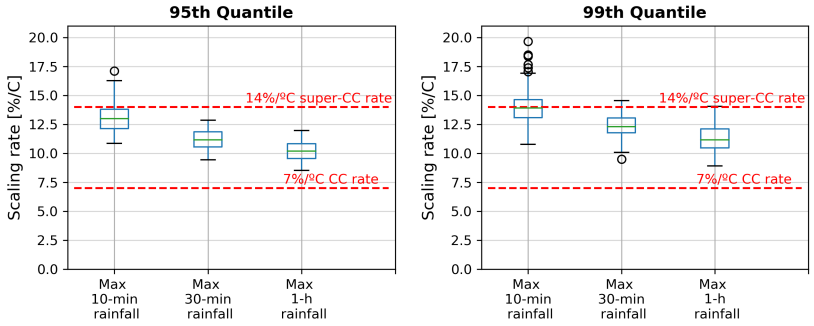


Figure 8.2: Boxpot of scaling rate of peak rainfall intensity at varying temporal resolution: 10 minute, 30 minute and 1 hour rainfall. Computed for 98 stations through the quantile regression model.

In a future warming world, the change rate of heavy precipitation amounts are generally expected to increase more than that of annual mean precipitation (Fowler et al., 2007). Therefore, I believe that the study of scaling rates should concentrate in better understanding the relationship between storms and temperature instead of trend analysis on hourly observations. Within this project, the study of storm properties has given a first indication on the potential increase of total rainfall, or mean intensity with dew point temperatures. This information can now be used, for example, in stochastic weather simulators, where different storm characteristics are added as probability density functions in order to generate random storms. With the scaling rates obtained, stochastic weather simulators could now simulate, for example how the influence of having different type of storms together would affect the flooding (given that in my research a larger sensitivity to shorter-duration storms has been shown).

8.1. Key findings

Apart from the main question, in this project the relationship between extreme rainfall with dew point temperature has been studied from multiple perspectives. With the aim to address the limitations that exist when estimating scaling rates, the following section addresses the questions related to: (i) Sensitivity analysis of existing methods to estimate scaling rates and (ii) performance of dew point temperature in sub-Saharan Africa.

Based on the sensitivity analysis of existing methods to estimate scaling rates, the main findings are:

- Binning precipitation in temperature bins can result in biased scaling rates, caused mostly by the undersampling of precipitation observations at the extreme temperature bins.
- Quantile regression methods are unbiased and provide scaling rates based on the total amount of observations instead of limited observations per bin.

- A comparative between quantile regression models and binning method where bins have equal number of observations has showed very similar scaling rates, although quantile regression has the smallest confidence interval from all the methods.
- For 80% of the stations analysed in Germany (out of 350 stations), a change-point model is preferred over quantile regression, based on the Bayesian Information Criteria (BIC). An analysis on the change-points obtained, indicate that most stations have a change-point between 8º-13ºC. From this model, it is also observed that the scaling rate for the range of temperatures before the change-point is constant with values slightly lower than the CC rate, whereas slopes after the change-point indicate steeper slopes that exceed, in most cases, super-CC rates (> 14%/ºC).
- The results from the change-point model indicate that while there is change-point detected, the confidence interval (mostly for the slope after the change-point) is still very large.
- When comparing the slopes between change-point models and quantile regression, this study shows that using quantile regression alone results in large increases in scaling rates with increasing quantile. When change-point models are considered, the slopes before the change-point occurs stay almost constant over increasing quantiles.
- With a focus on the Africa analysis, this study shows that, in order to use existing methods for measuring scaling rates, it is vital to check that there are enough number of observations in the extreme bins, and concentrate only on fitting the regression on those bins that are represented by a large amount of observations, to avoid unrealistic scaling rates driven by data undersampling.

Based on the performance of dew point temperature for estimating scaling rates in sub-Saharan Africa:

- There is a consistent negative scaling rate observed for most stations in the countries studied. From all the countries analysed, Kenya is the one that has a larger range of scaling rates, with both positive and negative slopes identified.
- A study on the range of temperatures where precipitation occurs indicates that, for lower dew point temperatures (found mostly in Kenya), the scaling rates are positive, whereas for larger dew point temperatures (Ghana and Tanzania), the scaling rates are negative. This result suggest that it is the very high dew point temperatures that influence the negative slopes.
- Looking at the relationship between relative humidity and surface air temperature, it is shown that a decay occurs at the very large temperature. The fact that using dew point temperature (which is a proxy for relative humidity) instead of surface air temperature does not lead to more positive scaling rates suggest that dew point temperature might not be the only reason causing the decay at larger temperatures.
- To study the patterns of atmospheric moisture over the tropics, a study differentiating between coastal and inland stations will provide better insight on the actual performance of dew point temperature as predictor.
- Based on recent research in the field, the exploration of integrated water vapour and its relationship with SAT and precipitation is encouraged for the sub-Saharan African stations.

8.2. Further research and recommendations

Some of the suggestions for further research have already been provided during the discussion of results. However, there are addition points worth to mention. For any person thinking about researching in the topic of scaling rates, it is important to think that there is uncertainty. Not only on the methods used and the scaling rates obtained, but on the general understanding of useful applications for climate change studies. Future research should look more into the role of covariates, combining several sets of properties and testing what combination can best describe the relationship between precipitation and temperature. In this project, this has been mainly done by conditioning storm properties to duration, but there are many other variables that can be included, such as time of the year (to infer changes in seasons) or intermittency.

Concerning the analysis on data-scarce regions with complex climates, such as in sub-Saharan Africa, one of the main recommendation for future research is to use the TAHMO dataset to characterise the

storms for each country, instead of fixed observations. So far, only 3-4 years of data are available and, while the study of scaling rates is not enough, having data at very high resolution (5-min) allows for an very detailed analysis on storm characteristics in the country. For example in Ghana and Kenya, there are more than 55 stations with data available, which offers the possibility to combine all datasets and obtain storm characteristics such as those analysed in this project: Storm duration, maximum precipitation rainfall at different time aggregations,mean intensity of rain and total rainfall depth. This variables together would give information on the time of the year different types of storms occur, and help better characterise the storms in the country, which could be translated into applications from agriculture and water resources management

Concerning the study of storm events for different locations, the straightforward outcome of this project is on the application and integration of the variables analysed in the field of flood hydrology. While in this project a conceptual application is presented, based on merging the results with existing stochastic weather generators, future research should look at the implications that super-CC behaviours in certain scaling rates could have when characterising extreme storms in a small location.

In this project, no differentiation has been directly made between summer and winter periods. For the case of Germany, summers are characterised by convective storm events whereas for winter large stratiform storms are more frequeny. Indirectly, this differentiation has been studied in the storm event block when conditioning storms by duration. However, future work should look at how storms properties differ between seasons, as well as the overall scaling rate. For areas with arid or dry climate, this could be specially relevant as the lack of moisture availability can vary between seasons and locations (inland versus coastal).

In existing literature, projects normally compare results from observational studies with models. For the scope of this project, validation of results was done through statistical methods. It would be interesting, mostly for sub-Saharan Africa stations, to compare the results with high-resolution models RCM (10-50 Km grid), which better capture the convection storms.

To sum up, this project has addressed the limitations that existing research on scaling rates have, mostly those aspects referred to the methodology. With the use of high resolution dataset, it has been shown that the sources of error can not only come from the method itself, but based on the length and quality of observations. Therefore, for future research, more focus should be put on developing methodologies that can allow to obtain, with only some years of data, similar results than with long-term observations.

Appendix

In the following Appendix, a more detailed collection of figures and calculations done during the project are presented, as well as complementary work. The section contains information from Part 2 and Part 3 of the research.

A. Storm event analysis

A.1 Sensitivity analysis of storm duration classes

In the research, 3 storm duration classes have been established: short (up to 2 hour events), medium (between 2 and 10 hours) and long (larger than 10 hours). While this decision is based on existing research and after checking the histogram of precipitation it is not a fixed decision, since different types of events do not have always the same duration. In this section a small sensitivity analysis is performed with slight variations on the duration classes, in order to check if those changes really influence the results obtained. This is meant to give consistency to the outcomes presented in the storm event analysis. In particular, 2 new duration classes are considered:

- Scenario 1: Short duration class reduced up to 1 hour events. The medium duration class takes from 1 hour to 10 and the long duration stays like in the baseline scenario.
- Scenario 2: Medium duration class between 1 and 8 hours. The long duration increases with 2 more hours compared to the baseline while the short duration reduces 1 hour.

In Figure 8.3 the results for the 2 scenarios considered are applied to the variable of total duration length. The results show that changes in the storm duration do not significantly change the overall scaling slope. The only differences are the shift of the slope upwards or downwards which is associated to the fact that a new distribution of duration classes changes the proportion of events at each group. However, this change in storm distributions do not have a large impact on the scaling rate slope. The same analysis is done for the peak rainfall intensity, which shows no significant difference either.

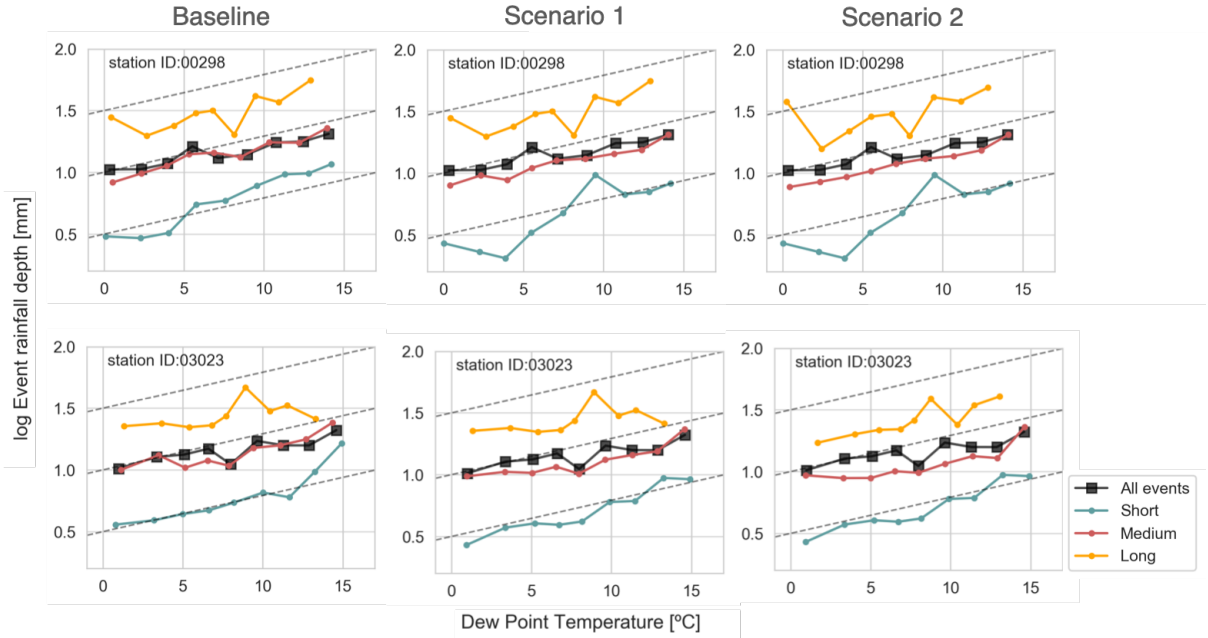


Figure 8.3: Examples of the changes in the relationship between total rainfall depth with dew point temperature conditioned to 3 duration classes for the 3 scenarios considered.

B. Scaling rates in sub-Saharan Africa

B.1 TAHMO Dataset results

Following the analysis of the scaling rates in the 4 stations with the maximum number of observations per country, in this section the results from the total number of stations that passed the quality test is presented. In particular, results are shown for the Quantile regression, given that in the outcomes of Block 1 indicated this method to outperform the binning methods. Figure 8.4 shows some example results of the quantile regression applied to the study areas.

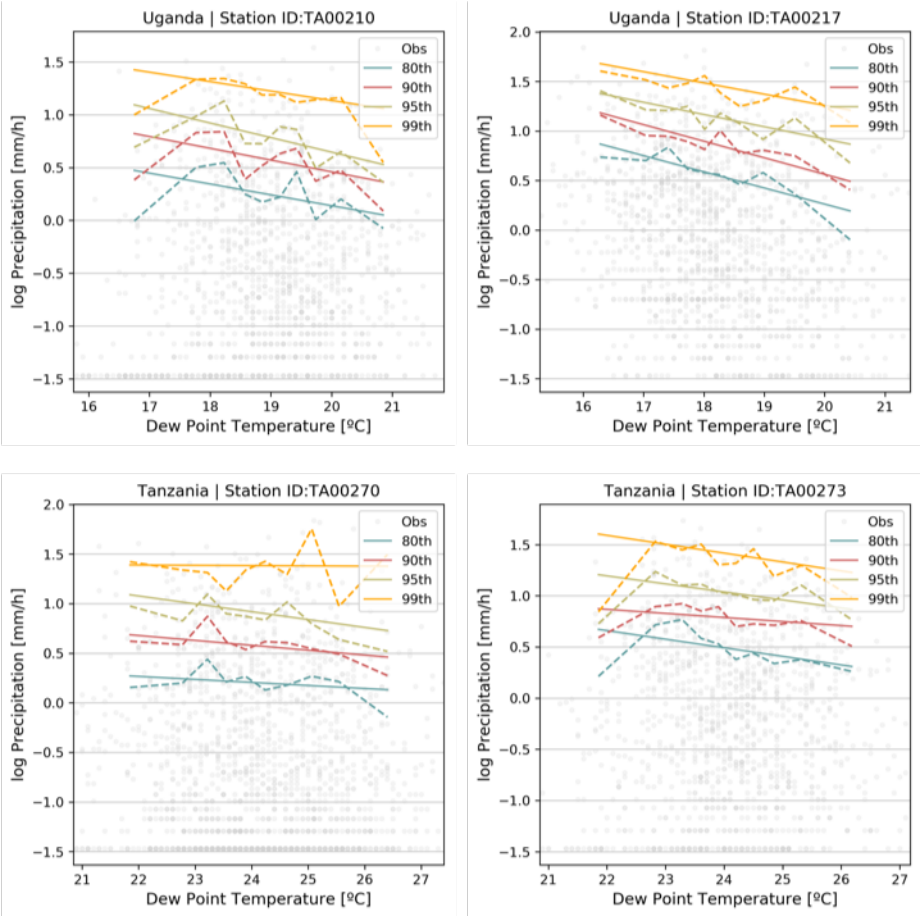


Figure 8.4: Examples of scaling rates for stations in Ghana, Kenya, Uganda and Tanzania. Solid lines represent the quantile regression fit. The dashed lines correspond to the binning method with equal length of observations

B.2 Overall scaling rates and range of dew point temperatures

One outcome from exploring the scaling rates of all the stations in the TAHMO dataset is that the range of dew point temperatures where precipitation occurs can be connected to the sign of the slope. Table 8.1 shows the range of dew point temperatures obtained from most stations and the associated sign of the slope.

Table 8.1: Range of dew point temperature for which precipitation occur and the associated slope sign.

	Range of dew point temperature	
Slope	**Negative slope**	**Positive slope**
Ghana	22-27 | 20-24	-
Kenya	17-21 | 17-21 | 19-23 | 17-21	11-16 | 14-18 | 10-14 | 12-17 | 12-18
Tanzania	22-26	-
Uganda	16-20 | 17-21	15-18

B.3 Relationship between relative humidity and surface air temperature for a set of stations

The following figure plots the relationship between relative humidity of hourly observations with the hour surface air temperature. Results vary across stations. The figure represents a sample of the main patterns observed.

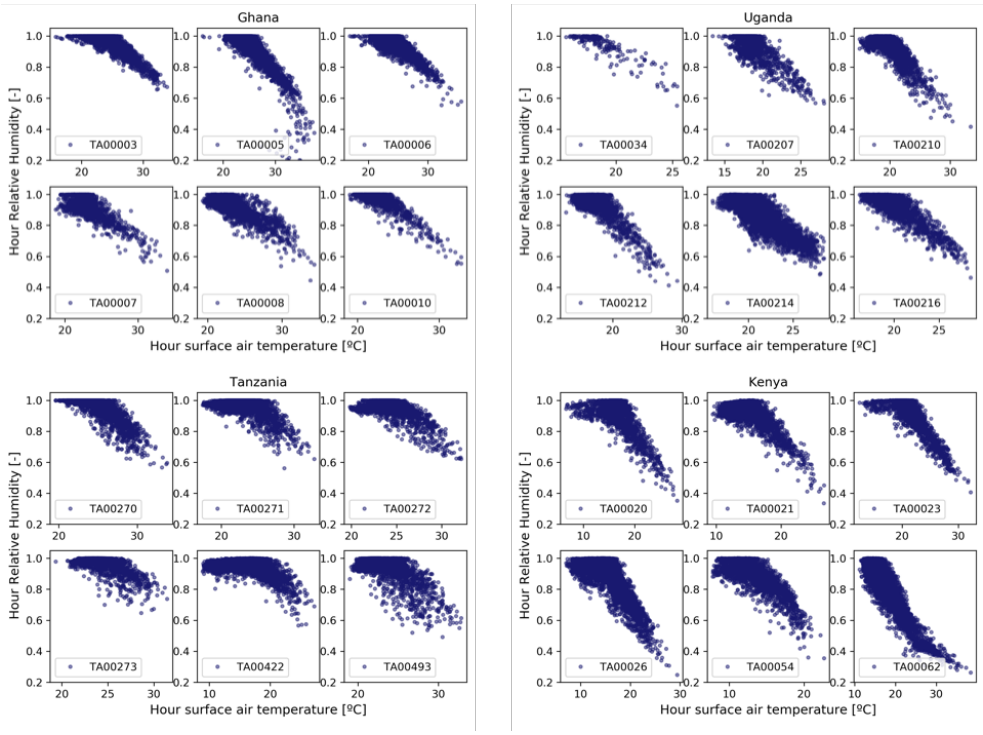


Figure 8.5: Relative humidity on wet hours with surface air temperature for a selection of station from the 4 countries.

B.4 Spatial distribution of scaling rates form the TAHMO stations

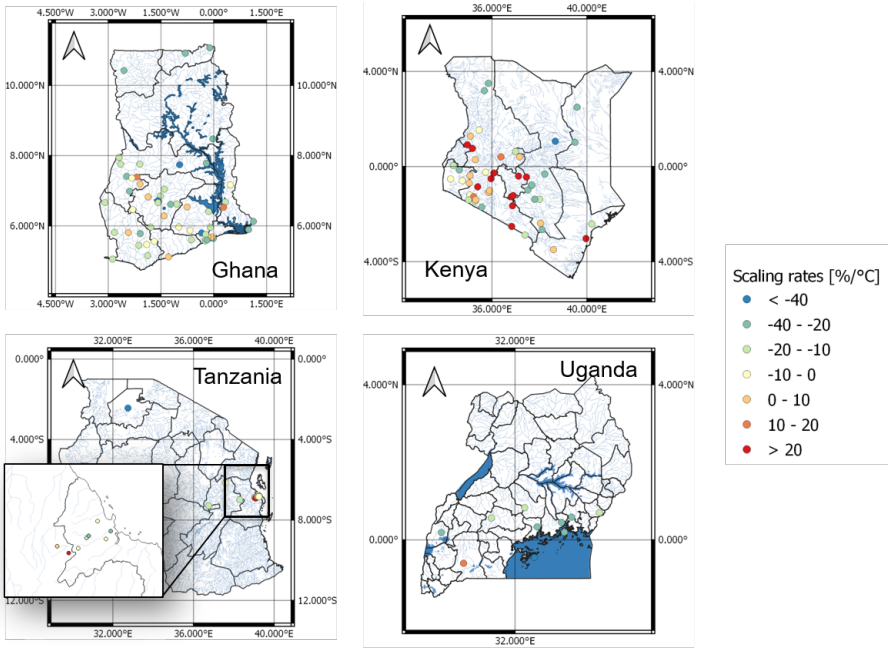


Figure 8.6: Spatial distribution of scaling rates calculated with the quantile regression for the 95th quantile.

B.5 GSOD Dataset results

The following figure shows 4 examples of the scaling rate results obtained from applying the Quantile regression on stations from the GSOD dataset, which are based on daily observations.

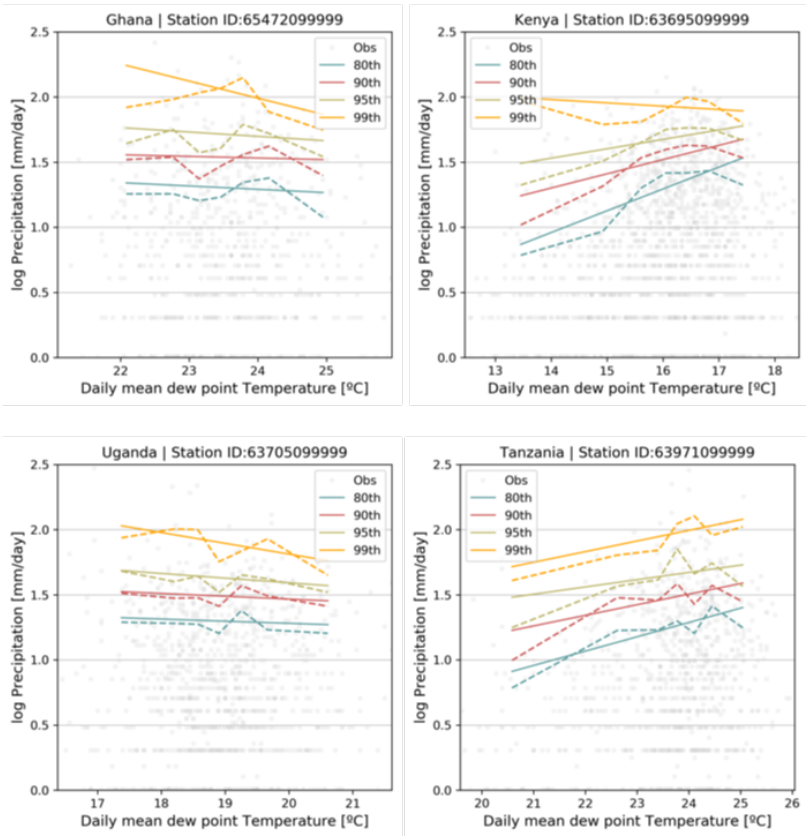


Figure 8.7: Selection of 4 example stations where the quantile regression is applied. In dashed lines the percentiles for each bin with equal number of observations per bin is shown.

For this project, the GSOD stations stations selected have less than 15% of missing data per year and a total length of at least 10 years. The total number of stations that satisfied this condition for sub-Saharan countries were Ghana: 1 station, Kenya 20 stations, Tanzania 14 and Uganda 1 station. In the following figure, the range of results for the countries with more than one station are shown. For Ghana and Uganda, the results of applying the quantile regression method to the only station that passed the quality test gave, for the 95th percentile a scaling rate of -7.44%/ºC and -7.912%/ºC.

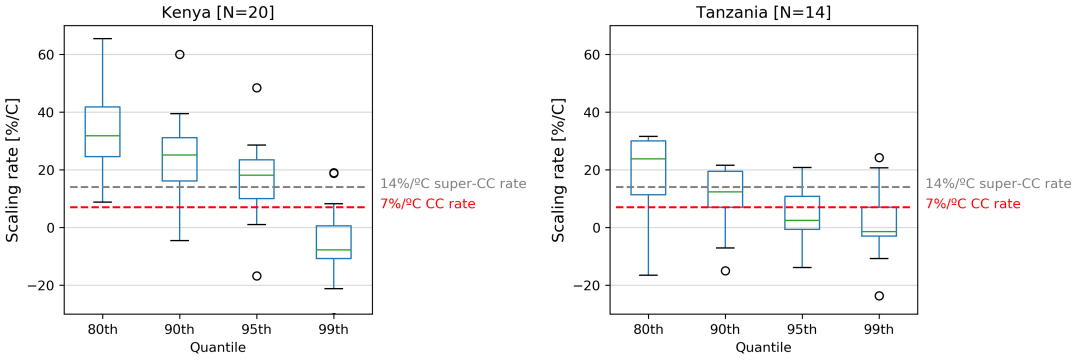


Figure 8.8: Boxplot of the 95th percentile of scaling rates obtained for Kenya and Tanzania through the quantile regression method.

References

Ali, H., Fowler, H. J., & Mishra, V. (2018). Global Observational Evidence of Strong Linkage Between Dew Point Temperature and Precipitation Extremes. *Geophysical Research Letters*, *45*(22), 12,320–12,330. doi: 10.1029/2018GL080557

Ali, H., & Mishra, V. (2017). Contrasting response of rainfall extremes to increase in surface air and dewpoint temperatures at urban locations in India. *Scientific Reports*, *7*(1), 1–15. doi: 10.1038/s41598-017-01306-1

Allen, M., & Ingram, W. J. (2002). Constraints on future changes in climate and the hydrology cycle. *Nature*, *419*, 224–232.

Asare-Nuamah, P., & Botchway, E. (2019). Understanding climate variability and change: analysis of temperature and rainfall across agroecological zones in Ghana. *Heliyon*, *5*(10). doi: 10.1016/j.heliyon.2019.e02654

Ban, N., Schmidli, J., & Schär, C. (2015). Heavy precipitation in a changing climate: Does short-term summer precipitation increase faster? *Geophysical Research Letters*, *42*(4), 1165–1172. doi: 10.1002/2014GL062588

Beck, H. E., Zimmermann, N. E., McVicar, T. R., Vergopolan, N., Berg, A., & Wood, E. F. (2018). Present and future köppen-geiger climate classification maps at 1-km resolution. *Scientific Data*, *5*, 1–12. doi: 10.1038/sdata.2018.214

Berg, P., Haerter, J., Thejll, P., Piani, C., Hagemann, S., & Christensen, J. (2009a). Seasonal characteristics of the relationship between daily precipitation intensity and surface temperature. *Journal of Geophysical Research: Atmospheres*, *114*(D18).

Berg, P., & Haerter, J. O. (2013). Unexpected increase in precipitation intensity with temperature - A result of mixing of precipitation types? *Atmospheric Research*, *119*, 56–61. doi: 10.1016/j.atmosres.2011.05.012

Berg, P., Haerter, J. O., Thejll, P., Piani, C., Hagemann, S., & Christensen, J. H. (2009b, sep). Seasonal characteristics of the relationship between daily precipitation intensity and surface temperature. *Journal of Geophysical Research*, *114*(D18), D18102. Retrieved from `http://doi.wiley.com/10.1029/2009JD012008` doi: 10.1029/2009JD012008

Berg, P., Moseley, C., & Haerter, J. O. (2013). Strong increase in convective precipitation in response to higher temperatures. *Nature Geoscience*, *6*(3), 181–185. doi: 10.1038/ngeo1731

Bigg, G. R., Jickells, T. D., Liss, P. S., & Osborn, T. J. (2003, aug). The role of the oceans in climate. *International Journal of Climatology*, *23*(10), 1127–1159. doi: 10.1002/joc.926

Boessenkool, B., Bürger, G., & Heistermann, M. (2017). Effects of sample size on estimation of rainfall extremes at high temperatures. *Natural Hazards and Earth System Sciences*, *17*(9), 1623–1629. doi: 10.5194/nhess-17-1623-2017

Bui, A., Johnson, F., & Wasko, C. (2019). The relationship of atmospheric air temperature and dew point temperature to extreme rainfall. *Environmental Research Letters*, *14*(7), 074025. doi: 10.1088/1748-9326/ab2a26

Dahm, R., Bhardwaj, A., Weiland, F. S., Corzo, G., & Bouwer, L. M. (2019). A temperature-scaling approach for projecting changes in short duration rainfall extremes from GCM data. *Water*, *11*(2), 1–15. doi: 10.3390/w11020313

Danielson, J. J., & Gesch, D. B. (2011). *Global multi-resolution terrain elevation data 2010 (gmted2010)*. US Department of the Interior, US Geological Survey.

Drobinski, P., Alonzo, B., Bastin, S., Da Silva, N., & Muller, C. (2016). Scaling of precipitation extremes with temperature in the French Mediterranean region: What explains the hook shape? *Journal of Geophysical Research*, *121*(7), 3100–3119. doi: 10.1002/2015JD023497

DWD. (2018). *Climate Data Center (CDC): Historical hourly station observations of precipitation for Germany, version v006* (Tech. Rep.).

Fowler, H. J., Blenkinsop, S., & Tebaldi, C. (2007). Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling. *International Journal of Climatology*, *27*(12), 1547–1578. doi: 10.1002/joc.1556

Haerter, J. O., Berg, P., & Hagemann, S. (2010). Heavy rain intensity distributions on varying time scales and at different temperatures. , *115*(October 2009), 1–7. doi: 10.1029/2009JD013384

Hardwick Jones, R., Westra, S., & Sharma, A. (2010). Observed relationships between extreme sub-daily precipitation, surface temperature, and relative humidity. *Geophysical Research Letters*, *37*(22), 1–5. doi: 10.1029/2010GL045081

Hofstätter, M., Lexer, A., Homann, M., & Blöschl, G. (2018). Large-scale heavy precipitation over central europe and the role of atmospheric cyclone track types. *International Journal of Climatology*, *38*, e497–e517.

Kisembe, J., Favre, A., Dosio, A., Lennard, C., Sabiiti, G., & Nimusiima, A. (2019). Evaluation of rainfall simulations over Uganda in CORDEX regional climate models. *Theoretical and Applied Climatology*, *137*(1-2), 1117–1134. doi: 10.1007/s00704-018-2643-x

Koenker. (2005). *Quantile regression (econometric society monographs; no. 38)*. Cambridge university press.

Koenker, R., & Bassett Jr, G. (1978). Regression quantiles. *Econometrica: journal of the Econometric Society*, 33–50.

Koenker, R., Portnoy, S., Ng, P. T., Zeileis, A., Grosjean, P., & Ripley, B. D. (2012). *Package 'quantreg'.* Technical Report Last accessed April 21th.

Lawrence, M. G. (2005). The relationship between relative humidity and the dewpoint temperature in moist air: A simple conversion and applications. *Bulletin of the American Meteorological Society*, *86*(2), 225–233. doi: 10.1175/BAMS-86-2-225

Lee, O., Sim, I., & Kim, S. (2019). Application of the non-stationary peak-over-threshold methods for deriving rainfall extremes from temperature projections. *Journal of Hydrology*(June), 124318. doi: 10.1016/j.jhydrol.2019.124318

Lenderink, G., & Attema, J. (2015). A simple scaling approach to produce climate scenarios of local precipitation extremes for the Netherlands. *Environmental Research Letters*, *10*(8). doi: 10.1088/1748-9326/10/8/085001

Lenderink, G., Mok, H. Y., Lee, T. C., & Van Oldenborgh, G. J. (2011). Scaling and trends of hourly precipitation extremes in two different climate zones - Hong Kong and the Netherlands. *Hydrology and Earth System Sciences*, *15*(9), 3033–3041. doi: 10.5194/hess-15-3033-2011

Lenderink, G., Van Den Hurk, B. J., Klein Tank, A. M., Van Oldenborgh, G. J., Van Meijgaard, E., De Vries, H., & Beersma, J. J. (2014). Preparing local climate change scenarios for the Netherlands using resampling of climate model output. *Environmental Research Letters*, *9*(11). doi: 10.1088/1748-9326/9/11/115008

Lenderink, G., & Van Meijgaard, E. (2008). Increase in hourly precipitation extremes beyond expectations from temperature changes. *Nature Geoscience*, *1*(8), 511–514. doi: 10.1038/ngeo262

Lenderink, G., & Van Meijgaard, E. (2010). Linking increases in hourly precipitation extremes to atmospheric temperature and moisture changes. *Environmental Research Letters*, *5*(2). doi: 10.1088/1748-9326/5/2/025208

Lenderink, G., & van Meijgaard, E. (2009, jun). Unexpected rise in extreme precipitation caused by a shift in rain type? *Nature Geoscience*, *2*(6), 373–373. Retrieved from `http://www.nature.com/articles/ngeo524` doi: 10.1038/ngeo524

Li, C., Zwiers, F., Zhang, X., & Li, G. (2019). How Much Information Is Required to Well Constrain Local Estimates of Future Precipitation Extremes? *Earth's Future*, *7*(1), 11–24. doi: 10.1029/2018EF001001

Lochbihler, K., Lenderink, G., & Siebesma, A. P. (2019). Response of Extreme Precipitating Cell Structures to Atmospheric Warming. *Journal of Geophysical Research: Atmospheres*, *124*(13), 6904–6918. doi: 10.1029/2018JD029954

Lu, S., Veldhuis, M.-c., Giesen, N. V. D., & Heemink, A. (2020). Precipitation Regime Classification Based on Cloud-Top Temperature Time Series for Spatially-Varied Parameterization of Precipitation Models. , 1–18. doi: 10.3390/rs12020289

Mishra, V., Wallace, J. M., & Lettenmaier, D. P. (2012). Relationship between hourly extreme precipitation and local air temperature in the United States. *Geophysical Research Letters*, *39*(16). doi: 10.1029/2012GL052790

Molnar, P., Fatichi, S., Gaál, L., Szolgay, J., & Burlando, P. (2015). Storm type effects on super Clausius-Clapeyron scaling of intense rainstorm properties with air temperature. *Hydrology and Earth System Sciences*, *19*(4), 1753–1766. doi: 10.5194/hess-19-1753-2015

Mudelsee, M. (2020). *Statistical analysis of climate extremes*. Cambridge University Press.

Muggeo, V. (2015). *Segmented: regression models with breakpoints/changepoints estimation. r package version 0.5-1.4.*

Mumo, L., Yu, J., & Ayugi, B. (2019). Evaluation of spatiotemporal variability of rainfall over Kenya from 1979 to 2017. *Journal of Atmospheric and Solar-Terrestrial Physics*, *194*(July), 105097. doi:

10.1016/j.jastp.2019.105097

Nimusiima, A., Kisembe, J., & Nakyembe, N. (2019). Evaluation of Past and Future Extreme Rainfall Characteristics over Eastern Uganda. *Journal of Environmental and Agricultural Sciences (JEAS*, *18*(July), 38–49.

O'Gorman, P. A., & Schneider, T. (2009). The physical basis for increases in precipitation extremes in simulations of 21st-century climate change. *Proceedings of the National Academy of Sciences*, *106*(35), 14773–14777. doi: 10.1073/pnas.0907610106

Oueslati, B., Yiou, P., & Jézéquel, A. (2019). Revisiting the dynamic and thermodynamic processes driving the record-breaking January 2014 precipitation in the southern UK. *Scientific Reports*, *9*(1), 1–7. doi: 10.1038/s41598-019-39306-y

Panthou, G., Mailhot, A., Laurence, E., & Talbot, G. (2014). Relationship between surface temperature and extreme rainfalls: A multi-time-scale and event-based analysis. *Journal of Hydrometeorology*, *15*(5), 1999–2011. doi: 10.1175/JHM-D-14-0020.1

Pumo, D., Carlino, G., Blenkinsop, S., Arnone, E., Fowler, H., & Noto, L. V. (2019). Sensitivity of extreme rainfall to temperature in semi-arid Mediterranean regions. *Atmospheric Research*, *225*(November 2018), 30–44. doi: 10.1016/j.atmosres.2019.03.036

R Core Team. (2017). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from `https://www.R-project.org/`

Roderick, T. P., Wasko, C., & Sharma, A. (2019). Atmospheric Moisture Measurements Explain Increases in Tropical Rainfall Extremes. *Geophysical Research Letters*, *46*(3), 1375–1382. doi: 10.1029/2018GL080833

Schär, C., Ban, N., Fischer, E. M., Rajczak, J., Schmidli, J., Frei, C., … Zwiers, F. W. (2016). Percentile indices for assessing changes in heavy precipitation events. *Climatic Change*, *137*(1-2), 201–216. doi: 10.1007/s10584-016-1669-2

Schleiss, M. (2018). How intermittency affects the rate at which rainfall extremes respond to changes in temperature. *Earth System Dynamics*, *9*(3), 955–968. doi: 10.5194/esd-9-955-2018

Schroeer, K., & Kirchengast, G. (2018). Sensitivity of extreme precipitation to temperature: the variability of scaling factors from a regional to local perspective. *Climate Dynamics*, *50*(11-12), 3981–3994. doi: 10.1007/s00382-017-3857-9

Sharma, A., Wasko, C., & Lettenmaier, D. P. (2018). If Precipitation Extremes Are Increasing, Why Aren't Floods? *Water Resources Research*, *54*(11), 8545–8551. doi: 10.1029/2018WR023749

Singer, M., Michaelides, K., & Hobley, D. (2018). Storm 1.0: a simple, flexible, and parsimonious stochastic rainfall generator for simulating climate and climate change. *Geoscientific Model Development*, *11*, 3713–3726.

Suleiman, R. (2018). *Local and regional variations in conditions for agriculture and food security in Tanzania: A review. AgriFoSe2030 reports 10* (Tech. Rep. No. October). doi: 10.13140

Tomal, J. H., & Ciborowski, J. J. (2017). Statistical methods for estimating ecological breakpoints and prediction intervals. *arXiv preprint arXiv:1709.07107*.

Trenberth, K. E., Dai, A., Rasmussen, R. M., & Parsons, D. B. (2003). The changing character of precipitation. *Bulletin of the American Meteorological Society*, *84*(9), 1205–1217+1161. doi: 10.1175/BAMS-84-9-1205

Turner, H. A., & Berentsen, W. H. (2020). *Climate in Germany.*

Utsumi, N., Seto, S., Kanae, S., Maeda, E. E., & Oki, T. (2011). Does higher surface temperature intensify extreme precipitation ? , *38*(August), 1–5. doi: 10.1029/2011GL048426

Van de Vyver, H., Van Schaeybroeck, B., De Troch, R., Hamdi, R., & Termonia, P. (2019). Modeling the scaling of short□duration precipitation extremes with temperature. *Earth and Space Science*(April), 2019–2020. doi: 10.1029/2019ea000665

van de Giesen, N., Hut, R., & Selker, J. (2014). The Trans-African Hydro-Meteorological Observatory (TAHMO). *Wiley Interdisciplinary Reviews: Water*, *1*(4), 341–348. doi: 10.1002/wat2.1034

Wasko, C., Lu, W. T., & Mehrotra, R. (2018). Relationship of extreme precipitation, dry-bulb temperature, and dew point temperature across Australia. *Environmental Research Letters*, *13*(7). doi: 10.1088/1748-9326/aad135

Wasko, C., & Sharma, A. (2014). Quantile regression for investigating scaling of extreme precipitation with temperature. *Water Resources Research*, *50*(4), 3608–3614. doi: 10.1002/2013WR015194

Wasko, C., Sharma, A., & Johnson, F. (2015). Does storm duration modulate the extreme precipitation-temperature scaling relationship? *Geophysical Research Letters*, *42*(20), 8783–8790. doi: 10

.1002/2015GL066274

Westra, S., Fowler, H. J., Evans, J. P., Alexander, L. V., Berg, P., Johnson, F., … Roberts, N. M. (2014). Future changes to the intensity and frequency of short-duration extreme rainfall. *Reviews of Geophysics*, *52*(3), 522–555. doi: 10.1002/2014RG000464

WFP. (2013). *Comprehensive food security & vulnerability analysis (cfsva), tanzania,(2012).* World Food Programme in collaboration with the World Bank Rome.

Zhang, W., Villarini, G., & Wehner, M. (2019). Contrasting the responses of extreme precipitation to changes in surface air and dew point temperatures. *Climatic Change*, 257–271. doi: 10.1007/s10584-019-02415-8

Zhang, X., Wan, H., Zwiers, F. W., Hegerl, G. C., & Min, S. K. (2013). Attributing intensification of precipitation extremes to human influence. *Geophysical Research Letters*, *40*(19), 5252–5257. doi: 10.1002/grl.51010

Zhang, X., Zwiers, F. W., Li, G., Wan, H., & Cannon, A. J. (2017). Complexity in estimating past and future extreme short-duration rainfall. *Nature Geoscience*, *10*(4), 255–259. doi: 10.1038/ngeo2911