

Estimation of
Convective Extreme Events over Germany
based on
Large Scale Parameters

Inaugural Dissertation

zur

Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät
der Universität zu Köln

vorgelegt von

Sven Ulbrich

aus Köln

Köln, 2016

Berichterstatter:

PD Dr. Joaquim G. Pinto
Prof. Dr. Yaping Shao

Tag der mündlichen Prüfung: 26. Oktober 2016

Abstract

Strong convective events can produce extreme precipitation, hail, lightning or gusts, potentially inducing severe socio-economic impacts. These events have a relatively small spatial extension and, in most cases, a short lifetime. In this study, a model is developed for estimating convective extreme events based on large scale conditions. It is shown that strong convective events can be characterized by a Weibull distribution of radar-based rainfall with a low shape and high scale parameter value. A radius of 90km around a station reporting a convective situation turned out to be suitable. A methodology is developed to estimate the Weibull parameters and thus the occurrence probability of convective events from large scale atmospheric instability and enhanced near-surface humidity, which are usually found on a larger scale than the convective event itself. Here, the probability for the occurrence of extreme convective events is estimated from the KO-index indicating the stability, and relative humidity at 1000hPa. Both variables are computed from ERA-Interim reanalysis. In a first version of the methodology, these two variables are applied to estimate the spatial rainfall distribution and to estimate the occurrence of a convective event. The developed method shows significant skill in estimating the occurrence of convective events as observed at synoptic stations, lightning measurements, and severe weather reports. In order to take frontal influences into account, a scheme for the detection of atmospheric fronts is implemented. While generally higher instability is found in the vicinity of fronts, the skill of this approach is largely unchanged. Additional improvements were achieved by a bias-correction and the use of ERA-Interim precipitation. The resulting estimation method is applied to the ERA-Interim period (1979-2014) to establish a ranking of estimated convective extreme events. Two strong estimated events that reveal a frontal influence are analysed in detail.

As a second application, the method is applied to GCM-based decadal predictions in the period 1979-2014, which were initialized every year. It is shown that decadal predictive skill for convective event frequencies over Germany is found for the first 3–4 years after the initialization.

Zusammenfassung

Konvektive Ereignisse können starke sozioökonomische Auswirkungen durch aussergewöhnlichen Niederschlag, Hagel, Blitze oder Böen verursachen und sind somit von besonderem Interesse. Diese Ereignisse haben typischerweise keine große horizontale Ausdehnung, haben meist eine kurze Lebensdauer und sind schwierig vorherzusagen. In dieser Arbeit wird ein Modell zur Abschätzung von konvektiven Ereignissen basierend auf den großskaligen Bedingungen entwickelt. Es wird gezeigt, dass sich konvektive Ereignisse typischerweise in den Weibull-Parametern des Radar-basierten Niederschlags durch einen niedrigen Shape-Parameter und einen hohen Scale-Parameter auszeichnen. Ein Radius von 90km um eine Station, die ein konvektives Ereignis meldet, erwies sich dabei als geeignet. Eine Methode wird entwickelt um die Weibullparameter und damit die Auftretswahrscheinlichkeit von konvektiven Ereignissen basierend auf einer instabilen atmosphärischen Schichtung und erhöhter bodennaher Luftfeuchte abzuschätzen. Diese beiden atmosphärischen Größen weisen typischerweise größere Muster auf als ein konvektives Ereignis. In der vorliegenden Studie wird die Wahrscheinlichkeit des Auftretens von konvektiven Extremereignissen basierend auf dem KO-index und der relativen Feuchte auf 1000hPa abgeschätzt. Beide Parameter werden basierend auf ERA-Interim Reanalysen berechnet, und genutzt um zunächst die räumliche Niederschlagsverteilung abzuschätzen. Basierend auf dieser wird das Auftreten eines konvektiven Ereignisses identifiziert. In der Vorhersage von konvektiven Ereignissen, die durch Beobachtungen an Wetterstationen, Blitzmessungen oder Extremwetterbeobachtungen definiert sein können, weist die Methode einen bedeutenden Skill auf. Um einen Einfluss durch Fronten zu berücksichtigen, wurde eine automatisierte Frontanalyse angepasst. Obwohl die Stabilität der Atmosphäre in der Nähe einer Kaltfront geringer

ist, bleibt der Skill der entsprechenden Methode nahezu unverändert. Ein zusätzlicher Performancegewinn konnte durch eine Biaskorrektur und Nutzung des ERA-Interim Niederschlags erzielt werden. Die entwickelte Methode wird auf den ERA-Interim Zeitraum (1979-2014) angewandt um ein Ranking der abgeschätzten konvektiven Extremereignisse zu erzeugen. Zwei starke Ereignisse werden detailliert analysiert. Beide weisen einen frontalen Einfluss auf. Als zweites wurde das Verfahren auf GCM-basierte dekadische Vorhersagen für den Zeitraum 1979-2014 angewandt, welche jährlich initialisiert werden. Für konvektive Ereignisse in Deutschland zeigt sich ein dekadischer Vorhersageskill für die ersten 3-4 Jahre nach der Initialisierung.

Contents

Abstract	i
Zusammenfassung	iii
Abbreviations	ix
1 Introduction	1
1.1 Motivation	1
1.2 Outline and overview	4
2 Scientific background	5
2.1 What is a convective event?	5
2.1.1 Requirements of an convective event	6
2.1.2 Triggering events	10
2.1.3 Modelling of convective events	11
2.2 Convective events in Germany	12
2.3 Rainfall properties of events	14
2.4 Fronts	16
2.4.1 Contribution of fronts to convective events	17
3 Data	19
3.1 Station data	19
3.2 Radiosounding data	22
3.3 Lightning data	23
3.4 ESWD database	24

3.5	Radolan data	25
3.6	Reanalysis dataset ERA-Interim	26
3.7	Time periods of different datasets	27
4	General approach	29
5	Methods	31
5.1	Rainfall distribution functions	31
5.1.1	Poisson distribution	32
5.1.2	Negative exponential distribution	32
5.1.3	Weibull distribution	33
5.2	Non-linear regression	34
5.3	Measures of skill	34
5.4	Analysis of fronts	36
6	Results	39
6.1	Analysis of rainfall in reanalysis	39
6.2	Spatial rainfall distribution associated with convective events	41
6.2.1	Representative radius for convective weather phenomenon	41
6.2.2	Lightning in the vicinity of stations reporting a convective weather situation	44
6.2.3	Identification of convective events from the spatial rainfall characteristics	46
6.3	Influence of stability parameters on station rainfall distribution	54
6.4	Improvement of estimation of rainfall parameters	63
6.4.1	Fronts	64
6.4.2	Bias-correction	72
6.4.3	Additional improvement by inclusion of model rainfall	79
6.4.4	Summary of improvements	81
6.5	Model combination and validation	82
6.6	Ranking of estimated events	86
6.7	Case studies	89
6.7.1	Case of 21 July 2014	89
6.7.2	Case of 21 June 2007	92
6.8	Application to decadal predictions	95
7	Summary and discussion	99

7.1	Summary	99
7.2	Discussion and conclusion	101
A	Stability/humidity indices	107
A.1	Vertical totals	107
A.2	Total totals	107
A.3	KO-index	108
A.4	S index	109
A.5	Potential instability index	109
A.6	Lifted index	109
A.7	Showalter index	110
A.8	Deep convection index	110
A.9	CAPE	111
A.10	CIN	111
A.11	K-index	112
A.12	SWEAT	113
A.13	SWISS ₁₂	113
A.14	DSI	113
A.15	Helicity	114
A.16	Storm-relative environmental helicity	114
A.17	Energy helicity index	114
A.18	SHUM _{1000→500}	115
B	Estimating transformed rainfall intensity distribution parameters from the KO/RH index	117
	References	119

Abbreviations

DWD	D eutscher W etter d ienst
ECMWF	E uropean C entre for M edium- R ange W eather F orecasts
ERA-40	E CMWF 45-year r eanalysis project
ESSL	E uropean S evere S torms L aboratory
ESWD	E uropean S evere W eather D atabase
GCM	G eneral c irculation m odel
MCS	M esoscale C onvective S ystem
NCAR	N ational C enter for A tmospheric R esearch
NCEP	N ational C enters for E nvironmental P rediction
PDF	P robability d ensity - function
Radolan	R adar- O nline- A neichung
RCM	R egional c limate m odel

1 Introduction

1.1 Motivation

Intense atmospheric convection can lead to extreme precipitation, hail, lightning, and gusts. These impacts can cause enormous losses and fatalities. For example, the German Insurance Association ('GdV') announced that hailstorms in Germany generated a total loss of more than 4 billion Euro in 2013 (source: *German Insurance Association (2014a)*). The 'Naturgefahrenreport 2014' of the GdV states that 28% of the losses in 2013 at insured buildings were caused in summer by extreme winds and hail. In fact, convective events are responsible for a large amount of damage during the summer in Germany (e.g. *Kunz (2007)*). On a longer time-scale, like e.g. for the period from 2002-2013, 12% of damages at insured residential buildings occurred in summer due to hail and extreme winds, which is in the same order of magnitude as losses caused by extreme winds and hail in winter (15%) (source: *German Insurance Association (2014b)*). Beside damages of residential buildings, losses produced by convective events are reported for example as damages of facilities or cars. Companies may experience losses related to convective events caused by business interruptions due to early warnings. Extreme precipitation events (e.g. flash floods) are another major hazard associated with intense atmospheric convection. Because of their rapid development and their rather regional scale characteristics, they are difficult to predict (*Gallus Jr. (1999)*; *Roberts & Lean (2008)*). Additionally, intense convective precipitation can contribute to large river floods (e.g. *Bronstert et al. (2002)*) and are of strong importance for floods in smaller catchments.

Events with a high loss are a risk for insurance companies for different reasons. For example, the underestimation of such risk is a problem for a company. Insurance companies commonly obtain an estimation of loss frequencies based on past events. Such database of past losses can only comprise reported events, but some extreme convective events do not hit any insured facilities and are therefore not included in such database. Insurance companies desire to estimate frequencies of losses due to extreme convective events, which requires a larger dataset. An estimation of such extreme events based on meteorological observations for a climatological timespan could contribute to a better estimation of loss frequencies.

In general, a rough estimation of the frequency of extreme convective events can be obtained from synoptic observations, which provide the longest measured meteorological time-series. Extreme convective events are often accompanied by hail, extreme gusts, lightning and extreme precipitation, which are continuously reported by weather observation stations. They provide information about different weather phenomena, precipitation type and in-situ measurements. The different impacts/hazards (intense short rainfall, lightning, and gusts) are often related, but sometimes show different intensities. For example, it is possible that a convective event produces intense rainfall, while e.g. no lightning occurred. Dry convection is usually not strong enough to cause major impacts. Problematic in the analysis of such events is the relatively small database of observed extreme convective events. This is not only due to their rarity, but also due to their usually relatively small spatial extension and their short lifetime. Therefore, many extreme convective events were not observed and hence not reported.

Besides the possibility of in-situ measurements, rainfall and lightning have the advantage that their position and intensity can be obtained or estimated from remote sensing data. Unfortunately, these datasets are just starting to have a sufficient length to provide statistically significant frequencies of extreme convective events (*Junghänel et al. (2015)*). Lightning strike positions are identified as 3-dimensional positions and collected in a database for Germany since 2007. Since June 2005, the meteorological radar coverage increased and provides a detailed information of estimated precipitation rates. In contrast to lightning, which appears only for some intense convective events, convective precipitation occurs also for weak convective events. It is therefore useful to identify characteristics of extreme convective and weak/normal convective events to develop an identification of extreme convective events. In addition to precipitation intensities, hail occurrence can be estimated based on radar data (e.g. *Waldvogel et al. (1979)*; *Bringi*

et al. (1984); Aydin et al. (1986)).

To summarize, databases for strong convective events which may lead to high losses are still short or contain only local information. For socio-economic purposes, an area wide database of strong convective events for a long historical period would be beneficial. Some first time-series become a sufficient length for climatological investigations (e.g. Junghänel et al. (2015)). Further, methodologies for future climate scenarios and decadal predictions are desired in order to enable an estimation of changes in frequencies on these time scopes.

The aim of this study is to develop a methodology to estimate extreme convective events from large-scale parameters.

For this development, different data sources are combined. These are radar data, radiosoundings, reanalysis data, and lightning data. The central dataset is the radar dataset, which is only available for Germany. Therefore, the investigation area is Germany. In this study, a regression model is established to estimate the rainfall characteristics based on large-scale conditions like atmospheric stability indices and humidity. The final method is then applied to large-scale parameters from reanalysis and global circulation models (GCMs).

In summary, the following research questions are addressed in this study:

- Is it possible to establish a model to estimate the probability and intensity of extreme convective events based on large-scale parameters?
- How are local spatial rainfall properties related to atmospheric humidity or stability parameters?
- Do other conditions influence the probability of convective events, which can be obtained from large scale data?
- Are these resulting estimated events characterized by strong surface gusts?
- Do decadal predictions show a predictability of convective extreme events for reforecasts (also called hindcasts)?

1.2 Outline and overview

The current state of research is presented in the chapter 2. Chapter 3 describes the datasets used. The data comprises measured/observed data as well as derived parameters and models.

Chapter 4 introduces the general approach in more detail. Different data sources (station observations, precipitation based on radar data and lightning measurements) are combined. It is investigated to what extent spatial rainfall properties in the surrounding of a station are related to the same-time-observed weather phenomenon at the station. One key parameter providing information of the potential of the occurrence of extreme convective events is vertical stability of the atmosphere, as it is quantified by a number of available stability indices. Therefore, the influence of stability and humidity parameters on the spatial rainfall distribution is addressed.

The approach requires methods, which are presented in chapter 5. The presented methods comprise rainfall intensity distributions, statistics such as regression techniques, as well as an objective frontal analysis and its adjustments for this study. The results are presented in chapter 6: It includes a discussion of the spatial rainfall intensity distributions for convective events based on different data sources, a discussion of the influence of objective fronts on the observed weather phenomena and the observed rainfall parameter. The parameters for estimating convective events are identified and then applied and validated in this chapter. The method is applied to a reanalysis dataset and an event ranking is produced. Two of the most intense ranked events are discussed. Probabilities and intensities of convective events estimated from large scale atmospheric parameters are derived from decadal predictions. In the end, results are summarized and discussion and a short outlook are given.

2 Scientific background

This chapter reviews scientific findings regarding convective events. Starting from the definition of an convective event and a description of the associated effects (section 2.1), the preconditions for the occurrence of intense events are presented (section 2.1.1), which includes in particular indices of atmospheric stability indicating conditions favourable for convection are introduced (section 2.1.1). The initialization of a convective event is discussed (section 2.1.2), and the estimation of convective rainfall in numerical climate models is analysed (section 2.1.3). Then, aspects of convective events regarding the investigation area are introduced (section 2.2). Research results regarding rainfall properties during convective events are presented (section 2.3). In particular, their spatio-temporal rainfall properties are investigated and properties of rainfall which were identified based on other accompaniments of convective events. Finally, front identifications are discussed (section 2.4) and, using these schemes, frontal impact on convective events (section 2.4.1).

2.1 What is a convective event?

A convective event is a micro- to meso-scale atmospheric phenomenon characterized by intense vertical movements. Convection is usually connected with condensation of water vapour. More precisely, an intense convective event is usually associated with localized rainfalls of high intensity on short duration, strong gusts, and sometimes hail or lightning. This study focusses on convective extreme events, which can be divided in

four different kinds:

- single thunderstorm cell
- multi-cell system
- supercell
- mesoscale convective system

All these systems require a (potentially) unstably layered atmosphere (i.e. absolute, conditional or potential instability), and high humidity close to the surface. Also, an initiation mechanism is required. Differences appear in some conditions, mainly the vertical wind shear and moisture transports. The initial lifting, for example caused by gravity waves, frontal lifting, pre-existing density flow or the exceedence of the convective temperature, does not determine the type of thunderstorm. Nevertheless, the types of initial lifting are not independent of preconditions of the thunderstorm types, which have been discussed in the literature (e.g. *Wallace & Hobbs (2006)*). The different types of thunderstorms are related to different spatio-temporal characteristics of rainfall intensity. Several properties of thunderstorm types imply a diverse impact, such as spatial extension, lifetime, and vertical movements. A single thunderstorm cell is spatially the smallest convective system presented here. It has a short life-time, typically 30-60 minutes, and a diameter of up to few kilometres. The multi-cell system consists basically of multiple single thunderstorm cells. It is a moving system in which old cells are replaced by new cells. This development can progress for several hours. The spatial extension of a multi-cell system is more than 30km. Even more intense is the so called supercell, which leads to enormous precipitation, often to intense gusts and hail. The lifetime of a supercell is typically 1-8 hours, and its spatial extension is up to 50 km.

2.1.1 Requirements of an convective event

Atmospheric stability is of major importance for vertical movements in the atmosphere and particularly for the development of an convective system. Convective systems develop from an initial disturbance, which grows to become the convective system. This growth can only appear in an environment supporting vertical movements. A useful concept to investigate the stability of the atmosphere was introduced by *Petterssen et al. (1946)*. They introduced a concept for the analysis of the atmospheric stability based on an air parcel concept. The air parcel has no interaction with its surrounding except for adopting to surrounding pressure. Buoyancy is caused by a temperature difference between an air parcel and its surrounding. For the development of the air parcel, the dry adiabatic lapse rate for unsaturated air and the wet adiabatic lapse rate in case of

saturation are relevant parameters. Assuming a static case, the energy which an surface air parcel gains by its ascend can be calculated and is described by the parameter CAPE. The energy, which needs to be exceeded to reach the level of free convection (i.e. the level in which the parcel is warmer than its surrounding) is called convective inhibition (CIN). These are integrated measures of the atmosphere, which best represent the stability in a static case. Additional instability maybe influenced by processes like temperature advection or vertical movements. The presented parameters are only subset of a list of different stability parameters, which are presented in the following and more detailed in appendix A.

It is worth to mention that there are efforts to predict lightning based on a statistical-dynamical approach (e.g. *Sousa et al. (2013)*) and stability parameter based approach (e.g. *Bright & Wandishin (2005)*, *Santos et al. (2013)*). Particularly the later are of interest due to thunderstorms indicating a subset of intense convective events. Hence, the associated parameters found in these studies mark good conditions for the growth of convective events (e.g. CAPE was identified in both studies to be a relevant parameter). Different parameters have been developed in atmospheric science to indicate the probability of occurrence of different types of convective events and notably thunderstorms (e.g. *Showalter (1953)*, *Haklander & Van Delden (2003)*). Previous studies (e.g. *Kunz (2007)*; *Kunz et al. (2009)*) have revealed that in particular the parameters Convective Available Potential Energy (CAPE), Lifted Index (LI) and Deep Convection Index (DCI) are good proxies for the occurrence of severe thunderstorms.

Stability Indices

The atmospheric stability is complex as it depends on different variables (temperature, humidity, wind). In addition, the vertical extension of unstable layering of the atmospheric column can be relevant for the development of convective systems. Therefore, many stability indices have been developed. Each of them draws a simplified picture of instability or is focussing on specific situations like deep convection, which is convection in which air from the lower atmosphere is lifted to above 500hPa (*Giaiotti et al. (2006)*). In addition to static cases, additional parameters such as wind profiles contribute to some stability indices. These additional parameters are specific to the formation of particular types of thunderstorms or are designed to estimate particular effects of thunderstorms (e.g. gusts, lightning,...). For example, *van Delden (2001)* stated that the potential instability index is optimized to represent the thunderstorm frequency. Most stability parameters can be assigned to one of three levels. The first level is based on

stability index	Reference
vertical totals (VT)	<i>Miller (1972)</i>
total totals (TT)	<i>Miller (1972)</i>
convective index (KO)	<i>Andersson et al. (1989)</i>
S-index	<i>Reymann et al. (1998)</i>
potential instability index (PII)	<i>Van Delden (2001)</i>
lifted index (LI)	<i>Galway (1956)</i>
Showalter index (SI)	<i>Showalter (1953);Steinacker (1977)</i>
Deep convection index (DCI)	<i>Barlow (1993)</i>
convective available potential energy (CAPE)	<i>Moncrieff & Miller (1976)</i>
convective inhibition (CIN)	<i>Colby J. (1984)</i>
K-index	<i>George (2014)</i>
Sever Weather Threat Index (SWEAT)	<i>Bidner (1970)</i>
SWISS ₁₂	<i>Huntrieser et al. (1997)</i>
dynamical state index (DSI)	<i>Névir (2004)</i>
storm-relative environmental helicity (SREH)	<i>Davies-Jones et al. (1990)</i>
energy helicity index (EHI)	<i>Hart & Korotky (1991)</i>

Table 2.1: List of stability indices applied in this study and the corresponding references.

the temperature profile only, the second one includes additionally humidity and the third one also includes windshear. Hence, multiple stability indices have been developed in the past and are being developed. The most important parameters are listed in table 2.1 and introduced in more detail in the appendix A. Due to data requirements, not all presented parameters were investigated in this study.

Certain types of unstable atmosphere can be identified best in different levels and by different parameters. In addition, it is possible to distinguish some stability indices based on the underlying concept of stability. *Kunz (2007)* states that the potential instability is underlying concept for the potential instability index and the KO-index, whereas the latent instability is the basic idea behind the indices lifted index, deep convection index, showalter index, and CAPE. *Kunz (2007)* investigated the skill of different convective indices to predict isolated and severe thunderstorms. A huge number of convective indices was compiled for this region and compared to loss-, radar-, hail- and SYNOP-data. Applying different statistical indices, it was found that the most appropriate indices for the prediction of thunderstorms are lifted index (averaged values of lowest 100hPa lifted to 500hPa), deep convection index (surface air lifted to 850hPa), potential instability index and CAPE. *Kunz (2007)* compared the performance of the indices with underlying concepts of latent instability (LI, DCI, CAPE) with indices of potential instability (KO, PII). Focussing on severe thunderstorms, those indices of latent and potential instability perform best which include information on temperature and humidity of the lowest atmospheric levels. *Kunz (2007)* also found that indices considering additional dynamic

information like the SWISS index show a slightly better performance in Germany than the KO index (designed by the German weather service). For the decision between thundery and non-thundery days, there is no large advantage of any of the 3 groups of instability concepts. Additionally, *Kunz (2007)* investigated the skill with regard to different datasets related to convective events in Baden Württemberg (radar reflectivity, weather phenomenon reports and insurance data). For these data sets, the different indices revealed different skills depending on the choice of the thresholds used for the identification of a convective situation. Skills also depend on the thunderstorm intensity and dataset. Additionally, it should be taken into account that part of the indices considered by *Kunz (2007)* have been developed for particular regions and cannot always be applied to other regions without a loss of prediction skill, which can be understood from regional differences in the relevant dynamics of convective event generation.

For the same region, the South-West of Germany, *Kunz et al. (2009)* investigated recent trends of thunderstorms and hailstorm frequencies. The frequency of thunderstorm days does not have a significant trend in the mean, while number of hail days and hail damage to buildings has significantly increased in the period from 1986 to 2004). They confirm that a large part of the hail days is associated with a south-westerly flow, and that the increase in the frequency of this flow direction in the past decades explains the positive trend in hail days. No significant change was observed for the pattern most often associated with hail days. Another part of the work carried out by *Kunz et al. (2009)* concerns different convective indices based on temperature and moisture. Indices incorporating information from the lowest atmospheric levels show a significant positive trend in values above a certain threshold and in extreme values. In contrast, indices using only information of higher levels, do not show a trend. *Kunz et al. (2009)* suggest that this contrasting behaviour is explained by a strong positive trend in lower layers, while properties of the upper layering change only marginally.

Changes of stability indices on smaller temporal scales have also been investigated using regional climate models. For example, *Keil et al. (2014)* found the temporal change of CAPE to be a good indicator to estimate the intensity of a convective event. *Weisman & Klemp (1982)* suggested that the ratio of buoyant energy and vertical wind shear is a key factor for basic features of the convective storm structures. Using model studies, they found that short lived cells are generated when vertical wind shear is weak. At weak to moderate wind shear, secondary systems like in multi-cells appear, while at moderate to strong wind shear, split storms occur, which were the model equivalent

of observed super-cells. Another study based on numerical model simulations was performed by *Droegemeier et al. (1993)*, who investigated the influence of helicity. The storm relative energy helicity index (SREH) was found to exceed $250\text{m}^2/\text{s}^2$ for convective extreme events, which is higher than the threshold obtained based on observations $157\text{m}^2/\text{s}^2$ (*Davies-Jones (1986)*).

As indicated by the scientific work referenced in this chapter, parameters show different probability distribution functions (PDFs) for different types of convective systems. Hence, a regression based on multiple parameters can improve the predictability of occurrence of convective events. Such a regression has been applied in the development of stability indices incorporating dynamical and statical properties of the atmosphere, such as SWEAT (*CAeM Working Group on Advanced Techniques Applied to Aeronautical Meteorology (1992)*) or SWISS (*Huntrieser et al. (1997)*). Also, the occurrence probability and the intensity have been investigated separately in the past, e.g., by *Mohr et al. (2015a)*, who used a multiple logistic regression in determining the occurrence probability of hail. According to their study, the lifted index combined with the surface-near temperature (minimum and at 12UTC) and the circulation weather type perform best for this application.

2.1.2 Triggering events

An unstable atmosphere alone is not sufficient on its own for the occurrence of intense convection. An initial trigger is needed, which can be of different kinds. There are 3 main triggers: cold air moves below warm air, which rises above the lifting condensation level; horizontal moved air is forced to rise by an obstacle; and temperature is increased locally strongly enough to exceed the convective temperature (i.e. the surface temperature is increased over a small area so that the lifting condensation level is equal to the cumulus condensation level). Such trigger always induces a vertical displacement of air parcels with respect to the surrounding atmosphere. Note that also slantwise convection is generally possible.

The first important process initiating convective events is that cold air moves below the warm surrounding air and forces it to rise. In particular, this process appears at cold fronts. In general, it has the potential to trigger thunderstorms. In addition, cold air advection increasing with height can lead to an unstable atmospheric layering. On small scales, a cold, dense outflow of a pre-existing rainfall area, can induce this lifting. This

scale is subsynoptic and may therefore reduce the relationship of convective events and large-scale parameters.

Another effect is horizontally moving air rising to flow across orography, if it cannot go around the obstacle (this can be quantified using the Froude number). The orographic enhancement of precipitation over low mountain ranges has been investigated by, e.g., *Kunz & Kottmeier (2006a)*. They applied a diagnostic model, which is based on the linear theory of hydrostatic flow and calculates condensation rates from vertical lifting. In an idealized study, in which they defined the mountain to be bell shaped, they found the orographic (total) precipitation upstream of the mountain to be very sensitive to wind speed, static stability and temperature in 1000hPa. Applying their model to different situations in the black forest and comparing the results to precipitation measurements (in *Kunz & Kottmeier (2006b)*), they found the orographic precipitation maximum over the crests and some spillover in the lee of the mountains. Over their investigation area, the orographic precipitation could drift up to 50 km downstream. Rainfall in these two subregions (over the crests and downstream) are anti-correlated. The precipitation downstream of the mountains can be reduced due to the effect of evaporation. The importance of this mechanism increases with the slope of the terrain. As the local orography does not change with time on the regarded time scales, the local rainfall climatology includes this effect. Therefore, this process is implicitly included in the climatology of rainfall characteristics.

Another process initiating lifting of air parcels within one air mass is a heterogeneous surface, which can play a major role in triggering convective events due to heterogeneous surface heating. Such inhomogeneity may be due to surface material, shadow/increased heating (associated with orographic slopes). As a result, a convective cell is forming in locally unstable conditions due to this local heating and attracts warm air near the surface from the surrounding area. This process is important for the location of thunderstorm initiation within an air mass. However, it appears on a small scale and cannot be obtained from a large scale analysis. Thus, this effect largely produces the difficulties in the prediction of thunderstorm locations.

2.1.3 Modelling of convective events

The analysis of summer storms (predominantly thunderstorms) is a difficult task as most of the relevant effects happen on the sub-synoptic scale and are typically non-hydrostatic. That is why, high resolution models being capable to resolve non-hydrostatic processes

are required to get further insight into their dynamics. In models of low resolution (e.g. GCMs), convective events are not resolved, but their preconditions can be found. These preconditions are atmospheric instability and humid air close to the surface. Large-scale lifting of air masses, which can also trigger convection, may also be represented in GCMs.

High resolution simulations are one approach to understand dynamics involved in the development of convection better. Different studies have simulated extreme precipitation events with regional climate models (e.g. *Semmler & Jacob (2004)*; *Lenderink (2010)*; *Trapp et al. (2011)*). Such high resolution models are designed to explicitly compute convection without making use of the respective parametrization schemes. For example, *Chan et al. (2013)* found the spatial resolution of a regional climate model first to improve the simulated daily precipitation and an increasing bias above a certain threshold, which was 12km resolution in their study. This result indicates that the parametrization needs to be adjusted to the model's resolution. Currently, some convection resolving models are increasingly applied to overcome the deficiencies.

The boundary conditions for high resolution models can be provided by global reanalysis data, general circulation models (GCMs) or by relatively coarse RCMs. These do not resolve convection, but convective precipitation is needed in a parametrization scheme for the evaluation of the total precipitation in these datasets. Hence, precipitation is a forecasted also in reanalysis datasets. The modelled rainfall data and sometimes the modelled stability (as previously introduced in this chapter) are used for the investigation and the identification of convective events. Different reanalysis datasets have been compared with respect to the analysis of extreme precipitation over Europe by *Zolina et al. (2004)*. They found local disagreements in the patterns among different reanalysis datasets. Particularly during summer this led to opposite trends in summer for the same period (last 4 decades). According to their analysis, during winter, the correlation between the station data and the reanalysis data is 20-30% stronger.

2.2 Convective events in Germany

Thunderstorms in Germany often occur in summer within warm and moist air masses, but also at cold fronts, mostly in winter. Most thunderstorms appear in the South-West of Germany over the local mountain ranges, which is one of the regions in Europe most frequently affected by thunderstorms (see *van Delden (2001)*). The terrain of Germany can

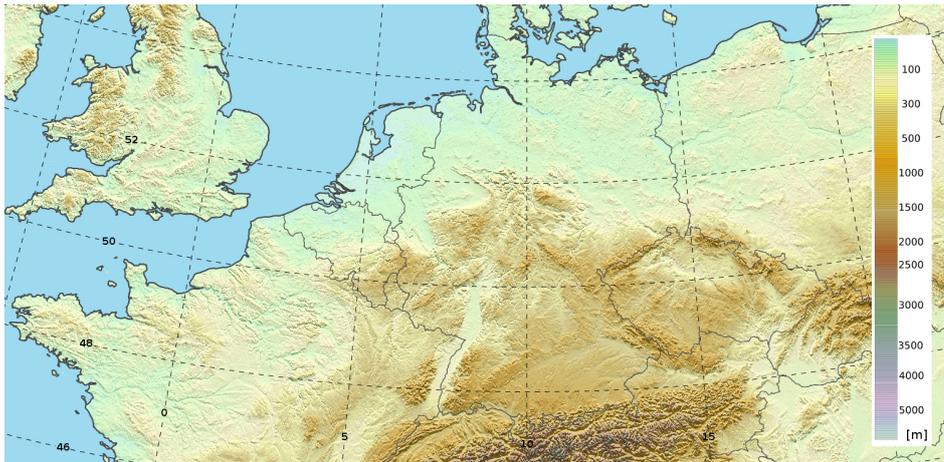


Figure 2.1: Orography of Germany and surrounding countries. The colour indicates the orography height (see legend). (map created with Ninjo)

be separated into 3 regions: lowlands in the North, uplands in the centre and Alps in the South. Often this is simplified by regarding Germany as two areas, the low orographically structured North and the strong orographically structured South). Some studies focus on one of either parts of Germany. In the South-West of Germany, thunderstorms are typically associated with winds from the South-West. This latter synoptic setting can be explained by advection of warm, moist air from the Mediterranean region to Germany and consequently the transport of latent heat. Accordingly, *Mohr et al. (2015a)* found the circulation weather types with winds from South-West to be dominant for hail occurrences in the South-West of Germany. This is similar for other possible accompaniments of convective events like gusts, lightning, and extreme rainfall. For the North of Germany, *Katarzyna (2013)* investigated for example the influence of atmospheric circulation on hail, which is associated with intense convective events. A low pressure system over northern Germany and Scandinavia was identified, implied by a largely negative anomaly of the sea level pressure and the 500hPa geopotential. The Azores high caused a north-western influx of cold, humid air. A negative temperature anomaly of 2°C was present for the hail cases, and increasing with height. In total, *Katarzyna (2013)* subdivided hail days into those occurring within air masses from northern to north-western and southern sources. The northern influxes caused negative anomalies of the temperature in the mid-troposphere. The author suggests that the southern circulation is connected to warm and dry air masses originating from the interior of the continent. Amongst other effects, convective events in this Germany or elsewhere can even produce tornadoes in rare cases (e.g. *Dotzek et al. (2003)*, *Rauhala & Schultz (2009)*). These appear very seldom in comparison to within the United States.

2.3 Rainfall properties of events

Extreme rainfall events including extreme convective rainfall events can be identified using statistical methods. Such has been performed in a huge number of studies (e.g. *Anagnostopoulou & Tolika (2012)*; *Früh et al. (2010)*; *Zolina et al. (2005)*). Based on different techniques, convective precipitation in such extreme events has been identified, investigated, and sometimes quantified. The identification techniques of convective precipitation can involve other phenomena of convective events such as lightning (e.g. *Gaál et al. (2014)*, *Rulfová & Kyselý (2013)*) or are based on the precipitation measurements itself (*Llasat (2001)*).

A phenomenon typical for all extreme convective events is intense local precipitation typically on short time periods. There have been efforts to identify convective precipitation into a convective and large scale precipitation component. For example the study of *Llasat (2001)* accumulated precipitation over different time scales. *Llasat (2001)* compared short term aggregated rainfall intensities (episodes of 5min aggregated precipitation exceeding a certain intensity threshold) and long term totals to identify convective events. The ratio between the time window is called convectivity index with a threshold of this index used to define a convective rainfall event. Similarly, *Casas et al. (2004)* used four different aggregation times to characterize the events. It was found that rainfall in the western Mediterranean area with a return period of more than 5 years can be assigned to four clusters: intense on short durations only, intense on mesoscale durations only, intense on 12-24hs durations and intense on all durations within the event. *Pinto et al. (2013)* performed a spatial aggregation of events defined from rainfall annual maxima at individual rain gauges. They used a cluster analysis based on different aggregation time periods. They found that the most intense events were characterized by extreme precipitation on all time scales up to 24 h, while the event clusters of moderate (daily) intensity could be distinguished with respect to the role of short- and long-lasting rainfall.

An alternative approach to the identification of convective events based on rainfall parameter properties only was presented by *Rulfová & Kyselý (2013)*. They disaggregated precipitation based on synoptic code for present weather available from reports at stations based on the WMO standard. They categorized the rainfall events into three major groups: large-scale rainfall, convective rainfall and other. On this basis, they were able to identify a positive trend in the amount of convective precipitation and in the number

of days with convective precipitation for the Czech Republic from spring to autumn over the period from 1982 to 2010. Their algorithm was able to disaggregate 95% of precipitation amounts, while performing better for moderate to heavy precipitation amounts. Additionally, they could obtain an improvement based on reports of cloud types, which is also part of the weather reports. They found convective precipitation to be dominant only in summer in the Czech Republic, while large-scale precipitation was found to be dominant at all stations from autumn to spring. *Rulfová & Kyselý (2013)* suggest that the increase of precipitation with altitude is mainly due to an increase of convective precipitation. Another result of their study is that the proportion of convective and large-scale precipitation depends on the mean daily surface air temperature in summer, while this dependence is much weaker during the other seasons. A similar approach by *Eggert et al. (2015)* disaggregated rainfall based on cloud observations only. Using a high resolution rainfall radar based dataset with a 1km resolution for Germany, they investigated the temporal and spatial scales of extreme precipitation. By aggregating rainfall on different temporal and spatial scales, they found that large-scale precipitation can be well characterized even when only data with coarse temporal resolution are available. For convective rainfall events, a characterization requires highly resolved data.

Spatially high resolved precipitation estimates from radar information can be used to identify convective extreme events. Such was performed by *Llasat et al. (2007)* using weather observations for the adjustment of the identification, in particular lightning. Another application of such high resolved data has been introduced by *Rigo & Llasat (2004)*, who developed a detection method of single/multiple cells within a convective system.

One phenomenon which can occur in the course of an intense convective event is lightning. The study of *Lang & Rutledge (2002)* found the lightning frequency to increase with the kinematic intensity of convection and the updraft volume, which was estimated based on sounding data. The electric impulse caused by a lightning strike can be measured, its position triangulated at a resolution of few hundred meters and its intensity obtained from the signal strength (e.g. *Inan et al. (1996)*, *Betz et al. (2007)*, *Betz et al. (2009)*). Investigating convective extreme rainfall events in Czech republic, *Gaál et al. (2014)* found a threshold of the 10-minute precipitation intensity, which is required to predict lightning occurrences. According to their results, the identified thresholds agree irrespective of the event frequency. They suggest that the event set of intense summer storms is characterized by high correlation coefficients between parameters describing

the main storm properties. Additionally, they found that a high percentage of the events can be considered as strongly convective with regards to the convectivity index of *Llasat (2001)*). Thus, it can be concluded that observations of rainfall intensity on short time scales can be used to estimate convective extreme events associated with lightning. A similar approach to estimate convective extreme event occurrences is conducted in this study, but is based on the spatial distribution for one time point rather than the temporal evolution at a single observation site.

2.4 Fronts

As outlined in chapter 2.1.2, fronts can induce vertical movement of air, and thus constitute a physical process able to induce convective events, which in turn can enhance the weather events associated with the front. Convergence lines, which appear relatively frequently ahead of a cold front, are also associated with vertical movement leading to convective events. Over Germany, frontal systems appear more often during winter (e.g. *Kunz (2007)*). If convective cells develop under these conditions, they are usually aligned along the fronts or convergence lines. Previous studies have compared rainfall statistics of precipitation particularly considering the influence of fronts (e.g. *Papritz et al. (2014)*; *Catto et al. (2012)*), revealing that a large number of intense precipitation events are associated with atmospheric fronts.

Fronts can be identified at the surface from changes in temperature, humidity, wind direction or wind speed. *Margules (1906)* assumed the front to be a surface without volume. In reality, the variables affected by the front do not change suddenly. The front has a horizontal extension in which variables such as temperature or dew point, wind-speed or direction change. In addition, the air density does not only change due to different temperatures of the two air masses but also due to their humidity. Thus, it is possible that some changes usually associated with fronts do not appear under certain conditions. An example is the so called masked cold front.

Most weather services analyse fronts subjectively. In this process, there are weather service specific rules. The German weather service is more frequently analysing frontal features with small horizontal extension as the British Met Office which prefers long interconnected fronts. The Met Office relatively often identifies sequences of closely neighbouring fronts. The choice of the German weather service to emphasize the 850hPa level (*DWD (2015)*) rather than considering the surface parameters follows the sugges-

tion by *Kurz (1990)* originally meant as a recommendation to use the wind direction in 850hPa as an indicator for the position of the surface front in cases as the masked cold front. This level is located above the planetary boundary layer (PBL) and typically reveals strong horizontal gradients in equivalent potential temperature and wet potential temperature at the fronts. Multiple procedures for an objective identification of fronts have been developed using various different parameters. *Hewson (1998)* used different predominantly temperature- and humidity-based parameters to obtain fronts, *Simmonds et al. (2011)* uses the change in wind direction. In addition, the mathematical procedure applied for front identification has huge impact on the shape and length of the identified structures (see *Hewson (1998)*). *Berry et al. (2011)* used a nearest neighbour algorithm and applied a front length filter. The resulting highest front frequencies are found over the ocean in the mid-latitudes storm tracks and are co-allocated with maxima of cyclonic vorticity centres. The same authors found cold fronts in the mid-latitudes to appear more often further north than warm fronts. With respect to precipitation, they found that frontal systems/front segments without precipitation appear with different frequencies in different regions of the world. While over oceanic regions less than 40% of grid points close to a front have no rainfall, in South Africa and Australia more than 70% are dry.

2.4.1 Contribution of fronts to convective events

At cold fronts, an increased number and intensity of accompaniments of convective events could be identified, i.e. precipitation (e.g. *Papritz et al. (2014)*), hail (e.g. *Browning & Harrold (1970)*), gusts (e.g. *Chromow (1942)*), and lightning (e.g. *Zhou et al. (2002)*). A particular effect is that atmospheric fronts are often associated with intense precipitation, and impact the local distribution of rainfall in space and time (*Catto et al. (2012)*). In this study, in particular the major role of fronts in heavy precipitation events is of importance. For instance, *Catto & Pfahl (2013)* found on average 51% of global precipitation extremes to be associated with fronts, with 75% in the mid-latitudes and 31% in the sub-tropics. By combining an objective front analysis and satellite based global precipitation data, *Catto et al. (2012)* confirmed the importance of fronts to the precipitation climatology in the mid-latitudes, while the importance differs amongst oceanic and continental regions. While in the major oceanic storm tracks over 90% of the annual rain comes from fronts, over parts of the Northern hemisphere continents 30% of the precipitation comes from warm fronts during most of the year. They also state that in a warming climate, the proportion of precipitation associated to fronts could change due to a shift of the storm tracks. Heavy convective precipitation is often associated to cold fronts. *Browning et al. (1975)* separated the regions relative to

a cold front into 3 regions: pre-frontal, at the surface cold front, and post-frontal. They found the orographic effects behind the front to be well defined, but rather slight. At the surface front heavy precipitation appeared regardless of topography, while ahead of the front, orographic effects vary from small to very large, depending on the presence of the 'seeder-feeder' effect. *Wilby et al. (1995)* highlighted the added value of subdividing daily rainfall data sets by frontal influence, while frontal activity may be just one amongst several relevant processes affecting rainfall probabilities. Summarizing fronts have an important role in the precipitation climatology by impact on the atmospheric dynamics. Consequently, rainfall properties differ from frontal influenced situations to non-frontal situations. On an event basis, for Germany, a cold frontal case and another convective situation were compared by *Weijenborg et al. (2015)*. Their study analysed small scale PV structures of convective systems. They found the PV structure to be organized and vertically tilted. The PV structure of the cold front was found to interact with the PV structures of the cells and also to favour the generation of new convection cells.

3 Data

Several data sources have been applied for different purposes in the context of the identification and quantification of extreme convective events (table 3.1). The different datasets are introduced in detail in the following.

Data	period	purpose
Radiosoundings	2001-2015	Validation of stability indices from reanalysis
Reanalysis	1979-2015	Identification of fronts; computation of stability indices; training and validation of regression model
Radolan dataset	2005-2014	Training and validation of regression model
Station data	2000-2010	convective event-dataset
Lightning dataset	2007-2014	convective event-dataset
ESWD	1979-2014	convective event dataset

Table 3.1: List of datasets applied in this study, the available time period, and their purpose.

3.1 Station data

The synoptic station data, which satisfies the standard of the world meteorological organisation (WMO) and has been archived by the Spanish weather service, is used in this study for the identification of convective events. The data is provided on the OGIMET-website (www.ogimet.com). The dataset consists of synoptic data like temperature, dew point, cloud cover, wind strength, wind direction, precipitation amount, and reported weather phenomenon during the past six, three, and one hour. Only the weather phenomenon has been used in this study. The weather phenomenon can be one of 100 different types, such as light drizzle, lightning, heavy shower, providing a detailed picture of the observed weather. A detailed list can be found in the manual of codes of

WMO (1974).

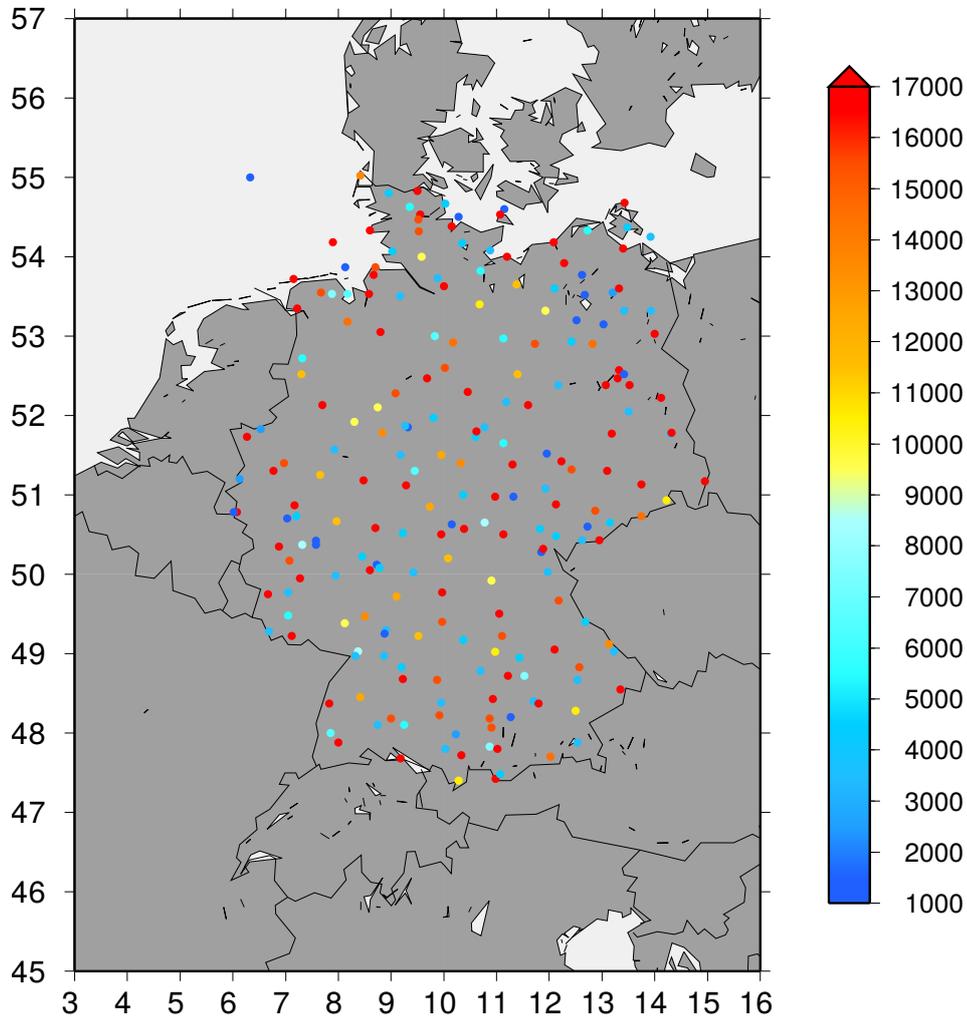


Figure 3.1: German weather stations in the OGIMET database. The colour indicates the number of weather reports within the period of 2000 to 2010.

Within this study, 292 German weather stations were considered. Figure 3.1 shows the location of German weather stations, which measured in the period of 2000 to 2010, and their data coverage for this period. The measurements are performed at most stations at 0,6,12, and 18 UTC. The number of reports in the period differs between the stations. Stations with availability of about 90% of the period 2000-2010 (marked by red colours in Figure 3.1) are relatively homogeneous scattered all over Germany. Stations, which cover less than 50% for this period, are shown in blue colours in Figure 3.1 and also distributed over entire Germany.

code	reported weather
25	rain showers in past hour, currently no prec.
26	snow showers in past hour, currently no prec.
27	graupel or hail showers in past hour, currently no prec.
29	thunderstorm in past hour, currently no prec.
80	light rain showers
81	moderate or intense showers
82	very intense rain showers
83	light sleet showers
84	moderate or intense sleet showers
85	light snow showers
86	moderate or intense snow showers
87	light graupel showers
88	moderate or intense graupel showers
89	light hail showers
90	moderate or intense hail showers
91	thunderstorm in last hour, currently light rain
92	thunderstorm in last hour, currently moderate or intense rain
93	thunderstorm in last hour, currently light snow/snow and rain/graupel/hail
94	thunderstorm in last hour, currently moderate/intense snow/snow and rain/graupel/hail
95	light or moderate thunderstorm including rain or snow
96	light or moderate thunderstorm including graupel oder hail
97	intense thunderstorm with rain or snow
98	intense thunderstorm with sand storm
99	intense thunderstorm with graupel or hail

Table 3.2: List of reported weather codes which have been assigned to the group of convective events. Thunderstorm events are marked in red.

For this study, the dataset was post-processed manually to correct the most obvious digitalization errors (e.g. wrong format) and in some cases, where the decoding was not possible, the corresponding records have been removed from the database. The reported weather phenomenon of the past hour is used within this study to characterize rainfall types. Primarily, convective rainfall situations are of interest, but also large-scale rainfall is important for comparison purposes. Hence, the reported weather phenomena are clustered in different groups:

- convective rainfall
- large-scale rainfall
- other

The members of the two first groups are listed in table 3.2 and 3.3. For parts of this study, lightning events, which is a subgroup of convective events, are regarded separately. The weather phenomena reports can be used to analyse the seasonal cycle of convective events purely based on observations. However, these observations are subjective and have their limitations. The fraction of reported thunderstorms as coded in the

code	reported weather	code	reported weather
50	intermittent light drizzle	65	continuous heavy rain
51	continuous light drizzle	66	light freezing rain
52	intermittent moderate drizzle	67	moderate to heavy freezing rain
53	continuous moderate drizzle	68	light rain and snow
54	intermittent heavy drizzle	69	moderate to heavy rain and snow
55	continuous heavy drizzle	70	intermittent light snow
56	light freezing drizzle	71	continuous light snow
57	moderate to heavy freezing drizzle	72	intermittent moderate snow
58	light drizzle and rain	73	continuous moderate snow
59	moderate to heavy drizzle and rain	74	intermittent heavy snow
60	intermittent light rain	75	continuous heavy snow
61	continuous light rain	76	diamond dust
62	intermittent moderate rain	77	snow grains
63	continuous moderate rain	78	snow crystals
64	intermittent heavy rain	79	ice pellets

Table 3.3: List of reported weather codes which have been assigned to the group of large-scale events.

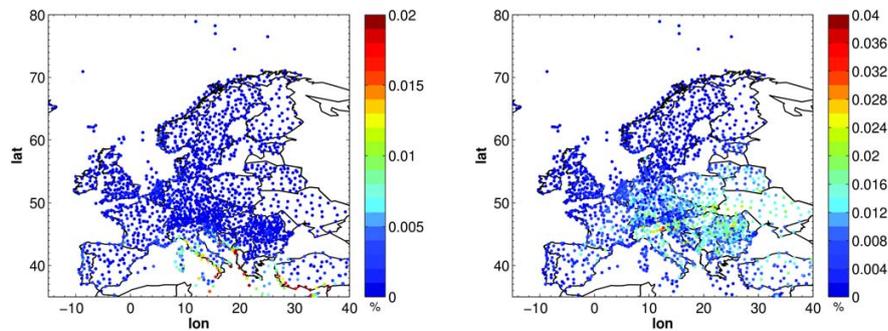


Figure 3.2: ratio of lightning reports for winter (left) and summer (right) for each OGIMET station indicated by its colour (see legend).

synoptic reports (ww=29 or ww=9X) is depicted for the whole of Europe in Figure 3.2 for both, summer and winter half years. The number of thunderstorm events is higher in summer in comparison to winter. Within this study, the reported weather phenomena have been used to identify convective situations based on observations.

3.2 Radiosounding data

Radiosoundings of German stations are used in this study for validation purposes. The stations are maintained by the German weather service (DWD). Radiosondes provide a vertical picture of the atmospheric layering by ascending probes with weather balloons (radiosondes). The radiosondes measure, among other atmospheric variables, temperature and dew points at different heights.

In this study, data for the period from 2001 to 2015 is used. Radiosondes are ascended at some German stations regularly, for every 6, 12 or 24 hours. A map of the stations in Germany and nearby countries performing these weather balloon ascents is depicted in Figure 3.3.

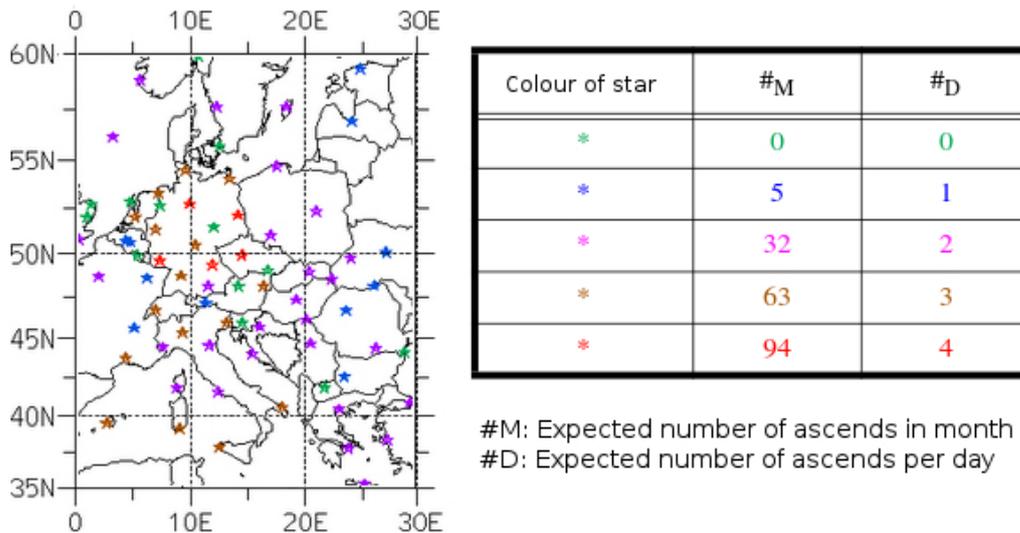


Figure 3.3: Radiosounding stations in Europe. The colour code identifies the frequency of measurements. This map has been compiled by Volker Ermert.

The atmospheric stability is calculated based on the radiosounding data in this study as a reference. The stability and the measured humidity obtained from the radiosoundings are used to validate those retrieved from the later introduced reanalysis dataset.

3.3 Lightning data

Lightning is a phenomenon, which appears only in strong convective cells and under particular conditions, like a particular temperature at a certain height. A dataset of lightning information over Germany from Nowcast GmbH is used in this study. It contains highly resolved lightning positions, intensity [kA] and time. The data covers the period 2007 to 2014.

Lightning positions are obtained by triangulation (see <http://www.nowcast.de>). A station network measures the electromagnetic wave, which is produced by a lightning strike. Since the propagation speed of an electromagnetic wave is the speed of light, the position is triangulated based on the time differences of the signal arriving at different

stations. These stations (part of the so-called LINET network) have a spacing of 150 to 250km. Based on the amplitude of the signal, the lightning intensity can be obtained.

The network of the nowcast GmbH is able to distinguish amongst intra-cloud strikes and cloud-to-ground strikes (see www.nowcast.de). The basic concept is that the signal of a cloud-to-ground strike originates closer to the surface. Hence, it reaches a station faster in comparison to the signal of an intra-cloud strike. These differences in runtime allow the LINET network to distinguish both types of lightning strikes.

The lightning dataset is used in this study to define a subset of intense convective situations. In particular, a situation at a location is regarded convective, if the closest lightning position undercuts a certain distance.

3.4 ESWD database

Several reports of different observed extreme weather phenomena are provided in the 'European severe weather database' (www.eswd.eu). This project is carried out by the European severe storms laboratory (ESSL, *Dotzek et al. (2009)*). Reported phenomena are tornadoes, severe winds, large hail, heavy rain, funnel clouds, dust devils, ice accumulation, damaging lightnings, and avalanches. The event report includes location and time. The available dataset covers the period from 1979 to present. Within this study, we focus on reports in Germany.

The reports are collected from hydrometeorological services, voluntary observer networks (Skywarn and Ketaunos), and public. Hence, event reports need to undergo a quality check to be validated. The database has different subsequent levels for the quality control reaching from 'as received' to 'event fully confirmed'. This ensures that even though a report is not necessarily reliable, it appears within the database and may reach a higher level of event confirmation due to later reports to the database. Within this study, only information for extreme convective events have been taken into account. These are hail, extreme precipitation, and tornadoes. Although these phenomena are not always related, it is not distinguished between these within this study. Summarizing, this dataset can provide a good basis for damaging, intense convective events.

In this study, this dataset defines an alternative subset of convective extreme events. Like for the lightning dataset, a situation was regarded convective, if the closest lightning position undercuts a certain distance.

3.5 Radolan data

The 'Radar-Online-Aneichung' (Radolan) data is provided by the German Weather service (DWD). This dataset contains gridded precipitation data, which is based on radar reflectivity fitted to measurements. The dataset has a spatial resolution of 1km x 1km and covers the area of Germany (area within circles around radar stations in Figure 3.4 a). It is provided in different temporal resolutions ranging from 1h to 24 hours. The product version (RW) with 1h-ly resolution is used here for the period from 2005 to 2014. The database is continuously being updated.

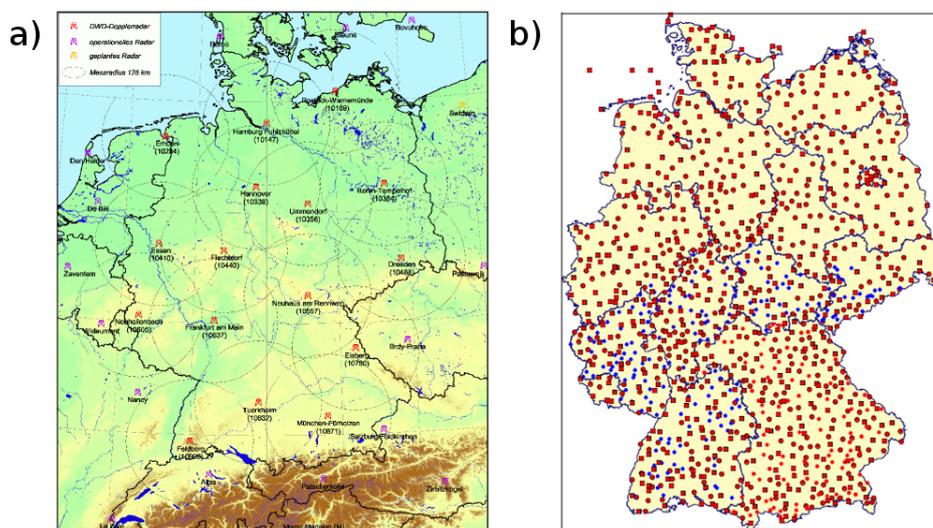


Figure 3.4: (a) Radar stations and (b) Regional distribution of the rain gauges (source: Weigl et al. (2004))

The radar data provides an objective information of rainfall. Due to orography and buildings, the radar cannot provide information of precipitation at the ground. Therefore, it may happen that radar suggests rainfall at a location, in which rain drops do not hit the ground. However, on a larger scale, it is a good measure for the spatial distribution of precipitation.

The radar stations are distributed fairly homogeneous over Germany (see Figure 3.4 a). They are calibrated to rain gauges (ombrometers), which have a high spatial coverage in Germany (see Figure 3.4 b). The effects of shadowing due to elevated objects is reduced based on a topographic database, which has been developed for military purposes for the "Amt für militärisches Geowesen". A number of different adjustments and methods have been applied for the development of this product, which can be found in the final report of the Radolan project (see Weigl et al. (2004)) and publications validating rainfall and

discharge datasets (e.g. *Grasselt et al. (2008)*).

This database provides rainfall information with a high spatial and temporal resolution and is hence very useful for this study for the investigation of rainfall intensity distributions of precipitation events, particularly convective events. The radar's spatial coverage allows the investigation of rainfall intensities not only at the station directly. Therefore, rainfall of thunderstorm cells, which do not necessarily strike the station directly and would be missed without the radar information, can contribute to the rainfall statistics.

3.6 Reanalysis dataset ERA-Interim

The reanalysis data contributing to this study is the ERA-Interim dataset (*Dee et al. (2011)*). The reanalysis dataset is based on a four dimensional variational data assimilation system, i.e. it uses the forecast model for each analysis time step to compile the state evolution. The ERA-Interim data is the latest reanalysis dataset provided by the European Centre for Medium-Range Weather Forecasts (ECMWF). It is available for the period from 1979 to recent. Here, the period from 2000 to 2014 has been applied to match the available period of different data sources.

This dataset includes upper air as well as surface information without gaps and has been homogenized by the ECMWF. This is necessary in order to gain information about the atmospheric stability and for the identification of the large-scale atmospheric conditions (including the identification of fronts). For such calculations, the temperature, specific humidity, horizontal wind fields and geopotential are needed, which are included in this dataset on a 6-hourly basis with a spatial resolution of 0.75° resolution (this corresponds to 80km resolution and to a T255 resolution in the spectral space.) The data is available on model levels as well as pressure levels. The precipitation data included in this dataset is calculated with an integrated forecast scheme and is separated into parametrised grid scale and subgrid scale (convective) precipitation.

The main purpose of this dataset in this study is to provide information on temperature and humidity at different levels. This information is used for the identification of fronts as well as the calculation of stability and humidity indices and the development of the regression model.

3.7 Time periods of different datasets

The presented datasets cover different time periods, and thus cannot be easily combined for a long common period. The station database comprises information from 2000 to 2010, with a 6-hourly resolution. The ERA-Interim reanalysis is available between 1979 and 2014. In addition, the hourly resolved Radolan data used as the basis for the spatial rainfall information is available from July 2005 to December 2014. Associating reanalysis, Radolan and the station dataset is possible for the 2005 to 2010 period. Thus, the station-based time-series for the time period 2005 to 2010 is used for the development of the regression model. The validation of the regression model is not restricted to station data and thus a longer period (2011 to 2014 in addition to the reference period 2005-2010) can be used for this purpose. When the methodology is developed and tested, it can be applied only using the reanalysis data, as they provide all needed information about stability and fronts. In addition to the datasets required for the development of the methodology, the lightning database is used. It is available from 2007 to 2014, and is separated into a reference period (2007-2010) and a validation period (2011-2014).

4 General approach

The pursued aim in this thesis is focusses on the following questions:

- **In how far can stability/humidity indices be used as indicators for the spatial intensity distribution in an area?** In order to answer this question, radar data and humidity/stability indices from reanalysis data are used.
- **What is the quantitative influence of a nearby front on the rainfall properties and on the relationship between local stability/humidity and precipitation?** For this purpose an objective front identification scheme is implemented and applied to the reanalysis dataset. The impact on the rainfall intensity histogram is analysed. It is also explored to what extent the distance from a front is relevant.

Thus, the general aim pursued in this thesis is to estimate the probability and intensity of convective extreme events based on large-scale parameters. The intensity is characterized in this thesis based on rainfall characteristics. Therefore, it consists of 2 steps to infer from large scale weather parameters to convective events:

- First, a method to identify convective extreme events from local rainfall characteristics is developed. For this purpose, rainfall intensity histograms are compiled from radar based rainfall estimates for the vicinity of several synoptic stations across Germany. The optimal radius around a station is identified, where the average rainfall intensity histograms differ most between convective and non-convective types of reported weather. Based on these estimations, convective extreme events can be identified based on the spatial rainfall characteristics.
- Second, a method to estimate the local rainfall characteristics from the large scale atmospheric conditions is established. Various stability and humidity indices are

explored with respect to their relationship with rainfall intensity distributions, aiming at an identification of the convective events from the large-scale parameters like atmospheric stability and low level humidity. In addition, a procedure for the objective identification of cold fronts is implemented. They are a known factor for the occurrence of convective events, and may therefore influence the relationship of rainfall intensity distribution parameters and stability indices. It is analysed in how far the presence/absence of atmospheric fronts modifies the relationship of rainfall properties and the stability/humidity indices.

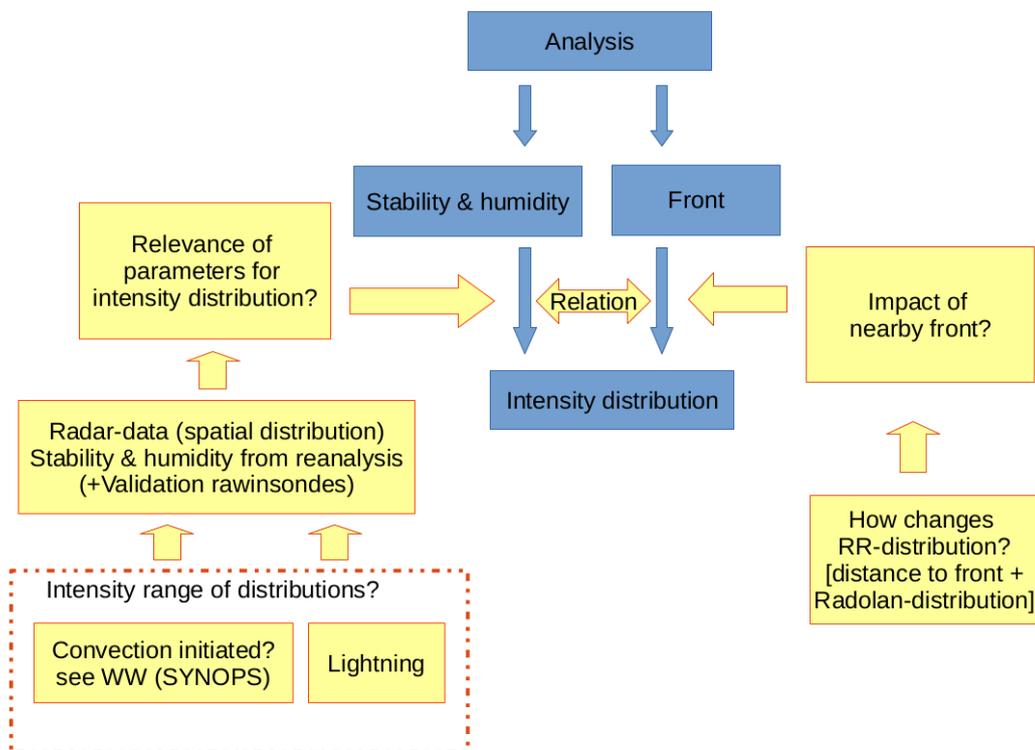


Figure 4.1: Concept for the development of an empirical model for the intensity distribution in convective events based on large-scale parameters. Blue shaded boxes indicate the downscaling procedure, while yellow boxes indicate scientific questions and datasets used to answer these questions.

Figure 4.1 shows schematically the general procedure in blue and the associated scientific questions and related datasets in yellow.

5 Methods

5.1 Rainfall distribution functions

In the literature, different rainfall distribution functions have been used to characterize time-series of observed precipitation data at individual locations. Following e.g. *Eggert et al. (2015)*, also the spatial rainfall intensity distribution in the vicinity of a station may be described with the same kinds of functions.

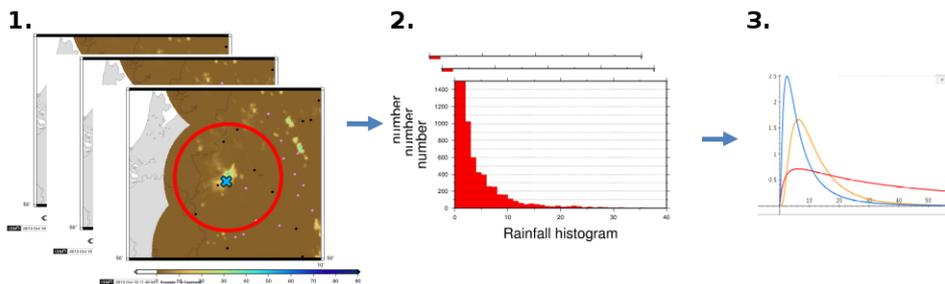


Figure 5.1: Concept to produce theoretical distribution parameters from radar-based rainfall data around a synoptic station. For each time step, a intensity histogram of rainfall within a circular area around each station is compiled. A theoretical distribution is fitted to this histogram.

In the present study, it is first explored which area around a station is optimal to distinguish between convective and non-convective events from the spatial rainfall intensity distribution provided by the Radolan dataset (introduced in section 3.5). The characteristics of the rainfall distribution function can then be derived for any location (and its vicinity) and any particular point in time. Figure 5.1 displays this procedure schematically. In the literature, several distribution functions have been used to describe the

frequency of rainfall intensity. Logarithmic, Poisson and Weibull distributions are just a small sample of many applied theoretical distributions.

5.1.1 Poisson distribution

The Poisson distribution is a discrete probability distribution. It provides the probability of a given number of events (k) for a fixed number of tries, with the chance of an event being constant. For the Poisson distribution, the probability of an event is defined as:

$$P(k; \lambda) = \frac{\lambda \cdot e^{-\lambda}}{k!}$$

where k is an integer and λ is equal to the mean.

The cumulative density function (CDF) is then defined as $CDF(k) = \frac{\gamma(|k+1|, \lambda)}{|k|!}$

The Poisson Distribution has been used to model rainfall in space and time (e.g. *Cox & Isham (1988)*, *Rodriguez-Iturbe et al. (1987)*, *Rodriguez-Iturbe et al. (1988)*). For the application of rainfall distributions, it is also sometimes used together with other distributions, for example to estimate at the same time rainfall occurrence and quantity of rainfall (e.g. *Dunn (2004)*).

5.1.2 Negative exponential distribution

The negative exponential distribution describes the behaviour of the time intervals between events in a Poisson process. This theoretical distribution is defined as:

$$f(x; \lambda) = \begin{cases} \lambda \cdot e^{-\lambda x} & \text{if } x > 0 \\ 0 & \text{if } x \leq 0 \end{cases}$$

with $1/\lambda$ equal to the mean.

This distribution finds application in rain drop size distributions (e.g. *Uijlenhoet & Stricker (1999)*), but also in modelling rainfall occurrences. For example, *Li et al. (2013)* found mixed exponential distributions to perform best for modelling precipitation probabilities amongst their tested distribution functions. It is also a simple distribution to model runoff (e.g. *Moore & Clarke (1981)*).

5.1.3 Weibull distribution

The Weibull distribution is described by the following PDF:

$$f(x; \lambda, k) = \begin{cases} \frac{k}{\lambda} \left(\frac{x}{\lambda}\right)^{k-1} e^{-(x/\lambda)^k} & \text{if } x > 0 \\ 0 & \text{if } x \leq 0 \end{cases}$$

Here, k is the shape parameter and λ is the scale parameter. Both are positive definite. The Weibull distribution is related to the exponential distribution and the Rayleigh distribution. These two one-parameter distributions are special cases of the Weibull distribution.

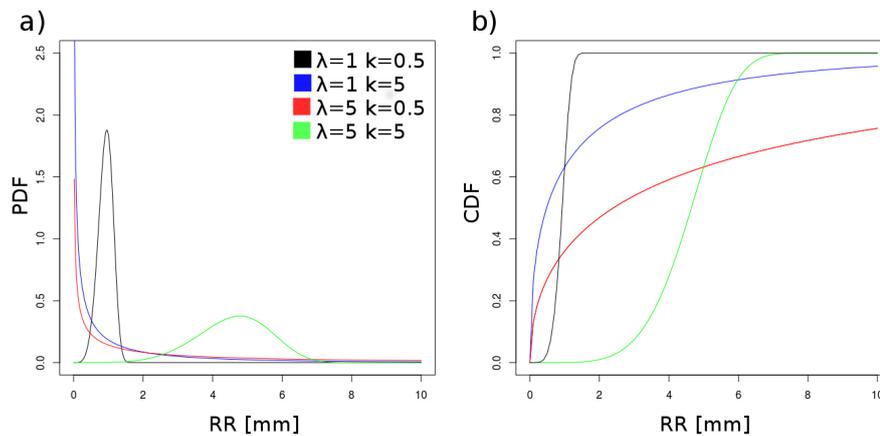


Figure 5.2: (a) Probability distributions of the Weibull distribution for different scale (λ) and shape (k) parameters (see legend); (b) as (a) but showing the cumulative distribution functions.

Both mean intensity and distribution's width can be directly computed from the distribution parameters. The mean intensity is $\lambda \cdot \Gamma(1 + 1/k)$, the variance as a measure of the distribution's width is $\lambda^2 \left[\Gamma(1 + \frac{2}{k}) - (\Gamma(1 + \frac{1}{k}))^2 \right]$. The probability distributions for particular combinations of λ and k are shown in Figure 5.2. The Weibull distribution has per definition a value of zero at a rainfall intensity of zero. Rainfall frequencies of zero (no rainfall) occur quite often, however. Therefore, all rainfall intensities are formally shifted by 1mm/h in order to avoid this problem.

Convective situations are characterized by large differences between rainfall amounts at neighbouring locations, and thus by a large width of the distribution. Instead, non convective events are usually characterized by more similar rainfall amounts. The difference will thus predominantly be reflected by the shape parameter.

5.2 Non-linear regression

In order to estimate the rainfall intensity distribution from parameters of the large scale atmospheric environment, the non-linear nature of these relationships must be taken into account. It is therefore necessary to use a non-linear regression within this study. Here the tool 'Non-linear Least Squares' (nls) from the program container of R is used. The tool is based on the Gauss-Newton algorithm starting with given initial conditions for the parameters. This procedure is described in detail in *Bates & Watts (1988)* and *Bates & Chambers (1992)*.

5.3 Measures of skill

A number of skill scores is designed to analyse the performance of a event prediction. In a perfect deterministic prediction each occurrence or non-occurrence of an event is correctly predicted. The deviation from this optimal situation can be quantified from the numbers in a contingency matrix as shown in table 5.1.

		event observed		
		Yes	no	marginal total
event predicted	Yes	a	b	a+b
	No	c	d	c+d
	marginal total	a+c	b+d	a+b+c+d=n

Table 5.1: Contingency table of event forecasts and observations

The **false alarm rate** is defined as $b/(a+b)$ (compare table 5.1). It can range from zero to one, where 0 indicates that all predicted events were also observed. A value close to one indicates that almost none of the predicted events actually occurred.

The **critical success index** is defined as $a/(a+b+c)$ (compare table 5.1). It can range from zero to one. A value 1 indicates a perfect prediction meaning that b and c are zero. A critical success index of zero would indicate that none of the events was correctly predicted. This measure does not take the cases into account, when a non-occurrence was correctly predicted. This property makes the critical success index an adequate measure for rare events, in which a huge number of correctly predicted non-events results in a false impression of the performance of the prediction procedure.

The **Heidke Skill Score** is often used as a measure for the quality of a forecast scheme, relating it to a climatology-based standard forecast. It was defined by *Heidke (1926)* as:

$$HSS = \frac{(sv - sf)}{(ps - sf)}$$

with

sv: score value $((a + d)/n)$

sf: score of standard forecast $((a + b) \cdot (a + c) + (b + d) \cdot (c + d))/n^2$

ps: perfect score (100%)

The score value is equal to the percentage of correctly forecasted events and non-events (i.e. the relative number of events where the forecast and the observations match). Perfect score is specified as the optimal forecast and therefore defined as 100%. The score of the standard forecast is the (average) relative number of correct forecasts by chance.

In total the HSS ranges from -infinity to 1. A value of 1 represents a perfect forecast and a value of 0 represents the performance of a random forecast. A negative value represents a forecast, which performs worse than the random based forecast. The HSS can be expressed based on the contingency table 5.1 as:

$$HSS = 2 \cdot \frac{(a \cdot d - b \cdot c)}{(a + c) \cdot (c + d) + (a + b) \cdot (b + d)}$$

The **mean squared error skill score (MSESS)** relates a forecast to a reference forecast. It based on the mean squared error (MSE) and is a skill score which is not based on a categorized forecast in contrast to the HSS.

The MSESS is defined as:

$$MSESS = 1 - \frac{MSE_{forecast}}{MSE_{reference}}$$

$MSE_{forecast}$ and $MSE_{reference}$ are the mean squared errors of the forecasts of interest and a reference forecast, with mean squared error defined as:

$$MSE = \sum_{i=1}^n (y_{observation} - y_{prediction})^2$$

n is the number of individual forecasted values, $y_{observation}$ defines the observed value, and $y_{prediction}$ the forecasted value.

Positive values of the MSESS indicate that the forecast is better than the reference forecast. A value of 1 indicates a perfect prediction. A value of 0 indicates that both predictions perform equally well. If the reference forecast is better than the new forecast, the MSESS is negative.

5.4 Analysis of fronts

Fronts in meteorology are moving air mass boundaries, and are often identified manually (*Hewson (1998)*). Such subjective identification often takes a number of different parameters into account, like a strong temperature gradient, high vorticity, a humidity gradient, rainfall and clouds as indicators. The objective front identification method applied in the present work has been originally developed by *Hewson (1998)* and is currently used at the British weather service, contributing there to the manual identification of fronts (*Dominy (2006)*). For the present thesis, a Fortran-Program was coded and applied.

In contrast to objective identification methods based on the change of wind direction (e.g. *Papritz et al. (2014)*, *Rudeva & Simmonds (2015)*), the present procedure is based on the identification of strong gradient of equivalent potential temperature on one vertical level (τ). It defines the front where the change of the gradient reaches a maximum. This means that the spatial derivative of the temperature gradient change is zero in the direction of the gradient (\vec{s}). Thus, the basic equation is formed as follows:

$$\frac{\partial \left(\vec{\nabla} \left| \vec{\nabla} \tau \right| \right)_s}{(\partial s)} = 0 \quad (5.1)$$

with $\vec{s} = \pm \vec{\nabla} \left| \vec{\nabla} \tau \right| / \left| \vec{\nabla} \left| \vec{\nabla} \tau \right| \right|$. Using reanalysis, the gradient at a grid point is computed using the four adjacent grid points in a cross-front direction (defined from the temperature gradient vectors) and a finite difference approach.

This procedure provides candidates for fronts. To identify a front, two more criteria must be fulfilled:

First, the rate of change of the temperature gradient across a front has to be greater than a predefined value K_1 , i.e.

$$-\vec{\nabla} |\vec{\nabla} \tau| \cdot \left(\frac{\vec{\nabla} \tau}{|\vec{\nabla} \tau|} \right) > K_1 \quad (5.2)$$

Second, the strength of the temperature gradient must exceed a threshold K_2 , as not every local temperature gradient maximum is of interest. This condition can be written as:

$$|\vec{\nabla} \tau|_{(x,y)} + m\chi \left| \left(\vec{\nabla} |\vec{\nabla} \tau| \right)_s \right|_{(x,y)} > K_2 \quad (5.3)$$

where χ is the grid length and m an associated fixed factor. An unrealistically large number of extremely short fronts resulting from a simple application of this procedure is reduced by initially applying a smoothing function on the original temperature field.

The procedure coded for the present thesis differs from the procedure described by *Hewson (1998)* in several details: *Hewson (1998)* set m to $1/\sqrt{2}$, which is suitable when using an UTM grid. Using this fixed value with a lat-lon grid resulted in many meridional fronts, which was avoided introducing a spatially varying factor. *Hewson (1998)* used a graphical commercial program for connecting the frontal points into frontal lines. As this software was not available, the identification of the frontal lines (generally running between grid points) was coded as follows. For front points, equation 5.1 must be fulfilled. Neighbouring grid points show an opposite sign of $\frac{\partial (\vec{\nabla} |\vec{\nabla} \tau|)}{(\partial s)}$ in case, a front is between these. Therefore, every two neighbouring grid-points were investigated if this value has an opposite sign. If they have an opposite sign, the position of the zero between them was estimated by using a linear weighting function. In case this point fulfilled both frontal criteria (formula 5.2 and 5.3), it was defined as a frontal point. Figure 5.3 shows exemplary the result of a manual analysis and the corresponding result of the front analysis. The date was chosen to be comparable to the figures of *Hewson (1998)* showing this case study.

Hewson (1998) computed geostrophic wind based on temperatures in order to distinguish between warm and cold fronts. Avoiding this rather uncertain approach, the wind information available in the reanalysis is used. In spite of the smoothing of temperatures mentioned above, the number of very short frontal structures remained high. Therefore, a procedure to connect short frontal pieces to one realistic front and to calculate the front

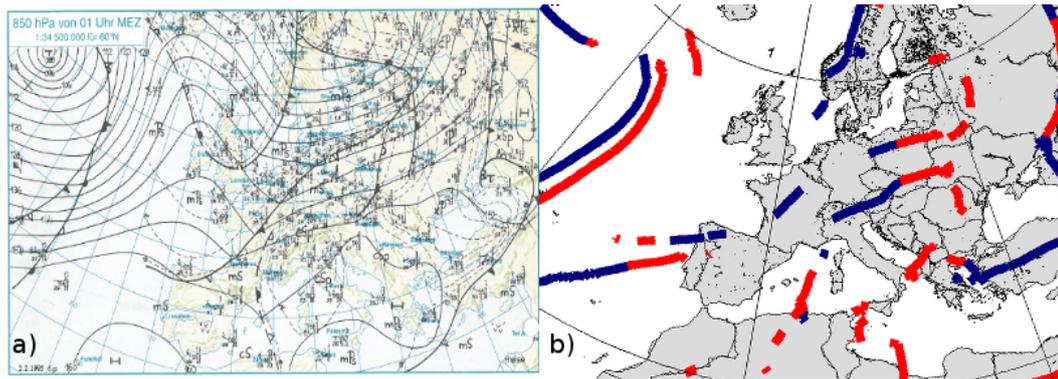


Figure 5.3: Fronts identified: (a) by Berliner Wetterkarte for 850hPa for 2 February 1995 and (b) by the objective front identification method. Blue line: cold front; red line: warm front.

length was developed. In order to connect front segments, a four grid point rectangle (directly neighbouring grid points) around each grid point was searched for front points. In case, multiple front points were found on the edges of this rectangle, they were connected by a line, which defines a front segment. In case one point was part of a warm front and the other one part of a cold front point, an additional point was set belonging to both warm and cold front. Such points appear, for example, in the case of secondary cyclogenesis. The total length of the front is then calculated summing up the respective grid point distances.

6 Results

6.1 Analysis of rainfall in reanalysis

The aim of this and the following section is to analyse the characteristics of the spatial rainfall distribution associated with convective events. The superordinate objective of this is an event definition based on such characteristics.

For a first impression, the ERA-Interim dataset is investigated, which contains 6-hourly forecasts of rainfall separated into large-scale and convective rainfall. These fractions originate from the respective parametrization schemes in the numerical model used for reanalysis. Figure 6.1 shows the rainfall climatology according to the ERA-Interim database. Enhanced precipitation is visible over the Alps (in the South of Germany). Please note that no precipitation minimum was applied here. The fraction of convective precipitation is low (40%) in this region and increases towards the North of Germany (up to 52%). Observational estimates of extreme convective precipitation intensities from radar data reveal higher values for the South of Germany compared to the North in both summer and winter term (see *Eggert et al. (2015)*). For ERA-Interim, an increased amount of convective precipitation is also found over the North Sea near the Netherlands. In general, over Europe a decrease of the fraction of convective precipitation can be identified from South to North (see Figure 6.2).

On average, the convective precipitation amounts about 50% of the total precipitation for the ERA-Interim period (compare Figure 6.1). In agreement with observational results, the fraction of convective precipitation follows an annual cycle (see table 6.1) over Germany (for simplicity in this paragraph defined as the area $5^{\circ}E - 15^{\circ}E$ and

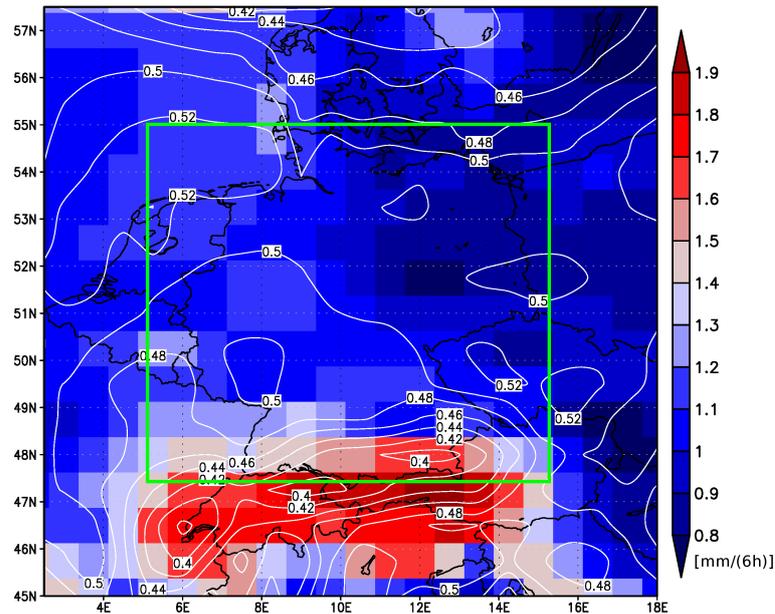


Figure 6.1: Average rainfall rate for Germany and neighbouring areas from the ERA-Interim database for the period 1979 to 2014 as filled grid-boxes [mm/(6h)]. The contour lines indicate the ratio of convective precipitation to the total precipitation amount. The green box highlights the area, which has been regarded as Germany in this section.

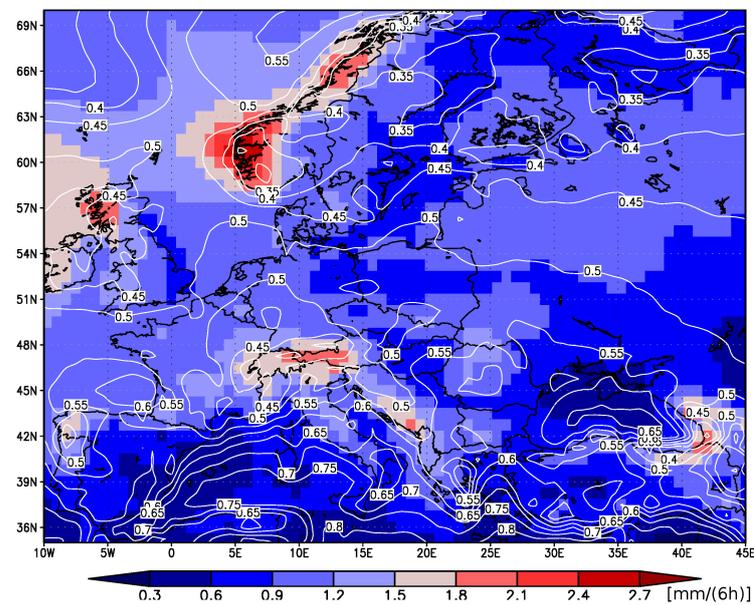


Figure 6.2: Average rainfall rate for Europe from the ERA-Interim database for the period 1979 to 2014 as filled grid-boxes [mm/(6h)]. The contour lines indicate the ratio of convective precipitation to the total precipitation amount.

variable	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug.	Sep	Oct	Nov	Dec
conv. RR [mm]	0.32	0.31	0.39	0.45	0.65	0.81	0.86	0.79	0.57	0.47	0.39	0.37
large scale RR [mm]	0.73	0.62	0.59	0.40	0.41	0.42	0.43	0.40	0.48	0.56	0.69	0.83
ratio conv. RR [%]	30.7	33.5	40.2	53.4	61.5	66.2	66.8	66.7	54.2	44.6	35.0	30.2

Table 6.1: Monthly averages of the ERA-Interim precipitation amount of convective precipitation [mm], large scale precipitation [mm], and the ratio of the convective precipitation to the total precipitation amount for the area of Germany and surrounding ($5 - 15^{\circ}E$ and $47.5 - 55^{\circ}N$).

$47.5^{\circ}N - 55^{\circ}N$). It is least during winter (about 37%) and highest in summer (about 70%). The high fraction of convective precipitation in summer generally supports the necessity of a good representation of convective rainfall to have a well estimated total rainfall amount.

6.2 Spatial rainfall distribution associated with convective events

In this section, the spatial variation of rainfall characteristics for convective situations is investigated based on the Radolan data. The aim is to establish an identification method for intense convective events based on rainfall properties.

6.2.1 Representative radius for convective weather phenomenon

Following *Eggert et al. (2015)*, it should be possible to identify convective rainfall events from the spatio-temporal rainfall distribution around a station (see chapter 2.3). Other than stratiform rainfall, convective rainfall is characterized by a strong spatial variability in the vicinity of a station, which includes both locations with low or no precipitation and locations with very intense rainfall at a specific time. Figure 6.3 shows exemplary the relation between a precipitation field and observed weather at individual stations for 26 July 2008 18UTC. The fact that the rainfall intensities around a station are related to the observed weather at the station, shall be used for the identification of convective events. Rainfall intensity statistics for a single time point in the surrounding of a station are compiled using the Radolan data. Distinguishing convective and non-convective events from the available weather reports, it is investigated which radius around a station should be taken to make the distinction just from the rainfall data.

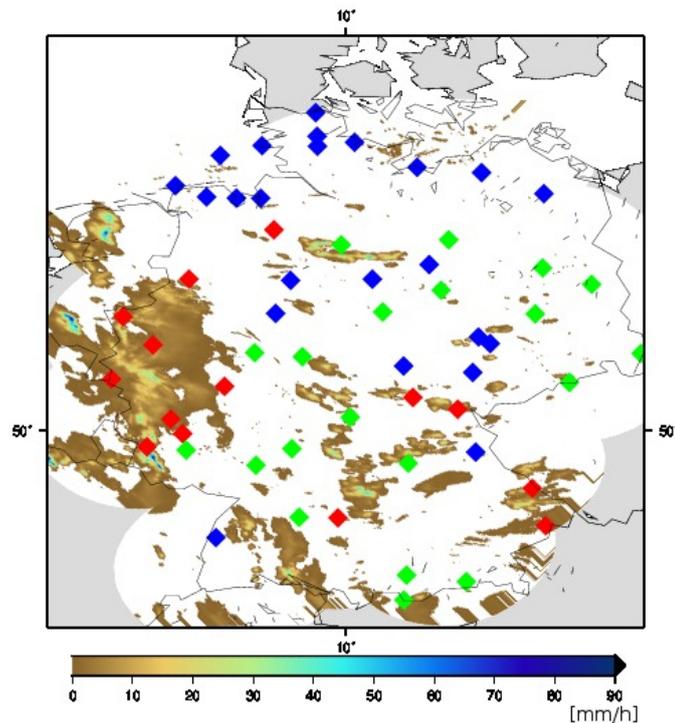


Figure 6.3: Rainfall [mm/h] in Germany based on Radolan for 26 July 2008 18UTC. Dots indicate the corresponding observed station weather at 18 UTC: convective (red), other (green), and no significant type of weather observed (blue).

The Radolan rainfall data within a circle of arbitrary radius around the station are thus considered separately for the group of convective and large scale station SYNOP reports (cf. table 3.2 and 3.3). Characteristics within a suitable radius should be distinctly different for convective and large-scale rainfall events. A suitable choice of radius should fulfil two requirements: First, it should be large enough to allow a reliable estimation of rainfall properties like spatial variability of rainfall. Second, the rainfall properties within the radius should reflect the observed weather at the station rather than the large scale situation. Due to limited computation time and storage amount, six radii of 50km, 60km, 90km, 120km, 150km, and 180km have been tested for the station of Essen only.

The observations in the OGIMET data are available on a 6 hourly basis. For this analysis only values of daytime events (12 and 18 UTC have been used. This limitation has been introduced, because mechanisms at night may differ from those at daytime. For example, relative humidity at low levels systematically increases at night. The relative humidity is one large-scale parameter used later this thesis. The most significant weather at the actual time or within the hour preceding this moment is reported. Therefore, the corresponding Radolan rainfall information for the hour before the report is assigned to

the weather report. Radolan data for other time periods is not used. For each of the three groups of weather phenomena (convective, large-scale, other), an average rainfall intensity histogram is calculated based on Radolan data (see chapter 3.1). The histograms for each radius and group are shown in Figure 6.4a for the period 2007-2010.

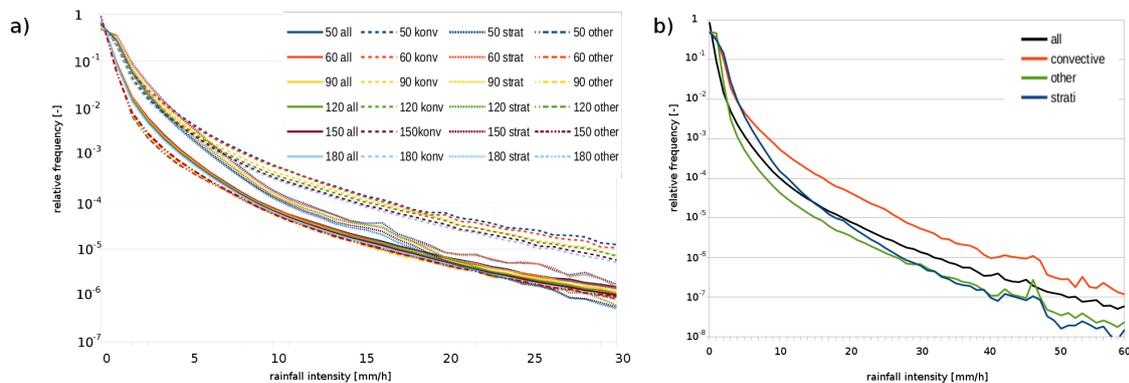


Figure 6.4: (a) Climatologies of Radolan rainfall intensity distribution around the station of Essen for the different radii (50km, 60km, 90km, 120km, 150km, and 180km) and for different groups of reported weather phenomena (all, convective (konv), stratiform (strat), other) for the period 2007-2010; (b) as (a) but for the radius of 90km, the period from 2005-2010, and as an average over all OGIMET stations.

For all radii, an increased number of grid points with high rainfall intensities is found for convective events. This supports the suitability of the approach to identify intense convective events from the spatial rainfall intensity distribution. Large scale rainfall events exhibit a lower spatial variability of the rainfall amounts, although some intense rainfall intensities are included. The study of *Eggert et al. (2015)* also identified some intense rainfall for large-scale precipitation using Radolan data and suggested convective cells embedded in large-scale rainfall to be responsible for this. As expected, the group of other weather phenomena shows the highest percentage of no precipitation in the vicinity of the station itself (see value for zero precipitation at logarithmic scale in Figure 6.4). However, this group shows also some very intense precipitation within the radius considered.

Rainfall intensity distributions for the different radii around the station of Essen are shown in Fig. 6.4. While the curves are rather smooth for low intensities, large variations can be seen at the high intensities (>20mm/h). Particularly, results for a 50 and 60 km radius show such less smooth behaviour, which indicates a larger uncertainty. This is apparently related to the lower number of data entering the statistics, associated with the smaller radii. In addition, local effects influencing the climatology, for example

orographic rainfall, may play a larger role for the small areas. The radii of 90 km and above provide a more stable result for high intensities, apparently providing a sufficient coverage of the intense rainfall amounts. The representation of the reported weather at a station should be based on the rainfall characteristics from the smallest area possible. Thus, a radius of 90 km is a good choice. Further analysis is performed based on this radius.

In order to obtain a sufficiently large basis for the analysis, the described procedure is applied to all 292 stations in Germany for the extended period 2005 to 2010. The average rainfall intensity histogram for each type of weather is depicted in Figure 6.4b. The histogram properties basically remain the same as determined for Essen and the shorter period. Large-scale precipitation is associated with low frequencies of the high rainfall intensities, while convective situations are associated with higher frequencies of grid-points with higher rainfall intensity. It is concluded that the radius of 90km can be applied as a standard. It should be noted, however, that the optimal radius for some individual stations may differ from this choice.

6.2.2 Lightning in the vicinity of stations reporting a convective weather situation

Beside using the reported weather at a station, it is possible to identify intense convection based on lightning, as recorded in the NOWCAST dataset (see section 3.3). Figure 6.5 indicates exemplary for 26 July 2008 that heavy precipitation (Figure 6.5a) is often collocated with positions of lightning measurements (Figure 6.5b) for convective situations. Because lightning appears only for part of the intense convective cells and is not necessarily associated with local rainfall (e.g. "lightning without precipitation"), it can only serve as an indicator for the most intense convective events. In order to analyse if the rainfall distribution characteristics of events with lightning differs from convective situations according to synoptic observations, only the subset of events with rainfall within the radius (according to the Radolan data) is evaluated.

A point in time and space is assigned the flag 'lightning event' if any lightning strikes occurred within the radius of 90km around the station within the preceding hour. Radolan intensity histograms are then subdivided into no-lightning events and lightning events. The average PDFs are shown in Figure 6.6a for the station of Essen and in Figure 6.6b as merged result from all German synoptic stations in the period 2005-2010. As expected, events with high rainfall intensities within the 90 km radius are more frequent

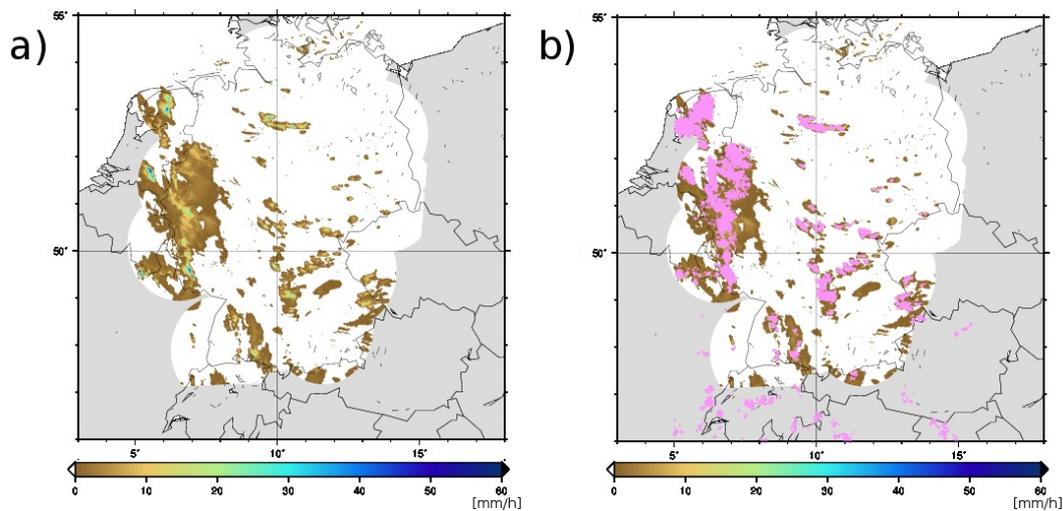


Figure 6.5: (a) Rainfall intensity from Radolan for 26 July 2008 integrated from 18 to 19 UTC; (b) as (a) additional pink triangles mark the lightning positions of the corresponding hour (18UTC-19UTC).

(by almost 2 orders of magnitude) for lightning convective events in comparison to their no-lightning counterparts. This finding is in line with the result of *Gaál et al. (2014)*, who analysed rainfall intensities associated with lightning events.

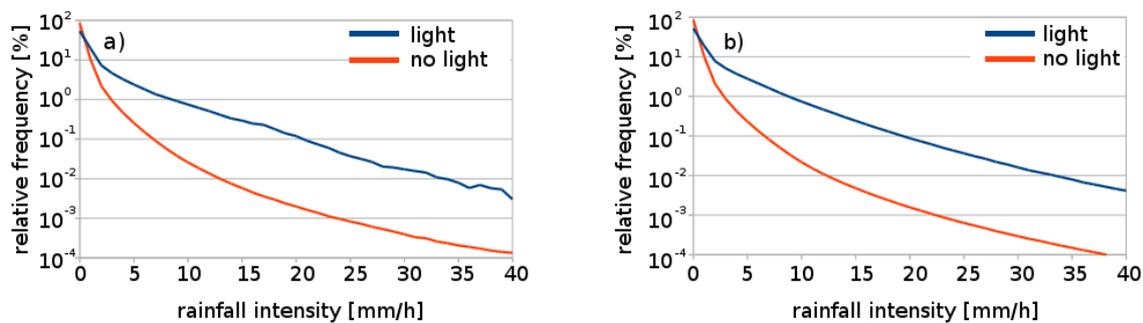


Figure 6.6: Climatology of rainfall intensity distribution averaged over (a) the station of Essen and (b) all stations based on the information if lightning ('light') occurred within the 90km radius around the station or not.

It is confirmed that the difference in the rainfall histograms is stronger between lightning and no-lightning events than amongst the different groups of weather reports (cf. 6.7). Unexpectedly, the events with lightning detected by the observer have lower intensities than those with measured lightning in a 90km radius (compare thick line in Figure 6.7 to dashed red line). It was tested if the observed events are a subgroup of the measured lightning events. This was not always the case, but the reasons for this result are not clear.

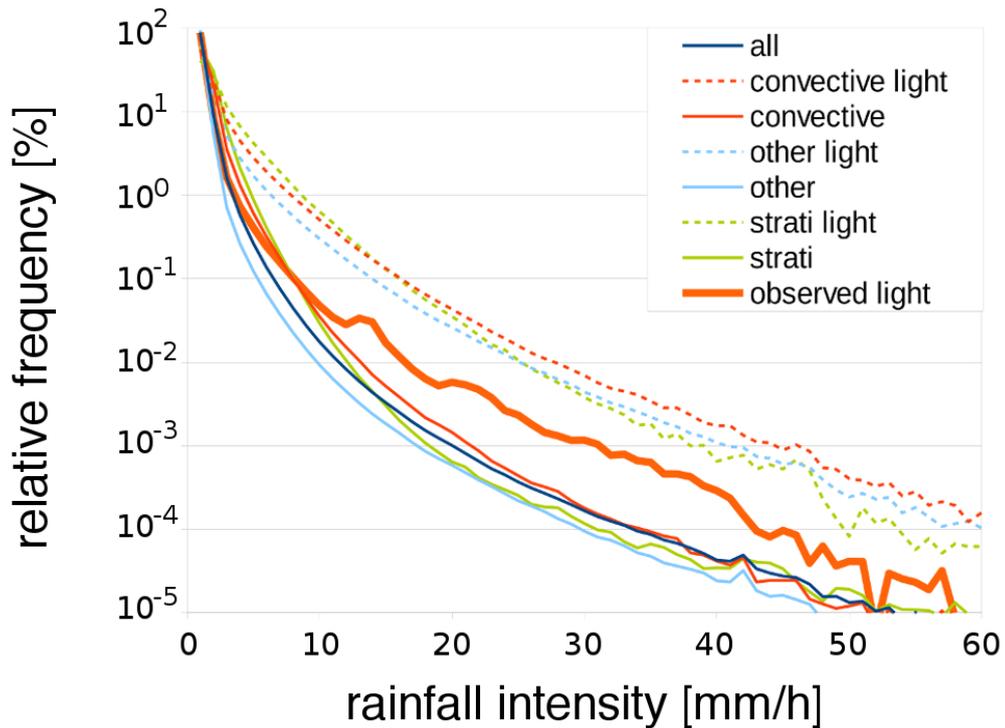


Figure 6.7: Climatology of rainfall intensity distribution averaged over all stations for different reported weather types and based on the information if lightning ('light') occurred within the 90km radius around the station for all stations. It also includes the climatology of rainfall intensity distribution averaged over all stations for situations with lightning reported by the observer (thick orange line).

It is concluded that the identification of very intense convective precipitation events should be based on rainfall using the likelihood of an occurrence of lightning as determined from the rainfall intensity distribution.

6.2.3 Identification of convective events from the spatial rainfall characteristics

As rainfall characteristics for convective and large-scale rainfall (according to weather observations) differ, it should be possible to identify a convective (extreme) event from its rainfall intensity histogram. Here, first weather reports are used, while measured lightning is considered later in this chapter.

In order to find a simple analytical function representing the rainfall intensity distribution, theoretical rainfall distributions are fit to the individual (1h aggregated) rainfall histograms of all stations. The theoretical distributions applied in this study are the Poisson

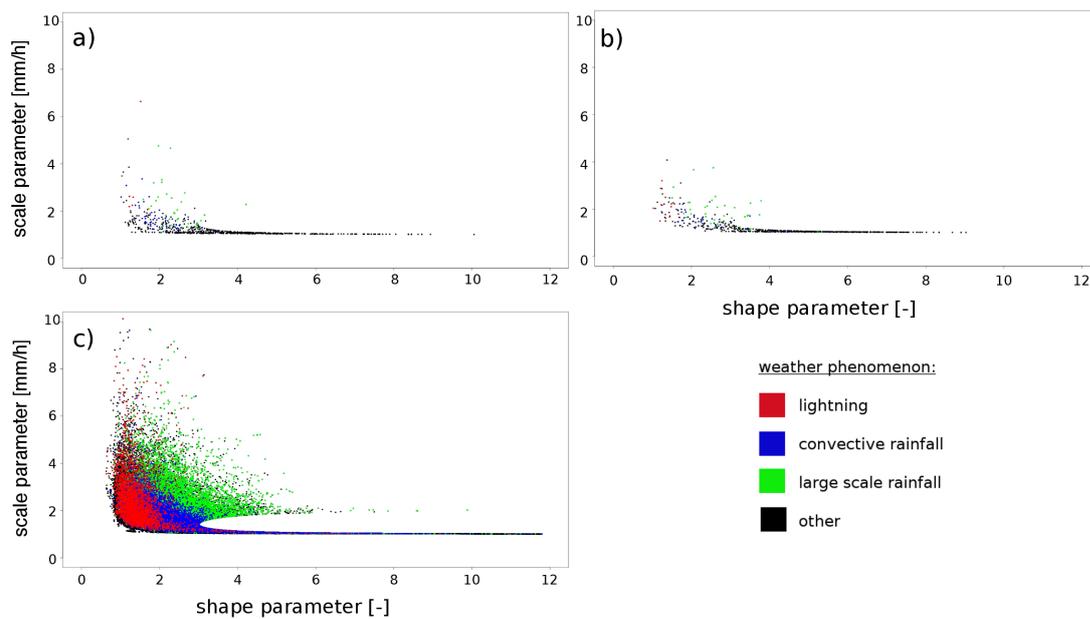


Figure 6.8: (a) Weibull distribution parameters of rainfall distributions for station Essen (summer, 12+18UTC). The shape parameter is depicted on the x-axis, the scale parameter on the y-axis. The associated reported weather phenomenon is indicated by the colour code (red: lightning, blue: other convective events, green: large scale rainfall, black: other situations); (b) as (a) but for Lindenberg; (c) as (a) but for all stations.

distribution, the exponential distribution, and the Weibull distribution. The characteristics of these distributions have been introduced in chapter 5.1. Amongst these distributions, the Weibull distribution showed the best fit results and is therefore applied in this study. This outcome is apparently due to the characteristics of the different distributions. The Weibull distribution can decrease monotonously or may have an absolute maximum in the probability function different from zero, corresponding to the most frequent rainfall intensity in the 90 km vicinity around a station. In contrast, the Poisson distribution must have a defined maximum, while the exponential distribution must have a maximum probability at zero (rainfall), which is not appropriate for most large-scale rainfall situations. In weather situations with single unorganized thunderstorm cells there are many grid points with few or even without precipitation around a station. The advantage of a two parameter distribution (such as the Weibull) is its ability to assign the distribution in one parameter (scale) mainly representing average rainfall intensity in the area, and one (shape) representing the variability within the area. It is expected that in particular the variation is high for convective events (due to high localized rainfall intensities, see e.g. Casas et al. (2004), Pinto et al. (2013), Llasat (2001)).

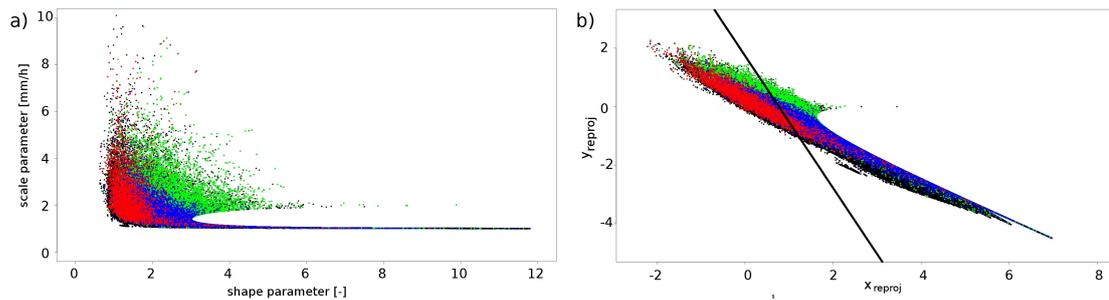


Figure 6.9: (a) Original (identical with Figure 6.8c) and (b) transformed rainfall distribution parameters around a station for single time points and for all individual stations. The colour indicates the associated reported weather phenomenon (see legend in Figure 6.8). The black line indicates the optimal linear border line for the prediction of a convective event in the associated rainfall parameters.

The Weibull parameters are determined for rainfall events using the R routine ‘fitdistr’ of the package ‘MASS’ (cf. *Venables & Ripley (2002)*). Figure 6.8a shows the two Weibull parameters for the station Essen and Figure 6.8b for the station Lindenberg. Each dot represents a rainfall event in the period from 2005 to 2010. The intensity is predominantly reflected by the scale parameter of the Weibull distribution (y-axis). The shape parameter (x-axis) is mainly reflecting the width of the rainfall intensity distribution. A small value of the shape parameter represents a wide distribution and thus a high variability of the rainfall intensities. For both stations, these groups of reported weather (lightning, convective, large scale, and other) show similar parameter distributions. The result for all stations is depicted in Figure 6.8c. The largest scattering of rainfall intensity (and thus a small shape parameter) is found for events of observed lightning (by a weather observer, red dots). This result is consistent with Figure 6.6b showing that there are more frequently intense rainfalls for lightning events than for no-lightning events. From Figure 6.8c, it can also be concluded that thunderstorm events are often characterized by rainfall intensity distributions with a relatively large amount of high intensities. As expected, the majority of events with the lowest scattering of rainfall intensities are large-scale rainfall events, which are at the same time characterized by a medium to high scale parameter. Convective situations without lightning observed at the station are generally found between lightning events and large-scale rainfall events in the diagrams, confirming the results of *Rulfová & Kyselý (2013)*.

The obtained systematic relationships confirm the possibility to identify extreme convective precipitation based on rainfall properties. In order to obtain a simple function describing the distribution of points in Figure 6.8c for the different groups of reported weather, a projection of the Weibull parameters has been performed first (see Figure 6.9).

The projection is defined by:

$$x' = \ln(x_{shape}) - \ln(y_{scale} - 1) \quad (6.1)$$

$$y' = \ln(y_{scale} - 1) \quad (6.2)$$

where x_{shape} is the shape parameter and y_{scale} is the scale parameter of the Weibull distribution.

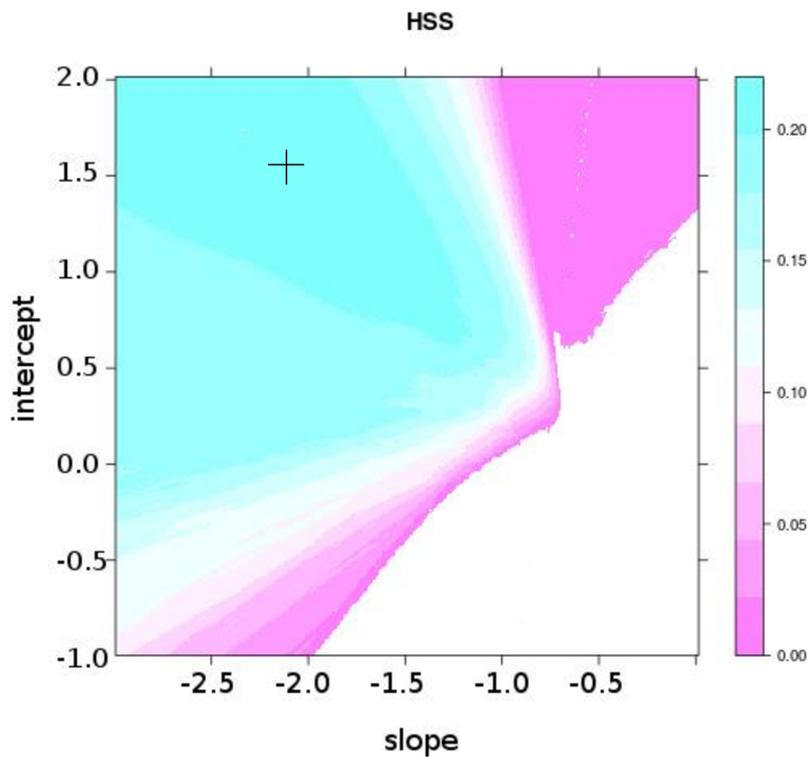


Figure 6.10: Heidke skill score of predicting a convective weather report in dependence of a linear border line in the associated transformed rainfall parameters. The slope is depicted on the x-axis and the axis intercept is depicted on the y-axis. The black '+' indicates the maximum HSS.

The result of this transformation shows that the events of each group are predominantly aligned in specific bands (cf. Figure 6.9b). The optimum boundary line separating convective and non convective event can be defined making use of the Heidke Skill score (HSS). Assuming a linear boundary line, the optimal boundary line was found in terms of HSS by varying the slope and the axis intercept. The HSS as a function of both parameters can be found in Figure 6.10. The associated maximum HSS has a value of 0.22. The equation is $y' = -2.15x' + 1.57$. Thus, an event rated convective, if $y' < -2.15x' + 1.57$.

The corresponding matrix for the prediction is showing the following result:

	observed	not observed
predicted	759	3426
not predicted	1410	35171

Table 6.2: Number of observed weather situations for all stations separated into both cases for convective event prediction and occurrence for the period 2005-2010 during the summer months (April-September) at 12 and 18UTC.

This confirms that the majority of situations are correctly categorized as ‘no convective event’. However, the false alarm rate is very high (~82%), with the alarm is defined from the rainfall distribution and the event defined as a convective weather observation. This can be related to a number of weak events not regarded as convective by the observers. A wrong categorization of the observer might, for example, be associated with the fact that 90 km is a distance larger than the observer’s sight field. Hence, more distant convection within the radius may have contributed to a biased estimate of the high false alarm rate. Irrespective of possible reasons, the high false alarm rate is not satisfactory. Thus, it is tested if this deficiency can be overcome using the NOWCAST lightning data set for the definition of a convective event. As mentioned before, this definition will lead to a focus of the investigation on intense convection.

Analysing the Weibull parameters of the cases characterized by at least one measured lightning within the 90 km radius, it becomes visible that most lightning events are associated with Weibull parameters predominantly defined by a low shape parameter (x-axis) in Fig. 6.11a. Using all grid points over Germany rather than just the weather observing stations, it is found that only about 2% of lightning events are weak rainfall events (scale parameter <1.1). Figure 6.11b shows the projection of the Weibull parameters. In addition, for boxes of shape and scale parameter values the number of events was calculated (see Figure 6.11c) and repeated for the transformed parameters (see Figure 6.11d). As expected, it becomes visible that weak rainfall vents in terms of the scale parameter occur most often. For each box, the relative number of events has been calculated (see Figure 6.11e+f). Very intense rainfall with a high spread in the intensity is always associated to a measured lightning event. The percentage of no-lightning events increases with a increasing shape parameter. The result once more confirms that a low shape parameter is typical for lightning events and also suggests a linear borderline between the ‘certain’ and the ‘improbable’ lightning events.

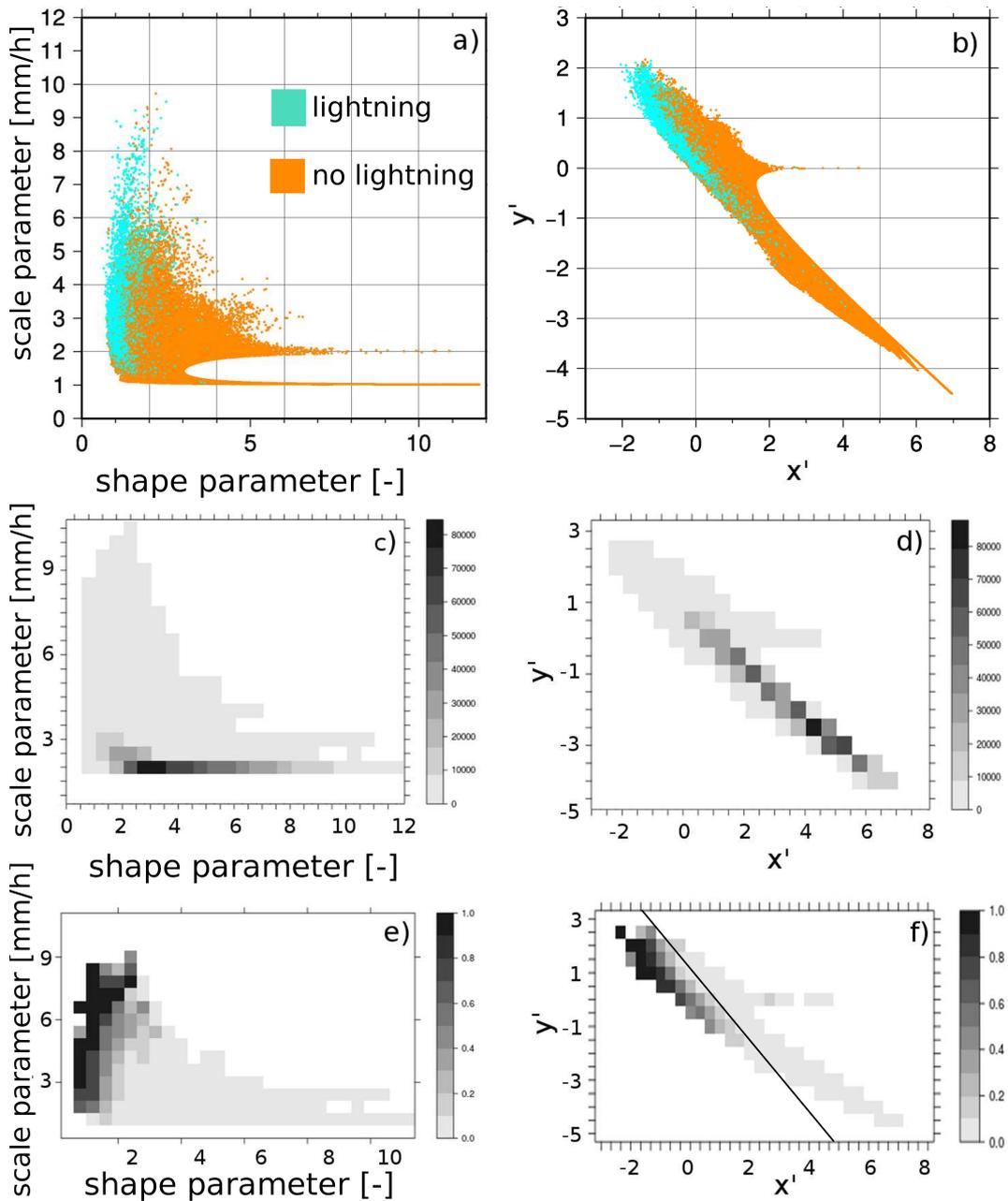


Figure 6.11: (a) Distribution of rainfall parameters around German weather stations for single time points and the information if lightning occurred within the 90km radius; (b) as (a) but for the transformed distribution parameters; local density of points from (a) are depicted in (c); local density of points from (b) are depicted in (d); (e) and (f) show the fraction of events compared to all situations within particular boxes of (a)/(b).

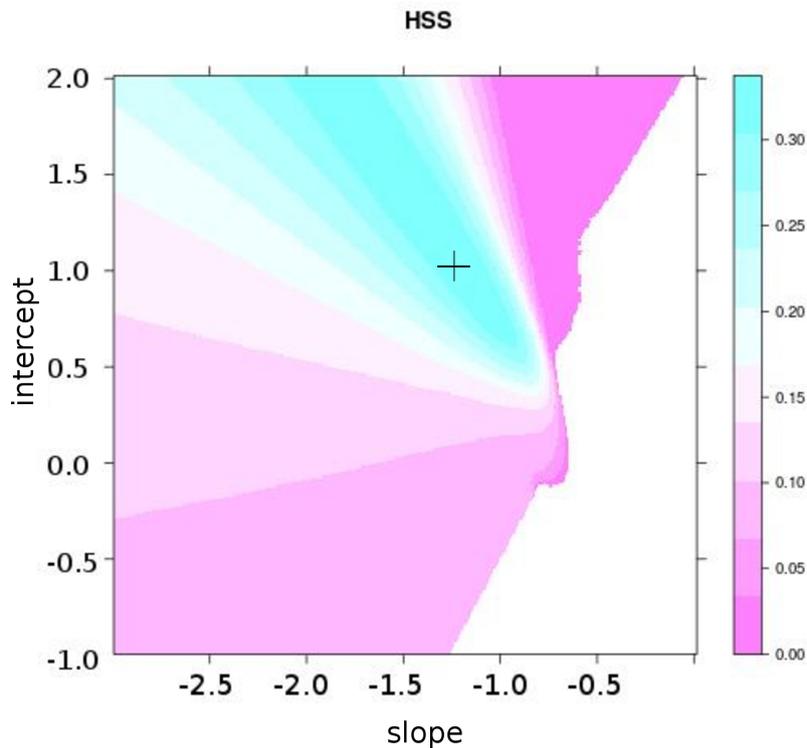


Figure 6.12: Heidke skill score of predicting a lightning event in dependence of a linear border line in the associated transformed rainfall parameters. The slope is depicted on the x-axis and the axis intercept is depicted on the y-axis. The black '+' indicates the maximum HSS.

Based on the analysis of typical rainfall distribution parameters, again a procedure for an objective event classification can be developed. The dependence of the Heidke Skill score on the coefficients on the linear border line is shown in Figure 6.12. The best result based on all stations in terms of HSS is obtained for the linear borderline defined as: $y' = -1.22 \cdot x' + 1.02$. The corresponding HSS has a value of 0.34. The corresponding contingency table shows the result in more detail:

	observed	not observed
predicted	19056	44141
not predicted	1889	244200

Table 6.3: Number of situations separated into both cases for lightning event prediction and occurrence using hourly information of the summer months (April-September) from the Radolan and Nowcast database for all OGIMET station positions.

The corresponding false alarm rate (43%) is much lower than for the observer based definition of a convective event (~82%). This result indicates that the objective lightning dataset is generally indicating more extreme events than the observed weather at the station. It is interesting to see that the optimal slope is close to -1. This value would

correspond to a dependence only on the shape parameter. To predict a lightning event based on the Weibull parameters, the most simple and yet efficient way would be then based only on the Weibull shape parameter. As visible in Figure 6.12, this leads only to a minor reduction in terms of the HSS value. By variation of the shape parameter only, a good performance in the event prediction can be found. The respective performance in dependence is depicted by skill scores in Figure 6.13. Based on the HSS, the optimum shape parameter for the prediction of lightning events is 2.07, when revealing a HSS value of 0.32 (see yellow line in Figure 6.13). All values below this threshold of the shape parameter are predicted as an lightning event. The corresponding false alarm rate is 72% and thus the corresponding hit rate 28%.

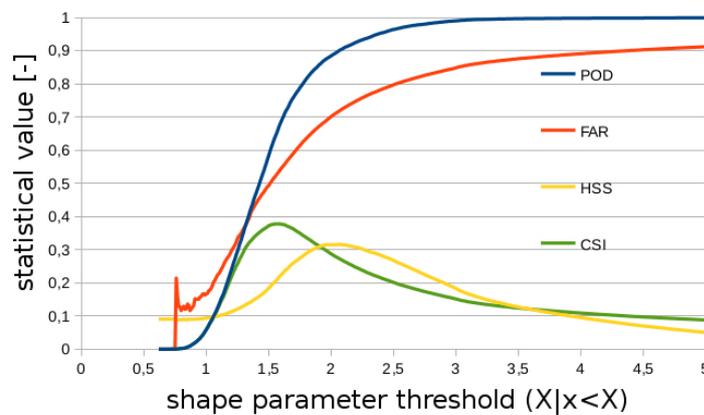


Figure 6.13: Different skill scores (POD, FAR, HSS, CSI see legend; value on y-axis) of the prediction of a lightning event based on undercutting a given shape parameter (x-axis) within the period 2007-2010 based on station positions.

The same procedure as for the probability of lightning events was repeated for convective events within the ESSL database. The optimal linear border line follows the equation $y' = -1.5x' + 0.5$ and results in an HSS value of only 0.11. The associated false alarm rate of 92% can be explained by the observations, which strongly depend on the population density. Although the ESSL database is relatively inhomogeneous, the convective events according to ESSL reports (hail, heavy rain, damaging lightning, and tornadoes) can be identified from the rainfall distribution (cf. Figure 6.14) using almost the same parameter thresholds as found for the lightning events (see Figure 6.11e) and for observed convective events (Fig. 6.8c). Thus, a high fraction of ESSL events appear in a sub-section of the area covered by measured lightning events which is characterized by particularly low values of the shape parameter (i.e. a wide rainfall distribution) and a high scale parameter (i.e. a high rainfall intensity).

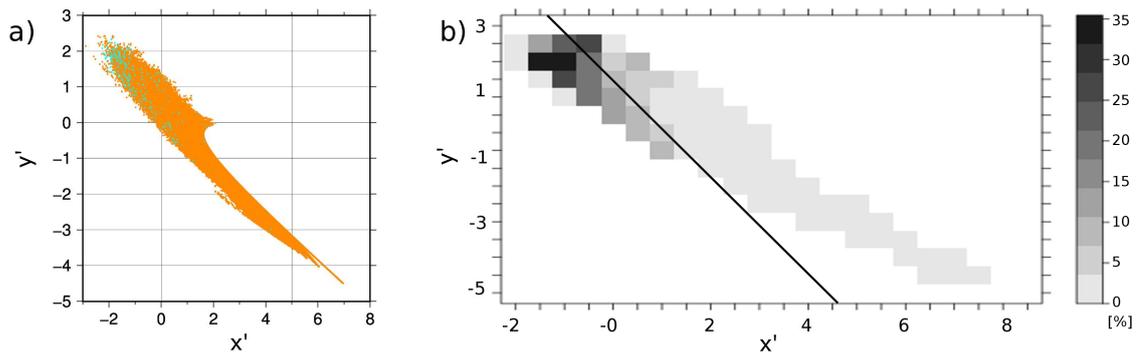


Figure 6.14: ESSL events reported in 90km radius around a station. (a) Distribution of rainfall parameters of ESSL and no-ESSL events for single time points and the associated information and (b) the corresponding event-density.

In conclusion, based on the results for the different sources of convective events (OGI-MET, Nowcast, and ESSL), it is stated that convective events can be described by characteristic values of the rainfall distribution parameters. Convective events are typically associated with a large spread of rainfall intensities in the area considered, implying a low value of the shape parameter. The best result could be obtained for the lightning events. There is a range of distribution parameter combinations for which lightning is probable according to the dataset. Therefore, the method developed in this study is constructed based on the results of the prediction of lightning events based on the rainfall distribution. Additionally, it was found that using only a shape parameter threshold the skill to predict convective extreme events indicated by lightning events, the skill is minor reduced compared to using both parameters. Therefore, the event estimation method subsequently used in this study is based on the threshold for the shape parameter and makes use of the validation results of the lightning prediction.

6.3 Influence of stability parameters on station rainfall distribution

In section 6.2, it has been shown that the occurrence of a (extreme) convective event is reflected by the rainfall distribution parameters. The next step within the analysis is the investigation in how far the rainfall distribution parameters can be estimated from stability and humidity parameters. This investigation is necessary for the development of a model describing the rainfall distribution based on large-scale parameters.

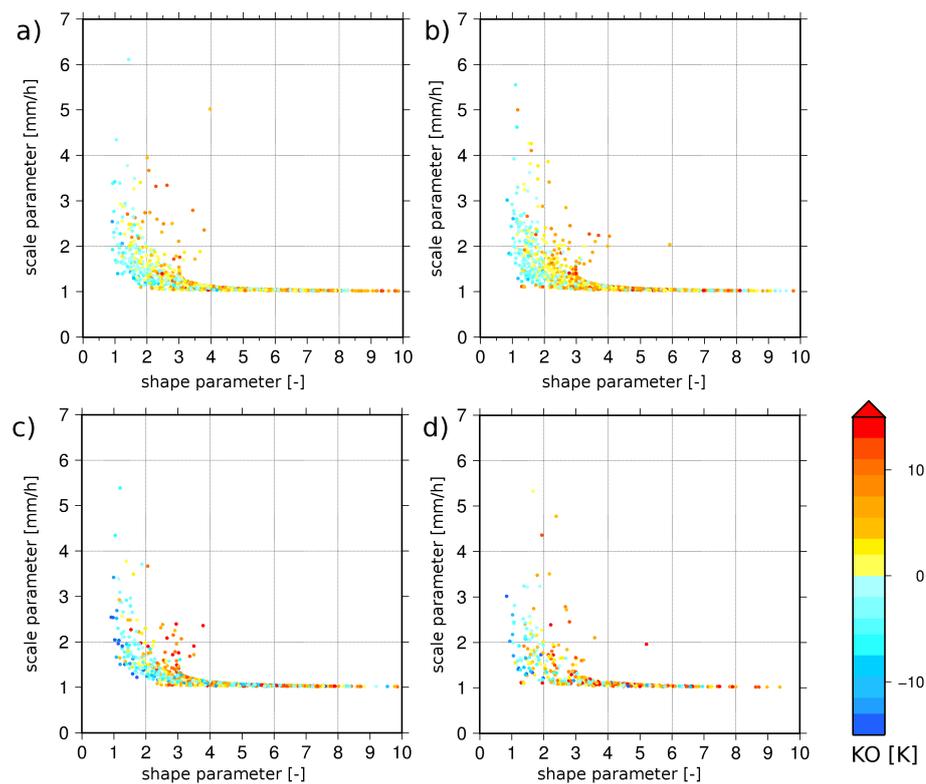


Figure 6.15: Distribution of rainfall parameters around the station of (a,c) Lindenberg and (b,d) Essen for single time points and in colour the associated KO index (a,b) of the closest ERA-Interim grid point and of (c,d) the associated radiosounding.

The stability and humidity indices are initially calculated for the stations Essen and Lindenberg, for which both surface and radiosonde observation are available. The indices are computed using ERA-Interim data (described in chapter 3.6) at the closest grid point. The association of stability parameters and the distribution parameters of the rainfall around the station can be displayed in a scatter plot for single stations. Figures 6.15a,b depict exemplarily the rainfall distribution parameters for the stations Essen and Lindenberg and the associated KO index derived from the nearest neighbour grid point. As throughout this work, only values of daytime events (12 and 18 UTC) have been used. A systematic relationship between the KO index and the distribution parameters is found for both stations. For a particular scale parameter (depicted on the y-axis), the KO index generally increases with increasing shape parameter (depicted on the x-axis). This also means that for lower (i.e. more negative) KO-values, the rainfall shape parameter is on average lower, indicating a broad rainfall distribution function. This result is not unexpected as static stability parameters like the KO index are designed to indicate the probability of intense convection, which is in turn related to a strong spatial variation of rainfall intensity. The effect of ongoing convection, which decreases static

instability (e.g. *Keil et al. (2014); Weijenborg et al. (2015)*) though vertical transports, is apparently not erasing this relationship, but may be reason for part of the variability between individual events. Remarkably, the systematic relationship between stability parameter and rainfall intensity spread can even be found for stable weather situations (values of KO larger than zero), even though these situations are typically characterized by large-scale rainfall. Under stable situations, large scale lifting is needed to produce precipitation. The intensity of such lifting is not directly depending on the stability, but the stability, as an antagonist of (small scale) vertical movements embedded in the large scale dynamics might explain the range of the systematic behaviour of the stability within the rainfall parameters. Stability also influences the Froude number and thus the relationship of flow over and around orographic structures. Similar systematic relationships can also be found for all other tested stability and also for humidity indices (not shown here). The relationship seems to be strongest for the KO index.

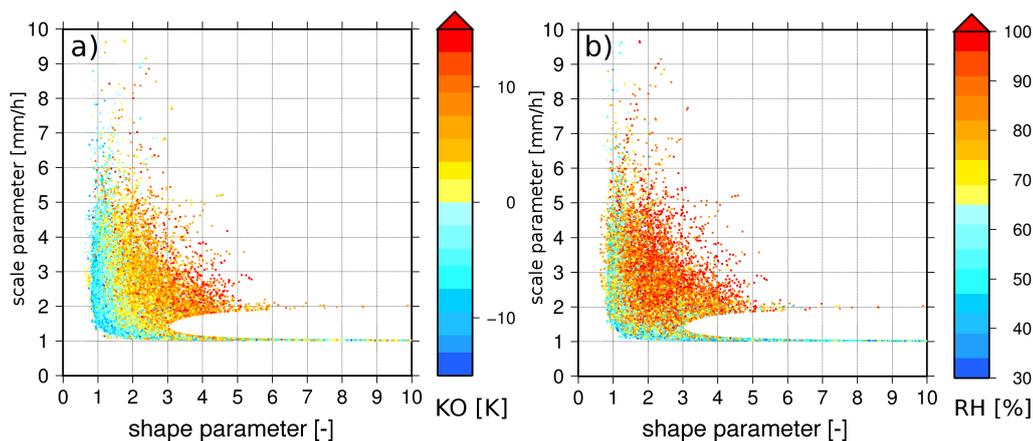


Figure 6.16: Rainfall distribution parameters based on information from all stations. The colour indicates the value of the associated index obtained from the nearest grid point for (a) KO and (b) RH.

Replacing the reanalysis-based stability and humidity parameters with those computed from radiosoundings at the corresponding stations (*Essen and Lindenberg, Figure 6.15c,d*) confirms the relationship. It still needs to be tested, however, if the relationship remains the same for other stations without radiosondes which entered the re-analysis.

For this purpose, results from all synoptic stations are merged (see *Figure 6.16*). The overall relationship does not change and seems to become even clearer by the use of multiple stations. The lowest shape parameters appear for negative KO values, which represent a static instability of the atmosphere. Accordingly, a stable atmosphere favour-

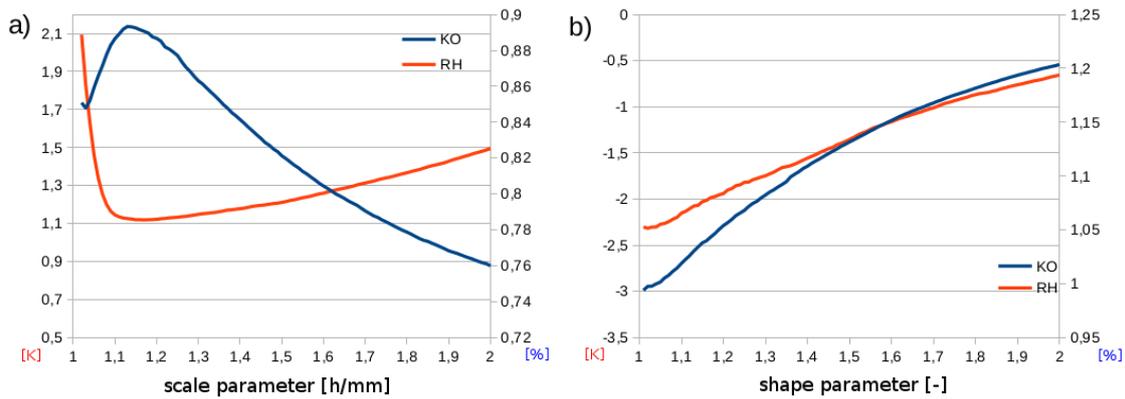


Figure 6.17: (a) Difference between the average value of RH (red) and KO (blue) from ERA-Interim for all events above a threshold of the scale parameter and those events below this threshold, as given on the abscissa. (b) as (a) but for the shape parameter.

ing homogeneous rainfall is reflected by a large value of the shape parameter. This result is in accordance to *Kunz (2007)*, who found the large-scale atmospheric stability to play a major role for the type of precipitation.

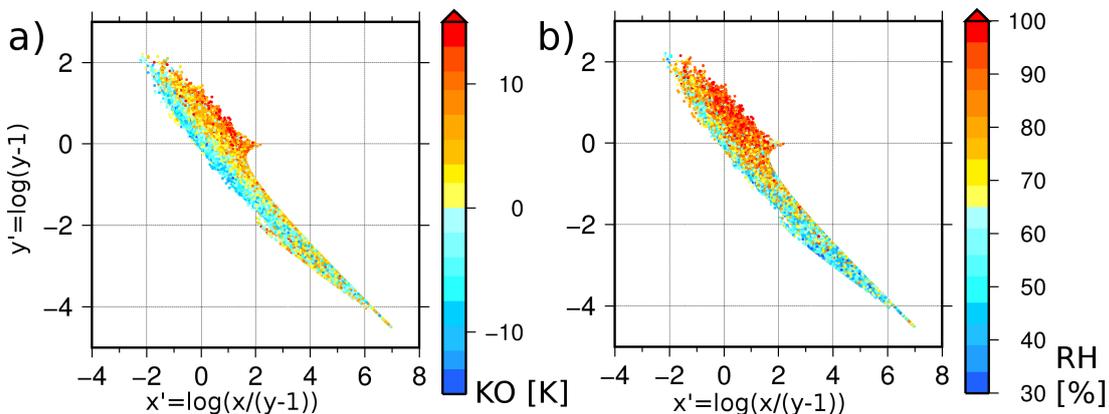


Figure 6.18: Relationship between stability/humidity parameters and transformed rainfall parameters. The colour code indicates the values of (a) KO index and (b) relative humidity.

Given that convective weather phenomena are associated with a low shape and sufficiently high scale parameter, Figure 6.16 reveals that these events are typically characterized by a low KO index and a low relative humidity. This means that there is a wider rainfall intensity distribution (i.e. a larger variation of rainfall intensity) for situations with a lower atmospheric static stability. Figure 6.16b indicates that, for rainfall with large values of scale and shape parameter, high values of the RH are observed. This may be explained by two causes: First, to produce large-scale rainfall, a sufficient amount of moisture is often required in a relatively large area in the lower levels. Second, homo-

geneous rainfall over a large area (indicated by large shape parameter) leads to a high humidity near the surface by evaporating moisture. In the case of a low shape parameter and a sufficiently high scale parameter, these causes are required only over smaller areas to produce rainfall within these, which leads therefore averaged for an entire ERA-Interim grid cell to a reduced surface-near RH.

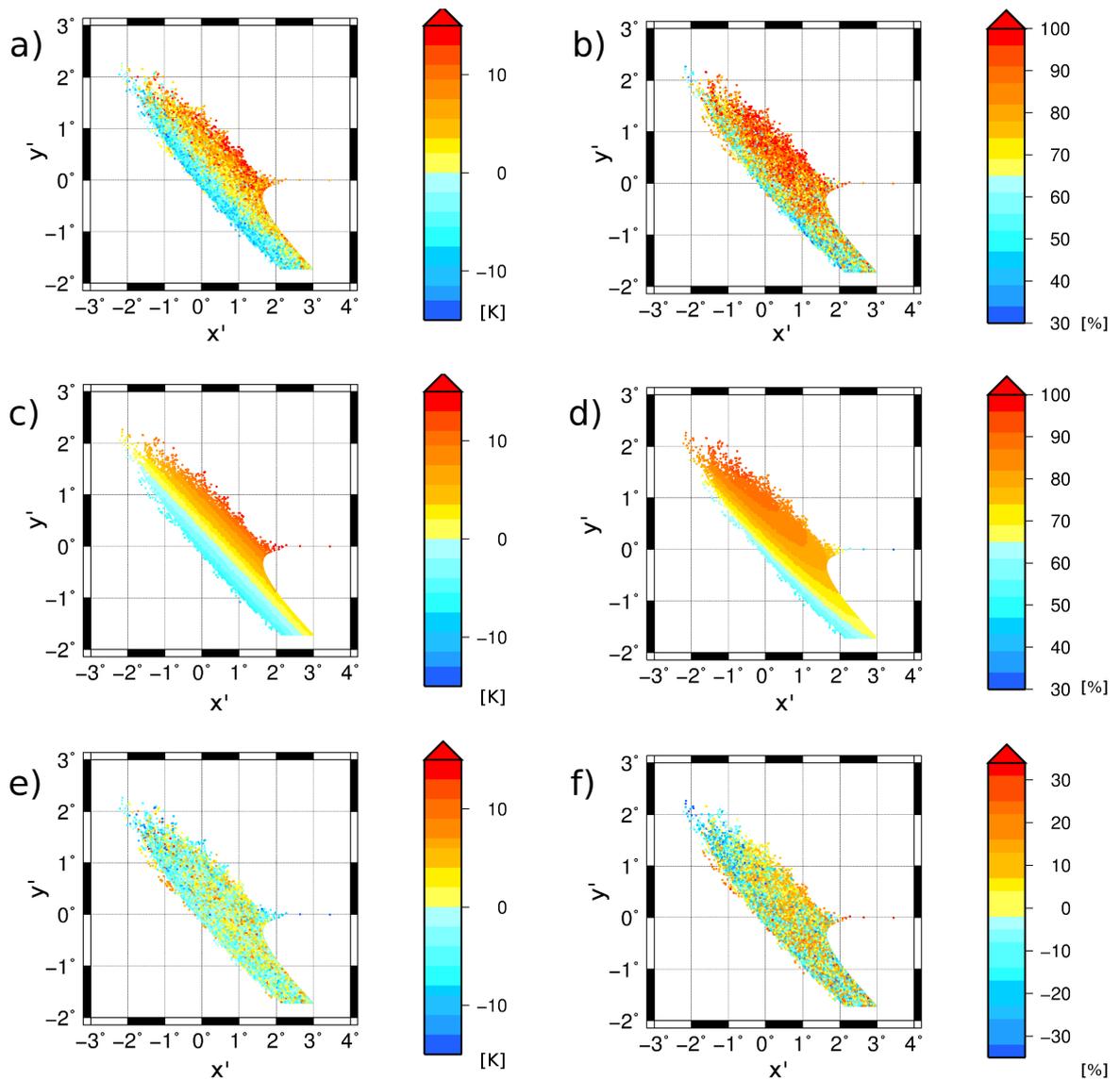


Figure 6.19: Relationship between the KO-index/RH-index and the transformed Weibull parameters: The coordinates indicate transformed parameters (see formula 6.1 and 6.2) describing the rainfall distribution. The colour code indicates the values of (a) KO index and (b) relative humidity, the fitted values based on the estimated relationship (c,d), and the difference between original stability value and fit (e,f).

Weak rainfall events have a different relationship with the large-scale atmospheric parameters. In addition, extreme convective events are the focus of this study. It is hence noteworthy to filter weak events out to obtain a clearer relationship between the rainfall distribution parameters and the large scale atmospheric parameters. For weak rainfall events, the uncertainty in the rainfall shape parameters is high due to the data resolution. Figure 6.17a indicates the difference of all average low-level RH and KO for situations below a certain Weibull scale parameter (threshold depicted on the x-axis) to the corresponding average of data points above this threshold (equivalent to a line parallel to the x-axis in Figure 6.16). A maximum difference for KO and a minimum difference are found for a scale parameter value of 1.18. This can be explained by weak rainfall events, which can be associated with a variety of conditions (beginning of intense events or mature stage of weak events), while more intense rainfall events contain more events at a mature stage. The identified threshold for the scale parameter of about 1.18 is typical for most large scale atmospheric parameters (not shown) and is hence applied. Events characterized by values above this value are regarded for the estimation and quantification the relationships. Figure 6.17b indicates that such threshold is not necessary for the shape parameter, for which the relationship is valid for the range of stability/humidity indices considered.

The application of the previously introduced transfer function (transforming the two Weibull parameters x_{shape} and y_{scale} to x' and y' , see formula 6.1 and 6.2, which were designed to obtain a linear border line between convective events and no-events) draws a simplified picture of the relationship between the large-scale parameters (given by the colours) and the transformed intensity distribution parameters (given by the location in the diagram). In this projection (Figure 6.16 compared to Figure 6.18), the influence of the large scale parameters appears almost the same at first sight, and is orthogonal to the general distribution dots in the diagram. The situations above the previously identified minimum scale parameter of 1.18, which corresponds to a $y' = -1.715$, show less variation around the mean relationship between the rainfall parameters and the stability parameters compared to the rainfall events below this threshold (see Figure 6.19a,b). For values above this threshold, the relationship between rainfall distribution parameters and stability index is approximately linear. Consequently, linear fits have been applied representing the first order relationship. Note that similar relationships can be found for other parameters. For some of them, there is a large variability around the mean relationship. The skill of the linear relationship can be quantified by the explained variance. The results are shown in table 6.4. It is found that the best agreement is found for the KO

index. Also, for the PII and the LI, the explained variances are rather high explained. However, the KO-index is kept as one key variable and is related to the transformed shape and scale parameters (x' and y') according to:

$$KO = a_1 + a_2 \cdot x' + a_3 \cdot y' \quad (6.3)$$

with $a_1 = -9.777$; $a_2 = 20.52$; $a_3 = 27.112$

This formula associates a specific KO index value with a range of likely combinations of scale and shape parameter of the rainfall intensity distribution, but not with an unambiguous specific combination. Thus, a second variable is needed, which has

- a low correlation with the KO-index, but still
- a strong relationship with the rainfall parameters.

In particular, the gradient of the KO index in the rainfall parameters (see 6.19a,c) has to have a different direction than the gradient of the second variable. Based on the two requirements, the low level relative humidity (RH in 1000hPa, which is extrapolated if below surface) was found to be suitable (see Figure 6.19b,d). Please note that the specific humidity of low vertical levels has also been analysed, but revealed a higher correlation with the KO index.

variable	standarddeviation	deviation to plane	explained variance
TT	5.954	5.234	22.7%
KO	4.442	3.581	35.0%
K-index	7.274	7.151	3.4%
CAPE	35.878	35.282	3.3%
S index	17.902	16.786	12.1%
VT	2.568	2.384	13.9%
CAPE _{smith}	223.287	209.123	12.3%
LI	3.079	2.653	25.7%
SREH	34.168	33.795	2.2%
EHI	5.901	5.892	0.3%
PII	0.034	0.028	31.9%
RH(1000)	54.927	51.742	10.9%
SHUM(1000)	0.00229	0.00215	12.4%
SHUM(850)	0.00187	0.00177	10.0%
SHUM(1000→500)	0.00183	0.00172	11.2%

Table 6.4: Information for each stability/humidity parameter (for details see chapter A) about standard deviation, the deviation to the fitted plane and the explained variance.

The attempt to estimate the rainfall intensity parameters from formula 6.3 and a linear relationship for relative humidity was not successful as for many cases an extreme over-

estimation of rainfall intensities occurred. Therefore, a non-linear fit was applied to RH (including second order Taylor coefficients). The respective formula is:

$$RH = b_1 + b_2 \cdot x' + b_3 \cdot y' + b_4 \cdot x'^2 + b_5 \cdot y'^2 + b_6 \cdot x' \cdot y' \quad (6.4)$$

with $b_1 = 64.842$; $b_2 = 30.943$; $b_3 = 42.918$;
 $b_4 = -14.379$; $b_5 = -18.892$; $b_6 = -32.293$

This increases the explained variance of the identified relationship (cf. Figure 6.19). The estimations based on the fits for the KO index and the relative humidity are depicted in Figure 6.19c,d. The difference between the estimated values and the original values, providing the uncertainty (noise) and systematic biases are visible in Figure 6.19e,f. The noise is as expected non-negligible and stronger for RH.

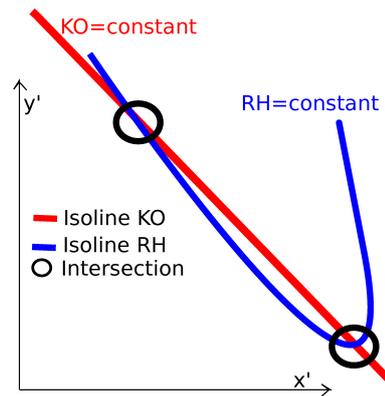


Figure 6.20: Scheme of the detection of the two intersections for given values of KO (red) and RH (blue). The solid lines indicate the isolines within the transformed parameters. Black circles mark the intersections. Black arrows indicate the axes for the transformed Weibull parameters.

Due to the non-linearity in the relationship of RH and the Weibull parameters, however, the curves have either no (for high RH and low KO values) or two intersections. The corresponding formulas are presented in the appendix B. The schematic curves shown in Figure 6.20 also illustrate a source of uncertainty in the estimation. Where the lines associated with a specific KO and RH value are almost parallel (i.e. the gradients are pointing in a similar direction) the resulting intersection point and thus the derived combination of scale and shape parameter is uncertain. This situation is found for a high instability and a low humidity (compare Figure 6.19c,d). For the case of two intersections, it is sometimes found that one of the two intersections represents an unrealistically high amount of precipitation according to the estimated scale parameter, i.e. it is out of

the range of typical climatological values (not shown). In this case, the other combination is chosen. Cases for which the use of both of the two intersections leads to unrealistic rainfall amounts were not found. In the cases of no intersection (7 % of the cases considered, associated with particularly high humidity values or a very unstable atmosphere), the rainfall parameters are estimated using the point on the KO-index curve closest to the corresponding curve for the RH value.

	all		scale (obs) >1.18	
	obs	est	obs	est
mean x'	3.429	2.964	1.297	1.565
var x'	3.516	5.279	0.821	5.594
correlation x'	0.497		0.288	
mean y'	-2.124	-1.786	-0.601	-0.705
var y'	1.760	3.259	0.479	3.742
correlation y'	0.494		0.275	

Table 6.5: Different statistic measures of the transformed rainfall distribution parameters in the observations and the corresponding estimation for rainfall events with an estimated shape parameter below 9.

In order to evaluate the quality of the approach, the rainfall distribution parameters obtained from the model application (using RH and KO) are compared with those obtained from the Radolan data at the locations with station observations. The following evaluations exclude homogeneous rainfall events (here, it is defined as estimated shape parameters larger than 9). Results are shown based on the transformed distribution parameters x' and y' in table 6.5, their counterparts based on shape and scale parameters are shown table 6.6. Considerable deviations between observed and estimated distribution parameters are found in terms of both means and variance. Correlations are 0.4 for the Weibull shape parameter and 0.26 for the scale parameter and approximately 0.5 for both transformed parameters when all situations are concerned. They are a bit lower when restricting the evaluation to events with observed scale parameters larger than 1.18 (i.e. events, which have been used as training dataset to determine the relationships of the individual humidity/stability indices and the transformed rainfall parameters). For the transformed rainfall parameters, the correlations are 0.29 for x' and 0.28 for y' . Also, for this dataset ($y_{scale} > 1.18$), the correlation for the shape parameter (0.31) is higher than for the scale parameter (0.17).

Summarizing, it is concluded that this first estimation (development state 1, called DS1 in the following) shows a certain skill in estimating both transformed rainfall parameters, which reveals in the correlation between observation and estimation. As the model uses

	all		scale (obs) >1.18	
	obs	est	obs	est
mean shape	4.326	3.722	2.113	2.672
var shape	5.529	3.423	0.450	1.897
correlation shape	0.401		0.313	
mean scale	1.298	2.636	1.711	4.541
var scale	0.237	46.139	0.377	102.759
correlation scale	0.256		0.171	

Table 6.6: Different statistic measures of the rainfall distribution parameters in the observations and the corresponding estimation for rainfall events with an estimated shape parameter below 9.

the transformed distribution parameters, the correlation is obviously higher for these compared to the Weibull shape and scale parameters. The Weibull shape parameter reveals a higher correlation between values obtained from estimations and observations than the result for the scale parameter. This indicates that the model performs better in describing the shape of the rainfall intensity distribution than the intensity. However, the performance may be improved using additional parameter's information like fronts, which is tested in the following section.

6.4 Improvement of estimation of rainfall parameters using large scale atmospheric conditions

The first model (called DS1) for the estimation of rainfall intensity distribution parameters from KO index and RH was developed in the previous section. Some parameters and features of the large-scale atmospheric conditions are tested in terms of their potential contribution to improve the model beyond DS1:

- fronts
- differences in local properties of the climatological rainfall distribution values (e.g. mean, variance)
- ERA-Interim precipitation forecast.

These are tested and, if leading to an improvement of the model, included into the model as a new development stage (DS2, DS3,...)

6.4.1 Fronts

Influence of a front on convection

Fronts can have major impact on the probability of thunderstorm occurrence and the local rainfall intensity distribution (e.g. *Berry et al. (2011)*, *Papritz et al. (2014)*). The potential impact of fronts on the rainfall distribution shall be estimated in the following. Additionally, it shall be explored if the presence of a front modifies the relationship between rainfall parameters and stability/humidity indices investigated in the previous section of the thesis. The tool used is the objective front analysis as presented in chapter 2.4. Previous studies have shown that cold fronts support temporally and spatially heterogeneous rainfall, whereas warm fronts favour large scale rainfall. Hence, the different front types are distinguished in this study. The front identification scheme has been applied to the complete ERA-Interim dataset.

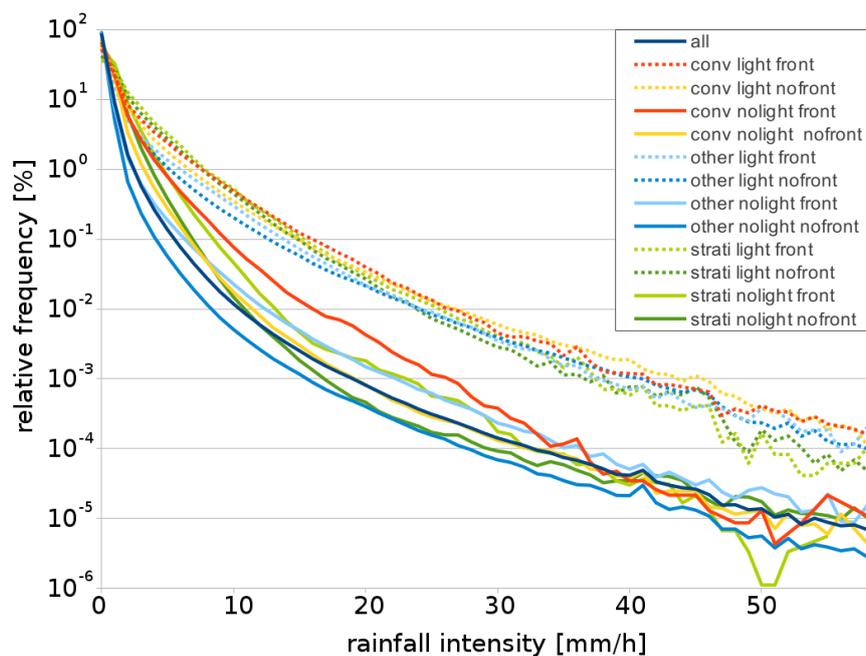


Figure 6.21: Climatology of rainfall intensity distribution averaged over all stations for different reported weather types, based on the information if lightning ('light'/'no light') occurred within the 90km radius around the station, and if a cold front was measured within a 300km radius.

A front is regarded to possibly influence the local weather situation, if the distance to the front is lower than 300km. For each station, the frontal influence was determined for each time step resulting in a time-series for each station. Four different groups of

frontal influence are possible in this series:

- no front
- warm front
- cold front
- both warm and cold front.

First of all, the rainfall intensity distributions assigned to the occurrence of the different weather phenomena (see section 6.2.1) are further distinguished with respect to the presence of a front. First, only results for cold fronts are discussed, because they are regarded to contribute to the genesis of intense convective events. The average intensity distributions are depicted in Figure 6.21, additionally distinguishing if a lightning event was measured within the 90km radius for the corresponding hour ('light front'). They reveal a certain difference depending on the presence (or absence) of a cold front in terms of a larger (lower) relative frequency of grid points of high rainfall intensity. The presence of a lightning measurement within a distance of 90km of the station, however, is clearly more important than the presence of a cold front within 300 km. For example, the light blue line in Figure 6.21 (other weather phenomena with cold front) is for all intensities above the dark blue line (other weather phenomena without cold front), indicating that a larger area within the 90 km radius is covered with intense rainfall for cold frontal influenced situations in comparison to other situations. A particularly strong influence of a cold front is found when considering stations which are reporting convective-weather-phenomena without associated lightning. This systematic indicates that the applied front identification scheme can potentially contribute to improve the estimation of the rainfall distribution parameters.

The frequency of occurrence of the different classes of fronts separated for the reported weather phenomena at all synoptic stations is given in table 6.7.

weather phenomenon	#stratiform	#convective	#lightning	#all
both fronts	25	14	12	377
cold front	1081	815	372	16276
warm front	891	202	125	9737
no front	8430	6834	1323	159461

Table 6.7: Absolute number of weather phenomenon reports at synoptic stations in dependence of the presence of a detected front in the vicinity (closer than 300 km) and presence/absence of fronts of a minimum length of 300km.

A strong influence of fronts on the occurrence of weather phenomena becomes obvious considering the share of reported phenomena (stratiform, convective, lightning) in all reported events. Cases classified as convective or lightning are clearly more frequent with cold fronts ($815/16276=5.0\%$) than with both types of fronts ($14/377=3.7\%$), without frontal influences ($6834/159461=4.3\%$), or warm fronts ($202/9737=2.0\%$). Cold frontal events show a slightly higher share of stratiform events ($1081/16276=6.6\%$) than convective situations ($815/16276=5.0\%$), which was not expected. Expectedly, the fraction of stratiform events is clearly higher for warm front situations ($891/9737=9.2\%$).

Figure 6.22 shows, separated for the different frontal influences, the relative frequencies of the shape parameter obtained from Radolan for all synoptic stations in Germany. It becomes visible that the distribution appears to be generally similar for the cases of no front, warm front and cold front, but the frequencies differ (most often no front influence was found). However, in the relative frequencies (bars as horizontal lines), it becomes visible that the warm and in particular the cold frontal events show a higher fraction of events characterized by a low shape parameter in comparison to those events without frontal influence. As small shape parameters are indicating convective events, the presence of cold fronts enhances their probability, confirming the importance of cold fronts.

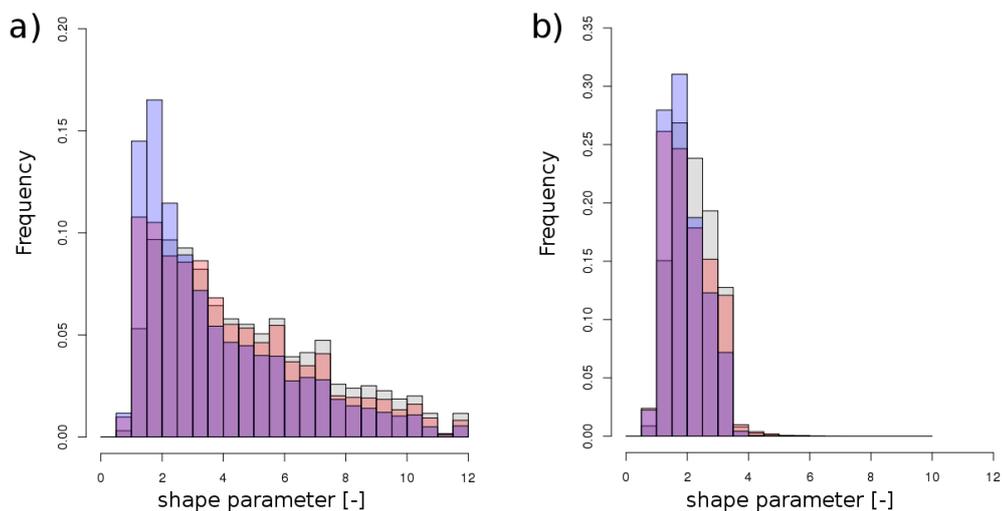


Figure 6.22: Histograms showing the relative frequencies of the shape parameter for the different frontal influences of cold front (blue), warm front (red), and no front in the 300km radius (grey). The result is based on (a) all rainfall events and (b) those with a scale parameter larger than 1.18 according to radar observations. The overlap of histograms is reflected by the respective colour.

The fact that not only cold front, but also warm fronts show a relatively high number of rainfall events with a large spread in rainfall intensity values around the station is surprising. It may be a result of the chosen large 300km radius of frontal influence, or indicative of a deficiency in the front analysis. This analysis uses wind direction and frontal temperature gradient to distinguish warm/cold fronts. In particular, the wind direction is typically weak during summer due to decreased large-scale pressure gradients, potentially inhibiting a correct distinction. An alternative approach has been suggested by *Berry et al. (2011)*. They used low cross-front wind speeds to establish a group of almost constant fronts. However, such procedure requires additional adjustments and has therefore not been applied in the present study. Nevertheless, the percentage of large-scale rainfall events is significantly lower for cold front events indicating that the procedure has a positive effect for our purpose.

Interactions of fronts and atmospheric parameters

A front influences not only the rainfall properties, but also the stability of the atmosphere. For cold frontal events, lower KO values are observed, which indicates more unstable atmospheric conditions. The general characteristics are evaluated quantitatively in this section based on regression results, which indicate the average pattern of the different humidity/stability indices in the rainfall distribution parameters and its strength.

case front	a_1	a_2	a_3
cold	-10.0	25.3	19.3
warm	-9.3	23.4	17.8
both	-9.4	23.5	18.1
no	-9.7	27.4	20.7

Table 6.8: Coefficients of gradient fits (compare formula 6.3) for KO index based on data for different frontal influences.

Figure 6.23 indicates, for example, that under cold-frontal situations the number of stable situations is lower than under warm front situations or the absence of a nearby front. In order to quantify such differences between the different groups of frontal influences, the rainfall events described by rainfall parameters are (as before) associated with the KO index obtained from the nearest grid point. These events are then separated into the 4 different groups of frontal influence (i.e. cold front, warm front, both fronts and no front; see Figure 6.23a,b,c,d. For each class of frontal influence separately, a regression is performed using the previous applied approach for describing relationship of the KO index

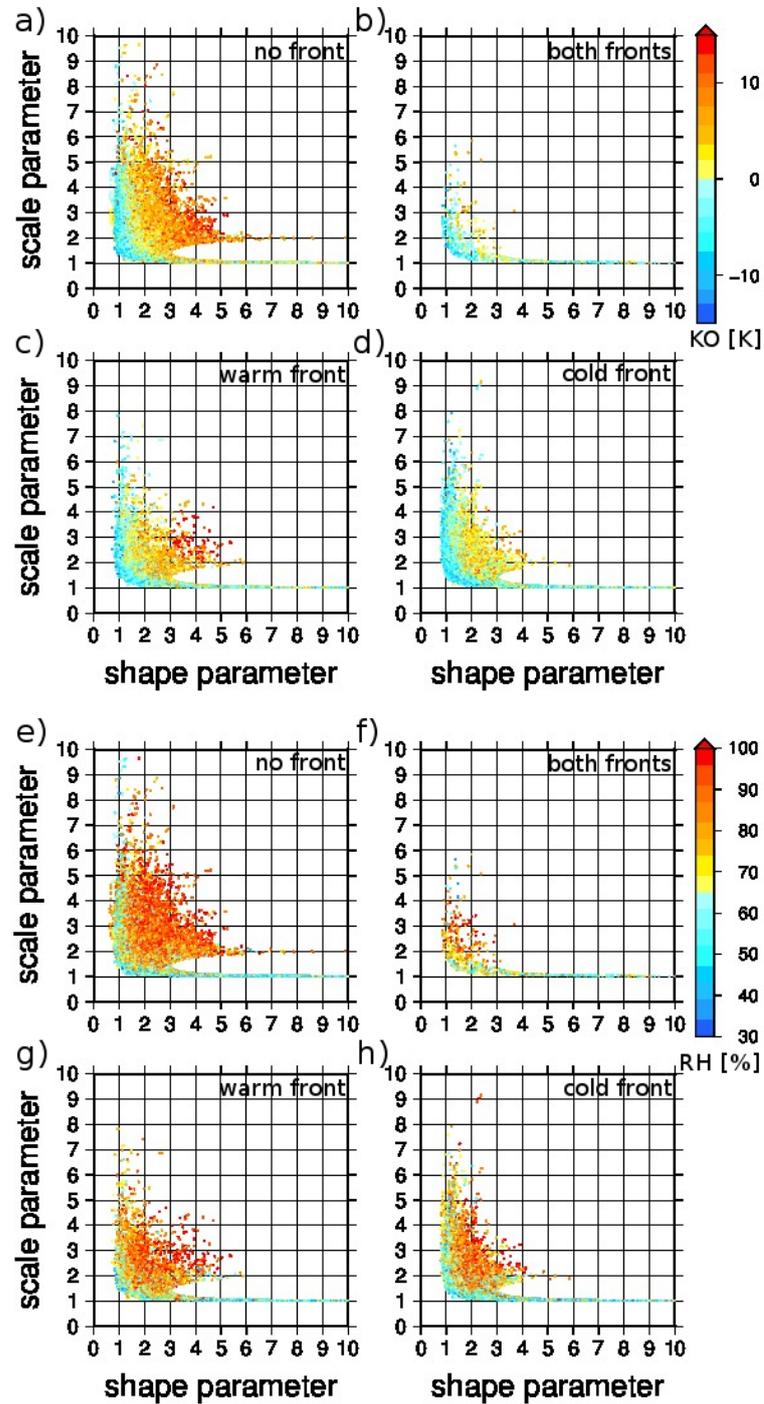


Figure 6.23: Distribution of rainfall parameters around the station for single time points and the associated KO index (a-d; for RH see e-h) of the closest ERA-Interim grid point for all German OGIMET stations separated for the four different frontal influences for a minimum front length of 300 km.

	all		scale (obs) >1.18	
	obs	est	obs	est
mean x'	3.437	2.919	1.300	1.501
var x'	3.505	5.446	0.819	5.764
correlation x'	0.494		0.293	
mean y'	-2.128	-1.748	-0.603	-0.649
var y'	1.754	3.377	0.478	3.874
correlation y'	0.491		0.277	

Table 6.9: Estimation based on front identification: Different statistic measures of the transformed rainfall distribution parameters in the observations ('obs') and the corresponding estimation ('est') for estimated shape parameter below 9 and synoptic stations and time steps with a **corresponding station report** (2005-2010).

with the transformed rainfall parameters. The resulting coefficients are shown in table 6.8. The results of the fitted curves in the rainfall distribution parameters show rather similar coefficients. The found small differences in the coefficients show nevertheless in the correct direction, as for example a lower intercept is found for the group of nearby cold fronts.

Improvement of rainfall distribution estimation by front identification

The frontal influence is included now into the model (DS2). The performance of DS2 should perform better than the previous development stage (DS1) in terms of the rainfall distribution parameters' estimation. As the different identified relationships between the transformed rainfall parameters and KO and respectively RH are not strongly changed, only a minor improvement can be expected. This is confirmed in table 6.10 (and for completeness in table 6.9). The average values for both estimated shape and scale factor remain similar and still differ from observations showing only a negligible improvement. Their variability (in particular for the scale parameter) remains unchanged but by far exceeds the observed variability. For situations the method has been originally designed for (i.e. $y_{scale} > 1.18$), this model shows a small improvement in the correlation with the Radolan based rainfall parameters compared to the model without distinguishing fronts (see table 6.5 and 6.10). Therefore, this method distinguishing between different frontal influences is used further on.

	all		scale (obs) >1.18	
	obs	est	obs	est
mean shape	4.333	3.704	2.115	2.655
var shape	5.528	3.413	0.451	1.890
correlation shape	0.398		0.316	
mean scale	1.296	2.816	1.708	4.934
var scale	0.234	68.099	0.374	157.239
correlation scale	0.232		0.151	

Table 6.10: Estimation based on front identification: Different statistic measures of the two rainfall distribution parameters in the observations ('obs') and the corresponding estimation from DS2 ('est') for estimated shape parameter below 9 and synoptic stations and time steps with a **corresponding station report** (2005-2010).

In order to investigate, why the improvement of DS2 compared to DS1 is relatively weak, we analyse the results of DS1 separated for different frontal influences. The analysis of the cold frontal events within the established estimation method DS1 (see blue histogram in Figure 6.24; note that it is purple, where it intersects with the red histogram) shows that events with cold frontal influence are on average characterized by a low shape parameter values. This behaviour can be found no matter if the consideration is restricted to events which are intense according to observations (Radolan derived scale parameter larger than 1.18) or not (see Figure 6.24a,b). This shows that the estimation DS1 already mirrors the presence a cold front. For warm frontal influenced situations (see red histogram in Figure 6.24), even a higher percentage of low DS1-estimated shape parameters are found compared to cold frontal influenced situations, which was not expected.

	all		scale (obs) >1.18	
	ref.	val.	ref.	val.
mean shape (obs)	4.331	5.046	2.115	2.209
mean shape (est)	3.697	3.636	2.652	2.677
var shape (obs)	5.539	8.623	0.454	0.463
var shape (est)	3.410	3.401	1.888	1.987
correlation shape	0.401	0.384	0.315	0.356
mean scale (obs)	1.297	1.260	1.710	1.673
mean scale (est)	2.830	2.746	4.949	4.653
var scale (obs)	0.235	0.221	0.374	0.386
var scale (est)	67.676	52.385	155.454	112.303
correlation scale	0.234	0.193	0.152	0.105

Table 6.12: Estimation based on front identification: Different statistic measures of the rainfall distribution parameters in the observations ('obs') and the corresponding estimation from DS2 ('est') for estimated shape parameter below 9 and time steps with a **Radolan** observation (for reference period 2005-2010, and validation period 2011-2014).

	all		scale (obs) >1.18	
	ref.	val.	ref.	val.
mean x' (obs)	3.433	3.860	1.298	1.399
mean x' (est)	2.911	2.833	1.494	1.499
var x' (obs)	3.515	4.223	0.819	0.773
var x' (est)	5.469	5.287	5.776	5.565
correlation x'	0.495	0.470	0.292	0.275
mean y' (obs)	-2.126	-2.426	-0.602	-0.656
mean y' (est)	-1.742	-1.682	-0.643	-0.641
var y' (obs)	1.759	2.130	0.478	0.462
var y' (est)	3.394	3.261	3.885	3.700
correlation y'	0.492	0.467	0.276	0.237

Table 6.11: Estimation based on front identification: Different statistic measures of the transformed rainfall distribution parameters in the observations ('obs') and the corresponding estimation from DS2 ('est') for estimated shape parameter below 9 and time steps with a **Radolan** observation (for reference period 2005-2010, and validation period 2011-2014).

The procedure using the relationship obtained from data of the period 2005-2010 with station observations has been applied to the entire reference period 2005-2010, for which Radolan and ERA-Interim data were available (not limited to station data availability, in contrast to results in tables 6.9 and 6.10), and the validation period 2011-2014. Please note that the computation for the reference period contains the previously used data pairs and is therefore from previously introduced results. The results for the extended reference period and the validation period are shown in table 6.11 for the transformed Weibull parameters and in table 6.12 for the Weibull shape and scale parameter. For the events with an estimated shape parameter lower than 9 (compare section 6.3), the correlation of both observed and estimated shape parameter is in the order of 0.3 to 0.4, irrespective of the time period considered, and the possible restriction for observed scale parameter (larger than 1.18, see section 6.3). The correlations between estimated and observed scale parameter are lower (0.2 for all events, 0.1 to 0.15 for those with larger shape parameters, see table 6.12). Correlations for the validation period are only slightly reduced compared to the training period, indicating the stable performance of this stage of the developed method. There are still some events with an excessively high estimate of the scale parameter, causing an extremely high variance of the scale parameter around its mean. These are events of very high instability/humidity for which the intersection method failed.

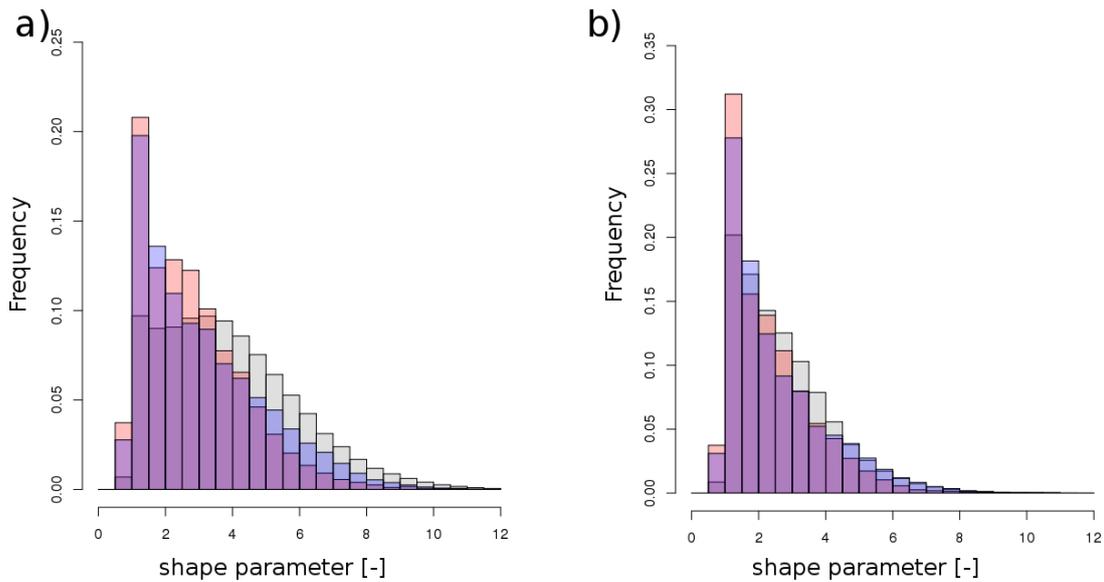


Figure 6.24: Histograms showing relative frequencies of the shape parameter obtained from the model DS1 for the different frontal influences of cold front (blue), warm front (red), and no front in the 300km radius (grey). The result is based on (a) all rainfall events and (b) those with a scale parameter larger than 1.18 according to radar observations. The overlap of histograms is reflected by the respective colour.

Summarizing, the inclusion of the front identification to the model led to an improvement of the model. Therefore, this approach has been included in the current development state of the model (now called DS2), which is further improved in the following.

6.4.2 Bias-correction

A major issue in the estimation remains the scale parameter. Its estimations can be improved using a bias-correction taking into account the location of the stations within Germany. Figure 6.25 shows the average Weibull parameters for Germany derived from Radolan rainfall for the ERA-Interim grid points¹. High values of the shape parameter are observed as a climatological mean for the North of Bavaria, the South of Mecklenburg, and the Northwest of Lower Saxony. The scale parameter shows climatologically low values over south Germany and a slightly reduced average over western Germany (almost the region of North Rhine-Westphalia). Low values of the shape parameter can be either explained by frequent more intense but heterogeneous rainfall intensities (e.g. by strong orographic structured terrain, which often initiates convection) or by many very weak rainfall events (due to few grid points exceeding the minimum rainfall intens-

¹In the Radolan dataset no data beyond the national boundaries of Germany is available. Thus, for grid points at the German boundaries an estimate based on the available data is used.

ity, the rainfall distribution results in a low shape parameter). To take such local effects into account requires a bias-correction.

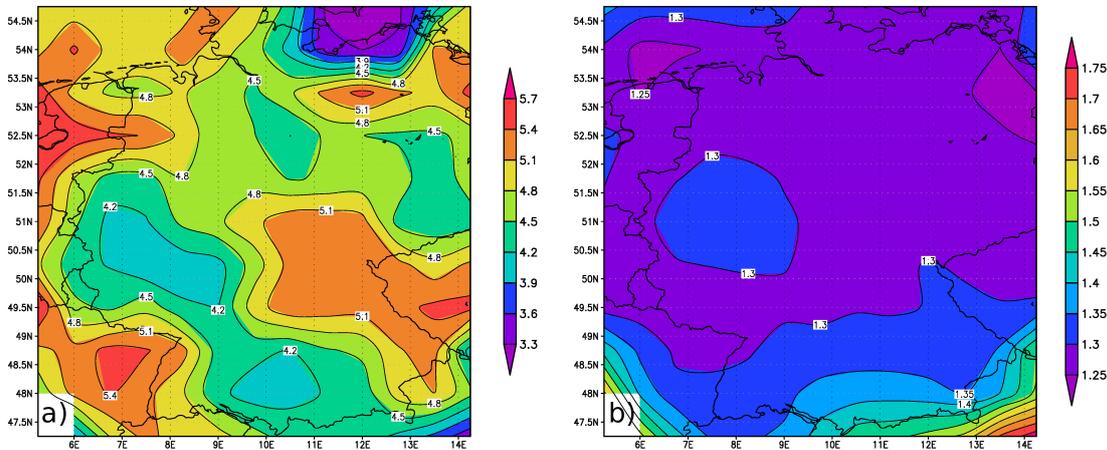


Figure 6.25: Spatial distribution of the mean (a) shape and (b) scale Weibull parameters obtained from Radolan for the period 2005-2014 for the ERA-Interim grid points.

A bias-correction is applied to the estimated rainfall parameters from DS2 (i.e. based on the KO-index, humidity, and the presence of a type of frontal structure). This procedure should not only improve the local estimation of the parameters, but also help to adjust those events which could not be estimated using the intersection method due to extreme instability/humidity. As an example, Figure 6.26a shows the scatter-plots and the associated histograms of the transformed rainfall parameters derived from Radolan for the station of Essen and Figure 6.26b the corresponding scatter plots and histograms for the estimation model DS2 (which uses using KO, RH and the different frontal influences). It becomes evident that there is a group of events estimated with a high transformed parameter y' (larger than 2). These are events estimated without intersections. For this subgroup, the estimated values deviate stronger from the observed parameters in terms of average value and the correlation in comparison to the subgroup of estimations with intersections. This motivates the bias-correction to regard both estimation situations (estimation with and without intersections) separately. In addition, the result shows a difference in the skewness of these estimated rainfall parameters to those obtained from Radolan. Again, this is in particular evident in the example for the estimations not based on intersections. For the station of Essen shown in Figure 6.26, the skewness of these estimated transformed parameter y' is about 0.99, while it is only half of this value for the transformed parameter y' from the observations (skewness(y')=0.46). A similar result is obtained for the transformed parameter x' (skewness from observation: -0.35; skew-

ness from DS2: -0.75). Such relatively strong differences motivate the bias-correction, in which three different statistical values are corrected. These are the mean, standard deviation and skewness. Again, the years 2005-2010 of the ERA-Interim period serve as reference period for the training of the bias-correction and the period 2011-2014 as validation period. Please note that the bias-correction has not been developed station-wise for each type of front due to the lack of sufficient data.

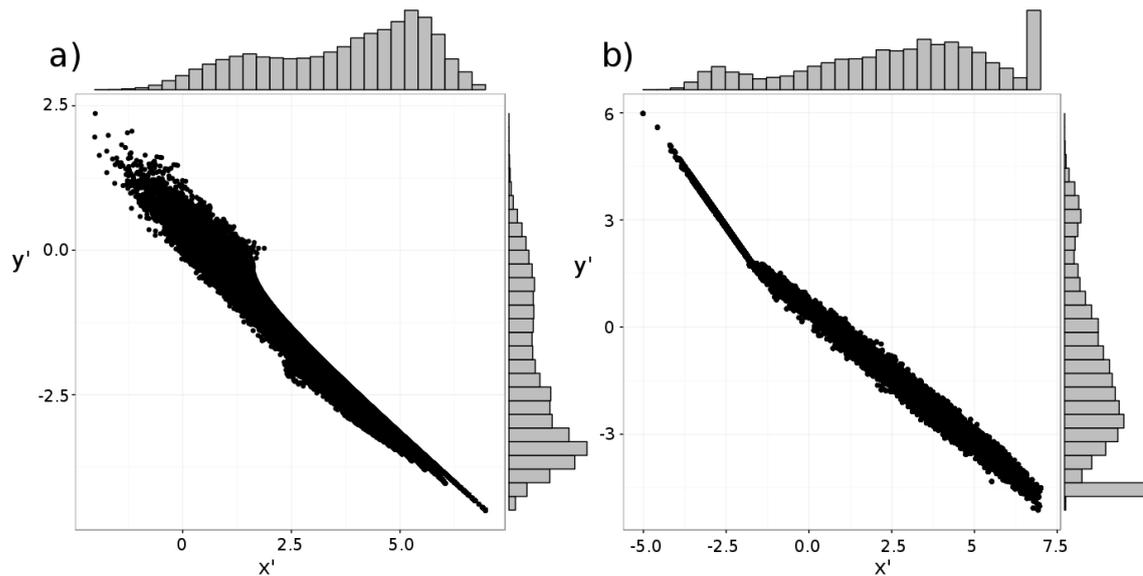


Figure 6.26: Scatter plot of the transformed Weibull parameters, x' (x-axis) and y' (y-axis), obtained from Radolan for the period 2005-2014 for the ERA-Interim grid points obtained from Radolan observations (a) and the estimation DS2 (which uses KO, RH, and fronts). Histograms for each of the estimated transformed parameters are depicted at the boundary of the scatter plot.

For the correction, the estimated local rainfall parameter distribution needs to be adjusted to fulfil the properties of the station's rainfall climatology. It is performed for the transformed rainfall parameters, which ensures that the results of the bias-correction are physically correct (i.e. shape parameter larger than 0 and scale parameter larger than 1). The procedure consists of two steps. In a first step, the skewness is corrected by deformation of the parameter space. For this approach, a function is searched with the following property: A strictly monotonic, continuous derivative is required (to conserve the order of the estimated values and to enable the variable correction of skewness by adding a constant) ranging from 0 to infinity (in order to enable create all stronger skewed distributions). The logarithmic and the exponential function are the canonic solutions fulfilling these requirements. The logarithmic function was chosen due to numerical aspects. The bias-correction was performed involving the logarithmic

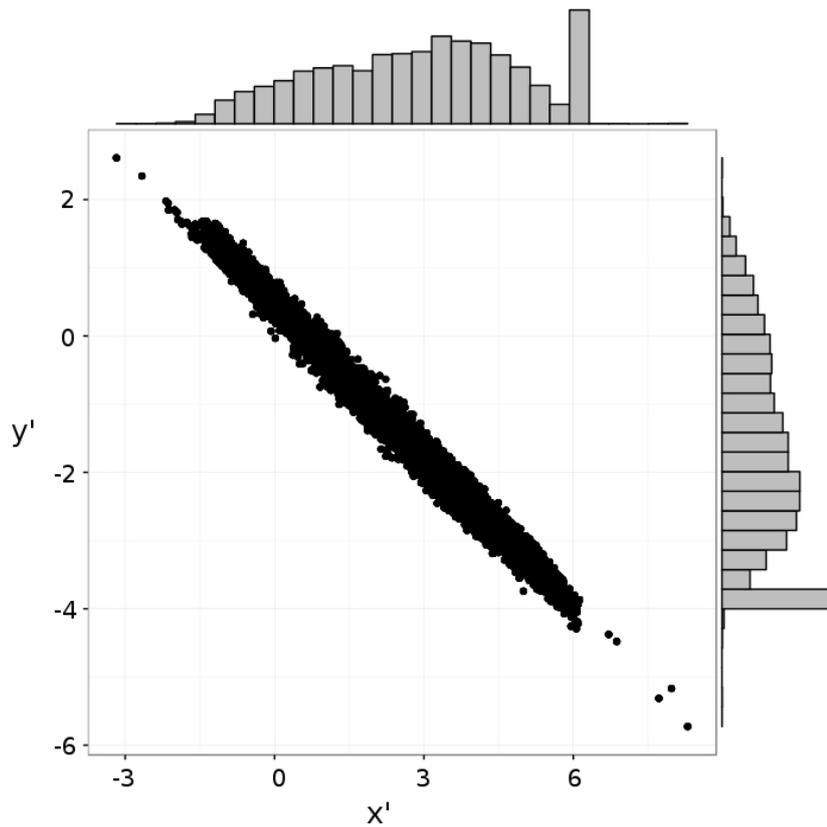


Figure 6.27: Scatter plot of the both transformed Weibull parameters, x' (x-axis) and y' (y-axis), for the period 2005-2014 for the station of Essen obtained the estimation model DS3 (bias-corrected estimation using KO, RH, and fronts). Histograms are depicted for each of the estimated transformed parameters are depicted at the boundary of the scatter plot.

function $x'' = \ln(x' + c)$, if the observation is more left skewed than the observation and $x'' = -\ln(-x' - c)$ otherwise. This procedure conserves the order of the values, and changes the distance amongst different points (and of course mean and variance). For dominantly linear relationship (estimation close to observation), the added constant becomes very large, leading to numerical problems with the resolution. Afterwards, the mean is adjusted by setting the estimated average to the stations average. Additionally, each value at the station is corrected, so that the standard deviation of the estimation for the station is the same as the station's standard deviation from the rainfall distribution parameters obtained from Radolan. The formula to calculate the bias-corrected value (X_{biasc}) is thus:

$$X_{biasc} = Mean(X_{clim}) + \frac{Std(X_{clim})}{Std(\ln(X'_{pred} + c))} \cdot \left(\ln(X'_{pred} + c) - Mean(\ln(X'_{pred} + c)) \right)$$

This procedure adjusts the climatology of the estimated rainfall parameters to fit the observed climatological properties. This additional correction to the output of the model DS2 defines in total the new model DS3. The bias-correction is performed to the complete climatology. The result is depicted in Figure 6.27 and agrees not much better with the scatter plot obtained from using Radolan data (see Figure 6.26a).

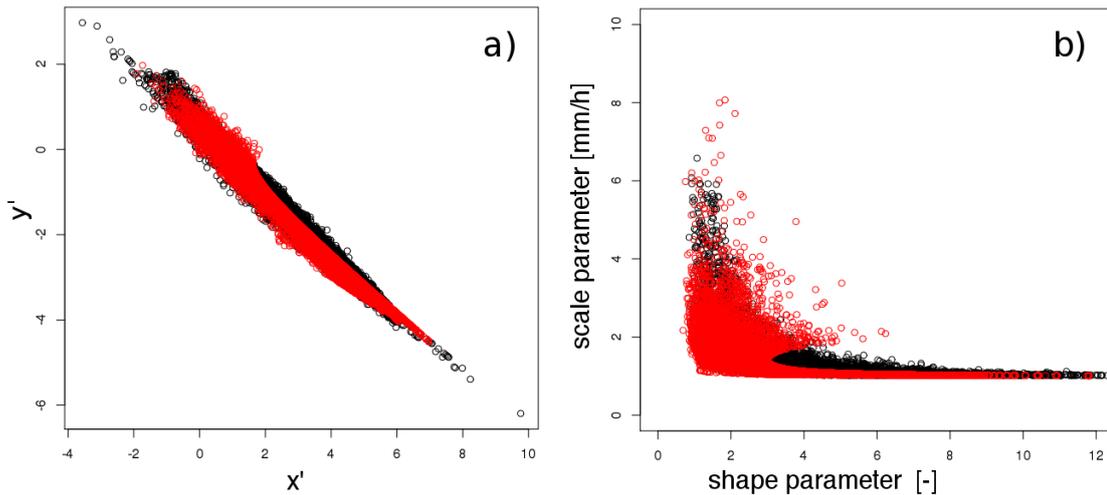


Figure 6.28: Estimated rainfall distribution parameters (black dots, based on DS3) with those obtained from Radolan (red dots) for the station of Essen for (a) the transformed Weibull parameters and (b) the original Weibull parameters.

As an example, the bias-corrected estimation and the Radolan based transformed Weibull parameters for the station of Essen are shown in Figure 6.28. It shows that the estimated (black) and the Radolan based result (red) match well in terms of the covered area. In both, the Radolan based and the estimated data, there is the defined border to events which cannot occur owing to the data resolution (for specific fixed but unknown low y' the values of x' show a maximum). Consider for example the case with zero precipitation at all Radolan grid points, but few above, then one additional value above zero increases not only the field average but also the variance. This explains why for weak rainfall events certain combinations of the two Weibull parameters are not possible.

As a result of the bias-correction, the deviation between estimation and observation in the mean scale parameter averaged over Germany is reduced from 1.6 (before the bias-correction) to 0.6 for the analysed events (naturally, the bias is zero for the calibration period after correction). However, the shown results in tables 6.13 and 6.14 are again limited to results, which are characterized by an estimated shape parameter lower than 9. The usage of this subset of data explains that there are remaining differences in

mean and variance between the computed and the observed parameters in spite of the bias-correction. The bias-correction altogether shows a good improvement of the spatial rainfall distribution (compare tables 6.13/6.14 and 6.11/6.12).

	all		scale (obs) >1.18	
	ref.	val.	ref.	val.
mean x' (obs)	3.515	3.998	1.372	1.427
mean x' (est)	3.067	3.043	2.245	2.286
var x' (obs)	3.391	4.146	0.794	0.764
var x' (est)	2.954	2.880	2.903	2.840
correlation x'	0.372	0.329	0.157	0.177
mean y' (obs)	-2.171	-2.518	-0.617	-0.656
mean y' (est)	-1.827	-1.814	-1.245	-1.285
var y' (est)	1.738	2.108	0.470	0.464
var y' (est)	1.614	1.570	1.604	1.545
correlation y'	0.362	0.316	0.196	0.203

Table 6.13: Different statistic measures of the transformed rainfall distribution parameters in the observations ('obs') and the corresponding estimations from DS3 ('est') **at the weather station positions** for estimated shape parameter below 9 and time steps with a Radolan observation (for reference period 2005-2010, and validation period 2011-2014)(for reference period 2005-2010, and validation period 2011-2014)).

	all		scale (obs) >1.18	
	ref.	val.	ref.	val.
mean shape (obs)	4.443	5.244	2.241	2.268
mean shape (est)	3.877	3.832	3.053	3.065
var shape (obs)	5.443	8.792	0.489	0.470
var shape (est)	3.237	3.230	2.389	2.442
correlation shape	0.352	0.314	0.190	0.222
mean scale (obs)	1.284	1.244	1.696	1.672
mean scale (est)	1.386	1.380	1.615	1.581
var scale (obs)	0.216	0.198	0.347	0.355
var scale (est)	0.511	0.438	0.896	0.702
correlation scale	0.245	0.226	0.147	0.163

Table 6.14: Different statistic measures of the rainfall distribution parameters in the observations ('obs') and the corresponding estimations from DS3 ('est') **at the weather station positions** for estimated shape parameter below 9 and time steps with a Radolan observation (for reference period 2005-2010, and validation period 2011-2014).

Notably, the performance of the procedure for estimating the two Weibull parameters is much better for the shape parameter than for the scale parameter. For the shape parameter, the correlation between bias-corrected estimation and observations is 0.19 for

events with scale parameter >1.18 and 0.35 for all events. The respective correlations for scale parameter are 0.15 and 0.25 . Thus, the method shows some more skill for the shape parameter. This indicates that there is still room for further improvement, although not based on this simple parameter based approach. Therefore, the development stage 3 (DS3) of the method is validated in more detail in the following section and further improved afterwards.

Validation of application to gridded dataset

In a next step, the development stage of the model (DS3) (which uses KO, RH, front information and the bias-correction) is validated for all grid point positions in order to ensure that the model does not decrease in performance for locations other than the station positions, because the relationship has been developed based on highly resolved rainfall around station locations. The performance of DS3 for the grid points (see table 6.15) is very similar to its performance for the station positions (see table 6.13). Again, the correlation for the shape parameter reveals a better result than for the scale parameter, see table 6.16. This indicates that particularly the estimation of the scale parameter can be improved further. The estimation of the shape parameter is therefore defined as the final estimation, while the estimation of the scale parameter is attempted to be improved in the next section.

	all		scale (obs) >1.18	
	ref.	val.	ref.	val.
mean x' (obs)	3.368	3.834	1.325	1.402
mean x' (est)	3.425	3.374	2.537	2.497
var x' (obs)	3.594	4.307	0.832	0.785
var x' (est)	3.051	2.938	3.076	2.975
correlation x'	0.440	0.421	0.197	0.190
mean y' (obs)	-2.063	-2.392	-0.596	-0.633
mean y' (est)	-2.127	-2.083	-1.487	-1.444
var y' (obs)	1.823	2.190	0.493	0.480
var y' (est)	1.599	1.541	1.676	1.620
correlation y'	0.447	0.429	0.232	0.217

Table 6.15: Different statistic measures of the transformed rainfall distribution parameters in the observations ‘obs’) and the corresponding estimations from DS3 (‘est’) **at the centres of ERA-Interim grid points’ positions** for events with estimated shape parameter below 9 (for reference period 2005-2010, and validation period 2011-2014).

	all		scale (obs) >1.18	
	ref.	val.	ref.	val.
mean shape (obs)	5.098	2.223	2.192	2.269
mean shape (est)	4.106	3.228	3.226	3.232
var shape (obs)	8.888	0.527	0.533	0.515
var shape (est)	3.762	2.667	2.665	2.671
correlation shape	0.386	0.364	0.186	0.211
mean scale (obs)	1.318	1.274	1.722	1.694
mean scale (est)	1.316	1.317	1.536	1.542
var scale (obs)	0.293	0.238	0.484	0.406
var scale (est)	0.771	0.596	1.280	1.161
correlation scale	0.209	0.228	0.135	0.140

Table 6.16: Different statistic measures of the rainfall distribution parameters in the observations ‘obs’ and the corresponding estimations from DS3 (‘est’) **at the centres of ERA-Interim grid points’ positions** for events with estimated shape parameter below 9 (for reference period 2005-2010, and validation period 2011-2014).

6.4.3 Additional improvement by inclusion of model rainfall

The estimation method (DS3, based on KO, RH, fronts, and the bias-correction) showed an higher correlation with the observations for the shape parameter compared to the scale parameter. The shape parameter is of more importance to identify convective events (see chapter 6.2). Nevertheless, the rainfall intensity is of importance to identify only events of a certain relevance in terms of precipitation. The performance of the estimation of the scale parameter is not yet satisfying and can be improved using the ERA-Interim rainfall. The average rainfall is associated to the two Weibull parameters. Therefore, it is possible to replace the scale parameter by calculating it from the ERA-Interim rainfall representing the mean rainfall intensity and the previously estimated shape parameter:

$$y_{RR} = \frac{RR \cdot 1h/mm + 1}{\Gamma(1 + 1/x_{shape})} \quad (6.5)$$

with RR being the ERA-Interim precipitation rate for the grid point, x_{shape} being the estimated shape parameter, and y_{RR} being the new scale parameter. The shift of the rainfall amount for 1 in the numerator ($RR \cdot 1h/mm + 1$) is related to the procedure of the distribution fit, which has been explained in section 5.1.3.

This newly estimation DS4 (based on RH, KO, fronts, bias-correction and rainfall from the ERA-Interim forecast) reveals an increased performance for the estimation of the scale parameter (see table 6.18 compared to table 6.16 and for completeness table 6.17

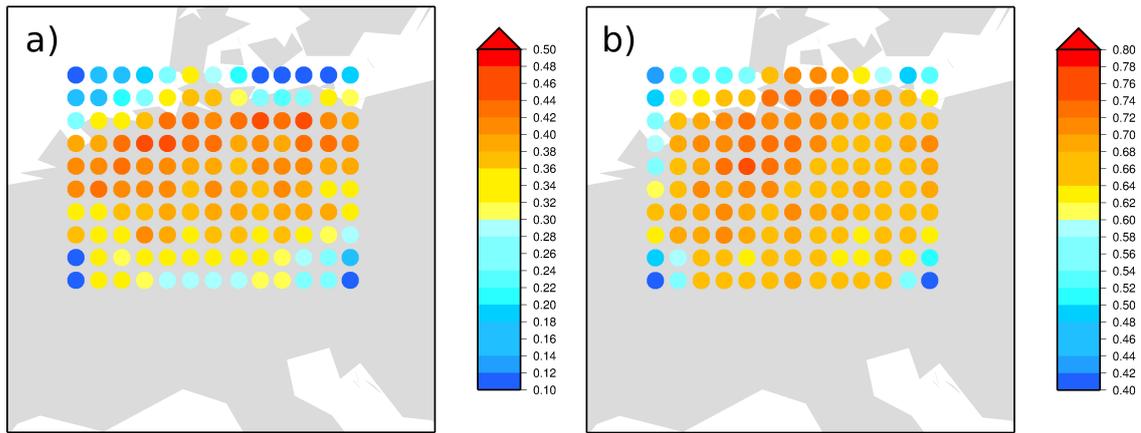


Figure 6.29: Correlation between observations and estimation based on RH, KO and ERA-Interim rainfall prediction for (a) shape and (b) scale parameter. The points cover the same area as in Figure 6.25.

	all		scale (obs) >1.18	
	ref.	val.	ref.	val.
mean shape (obs)	4.328	5.098	2.192	2.269
mean shape (est)	4.142	4.106	3.226	3.232
var shape (obs)	5.703	8.888	0.533	0.515
var shape (est)	3.843	3.762	2.665	2.671
correlation shape	0.386	0.364	0.186	0.211
mean scale (obs)	1.318	1.274	1.722	1.694
mean scale (est)	1.427	1.436	1.831	1.900
var scale (est)	0.293	0.238	0.484	0.406
var scale (obs)	0.650	0.709	1.265	1.486
correlation scale	0.554	0.579	0.450	0.483

Table 6.17: Different statistic measures of the rainfall distribution parameters in the observations ‘obs’ and the corresponding estimations from DS4 (‘est’) for the validation period for events with estimated shape parameter below 9.

to table 6.15). The correlation between the Radolan based scale parameter and the new scale parameter (from formula 6.5) is in the order of 0.56, which shows an improvement to the correlation of the scale parameter without the ERA-Interim rainfall. Analysing the performance for specific grid points, a correlation of 0.7 between the Radolan-based scale parameter and its estimation using the final model DS4 is found with the highest correlation values in the North(west) of Germany (see Figure 6.29b). There are boundary effects where the Radolan data does not cover the entire circle around the grid point, but also over the ocean lower correlations are found. This indicates that the actual rainfall distribution cannot be estimated well from only a small part of the circle covered by Radolan information. Correlations of the estimated shape parameter (using DS4, identical

with DS3 for this parameter) have a correlation with the shape parameter of Radolan of about 0.4 (see also Figure 6.29a). They are highest in the North of Germany.

	all		scale (obs) >1.18	
	ref.	val.	ref.	val.
mean x' (obs)	3.368	3.833	1.325	1.402
mean x' (est)	2.943	2.928	1.904	1.829
var x' (obs)	3.593	4.307	0.831	0.785
var x' (est)	1.990	1.953	1.845	1.837
correlation x'	0.645	0.640	0.444	0.437
mean y' (obs)	-2.063	-2.391	-0.596	-0.633
mean y' (est)	-1.642	-1.634	-0.851	-0.774
var y' (obs)	1.823	2.191	0.493	0.480
var y' (est)	1.199	1.205	1.354	1.376
correlation y'	0.653	0.650	0.498	0.481

Table 6.18: Different statistic measures of the rainfall distribution parameters in the observations ('obs') and the corresponding estimations from DS4 ('est') for the validation period for events with estimated shape parameter below 9.

Summarizing, it can be concluded that using the ERA-Interim precipitation for computing the scale parameter has a positive effect on the model performance. Summarizing it can be concluded that all previously introduced steps had a positive effect on the performance. The final method contains all these steps and is recapitulated in the next section.

6.4.4 Summary of improvements

The developed model **DS1**, which estimates the rainfall distribution parameters from KO and RH, was further improved by including additional parameters:

DS2: like DS1 but additionally uses information about relationship between rainfall distribution parameters and stability parameters for no/warm/cold/both fronts

DS3: like DS2 but additionally includes information about the local properties of rainfall distribution parameters

DS4: uses the shape parameter obtained from DS3 and the precipitation from ERA-Interim to replace the scale parameter.

The work-flow of the final method is schematically depicted in Figure 6.30.

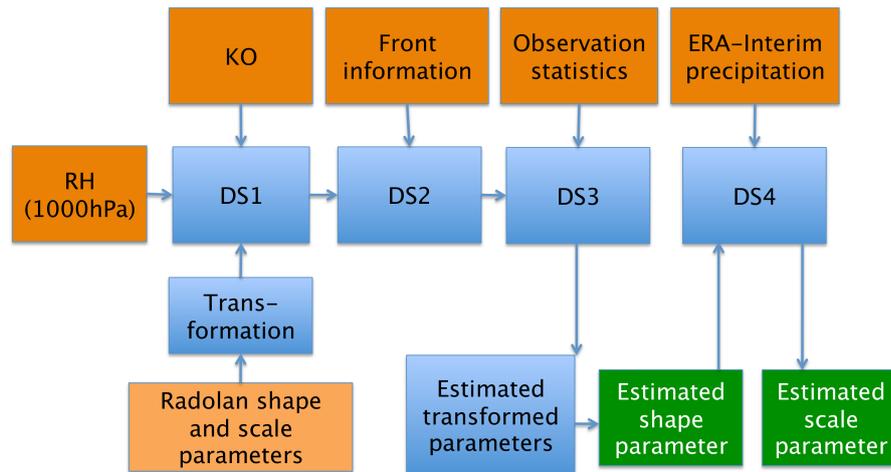


Figure 6.30: Scheme of developed model and used variables to obtain the final estimated rainfall distribution parameters. Input variables are indicated by an orange colour. Methods are indicated by an blue colour. The final output variables are indicated by green rectangles.

6.5 Model combination and validation

Relationships between the large scale atmospheric state and rainfall parameters on one hand, and on the other hand between observed rainfall parameters and intense convective events have been determined. The former are used to estimate the rainfall parameters with the final model DS4 (see previous section 6.4.3). The latter determines if an estimated convective extreme event is present as the Radolan based thresholds are exceeded (see section 6.2.3). The performance of the combined procedure, i.e. estimating the occurrence of intense convective events from the large-scale parameters, is evaluated in the following.

Extreme convection (identified from lightning occurrence) is assigned to a threshold of 2 for the shape parameter (see yellow curve in Figure 6.13). The skill of this estimation as measured using the HSS is 0.033, which is much lower than the HSS found when using the Radolan-based shape parameter. The same is valid for the threshold of 1.5, which is based on the analogous application to the grid-point positions. Also, for this threshold, the corresponding HSS is 0.035 and thus much lower than the skill of 0.39 found when using the Radolan-based shape parameter (red curve in Figure 6.31). As both threshold values may not be the optimum choice, the link is tested for a range of threshold values, computing the skill for each of them (yellow curve in Figure 6.31). It is found that indeed the shape parameter of 2 is not the optimum choice. Instead, for the large-scale parameter dependent approach the skill is low largely independent of the threshold. The

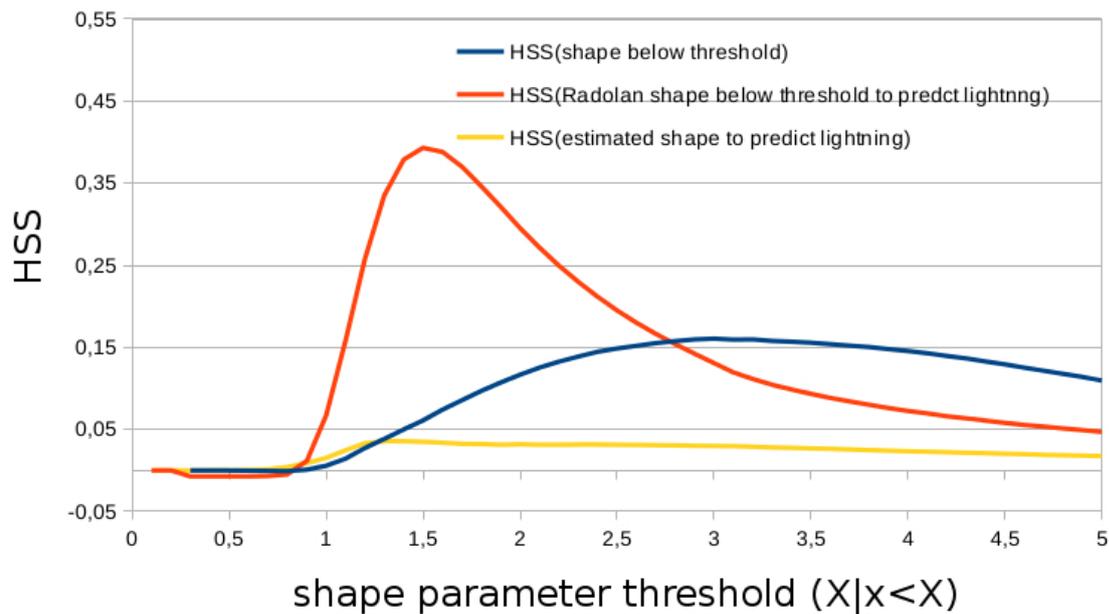


Figure 6.31: HSS for the validation period for (red) undercutting a given threshold and (yellow) the prediction of a convective event by undercutting an event based on the shape parameter from Radolan data. The performance of DS4 and Radolan to undercut the same threshold (x-axis) is depicted in blue.

HSS values are above 0.025 for a shape parameter between 1.2 and 3.7. As this skill is rather low, a different approach is chosen to define a suitable threshold for the shape parameter. The skill of the procedure in terms of undercutting a specific threshold of the shape parameter is computed (blue curve in Figure 6.31). Please note that this curve has been obtained from ERA-Interim grid points with 95% of the 90km radius covered with radar data, which explains a higher HSS of 0.39 compared to 0.32 in Figure 6.13. High skills are found for a threshold in the range from 2.5 to 4. A notable predictive skill for lightning events (representing a subset of convective events) is found for lower thresholds indicating a convective character of the rainfall. Therefore, a threshold of 2.5 is chosen, which is also acceptable in terms of the convective character of precipitation.

In order to ensure a relevant precipitation amount, a minimum scale parameter of 1.18 for a ERA-Interim grid cell is chosen to define events, which corresponds on average to precipitation amount of 1 mm in 6 hours. This additional threshold impacts the HSS values. In particular, for a shape parameter threshold of 2.5, the HSS for increases from 0.15 (compare blue line in Figure 6.31) to 0.28 (no figure shown).

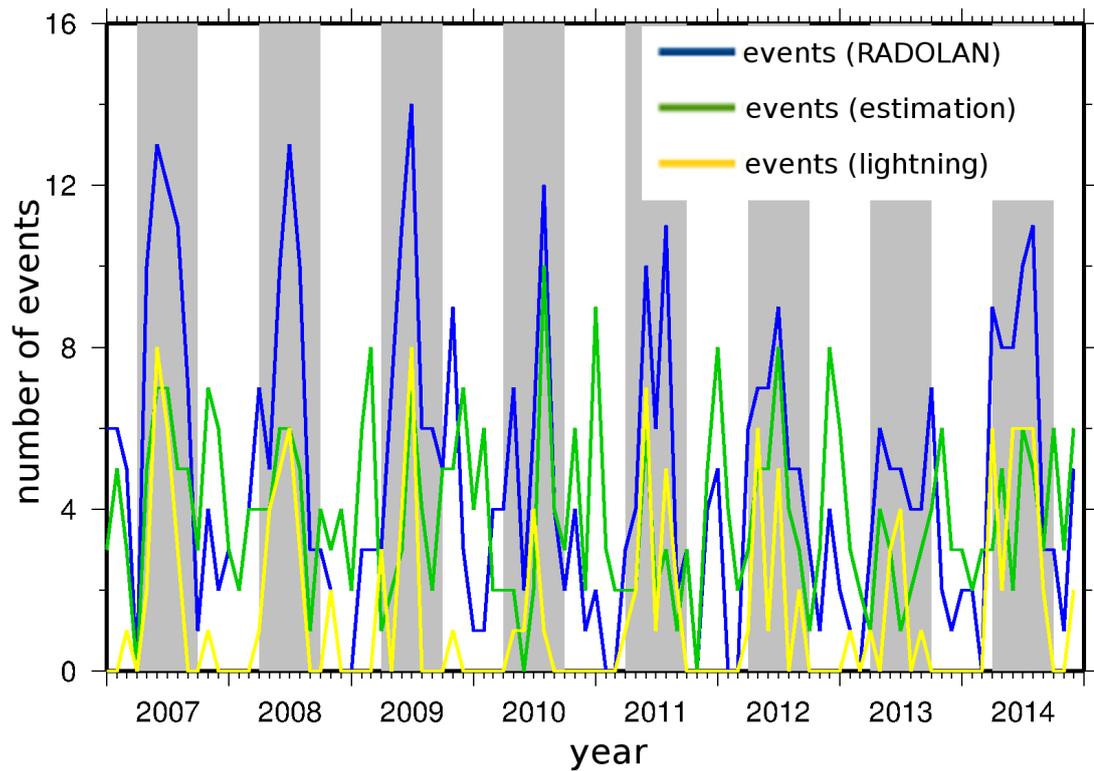


Figure 6.32: Time-series of monthly number of (lightning) events for one grid point (with centre at 8.25E, 51.75N): observed (yellow); obtained based on Radolan rainfall distribution (blue); estimated events based on rainfall estimation (green). The summer half year is marked in grey. For details see main text.

Time-series of the monthly frequencies of estimated events from the large-scale parameters using the above thresholds to estimate events (DS4), estimated events from the Radolan data using the same thresholds, and measured lightning events are exemplarily shown for the grid-point (8.25°E, 51.75°N) in Figure 6.32.

The three time-series are well correlated for summer (AMJJAS). The respective correlations are: DS4-Radolan: 0.77; DS4-lightning: 0.50; Radolan-lightning: 0.78. As expected, the estimation based on the ‘observed’ rainfall distribution parameters (blue) shows a higher correlation with the number of lightning events (yellow) than the estimated intense events based on the estimated rainfall parameters from the final method DS4 (in green). The annual cycle with its maxima in summer is well reflected by both estimations. Also, time of peaks in the occurrence of the occurrences of convective extreme events in summer are generally well captured. The estimated number of events based on the Radolan data (blue) is higher than the estimation based on DS4 (green). For winter, an overestimation in the number of events is visible for the estimation based on DS4. This

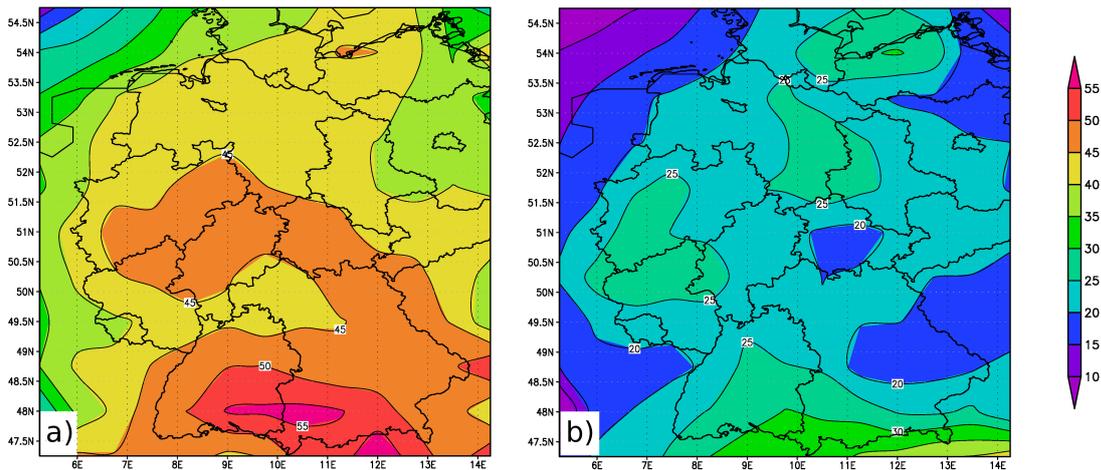


Figure 6.33: Estimated annual frequency of events based on (a) Radolan data and (b) the ERA-Interim data using DS4.

can be understood from the fact that the model was trained for summer. It is noteworthy that the number of lightning observations (yellow) shows a minimum in winter. In contrast, events estimated from Radolan-based rainfall parameters and from DS4 reveal a second event maximum during winter. It may be possible that the atmospheric properties during winter are not favouring lightnings even if observed rainfall properties do.

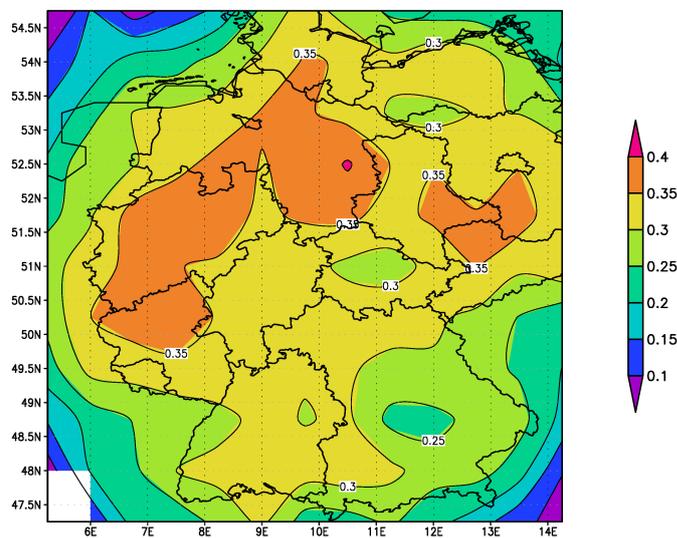


Figure 6.34: HSS comparing the estimated event/no-event situations using estimations based on ERA-Interim (DS4) and rainfall distribution parameters obtained from Radolan (observations).

Figure 6.33b shows the estimated convective extreme event frequency per summer. With respect to the spatial distribution of event frequencies, the comparison with the frequency

of events from Radolan (Figure 6.33a) reveals that there is an underestimation all over the investigation area, in particular in the South-East of Germany (north of the Alps) and a weaker underestimation in Northern and Western Germany. The performance of the event estimations is measured with the HSS and is depicted in Figure 6.34. It ranges between 0.2 and 0.4 with the highest skill in the North-West of Germany (with maximum value of 0.4) and shows a small decrease towards the South of Germany. A strong decrease in the HSS at the outer grid points of the analysed area is identified and indicates a clear reduction in the performance of the prediction. For these grid points, the Radolan data does not cover the entire area. Consequently, calculating the rainfall parameters for these grid points is problematic and thus the reason for this decreased performance.

6.6 Ranking of estimated events

Based on a grid point basis, for Germany, events are identified by application of DS4 to the ERA-Interim data. The combination of the corresponding information from the time-series can be used to obtain an estimated footprint of an event. For ranking dates of estimated events (i.e., shape parameter <2.5 and scale parameter >1.18) in Germany based on their spatially integrated intensity, the 6-hourly rainfall intensity of grid points experiencing an (ERA-Interim) estimated event are integrated. Only the maximum out of the 12 and 18UTC rainfall intensity is used. The corresponding ranking for the whole ERA-Interim period (1979-2014) is shown in table 6.19. Additionally, this table presents the related area of events as percentage of grid points over Germany and the average rainfall intensity of event grid points. In addition, circulation weather types (CWTs, e.g. *Moemken et al. (2016)*) have been calculated for the respective days. For the most intense events, the cyclonic weather type (marked as c) is most frequent, with wind directions predominantly ranging from south-westerly to westerly (numbers: 15-16 in table 6.19). This corroborates the results of previous studies, which found a predominance of south-westerly wind directions for convective events in Germany (e.g. *Mohr et al. (2015b)*).

The ranking of extreme events (table 6.19) can also be compared to a ranking with respect to the occurrence of extreme gusts. In an earlier study by *Seregina et al. (2014)*, the station-wise most extreme gusts at German stations in the period 2001-2011 were ranked. 8 out of the top 28 events occurred in summer (marked as 'summer*' in table 6.20). As visible, most of these wind gust events have also a high ranking (top 10%)

in the list of convective events (see last column in table 6.20). This indicates that there might be a relationship between intense gusts on the very local scale and the convective extreme events on the regional scale.

Rank	Date	event sum [mm]	average in- tensity [mm]	% event grid points	CWT
1	19970718	415.56	3.774	78.1	c
2	20140721	393.11	4.784	59.7	c
3	20140708	385.43	5.018	54.2	c
4	20020717	375.56	5.636	46.4	11c
5	19960828	374.61	4.708	54.7	c
6	19890827	370.97	4.282	60.9	c
7	19840715	369.60	3.828	69.2	16
8	19860723	353.29	5.416	46.1	15c
9	19870818	351.33	3.333	75.0	16c
10	20100530	329.24	3.603	64.2	15c
11	20110713	329.17	4.662	50.9	c
12	20110622	328.72	5.735	40.1	16
13	19800720	319.84	4.857	46.3	16
14	20070821	319.86	4.068	54.7	c
15	20060828	312.26	3.639	60.9	17
16	19860812	310.29	3.414	65.6	c
17	20080822	308.08	4.787	45.2	16a
18	19980607	306.62	3.925	55.0	c
19	19930521	305.65	5.552	38.8	c
20	19900930	305.68	3.602	59.9	15

Table 6.19: Ranking of estimated convective events for Germany. The ranks are given in the first column. The second column provides the associated date. The third column gives the integrated intensity, which has been used for the ranking. The fourth and fifth columns provide information about the mean local intensity and the spatial extension of the events. The last column indicates the associated circulation weather type (c represents cyclonal, a represents anticyclonal wind; the numbers indicate the associated main wind direction ranging from 11: north-eastern to 18: northern wind direction)

Another issue is the analysis of the footprint for individual estimated events. The footprints can be validated and depict the performance of the developed model for single extreme events. The four summer events, which have the highest return periods according to *Seregina et al. (2014)*, are associated with a spatially integrated convective rainfall over Germany ranking under the top 10% of the time period considered. However, the two rankings are not comparable for individual events, since the ranking in *Seregina et al. (2014)* is based on the return periods of wind gusts from station measurements, while the ranking presented here is based on Germany-wide integrated convective rainfall. Nevertheless, a potential relation between both rankings is evident, but requires more detailed investigations.

ΔX_n ($y = au^b$)	max.return period (years)	wind gust speed ms^{-1}	date (dd.mm.yyyy)	event names	summer rank
(7.32)	$1.7 \cdot 10^4$	59	08.06.2003	summer*)	45
(4.88)	$1.5 \cdot 10^3$	53	29.07.2005	summer*)	38
(3.94)	566	36	18.01.2007	Kyrill)	
2.26	105	65	25.11.2002	Andrea (Foehn storm)	
1.96	78	42	01.03.2008	Emma	
1.78	65	35	28.02.2010	Xynthia	
1.63	46	33	12.02.2005	Ulf	
1.47	48	32	26.02.2002	Anna	
1.40	35	28	27.10.2002	Jeanett	
1.23	38	36	08.07.2004	summer*)	22
1.02	31	35	28.01.2002	Jennifer	
0.98	29	30	02.01.2003	Calvann	
0.65	21	41	08.01.2005	Erwin	
0.52	18	34	20.03.2004	Oralie	
0.41	16	29	06.07.2001	summer*)	80
0.35	16	35	20.05.2006	Gertrud	
0.33	15	32	16.12.2005	Dorian	
0.26	14	33	12.01.2004	Gerda	
0.22	14	31	31.01.2002	Lydia	
0.15	13	37	31.12.2006	Karla	
0.09	12	32	22.02.2008	Annette	
-0.09	10	22	16.11.2002	Uschi (Foehn storm)	
-0.11	9	29	18.06.2001	summer*)	106
-0.18	9	35	18.11.2004	Pia	
-0.41	7	30	12.08.2004	summer*)	210
-0.58	6	32	31.07.2001	summer*)	543
-0.66	6	34	03.08.2001	summer*)	297
-0.69	6	31	14.12.2003	Fritz	

Table 6.20: Maximum return periods based on measurements. Brackets indicate values with high uncertainty due to deviation from the Gumbel-distribution. Stars indicate summer events (source of table and caption: *Seregina et al. (2014)*). An additional column has been inserted, indicating the summer event ranking of the data for the period 2001-2011.)

In summary, a reliable ranking was obtained. The moderate agreement with locally obtained gusts indicates that the ranking is useful and performs well. Convective rainfall intensities and the extension of corresponding areas contribute to the intensities of events.

6.7 Case studies

This section discusses two cases, namely the events on 21 July 2014 and 21 June 2007 that are high ranked and within the period of Radolan observations. The first event is the most intense in the period covered by Radolan data and is the second-strongest in table 6.19. The second event ranks 55th and has been randomly chosen.

6.7.1 Case of 21 July 2014

The rainfall event on 21 July 2014 was estimated as the second most intense convective summer rainfall event in the ERA-Interim period according to the developed ranking (cf. table 6.19). This day was characterized by low pressure over the Alps and a cyclone core ('Paula') centered over the Czech Republic. Associated with this cyclone, a front was extending over Germany towards the North Sea (cf. Figure 6.35). In addition, along this frontal structure, low pressure disturbances and nearby convergence lines evolved later on this day (not shown, source: Berliner Wetterkarte, 2014). This situation led to strong convective precipitation across Germany.

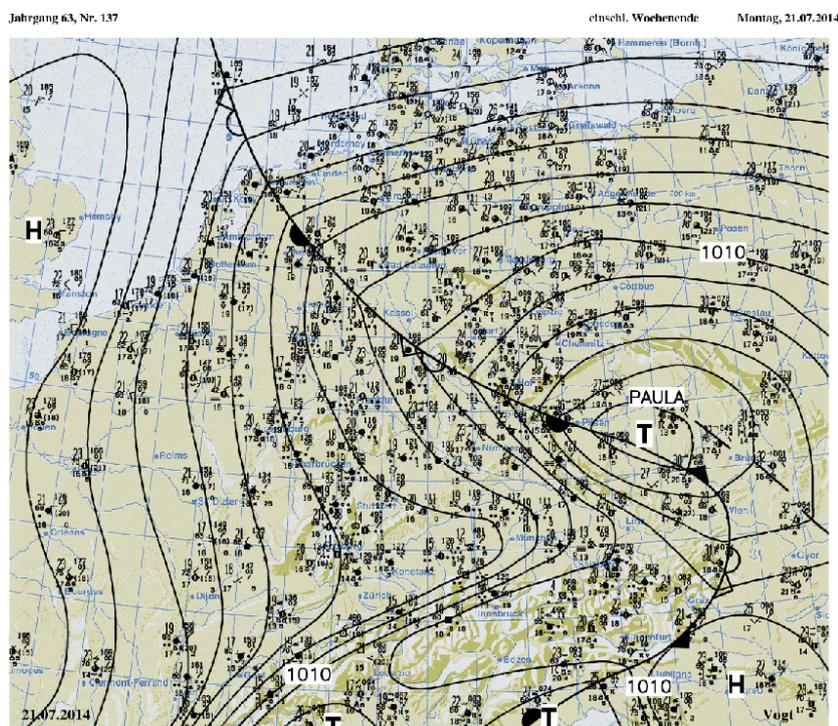


Figure 6.35: Surface weather chart of 21 July 2014 06 CET for Central Europe (source: Berliner Wetterkarte).

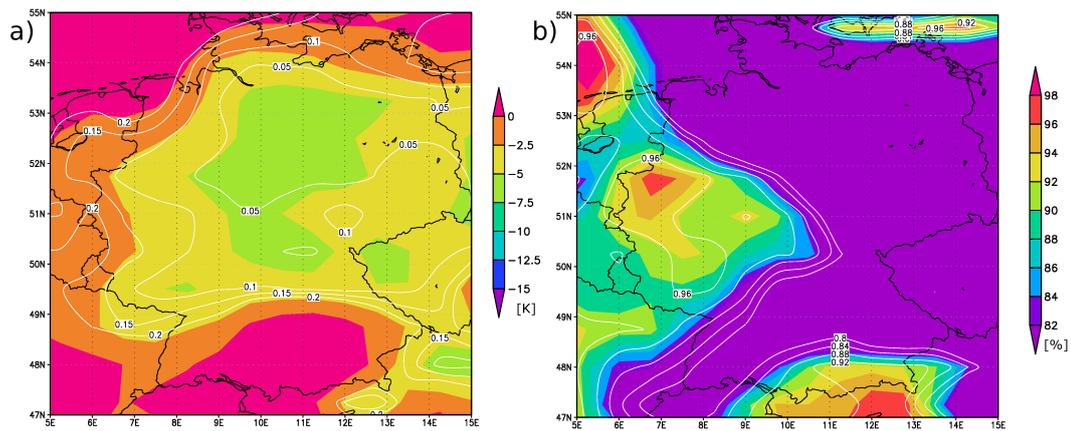


Figure 6.36: (a) KO index and (b) 1000hPa relative humidity for 2014-07-21 18 UTC as obtained from the ERA-Interim dataset. The shaded areas indicate the absolute values, contours indicate the local percentiles with respect to the period from 1979 to 2014.

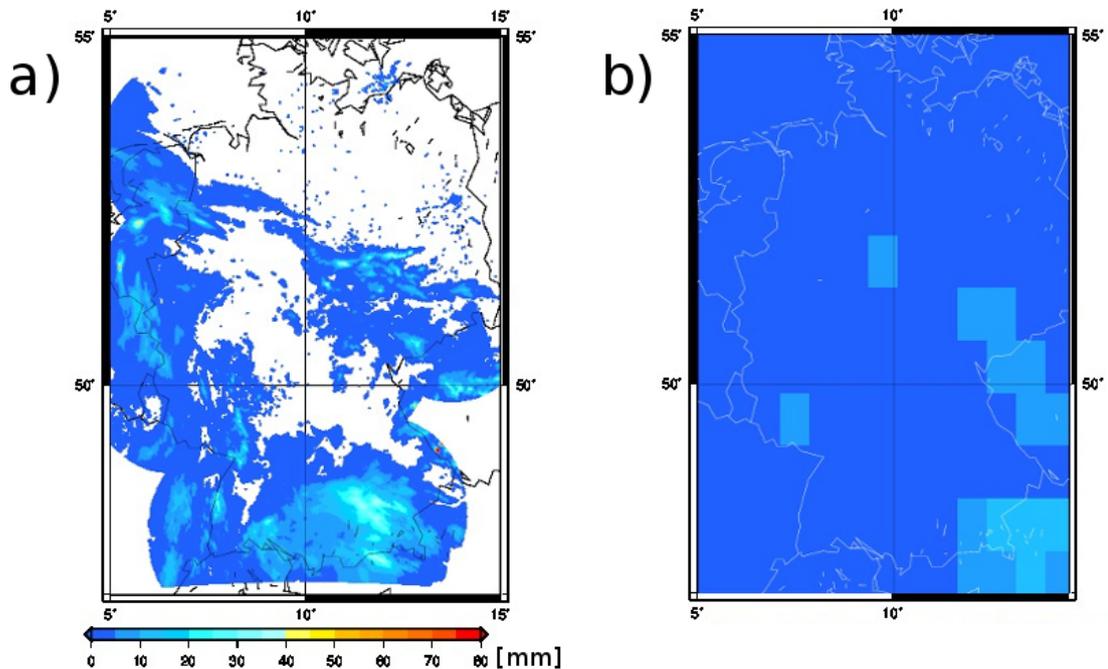


Figure 6.37: Precipitation amount of 2014-07-21 12-18 UTC for (a) the Radolan data and (b) for the ERA-Interim precipitation forecast.

This day was part of a period of widespread heat and high thunderstorm potential (19.7.2014-22.7.2014 according to KIT, www.wettergefahren-fruehwarnung.de). With this respect, high instability values are found over large parts of Germany for this day (i.e. a low KO index, Figure 6.36a) and high 1000hPa relative humidity values are found over western and southern Germany (cf. Figure 6.36b). Warnings of thunderstorm and high-intensity rainfall were issued for large parts of Germany for this day (e.g.

www.wettergefahren-fruehwarnung.de). In south-western Germany, where many thunderstorms occurred, an average temperature drop of 10°C was observed (source: Berliner Wetterkarte). Highest rainfall amounts for the 24h-period from 21.7.2014 06UTC to 22.7.2014 06UTC were observed in several locations in the South of Germany (e.g. 94mm on the Feldberg (Baden-Württemberg), 88mm in Ottobeuren (Bavaria), 84.5mm in Marklkofen (Bavaria)). This shows some agreement with the rainfall pattern from Radolan (cf. Figure 6.37a) and ERA-Interim (cf. Figure 6.37b).

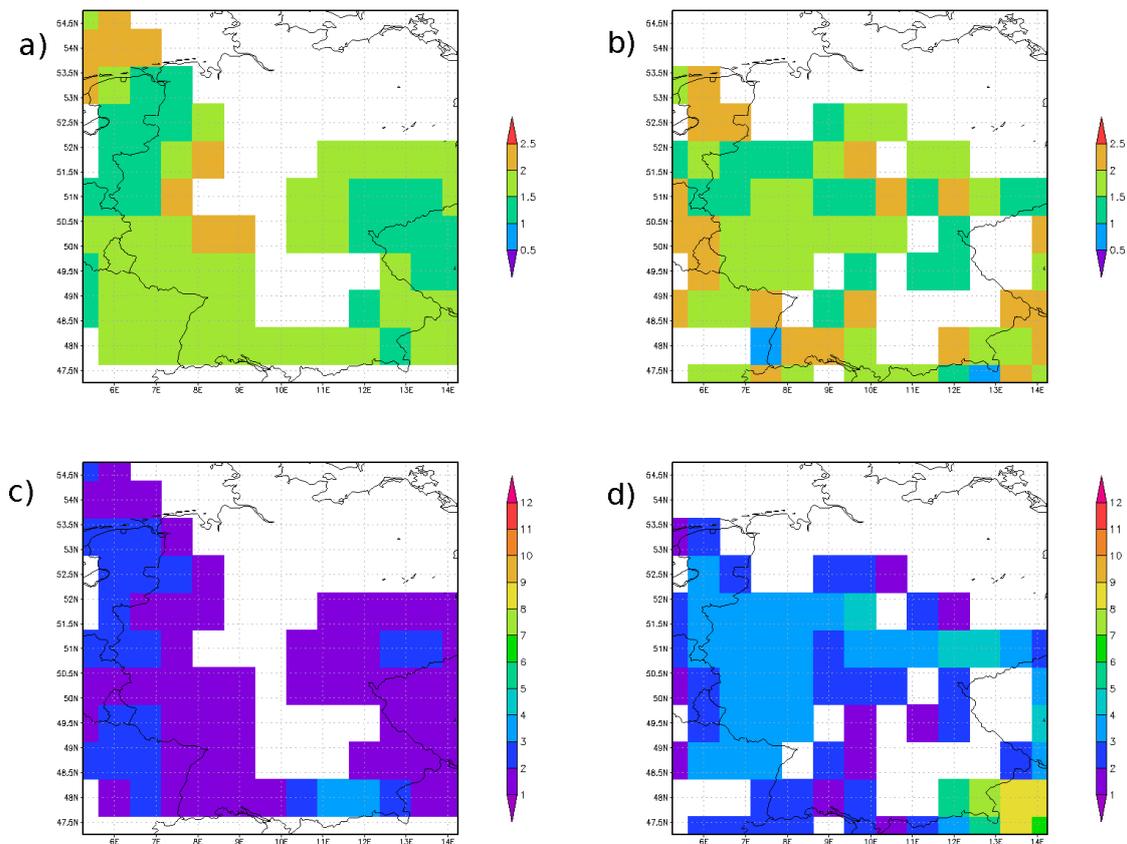


Figure 6.38: Weibull shape (a+b) and scale parameters (c+d) of event grid points as obtained from the Radolan data (a+c) and estimated from DS4 (b+d) at 2014-07-21 18UTC.

The locations of convective events at 21 July 2014 18UTC, based on the rainfall distribution parameters (shape and scale parameter; Figure 6.38), show a reasonable spatial agreement between those obtained from Radolan and those estimated by the developed method (DS4). In particular, in both datasets, the majority of grid points showing an event are in the south-western half of Germany (Figure 6.38). Note that the lowest KO-values occur east of the precipitation area at the western border of Germany (yellow

and green area in Figure 6.36a). The instability values in this region led together with the high 1000hPa relative humidity to an overestimation of the spatial extension of the event area in western Germany compared to the Radolan observations. Nevertheless, the largest scale parameter values (indicating the rainfall intensity) are found in the South-East of Germany near the Alps for Radolan observations and for the estimation from DS4 (see Figure 6.38c,d). Comparing the estimated shape and scale parameter for the event grid points and datasets, a better agreement between Radolan observations and the estimation from DS4 is found for the scale parameter. This is in line with the previously found higher correlation for the scale parameter compared to the shape parameter (see section 6.4.3). However, the number of grid points with estimated events (DS4) is smaller than for those estimated using Radolan data. This is in agreement with the previously found frequencies of estimated events using Radolan information and using DS4 (cf. Figure 6.33).

6.7.2 Case of 21 June 2007

The rainfall event on 21 June 2007 was estimated as the 55th most intense convective summer rainfall event in the ERA-Interim period. The weather situation for this day was dominated by two cyclones ‘Steven‘ and ‘Thies‘. Ahead of a trough associated to cyclone ‘Steven‘ (located over Ireland in Figure 6.39), subtropical warm air was transported to Germany. A cold front occlusion was located over Germany from North-East to South-West. At this occlusion, which was associated to the cyclone ‘Steven‘, wave disturbances developed and favoured the development of showers and thunderstorms. There were two centres of convective rainfall activity over Germany (Figure 6.40a): One was located over Brandenburg at the occlusion point of the fronts associated to ‘Steven‘, and the other one over western Germany related to the cyclone ‘Thies‘. The largest precipitation amount was observed at Frankfurt airport with a daily integrated precipitation amount of 46mm (according to Berliner Wetterkarte). This rainfall was embedded in a region of thunderstorms, which moved northward from the South-West of Germany. The first thunderstorms associated to this cluster occurred in the evening before the 21 June 2007. Under the influence of a short wave trough in the upper troposphere over France/Switzerland, a second cluster developed, which also moved northward.

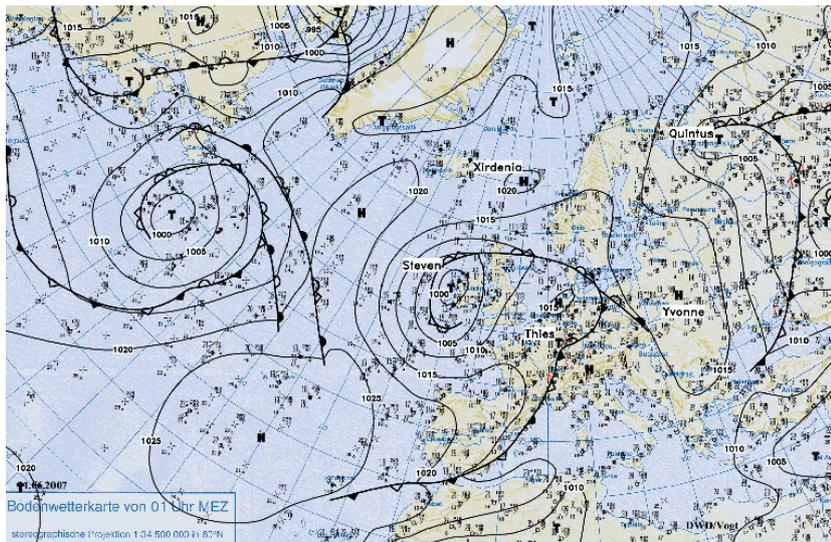


Figure 6.39: Surface weather chart of 21 June 2007 01 CET for the European/Atlantic region (source: Berliner Wetterkarte).

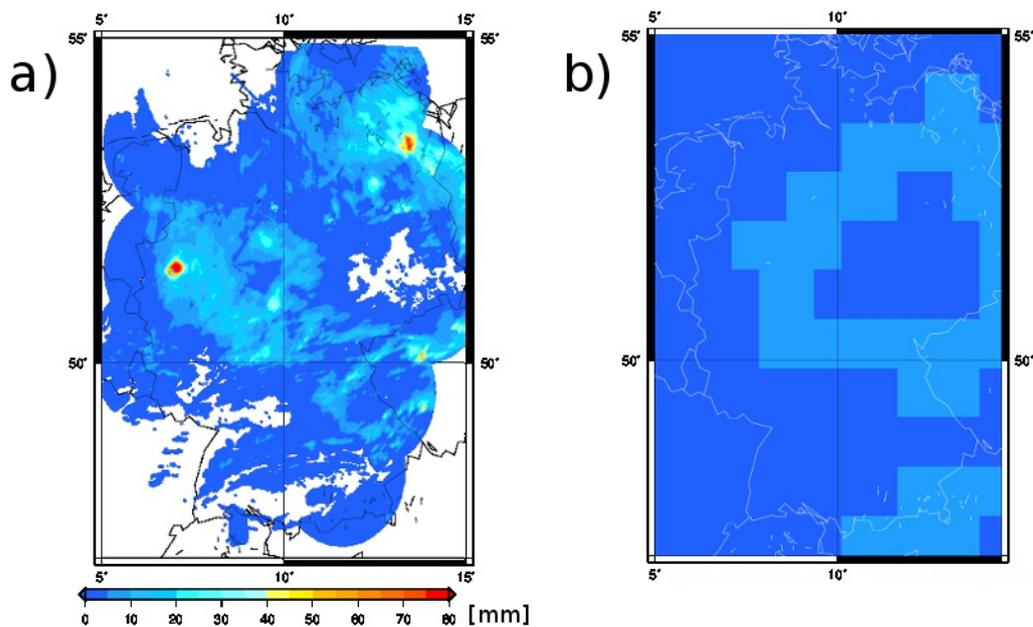


Figure 6.40: Precipitation amount of 2007-06-21 12-18 UTC for (a) the Radolan data and (b) for the ERA-Interim precipitation forecast.

For both rainfall datasets (ERA-Interim and Radolan observation, Figure 6.40a,b), a larger circular structure of enhanced rainfall is found over north-eastern Germany. However, the position of the inside area, which is characterized by a reduced precipitation amount, differs between observation and ERA-Interim. The corresponding KO index indicates stronger atmospheric instability over eastern and southern Germany (Figure 6.41a).

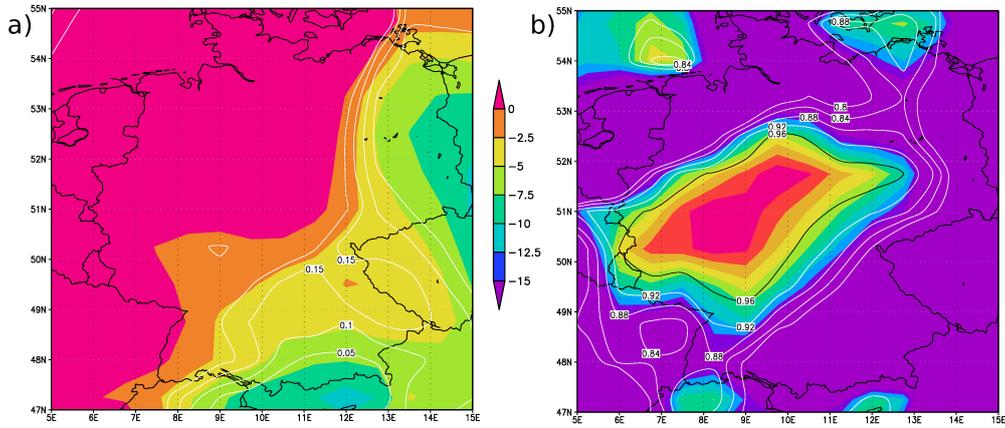


Figure 6.41: KO index (a) and 1000hPa relative humidity (b) for 2007-06-21 12 UTC as obtained from the ERA-Interim dataset. The shaded areas indicate the absolute values, contours indicate the local percentiles with respect to the period from 1979 to 2014.

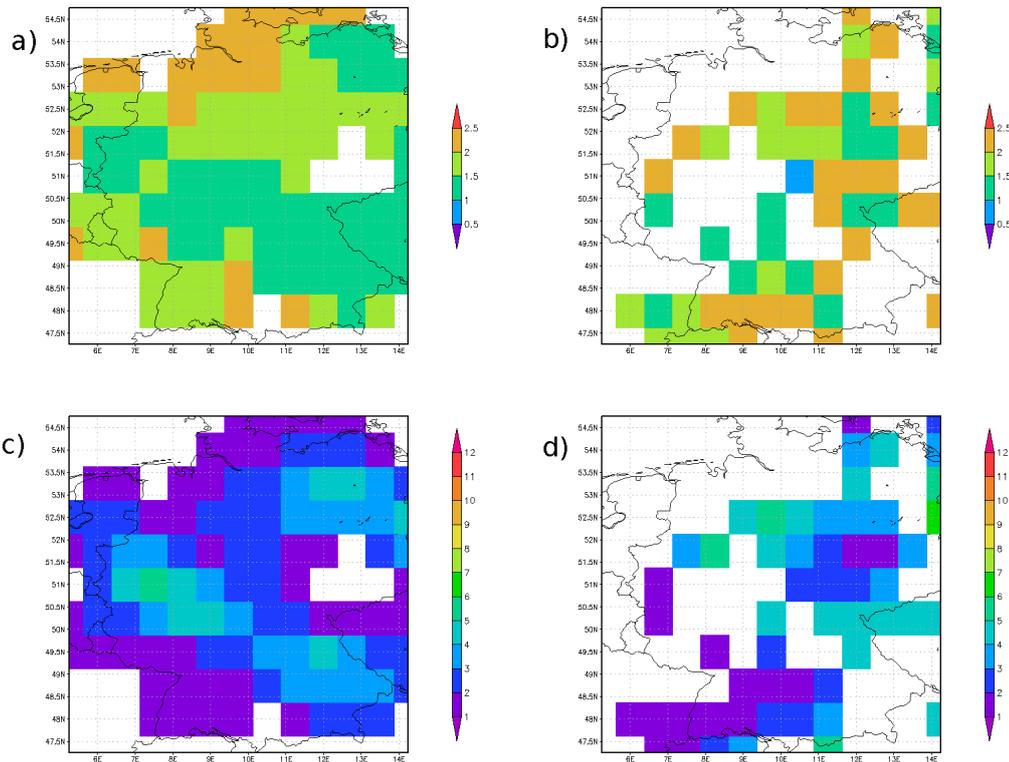


Figure 6.42: Weibull shape (a+b) and scale parameters (c+d) of event grid points as obtained from the Radolan data (a+c) and estimated from DS4 (b+d) at 2007-06-21 12UTC.

High values of low-level humidity are found over central Germany (Figure 6.41b). Both fields together suggest that the estimated convective events are preferably located in the South-West of Germany. According to the rainfall distribution parameters obtained from Radolan data, almost entire Germany is covered by estimated convective events (Figure 6.42a,c). DS4 shows for this date less grid points with an event (Figure 6.42b,d), which can be explained by the KO index that indicates a stable atmosphere in the north-western half of Germany. In western Germany, a convective rainfall maximum is revealed based on the Radolan observations, but only few grid points indicate modelled (DS4) events in this region. Although this area is characterized by a high 1000hPa relative humidity, the atmospheric stability (indicated by ERA-Interim) is a possible explanation for the small number of estimated (DS4) events in this area.

In total, the rainfall estimation (DS4) for convective events reveals the general event positions roughly in the correct place (cf. Figure 6.42c,d). In comparison to the shape parameter, the modelled scale parameter appears to agree better to the one obtained from Radolan (6.42c,d). However, differences between rainfall observations and forecasts (particularly over eastern Germany) result in differences of the scale parameter between the estimation (using DS4) and the Radolan observation. In general, the main pattern of the convective event area is also represented by the DS4 model for this situation. Nevertheless, the area of grid points with estimated events is smaller (again in agreement to a lower frequency of estimated events). Furthermore, the estimated event area is more fractured, which is related to the lower performance in predicting the shape parameter in comparison to the scale parameter.

6.8 Application to decadal predictions

The planning horizon of many political and economic institutions is commonly within 5 to 10 years, i.e., in the decadal forecast range. The development of decadal predictions thus opens a new service that meteorology can provide to society. One particular aspect of decadal forecasting is the planning of risk reduction measures with respect to intense convective events. Based on skillful forecasts on the decadal time scale, their application could be focused on periods with an enhanced risk. Decadal predictions are based on coupled ocean-atmosphere GCMs. They are generated by initializing an ensemble of model runs from the observed climate state (Meehl et al. (2009)). The ocean initialization is of primary importance, as the ocean is a climate sub-system with long memory. Estimations of future boundary conditions like the development of greenhouse gas for-

cing are another key factor for these predictions, which are thus equally depending on both initialization and forcing. The evaluation of the forecast skill of decadal forecast systems is based on so-called "hindcasts", which are forecasts starting in the past, so that their results can be compared with the observed climate development.

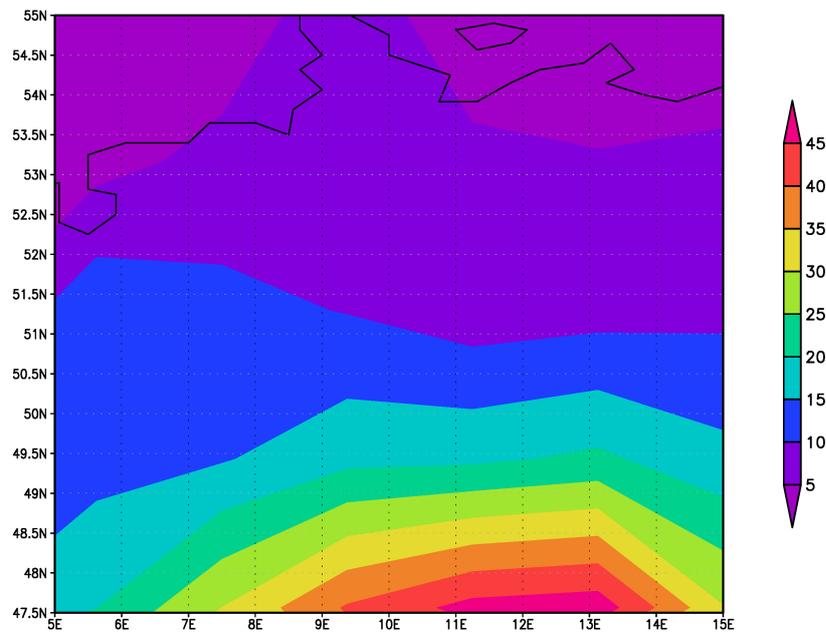


Figure 6.43: Estimated number of events per year and grid point based on MIKLIP hindcasts (3600 years representing the period 1960-2014).

The German research project MIKLIP (cf. *Marotzke et al. (2016)*) is one of the current initiatives to develop a decadal prediction system. This system is based on the MPI-ESM model system (*Jungclaus et al. (2013)*). In the MIKLIP model generation 1 (called ‘baseline 1’), which is evaluated in the current study, the prediction system is initialized with observed anomalies of ocean parameters that are superimposed on the climatology of the model ocean. This procedure avoids the occurrence of a major model drift, which is found when alternatively applying a full field‘ initialization (*Smith et al. (2007)*). The initialization basis for the ocean in baseline 1 runs is the ORAS4 ocean reanalysis (*Balmaseda et al. (2013)*). For the initialization of the atmospheric component, ERA40 (*Uppala et al. (2005)*) is used for 1960-1989 and ERA-Interim (*Dee et al. (2011)*) for 1990-2014. The decadal predictions are started each year between 1960 and 2014 with 10 realisations and then running freely for 10 years. While MIKLIP also includes regionalization efforts, only the GCM output is used here.

To analyse the performance of the MIKLIP system in predicting the frequency of convective events, the event estimation method DS4 was applied to the MIKLIP GCM output comprising 3600 forecasted years in total, and covering the period 1979-2014. In order to account for systematic errors of the model system that are found for atmospheric stability over Europe (see *Pistotnik et al. (2016)*), a bias-correction was applied to the rainfall intensity parameters resulting from these computations. The procedure used for bias-correction is identical to that used in chapter 6.4.2, i.e. mean, standard deviation and skewness are corrected by modifying the rainfall distribution parameters of the individual events in order to meet observational results. In addition, thresholds in the front analysis had to be adapted to account for the spatial resolution of the MIKLIP system (1.875°), which is different to that of ERA-Interim (0.75°). For the MIKLIP ensemble, the spatial distribution of the event frequencies (Figure 6.43) is similar to its ERA-Interim counterpart (Figure 6.33). However, there is an overestimation in the South of Germany and an underestimation in the North when compared with ERA-Interim. Moreover, the event frequency maxima related to orography are not well represented, except for the maximum related to the Alps. This result is not unexpected because of the coarse spatial resolution of the MPI-ESM.

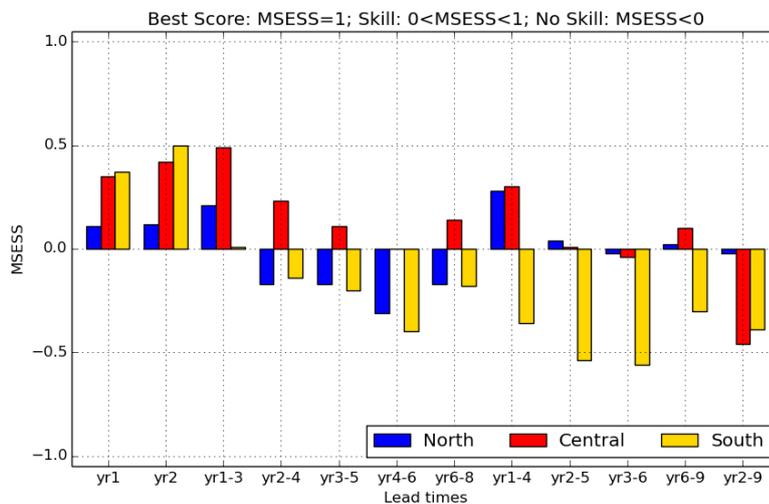


Figure 6.44: MSESS for the spatial averages of the ensemble mean event days in summer with convective rainfall events in North- (blue), Central- (red), and South-Germany (yellow) for different lead times and prediction periods in baseline1.

The skill of the baseline 1 hindcasts is quantified by the MSESS (see chapter 5.3), which is computed for the period 1979-2014 using the full 10 member ensemble, each of them starting on January 1. An ensemble of three climate runs (so-called historical runs) is

used as reference forecast for computing the MSESS. The historical runs are simulations based on the same model system as the hindcasts, which use the same greenhouse gas and aerosol forcings, but are not initialized every year. Instead, they are started from an arbitrary pre-industrial initialization and run until the final year that is used for evaluation. The analysis is performed on basis of larger areas (North Germany [7.5-15°E; 52-53°N], Central Germany [7.5-15°E; 50-52°N] and South Germany [7.5-15°E; 48-50°N]). The skill is computed for different forecast horizons, i.e., for different lead years and prediction periods, starting with lead year 1 immediately after initialization. Results are shown in Figure 6.44. A small positive prediction skill is found for all three regions in the first three years after the initialization. For North and Central Germany there is even skill when considering the first four years of the hindcasts. After this period, the skill of the decadal predictions decreases. For longer forecast periods, the skill for the South of Germany is particularly low. A possible explanation is that the low-level relative humidity and the calculation of the KO index are based on 1000hPa level data. These variables largely determine the estimated shape parameter and consequently the estimated occurrence of convective events. In the South of Germany, which is characterized by high orography, the 1000hPa level is often below the surface. Therefore, the associated variables are hence extrapolated. Consequently, the estimation of convective events in this region often depends on extrapolated data, which has a high potential to reduce the performance of the developed model. Even with this aspect leaving room for improvement, the positive skill that is found for the first couple of forecast years is a promising result, and a basis for future research on this topic.

7 Summary and discussion

7.1 Summary

In the present study, a method based on the large-scale state of the atmosphere was established to estimate the occurrence of intense convective events for Germany. The development of this methodology required a chain of research steps: In order to identify and characterize convective extreme events (during summer and at daytime), rainfall data available from calibrated hourly radar data (Radolan) were used, computing the spatial intensity distribution around individual locations. Convective events were identified using weather observations from synoptic stations, extreme weather reports (ESSL) and lightning positions from a measurement network (Nowcast). This approach requires a definition of a radius around a station in which the spatial rainfall distribution is an optimal indicator for this decision. A radius of 90km turned out to be a suitable choice. Thus, typical rainfall distributions indicating the occurrence of convective events could be identified. Lightning within the 90 km radius is a good indicator for extreme convective events. They are characterized by particularly large rainfall intensities at part of the grid points and thus a low Weibull shape parameter. Particularly, undercutting a certain threshold of the shape parameter of 1.5 provides a relatively good prediction of lightning occurrence (HSS: 0.32; see section 6.2.3). The fact that lightning can be particularly well identified from the rainfall distribution (in particular better than ESSL events and observations classified as convective) is in agreement with a study by *Gaál et al. (2014)*.

The relationship between the spatial rainfall distribution (described by the Weibull scale and shape parameter) and the large scale atmospheric conditions was explored trying out

a number of different large scale atmospheric indices computed from reanalysis data. Due to the complex statistical relationship between the two Weibull distribution parameters, it turned out to be useful to apply a mathematical transformation leading an almost linear climatological relationship of transformed Weibull parameters. It was shown that events with a low (high) atmospheric stability were found on systematically shifted lines parallel to the mean orientation of the transformed Weibull parameters. Thus, it was possible to describe the relationship between the transformed rainfall distribution parameters and the stability indices with a 2-dimensional linear fit. The strongest link was found for the KO index, which was used subsequently. Thus, a given KO index defines a specific line relating the two Weibull parameter to each other. As two parameters must be used for estimating an actual pair of Weibull parameters, a second large scale variable is required. The near surface relative humidity was used, but required second order polynomial fit. This resulted in the first development stage of the method (DS1) making it possible to estimate rainfall distribution parameters based on the given KO and RH indices. Evaluating the resulting estimation method DS1, it turned out that situations with very unstable atmospheric conditions were extremely sensitive to the exact value of the input parameters, indicating large uncertainty, particularly for these events. Different approaches for improving the above estimation were examined. A first improvement was obtained identifying nearby frontal structures. It was confirmed rainfall distributions between frontal and non-frontal situations differ considerably. Thus, the relationship between the two large-scale parameters and the rainfall distribution may be assumed to depend on their presence or absence. The fits were performed for rainfall events with cold, warm, or no front within a 300km range. The respective model version (DS2) showed, however, only slightly minor improvements in the correlations compared to DS1. Apparently, much of the frontal influence is already incorporated in the two large scale atmospheric parameters. A further improvement (DS3) was achieved by applying a bias-correction to the results of DS2, using with the properties of the local climatology of the rainfall distribution parameters. In addition, the bias-correction limits the rainfall distribution parameters to a realistic range. Again, against the previous version, the development stage shows only a modest improvement. As the model was found to be considerably more skilful in estimating the shape parameter than in estimating the scale parameter, a hybrid methodology was finally introduced which uses the ERA-Interim precipitation and the shape parameter from the aforementioned methodology for estimating the scale parameter. This approach (DS4) indeed leads to a major improvement of the scale parameter estimation.

The aim of this study is the estimation of the occurrence of convective events in Germany based on large-scale parameters. A validation of DS4 was performed for the subgroup of results indicating a strong convective event according to the estimated shape parameter (smaller than 2.5) and scale parameter (larger than 1.18). The prediction of such events was compared to observed events according to Radolan. The performance in total is relatively good, showing a HSS of up to 0.4 for the prediction of convective events based on rainfall distribution thresholds.

As application of the method, a ranking for convective extreme events was developed. It is based on the summation of rainfall intensity at all grid points of a day belonging to the aforementioned subgroup. The two top events within the observation period were investigated as case studies. In particular, For these events, the agreement of the shape and scale parameters in terms of their spatial distributions was considered. In both cases, the basic spatial structure of the event locations was reproduced. In addition to this analysis, the ranking was compared to locally extreme wind gusts in Germany. 4 of the 5 highest ranked summer gust events were found in the top 10% of the estimated events.

Finally, the methodology was applied to decadal hindcasts performed with a GCM. A small positive skill of the hindcasts in terms of the likelihood of extreme convective events was found for central Germany in the first three to four prediction years.

7.2 Discussion and conclusion

With respect to their influence of large-scale stability on the rainfall distribution parameters, the KO index revealed a stronger relationship than other parameters. For example, the potential instability index (PII) and the lifting instability index LI show a stable, yet weaker, relationship with the rainfall parameters. Quantities which have a lower boundary of their values (like CAPE) are not as tightly connected with the rainfall parameters. This is partly due to the fact that situations with zero values do not contribute to quantify the regression sought. All these parameters are modified by a developing convective event itself. Thus, the temporal evolution of rainfall and instability is a factor, that was ignored in the present study. Including it, may lead to significant improvements of the methodology.

Also, vertical wind shear is not used. It is incorporated in several convection parameters, like, e.g., those introduced by *Huntrieser et al. (1997)*, *Miller (1972)*, and *Thompson et al. (2003)*). Its influence appears to be non-linear. While *Westermayer et al. (2016)* found a reduced frequency of lightning events for a deep layer shear of 10 m/s and higher values for stronger and weaker wind shear, *Půčik et al. (2015)* found an large vertical wind shear for the production of hail, which, as lightning, requires strong vertical movements. In the course of this study, some additional tests have been performed replacing the KO-index with parameters incorporating vertical wind shear (EHI, Cape*windshear), but did not show an improved result. As a strong vertical wind shear is commonly associated with fronts, its impact on the convective situations is partly incorporated in the methodology developed here. In addition, the study of *Půčik et al. (2015)* showed that the frequency of heavy rain is much more closely related to the stability than to the deep layer shear. Also, the decreased frequency of lightning events for a deep layer shear of 10 m/s found by *Westermayer et al. (2016)* is not as strong as the dependence of the lightning frequency on the stability (see their Figure 9). Also the study of *Mohr & Kunz (2013)* found a reduced performance for almost all tested stability indices including dynamical properties (except CAPE*shear) for Germany, which is in accordance to the results found here.

The comparison between the Weibull parameters and the occurrence of convective extreme events derived from other sources showed that these events are characterized by a low shape parameter and high values of the scale parameter. This behaviour emerged particularly clearly for the instrumentally detected lightnings. The relation to lightnings can be understood from the fact that the generation of lightning requires a minimum strength of convection. Usage of other indicators of convective activity, like lightning observed at synoptic stations, or ESSL reports, produced a weaker relationship to the shape parameter. Thus, the choice of the rainfall distribution associated with instrumentally detected lightning is particularly suitable for the detection of convective extreme events.

The applied method (from large-scale parameters to rainfall parameters to estimate the occurrence of convective events) performs better in the northern half of Germany compared to the south (see sections 6.4.3 and 6.5). One possible reason may be related to problems with the used values of the humidity obtained for 1000hPa, which will occasionally be extrapolated, in particular in the presence of orography, which is more

pronounced in the South of Germany.

An analysis of the ranked events (c.f. table 6.19) often reveals a convergence zone ahead of a cold front. This explains a large extension of the convective area in these cases, but also shows that the front analysis is, in principle, an important ingredient of the methodology. Convective rainfall typically occurs at a cold front, while at a warm front usually large scale precipitation is observed. The applied front identification scheme resulted in a unexpected high number of convective events at warm fronts. This suggests that further improvement of the method of distinguishing warm and cold fronts should be attempted. If an improvement of the frontal detection scheme would lead to the ability to identify convergence lines, they can directly be accounted for, rather than including their potential role by choosing a large influence radius of a front. As a consequence, a reduced radius of frontal influence could be used, leading to a better representation of the identified relationship for a front type and thus an improved estimation. Distinguishing pre-/postfrontal situations may also be beneficial.

It may be possible to modify the bias-correction applied, relating it to frontal influences rather than to local rainfall climatologies alone. For example, the estimation of scale and shape parameter, as an alternative to the procedure performed here, may be based on the ERA-Interim rainfall and the identified KO-relationship alone. With such an approach, there is no need to use relative humidity as one parameter. This may be beneficial, as the KO index reveals a much stronger relationship with both rainfall distribution parameters.

In the presented work, it was shown that the locations of convective extreme events can be estimated from the small shape and high scale parameters at ERA-Interim grid points. Both of these Weibull parameters are estimated by the model DS4 performing better for the scale parameter than for the shape parameter. The final estimation method, which requires only few coarse resolved atmospheric variables, shows skill to estimate the occurrence of convective extreme events and to estimate their convectivity and intensity, and thus has the potential to estimate convective events for a large number of climate scenario runs in the future.

Appendices

A Stability/humidity indices

A.1 Vertical totals

The vertical totals index was introduced by *Miller (1972)* and is estimating the instability using the temperature difference between 850hPa and 500hPa without any humidity information.

$$VT = T_{850} - T_{500}$$

Hence, this index represents the average temperature gradient between roundabout the top of the boundary layer extending through half of the air mass. An increased gradient is related to a less stable or more unstable atmosphere and consequently a higher probability of thunderstorms. 26 K defines a threshold which is usually assumed to separate best between thunderstorm prone weather and weather that cannot produce thunderstorms. VT=26K corresponds approximately to a vertical temperature gradient of 0.65K/100m, which is the average observed lapse rate.

A.2 Total totals

The total totals index is used to predict the occurrence of thunderstorms and was first introduced by *Miller (1972)*. It is fairly easy to calculate and is hence widely used.

$$TT = T_{850} - T_{500} + T_{d,850} - T_{500} = T_{850} + T_{d,850} - 2 \cdot T_{500}$$

A value of the total totals within specific ranges indicates the probability of thunderstorms. If the value is below 44, thunderstorms are unlikely to occur. With increasing value, the instability increases and consequently the probability of thunderstorms.

TT value	Thunderstorm occurrence
<44	Convection not likely
44-50	Likely thunderstorms
51-52	Isolated severe storms
53-56	Widely scattered severe
>56	Scattered severe storms

Table A.1: values of total totals and the associated expected weather. (Source: www.theweatherprediction.com).

Due to the weak influence of humidity in the different levels, this measure is not the optimum in case of lifting of the atmosphere. Such situation changes the temperatures a lot using latent heat.

A.3 KO-index

The KO-Index was presented in Andersson *et al.*, 1989. It represents the potential stability of an air mass. The average vertical equivalent potential temperature gradient from the low to the mid-troposphere is represented by information of different pressure levels (1000hPa, 850hPa, 700hPa and 500hPa) it is defined as:

$$KO = 1/2 \cdot (\Theta_{e,500} + \Theta_{e,700}) - 1/2 \cdot (\Theta_{e,850} + \Theta_{e,1000})$$

The probability of thunderstorms increases with a decrease of KO.

KO-index value	Thunderstorm occurrence
> 6	None
2 to 6	Isolated thunderstorms
< 2	Numerous thunderstorms

Table A.2: KO-index values vs. Airmass Thunderstorm Probability. (Source: Online-Lexikon des DWD).

A.4 S index

The S index uses the total totals index and the temperature and dew point information of 700hPa. It is defined as:

$$SI = TT - (T_{700} - T_{d,700}) - x$$

with $x=0$ if $VT > 25$
 $x=2$ if $25 > VT \geq 22$
 $x=6$ if $22 > VT$

The German Military Geophysical Office (*Reymann et al. (1998)*) introduced this index for the period from April to September. It is similar to the K index, but penalizes low values of VT. SI can be reformulated to $K - T_{500} - \zeta$ where ζ is the penalty coefficient.

A.5 Potential instability index

The potential instability index has been introduced by *Van Delden (2001)* and is a measure of the potential instability between the 925hPa and 500hPa layer.

$$PII = \frac{\theta_{e,925} - \theta_{e,500}}{z_{500} - z_{925}}$$

A.6 Lifted index

The lifted index provides information of the stability using 2 layers. In most cases these are the surface height and the 500hPa level. It has been first presented by *Galway (1956)*. The concept of adiabatic lifting is the main idea behind this concept. To make the temperature of two layers comparable, the lower layer is lifted to the higher one. If the lifted air is colder, the temperature difference is positive. This means that the air is stable. If the lifted air is warmer, then it would have a lower density than the upper air mass when lifted to that level. This means that the air would be unstable.

LI value	Thunderstorm Probability
≥ 2	No significant activity
0 to 2	Showers/thunderstorms possible with other source of lift
-2 to 0	Thunderstorms possible
-4 to -2	Thunderstorms more probable, but few if any severe
< -4	Severe Thunderstorms possible

Table A.3: The risk of thunderstorms and severe weather activity as indicated by the lifted index. (Source: online glossary of NOAA).

A.7 Showalter index

The Showalter index is calculated similarly to the vertical totals. It is a temperature difference from the temperature in 500hPa to the temperature of an air parcel ascending from 850hPa to 500hPa. It has been developed by *Showalter (1953)* and been modified by *Steinacker (1977)* for mountainous regions.

$$SI = T_{500} - T'_{850 \rightarrow 500} \quad (\text{A.1})$$

SI value	Thunderstorm Probability
≥ 3	No significant activity
1 to 3	Showers/thunderstorms possible with other source of lift
-2 to 1	Thunderstorms possible (generally weak)
-3 to -2	Thunderstorms more probable (possibly strong)
-6 to -3	Strong or severe thunderstorms possible
< -4	Any thunderstorm likely to be strong or severe

Table A.4: The risk of thunderstorms and severe weather activity as indicated by the lifted index. (Source: online glossary of NOAA).

A.8 Deep convection index

The deep convection index (*Barlow (1993)*) makes use of the lifted index and the temperature and the dew point in 850 hPa. In particular, its intention is to predict the occurrence of convection reaching high levels. Such convection is usually associated with strong thunderstorms.

$$DCI = (T + T_d)_{850} + LI \quad (\text{A.2})$$

Thunderstorms are assumed to may occur if DCI exceeds 30°C .

A.9 CAPE

The convective available potential energy (CAPE) describes the energy that an air parcel is gaining when lifted over the level of free convection to the level of zero buoyancy. In this concept the ascending air parcel does not exchange temperature or humidity with the surrounding air. Also, developing drops due to condensing water do not provide earthward impulse to the air parcel. In general, there exist 3 approaches for obtaining the initial conditions of the ascending air parcel. The first one used surface based temperature, pressure and humidity and the corresponding CAPE is therefore called SB CAPE (surface based). Problematic in this concept is a very high humidity close to the surface. A dry boundary layer may then lead to an overestimation of the lability. The second one decreases the effect by assuming the air parcel to be a mixture of the lowest levels. This is called ML CAPE (mixed layer) and uses the lowest 50-100hPa of the atmosphere. A third concept assumes that the air parcel comes from the layer with maximum CAPE and is called MU CAPE (maximum unstable CAPE). This CAPE is of importance for elevated convection, which can appear at night close to the surface and at a warm front. The ascent of the parcel can be assumed in different ways. First as pseudo-adiabatic process, second as a reversible process and third as an ascent including mixture with the surrounding, which leaves partly the air parcel concept.

CAPE value	Atmospheric Instability
$0 \text{ J/kg} < \text{CAPE} < 1000 \text{ J/kg}$	marginal instability
$1000 \text{ J/kg} < \text{CAPE} < 2500 \text{ J/kg}$	moderate instability
$2500 \text{ J/kg} < \text{CAPE} < 3500 \text{ J/kg}$	strong instability
$3500 \text{ J/kg} < \text{CAPE}$	extreme instability

Table A.5: Values of CAPE and indicated instability. (Source: online glossary of NOAA).

A.10 CIN

The convective inhibition is the energy a lifted air parcel needs to overcome to reach its level of free convection. It is defined as CAPE, because it makes use of the buoyancy. Usually this value is non-negative due to a stable situation in low levels. CIN is often present due to a thermal inversion above the boundary layer. CIN is often overcome by dynamical lifting or locally high or humid air near surface. Graphically, CIN can be

obtained as the area between the air parcel trajectory and the temperature profile in a T-log(p) diagram below the LFC.

$$CIN = g \cdot \int_0^{LFC} \frac{\theta_{v,p} - \theta_{v,e}}{\theta_{v,e}} dz$$

If CIN is very low, convection can start easily. In case it becomes very strong, it cannot be overcome (except for cases of forced lifting e.g. mountains or large scale forcing as confluence)

CIN value	Thunderstorm Occurrence
0 J/kg < CIN < 15 J/kg	only minor cumuli develop
15 J/kg < CIN < 50 J/kg	single cell thunderstorms possible
50 J/kg < CIN < 200 J/kg	multi-cell thunderstorms possible
200 J/kg < CIN	stability of stratification to overcome. no thunderstorms develop.

Table A.6: CIN and the associated potential of thunderstorms (Source: online glossary of NOAA).

A.11 K-index

The K-index was defined by *Gerorge (1960)* to predict the occurrence of thunderstorms. This index can be calculated from temperature and dew points (both in °C) in standard levels of radiosondes ascends. The K-index is defined as:

$$K = \underbrace{T_{850} - T_{500}}_A + \underbrace{T_{d,850}}_B - \underbrace{(T - T_d)}_{700} \quad (A.3)$$

A provides information about the dry-static stability between 850hPa and 500hPa. Term B gives the specific humidity in 850hPa and Term C is a measure of the humidity in 700hPa.

K-index value	Thunderstorm Probability
Less than 20	None
15 to 20	Isolated thunderstorms and showers possible (<20%)
20 to 25	Isolated thunderstorms and showers possible (20-40%)
26 to 30	Widely scattered showers (40-60%)
31 to 35	Showers possible (60-80%)
31 to 35	Numerous showers and thunderstorms probable (80-90%)
Above 35	Numerous showers and some heavy thunderstorms (>90%)

Table A.7: K-index values vs. Airmass Thunderstorm Probability. (Source: glossary of www.srh.noaa.gov).

A.12 SWEAT

Bidner (1970) developed an index for the prediction of severe thunderstorms. The Severe Weather Threat Index has been developed for the Great Plains in the United States.

$$SWEAT = 12 \cdot T_{d,850} + 20 \cdot (TT - 49) + 2 \cdot f_{850} + f_{500} + 125 \cdot [\sin(d_{500} - d_{850})] + 0.2$$

SWEAT value	Convective Potential
SWEAT > 400	tornadic possible
400 > SWEAT > 300	Severe possible
300 > SWEAT > 150	Slight severe

Table A.8: SWEAT values and indicated convective potential. (Source: www.theweatherprediction.com).

A.13 SWISS₁₂

The SWISS₁₂ index has been developed by *Huntrieser et al. (1997)* and adjusted to the situation in Switzerland. It includes not only static information but also wind shear. It is the follow-up to the SWISS₀₀ index, which used the information from 00UTC for Switzerland. SWISS₀₀ is defined as:

$$SWISS_{12} = ShowI - 0.4WSh_{3-6km} + 0.1(T - T_d)_{600}$$

The SWISS₁₂ index only slightly differs, and uses 12UTC values instead of 00UTC values. The SWISS₁₂-index uses the 12UTC represents better the convective preconditions the thunderstorms, which occur most often in the afternoon.

A.14 DSI

The dynamical state index has been developed by Peter Nevier. It is a rather complicated index, but represents mainly ageostrophy. It is based on the geopotential, the potential vorticity and Bernulli stream function (*Névir (2004)*).

$$DSI = -g \frac{\partial (\Pi, \theta, B)}{\partial (x, y, p)}$$

A.15 Helicity

The helicity H in (J/kg) describes the transport of vorticity of the surrounding into an air parcel, which is experiencing a convective movement.

$$H = - \int_{z_0}^z \vec{v}_h \cdot \vec{\zeta}_h dz$$

where \vec{v}_h and $\vec{\zeta}_h$ are the horizontal vectors of velocity and vorticity. In case the Wind changes only its strength with height, the horizontal axis of the vorticity is perpendicular to the flow or v_h (vorticity produced by wind shear) and the helicity is consequently zero. In case there is a change of wind direction with height, H is not zero and a horizontal axis of rotation exists in the stream direction (streamwise vorticity).

A.16 Storm-relative environmental helicity

The 'storm-relative environmental helicity' (SREH) describes the relative influx into the thunderstorm cluster based on the horizontal wind vector. It is defined as:

$$SREH = - \int_{z_0}^z \vec{k} \cdot (\vec{v}_h - \vec{c}) \times \frac{\partial \vec{v}_h}{\partial z} dz$$

c defines the (horizontal) velocity of the thunderstorm cell/cluster. The integration is performed for the influx height, which is often regarded to be the lowest 3km. As visible from the formula, SREH corresponds to 2 times the area defined by the wind shear vector and the influx vector in the hodogram. The value of SREH provides a measure of the tendency in a supercell to produce rotating updrafts.

A.17 Energy helicity index

The „Energy Helicity Index“ (EHI) was defined by *Hart & Korotky (1991)* and is defined by the product of CAPE and relative helicity with a normalization factor. In case of a stable atmospheric layering, rotating supercells and tornadoes don't occur under high values of SREH(intense vertical wind shear). That is why, empirical indices connect CAPE and SREH. One of these is often used in the United States of America to predict tornadoes and is called 'energy helicity index' (EHI). A defined threshold of CAPE needs to be exceeded to enable the possibility of tornadoes in this model. EHI is defined as:

$$EHI = \frac{(CAPE * SREH)}{CAPE_{ts}}$$

The following relationship between tornadoes and EHI is found and applied by the US Storm Prediction Center (SPC):

EHI value	Thunderstorm Potential
EHI > 1	thunderstorms possible / potential for super cells
5>EHI>1	up to F2, F3 tornadoes possible
EHI>5	up to F4, F5 tornadoes possible

Table A.9: EHI values and indicated potential of tornadoes/supercells. (Source: www.theweatherprediction.com).

EHI is calculated operationally for a 1 km (EHI1) and 3 km (EHI3) inflow height. Maps of the situation here is available under:

http://www.spc.nssl.noaa.gov/exper/mesoanalysis/new/meso_mob.php?sector=&parm=ehi3

A.18 SHUM_{1000→500}

This parameter is actually not a stability index, but a moisture parameter SHUM_{1000→500} has been developed based on the concept of the air parcel theory. It provides the information about reduction of specific humidity [kg/kg] for an air parcel loses, which lifted adiabatically from 1000 to 500hPa. This corresponds to the heights, which are compared in the KO-index. Related to this moisture parameter are fore example SHUM(1000) and SHUM(850), which indicate the specific humidity at a certain pressure level (in hPa). Analogous, the relative humidity at 1000hPa is indicated by RH(1000).

B Estimating transformed rainfall intensity distribution parameters from the KO/RH index

The following formula gives the intersections based on the formulas describing the obtained relationships between KO/RH and the rainfall intensity distribution parameters.

$$x'_{1,2} = \left(\pm \sqrt{\frac{(b_2 \cdot a_2^2 - b_3 \cdot a_3 \cdot a_2 + 2 \cdot b_5 \cdot a_1 \cdot a_3 - 2 \cdot b_5 \cdot a_3 \cdot KO - b_6 \cdot a_1 \cdot a_2 + \dots}{b_6 \cdot a_2 \cdot KO)^2 - 4 \cdot (b_4 \cdot a_2^2 + b_5 \cdot a_3^2 - b_6 \cdot a_3 \cdot a_2) \cdot \dots}}{2 \cdot (b_4 \cdot a_2^2 + b_5 \cdot a_3^2 - b_6 \cdot a_3 \cdot a_2)} \right) / \dots$$

$$y' = (a_1 + a_2 \cdot x') / a_3 \quad (\text{B.1})$$

with $a_1 = -9.777$; $a_2 = 20.52$; $a_3 = 27.112$; $b_1 = 64.842$; $b_2 = 30.943$; $b_3 = 42.918$; $b_4 = -14.379$; $b_5 = -18.892$; $b_6 = -32.293$.

References

- ANAGNOSTOPOULOU, C., & TOLIKA, K. 2012. Extreme precipitation in Europe: statistical threshold selection based on climatological criteria. *Theoretical and Applied Climatology*, **107**(3-4), 479–489.
- ANDERSSON, T., ANDERSSON, M., JACOBSSON, C., & NILSSON, S. 1989. Thermodynamic indices for forecasting thunderstorms in southern Sweden. *Meteorological Magazine*, **118**(1404), 141–146.
- AYDIN, K., SELIGA, T.A., & BALAJI, V. 1986. Remote sensing of hail with a dual linear polarization radar. *Journal of Climate and Applied Meteorology*, **25**(10), 1475–1484.
- BALMASEDA, MAGDALENA ALONSO, MOGENSEN, KRISTIAN, & WEAVER, ANTHONY T. 2013. Evaluation of the ECMWF ocean reanalysis system ORAS4. *Quarterly Journal of the Royal Meteorological Society*, **139**(674), 1132–1161.
- BARLOW, W. 1993. A new index for the prediction of deep convection. *Pages 129–132 of: 17th conference on severe local storms, St. Louis, MO, Amer. Meteorol Soc.*
- BATES, D.M., & CHAMBERS, J.M. 1992. Nonlinear models. *Chapter 10 of Statistical Models in S eds J. M. Chambers and T. J. Hastie, Wadsworth & Brooks/Cole.*
- BATES, D.M., & WATTS, D. G. 1988. *Nonlinear Regression Analysis and Its Applications.* Wiley.
- BERRY, G., REEDER, M.J., & JAKOB, C. 2011. A global climatology of atmospheric fronts. *Geophysical research letters*, **38**, L04809.

- BETZ, H.D., SCHMIDT, K., FUCHS, B., OETTINGER, W.P., & HÖLLER, H. 2007. Cloud lightning: detection and utilization for total lightning measured in the VLF/LF regime. *Lightning Res*, **2**, 1–17.
- BETZ, H.D., SCHMIDT, K., LAROCHE, P., BLANCHET, P., OETTINGER, W.P., DEFER, E., DZIEWIT, Z., & KONARSKI, J. 2009. LINET—an international lightning detection network in Europe. *Atmospheric Research*, **91**(2), 564–573.
- BIDNER, A. 1970. The Air Force Global Weather Central severe weather threat (SWEAT) index - a preliminary report. *Air Weather Service Aerospace Science Review*, **AWS RP 105-2, No 70-3**, 2–5.
- BRIGHT, D.R., & WANDISHIN, M.S. 2005. A physically based parameter for lightning prediction and its calibration in ensemble forecasts. *In: Preprints, Conf. on Meteor. Applications of Lightning Data*.
- BRINGI, V.N., SELIGA, T.A., & AYDIN, K. 1984. Hail detection with a differential reflectivity radar. *Science*, **225**(4667), 1145–1147.
- BRONSTERT, A., NIEHOFF, D., & BÜRGER, G. 2002. Effects of climate and land-use change on storm runoff generation: present knowledge and modelling capabilities. *Hydrological Processes*, **16**(2), 509–529.
- BROWNING, KA, & HARROLD, TW. 1970. Air motion and precipitation growth at a cold front. *Quarterly Journal of the Royal Meteorological Society*, **96**(409), 369–389.
- BROWNING, K.A., PARDOE, C.W., & HILL, F.F. 1975. The nature of orographic rain at wintertime cold fronts. *Quarterly Journal of the Royal Meteorological Society*, **101**(428), 333–352.
- CAEM WORKING GROUP ON ADVANCED TECHNIQUES APPLIED TO AERONAUTICAL METEOROLOGY. 1992. Methods of interpreting numerical weather prediction output for aeronautical meteorology. *Secretariat of the World Meteorological Organisation, Geneva, Switzerland*, **770(195)**, 89ff.
- CASAS, M. C., CODINA, B., REDAÑO, A., & LORENTE, J. 2004. A methodology to classify extreme rainfall events in the western mediterranean area. *Theoretical and Applied Climatology*, **77**(3-4), 139–150.
- CATTO, J.L., & PFAHL, S. 2013. The importance of fronts for extreme precipitation. *Journal of Geophysical Research: Atmospheres*, **118**(19), 10–791.

- CATTO, J.L., JAKOB, C., BERRY, G., & NICHOLLS, N. 2012. Relating global precipitation to atmospheric fronts. *Geophysical Research Letters*, **39**(10).
- CHAN, S.C., KENDON, E.J., FOWLER, H.J., BLENKINSOP, S., FERRO, C.A.T., & STEPHENSON, D.B. 2013. Does increasing the spatial resolution of a regional climate model improve the simulated daily precipitation? *Climate dynamics*, **41**(5-6), 1475–1495.
- CHROMOW, SP. 1942. Die Kaltfront. *Pages 287–294 of: Einführung in die synoptische Wetteranalyse*. Springer.
- COLBY J., FRANK P. 1984. Convective inhibition as a predictor of convection during AVE-SESAME II. *Monthly weather review*, **112**(11), 2239–2252.
- COX, D. R., & ISHAM, V. 1988. A Simple Spatial-Temporal Model of Rainfall. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, **415**(1849), 317–328.
- DAVIES-JONES, R., BURGESS, D., & FOSTER, M. 1990. Test of helicity as a tornado forecast parameter. *Pages 588–592 of: Preprints, 16th Conf. on Severe Local Storms, Kananaskis Park, AB, Canada, Amer. Meteor. Soc.*
- DAVIES-JONES, ROBERT P. 1986. Tornado dynamics. *Thunderstorm morphology and dynamics*, **2**, 197–236.
- DEE, D.P., UPPALA, S.M., SIMMONS, A.J., BERRISFORD, P., POLI, P., KOBAYASHI, S., ANDRAE, U., BALMASEDA, M.A., BALSAMO, G., BAUER, P., *et al.* . 2011. The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, **137**(656), 553–597.
- DOMINY, P. 2006. Creating the daily analysis charts for the Weather log. *Weather*, **61**(6), 167–170.
- DOTZEK, N., GRIESER, J., & BROOKS, H.E. 2003. Statistical modeling of tornado intensity distributions. *Atmospheric research*, **67**, 163–187.
- DOTZEK, N., GROENEMEIJER, P., FEUERSTEIN, B., & HOLZER, A.M. 2009. Overview of ESSL's severe convective storms research using the European Severe Weather Database ESWD. *Atmospheric Research*, **93**(1), 575–586.

- DROEGEMEIER, K.K., LAZARUS, S.M., & DAVIES-JONES, R. 1993. The influence of helicity on numerically simulated convective storms. *Monthly weather review*, **121**(7), 2005–2029.
- DUNN, P.K. 2004. Occurrence and quantity of precipitation can be modelled simultaneously. *International Journal of Climatology*, **24**(10), 1231–1239.
- DWD. 2015 (April). *Lexikon: Frontenklassifikation @ONLINE*.
- EGGERT, B., BERG, P., HAERTER, J.O., JACOB, D., & MOSELEY, C. 2015. Temporal and spatial scaling impacts on extreme precipitation. *Atmospheric Chemistry and Physics*, **15**(10), 5957–5971.
- FRÜH, B., FELDMANN, H., PANITZ, H.-J., SCHÄDLER, G., JACOB, D., LORENZ, P., & KEULER, K. 2010. Determination of Precipitation Return Values in Complex Terrain and Their Evaluation. *Journal of Climate*, **23**(9), 2257–2274.
- GAÁL, L., MOLNAR, P., & SZOLGAY, J. 2014. Selection of intense rainfall events based on intensity thresholds and lightning data in Switzerland. *Hydrology and Earth System Sciences*, **18**(5), 1561–1573.
- GALLUS JR., WILLIAM A. 1999. Eta simulations of three extreme precipitation events: Sensitivity to resolution and convective parameterization. *Weather and Forecasting*, **14**(3), 405–426.
- GALWAY, J.G. 1956. The lifted index as a predictor of latent instability. *Bull. Am. Meteorol. Soc.*, 528–529.
- GEORGE, J.J. 2014. *Weather forecasting for aeronautics*. Academic press.
- GERMAN INSURANCE ASSOCIATION. 2014a. Naturgefahrenreport 2014. *Gesamtverband der Deutschen Versicherungswirtschaft e. V.*, Stand: September 2014.
- GERMAN INSURANCE ASSOCIATION. 2014b. Online-Serviceteil zum Naturgefahrenreport 2014. *Gesamtverband der Deutschen Versicherungswirtschaft e. V.*, Stand: September 2014.
- GERORGE, J.J. 1960. *Weather Forecasting for Aeronautics*. Academic Press, New York, 407–415.
- GIAIOTTI, D.B., STEINACKER, R., & STEL, F. 2006. *Atmospheric Convection: Research and Operational Forecasting Aspects*, vol. 475-CISM Courses and Lectures.

- GRASSETT, R., SCHÜTTEMEYER, D., WARRACH-SAGI, K., AMENT, F., & SIMMER, C. 2008. Validation of TERRA-ML with discharge measurements. *Meteorologische Zeitschrift*, **17**(6), 763–773.
- HAKLANDER, A.J., & VAN DELDEN, A. 2003. Thunderstorm predictors and their forecast skill for the Netherlands. *Atmospheric Research*, **67**, 273–299.
- HART, J.A., & KOROTKY, W. 1991. The SHARP workstation v1. 50 users guide. *NOAA/National Weather Service*.
- HEIDKE, P. 1926. Berechnung des Erfolges und der Güte der Windstärkevorhersagen im Sturmwarnungsdienst. *Geografiska Annaler*, **8**, 301–349.
- HEWSON, T.D. 1998. Objective fronts. *Meteorological Applications*, **5**(1), 37–65.
- HUNTRIESER, H., SCHIESSER, H.H., SCHMID, W., & WALDVOGEL, A. 1997. Comparison of Traditional and Newly Developed Thunderstorm Indices for Switzerland. *Wea. Forecasting*, 108–125.
- INAN, U., FISHMAN, G.J., REISING, S.C., HORACK, J.M., *et al.* . 1996. On the association of terrestrial gamma-ray bursts with lightning and implications for sprites.
- JUNGLAUS, JH, FISCHER, NILS, HAAK, HELMUTH, LOHMANN, K, MAROTZKE, J, MATEI, D, MIKOLAJEWICZ, U, NOTZ, D, & STORCH, JS. 2013. Characteristics of the ocean simulations in the Max Planck Institute Ocean Model (MPIOM) the ocean component of the MPI-Earth system model. *Journal of Advances in Modeling Earth Systems*, **5**(2), 422–446.
- JUNGHÄNEL, T., BRENDDEL, C., WINTERRATH, T., & WALTER, A. 2015. Towards a radar- and observation-based hail climatology for Germany. *Meteorologische Zeitschrift*, 01, –.
- KATARZYNA, S. 2013. The influence of atmospheric circulation on the occurrence of hail in the North German Lowlands. *Theoretical and Applied Climatology*, **112**(3-4), 363–373.
- KEIL, C., HEINLEIN, F., & CRAIG, G.C. 2014. The convective adjustment time-scale as indicator of predictability of convective precipitation. *Quarterly Journal of the Royal Meteorological Society*, **140**(679), 480–490.
- KUNZ, M. 2007. The skill of convective parameters and indices to predict isolated and severe thunderstorms. *Natural Hazards and Earth System Science*, **7**(2), 327–342.

- KUNZ, M., & KOTTMEIER, C. 2006a. Orographic Enhancement of Precipitation over Low Mountain Ranges. Part I: Model Formulation and Idealized Simulations. *Journal of Applied Meteorology and Climatology*, **45**(8), 1025–1040.
- KUNZ, M., & KOTTMEIER, C. 2006b. Orographic Enhancement of Precipitation over Low Mountain Ranges. Part II: Simulations of Heavy Precipitation Events over Southwest Germany. *Journal of Applied Meteorology and Climatology*, **45**(8), 1041–1055.
- KUNZ, M., SANDER, J., & KOTTMEIER, C. 2009. Recent trends of thunderstorm and hailstorm frequency and their relation to atmospheric characteristics in southwest Germany. *International Journal of Climatology*, **29**(15), 2283–2297.
- KURZ, M. 1990. *Synoptische Meteorologie*. Selbstverl. des Dt. Wetterdienstes.
- LANG, T.J., & RUTLEDGE, S.A. 2002. Relationships between Convective Storm Kinematics, Precipitation, and Lightning. *Monthly Weather Review*, **130**(10), 2492–2506.
- LENDERINK, G. 2010. Exploring metrics of extreme daily precipitation in a large ensemble of regional climate model simulations. *Climate research (Open Access for articles 4 years old and older)*, **44**(2), 151.
- LI, Z., BRISSETTE, F., & CHEN, J. 2013. Finding the most appropriate precipitation probability distribution for stochastic weather generation and hydrological modelling in Nordic watersheds. *Hydrological Processes*, **27**(25), 3718–3729.
- LLASAT, M.-C., CEPERUELO, M., & RIGO, T. 2007. Rainfall regionalization on the basis of the precipitation convective features using a raingauge network and weather radar observations. *Atmospheric research*, **83**(2), 415–426.
- LLASAT, M.C. 2001. An objective classification of rainfall events on the basis of their convective features: application to rainfall intensity in the northeast of Spain. *International Journal of Climatology*, **21**(11), 1385–1400.
- MARGULES, M. 1906. Über Temperaturschichtung in stationär bewegter und ruhender Luft. *Meteorol. Z.*, **23**, 165–254.
- MAROTZKE, J., MÜLLER, W. A., VAMBORG, F.S.E., BECKER, P., CUBASCH, U., FELDMANN, H., KASPAR, F., KOTTMEIER, C., MARINI, C., POLKOVA, I., *et al.* . 2016. MiKlip—a National Research Project on Decadal Climate Prediction. *Bulletin of the American Meteorological Society*.

- MEEHL, GERALD A, GODDARD, LISA, MURPHY, JAMES, STOUFFER, RONALD J, BOER, GEORGE, DANABASOGLU, GOKHAN, DIXON, KEITH, GIORGETTA, MARCO A, GREENE, ARTHUR M, HAWKINS, ED, *et al.* . 2009. Decadal prediction: can it be skillful? *Bulletin of the American Meteorological Society*, **90**(10), 1467.
- MILLER, R.C. 1972. *Notes on analysis and severe-storm forecasting procedures of the Air Force Global Weather Central*. Tech. rept. DTIC Document.
- MOEMKEN, J., REYERS, M., BULDMANN, B., & PINTO, J.G. 2016. Decadal predictability of regional scale wind speed and wind energy potentials over Central Europe. *Tellus A*, **68**.
- MOHR, S., & KUNZ, M. 2013. Recent trends and variabilities of convective parameters relevant for hail events in Germany and Europe. *Atmos. Res.*, **123**, 211–228.
- MOHR, S., KUNZ, M., & KEULER, K. 2015a. Development and application of a logistic model to estimate the past and future hail potential in Germany. *Journal of Geophysical Research: Atmospheres*.
- MOHR, S., KUNZ, M., & GEYER, B. 2015b. Hail potential in Europe based on a regional climate model hindcast. *Geophysical Research Letters*, **42**(24).
- MONCRIEFF, M.W., & MILLER, M.J. 1976. The dynamics and simulation of tropical cumulonimbus and squall lines. *Quarterly Journal of the Royal Meteorological Society*, **102**(432), 373–394.
- MOORE, R.J., & CLARKE, R.T. 1981. A distribution function approach to rainfall runoff modeling. *Water Resources Research*, **17**(5), 1367–1382.
- NÉVIR, P. 2004. Ertel's vorticity theorems, the particle relabelling symmetry and the energy-vorticity theory of fluid mechanics. *Meteorologische Zeitschrift*, **13**(6), 485–498.
- PAPRITZ, L., PFAHL, S., RUDEVA, I., SIMMONDS, I., SODEMANN, H., & WERNLI, H. 2014. The role of extratropical cyclones and fronts for Southern Ocean freshwater fluxes. *Journal of Climate*, **27**(16), 6205–6224.
- PETTERSSSEN, S., KNIGHTING, E., JAMES, R.W., & HERLOFSON, N. 1946. *Convection in theory and practice*. Vol. XVI. Cammermeyer in Komm.

- PINTO, J.G., ULBRICH, S., PARODI, A., RUDARI, R., BONI, G., & ULBRICH, U. 2013. Identification and ranking of extraordinary rainfall events over Northwest Italy: The role of Atlantic moisture. *Journal of Geophysical Research: Atmospheres*, **118**(5), 2085–2097.
- PISTOTNIK, G., GROENEMEIJER, P., & SAUSEN, R. 2016. Validation of Convective Parameters in MPI-ESM Decadal Hindcasts (1971-2012) against ERA-Interim Reanalyses. *Meteorologische Zeitschrift*, *07*, –.
- PÚČIK, T., GROENEMEIJER, P., RÝVA, D., & KOLÁŘ, M. 2015. Proximity soundings of severe and nonsevere thunderstorms in Central Europe. *Monthly Weather Review*, **143**(12), 4805–4821.
- RAUHALA, J., & SCHULTZ, D.M. 2009. Severe thunderstorm and tornado warnings in Europe . *Atmospheric Research*, **93**(13), 369 – 380. 4th European Conference on Severe Storms 4ECSS 4th European Conference on Severe Storms.
- REYMANN, M., PIASECKI, J., HOSEIN, F., LARABEE, S., & WILLIAMS, G. 1998. *Meteorological techniques*. Tech. rept. DTIC Document.
- RIGO, T., & LLASAT, M.C. 2004. A methodology for the classification of convective structures using meteorological radar: Application to heavy rainfall events on the Mediterranean coast of the Iberian Peninsula. *Natural Hazards and Earth System Science*, **4**(1), 59–68.
- ROBERTS, N.M., & LEAN, H.W. 2008. Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events. *Monthly Weather Review*, **136**(1), 78–97.
- RODRIGUEZ-ITURBE, I., COX, D. R., & ISHAM, V. 1987. Some Models for Rainfall Based on Stochastic Point Processes. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, **410**(1839), 269–288.
- RODRIGUEZ-ITURBE, I., COX, D. R., & ISHAM, V. 1988. A Point Process Model for Rainfall: Further Developments. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, **417**(1853), 283–298.
- RUDEVA, I., & SIMMONDS, I. 2015. Variability and trends of global atmospheric frontal activity and links with large-scale modes of variability. *Journal of Climate*, **28**(8), 3311–3330.

- RULFOVÁ, Z., & KYSELÝ, J. 2013. Disaggregating convective and stratiform precipitation from station weather data. *Atmospheric Research*, **134**(0), 100 – 115.
- SANTOS, J.A., REIS, M.A., DE PABLO, F., RIVAS-SORIANO, L., & LEITE, S.M. 2013. Forcing factors of cloud-to-ground lightning over Iberia: regional-scale assessments. *Natural Hazards and Earth System Science*, **13**(7), 1745–1758.
- SEMMLER, T., & JACOB, D. 2004. Modeling extreme precipitation events—a climate change simulation for Europe. *Global and Planetary Change*, **44**(1), 119–127.
- SEREGINA, L.S., HAAS, R., BORN, K., & PINTO, J.G. 2014. Development of a wind gust model to estimate gust speeds and their return periods. *Tellus A*, **66**.
- SHOWALTER, A.K. 1953. A stability index for thunderstorm forecasting. *Bull. Am. Meteorol. Soc.*, 250–252.
- SIMMONDS, I., KEAY, K., & TRISTRAM BYE, J.A. 2011. Identification and Climatology of Southern Hemisphere Mobile Fronts in a Modern Reanalysis. *Journal of Climate*, **25**(6), 1945–1962.
- SMITH, DOUG M, CUSACK, STEPHEN, COLMAN, ANDREW W, FOLLAND, CHRIS K, HARRIS, GLEN R, & MURPHY, JAMES M. 2007. Improved surface temperature prediction for the coming decade from a global climate model. *science*, **317**(5839), 796–799.
- SOUSA, J.F., FRAGOSO, M., MENDES, S., CORTE-REAL, J., & SANTOS, J.A. 2013. Statistical–dynamical modeling of the cloud-to-ground lightning activity in Portugal. *Atmospheric Research*, **132–133**(0), 46 – 64.
- STEINACKER, R. 1977. Möglichkeiten von Gewitterprognosen im Gebirge. *Wetter Leben*, **29**, 150–156.
- THOMPSON, R.L., EDWARDS, R., HART, J.A., ELMORE, K.L., & MARKOWSKI, P. 2003. Close proximity soundings within supercell environments obtained from the Rapid Update Cycle. *Weather and Forecasting*, **18**(6), 1243–1261.
- TRAPP, R.J., ROBINSON, E.D., BALDWIN, M.E., DIFFENBAUGH, N.S., & SCHWEDLER, B.R.J. 2011. Regional climate of hazardous convective weather through high-resolution dynamical downscaling. *Climate dynamics*, **37**(3-4), 677–688.

- UIJLENHOET, R., & STRICKER, J.N.M. 1999. A consistent rainfall parameterization based on the exponential raindrop size distribution. *Journal of Hydrology*, **218**(3), 101–127.
- UPPALA, SAKARI M, KÅLLBERG, PW, SIMMONS, AJ, ANDRAE, U, BECHTOLD, V D, FIORINO, M, GIBSON, JK, HASELER, J, HERNANDEZ, A, KELLY, GA, *et al.* . 2005. The ERA-40 re-analysis. *Quarterly Journal of the Royal Meteorological Society*, **131**(612), 2961–3012.
- VAN DELDEN, A. 2001. The synoptic setting of thunderstorms in western Europe. *Atmospheric Research*, **56**(1–4), 89 – 110. Conference on European Tornadoes and Severe Storms.
- VAN DELDEN, A. 2001. The synoptic setting of thunderstorms in western Europe. *Atmospheric research*, **56**(1), 89–110.
- VENABLES, W.N., & RIPLEY, B.D. 2002. Random and mixed effects. *Pages 271–300 of: Modern Applied Statistics with S*. Springer.
- WALDVOGEL, A., FEDERER, B., & GRIMM, P. 1979. Criteria for the detection of hail cells. *Journal of Applied Meteorology*, **18**(12), 1521–1525.
- WALLACE, J.M., & HOBBS, P.V. 2006. *Atmospheric science: an introductory survey*. Vol. 92. Academic press.
- WEIGL, E., REICH, T., LAND, P., WAGNER, A., KOHLER, O., & GERLACH, N. 2004. Summarizing final report for the project duration from 1997 until 2004.
- WEIJENBORG, C., FRIEDERICHS, P., & HENSE, A. 2015. Organisation of potential vorticity on the mesoscale during deep moist convection. *Tellus A*, **67**(0).
- WEISMAN, M.L., & KLEMP, J.B. 1982. The dependence of numerically simulated convective storms on vertical wind shear and buoyancy. *Monthly Weather Review*, **110**(6), 504–520.
- WESTERMAYER, A.T., GROENEMEIJER, P., PISTOTNIK, G., SAUSEN, R., & FAUST, E. 2016. Identification of favorable environments for thunderstorms in reanalysis data. *Meteorologische Zeitschrift*, 07.
- WILBY, R.L., BARNESLEY, N., & O’HARE, G. 1995. Rainfall variability associated with Lamb weather types: the case for incorporating weather fronts. *International journal of climatology*, **15**(11), 1241–1252.

WMO. 1974. Manual on codes.

ZHOU, Y, QIE, X, & SOULA, S. 2002. A study of the relationship between cloud-to-ground lightning and precipitation in the convective weather system in China. *Pages 107–113 of: Annales Geophysicae*, vol. 20(1).

ZOLINA, O., KAPALA, A., SIMMER, C., & GULEV, S.K. 2004. Analysis of extreme precipitation over Europe from different reanalyses: a comparative assessment. *Global and Planetary Change*, **44**(1), 129–161.

ZOLINA, O., SIMMER, C., KAPALA, A., & GULEV, S. 2005. On the robustness of the estimates of centennial-scale variability in heavy precipitation from station data over Europe. *Geophysical Research Letters*, **32**(14).

Acknowledgements

I thank Prof. Dr. Michael Kerschgens giving me the opportunity to do my doctorate on this interesting topic. I wish to express my gratitude to PD Dr. Joaquim G. Pinto, who has been an excellent mentor for me from the start and who took over the supervision. I am grateful to Prof. Dr. Andreas H. Fink, who kindly assisted in mentoring me with his expertise on convective events. I would like to thank all three of them for their interest in this topic, many fruitful discussions, feedback and in general for their encouragement and support. I thank Prof. Dr. Yaping Shao for being the second reviewer of my thesis.

Special thanks to Dr. Patrick Ludwig, Dr. Melanie Karremann and in particular Dr. Mark Reyers for many discussions and their feedback during the last years. They provided many valuable suggestions, which contributed to the improvement of this thesis.

I would like to thank AON Benfield for providing me the opportunity to work in different projects and their financial support during the preparation of this thesis.

I am grateful to Tim Hewson, who kindly provided valuable suggestions for the development of the automatic front analysis.

The lightning data has been kindly provided by the nowcast GmbH. I would like to thank in particular Prof. Dr. Betz and Mr. Möhrlein for the possibility to use the data from 2007 to 2014. I thank the German Weather Service for providing the Radolan data, radiosounding data and for the opportunity to access the data server of the European Centre for Medium-Range Weather Forecasts (ECMWF), which kindly provided the ERA Interim data. I thank OGIMET for providing data of European synoptic stations. The

ESWD data has been kindly provided the Dr. Pieter Groenemeijer. The climate scenario runs and decadal hindcasts applied in this study were developed and provided by the MIKLIP project.

The German climate computing centre (DKRZ) in Hamburg and the University of Cologne's regional computing centre (RRZK) kindly provided the required computation and storage resources.

I thank my colleges within the research group for many valuable hints and for the nice and productive working atmosphere during the last couple of years.

I would also like to thank my parents and siblings for their love, encouragement and support. They were always there for me with their patience and understanding.

Erklärung zur Dissertation

Ich versichere, dass ich die von mir vorgelegte Dissertation selbstständig angefertigt, die benutzten Quellen und Hilfsmittel vollständig angegeben und die Stellen der Arbeit - einschliesslich Tabellen, Karten und Abbildungen -, die in anderen Werken im Wortlaut oder dem Sinn nach entnommen sind, in jedem Einzelfall als Entlehnung kenntlich gemacht habe; dass diese Dissertation noch keiner anderen Fakultät oder Universität zur Prüfung vorgelegen hat; dass sie - abgesehen von unten angegebenen Teilpublikationen - noch nicht veröffentlicht worden ist sowie, dass ich eine solche Veröffentlichung vor Abschluss der Promotion nicht vornehmen werde. Die Bestimmungen dieser Promotionsordnung sind mir bekannt. Die von mir vorgelegte Dissertation ist von PD Dr. Joaquim G. Pinto betreut worden.

Teilpublikationen liegen nicht vor.

Köln, den 3. September 2016

(Sven Ulbrich)