Data Analysis to Protect Against Climate-Driven Extremes

Elias Theis
Bergische Universität Wuppertal
Institute of Technologies and Management of the Digital Transformation
Submitted for the seminar "Selected Topics in Data-Science"
26.01.2025

***Abstract*—Extreme weather events, aggravated by climate change, have become increasingly frequent and severe, necessitating advanced preparedness strategies. Motivated by the flood disaster of 2021, this study examines precipitation radar data from the German Weather Service (DWD) to identify precipitation trends and develop methods to identify and localize extreme weather events. The analysis comes to the conclusion that not only heavy rainfall itself, but especially the accumulation of precipitation leads to weather extremes. However, the analysis of precipitation radar data is not sufficient to develop a robust, fast and accurate early warning system for extreme weather events.**

I. Introduction

Extreme weather events (WEs), such as floods and droughts, have become more frequent and severe due to the intensification of climate change [1]. These events pose significant risks to life, property, and infrastructure, underscoring the urgent need for effective methods of prediction and early warning. In this seminar paper, we aim to analyse long-term precipitation data from the German Weather Service (DWD) to uncover patterns and trends that can help identify and localize extreme weather events, with a particular focus on North Rhine-Westphalia (NRW). Throughout this study, the terms "extreme weather event" and "weather extreme" (WE) will be used interchangeably.

The motivation for this research arises from recent devastating weather events, such as the flood disaster of 2021, which caused immense damage in parts of Germany, particularly in NRW. While floods represent a prominent type of WE, this study recognizes that droughts are another critical concern aggravated by climate change. However, the scope of this analysis is limited to rainfall-related events, excluding droughts, to ensure a more focused and detailed investigation.

This paper is structured as follows: Section 2 describes the dataset and its characteristics, including spatial coverage and technical details. Section 3 outlines the methods used for the analysis, while Section 4 presents the results, focusing on precipitation trends and their relationship with WEs. Finally, Section 5 discusses the findings, highlighting their implications for early warning systems and offering recommendations for future work.

II. Dataset

The data set provided for this seminar paper comprises preprocessed precipitation radar data from the German Weather Service (DWD). It covers 19 years from 2006 to 2024 and shows the hourly precipitation over Germany with a resolution of about 1km per pixel (900×900 grid). The amount of precipitation is given in mm per m^2 per hour. The years 2023 and 2023 are represented in the data set with an increased resolution. In this period, precipitation is recorded every 10 minutes. As such a high resolution is not necessary for this analysis and in order to keep the data set homogeneous, the resolution of these years are reduced to that of the rest of the data set.

A. Spatial Coverage

The radar stations usually cover the whole of Germany (see figure 8 in the appendix). In the event of outages or other events, it is possible that part of Germany is not covered and no data is available for this area and period. These gaps are rare, but for the calculation, the corresponding area is excluded for the period.

B. Technical Details

The data comprises 9GB in compressed form and approx. 250GB in uncompressed form. As only 32GB of RAM is available for this analysis, it is not possible to load the entire dataset at once. Therefore, the data is loaded and processed in batches. The data set consists of 456 files. Each month consists of two files - one for precipitation and one for time information.

In order to optimally use the available system resources, several processes are started in parallel so that data is read from the hard disk and calculations are carried out simultaneously. In this way, the bottleneck, the hard disk, is used optimally.

The precipitation data is stored with a precision of 16 bits. For an exact calculation and to avoid overflow errors, all calculations are carried out with a precision of 64 bits.

C. Grid Systems

The dataset uses the *RADOLAN-RADVOR composite format description* [2] and is projected into the common *WSG84* (longitude and latitude) system for a meaningful visualisation. The distortion caused by the projection can be neglected here.

D. Further Data Sources

In addition to the precipitation radar dataset, this study uses two other data sources for the borders of Germany. The first polygon [3] describes the border of the Federal Republic of Germany and the second source [4] contains polygons for all 16 federal states. These boundaries are not only used for visualization, but also for the creation of masks. For example, a mask can be created and used to calculate the precipitation for the federal state of NRW.

III. Methods

The methodology for this study is divided into three main components: the analysis of precipitation trends, the detailed investigation of major weather extremes (WEs), and the evaluation of predictive approaches for these events. Each component is designed to address specific research objectives using the provided precipitation radar dataset.

A. Analysis of Precipitation Trends

The first step involves examining long-term precipitation trends to establish a foundational understanding of the dataset. Key metrics, such as annual and monthly precipitation averages, are calculated for the entirety of Germany and specifically for North Rhine-Westphalia (NRW). Visualizations, including line graphs and box plots, are used to highlight spatial and temporal variations in precipitation. These analyses provide insights into overall trends, such as seasonal rainfall patterns, annual fluctuations, and long-term changes in precipitation intensity or frequency.

B. Detailed Investigation of Major Weather Extremes

The second component focuses on identifying and analyzing significant weather extremes, particularly floods, that occurred during the dataset's time span (2006–2024). Extreme weather events are identified based on deviations from average precipitation levels and verified against documented events from external sources, such as historical records and news reports.

To better understand the relationship between precipitation and WEs, this analysis introduces the Cumulative Precipitation with Decay (CPD) metric. CPD models the accumulation of precipitation while factoring in runoff or decay over time. By applying this metric, the study explores whether accumulated precipitation can better capture the conditions leading to WEs compared to instantaneous rainfall levels. The discharge or decay parameter λ is set to 95% in the analyses. The recursive formula for calculating the CPD is defined as follows:

$$CPD_i = CPD_{i-1} \cdot \lambda + p_i \tag{1}$$

with

CPD_i	Cumulative Precipitation with Decay at time i
λ	decay factor (0.95)
p_i	precipitation in $\frac{mm}{m^2}$ at time i
CPD_0	first value (p_0)

The CPD is computed for both the average precipitation across Germany and for each 1-km pixel within the radar dataset. This pixel-wise computation enables spatial localization of extreme events and highlights regions particularly vulnerable to such phenomena.

C. Predictive Analysis of Weather Extremes

The final component evaluates the feasibility of predicting weather extremes based on the precipitation radar data. Using the CPD metric as a basis, the study examines its potential as a simple early indicator for WEs. Thresholds for CPD are defined to signal potential weather extremes, and their effectiveness is assessed by comparing predicted events against documented occurrences.

The limitations of relying solely on precipitation data for predictions are also addressed. Factors such as topography, river systems, and existing water levels, which play critical roles in the development of weather extremes, are acknowledged but not incorporated into the analysis due to data constraints. Recommendations for integrating these additional variables into future predictive models are discussed in the concluding sections of the paper.

IV. Results

A. Basic Precipitation Trends

First, simple precipitation trends are analysed using two figures. The first figure shows the annual rainfall (in mm per square metre) for Germany and NRW. These rainfall amounts are between 600mm and 1100mm. According to this data, 2007 was the year with the most precipitation and 2018 the year with the least. In most years, NRW records higher precipitation per square metre than Germany. Reasons for this are discussed in section V-A. Furthermore, a slight downward trend can be recognised, which includes an annual decrease of approx. 9mm. However, this trend has a Pearson correlation coefficient of 0.06, which indicates a very weak correlation and this observation could be a coincidence.

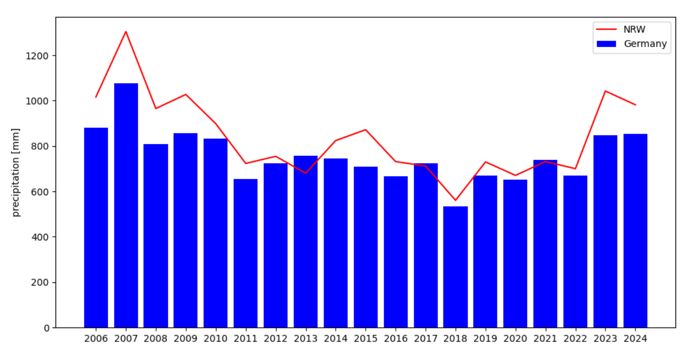

Fig. 1. Average yearly precipitation per m^2 in mm in Germany and NRW

Figure 2 shows the amount of precipitation broken down by month. The amount of precipitation for the whole of Germany is shown as a box plot so that the distribution can be recognised directly. The average amount of precipitation in NRW is also shown as a simple line in comparison.

Contrary to what one might expect, the rainiest months are in summer, with an average monthly rainfall of over 75mm. In the period from September to April, however, the amount of precipitation is relatively constant at around 50mm per square metre.

An important observation concerns the size of the area of these boxes. If we look at November, for example, it is noticeable that the amounts are between 10mm and 100mm. This is a very large range where the upper limit is 10 times the lower limit. This observation shows how strongly rainfall can fluctuate in Germany.

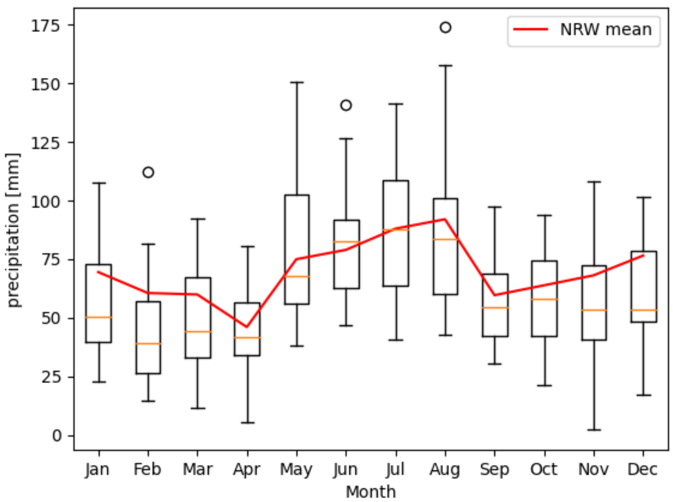

Fig. 2. Average monthly precipitation per m^2 in mm in Germany and NRW

The two previous graphs provide a good overview of the development of precipitation within a year and over the years. The next step is to analyse the rainfall distribution within Germany.

The 3. figure shows the average monthly precipitation as a heat map over Germany. Increased precipitation can only be seen in the north on the coast.

Almost all other areas with increased values can be traced back to the location of the radar measuring station. For example, the measuring station in Memmingen (city in Bavaria) can be easily recognised by a dark ring and a few outgoing rays. All measuring stations operated by the German Weather Service (DWD) can be seen in the figure in the appendix.

You can also see that there are several other small stations outside Germany. For all analyses, however, the data was limited to exactly the area within the borders of the Federal Republic of Germany.

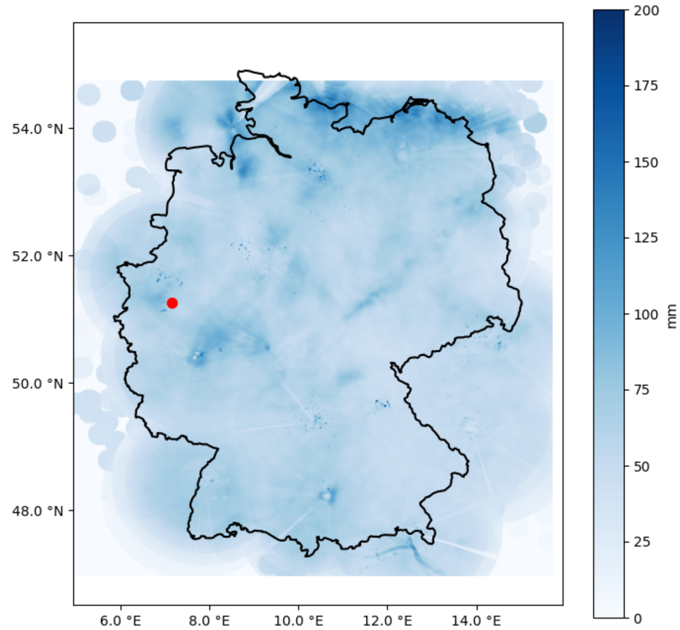

Fig. 3. Average monthly precipitation per m^2 in mm as a heatmap of germany. Wuppertal is highlighted in red.

The previous analyses relate to general trends and in the next section extreme weather events in the period covered by this data set are examined in more detail.

B. Investigate Extreme Weather Events

It is not trivial to define the term ‘extreme weather event’ or ‘weather extreme’ precisely. In each case, however, we are concerned with major deviations from the average with potential damage to property and personal injury. Thus, all weather extremes that are represented in lists from Wikipedia [5], as articles from popular news media or similar are compiled in the context of this work. A list of these events can be found in the appendix.

In order to analyse the relationship between precipitation and WEs, the daily precipitation in combination with all WEs is shown in the following figure. It quickly becomes apparent that the WEs are not very easy to recognise from the precipitation: *WEs are not always represented by very high precipitation and very high precipitation does not always lead to WEs*.

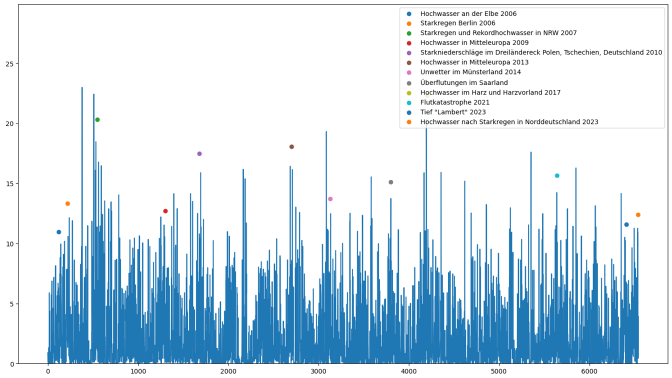

Fig. 4. Daily precipitation per m^2 in mm in Germany along weather extremes

These findings suggest that not only very heavy rainfall alone leads to WE, but that the accumulation of this leads to WEs. To test this hypothesis, the metric 'Cumulative Precipitation with Decay' (CPD) is introduced. This is intended to provide a method for modelling the accumulated precipitation subtracting a regular percentage runoff. The real accumulation of rainfall is much more complex, depends on many factors and is significantly influenced by lakes, rivers and reservoirs. A metric such as CPD cannot do justice to this and only attempts to depict these effects in a highly simplified manner.

Figure 5 shows the CPD values. As the values are relatively noisy and to make it easier to recognise WEs, the CPD values were squared after the calculation. Furthermore, three times the average was used as the threshold value for recognising WEs.

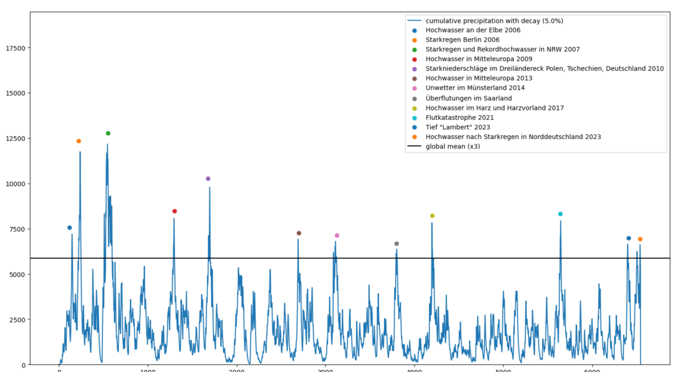

Fig. 5. Squared Cumulative Precipitation with Decay (CPD), as in (1) along weather extremes

This approach works quite well to recognise weather extremes based on precipitation radar data. The data previously used for this method was the average for the whole of Germany and does not indicate where in Germany a WE has occurred. This is changed in the following by calculating the CPD values for each pixel in parallel and independently of each other. The material accompanying this paper includes an animation of these values over time. Figure 6 shows the temporal average of this calculation. This map basically shows two different areas that are particularly frequently affected by WEs.

The first area covers all regions in the south of Germany that are approximately 800 metres or higher above sea level. It is already known that there is significantly more precipitation in the Alps [6]. Consequently, WEs also occur more frequently in these regions.

The second area that is more affected is in NRW. It is difficult to find the precise reasons why exactly this area is so much more affected by WEs than the rest of Germany.

It is also noticeable that the area of the former GDR is visibly less affected by WEs than the rest of Germany. This observation that the federal states of Mecklenburg-Western Pomerania, Brandenburg, Berlin, Saxony-Anhalt, Saxony and Thuringia are less affected is probably also due to chance.

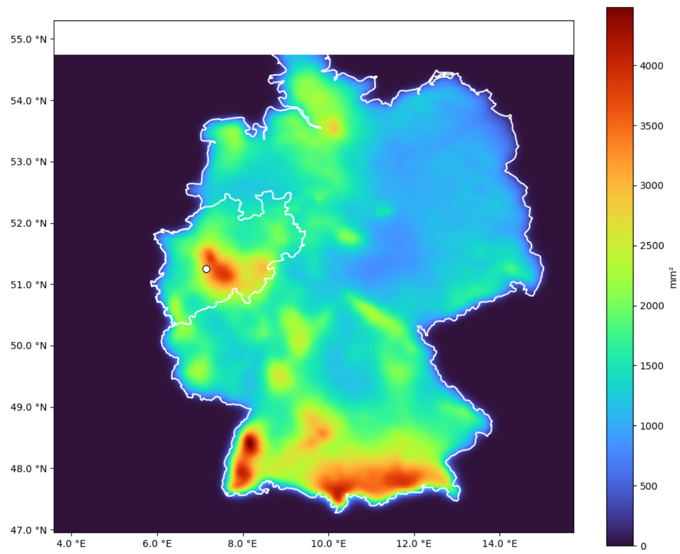

Fig. 6. Temporal mean of squared Cumulative Precipitation with Decay (CPD), as in (1) as a heatmap of germany. Wuppertal and the border of NRW are highlighted in white.

This form of visualisation also allows the areas affected by the 2021 flood disaster to be depicted relatively well. However, it should be noted that this method does not take into account many important factors such as topography and rivers. Nevertheless, Figure 7 shows the CPD values at the peak of the flood disaster.

This map shows that not only NRW is severely affected, but also large parts of Baden-Württemberg and Bavaria.

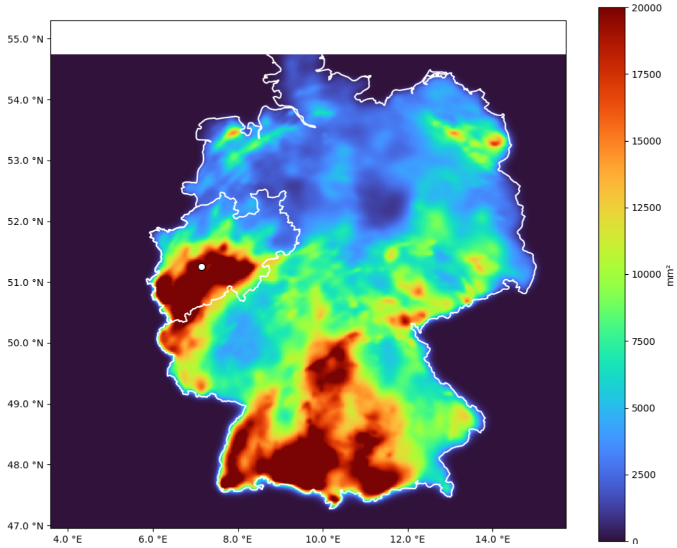

Fig. 7. Squared Cumulative Precipitation with Decay (CPD), as in (1) as a heatmap of germany on the 15.07.2021. Wuppertal and the border of NRW are highlighted in white.

V. Discussion

A. Increased Precipitation in NRW

On average, it rains more per square metre in North Rhine-Westphalia (NRW) than the German average because the geographical and meteorological conditions in NRW are strongly influenced by differences in altitude and specific weather conditions. One key factor is the *windward and leeward effect* of the low mountain ranges. In regions such as the Sauerland, Siegerland and Bergisches Land, moist air masses from westerly and south-westerly directions accumulate on the slopes facing the wind (windward). These are forced to rise, cool down and lead to increased cloud formation and precipitation. The prevailing westerly to south-westerly weather conditions bring humid air masses from the Atlantic, which meet the low mountain ranges in NRW and lead to the effects described above. [7]

B. Building an Early Warning System

In this section, the extent to which these analyses can help to develop an early warning system for extreme weather events is examined in more detail. Although CPD offers a simple and fast method to model the accumulation of precipitation in a highly simplified way and to recognise potential WEs, this approach lacks a great deal of information to correctly map the entire situation. For example, topography plays a very important role. If more rain falls than the ground can absorb, this water collects and flows in the direction of steepest descent. Accordingly, areas located in valleys are significantly more affected than those located on mountains. Another important factor concerns different bodies of water. For example, the filling levels of reservoirs and the water levels of rivers can strongly influence the consequences of heavy rain. Even snow that has already fallen and will only melt at a later date is not taken into account here. If all this information were combined with precipitation and forecasts of precipitation, it would probably be possible to develop an early warning system. Without this information, however, an early warning system would be too slow and inaccurate.

VI. References

References

[1] IPCC, “Beitrag von arbeitsgruppe i: Naturwissenschaftliche grundlagen.” https://www.de-ipcc.de/media/content/Hauptaussagen_AR6-WGI.pdf.

[2] D. Wetterdienst, “Radolan-radvor-kompositformatbeschreibung.” https://www.dwd.de/DE/leistungen/radolan/radolan_info/radolan_radvor_op_komposit_format_pdf, 2024.

[3] SimpleMaps.com, “Germany - simplemaps.com.” https://simplemaps.com/static/svg/country/de/all/de.json.

[4] SimpleMaps.com, “Germany federal states - simplemaps.com.” https://simplemaps.com/static/svg/country/de/admin1/de.json.

[5] wikipedia.org, “Liste von hochwasser-ereignissen - wikipedia.” https://de.wikipedia.org/wiki/Liste_von_Hochwasser-Ereignissen.

[6] C. Wastl, “Klimatologische Analyse von orographisch beeinflussten Niederschlagsstrukturen im Alpenraum,”

[7] U. u. V. N.-W. Landesamt für Natur, “LANUV-Fachbericht 27,”

[8] D. Wetterdienst, “Radar - network of deutscher wetterdienst.” https://www.dwd.de/EN/research/weatherforecasting/num_modelling/bilder/02_datenassimilation_abb09_en.html, 2024.

VII. APPENDIX

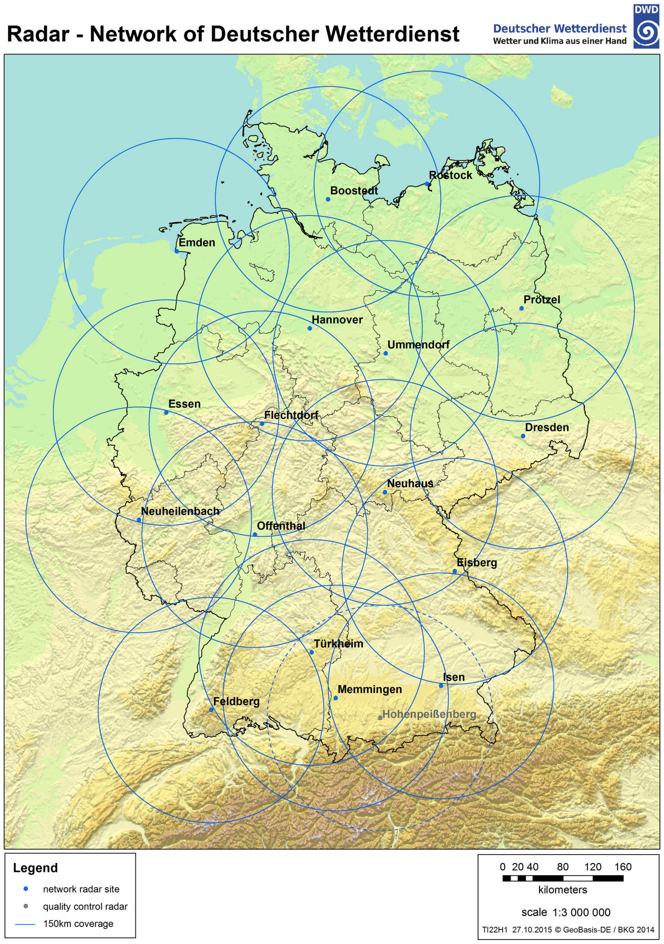

Fig. 8. Radar - Network of Deutscher Wetterdienst [8]

Time	**Event**	**Source / Reference**
April 2006	Floods on the Elbe	https://earth.esa.int/eogateway/gallery/flood-in-europe-germany-2006
August 2006	Heavy rain in Berlin	https://www.faz.net/aktuell/sport/fussball-wm-2006/deutschland-und-die-wm/unwetter-berlin-im-ausnahmezustand-1357374.html
August 2007	Heavy rain in North Rhine-Westphalia	https://wetterkanal.kachelmannwetter.com/august-2007-starkregen-und-rekordhochwasser-in-nrw
June 2009	Floods in Central Europe	https://de.wikipedia.org/wiki/Hochwasser_in_Mitteleuropa_2009#Deutschland
August 2010	Heavy precipitation in the border triangle Poland, Czech Republic, Germany	https://www.dwd.de/DE/leistungen/besondereereignisse/niederschlag/20100901_sachsen_hochwasser.pdf
June 2013	Floods in Central Europe	https://de.wikipedia.org/wiki/Hochwasser_in_Mitteleuropa_2013#Deutschland
July 2014	Storm in the Münsterland	https://www1.wdr.de/nachrichten/westfalen-lippe/uebersicht-hochwasser-muenster-100.html
May 2016	Flooding in the Saarland	https://www.gdv.de/resource/blob/22200/7e39c831f9f98a7ad11057600122ac3d/publikation-naturgefahrenreport-2016-data.pdf
July 2017	Floods in the Harz and Harz foreland	https://de.wikipedia.org/wiki/Hochwasser_im_Harz_und_Harzvorland_(2017)
July 2021	**Floods in Western and Central Europe** (Motivation of the associated project)	https://de.wikipedia.org/wiki/Hochwasser_in_West-_und_Mitteleuropa_2021
June 2023	Heavy rainfall events in western and central German regions	https://www.dwd.de/DE/Home/_functions/aktuelles/2023/20230627_starkniederschlaege_tief-lambert.html
December 2023	Floods in Lower Saxony	https://www.umwelt.niedersachsen.de/startseite/aktuelles/aktuelles-hochwasser-in-niedersachsen-228435.html

TABLE I
LIST OF MAJOR PRECIPITATION-RELATED WEATHER EXTREMES IN GERMANY BETWEEN 2006 AND 2024