Reply to Reviewer 1 for manuscript "Interannual variations in the seasonal cyle of extreme precipitation in Germany and the response to climate change"

Madlen Peter, Henning W. Rust and Uwe Ulbrich
Institute of Meteorology, Freie Universität Berlin, Germany

August 30, 2023

Reviewer 1: This manuscript presents an analysis of the seasonal and interannual variations of extreme precipitation at stations in Germany, employing a non-stationary block maxima approach. Additionally, it investigates the impact of climate change on the seasonal cycle of extreme precipitation in Germany, which is a crucial topic in climate change research. The paper is well-structured and complemented by visually appealing figures. However, there are several issues that require attention and improvement before this work can be considered for publication.

Main comments

- **Reviewer 1:** For introduction, according to the objective of the paper, it is important to address what previous studies have specifically accomplished, identify the existing gap or problem in the research, and emphasize why this problem is of significant concern. It is crucial to provide clarity on these aspects before describing the approach or research you intend to use in your study. For example, the third paragraph of the introduction discusses previous analyses conducted on extreme precipitation in Germany across different seasons. "Analyses of extreme precipitation in Germany for different seasons has already been done (Zolina et al., 2008; Łupikasza, 2017; Fischer et al., 2018; Zeder and Fischer, 2020; Ulrich et al., 2021)." More details of what previous studies have done are needed before you introduce the two main new aspects you will do in this study. In addition, the second question is "RQ2 How important is a flexible shape parameter to reflect recorded variations?". However, you did not add any descriptions or previous studies about shape parameter in the introduction. Therefore, My suggestion is to rewrite the introduction.

Answer: We have reworked the introduction and added some detailed information about previous studies: “Zolina et al. (2008) and Łupikasza (2017) analysed quantiles of daily precipitation sums separately for the seasons DJF, MAM, JJA and SON, while Fischer et al. (2018, 2019) used available data more efficiently by modelling monthly maxima of daily precipitation sums for all months simultaneously. This approach has been proven to lead to more robust and reliable results than considering months separately. Ulrich et al. (2021) extended this method by including different durations to efficiently estimate intensity-duration-frequency curves. Furthermore, Zeder and Fischer (2020) analysed the effect of climate change on seasonal extreme precipitation and found a positive connection to the north-hemispheric temperature rise. In our approach we combine the simultaneous modelling of available data for all months with interannual variations, thus accounting for potential changes of the seasonality due to climate change and natural variability.”

Additionally, we have added some sentences about the importance of the shape parameter: “The goal of this paper is to assess the performance of the seasonal-interannual modelling with a special attention to a flexible shape parameter ξ. This parameter is difficult to estimate as it interferes with the scale parameter (Ribereau et al., 2011) and requires long records for reliable results (Papalexiou and Koutsoyiannis, 2013). Nevertheless, it describes the behaviour of the very rare events and consequently plays an important role for assessing extreme precipitation changes.”

- **Reviewer 1:** For method: return level, the return period T can be written as $T = \mu/(1-p)$, where, p is the non-exceedance probability. μ is the mean interarrival time between two successive events, which is defined as one divided by the number extreme events per year. When considering annual maxima, μ corresponds to 1 year. However, in your study, when calculating the return period T, are we utilizing the annual or monthly maximum or non-exceedance probabilities? If we are using the monthly maximum time series or non-exceedance probabilities, μ should not be equal to 1.

  Answer: We calculate the January (or February or...) return levels expressed in frequency per January (or February or ...), for example, the one in ten Januarys return level (10-January return level). This should not be mistaken with the annual return levels obtained with annual maxima. We have added a paragraph to Section 3.5 to clarifiy which return levels are considered in our manuscript: ”As we consider monthly maxima we calculate as well monthly return levels. Similar to e.g. 100-year return levels obtained with annual maxima, we determine the 100-January return levels, the 100-February return levels, and so on. In the following we state them as monthly 100-year return levels instead of naming respective months. This should not be confused with annual return levels. However, they can be calculated as well with monthly maxima leading to more accurate and reliable annual results (Maraun et al., 2009; Fischer et al., 2018)."

- **Reviewer 1:** In addition, when applying the GEV to the monthly maximum, if two extreme events occur on the last day of the month and the first day of the next month, these two events are often treated as a single individual event. When applying the GEV to non-exceedance probabilities, precipitation occurrences are highly clustered in time and space. Therefore, the independence of the extreme values should be taken into account prior to modeling.

  Answer: Indeed, it could happen that precipitation maxima of two successive months belong to the same precipitation event and maxima are not completely independent in time. However, for our dataset, this is the case for only about 0.6% of the data. Here, the temporal dependence is neglected and independence is assumed. We added to Section 3.1 the sentence: "This requires independent block maxima of successive months. However, this assumption can be violated if two monthly maxima belong to the same precipitation event, e.g. if one maximum occurs at the end of the month and the second one at the beginning of the next month. For the given records, about 0.6% of the monthly maxima have been registered at successive days. Since this fraction is low, we neglect temporal dependances and assume independent monthly maxima."

- **Reviewer 1:** The fitted return period distribution may exhibit uncertainties due to the limited sample size of the data. The short time period of the datasets may introduce uncertainty in the distribution model fitting. Therefore, for the question "RQ1: Can a model with interannual variations better represent the observations than a seasonal-only model?" how can you distinguish the difference or bias from the uncertainty in distribution model fitting or from the model with or without interannual variations? As shown in the paper "the total QSS for different non-exceedance probabilities (return periods). Skill is positive but small<= 2%, increasing with non-exceedance probability (return period)." The larger bias for higher return period is very likely caused by large uncertainties for higher return period in model fitting.

  Answer: Indeed, quantile estimates for higher non-exceedance probabilites are related with higher uncertainties. This is also reflected in the QSS as you stated. Thus, we have mentioned that evaluations for return levels with a return period larger than the observation period need to be treaten cautiously. We have written in the manuscript: "The latter has to be interpreted with care as there are very few observations in the range of the upper quantiles. Return levels with a return period higher than the time range of the data should be treated cautiously, since the quantile score can not reasonably evaluate those values (Fauer and Rust, 2023)." Here, this is only the case for non-exceedance probabilities of 0.99 and 0.995 (return period of 100 and 200 years). We have added to the text:

"As we consider for each station at least 80 years of observations, this only matters for non-exceedance probabilities (return periods) of 0.99 and 0.995 (100 and 200 years).". Uncertainties exist in both models, a seasonal-interannual model and a seasonal-only model. However, we have analysed in a cross-validated approach, that the interannually varying model is beneficial. Here, the outcome "Skill is positive but small $\lesssim 2\%\%$" refers to the overall investigation summarising the skill for all months and stations. However, skill is more pronounced for single stations and different months, which is illustrated in more detail in Fig. 7, Fig. 8 and Fig. 10.

Other comments

- **Reviewer 1:** Lines 21-25: Please maintain consistency in the usage of terminology such as 'heavy rain' or 'heavy precipitation' throughout the entire

  Answer: Many thanks for the hint. We have changed it to 'heavy precipitation' for the entire document.

- **Reviewer 1:** Line 77-79, "The four stations Krümmel (1899-01-01 until 2021-12-31), Mühlhausen / Oberpfalz-Weiherdorf (1931-01-01 until 2021-12-31), Rain am Lech (1899-01-01 until 2021-12-31) and Wesertal-Lippoldsberg (1931-01-01 until 2021-12-31) are highlighted in Fig. 1 and will be discussed exemplarily in this study". Why do you choose these four stations? It would be beneficial to include a brief introduction explaining the reasons behind selecting these stations. Although you provide more details about the stepwise selection process in Section 4.1, adding an introductory explanation would provide context for the readers.

  Answer: We have added the sentence: "We have selected these stations as they are characterised by different changes in seasonality (see Sec. 7) represented by divergent model setups (see Sec. 4.1). Additionally, their interannual changes are more pronounced than for other stations."

- **Reviewer 1:** The sample size of the data in model fitting. For example, for figure 7 and figure 8, the GEV was applied to each station, especially for each month of each station, how many extreme values (sample) are you collected for each station for each month? Are the number of samples enough for distribution model fitting?

  Answer: Since we consider stations with at least 80 years of observations, there should be at least 80 data points for each station and month. However, missing values within the observation period are allowed such that less than 80 data points might be available. The minimum number of maxima for one month and station is 78. We have checked exemplarily the stationary GEV for this month by using model diagnose (qq-plot, pp-plot, return level plot and histogram), revealing that a sample size of 78 data points is sufficiently large enough for model fitting.

- **Reviewer 1:** Figure 12, it is better to add a,b,c,d into the each figure.

  Answer: added

- **Reviewer 1:** In Section 7, "Impact of climate change on the seasonality of extreme precipitation," it is important to note that the trend of the time series is significantly influenced by the chosen start year. Although the study mentions comparing the time period from 1941 to 2021, where all stations have data, the start year appears to be different in Figure 15. Could you please provide further clarification on the discrepancy?

  Answer: We describe the problem of calculating linear trends for fixed time periods in detail in the first paragraph of Section 7 and in Appendix A. To highlight that this is not done for the analyses of the example stations, we have added to the paragraph "We compare the time period from 1941 to 2021 where all stations have data. Note that estimating linear trends for fixed (and short) periods of time can yield very different results depending on the considered time period due to decadal variability. Thus, the trend estimates presented here for the given time period serve as a rough indicator for climate change effects; for a more detailed analysis the whole datasets should be taken into account for each station" the following part: ", which is done in Fig. 15.".

- **Reviewer 1:** In section 7, the return period was calculated for each station for each year? is the sample size enough for model fitting?

  Answer: The non-stationary model, which considers maxima of all months and years simultaneously, allows for varying return within the year and throughout the years. Since we only consider stations with at least 80 years of observations, a minimum of 960 data points (80 times 12) are available for each station which is sufficient for model fitting. Due to the interannual component in modelling we are able to estimate different return levels for each year.