Dear Reviewer,

thank you for your comments and suggestions. Please find below the answers on the comments you have made.

- The Introduction does not clearly present what the current research gaps are. It indicates numerous studies performed in the past, but the novel contribution of this manuscript is not clearly formulated.

  In the updated manuscript we will reformat the introduction and give a clear overview on the research gaps and on the novel contributions of this paper. Below we list the novelty contribution from our paper:

  - The German Weather Service has a rather dense network of long observations (~60-70 years) and high temporal resolution (1min), which gives us an excellent opportunity to compare and validate different methods for the estimation and regionalisation of the DDF curves.
  - Typically, studies either use the Koutsoyiannis approach or multi/simple scaling approach of GEV parameters to generalise the extremes over different durations. To our knowledge there is no comparison of the two approaches in the literature. The Koutsoyiannis approach has been implemented in Germany by Ulrich et al. 2020, but on a shorter available 1 min dataset (up to 14 years), while Fischer and Schumann (2018) have implemented the multi scale approach only at a long station (~85 years). Other applications of these methods, are either in daily timeseries, or hourly but short time series, and hence it becomes interesting to investigate which of the methods is more appropriate when a long and high-resolution network is available.
  - The same is true also for the regionalisation approaches: to our knowledge there is no comparison between kriging and index-based regionalisation. Although many studies implement them, it is either one method implemented or the other (this is valid also for other regression models). Naturally, it is interesting to see which of the methods is more appropriate when validated on a long and high-resolution network, and where lie the advantages and disadvantages of each method.
  - One of the main novelties that we include different networks, and investigate how to integrate them the best, while validating on a long and high-resolution network. This integration is performed at various points like for instance: the long automatic and short digital network are homogenised by considering sensor changes, the daily network is disaggregated based on radar parametrization, all extreme series are described by a fixed shape parameter (taking into the account knowledge from literature), and lastly, we integrate the dataset differently by applying for regionalisation ordinary kriging or external drift kriging. There is enough evidence in the literature that the GEV parameters of long and short series are different from one another, however the regionalisation (either by index or kriging) is done with all available data pooled together. The clear message from this paper is that the information from long series is more important, and should have other weights than the short series.

  Following your suggestion, we will reformat the introduction and literature review, to point better the research gaps and the novelty mention above.

- Importantly, I missed a discussion section that links the results with previous findings and shows how this study goes beyond the previous research and what the main scientific novelty is.

yes, we will include a section at the end of the results to discuss about the previous research and the outcome from our research. The discussion will be focused at these points:

- The three different data sets implemented here, distinguish from one another based on the parameter values (as shown in Appendix A3) (also on the spatial dependency, variograms, that will be included in the updated version of the manuscript). When fixing the shape parameter to 0.1, the location and Koutsoyiannis parameters of LS and SS, are in similar range, and the main difference is seen at the scale parameter (where the SS has high values of the scale parameter than LS). This gives a tendency of the short durations to estimate bigger rainfall volumes for higher return periods. Typically, this is treated by index-based regionalisation, where extremes within a region are pooled together to estimate the DDF curves at an unknown location. However, we show here that integrating the LS and SS with external drift kriging, hence accounting for the spatial dependency of the extremes, delivers better performance than grouping them together in the index-based regionalisation (also valid for the LS and DS integration).
- We have fixed the shape parameter to 0.1 for whole Germany, which is very similar to results obtained by Ulrich et al. 2021 (shape parameter as 0.11 from the annual GEV approach) and validates our approach. The spatial distribution of $\eta$ parameter follows the spatial structure of the annual rainfall sum in Germany, the distribution of the location ($\mu$) parameter follows the information from the elevation, while the scale ($\sigma$) and $\theta$ parameter don't seem to be influences by any climatologic or site characteristic. This is also seen at Van der Vyver 2012, where annual rainfall and elevation is concluded as important covariates, mainly for the location ($\mu$) parameter, while the scale ($\sigma$) parameter didn't have meaningful covariates and the shape parameter didn't show any spatial structure but was kept constant over Belgium. These results agree to a certain extend with the results obtained here. However, here we consider rainfall statistics extracted from short network or daily network as more important than the annual rainfall (which itself is an interpolation from point observation). Thus, interpolation of long datasets, should include extreme statistics from short or daily network rather than annual rainfall as an additional information.
- In the existing KOSTRA maps, all durations are dependent on elevation. Here, the elevation itself didn't show much effect on the scale ($\sigma$) and $\theta$ parameter, only to some extend on the location ($\mu$) and $\eta$ parameter. This means that the extremes of longer duration (affected by the $\eta$ parameter) and of low return period (affected by the location parameter) will show a pattern resembling the elevation. This is not true for short durations (affected by the $\theta$ parameter) and high return periods (affected by the scale parameter). This as well agrees with other studies, that report a weak dependence of short duration rainfall (shorter than 1 or 2 hours) with the elevation (Langefeld et al. 2019).
- Lastly, the kriging interpolation, open the possibility to capture better the uncertainty – not only the sample uncertainty which is typically done by bootstrapping the points statistics, but accounting as well the spatial structure of extremes by considering spatial simulations. This results in estimates that will be more precise near to the location of long time series, and less precise in regions far from long time series. We are actually working on a second manuscript that investigates these uncertainties, and will be submit it soon for publication in HESS.

- The study presents a large number of evaluations and results and uses a very dense regional dataset. Some more effort to link the results with regional characteristics and hydrologic and climate processes can add more significance to the results and add a new dimension to the demonstration of the generality of findings.

  We will try to discuss more in the updated version of the manuscript, the results and the suitability of the methods on some regional and climate characteristics of Germany.

References:


Fischer, S. and Schumann, A. H. 2018: Berücksichtigung von Starkregen in der Niederschlagsstatistik, Hydrol. und Wasserbewirtschaftung, 62(4), 221–240, doi:10.5675/HyWa.

Koutsoyiannis, D., Kozonis, D., and Manetas, A. 1998: A mathematical framework for studying rainfall intensity- duration-frequency relationships, J. Hydrol., 206, 118–135, https://doi.org/10.1016/S0022-1694(98)00097-3.

Lengefeld, K., Winterrath, T., Junghänel, T., Hafer, M., Becker, A. 2019: Characteristics spatial extent of hourly and daily precipitation events in Germany derived from 16 years of radar data. Meteorol. Z. Vol. 28, No. 5, 363-378. DOI 10.1127/metz/2019/0964

Van de Vyver, H. 2012: Spatial regression models for extreme precipitation in Belgium. Water Reources Research, Vol. 48, W09549, doi:10.1029/2011WR011707

Ulrich, J., Jurado, O. E., Peter, M., Scheibel, M., and Rust, H. W. 2020: Estimating IDF Curves Consistently over Durations with Spatial Covariates, Water, 12, 3119, https://doi.org/10.3390/w12113119.

Ulrich, J., Fauer, F. S., Rust, H. W. 2021: Modeling seasonal variations of extreme rainfall on different timescales in Germany. Hydr. Earth Syst. Sci. 25, 6133-6149, https://doi.org/10.5194/hess-25-6133-2021