

General comments

This work deals with the characterization of uncertainty in regionalized design rainfall curves, (so-called depth-duration-frequency (DDF) curves) constructed for the entire Germany in an accompanying work by a larger research group, also involving the authors of the present paper. The former paper is currently under review for the same journal (Shehu et al., 2022). A lot of choices made in the current paper are dictated by the way the subject was approached in the first paper and, to a certain extent, constrained by it. Although the first paper is currently still under review, I find the approach carried out in it reasonable and I will limit my review to how uncertainty in these choices was handled, not intending to extend my review to both works.

This paper focuses on identifying and exploring various components of uncertainty present in the regionalized design rainfall curves, with a particular emphasis on the spatial structure of uncertainty. This is an important, challenging and under-explored research subject and the efforts that the authors have undertaken to approach it are considerable and worth of appraisal.

The paper is very well-written, the results are discussed thoroughly and the authors strive to provide insights into a challenging subject. A great amount of computational work is obviously involved. I wish to congratulate the authors for these efforts and I suggest publication in HESS upon some revisions/clarifications that can be addressed by minor alternations/additions to the manuscript. I discuss these below along with some suggestions from improvements that might be found useful.

Specific comments

-On uncertainty of DDF parameters (referred to as 'local' uncertainty)

By the investigation carried out in the first paper, the authors conclude on using a common, fixed value of the shape parameter, based on the literature for return periods up to 100 years (in particular, following Koutsoyiannis' work (2004)). Using a common value of the shape parameter is a reasonable choice given the high uncertainty involved in its estimation from single stations. However, fixing the value of the shape parameter based on the literature, instead of estimating it, is not devoid of uncertainty either. It may not be straightforward to assess the uncertainty of a fixed parameter, but on the other hand, neglecting the shape parameter from the uncertainty assessment is a limitation that I think should be discussed. Perhaps the authors could add a brief discussion of the uncertainty associated with the shape parameter, and comment on its expected impact on local uncertainty, which presumably would be exacerbated if this was also accounted for.

-On regionalization uncertainty

The quantification of the spatial uncertainty presents the greatest challenge and also, the major contribution of the work. The authors could elaborate more on the reasons why they employ

spatial simulations for assessing regionalization uncertainty, since this is the only component of the uncertainty analysis for which simulation is employed. For instance, Lines 87-90 could be better explained; the reasons that the authors do not employ kriging for spatial uncertainty are less clear to someone not familiar with the provided geostatistical literature. In particular, given that they consider kriging variance to be a measure of uncertainty for the unobserved locations (Line 88), why is it later claimed to be only efficient for the local (I understand this as 'at-station') uncertainty and not the spatial one? The question also arises since the Sequential Gaussian Simulation that the authors finally choose for the assessment of spatial uncertainty is based on the kriging mean and variance. Do the authors use spatial simulations mainly because they seek to produce richer spatial patterns than the ones obtained by a kriging interpolation? A brief explanation of the terms 'local' and 'spatial uncertainty' would also be helpful to the reader at this point (a definition is provided later on, in Section 3).

My understanding is that in the parts of Experiments 3-5 that involve spatial simulations, the authors produce prediction intervals, which are subtly different than the confidence intervals obtained by the bootstrapping procedure followed for the assessment of local and variogram uncertainty. Could the authors comment on this difference and clarify, if appropriate?

Line 306-307: Could the authors explain the rationale for performing the conditional simulation, i.e. maintaining in the simulation a set of observed values for the long stations, and also, explain how this set is chosen?

-On the comparison among different components of uncertainty

The authors find that 'the spatial simulations add to the regionalisation the biggest uncertainty', and based on that conclude that 'the spatial uncertainty is the main source of uncertainty when regionalising the DDF curves'. Although this finding is reasonable and supported by the results presented, I also deem the estimation method important in the context of the inter-comparison of the different uncertainty intervals. Perhaps, this might not be the determining factor, but still could the observed difference between the spatial uncertainty and the other sources be, in part, due to the fact that (spatial) simulation produces wider intervals than bootstrapping?

-Conclusions

From the large experience gained through this work, the authors could perhaps comment on whether/how the observed uncertainty could be constrained by targeted data collection. For instance, what do they think should be prioritized in terms of data collection, and could there be an added benefit from incorporating gridded precipitation/satellite products in their framework?

Technical/minor comments

Line 59: The phrase 'non-representativeness of point-measures' appears quite vague; I suggest adding a short explanation of what is meant by it.

Line 188: Typo inside the parenthesis, μ should be replaced with η .

Line 196: Do the authors mean a 5 km x 5 km grid here?

Line 201: I suggest '...have an inadequate length for' instead of '...too little observation years for'.

Line 208: Probably 'deviate' is meant instead of 'denote' here.

Line 222: 'method' instead of 'moment'.

Line 297: Could the authors justify the choice of 133 stations here?

Line 365: 'fixing' instead of 'fixating'.

Line 508-509: Could the authors clarify what is implied here for the behavior of the Koutsoyiannis' parameters? Does the explanation provided in Lines 529-530 apply here as well?

Figure 12: Please check subplots' numbering and consider replacing 'volume' with 'depth'.

Figure 13: 'with' instead of 'will'.

Please check numbering of the sections; there are two sections '2.1'.

With best regards,

Theano Iliopoulou

References

Shehu, B., Willems, W., Stockel, H., Thiele, L.-B. and Haberlandt, U.: Regionalisation of Rainfall Depth-Duration-Frequency curves⁷⁸⁴ in Germany, *Hydrol. Earth Syst. Sci.*, [preprint], 2022.

Koutsoyiannis, D.: Statistics of extremes and estimation of extreme rainfall: II. Empirical investigation of long rainfall records, *Hydrol. Sci. J.*, 49(4), 591–610, doi:10.1623/hysj.49.4.591.54424, 2004b