Second response to comments for egusphere-2024-1463 ‘Downscaling the probability of heavy rainfall over the Nordic countries’

Dear editor, We are grateful for the opportunity to respond to the second round of reviews, and we think the discussion now is more on semantics and peripheral details, rather than methods or results being flawed or misleading. It is nevertheless a good opportunity to have a scientific debate about such issues. From our discussion of this manuscript, it appears that it is indeed a contribution to a multi-disciplinary approach that broadens the hydrological perspective and provides a cross-fertilization across disciplinary boundaries. We believe that downscaling is an important concept for the part of the hydrological community who is concerned with climate change, connecting small-scale processes to the global-scale processes underpinning hydro-climatology. Hence, it’s worthwhile elaborating on the topic of downscaling. Our results are also relevant for understanding floods, droughts, desertification, land degradation, eutrophication, and other aspects of global change. However, it is beyond the scope of this paper to teach downscaling from scratch, and we refer those readers who want to learn more about downscaling to an appropriate introductory text (Oxford Research Encyclopedia) and text books (Benestad et al., 2008; Maraun and Widmann, 2018) (see lines 103-105 in our revised manuscript). Rather than providing another introduction to downscaling, our paper presents state-of-the-art within empirical-statistical downscaling, building on previous work, and we follow the scientific tradition where our work continues where previous work stopped. It’s not common to start from scratch for every new step published in a scientific paper.

Anonymous referee #1

The reply letter of the authors is very general and there is no point-by-point response to my comments. The reply letter has only addressed 2 points of my previous review, which are 1) the definition of the term “downscaling” and 2) the fact that the manuscript was too reliant on supplementary material and therefore not self-contained.

We provided much of the points in the revised paper itself because we wanted to explain them to any reader who questioned downscaling. We also mistakenly thought it was obvious that the structure of our manuscript had changed so that the reader doesn’t have to read the supplementary material to understand our work. We nevertheless keep the supplementary material to provide transparency and make it possible for anyone to replicate our work, even if this is not standard practice. Below we provide specific details about our response and refer to the lines in our manuscript where they have been addressed. Also, we have added a track-changes version where it is possible to see the revisions.

The points brought up in the first report (DOI:10.5194/egusphere-2024-1463-RC1) beyond those addressed in our reply letter were addressed in the interactive discussion (DOI:10.5194/egusphere-2024-1463-AC1), and we thought it made no sense repeating them when wrapping up the discussion. We hope that we managed to address them in an acceptable way and expected an ongoing interactive discussion if that were not the case.

For addressing points 1) and 2), the authors have added an extensive discussion in appendix A. But let's face it: people rarely read appendixes. Thus, while I find this appendix useful, it is also in my opinion convoluted, and mostly focused on what the authors are not doing (estimating values on a grid with smaller pixels) rather than on what they are doing (estimating statistics at the stations locations). While both approaches are worthy of investigation and legitimate, I would like to see a clear statement in the body of the manuscript itself, of what is actually done in the article. Essentially, if the reviewers are confused on the nature of the work done, readers will be confused as well (here the authors have argued that the reviewers are not familiar enough with the community, but this argument is invalid because reviewers are supposed to be representative of the readers).

Thanks for this comment. We are surprised by the point raised on the particular definition of the term downscaling, and we elaborated the concept of downscaling in Appendix A in our revised manuscript. It may be incomplete as we missed one piece of information when it came to good references to the definition of downscaling being "estimating values on a grid with smaller pixels". We have been involved in the WCRP CORDEX programme over years and never come across this stringent definition before. We also have quite an extensive list of scientific publications on downscaling without our use of downscaling having been a problem. As we have pointed out previously, there are many papers published in the scientific literature where term downscaling aligns with our use of the term herein, and to iterate on this, downscaling is described in the most recent IPCC assessment report (AR6 Box 10.1, p.1377) with the following passage: "*In WGI Chapter 9 (Flato et al., 2014), regional downscaling methods were addressed as tools to provide climate information at the scales needed for many climate impact studies*". This is in accordance with our definition of downscaling. So it would be great to have some good references of the term downscaling if it really differs from the interpretations that we are used with. We have nevertheless acknowledged different takes on the term downscaling in section *A1 Interpretations of the concept of downscaling* (lines 335-359), but as stated above, we lack a good reference to the definition proposed by Anonymous referee #1. Using [DOI:10.5194/egusphere-2024-1463-RC1] does not serve this purpose well.

The point of placing this discussion in the appendix, as Anonymous referee #1 rightly pointed out, is that most people do not read them. The reason for this choice is also that most of our colleagues have a similar understanding of downscaling as CORDEX, the IPCC, and papers published in the scientific literature. We don't want superfluous text to make the paper more cumbersome for readers who already are familiar with downscaling. But, if the readers are not familiar with downscaling, then it's a better reason for presenting it in HESS and in an appendix. Downscaling is often relevant to hydrological studies. We have included some text referring the reader to introductory material on downscaling (R. E. Benestad, Hanssen-Bauer, and Chen 2008; Maraun and Widmann 2018; R. Benestad 2016).

The nature of appendices is influenced by the need to provide a more introductory description of downscaling, as suggested by the reviewer and the discussion about what downscaling really entails. They also provide information requested through previous reviews, however, we have tried to structure them in a way that they also best provide specific details about the method used in our case. They bear a character of discussions connected to previous review comments, but we do not see this as a problem, as other readers who may have similar concerns as Anonymous Reviewer #1 may find this discussion useful.

My other comments seem to have not been addressed, or at least not replied to clearly. I therefore paste them again below as a numbered list:
3) A schematic of the methodology would be helpful.
Fig. A1. provides a schematic, showing how the predictor is represented in terms of common EOFs. We have called it a 'graphical illustration', but it's really a schematic with more realistic graphics.

4) the data are only produced at the stations locations, meaning that it is less effective in areas with a lower stations density (such as northern Norway/Finland). This should be at least discussed.
This is a good point, and we have added some discussion about the effect of variable density of the observational network, mainly in Appendix A with detailed information about the methods (lines 419-424), but also in the discussion section (lines 130-131).

5) Very few details are given on the kriging procedure used. Which variogram parameters were used, and what does this tell us about the spatial dependance - which is more discussed from the angle of the variance carried by the principal components, a very indirect way of looking at spatial correlation. Furthermore, I expect the spatial dependance to be non-stationary, e.g. very different in western Norway than in other parts of the domain. How is this addressed?
We have explained the kriging procedure in the appendix, and it was developed at UCAR for creating climatological maps such as those shown in our figures. We provide references to the kriging method which was based on Markov random fields, and we are unsure whether the question about variogram parameters is relevant here. Lines 425-435 in the appendix provides more details about the kriging, however, we are no experts on kriging and our focus in this study is on downscaling daily precipitation statistics, and our focus should be clearly expressed in the abstract of our paper.

The kriging was applied to the spatial weights of the leading modes from PCA which represent coherent spatial variability and large-scale covariance structure. Hence, neighbouring weights are often similar, and we don't expect that a different kriging method would give radically different results. There are other factors that are associated with greater uncertainties, e.g. emission scenarios and the approximate methods used together with downscaling (this is of course discussed in the paper). Each PCA mode is stationary over the calibration period, and the combination of different modes takes care of non-stationarity. This is now explained in line 130 in the revised manuscript, but the point about stationarity is elaborated on in Appendix A. In other words, the choice of kriging does not have a big consequence for our general conclusions.

6) For instance, section 2.2 mentions the use of EOFs for the large-scale data and PCA for the station data. It is not clear why a different analysis was done for each data type.
Thanks for asking this. The difference between EOFs and PCA is that the former represents data on regular longitude-latitude grids, where the data is weighted by the area of their respective grid box area, whereas PCA represents data with coordinates that have irregular structures. This is now explained in Appendix A (lines 419-424). In the R-package 'esd' used for carrying in the analysis, we refer to the function used for a group of station data (with an irregular geographical scatter) as PCA and the function for field data on a regular longitude-latitude grid as EOF, but PCA and EOF analyses are essentially the same thing and the same method (SVD) is used for both. The only difference is really that the data are prepared in different ways and you get slightly different types of output (spatial patterns on a regular vs irregular grid).

7) The validation of the results is mostly done visually, with a few assessment metrics given along the text in section 2.3. I would expect results of the cross-validation to be given as detailed tables.
Cross-validation correlation scores are presented in Tables B1-B2 in Appendix B.

8) Figure 5: are these statistics calibrated? In other terms, for a Normal distribution I would expect about 64% of the data to fall with the 1-standard deviation confidence interval. This does not seem to be the case here. Can this be commented?
No, the results are the raw predictions from the equation that expresses the probability in terms of the wet-day frequency and wet-day mean precipitation. We have expanded the figure caption to explain this.

9) The discussion section is long and windy. Some of it could be moved to the introduction (e.g. ll.290-300), or to the conclusion, or removed.
Thanks for this comment. We have shortened and restructured the discussion in our revised version of the manuscript.

Anonymous referee #3

The authors have now addressed my comments. Overall the revised version is much more clear.

The references do not appear on the pdf of the revised version. I believe that this is a typesetting error that the authors need to fix.
Thanks for pointing this out. The references are fixed in the revised version.

I would also encourage the authors, while they adopt the use of the exponential distribution, to acknowledge the work that has done clearly showing that rainfall extremes do indeed in most cases have heavy tails, and clearly state that as a limitation of the study.
Thanks. We have acknowledged this fact and cited a couple of papers.

References

Benestad, Rasmus. 2016. *Downscaling Climate Information*. Vol. 1. Oxford University Press. https://doi.org/10.1093/acrefore/9780190228620.013.27.

Benestad, Rasmus E., Inger Hanssen-Bauer, and Deliang Chen. 2008. *Empirical-Statistical Downscaling*. World Scientific. https://doi.org/10.1142/6908.

Maraun, Douglas, and Martin Widmann. 2018. *Statistical Downscaling and Bias Correction for Climate Research*. Cambridge University Press.