Mudelsee M, Bermejo MA (2017) Optimal heavy tail estimation, Part I: Order selection. Nonlinear Processes in Geophysics Discussions.

https://www.nonlin-processes-geophys-discuss.net/npg-2017-25/

Author's Response to Review Comments, 24 October 2017

Point-by-Point Response to Review by Anonymous Referee #1

Referee Comment 1. The manuscript presents order selection in optimal heavy tail estimation, which is interesting. The subject addressed is within the scope of the journal. 2. However, the manuscript, in its present form, contains several weaknesses. Appropriate revisions to the following points should be undertaken in order to justify recommendation for publication. 3. For readers to quickly catch your contribution, it would be better to highlight major difficulties and challenges, and your original achievements to overcome them, in a clearer way in abstract and introduction.

Author's Response We first note that Anonymous Referee #2 has not raised these points. We think that we clearly highlighted the major methodological challenge (the selection of the order) and our original contribution to overcome this (the introduction of a brute-force order selector) in the Abstract (p. 1, l. 3 to 7), the Introduction (p. 1, l. 14 to 27)—and the title.

Author's Changes in Manuscript None.

Referee Comment 4. It is shown in the reference list that the authors have several publications in this field. This raises some concerns regarding the potential overlap with their previous works. The authors should explicitly state the novel contribution of this work, the similarities and the differences of this work with their previous publications.

Author's Response We first note that Anonymous Referee #2 has not raised this point. We publish on climate and statistics since the year 1989 and think that since that year we have been rather careful to avoid dual publications and hidden overlap with previous own work, and we think that also this NPGD contribution is fine is this regard. The main novel contribution is the order selector (see Review by Anonymous Referee #1, Author's Response to Comments 1 to 3, above). We cite previous work on other issues needed in the paper: statistics (Mudelsee, 2014), river Elbe floods (Mudelsee et al. 2003) and long memory of river runoff (Mudelsee 2007).

Author's Changes in Manuscript None.

Referee Comment 5. It is mentioned in p.1 that a data-adaptive order selector is adopted for optimal heavy tail estimation. What are the other feasible alternatives? What are the advantages of adopting this particular approach over others in this case? How will this affect the results? More details should be furnished.

Author's Response We first note that Anonymous Referee #2 has not raised this point. Other feasible selectors have already been mentioned in the manuscript (p. 5, l. 1 to 13). The advantage of our new approach over the other approaches is that it gives more accurate estimation results of the heavy-tail index parameter. This is shown in the Monte Carlo experiment (Section 3). We think that all this is clear and detailed.

Author's Changes in Manuscript None.

Referee Comment 6. It is mentioned in p.1 that the river Elbe is adopted as the case study. What are other feasible alternatives? What are the advantages of adopting this particular case study over others in this case? How will this affect the results? The authors should provide more details on this.

Author's Response We first note that Anonymous Referee #2 has not raised this point. Feasible alternatives would include other rivers. The advantage of utilizing the Elbe: here are available very long, high-quality runoff data. Results for the Elbe are likely at least as accurate as for many other rivers. Please bear in mind that this is a methodological paper, the case of the Elbe serves to illustrate the method. We think what is in the manuscript is already detailed enough.

Author's Changes in Manuscript None.

Referee Comment 7. It is mentioned in p.2 that the Hill estimator is adopted for statistical estimation of the heavy tail index. What are other feasible alternatives? What are the advantages of adopting this particular estimator over others in this case? How will this affect the results? The authors should provide more details on this.

Author's Response We first note that Anonymous Referee #2 has not raised this point. The Hill estimator of the heavy-tail index parameter is one of the most widely employed estimators. We also have the intuition that is it is a rather accurate estimator. There exist alternative estimators, we mention (p. 10, l. 32) Pickands estimator. The cited book (Resnick, 2007) lists more estimators. We think that this paper is not the place to study more estimators. As already said in the paper (p. 10, l. 29 to 32), we plan to study other estimators (to test our intuition), the results of which we plan to publish in a sequel to this paper.

Author's Changes in Manuscript None.

Referee Comment 8. It is mentioned in p.2 that a first-order autoregressive process is adopted in this study. What are the other feasible alternatives? What are the advantages of adopting this particular process over others in this case? How will this affect the results? More details should be furnished.

Author's Response We first note that Anonymous Referee #2 has not raised this point. In our view on climate and hydrology, the first-order autoregressive (AR1) process is by far the most important model of persistence. One of the reasons is its simplicity, while it still is able to capture in many situations to a good degree the serial dependence of the data-generating system. Another reason is that it avoids the so-called embedding problem and can therefore be readily used for unevenly spaced paleoclimate time series (from natural archives). Our book (Mudelsee, 2014), cited in the manuscript (p. 3, l. 10) gives the mathematical details and presents many applications. We do mention (p. 9, l. 5 to 9) the long-memory model as an alternative to the AR1 process. We think that this treatment of persistence in the manuscript is adequate.

Author's Changes in Manuscript None.

Referee Comment 9. It is mentioned in p.4 that the algorithm by Nolan (1997) is adopted to generate random values from a stable distribution. What are the other feasible alternatives? What are the advantages of adopting this particular algorithm over others in this case? How will this affect the results? More details should be furnished.

Author's Response We first note that Anonymous Referee #2 has not raised this point. However, this is an excellent point raised by Anonymous Referee #1! There are other distributions than s stable with a heavy-tail index, and our plan (communicated in the manuscript on p. 10, l. 30 to 31) is indeed to study alternatives. As regards the algorithm by Nolan (1997) to generate stable distributions, we have not performed an extensive research into alternative algorithms since we had it at hand and after we made some changes for implementation in Fortran 90, it worked well. We are convinced that using other algorithms would not affect the results significantly (within error bars). We think that more algorithmic work is beyond the scope of this manuscript.

Author's Changes in Manuscript None.

Referee Comment 10. It is mentioned in p.4 that asymptotic and bootstrap order selectors are adopted as benchmarks for comparison. What are the other feasible alternatives? What are the advantages of adopting these particular order selectors over others in this case? How will this affect the results? More details should be furnished.

Author's Response We first note that Anonymous Referee #2 has not raised this point. This point has already been treated above (Review by Anonymous Referee #1, Author's Response to Comment 5).

Author's Changes in Manuscript None.

Referee Comment 11. It is mentioned in p.4 that a Monte Carlo simulation experiment is adopted to compare the optimal order selector. What are the other feasible alternatives? What are the advantages of adopting this particular experiment over others in this case? How will this affect the results? More details should be furnished.

Author's Response We first note that Anonymous Referee #2 has not raised this point. Owing to the complexity of the data-generating process (heavy tails, serial dependence, and uneven spacing) and the estimation (order selection, Hill estimator), it is unfortunately not possible to use paper and pencil and derive and write down an analytical result on the properties (e.g., standard error, distributional shape) of the proposed method (brute-force order selection). Therefore one has to resort to numerical Monte Carlo computer simulations: there are no alternatives. (Let us point out that this is a ubiquitous situation in today's applied statistical research; a recent, very well written book on that matter is: Efron B, Hastie T (2016) *Computer Age Statistical Inference: Algorithms, Evidence, and Data Science.* Cambridge University Press, New York, 475 pp.) Please note that the Monte Carlo design parameters (Figure 1 of the manuscript) do already cover a certain range of possible settings. We therefore think that the results (superiority of the brute-force order selector) are robust in that sense, and we thus conclude that our treatment is detailed enough.

Author's Changes in Manuscript None.

Referee Comment 12. It is mentioned in p.4 that a gamma distribution is adopted to draw the prescribed uneven spacing. What are the other feasible alternatives? What are the advantages of adopting this particular distribution over others in this case? How will this affect the results? More details should be furnished.

Author's Response We first note that Anonymous Referee #2 has not raised this point. The selection of the gamma distribution for the spacing is based on the observation that this is a useful descriptive model; and we cite in the manuscript (p. 4, l. 11) our book (Mudelsee, 2014) as a reference for that. We believe that the selection of the distributional model for the time spacing may be an interesting thing to study for persistence time estimation, but it is hardly relevant here in the manuscript (which deals with heavy-tail index parameter estimation).

Author's Changes in Manuscript None.

Referee Comment 13. It is mentioned in p.8 that a quasi-brute force, two-step search method is adopted to find the optimal order. What are the other feasible alternatives? What are the advantages of adopting this particular method over others in this case? How will this affect the results? More details should be furnished.

Author's Response We first note that Anonymous Referee #2 has not raised this point. The quasi-brute force, two-step search method is used to avoid excessive computational costs associated with pure brute-force search. The quasi-brute force method has been attested to work well on artificial time series, as described in the software manual (manuscript p. 11, line 5). The manual was attached to the submission to NPGD and is available to Referees. It will be made publicly available with the software once the manuscript is accepted. We think that this treatment is detailed enough.

Author's Changes in Manuscript None.

Referee Comment 14. It is mentioned in p.8 that "…Although the observed time series has clearly more points (n = 38272) than the artificial (n = 5000), the error bar for the heavy tail index estimate is larger (RMSE_b = 0.13) than for the artificial (RMSE_b = 0.06). The reason is that the estimated..…" More justification should be furnished on this issue.

Author's Response We first note that Anonymous Referee #2 has not raised this point. The reason, explanation and justification has already been given in the manuscript (p. 8, l. 30 to p. 9, l. 2): persistence means a larger "equivalent autocorrelation coefficient" (as compared to a persistence-free situation), it reduces the effective data size (Mudelsee, 2014) and leads to larger error bars. We think that by writing this piece of text and citing our book for this (p. 8, l. 31), we did already enough to explain the error bars. Please note that the software manual (manual p. 15) shows more Monte Carlo results on the influence of persistence on estimation error bars. We think that this treatment is sufficient.

Author's Changes in Manuscript None.

Referee Comment 15. It is mentioned in p.10 that "…the study of the runoff series from the river Salt (Anderson and Meerschaert, 1998), which found..…(i.e., finite variance), in contrast to our finding..…" More justification should be furnished on this issue.

Author's Response We first note that Anonymous Referee #2 has not raised this point. At this point of the analysis, we can only report what other researcher found on other river runoff time series. We are not yet in a position to refute other findings. We therefore plan, as communicated in the paper (p. 10, l. 5 to 9) to do more studies on heavy-tails in river runoff series, involving (1) summer/winter comparisons, (2) sensitivity of removing the annual cycle, (3) comparisons among different stations along the Elbe and, finally, (4) a comparison with other rivers. The results we plan to publish in a sequel paper.

Author's Changes in Manuscript None.

Referee Comment 16. Some key parameters are not mentioned. The rationale on the choice of the particular set of parameters should be explained with more details. Have the authors experimented with other sets of values? What are the sensitivities of these parameters on the results? 17. Some assumptions are stated in various sections. Justifications should be provided on these assumptions. Evaluation on how they will affect the results should be made. 18. The discussion section in the present form is relatively weak and should be strengthened with more details and justifications.

Author's Response We first note that Anonymous Referee #2 has not raised these points. We are, however, not aware of unmentioned parameters or assumptions. We further think that we cannot agree that the discussion section in the present form is relatively weak. We would clearly have welcomed more accurate descriptions of those points 16 to 18 by Anonymous Referee #1.

Author's Changes in Manuscript None.

Referee Comment 19. Moreover, the manuscript could be substantially improved by relying and citing more on recent literatures about real-life case studies of contemporary optimization techniques in hydrologic engineering such as the followings:

Gholami, V., et al., “Modeling of groundwater level fluctuations using dendrochronology in alluvial aquifers”, Journal of Hydrology 529 (3): 1060-1069 2015.

Taormina, R., et al., “Data-driven input variable selection for rainfall-runoff modeling using binary-coded particle swarm optimization and Extreme Learning Machines”, Journal of Hydrology 529 (3): 1617-1632 2015.

Wu, C.L., et al., “Prediction of rainfall time series using modular artificial neural networks coupled with data-preprocessing techniques”, Journal of Hydrology 389 (1-2): 146-167 2010.

Wang, W.C., et al., "Improving forecasting accuracy of annual runoff time series using ARIMA based on EEMD decomposition," Water Resources Management 29 (8): 2655-2675 2015.

Chen, X.Y., et al., "A comparative study of population-based optimization algorithms for downstream river flow forecasting by a hybrid neural network model," Engineering Applications of Artificial Intelligence 46 (A): 258-268 2015.

Chau, K.W., et al., "A Hybrid Model Coupled with Singular Spectrum Analysis for Daily Rainfall Prediction," Journal of Hydroinformatics 12 (4): 458-473 2010.

Author's Response We first note that Anonymous Referee #2 has not raised this point. Let us first copy the paper title together with the abstracts of the suggested new citations.

Gholami, V., et al., "Modeling of groundwater level fluctuations using dendrochronology in alluvial aquifers", Journal of Hydrology 529 (3): 1060-1069 2015.
Abstract. Groundwater is the most important water resource in semi-arid and arid regions such as Iran. It is necessary to study groundwater level fluctuations to manage disasters (such as droughts) and water resources. Dendrochronology, which uses tree-rings to reconstruct past events such as hydrologic and climatologic events, can be used to evaluate groundwater level fluctuations. In this study, groundwater level fluctuations are simulated using dendrochronology (tree-rings) and an artificial neural network (ANN) for the period from 1912 to 2013. The present study was undertaken using the Quercus Castaneifolia species, which is present in an alluvial aquifer of the Caspian southern coasts, Iran. A multilayer percepetron (MLP) network was adopted for the ANN. Tree-ring diameter and precipitation were the input parameters for the study, and groundwater levels were the outputs. After the training process, the model was validated. The validated network and tree-rings were used to simulate groundwater level fluctuations during the past century. The results showed that an integration of dendrochronology and an ANN renders a high degree of accuracy and efficiency in the simulation of groundwater levels. The simulated groundwater levels by dendrochronology can be used for drought evaluation, drought period prediction and water resources management.

Taormina, R., et al., "Data-driven input variable selection for rainfall-runoff modeling using binary-coded particle swarm optimization and Extreme Learning Machines", Journal of Hydrology 529 (3): 1617-1632 2015.
Abstract. Selecting an adequate set of inputs is a critical step for successful data-driven streamflow prediction. In this study, we present a novel approach for Input Variable Selection (IVS) that employs Binary-coded discrete Fully Informed Particle Swarm optimization (BFIPS) and Extreme Learning Machines (ELM) to develop fast and accurate IVS algorithms. A scheme is employed to encode the subset of selected inputs and ELM specifications into the binary particles, which are evolved using single objective and multi-objective BFIPS optimization (MBFIPS). The performances of these ELM-based methods are assessed using the evaluation criteria and the datasets included in the comprehensive IVS evaluation framework proposed by Galelli et al. (2014). From a comparison with 4 major IVS techniques used in their original study it emerges that the proposed methods compare very well in terms of selection accuracy. The best performers were found to be (1) a MBFIPS–ELM algorithm based on the concurrent minimization of an error function and the number of selected inputs, and (2) a BFIPS–ELM algorithm based on the minimization of a variant of the Akaike Information Criterion (AIC). The first technique is arguably the most accurate overall, and is able to reach an almost perfect specification of the optimal input subset for a partially synthetic rainfall–runoff experiment devised for the Kentucky River basin. In addition, MBFIPS–ELM allows for the determination of the relative importance of the selected inputs. On the other hand, the BFIPS–ELM is found to consistently reach high accuracy scores while being considerably faster. By extrapolating the results obtained on the IVS test-bed, it can be concluded that the proposed techniques are particularly suited for rainfall–runoff modeling applications characterized by high nonlinearity in the catchment dynamics.

Wu, C.L., et al., "Prediction of rainfall time series using modular artificial neural networks coupled with data-preprocessing techniques", Journal of Hydrology 389 (1-2): 146-167 2010.
Abstract. This study is an attempt to seek a relatively optimal data-driven model for rainfall forecasting from three aspects: model inputs, modeling methods, and data-preprocessing techniques. Four rain data records from different regions, namely two monthly and two daily series, are examined. A comparison of seven input techniques, either linear or nonlinear, indicates that linear correlation analysis (LCA) is capable of identifying model inputs reasonably. A proposed model, modular artificial neural network (MANN), is compared with three benchmark models, viz. artificial neural network (ANN), K-nearest-neighbors (K-NN), and linear regression (LR). Prediction is performed in the context of two modes including normal mode (viz., without data preprocessing) and data preprocessing mode. Results from the normal mode indicate that MANN performs the best among all four models, but the advantage of MANN over ANN is not significant in monthly rainfall series forecasting. Under the data preprocessing mode, each of LR, K-NN and ANN is respectively coupled with three data-preprocessing techniques including moving average (MA), principal component analysis (PCA), and singular spectrum analysis (SSA). Results indicate that the improvement of model performance generated by SSA is considerable whereas those of MA or PCA are slight. Moreover, when MANN is coupled with SSA, results show that advantages of MANN over other models are quite noticeable, particularly for daily rainfall forecasting. Therefore, the proposed optimal rainfall forecasting model can be derived from MANN coupled with SSA.

Wang, W.C., et al., "Improving forecasting accuracy of annual runoff time series using ARIMA based on EEMD decomposition," Water Resources Management 29 (8): 2655-2675 2015.
Abstract. Hydrological time series forecasting is one of the most important applications in modern hydrology, especially for effective

reservoir management. In this research, the auto-regressive integrated moving average (ARIMA) model coupled with the ensemble empirical mode decomposition (EEMD) is presented for forecasting annual runoff time series. First, the original annual runoff time series is decomposed into a finite and often small number of intrinsic mode functions (IMFs) and one residual series using EEMD technique for a deep insight into the data characteristics. Then each IMF component and residue is forecasted, respectively, through an appropriate ARIMA model. Finally, the forecasted results of the modeled IMFs and residual series are summed to formulate an ensemble forecast for the original annual runoff series. Three annual runoff series from Biuliuhe reservoir, Dahuofang reservoir and Mopanshan reservoir, in China, are investigated using developed model based on the four standard statistical performance evaluation measures (RMSE, MAPE, R and NSEC). The results obtained in this work indicate that EEMD can effectively enhance forecasting accuracy and that the proposed EEMD-ARIMA model can significantly improve ARIMA time series approaches for annual runoff time series forecasting.

Chen, X.Y., et al., "A comparative study of population-based optimization algorithms for downstream river flow forecasting by a hybrid neural network model," Engineering Applications of Artificial Intelligence 46 (A): 258-268 2015.
Abstract. Population-based optimization algorithms have been successfully applied to hydrological forecasting recently owing to their powerful ability of global optimization. This paper investigates three algorithms,i.e. differential evolution (DE), artificial bee colony (ABC) and ant colony optimization (ACO), to deter-mine the optimal one for forecasting downstream river flow. A hybrid neural network (HNN) model,which incorporates fuzzy pattern-recognition and a continuity equation into the artificial neural net-work, is proposed to forecast downstream river flow based on upstream river flows and areal precipitation. The optimization algorithm is employed to determine the premise parameters of the HNNmodel. Daily data from the Altamaha River basin of Georgia is applied in the forecasting analysis. Discussions on the forecasting performances, convergence speed and stability of various algorithms are presented. For completeness' sake, particle swarm optimization (PSO) is included as a benchmark case for the comparison of forecasting performances. Results show that the DE algorithm attains the best performance in generalization and forecasting. The forecasting accuracy of the DE algorithm is comparable to that of the PSO, and yet presents weak superiority over the ABC and ACO. The Diebold–Mariano (DM) test indicates that each pair of algorithms has no difference under the null hypothesis of equal forecasting accuracy. The DE and ACO algorithms are both favorable for searching parameters of the HNN model, including the recession coefficient and initial storage. Further analysis reveals the drawback of slow convergence and time-consumption of the ABC algorithm. The three algorithms pre-sent stability and reliability with respect to their control parameters on the whole. It can be concluded that the DE and ACO algorithms are considerably more adaptive in optimizing the forecasting problem for the HNN model.

Chau, K.W., et al., "A Hybrid Model Coupled with Singular Spectrum Analysis for Daily Rainfall Prediction," Journal of Hydroinformatics 12 (4): 458-473 2010.
Abstract. A hybrid model integrating artificial neural networks and support vector regression was developed for daily rainfall prediction. In the modeling process, singular spectrum analysis was first adopted to decompose the raw rainfall data. Fuzzy C-means clustering was then used to split the training set into three crisp subsets which may be associated with low-, medium- and high-intensity rainfall. Two local artificial neural network models were involved in training and predicting low- and medium-intensity subsets whereas a local support vector regression model was applied to the high-intensity subset. A conventional artificial neural network model was selected as the benchmark. The artificial neural network with the singular spectrum analysis was developed for the purpose of examining the singular spectrum analysis technique. The models were applied to two daily rainfall series from China at 1-day-, 2-day- and 3-day-ahead forecasting horizons. Results showed that the hybrid support vector regression model performed the best. The singular spectrum analysis model also exhibited considerable accuracy in rainfall forecasting. Also, two methods to filter reconstructed components of singular spectrum analysis, supervised and unsupervised approaches, were compared. The unsupervised method appeared more effective where nonlinear dependence between model inputs and output can be considered.

Next, let us make a text search through the abstracts for the expressions "heavy", "tail", "distribut*" and "density". The results are zero. When browsing through the abstracts, this confirms the exercise from the text search, namely that none of these papers (certainly respectful work) deals with heavy-tail distributions, which, however, is the target of our present manuscript with NPGD. Instead the papers seem to focus on prediction and neural networks applied to runoff time series from various stations. Thus, the only study object common with the NPGD manuscript seems to be river runoff. With all due apologies, we think that this does not warrant inclusion of all of the suggested citations. However, we think that a paragraph on the "wider impacts" of our study should be added to the Conclusions section of our manuscript; this paragraph should include paleoclimatology, paleohydrology and dendrochronology; and we cite the first suggestion, Gholami et al. (2015) in that paragraph, together with other work.

Author's Changes in Manuscript Section 7 (Conclusions), at the very end of the manuscript (p. 10, after line 32) we add the following short paragraph (italicized).

The wider impact of optimal heavy tail estimation may be not only on the application to the area of instrumental environmental measurements, but also to reconstructed *variables from the areas of paleoclimatology (Cronin, 2010), paleohydrology (Gasse, 2009) and dendrochronology (D'Arrigo et al., 2011; Gholami, 2015). Furthermore, since extreme events in hydrology and related fields may also show the duration aspect (e.g., droughts, heatwaves), the estimation should not be restricted to measured or reconstructed variables. Rather, heavy tail index estimation should be a useful tool also for the analysis of index variables (Kürbis et al., 2009).*

Additional references:

Cronin, T. M.: Paleoclimates: Understanding Climate Change Past and Present. Columbia University Press, New York, 441 pp., 2010.

D'Arrigo, R., Abram, N., Ummenhofer, C., Palmer, J., and Mudelsee, M.: Reconstructed streamflow for Citarum river, Java, Indonesia: Linkages to tropical climate dynamics, Clim. Dynam., 36, 451–462, 2011.

Gasse, F.: Paleohydrology, in: Encyclopedia of Paleoclimatology and Ancient Environments, edited by: Gornitz, V., Springer, Dordrecht, 733–738, 2009.

Gholami, V., Chau, K. W., Fadaee, F., Torkaman, J., and Ghaffari, A.: Modeling of groundwater level fluctuations using dendrochronology in alluvial aquifers, J. Hydrol., 529, 1060–1069, 2015.

Kürbis, K., Mudelsee, M., Tetzlaff, G., and Brázdil, R.: Trends in extremes of temperature, dew point, and precipitation from long instrumental series from central Europe, Theor. Appl. Climatol., 98, 187–195, 2009.

Referee Comment 20. In the conclusion section, the limitations of this study and suggested improvements of this work should be highlighted.

Author's Response We first note that Anonymous Referee #2 has not raised this point. Please note that the present manuscript already achieves what is requested here by Anonymous Referee #1. First, we stress the need to do more studies of measured river runoff series (p. 10, l. 5 to 9), see also Review by Anonymous Referee #1, Author's Response to Comment 15, above). Second, we stress the need to do more studies on other methodical aspects (p. 10, l. 29 to 32). However, we stand by our manuscript that it tackles the most important methodical problem for heavy-tail estimation, namely order selection. And we show that our paper successfully solves that problem.

Author's Changes in Manuscript None.

Point-by-Point Response to Review by Anonymous Referee #2

Referee Comment This is an excellent manuscript and I thoroughly enjoyed reading it. I have one minor comment, which the authors may clarify. How do the authors consider duration of an extreme event, as that is a very important characteristic for hydrologic variable.

Author's Response Thanks a lot for the compliment, and thanks a lot for the interesting question! Anonymous Referee #2 is right that we have in our manuscript not dealt with the duration of an extreme event. Instead we have (Section 6 of the manuscript) analysed the daily runoff values (river Elbe, station Dresden). We completely agree that the duration aspect is an important characteristic index for a hydrological variable, for example, in the definition of droughts. The manuscript already stresses the need to do more studies of measured river runoff series (p. 10, l. 5 to 9), see also Review by Anonymous Referee #1, Author's Response to Comment 15, above)—and exactly here one should add a further study direction: droughts, and the various forms to define it (e.g., via the duration). Another field, for which the hint by Anonymous Referee #2 is helpful, is the climate variable temperature, which, together with duration of an extreme warm event, may be used to define heatwaves. We think it is worth to mention droughts and heatwaves as index variables, which can also be analysed by means of optimal heavy tail index estimation.

Author's Changes in Manuscript Section 7 (Conclusions), at the very end of the manuscript (p. 10, after line 32) we add a short paragraph; please see Review by Anonymous Referee #1, Author's Changes in Manuscript in response to comments 19 (above).