Dear Reviewer,

I would like to thank you for your valuable remarks on the MS. Please find my responses (plain font) to the comments (italic font) and suggested changes in response to the comments below.

Response to Reviewer #1

1. I have concerns when using the term "accurate" when comparing different statistical models (e.g. in the abstract and the last lines of the conclusion). I can understand the idea of comparing the models by estimating the differences in the low flow quantiles. As always for statistical models in hydrological statistics, it is almost impossible to compare the results to the "true" value, as we simply cannot observe this. We are limited to the observations we have and can only compare the results of the models. Hence, the term "more accurate" for one model compared to another is not suitable, even if one says that statistically the new model is more reasonable (to which I absolutely agree). The only way to justify something like accuracy would be simulations, as in this case the truth is known. I therefore recommend to weaken the statements in the manuscript and omit terms like "more accurate". Instead, the use of the term "gain" seems to be a very good idea to me and should be done throughout the whole manuscript. Similarly, the terms "error" (Figure 3) as well as under- and overestimation (e.g. ll. 275-276) seem to be misplaced here. I would also highly appreciate a simulation study that clearly demonstrates the benefits of the copula mixture model, but I can understand that this would increase the length of the manuscript too much.

RE: Thanks for this comment. I completely agree with this statement and it was actually my intention to avoid too strong statements, as the true distribution is indeed unknown, and superiority of an estimator over another is derived from theoretical considerations. Following the advices, I will revise the wording accordingly.

2. The author states that the Gumbel-Hougaard copula is used "as we are modeling extreme values". There exist many copula models for this purpose and I cannot see why only this should be suitable, especially for an application to whole Europe. I suggest to use a goodness-of-fit criterion such as AIC/BIC to select the best-fitting copula from a sample of possible models. It is not a priori clear to me why there should always be the same dependence structure for all catchments.

RE: The choice of the Gumbel-Hougaard model is based on a literature review from which I inferred its general suitability to model dependencies between extreme value distributions. Studies have either shown the direct suitability of the model (e.g. Klein et al., 2011; Poullin et al., 2007) or that the differences between different extreme value copula models to be mostly negligible (Genest and Favre, 2007). This was also confirmed for the data at hand in a preliminary analysis, and as the purpose of the study was less a comparative assessment of copula models than demonstrating the general suitability of mixed distribution approaches with respect to seasonal correlations, the priority was given to use the same model for all stations rather than using different models across the study area.

In response of the comment, I have checked the validity of the assumed Gumbel-Hougaard copula in greater detail, and assessed the sensitivity of the results of the study when using a possibly more adequate model. A first GOF assessment (Cramer-von-Mises test implemented in gofCopula of the copula package) shows that the GH copula was rejected in 15% of stations which is somewhat higher than the theoretical rejection rate of the hypothesis test (alpha = 0.05). In a further assessment a cross-validated log-likelihood criterion showed that the GH copula performs best among five EV copulas (according to the number of best performing cases), whereas the Gumbel copula has only

above-average rank (5.1) among eight other non-extreme value copula families and is outperformed by the Plackett (average rank 6.5) and Frank copula (average rank 6.5). This suggests a mixture of the three very contrasting models to be more adequate, given that there is no clear theoretical preference between them.

Following the suggestion of the reviewer I suggest to revise the paper using the station-wise best performing of three models (Gumbel, Plackett and Frank copula). In addition, I will use the Independence copula to model 30 cases with a negative Rho which are all insignificant and not plausible from a drought generation process perspective). All results will be adapted accordingly. A preliminary assessment shows that the results are not very sensitive to these changes. The regression effects of the considered predictors are even stronger and lead to clearer patterns across the study area, which supports the use of the best of different copulas at each case.

3. After the model description, the manuscript goes right into the details of the results. I am missing at least a short description of the data. As one major aspect here is the dependence between winter and summer low flows, it is of high interest how these events have been defined. What is done in a case when a summer low flow reaches into winter or there is no clear distinction between summer and winter? How is winter and summer defined and are the same seasonal thresholds used for whole Europe? These are aspects which make the application of seasonal mixture models to low flows so much harder compared to floods. Moreover, they motivate this whole study.

RE: A short description of the data will be added and include a statement about the range of the seasonal Rho values showing the various degrees of dependence between summer and winter lows.

Concerning the questions about how winter and summer low flows are defined and how to deal with the case that low flows reach from one season into the other, and what seasonal thresholds are used: I think there might be a misunderstanding of the event definition used in this paper (and I realize that I was not clear enough right from the beginning of the paper defining the low flow event, which I will improve in the introduction of the revised MS): I am not analyzing low-flows based on thresholds (i.e., Yevjevich's threshold level approach), where the effects of droughts continuing from one season to the other make a clear attribution of events hard to obtain. Such analysis would indeed be very sensitive to the defined summer and winter seasons and seasonal thresholds used across Europe.

However, I am focusing here on low-flow magnitude as described by annual and seasonal mina series, where the seasonal attribution of events into a winter and summer type are straightforward. Earlier studies (such as Laaha et al., 2017) have shown that defining a winter season from December to April is appropriate to attribute frost-generated annual minima uniquely to the winter season, and remaining events (generated by meteorological water-balance deficits) to the summer season. Ongoing events at the end of a season are explicitly considered in this paper by their seasonal correlation and enter into the copula model according to their station-specific relevance. For these reasons, I would argue that mixed distribution approaches are more straightforward for low-flows than for flood peaks where attribution to process types is much harder to obtain than for low-flows, which is also reflected in the much clearer mixed distribution plots for low-flows than those published for floods.

The definition of the low-flow index will be made clearer right from the beginning of the paper.

4. The indices under consideration should be explained. For example, it is not clear what SR or the circular seasonality index r are. In Table 1 it is stated that SR is the "ratio between mean summer and mean winter low-flow." If this is correct, why is it an indicator of seasonality? If there are few extreme low flows in summer and otherwise many large low flows while in winter there are many small low flows, the indicator would be highly impacted by the extreme low flows in summer. Wouldn't it be more reasonable to use the proportion of summer (or winter) low flows in the annual maximum series to investigate which has a greater impact on the AM? The index r only appears in Table 1 but is not mentioned in the text.

RE: I will add a paragraph explaining in which the considered indices will be briefly explained.

Concerning the question why SR is an indicator of seasonality, I refer to our earlier paper (see citation in MS above) where the skill of various seasonality indices was assessed in greater detail. Among them, the seasonality histograms (counts of low flow days per month) is the most detailed index, the circular seasonality index (mean day of occurrence and strength r) is more condensed, and the SR is the most condensed information. We found that seasonality patterns that we obtained when plotting the information for the Austrian study area are quite similar to each other, so all three indices were indeed appropriate to discern between summer, winter and mixed regimes in the study area. I agree that the index SR is more sensitive to absolute minima than durations or counts below a threshold, but all these indices are highly correlated. The proportion of summer (or winter) low flows in the annual minima series was used as a third indicator in the first paper, and termed mixture rate therein. It was used in an initial assessment but again where it led to similar results, but with somewhat weaker performance. As the SR and r index are more established we decided to show only results for SR and r here.

Technical remarks:

l. 19: use "than" instead of "as"

l. 58: What does "annual drought duration of volume" mean?

l. 62: I guess a "not" is missing here.

l. 149: (and throughout the text) I believe it should be "Elbersee gauge" instead of "gauge Elbersee"

l. 158: "of" instead of "or"?

l. 163: "that the mixed probability estimator"

l. 246: should be "in Table 2".

l. 308: "has" instead of "have"

Figure 1: The mixed Copula model should be added to the legend. Currently, it is only mentioned in the caption.

Figure 3: remove "as" from the first line in the caption

RE: Many thanks, done!