Reply to Reviewer 1 - Professor Chuck Kroll

Reviewer's general comments

Review of Non-stationary low-flow frequency analysis with mixed Weibull components and Copula-based dependence framework (https://doi.org/10.5194/egusphere-2026-1820) submitted to HESS.

Reviewer: Chuck Kroll, SUNY ESF, Syracuse, NY USA

Scientific significance: Good

Scientific quality: Good

Presentation quality: Good

I liked this manuscript, and recommend it be accepted for publication in HESS with minor revisions. It is well-written and generally clear. The manuscript provides an interesting framework to address both seasonality and trends in low flow frequency analysis. My primary suggestions are to improve some of the clarity of the manuscript and to further discuss the context of the analysis.

Response to Reviewer's general comments

Dear Professor Chuck Kroll,

Thank you very much for your careful reading of our manuscript, "Non-stationary low-flow frequency analysis with mixed Weibull components and Copula-based dependence framework" (https://doi.org/10.5194/egusphere-2026-1820), and for your constructive and encouraging comments. We are pleased that you found the manuscript interesting and suitable for potential publication in HESS after minor revisions. We have carefully considered your comments and revised the manuscript accordingly. Please find below our point-by-point response, where we refer to each comment in italic, followed by our response and the corresponding suggested revisions. In the revised text, additions and modifications are shown in blue, while deletions are shown in ~~red strikethrough~~.

Sincerely yours,

Farhana Sweeta Fitriana, Svenja Fischer, Gabriele Weigelhofer, Johannes Laimighofer and Gregor Laaha

Reviewer's specific comments

(1) Line 93. Do any of the ROBIN sites you are employing contain intermittent flows, i.e. zeros? If so, how does this impact your analysis?

Response to comment 1:

Thank you for drawing attention to this issue. In this study, we used 721 European catchments from the ROBIN Level-1 network and are therefore intended to provide high-quality, near-natural records suitable for analysing hydrological extremes, including low flows (Turner et al., 2025). To assess whether intermittent-flow conditions could affect our analysis, we checked the ROBIN catchments used against the intermittent-flow classification of Tramblay et al. (2021). No-flow conditions were defined as discharge values below 0.0001 $m^3\ s^{-1}$ lasting for at least five consecutive days and occurring in at least two different years. Based on this comparison, only 23 out of the 721 ROBIN catchments were classified as intermittent, corresponding to approximately 3.2% of the analysed catchments. This indicates that intermittent-flow conditions occur only in a small fraction of the study

sample and are therefore unlikely to substantially affect the overall conclusions. Nevertheless, we agree that catchments with repeated zero or near-zero flows require special consideration. Because the three-parameter Weibull distribution is continuous, it cannot explicitly represent a probability mass at zero. Such cases would require a modelling framework that accounts for zero-flow occurrences, for example, by treating zeros as censored observations within a mixed discrete-continuous L-moment framework. Extending the present non-stationary seasonal mixture Weibull-3 and copula framework to explicitly account for intermittent-flow conditions is beyond the scope of this study, but we acknowledge this as a potential direction for future work.

We have now added this clarification to the revised manuscript in (Lines 397-403) Finally, the possible influence of intermittent-flow conditions on the applicability of the WEI3-based framework was assessed by comparing the ROBIN catchments used in this study with the intermittent-flow classifications of Tramblay et al. (2021). Only 3.2,% of catchments were classified as intermittent, indicating that zero-flow conditions affect only a small fraction of the analysed dataset. Because the WEI3 distribution is continuous and cannot explicitly represent a probability mass at zero, catchments with repeated zero-flow occurrences would require a modelling framework that accounts for zero flows, for example by treating zeros as censored observations within a mixed discrete-continuous L-moment framework (Laaha, 2023a, b; Bhatti et al., 2019).

(2) Lines 112-113. The sentence reads awkwardly.

Response to comment 2:

Thank you for pointing this out. We agree that the transition between the two sentences was abrupt and have revised the wording to improve the flow and clarity. (Line 120-122) Figure 1a-b shows that the study catchments are predominantly small to medium-sized and located at low to mid elevations, with fewer large or high-elevation catchments. In addition to these physiographic characteristics, Figure 1c highlights clear spatial patterns of summer, winter, and mixed low-flow regimes across Europe.

(3) Lines 136-138. What is the importance of this clustering to your analysis?Is it to group sites for interpreting results? Would interpreting histograms of annual minimum 7-day flows produce similar groupings?

Response to comment 3:

Thank you for highlighting this point. We have revised the paragraph to clarify that (Lines 145-147) PAM clustering of seasonality histograms identifies whether low-flow events occur predominantly in summer, winter or both seasons, thereby providing regional evidence that annual low-flow minima may originate from different seasonal generating processes and supporting the need for a seasonal mixture framework. This seasonal information would not be captured by histograms of annual minimum 7-day flows, which mainly describe low-flow magnitudes rather than the timing of low-flow occurrence.

(4) Lines 152-155. These ranges of Kendall's tau seem arbitrary. Could you instead relate them to p-values for the Mann-Kendall test?

Response to comment 4:

We appreciate this comment, as it helps clarify the interpretation of our results. We agree that the interpretation of Kendall's τ should be linked to statistical significance. We have revised the text to make this procedure explicit and to clarify that the weak, moderate, and strong classes describe the strength of significant dependence, rather than replacing the significance test. (Lines 163-166) For each catchment, the dependence between the annual summer and winter low-flow series was assessed using Kendall's rank correlation coefficient (τ) (Kendall, 1976). Statistical significance was evaluated at the 5 % level, and

correlations with $p > 0.05$ were classified as non-significant. Only significant correlations were subsequently grouped, for interpretation purposes, into weak ($0.10 \leq \tau < 0.30$), moderate ($0.30 \leq \tau < 0.55$), and strong ($\tau \geq 0.55$) dependence classes.

(5) Line 168. I believe the Sen slope assesses the trend in the median and not the mean.

Response to comment 5:

Thank you for pointing this out, and we appreciate the correction. We have revised the sentence accordingly. (Lines 180-182) The magnitude of the trend is quantified using the Sen slope estimator, ~~a robust measure of the linear trend in the mean~~ a robust non-parametric estimate of the monotonic rate of change based on the median of all pairwise slopes.

(6) Line 173. While it is somewhat semantics, it is not the probability of a low flow event, but a probability of having a low flow event equal or smaller, i.e. a non-exceedance probability.

Response to comment 6:

Thank you for this helpful clarification. We agree that "non-exceedance probability" is the more precise term in this context, and we revised the wording accordingly. (Lines 186-187) Under the assumption that summer and winter events are independent, ~~the occurrence probability of a low-flow event~~ the probability that the low-flow event is equal to or smaller than magnitude q can be obtained from its seasonal non-exceedance probabilities.

(7) Line 181. While theoretically annual minimum flow series should follow a Weibull distribution, sometimes it has been shown to provide a poor fit to these series. Might you be able to show, or at least discuss, this distributional choice and its use at these catchments for these series? Could an L-moment diagram of these series motivate its use?

Response to comment 7:

Thank you for this very relevant comment. We are grateful to the reviewer for highlighting this issue. We developed the proposed framework for the WEI3 distribution deliberately, because it is theoretically motivated by the Fisher-Tippett-Gnedenko theorem as a limiting distribution for lower extremes with a finite lower bound, and is therefore well established in low-flow frequency analysis. It provides a suitable basis for extending stationary low-flow frequency analysis towards seasonal and non-stationary conditions.

At individual catchments, apparent departures from the fitted Weibull distribution may occur. In this study, we interpret such departures not only as a matter of distributional choice, but also as a possible indication of sample properties, especially process heterogeneity. For example, annual low-flow minima may be generated by different seasonal mechanisms, so that a single annual distribution may be less appropriate than a model that explicitly separates summer and winter low-flow processes. We therefore use distributional plots as diagnostic tools to identify such sample issues, rather than primarily as a basis for selecting alternative distributions with better goodness-of-fit. Previous studies by Laaha (2023a, b) provide evidence that explicitly accounting for seasonal process heterogeneity through mixed-distribution approaches improves the representation of low-flow extremes.

In response to your suggestion, we added an L-moment ratio diagram as an additional diagnostic for the stationary seasonal samples. The empirical L-skewness (τ_3) and L-kurtosis (τ_4) values of the summer and winter low-flow series are generally consistent with the theoretical Weibull curve, supporting the suitability of the Weibull distribution for the seasonal margins. However, this diagnostic cannot be directly transferred to the non-stationary setting, because standard L-moment diagrams

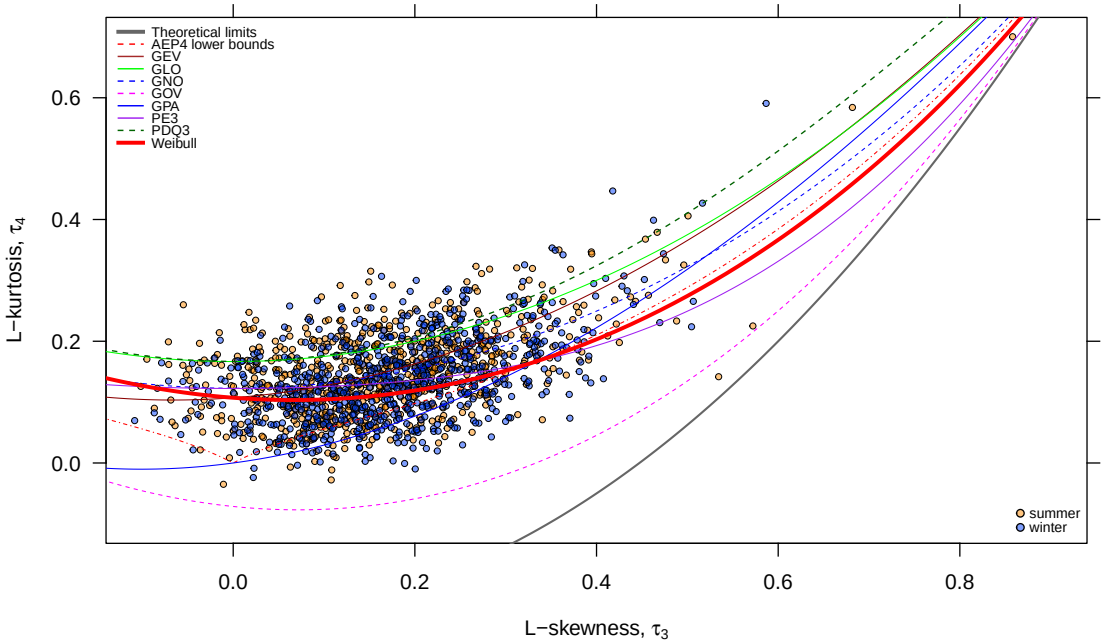

Figure 1. L-moment ratio diagram for summer and winter low-flow series. Points represent empirical L-skewness (τ_3) and L-kurtosis (τ_4) of the seasonal low-flow samples, while the curves show theoretical L-moment relationships for candidate distributions, including the Weibull distribution.

are designed for stationary samples. The use of the Weibull distribution in the non-stationary framework is therefore supported by both its theoretical basis for lower extremes and its reasonable agreement with the empirical L-moment characteristics of the seasonal samples.

In principle, the framework developed here could be extended to other distribution families. However, the aim of the present study was first to establish a non-stationary seasonal mixture and copula framework using the Weibull distribution, given its theoretical motivation for lower extremes and its use in low-flow frequency analysis. We also agree that L-moment diagrams can be valuable tools for motivating distributional choices in stationary frequency analysis. However, their direct application is not straightforward in the present setting, where the model is both seasonal and non-stationary, with time-varying parameters and copula-based dependence. To our knowledge, there is currently no established L-moment diagnostic framework for non-stationary seasonal mixture distributions of this kind. For this reason, we adopted a maximum-likelihood approach, which allows the stationary and non-stationary formulations to be compared consistently using likelihood-based criteria and likelihood-ratio tests.

We have now added a discussion of the choice of the Weibull distribution to the manuscript. (Lines 336-341) The use of the WEI3 distribution is theoretically motivated by the Fisher-Tippett-Gnedenko theorem, as the Weibull family represents a limiting distribution for lower extremes with a finite lower bound and is therefore well established in low-flow frequency analysis. Nevertheless, departures from the fitted distribution may occur at individual catchments. In this study, such departures are interpreted not only as a distributional issue, but also as a diagnostic indication of sample properties, particularly process heterogeneity between summer and winter low-flow generation. We also acknowledge that extending distributional diagnostics,

including L-moment-based tools, to non-stationary seasonal mixture models would be a valuable topic for future research, particularly for non-stationary regional frequency analysis.

(8) Line 219. Could you give further evidence of whether model parameters are stationary or not? You could split the record and fit the WEI3 to both halves and observe whether there are common trends in the parameters. While it is beyond the scope of this manuscript to consider other non-stationary moments, it might be good to assess this assumption.

Response to comment 8:

Thank you for this important comment. We agree that assessing whether model parameters are stationary is important. Splitting the record into two sub-periods and fitting the WEI3 distribution separately to each half could provide an exploratory indication of possible parameter changes. However, the parameter estimates would be affected by the shorter sample size in each sub-period, which is particularly relevant for low-flow extremes, where estimation uncertainty can already be substantial.

While the block-bootstrap Mann-Kendall test is applied separately to the seasonal series as a trend diagnostic, it does not require estimating a full Weibull parameter set from short sub-periods. By contrast, fitting separate WEI3 distributions to the first and second halves of the record would require estimating all distribution parameters independently for each season and sub-period, which may increase estimation uncertainty. For this reason, we assessed non-stationarity using the full record within a likelihood-based framework. Specifically, the stationary and non-stationary WEI3 formulations are compared using a likelihood-ratio test, where the non-stationary model allows temporal variation in the location parameter while keeping the scale and shape parameters constant. This approach uses all available data and provides a formal test of whether the additional time-varying parameter is supported by the data.

We agree that extending the framework to consider non-stationarity in other moments, particularly variance, would be valuable. This is the focus of our next research step, where we investigate whether the coefficient of variation of low-flow series remains approximately constant over time across European catchments. This follow-up work evaluates annual and seasonal low flows, associated with distinct generating processes, using both conventional moments and L-moments. The aim is to assess whether non-stationarity in low flows can be represented more parsimoniously, with changes mainly attributed to the mean, while avoiding unnecessary time variation in variance or shape parameters. If supported, this would imply that non-stationary models may focus primarily on changes in the mean, with changes in variance treated as largely consequential rather than independently modelled. However, allowing additional parameters, especially the shape parameter, to vary over time would substantially increase estimation uncertainty and the risk of overfitting. For this reason, the present study adopts a parsimonious formulation in which non-stationarity is introduced only in the location parameter of the seasonal distributions.

(9) Line 224. What motivates this model? Why linear in real space? The location parameter is essentially a lower bound. Have you looked at whether the lower bounds are increasing, as they seem to be in Fig. 4a? You might do a quantile regression to assess this.

Response to comment 9:

We are grateful to the reviewer for highlighting this issue. The motivation for our non-stationary mixture approach is that we first tested trend significance in the seasonal low-flow series using the block-bootstrap Mann-Kendall test. Where there is evidence of a significant trend in the seasonal series, the stationarity assumption may be violated and should therefore not be ignored in the frequency analysis.

The reason for limiting the model to a linear trend in the location parameter was explained in the manuscript (Lines 222-227), and we have now clarified this motivation further. In brief, the linear formulation provides a simple and parsimonious way to represent gradual temporal change while avoiding unnecessary model complexity. Given the limited record length, more

flexible non-linear formulations or time-varying scale and shape parameters would require stronger empirical support and could substantially increase estimation uncertainty and the risk of overfitting. The non-stationary formulation was then tested directly by comparing it with the simpler stationary mixture model using a likelihood ratio test. This test yielded somewhat lower significance rates than the individual seasonal trend tests, but the results were generally consistent with them.

While quantile regression could provide a useful additional exploratory diagnostic for temporal changes in individual seasonal series or in the annual series, we have not included it as an additional analysis in the present manuscript. This is because it would not directly test the full model structure considered here. In contrast, our likelihood-based framework evaluates non-stationarity within the seasonal mixture model and therefore allows the main assumptions motivating the study, namely process heterogeneity, seasonal dependence, and non-stationarity, to be assessed in a consistent framework.

We have now added a discussion of the model's motivation. The added discussion: (Lines 383-390) Its value lies not in universally replacing stationary models, but in identifying and representing catchments where seasonal low-flow regimes exhibit temporal change. In this case, the stationarity assumption may be violated and should not be ignored in the frequency analysis. To keep the framework parsimonious, non-stationarity was introduced as a linear trend in the location parameter and was evaluated formally by comparing the non-stationary and stationary models using a likelihood-ratio test. The linear formulation was chosen as a simple and parsimonious way to represent gradual temporal change. Given the limited record length, more flexible non-linear formulations or time-varying scale and shape parameters would require stronger empirical support and could substantially increase estimation uncertainty and the risk of overfitting.

(10) Line 251. It would be helpful to either include the likelihood function or provide a reference to where this could be obtained. Are you assuming your observations are independent?

Response to comment 10:

Thank you for this constructive comment. We have added the likelihood functions for the stationary and non-stationary WEI3 models in Appendix A and revised the main text to refer to this appendix. We also clarified that the likelihood is constructed from annual seasonal minima, which are assumed to be independent between hydrological years for maximum-likelihood estimation. Serial dependence is addressed separately in the trend-testing step using the block bootstrap Mann-Kendall test, which accounts for autocorrelation when evaluating trend significance. The revised text now states:

(Line 259) The likelihood functions for the stationary and non-stationary WEI3 models are provided in Appendix A.

(Lines 567-580) The parameters of the three-parameter Weibull (WEI3) distribution are estimated by maximum likelihood separately for the summer and winter low-flow series. Let $\{q_t\}_{t=1}^{n}$ denote a seasonal sample of low-flow minima.

In the stationary case, the location parameter μ is constant over time, and the log-likelihood function is given by

$$\ell(\xi,\sigma,\mu) = \sum_{t=1}^{n}\left[\ln(\xi) - \ln(\sigma) + (\xi-1)\ln\left(\frac{q_t-\mu}{\sigma}\right) - \left(\frac{q_t-\mu}{\sigma}\right)^{\xi}\right], \tag{1}$$

defined for $q_t > \mu$, with $\sigma > 0$ and $\xi > 0$.

In the non-stationary case, temporal variation in low-flow behaviour is introduced through a time-dependent location parameter,

$$\mu(t) = \beta_0 + \beta_1 t, \tag{2}$$

while the scale and shape parameters are assumed constant. The corresponding log-likelihood function becomes

$$\ell(\xi,\sigma,\beta_0,\beta_1) = \sum_{t=1}^{n}\left[\ln(\xi) - \ln(\sigma) + (\xi-1)\ln\left(\frac{q_t-\mu(t)}{\sigma}\right) - \left(\frac{q_t-\mu(t)}{\sigma}\right)^{\xi}\right], \tag{3}$$

with $q_t > \mu(t)$ for all t.

In both cases, parameter estimation is carried out by minimising the negative log-likelihood subject to the constraints $\sigma > 0$, $\xi > 0$, and valid support of the distribution.

(11) Line 260. Suggest replacing "occurrence" with "non-exceedance".

Response to comment 11:

Thank you for this helpful suggestion. We agree that "non-exceedance probability" is the more precise term in this context, and we have replaced (Lines 277-278) As low flows represent minima, return periods are defined as the reciprocal of the ~~occurence~~ non-exceedance probability, $T = 1/p$.

(12) Line 268. The rdT statistic is a measure of a relative change in a return period at the quantile estimate defined by the non-stationary model. I am not sure this is really a performance gain. I do not think you can really assess performance of these methods unless you have a controlled experiment, e.g. Monte Carlo, where the true statistic is known. If you fit the WEI3 to the first half and second half of the record, you would also see some shifts in rdT, but it would not necessarily be a performance gain.

Response to comment 12:

We appreciate your perspective on this point. We agree that rdT should not be interpreted as a formal measure of predictive performance in the strict sense, since such an assessment would require a controlled experiment, for example a Monte Carlo simulation, where the true statistic is known. You are also correct that fitting the WEI3 distribution to different sub-periods could lead to changes in rdT, but such changes would reflect differences between time periods rather than a performance gain.

Our intention was to use rdT as a model-based diagnostic of how estimated return periods change when a model that explicitly accounts for violated assumptions is compared with a simpler reference model. The proposed model preserves the full sample and treats non-stationarity explicitly, which should allow for a more appropriate parameterisation than a simple stationary model when the stationarity assumption is violated. Restricting the analysis to a subsample, for example a potentially more stationary period near the end of the observation record, could partly address the non-stationarity problem. However, this would also reduce the available sample size and may increase estimation uncertainty, particularly for extreme low-flow quantiles.

We have therefore clarified the wording in the manuscript to avoid implying that rdT is a general performance metric. In particular, the interpretation of the deviation as a gain is only intended in the context of nested model comparisons, where the more complex model extends the simpler reference model and is assumed to provide a more appropriate representation of the underlying low-flow process. We have revised the manuscript as follows: (Lines 271-276) This interpretation reflects the fact that the true low-flow distribution is unknown and that models incorporating additional hydrological process information are considered conceptually superior to simpler representations. It should be noted that the interpretation of the deviation as a gain depends on the assumption that the model provides a more appropriate representation of the underlying low-flow process than the alternative model. The terminology is therefore used in the context of nested model comparisons, where the more complex formulation extends the simpler alternative.

(13) Also related to this, the lower tail location and shape of the Weibull distribution will probably have a large impact on this statistic, indicating larger or smaller values of rdT. Ultimately, I think this is better presented as a distributional shift than a performance gain.

Response to comment 13:

Thank you for this important comment. We agree that the lower-tail location and shape of the Weibull distribution can influence rdT, and that this should be interpreted carefully. We also agree with your main point that, when comparing different time periods, the resulting change should be described as a distributional shift or return-period change rather than as a performance gain.

Our use of the term "gain" is therefore restricted to the context of comparing nested models, where the more complex model explicitly accounts for violations of the assumptions of the simpler reference model. In this sense, a model that accounts for non-stationarity and process heterogeneity can provide a more appropriate basis for estimating probabilities than a stationary model applied under violated assumptions. We agree that non-stationarity can be interpreted as a distributional shift, which is explicitly represented in our model. The resulting change in estimated return periods is therefore primarily a model-implied distributional shift; only under the assumption that the non-stationary model better represents the low-flow process can it be interpreted as a gain relative to the stationary alternative.

We also see your point that the lower-tail location and shape of the Weibull distribution can influence rdT. However, based on the findings of the companion paper (Laaha, 2023a, b), where both the model assumptions and their effects on the estimates were assessed, the differences between the mixed distribution and the extended mixed-copula estimator were found to be negligible for the return periods of main practical interest. This supports the use of the mixed-distribution approach as a valid estimator of low-flow events in this setting.

Since the terminology is defined in Section 3.4, we retain the term "gain", but have clarified its meaning to avoid suggesting that rdT represents a formal predictive performance measure. Specifically, we added the following clarification: (Lines 273-276) It should be noted that the interpretation of the deviation as a gain depends on the assumption that the model provides a more appropriate representation of the underlying low-flow process than the alternative model. The terminology is therefore used in the context of nested model comparisons, where the more complex formulation extends the simpler alternative, and (Lines 291-292) When rdT is used to compare two different time points within the non-stationary model, it should be interpreted as a distributional shift, expressed as a change in estimated return period, rather than as a performance gain.

(14) Line 300. What about precipitation patterns? It appears that southern Europe is experiencing decreasing precipitation. Zeder D. and Fischer, E.M. (2020), Observed extreme precipitation trends and scaling in Central Europe, Weather and Climate Extremes, doi:10.1016/j.wace.2020.100266.

Response to comment 14:

Thank you for this helpful comment. We agree that precipitation patterns are an important driver of low-flow changes and have revised the discussion to make this point clearer. The study by (Zeder and Fischer, 2020) focuses on extreme precipitation in Central Europe and reports an overall intensification of annual, summer, and winter precipitation extremes with Northern Hemisphere temperature. This does not necessarily contradict the observed intensification of summer low flows, because low-flow conditions are not controlled by extreme precipitation alone, but by the seasonal water balance, including precipitation deficits, evapotranspiration, soil moisture depletion, and groundwater storage. We have therefore revised the text to acknowledge that summer low flows may reflect the combined influence of warming, enhanced evapotranspiration, drought intensification, and regional precipitation deficits, particularly in central Europe. For winter low flows, we retain the interpretation that responses are more mixed, reflecting interactions among cold-season precipitation, snow accumulation and melt, and regional hydrological storage processes. We have revised (Lines 322-324) These patterns suggest that summer low flows

are shaped by the combined effects of warming, enhanced evapotranspiration, drought intensification, and regional precipitation deficits, particularly in central Europe (Zeder and Fischer, 2020). Whereas winter low flows reflect a more mixed response linked to snow processes and cold-season hydrological dynamics.

(15) Lines 308-310. This is a great motivation for the rest of the study. This might benefit from reporting a field significance test regionally.

Response to comment 15:

We appreciate the reviewer's careful reading and constructive feedback. We fully agree that, if Europe were interpreted as a single trend region, a field significance test would be needed to account for repeated testing and possible spatial dependence among catchments. However, the objective of this study is not to infer a region-wide European trend, but to determine whether non-stationarity should be considered in the frequency model for each individual catchment. Here, we focus on individual site models, where potential non-stationarity is tested and evaluated at the catchment level. Since the analysis is applied on an individual model basis, we consider individual trend tests to be sufficient and more appropriate for the purpose of this study. We also state in the manuscript (Line 361-362) that the non-stationary model is preferred for 12.2% of catchments in summer and 7.7% in winter.

(16) Line 314. What do you mean by succeeded? Converged? Did you assess distributional fit/significance? Something else? Please explain.

Response to comment 16:

Thank you for pointing this out. We agree that the word "succeeded" was imprecise. We have revised the sentence to clarify that we refer to the convergence of the maximum-likelihood optimisation to feasible parameter estimates. (Lines 344-345) The WEI3 maximum-likelihood optimisation ~~succeeded~~ converged to feasible parameter estimates in 76.1 % of catchments.

(17) Line 320. State significance level of LR test, 5%?

Response to comment 17:

Thank you for this helpful suggestion. We have revised the text to explicitly state the significance level used for the LR test. The revised sentence now reads, (Line 358) Based on the likelihood ratio test with 5% significance level.

(18) Line 321. This seems to be evidence that a stationary model may be adequate except in unique situations, approximately 10% of your sites. Sometimes sites classified as undisturbed are disturbed when analysing low flows, which are more greatly impacted by small disturbances to the system. Could there be other drivers to these systems?

Response to comment 18:

Thank you for this insightful comment. We agree that the results indicate that a stationary formulation remains adequate for most near-natural catchments over the analysed period, and that the non-stationary framework is most relevant for the subset of catchments where both trend evidence and model selection support temporal change. We also agree that, although the ROBIN Level-1 network is designed to minimise direct human influence, low flows can be sensitive to relatively small catchment

disturbances. We have therefore revised the discussion to acknowledge these possible drivers more explicitly: (Lines 376-381). The observed spatial structure therefore suggests that the detected non-stationarity may be linked to differences in low-flow generating processes across regions and catchment types. Although the ROBIN Level-1 network minimises direct human influence, low-flow regimes are strongly controlled by catchment storage and release processes, including soil and aquifer characteristics, recharge, evapotranspiration, vegetation, topography, and climate (Smakhtin, 2001). Therefore, residual effects from small abstractions, groundwater interactions, land-use change, local water management, reservoirs, or flow regulation cannot be fully excluded and may contribute to temporal changes in some catchments.

(19) Line 326. I would remove the words "of course also".

Response to comment 19:

Thank you for the helpful suggestion. We agree that the wording could be improved and have removed the phrase "of course also" to make the sentence more precise. The revised text now reads, (Lines 364-365) In addition ~~of course also~~ the Type-I error of statistical tests has to be considered.

(20) Line 338. I would soften the statement "is both relevant and informative", as only 10% of your sites appear to fit this situation.

Response to comment 20:

Thank you for this helpful suggestion. We agree that the original wording could imply broader applicability than is supported by the results. We have therefore softened the statement as follows: (Lines 382-383) Overall, these results suggest that the proposed non-stationary framework is ~~both relevant and informative eventhough it is not required for all catchments~~ particularly relevant for catchments with evident non-stationarity, even though it is not required for all catchments.

(21) Line 356. Figure 4b shows return periods up to 1000 years. Does this make practical sense? In the US, we rarely look at return periods greater than 10 years for design purposes, usually 7Q10 and 7Q2. Is a 50, 100 or 1000 year return period used in practice? Given the three-parameter Weibull was used, the lower tail behaviour will be strongly influenced by the location parameter, i.e. lower bound. A lot of your discussion is about the changes in "rarer events". Is that practical?

Response to comment 21:

Thank you for this important comment. We agree that very high return periods, particularly up to 1000 years, should be interpreted with caution and are not intended as direct operational design values. They are shown mainly for methodological illustration, to visualise the lower-tail behaviour of the fitted distributions and the divergence among the annual, seasonal-mixture, and copula-based models. Therefore, return periods beyond 100 years should be interpreted as illustrative extrapolations rather than direct operational design values. We have revised the manuscript to make this distinction clearer:

(Lines 416-420) The practical interpretation focuses mainly on the 100-year return period, which can be used as a stress-test or worst-case scenario in water management and infrastructure planning (Hodgkins et al., 2024). In this study, the 100-year return period is used to compare the sensitivity of low-flow estimates to seasonality, dependence, and non-stationarity, while higher return periods are shown only for illustrative comparison of lower-tail behaviour.

Regarding the lower limit of the Weibull distribution. We agree that, for the three-parameter Weibull distribution, the lower-tail behaviour is strongly influenced by the location parameter, which represents the lower bound. We believe that this issue is becoming increasingly relevant in a non-stationary world, where climate warming and changing hydrological functioning of

river systems may lead to changes in the lower limits of previously seemingly stationary low-flow regimes. Both aspects, namely changes in the lower-tail behaviour and non-stationarity in the location parameter, are accounted for in our extended modelling approach. We consider this important for practical applications, including stress testing and purpose-oriented monitoring tools

(22) Line 362. Is "strongly correlated" the correct term? Is this an increased seasonal dependence?

Response to comment 22:

Thank you, we appreciate the reviewer's careful reading. We agree that "increased seasonal dependence" is more appropriate than "strongly correlated" in this context. We have revised the wording accordingly. The revised text now reads: (Line 425-427) In this respect, this behaviour is consistent with Laaha (2023a, b) showing that under moderate seasonal dependence, the copula-based correction has only a limited influence, whereas larger differences would be expected for more ~~strongly correlated catchments.~~ increased seasonal dependence between summer and winter low-flow components

(23) Line 363. Only one parameter is varying, not all parameters.

Response to comment 23:

Thank you for pointing this out. We agree that the original wording was too general, as non-stationarity is not introduced by allowing all distribution parameters to vary over time. We have revised the sentence to clarify the specific parameters that vary: (Lines 429-430) When non-stationarity is explicitly taken into account, ~~by allowing the model parameters~~ the seasonal WEI3 location parameter is allowed to vary over time, while the scale and shape parameters are kept constant. For the copula-based formulation, the dependence parameter is also time-varying (Sect. 3.2.2). The effect of this formulation is illustrated by the non-stationary probability plot in Fig. 4c.

(24) Line 395. Table 1 appears to represent the variation at different quartiles and not at the beginning and end of the observation period.

Response to comment 24:

Thank you, we agree that this point deserves further clarification. We agree that the wording could be clearer. We have revised the text, (Lines 470-472) First, for catchments with evident non-stationarity, we compare return-period estimates for the beginning (1976) and end (2020) of the observation period. The resulting relative absolute deviations are summarised across catchments, and are reported in Table 1, thereby quantifying the temporal shift captured by the non-stationary framework. These quartiles, therefore, describe variability across catchments rather than temporal quartiles.

(25) Line 445. What is a "considerable subset"? Suggest you provide the statistics here, approximately 10%. Also, I would mention these are unregulated sites.

Response to comment 25:

We appreciate this constructive suggestion. We have replaced the qualitative phrase "considerable subset" with the actual percentages and clarified that the analysis is based on near-natural catchments. The revised text now reads: (Line 522-523)

The results show that non-stationary behaviour is evident ~~(in a considerable subset of European catchments~~ in approximately 10% of the near-natural European catchments used in this study.

(26) Line 447. Is the location parameter change a shift in the central tendency or the lower bound?

Response to comment 26:

Thank you for pointing that out. We appreciate the reviewer's careful reading. You are correct that the WEI3 location parameter represents the lower bound, not the central tendency. We have revised the wording accordingly. (Lines 523-525) In the present application, non-stationarity is introduced through the location parameter, so temporal change is represented mainly as a shift in the ~~central tendency~~ lower bound of the seasonal low-flow distribution.

(27) Aside: While this analysis is interesting, it is complex. Might you get a similar result by just fitting the WEI3 to the last 15-20 years of the record? Vogel and Kroll (2020), A comparison of estimators of the conditional mean under non-stationary conditions, Advances in Water Resources, doi:10.1016/j.advwatres.2020.103672, tried to address this issue looking at non-stationary conditional means.

Response to comment 27:

We appreciate the reviewer's constructive suggestion and the reference to Vogel and Kroll (2020). We agree that, in some cases, restricting the analysis to a shorter and more recent part of the record may partly address non-stationarity, because the selected sub-period may violate the stationarity assumption less strongly than the full record. However, this approach necessarily reduces the available sample size, which may increase estimation uncertainty, especially for low-flow extremes and longer return periods.

The motivation of our framework is therefore to retain the full sample while treating non-stationarity explicitly within one coherent model. We acknowledge that this introduces additional model complexity, but we consider this complexity necessary to represent hydrological systems that are increasingly affected by climate change. Importantly, the practical implementation of the model is not intended to be complex for users: the code is provided in a Zenodo repository, and further synthesis into a dedicated R package is planned.

We have now added Vogel and Kroll (2020) to the Introduction (Lines 39-42) Vogel and Kroll (2020) made an important contribution to the increasing awareness of non-stationarity in hydrological frequency analysis by considering parsimonious estimators that focus on a more homogeneous recent sub-period. Such approaches can reduce the influence of non-stationarity by using records that better represent current conditions.

References

Bhatti, S. J., Kroll, C. N., and Vogel, R. M.: Revisiting the probability distribution of low streamflow series in the United States, Journal of hydrologic engineering, 24, 04019 043, 2019.

Hodgkins, G. A., Renard, B., Whitfield, P. H., Laaha, G., Stahl, K., Hannaford, J., Burn, D. H., Westra, S., Fleig, A. K., Araújo Lopes, W. T., et al.: Climate driven trends in historical extreme low streamflows on four continents, Water Resources Research, 60, e2022WR034 326, 2024.

Kendall, M.: Rank Auto Correlation Methods, 4th Edn., Griffin, 1976.

Laaha, G.: A mixed distribution approach for low-flow frequency analysis–Part 1: Concept, performance, and effect of seasonality, Hydrology and Earth System Sciences, 27, 689–701, 2023a.

Laaha, G.: A mixed distribution approach for low-flow frequency analysis–Part 2: Comparative assessment of a mixed probability vs. copula-based dependence framework, Hydrology and Earth System Sciences, 27, 2019–2034, 2023b.

Smakhtin, V. U.: Low flow hydrology: a review, Journal of hydrology, 240, 147–186, 2001.

Tramblay, Y., Rutkowska, A., Sauquet, E., Sefton, C., Laaha, G., Osuch, M., Albuquerque, T., Alves, M. H., Banasik, K., Beaufort, A., et al.: Trends in flow intermittence for European rivers, Hydrological Sciences Journal, 66, 37–49, 2021.

Turner, S., Hannaford, J., Barker, L., Suman, G., Killeen, A., Armitage, R., Chan, W., Davies, H., Griffin, A., Kumar, A., et al.: ROBIN: Reference observatory of basins for international hydrological climate change detection, Scientific data, 12, 654, 2025.

Vogel, R. M. and Kroll, C. N.: A comparison of estimators of the conditional mean under non-stationary conditions, Advances in water resources, 143, 103 672, 2020.

Zeder, J. and Fischer, E. M.: Observed extreme precipitation trends and scaling in Central Europe, Weather and Climate Extremes, 29, 100 266, 2020.

Reply to Reviewer 2 - Anonymous reviewer

Reviewer's general comments

The manuscript presents a assessment of low flow freq analysis using non-stationary extension of Weibull and copula based dependence framework. The framework is used for evaluating around 700 European catchments from the ROBIN dataset and demonstrates the implications of non-stationarity on low-flow frequency estimates, particularly for rare events. The study addresses an important topic given increasing concerns regarding climate-change induced nonstationary impacts on hydrological droughts and low-flow conditions. The manuscript is generally well structured and methodologically rigorous. However, some aspects require a bit of attention and explanation before publication. Here are some points to consider by the authors:

Response to Reviewer's general comments

Dear Reviewer, thank you very much for your careful evaluation of our manuscript, "Non-stationary low-flow frequency analysis with mixed Weibull components and Copula-based dependence framework" (https://doi.org/10.5194/egusphere-2026-1820), and for your constructive and encouraging feedback. We appreciate your positive assessment of the relevance, structure, and methodological rigour of our study.

We have carefully considered all of your comments and suggestions and have revised the manuscript accordingly to improve the clarity of the methodological choices, assumptions, interpretation, and limitations of the proposed framework. Please find below our point-by-point responses to each comment in italic, followed by the corresponding revisions made to the manuscript. In the revised text, additions and modifications are shown in blue.

Sincerely yours,

Farhana Sweeta Fitriana, Svenja Fischer, Gabriele Weigelhofer, Johannes Laimighofer and Gregor Laaha

Reviewer's specific comments

(1) The introduction provides a good overview of non-stationary frequency analysis literature, however authors should more clearly explain why existing approaches (e.g. GAMLSS-based low-flow frequency analysis) are insufficient and specifically why the proposed extension of LFAmix and LFAmixC is necessary in the set up which they have used.

Response to comment 1:

We thank the reviewer for this helpful comment. We agree that the distinction between the proposed framework and existing approaches to non-stationary low-flow frequency analysis should be stated more clearly. GAMLSS-based approaches provide a flexible framework for representing temporal variation in distribution parameters through time or other covariates. However, in seasonal climates, non-stationarity is only one part of the modelling problem, because annual low-flow minima may originate from either summer or winter processes that differ substantially in their physical controls and statistical distributions, and these seasonal components may also be dependent. The stationary LFA_{mixC} framework was developed specifically to represent this seasonal process heterogeneity and the dependence between summer and winter low flows (Laaha, 2023b). However, it does not allow the seasonal marginal distributions or their dependence structure to change over time. The contribution of the proposed $NSLFA_{mixC}$ framework is therefore not simply to introduce time-varying distribution parameters, but to extend the established process-based seasonal mixture framework to non-stationary conditions. In this way, seasonal heterogeneity, inter-seasonal dependence, and temporal change can be represented consistently within the same low-flow frequency-analysis framework. The consideration of covariates, which makes the GAMLSS a frequently applied model, can also be realised for

the NSLFA$_{mixC}$ model by simply replacing the time parameter by any other covariate. A further practical distinction concerns the Weibull formulation.

We have revised the Introduction accordingly and added the following clarification (Lines 58-61):

GAMLSS-based approaches provide a flexible framework for representing temporal variation in the distribution. However, they do not address the additional process heterogeneity that arises when annual low flows originate from distinct seasonal generating processes, nor the dependence that may exist between these seasonal components.

(2) The manuscript adopts WEI3 throughout, but should the justification or advantages/disadvantages of using alternative distributions be included. Since low-flow frequency analysis often employs Lognormal, Generalized Logistic, Pearson Type III or GEV formulations, authors should discuss why WEI3 is uniquely suitable and whether results are sensitive to this choice.

Response to comment 2:

We appreciate the reviewer's thoughtful comment on the choice of marginal distribution. In the present study, we deliberately retain the WEI3 formulation because of its theoretical basis for lower extremes and its consistency with the established LFA$_{mixC}$ framework. We nevertheless agree that examining alternative marginal distributions would be useful for assessing the broader applicability of the framework in future work. In response to a related comment from Reviewer 1, we also analyse an L-moment ratio diagram for the stationary summer and winter low-flow samples (Fig. 1). The empirical L-skewness and L-kurtosis values show reasonable agreement with the theoretical Weibull relationship, further supporting the use of WEI3 for the seasonal margins. At the same time, apparent departures from a fitted Weibull distribution at individual catchments may not necessarily reflect the choice of distribution alone, but may also indicate sample properties such as process heterogeneity. In particular, annual low-flow minima may be generated by different summer and winter mechanisms, so that a single annual distribution may be less appropriate than a framework that explicitly represents the seasonal components. This is one of the main motivations for the NSLFA$_{mixC}$ approach, and previous studies by Laaha (2023a, b) have shown the benefit of explicitly accounting for seasonal process heterogeneity in low-flow frequency analysis. For this reason, the present study focuses on extending the theoretically motivated WEI3-based seasonal mixture and copula framework to non-stationary conditions, rather than conducting a catchment-specific comparison of alternative marginal distributions. We therefore retain WEI3 as the theoretically motivated marginal distribution in the present study. Extending the framework to other marginal distributions in future work would provide a useful test of the generality of the proposed modelling approach.

We have revised the manuscript accordingly. In Lines 341-343, we added: Alternative low-flow distributions may also provide suitable representations at individual catchments. However, the present study retains WEI3 because of its theoretical basis for lower extremes and its consistency with the established LFA$_{mixC}$ framework.

We also added the following statement to the limitations and future work section (Lines 551-554): In addition, the present framework is developed using WEI3, supported by its theoretical basis for lower extremes and its established use in low-flow frequency analysis. Extending the framework to alternative marginal distributions to consider local deviations from this assumption would provide a useful assessment of its broader applicability.

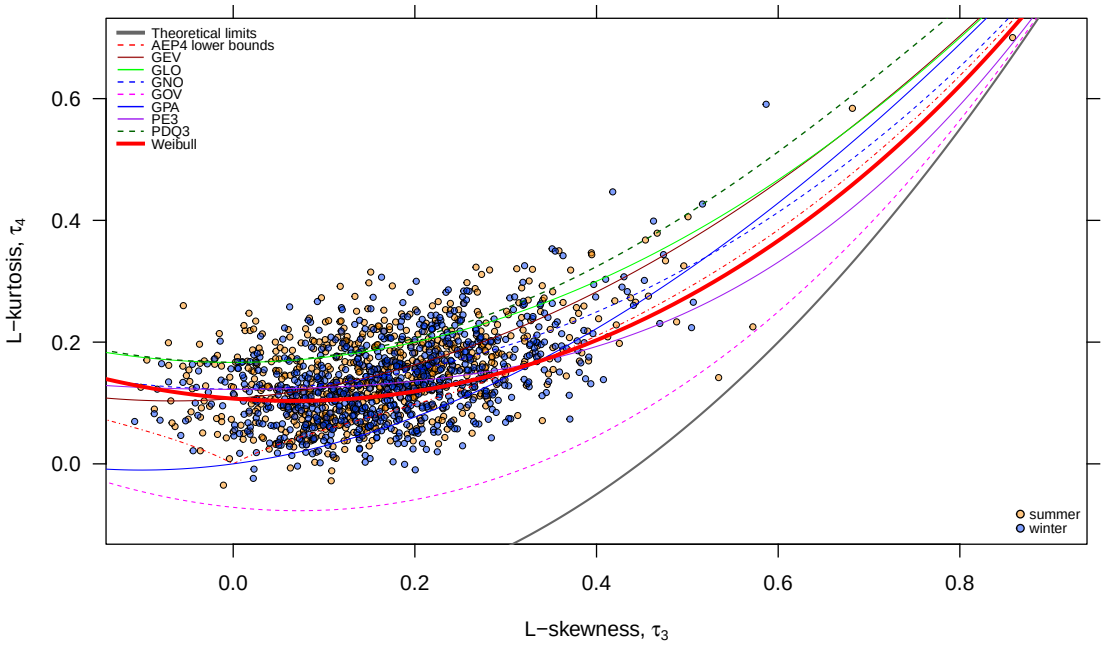

Figure 1. L-moment ratio diagram for summer and winter low-flow series. Points represent empirical L-skewness (τ_3) and L-kurtosis (τ_4) of the seasonal low-flow samples, while the curves show theoretical L-moment relationships for candidate distributions, including the Weibull distribution.

(3) Non-stationarity is only introduced through the location parameter. While the authors justify this using parsimony arguments, climate change can influence both variability and shape of low-flow distributions. Understandable that this will make the model very more sensitive and risky. . . .some discussion or such probabilistic sensitivity analysis is recommended.

Response to comment 3:

Thank you for highlighting this point. We appreciate the reviewer's thoughtful comment on how we represent non-stationarity. We agree that climate change may affect not only the location of low-flow distributions, but also their variability and potentially their tail behaviour. In the present study, however, we deliberately restrict non-stationarity to the location parameter as a parsimonious first-order formulation. This choice is motivated by both statistical and hydrological considerations. Allowing additional distribution parameters to vary independently over time would substantially increase model complexity and estimation uncertainty, particularly given the limited length of the annual seasonal low-flow series, as also noted by the reviewer. Moreover, previous scaling analyses have shown strong relationships among the moments of streamflow series (Vogel and Sankarasubramanian, 2000), suggesting that changes in variability may not necessarily occur independently of changes in the mean (Serago and Vogel, 2018). Allowing the scale parameter to vary independently with time could therefore introduce additional flexibility without accounting for these structural relationships. Particular caution is also required for the shape parameter, since temporal variation in shape would imply changes in tail behaviour. Such changes are more difficult to estimate robustly and to relate directly to physical climatic drivers, particularly when inference concerns rare low-flow events.

Introducing an independently time-varying shape parameter would therefore further increase parameter uncertainty and the risk of overfitting. We therefore fully agree with the reviewer that such extensions of the model would affect uncertainty and sensitivity of the model. At the same time, we agree that changes in variability deserve further investigation. An important next step is therefore to examine whether temporal changes in variability can be represented more parsimoniously as consequential changes associated with shifts in the mean, rather than by introducing an independently time-varying scale parameter. Such an approach could account for changing variability while preserving the structural relationships among streamflow moments and limiting unnecessary parameterisation. We have clarified this point further in the Discussion and Limitations sections by emphasising the parsimonious rationale for the location-only formulation and by identifying the investigation of variability changes associated with shifts in the mean as an important next step.

We have revised the manuscript accordingly. In Lines 545-548, we added: An important next step is to investigate whether temporal changes in low-flow variability can be represented parsimoniously as consequential changes associated with shifts in the mean, rather than by independently varying additional distribution parameters, while remaining consistent with the underlying physical processes driving change.

(4) The framework assumes a linear temporal evolution in location and dependence parameters. However, low-flow behaviour often exhibits regime shifts, accelerations, thresholds, or oscillatory behaviour linked to climatic variability. If not nonlinear relationships related analysis is done, at least this should be discussed in detail.

Response to comment 4:

We appreciate the reviewer's thoughtful comment on the assumed temporal form of non-stationarity. We agree that low-flow behaviour may exhibit more complex dynamics, including regime shifts, accelerations, thresholds, or oscillatory variability. The reason for limiting the model to a linear trend was already motivated in the manuscript and has now been clarified further. In brief, the linear formulation provides a simple and parsimonious representation of gradual temporal change while avoiding unnecessary model complexity. Given the limited record length, more flexible nonlinear formulations, or additional time-varying distribution parameters, would require stronger empirical support and could substantially increase estimation uncertainty and the risk of overfitting. The linear temporal component is also supported by the preceding trend analysis, in which monotonic change in the seasonal low-flow series was assessed using the block-bootstrap Mann-Kendall test. We then evaluated the non-stationary formulation directly against the simpler stationary model using a likelihood-ratio test, and the results were generally consistent. The linear formulation should therefore be interpreted as a parsimonious first-order representation of historical change rather than as an assumption that low-flow behaviour is intrinsically linear. We have expanded the discussion to clarify that more complex temporal dynamics may occur and that nonlinear formulations, change-point approaches, or climatic covariates could be considered where sufficiently long records and supporting evidence are available.

We have now added a discussion of the model's motivation. (Lines 383-393) Its value lies not in universally replacing stationary models, but in identifying and representing catchments where seasonal low-flow regimes exhibit temporal change. In this case, the stationarity assumption may be violated and should not be ignored in the frequency analysis. To keep the framework parsimonious, non-stationarity was introduced as a linear trend in the location parameter and was evaluated formally by comparing the non-stationary and stationary models using a likelihood-ratio test. The linear formulation was chosen as a simple and parsimonious way to represent gradual temporal change. Given the limited record length, more flexible non-linear formulations or time-varying scale and shape parameters would require stronger empirical support and could substantially increase estimation uncertainty and the risk of overfitting. More complex temporal dynamics, including regime shifts, accelerations, thresholds, or oscillatory behaviour linked to climatic variability, may also occur. Nonlinear formulations, change-point approaches, or climatic covariates could therefore be considered where sufficiently long records and supporting evidence are available.

(5) *The use of PAM clustering and seasonality histograms is interesting, but the physical interpretation of the resulting clusters needs stronger discussion. For example, how sensitive are the results to the chosen clustering approach, silhouette threshold (0.2), or seasonal definitions (April–October and November–March)?*

Response to comment 5:

Thank you for pointing this out, and we appreciate the suggestion. We have strengthened the discussion of the physical interpretation of the seasonality clusters and added sensitivity analyses in the Appendix. The PAM analysis identified two dominant seasonal patterns, with medoid peaks in February (winter-type) and August (summer-type). Catchments with silhouette widths below 0.2 were classified as mixed, resulting in 82.50% summer-type, 14.44% winter-type, and 3.06% mixed catchments. The classification was only weakly sensitive to the silhouette threshold, using 0.15 and 0.25 instead of 0.20 retained 98.47% and 98.89% agreement with the baseline, respectively, with differences mainly affecting the mixed class (Table. 1). As an additional robustness check, hierarchical clustering was applied. Agreement with PAM was high (96.67% agreement in the summer, winter, and mixed classifications), and both approaches yielded the same physical interpretation, with representative cluster profiles peaking in February and August. Differences were confined to weakly affiliated or mixed catchments, with no direct switches between summer- and winter-type regimes. The April-October and November-March seasonal definition follows the established continental-scale European convention of Laaha et al. (2017), as already stated in the manuscript Lines 116-118. This definition does not determine the PAM clusters, which are based on the complete 12-month seasonality histograms, but is used for their physical interpretation. Shifting either seasonal boundary by one month did not change the cluster interpretation, resulting in 100% agreement with the baseline (Table 2). Overall, the dominant summer-winter seasonality structure is robust to the methodological choices considered, while the small differences are limited to catchments with weak or transitional seasonality. We have added the following text to the Results and Discussion section: Lines 310-313: The classification is robust to the choice of silhouette-width threshold, with thresholds of 0.15 and 0.25 retaining 98.47% and 98.89% agreement, respectively, with the baseline threshold of 0.2. Changes primarily affect the proportion of catchments classified as mixed domain. Additional sensitivity analyses of the clustering approach and seasonal definition similarly support the robustness of the dominant summer-winter structure, with detailed results provided in Appendix B.

We have also added the detailed sensitivity results in Appendix B:

Table 1. Sensitivity of the low-flow seasonality classification to the silhouette-width threshold used to identify catchments with mixed seasonality. The baseline threshold is $s < 0.20$. Agreement refers to the percentage of catchments retaining the same classification as under the baseline threshold.

	Threshold	Winter (n)	Winter (%)	Summer (n)	Summer (%)	Mixed (n)	Mixed (%)	Agreement (%)
1	0.10	109	15.14	605	84.03	6	0.83	97.78
2	0.15	106	14.72	603	83.75	11	1.53	98.47
3	0.20	104	14.44	594	82.50	22	3.06	100.00
4	0.25	102	14.17	588	81.67	30	4.17	98.89
5	0.30	94	13.06	575	79.86	51	7.08	95.97

Table 2. Sensitivity of the physical interpretation of the PAM clusters to the seasonal definition. The baseline defines summer as April-October and winter as November-March.

	Seasonal Definition	Cluster 1	Cluster 2	Agreement (%)
1	Apr-Oct / Nov-Mar	Winter	Summer	100.00
2	May-Oct / Nov-Apr	Winter	Summer	100.00
3	Apr-Sep / Oct-Mar	Winter	Summer	100.00

(6) WEI3 fitting was successful in only 76.1% of catchments. This implies nearly one-quarter of stations could not be fitted successfully. May be the characteristics of these excluded catchments should be analysed and discussed, as this may introduce regional bias into the results.

Response to comment 6:

Thank you for this valuable suggestion. We further examined the catchments for which WEI3 fitting was unsuccessful to determine whether they show any clear spatial or hydrological pattern. Of the 721 catchments, WEI3 fitting was successful for 549 catchments (76.1%) and unsuccessful for 172 catchments (23.9%). The unsuccessful fits were distributed across the study domain rather than being concentrated in one particular region, while the successfully fitted catchments still provide broad spatial coverage across Europe (Fig. 2). We also compared catchment area, mean elevation, and low-flow characteristics, and distributional shape between the two groups (Table 3). Mean catchment elevation was similar, whereas unsuccessfully fitted catchments tended to be smaller, with a median area of 91.5 km^2 compared to 163.6 km^2 for successfully fitted catchments. These catchments also generally had lower specific low flows in the annual, summer, and winter series and less year-to-year variability. For example, median annual M7Q was 1.69 L s^{-1} km^{-2} for unsuccessful fits compared to 2.80 L s^{-1} km^{-2} for successful fits. We also examined which low-flow series were affected by fitting failures. The most common pattern was a winter-only failure, occurring in 54 catchments (31.4% of unsuccessful fits), followed by combined annual and summer failures in 49 catchments (28.5%). The remaining failures were distributed across the other seasonal combinations, indicating that the fitting difficulties were not limited to one particular series. The unsuccessful catchments also showed differences in distributional shape. Their low-flow series exhibited consistently higher positive L-skewness, compared with successfully fitted catchments. To further investigate whether distributional shape is related to fitting difficulties, we separately examined individual stationary WEI3 fits and estimated the Weibull shape parameter independently using L-moments. Among the series for which maximum-likelihood fitting was unsuccessful, 39-59% had an L-moment shape estimate below one, compared with approximately 3% of successfully fitted series. A shape parameter below one is relevant for the three-parameter Weibull because the likelihood can become increasingly difficult to optimise as the estimated location parameter approaches the lower endpoint of the observations. These results therefore suggest that the fitting difficulties occur more frequently for low-flow series with higher skewness and lower Weibull shape parameters. More generally, likelihood-based estimation can become non-regular when a distribution's support depends on estimated parameters (Coles et al., 2001). This issue is relevant to the three-parameter Weibull because its lower endpoint depends on the location parameter. Overall, the additional analysis shows that unsuccessful fits are spatially dispersed and do not indicate a clear regional concentration. They occur more frequently, however, in smaller catchments with lower low flows and more positively skewed low-flow distributions. The distributional-shape analysis provides a plausible explanation for the numerical difficulties encountered when estimating the three-parameter Weibull via maximum likelihood. Earlier low-flow studies have reported similar convergence issues (Vogel and Kroll, 1989). Although these results help explain the observed fitting difficulties, a more detailed investigation of the numerical behaviour of the WEI3 likelihood is beyond the scope of this study and would be an interesting topic for further methodological research.

We have added this discussion to the revised manuscript:Lines 345-355: The remaining optimisation failures are likely related to the constrained parameter space, which may yield infeasible parameter combinations. The unsuccessful fits were distributed throughout the study domain rather than being concentrated in a single geographical region, while the successfully fitted catchments retained broad spatial coverage across Europe. Unsuccessfully fitted catchments tended to be smaller and generally had lower low-flow values relative to catchment area in the annual, summer, and winter series, with less year-to-year variability. Mean catchment elevation was similar between the two groups. The unsuccessful catchments also showed consistently higher positive L-skewness, indicating differences in the distributional shape of their low-flow series. These characteristics may contribute to the greater difficulty in estimating the WEI3 parameters within the constrained parameter space. Similar convergence issues have been reported in earlier low-flow studies (Vogel and Kroll, 1989), suggesting that they arise from the estimation procedure rather than a general limitation of the WEI3 distribution itself. We also provide the corresponding spatial comparison and detailed catchment, low-flow, and distributional characteristics in Appendix D.

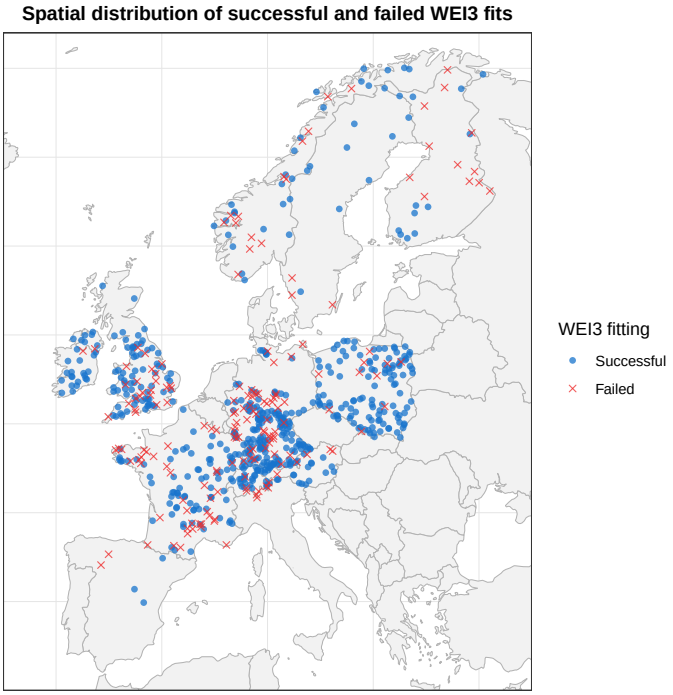

Figure 2. Spatial distribution of catchments with successful and unsuccessful WEI3 fits across the European study domain.

Table 3. Comparison of catchment characteristics, low-flow characteristics and distributional shape between catchments with successful and unsuccessful WEI3 fits. Values are reported as median [Q1-Q3]. Low-flow values are normalized by catchment area. p-values are based on Wilcoxon rank-sum tests.

Characteristic	Successful	Unsuccessful	p-value
Catchment area (km^2)	163.63 [83.60-494.50]	91.50 [44.15-247.21]	< 0.001
Mean catchment elevation (m)	349.82 [172.39-599.30]	350.00 [125.28-587.44]	0.291
Low-flow magnitude and variability			
Median annual M7Q ($L\,s^{-1}\,km^{-2}$)	2.80 [1.60-4.21]	1.69 [0.85-3.31]	< 0.001
Median summer M7Q ($L\,s^{-1}\,km^{-2}$)	2.97 [1.64-4.62]	1.93 [0.87-3.76]	< 0.001
Median winter M7Q ($L\,s^{-1}\,km^{-2}$)	4.86 [2.96-7.71]	3.27 [2.06-5.83]	< 0.001
IQR annual M7Q ($L\,s^{-1}\,km^{-2}$)	1.26 [0.86-2.00]	0.87 [0.51-1.39]	< 0.001
Distributional shape			
Annual L-skewness	0.13 [0.06-0.19]	0.22 [0.12-0.32]	< 0.001
Summer L-skewness	0.14 [0.08-0.20]	0.23 [0.13-0.32]	< 0.001
Winter L-skewness	0.16 [0.09-0.22]	0.26 [0.17-0.34]	< 0.001

(7) The paper focuses heavily on one Austrian catchment. Authors should consider presenting at least one additional example, preferably from a different characteristics dominated regime, to demonstrate framework performance across contrasting hydro-climatic conditions.

Response to comment 7:

Thank you for this suggestion. To complement the Austrian example and demonstrate the application of the NSLFA$_{mixC}$ framework under contrasting hydro-climatic conditions, we added a second example using Oksfjordvatn in Norway (Fig. 3). Stauf Vöckla at Traun represents a mixed seasonal low-flow regime (mixture rate = 0.52) with moderate inter-seasonal dependence (Kendall's τ = 0.48). At this catchment, non-stationarity is mainly associated with the summer component, while the winter component remains stationary. Oksfjordvatn provides a contrasting case: it is winter-dominated (mixture rate = 0.40), has weaker inter-seasonal dependence (Kendall's $\tau = 0.26$), and both the summer and winter components were identified as non-stationary. The two catchments also show contrasting temporal behaviour. At Stauf Vöckla, summer low-flow magnitudes decrease over time and the non-stationary Q_{100} estimates decline, indicating increasingly severe low-flow conditions. At Oksfjordvatn, both summer and winter low-flow magnitudes increase over the observation period, and the corresponding Q_{100} estimates also increase, indicating less severe low-flow conditions toward the end of the record. These two examples therefore illustrate how the framework can represent different seasonal regimes, dependence structures, and directions of temporal change.

We have added the following text to the manuscript:

Lines 452-463: To further illustrate the framework under contrasting hydro-climatic conditions, we analysed the catchment Oksfjordvatn in Norway as a second example. Unlike Stauf Vöckla, which represents a mixed seasonal regime (mixture rate = 0.52), Oksfjordvatn is winter-dominated (mixture rate = 0.40) and exhibits weaker inter-seasonal dependence (Kendall's $\tau = 0.26$, compared with 0.48 at Stauf Vöckla). The temporal behaviour also differs between the two catchments. At Stauf Vöckla, non-stationarity is mainly associated with declining summer low flows, whereas at Oksfjordvatn both summer and winter components were identified as non-stationary and low-flow magnitudes increase over time. Consequently, the non-stationary Q_{100} estimates decrease through time at Stauf Vöckla, indicating increasingly severe low-flow conditions, but increase at Oksfjordvatn, indicating less severe low-flow conditions toward the end of the observation period. The corresponding four-panel analysis for Oksfjordvatn is provided in Appendix E. Overall, both examples demonstrate how the NSLFA$_{mixC}$ framework captures seasonal contrasts, temporal shifts, and dependence effects within a single coherent analysis. It also illustrates that non-stationarity may alter not only the magnitude of design low flows, but also the relative importance of summer and winter contributions to the annual low-flow regime.

***(8)** Statements such as a 100-year event becoming a 43-year event are powerful but may be misinterpreted by readers. It should be clarified that these changes are statistical diagnostics based on the fitted non-stationary model and should not be interpreted directly as future projections.*

Response to comment 8:

We appreciate the reviewer for highlighting this point. We agree that the comparison between the 100-year return period in 1976 and the approximately 43-year return period in 2020 should be interpreted carefully. This comparison is based on the fitted non-stationary model and describes how the estimated return period of the same event changes between two points within the historical observation period. It is not intended as a projection of future event frequency. We have therefore clarified this interpretation in the manuscript by explicitly stating that the comparison represents a statistical diagnostic of historical change and should not be interpreted as a future projection.

We have revised the manuscript accordingly. In Lines 448-451, we added: In particular, an event corresponding to a 100-year return period in 1976 is associated with a return period of about 43 years by 2020, indicating that such an event became approximately on average 2.3 times as frequent according to the fitted non-stationary model. This comparison is a statistical diagnostic of change within the historical observation period and should not be interpreted as a projection of future event frequency.

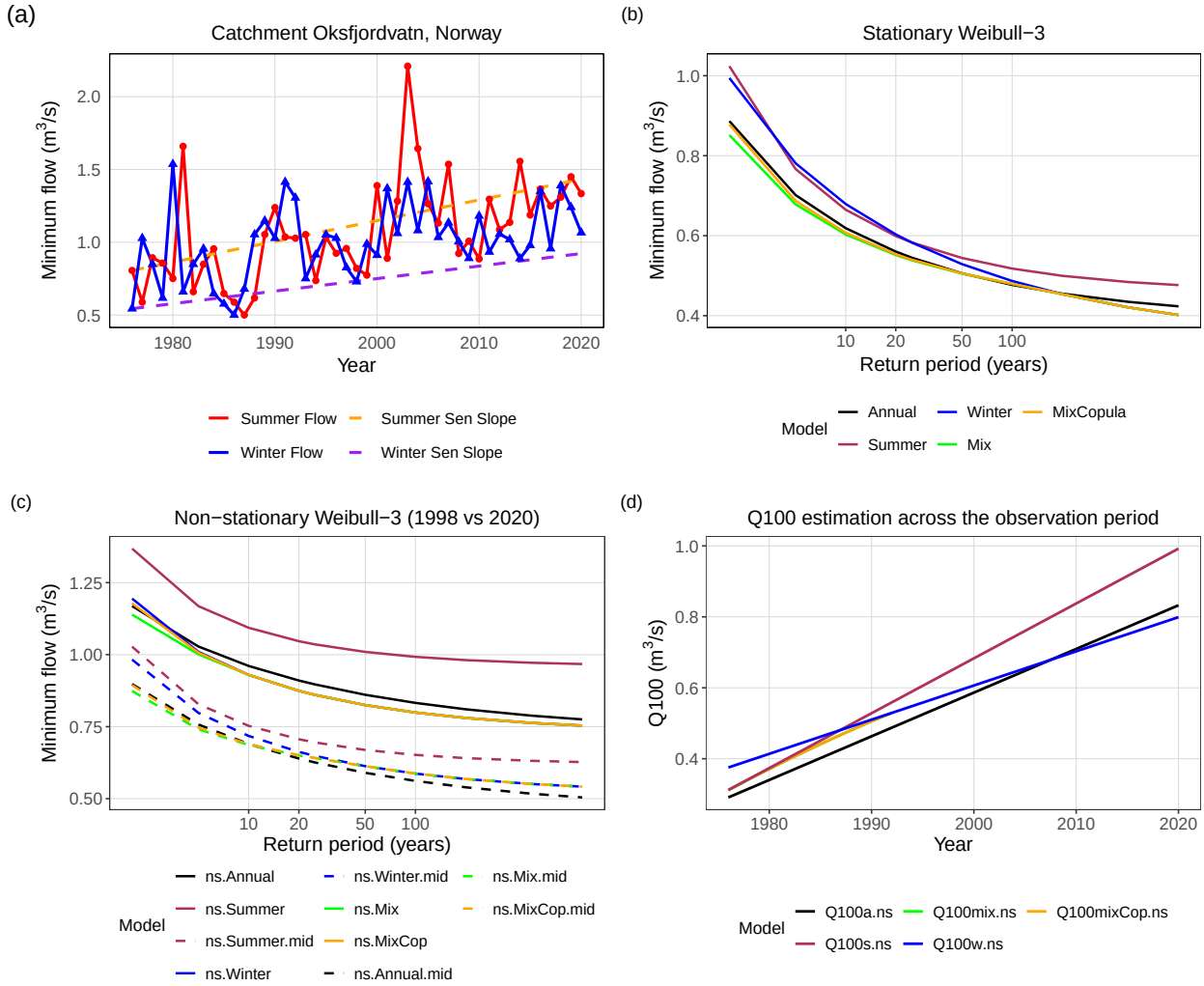

Figure 3. Application of the $\text{NSLFA}_{\text{mixC}}$ framework to Oksfjordvatn, Norway. (a) Observed summer and winter minimum flows with corresponding Sen's slope trends. (b) Stationary WEI3 low-flow frequency estimates for the annual, summer, winter, mixture, and mixture-copula models. (c) Non-stationary low-flow frequency estimates evaluated at the midpoint (1998) and end (2020) of the observation period. (d) Evolution of the estimated 100-year low-flow quantile (Q_{100}) over the observation period.

References

Coles, S., Bawa, J., Trenner, L., and Dorazio, P.: An introduction to statistical modeling of extreme values, vol. 208, Springer, 2001.

Laaha, G.: A mixed distribution approach for low-flow frequency analysis–Part 1: Concept, performance, and effect of seasonality, Hydrology and Earth System Sciences, 27, 689–701, 2023a.

Laaha, G.: A mixed distribution approach for low-flow frequency analysis–Part 2: Comparative assessment of a mixed probability vs. copula-based dependence framework, Hydrology and Earth System Sciences, 27, 2019–2034, 2023b.

Laaha, G., Gauster, T., Tallaksen, L. M., Vidal, J.-P., Stahl, K., Prudhomme, C., Heudorfer, B., Vlnas, R., Ionita, M., Van Lanen, H. A., et al.: The European 2015 drought from a hydrological perspective, Hydrology and Earth System Sciences, 21, 3001–3024, 2017.

Serago, J. M. and Vogel, R. M.: Parsimonious nonstationary flood frequency analysis, Advances in Water Resources, 112, 1–16, 2018.

Vogel, R. M. and Kroll, C. N.: Low-flow frequency analysis using probability-plot correlation coefficients, Journal of water resources planning and management, 115, 338–357, 1989.

Vogel, R. M. and Sankarasubramanian, A.: Spatial scaling properties of annual streamflow in the United States, Hydrological sciences journal, 45, 465–476, 2000.