EGUSPHERE-2024-2594
Detailed Responses to Editor Comments

Anne Martin, Élyse Fournier and Jonathan Jalbert

May 29, 2025

Specific comments

1. L3: I recommend moving the WMO definition from the abstract to the introduction.
   Done. The definition is now at L20.

2. L16: The subheading "Context" is not necessary; consider removing it.
   Done. Thank you for the suggestion.

3. L17–25: The introduction currently focuses too heavily on dam safety, whereas PMP has broader applications in hydrological risk assessment. As HESS is not specifically a journal for dam safety research, I suggest revising this section to present PMP in a broader hydrological context, with dam safety as one example among others.
   Done. Thank you for the suggestion (L15).

4. L35–37: The relevance of PMF is unclear. If it is not directly used in the model development or not specifically mentioned later in the manuscript, I recommend removing it. On a different, more general comment - the introduction could be revised to be less Canada-specific and more broadly applicable to international readers.
   (a) Done. There is stil one mention of the PMF to understand where the PMP can be used (L19).
   (b) We appreciate the suggestion and understand the value of making the introduction broadly applicable. However, we believe it is important to present the Canadian context in which our study was conducted, as it provides relevant background and motivation. We trust that international readers, with a clear understanding of this context, will be able to adapt the proposed method to their own settings.

5. L58–59: Eq. 2 may not be necessary. You could start discussing the ratio on line 60, referencing Eq. 1, which should suffice for clarity.
   Done. Thank you for the suggestion (L54).

6. Section 1.5: This section might be better integrated into the introduction. If the main point is to emphasize that climate models can be used to estimate changes in PMP (you

are not using climate models later in the paper, just discussing their use), I suggest shortening the section and incorporating it earlier in the text.

We have reduced the emphasis on modeling non-stationarity in the paper's objectives, as it was not required for our study (L114). However, we believe the paragraph on climate models remains important. Even without modeling non-stationarity, several sources cited in that section recommend using simulated data from climate models to address the scarcity of observational data for PMP estimation. We have incorporated elements of this discussion in the revised version of the manuscript (L99).

7. L108: Please clarify the acronym "CC" on first use.
Thank you for pointing that out.

8. L124–128: Consider removing this text; it does not appear essential.
Done. Thank you for the suggestion.

9. L128–129: The sentence regarding data availability would be more appropriately placed in the data availability section at the end of the manuscript. Can be removed here.
Thank you for pointing that out.

10. Structural suggestion: You might consider presenting the proposed PMP estimation model (Section 3) before introducing the case study data (Section 2). Since the model is general, introducing it first may improve the manuscript's logical flow. However, this is a suggestion; you may ignore it and choose to retain the current structure.
Thank you for the suggestion. The methodology is now the Section 2 and the simulation study appears in the Section 3.

11. Figures 1 and 2: These figures are not "a must" and could be moved to the supplementary material as Figures S1 and S2 (supplementary material in a separate file, not as an appendix).
Figures 1 and 2 were introduced in response to Reviewer 3's comments. These figures help justify that incorporating non-stationarity is not necessary for the analyzed data, as no trend is visible in either precipitation or precipitable water. While we agree that they are not essential, we believe they still have value and merit inclusion in the manuscript. Please let us know which option you consider preferable.

12. Table 2: This table could be merged with Table 1 to avoid redundancy in presenting station-specific information.
Both tables are merged into Table 2. Thank you for the suggestion.

13. L159–160: Please provide additional information for the EVT analysis: What threshold was selected? Why not use annual maxima? Which distribution was fitted (e.g., GEV)? Consider adding a brief sensitivity analysis on the threshold and distribution choice.
Additional information has been provided in Section 4.2, which now presents the EVT results. To summarize, while the block maxima approach could have been used, we adopted the Peaks-Over-Threshold (POT) methodology in accordance with the recommendations of the National Academies of Sciences, Engineering, and Medicine (2024).

The threshold was selected using the mean residual life plot, as suggested by Coles (2001) (L281). In the POT methodology, exceedances above the high threshold are modeled using the Generalized Pareto distribution (L280). We have also added uncertainty estimates for both the parameters (Table 4) and the return levels (L282, L349, L353).

14. Section 3.2.1: The method of moments is well established (see your references to Johnson et al.). Consider moving large parts of this section (or entirely) to the supplementary material unless some parts of it are critical to your main argument.
    Thank you for pointing that out. Section 2.2.1 of the revised version has been shortened. We retained the key points of the method of moments as applied to our model, the Pearson Type I distribution, which are:
    - the moments have closed-form analytical expressions;
    - the third and fourth moments do not depend on the upper bound, i.e., the PMP;
    - the estimates have tractable forms.
15. Section 5: Subsections are not strictly necessary here; the results can be presented as a single, continuous narrative.
    Done. Thank you for the suggestion.
16. L304–307 and L319–321: The parameter estimates would be more clearly presented in a table.
    The parameters estimates for Montréal and St-Hubert can now be found in Table 3. Thank you for the suggestion.
17. Figure 7: Please clarify why quantiles close to 100 are missing in panel (a), and why quantiles exceed 100 in panel (b).
    Figure 7 shows the quantile–quantile plots. For St-Hubert, the highest empirical quantile is 106 mm, which is why the axes exceed 100 mm. We chose to use quantile–quantile plots here because they place more emphasis on the tail, in contrast to probability–probability plots, where the axes are limited to the unit interval, emphasizing the bulk of the distribution.
18. L314–315: It is somewhat limiting to present the case study using only one of the three proposed estimation methods you described. You demonstrate them using the synthetic data, but it is better also to demonstrate using real case studies. I suggest including an additional case study, possibly using data from another region (you should not limit yourself to Canada), where all estimation methods are applied and compared. This would provide a more robust demonstration of the model's applicability.
    Estimates obtained using maximum likelihood and the Bayesian method have not been included, as doing so would contradict the conclusions of the simulation study. The simulation study demonstrated the poor performance of these two approaches due to a lack of parameter identifiability. For the Montréal data, the PMP maximum likelihood estimate exceeds 2×10^{13} mm. Elements of this discussion have been incorporated into the revised version of the manuscript in L368–381, where we address the regularized

likelihood method. The absurd PMP maximum likelihood estimate is specifically discussed there.
We are not comfortable adding a case study from another region. We do not necessarily have the expertise or contextual knowledge required to provide a critical analysis of results from other areas. We prefer to thoroughly describe the specific context of our study, which we understand well, and allow readers to assess whether the proposed method is adaptable to their own context. Furthermore, including a few examples from other regions would not, in itself, demonstrate the method's universality across all contexts. However, if you believe this is absolutely necessary, we would be willing to make the effort.

19. Section 6.2: The EVT results for the observed data are currently presented late in the manuscript. I recommend moving this material to Section 2.2, as it fits naturally with the presentation of other PMP estimates (see my above comments about the EVT).
For the manuscript flow, it is true that the EVT results should appear in the same table with the two other PMP estimation methods. It has now been moved to this section (Table 2 and L279-282).

20. L376–377 and for the second station later: Please summarize the GPD parameter estimates for each station in a table for clarity.
The parameter estimates can now be found in Table 4. Thank you for the suggestion.

21. L416–417: The recommendations of the "National Academies of Sciences, Engineering, and Medicine" are specific to the Canadian or U.S. context. I suggest rephrasing such statements to make them more general and globally relevant.
It is true that the recommendations address the North American context, but they stem from the broader observation that the definition and assumptions underlying PMP are outdated. This criticism is globally relevant, as the definition and assumptions of PMP are consistent worldwide. Therefore, we believe the recommendation can be considered general. Elements of this discussion have been added in the revised version of the manuscript (L400-403).

EGUSPHERE-2024-2594
Detailed Responses to Reviewer's 2 Comments

Anne Martin, Élyse Fournier and Jonathan Jalbert

May 29, 2025

I appreciate the authors' efforts in revising the manuscript, particularly their acknowledgment of the limitations of their proposed method for practical applications and their alignment with the National Academies of Sciences, Engineering, and Medicine (NASEM, 2024) recommendation to use an Extreme Value Theory (EVT)-based approach. However, the study's primary contribution lies more in its critical evaluation of PMP methodologies rather than in the statistical model itself. While previous studies have highlighted the flaws in the WMO-recommended PMP estimation methods, this study uniquely examines the limitations of physically based moisture maximization through a statistical framework before ultimately recommending an EVT-based approach. That said, I still have few concerns regarding the robustness of the conclusions, primarily due to the limitations of the EVT analysis and the inadequate discussion of sampling uncertainty.
Concerns:

1. The challenge of estimating very long return periods (e.g., 10,000 years) using only 75 years of data remains unresolved. Such extreme quantile estimates require substantial extrapolation, which increases uncertainty. Moreover, the stability of estimation may be sensitive to the threshold selection in the Peaks-Over-Threshold (POT) method. I suggest going over (NASEM, 2024) report about sampling uncertainty. The authors should explicitly discuss these uncertainties in the EVT approach to prevent misinterpretation by end-users. Otherwise, there is a risk of conveying an overconfident message about the reliability of these estimates.

   It is true that EVT-based PMP estimates are preferable, but they do not resolve all challenges. Extrapolating beyond the data range, especially for large return periods associated with PMP estimates, remains difficult and introduces substantial uncertainty. Such return level estimates should be accompanied by uncertainty evaluations (e.g., confidence intervals) to clearly communicate to end-users that PMP estimates carry substantial uncertainty inherent to extrapolation. The methodology based on simulated data, presented in Section 1.4 to address data scarcity, could also be adapted within the extreme value framework. Elements of this discussion have been added to the revised manuscript and confidence intervals have been added where it was not given in the previous version of the manuscript.

2. The authors acknowledge that identifiability issues affect the reliability of PMP estimates, yet the discussion remains largely theoretical. The manuscript would benefit from a practical demonstration in Section 6.3 of how alternative constraints—such as regularized maximum likelihood estimation or Bayesian priors—influence PMP estimates. If these methods were tested but found ineffective, the authors should clearly articulate why. This would strengthen the argument against using Pearson Type I for PMP estimation.

   Thank you for the suggestion. We added two paragraphs in Section 6.3 (L368–381) describing the method we developed to address non-identifiability issues. While the method proved effective, we do not encourage its use because it introduces a high level of subjectivity into the analysis; something we aimed to avoid with the proposed statistical model for PMP.