Responses to Review Comments to 'Characterising runoff processes for Australia: Insights from a parsimonious rainfall-runoff event identification method' (EGUSPHERE-2025-6241)

Our responses are in blue and revisions are underlined.

Reviewer 1

The manuscript introduces the Robust Variance-based Event Identification Method (RVEIM), a new and efficient approach for pairing rainfall and runoff events in hydrological studies. Traditional statistical methods often suffer from subjective parameter selection, leading to inconsistent results and physically impossible outcomes, such as runoff volumes exceeding rainfall. By utilizing streamflow variance and a time-variant search window, RVEIM minimizes these uncertainties and better reflects natural catchment responses. When tested across diverse Australian catchments, the method proved more stable than traditional local maxima benchmarks, showing significantly less variation in event characteristics. Overall, the method shows promise for applications in flood analysis and hydroclimatic studies. The study addresses an important methodological gap, and the proposed framework is potentially valuable. The manuscript is generally well written. However, several issues should be addressed to further strengthen the work prior to publication.

We thank the reviewer for the positive assessment of our work and for the constructive feedback provided.

Major Comments:

- The paper presents RVEIM as a significant improvement over existing methods, yet it primarily benchmarks it against a single "local maxima" method. While the authors justify this choice by citing the local maxima method's relatively lower uncertainty compared to others, the paper would be significantly strengthened by comparing RVEIM directly to the variance-based method by Fischer et al. (2021). Since RVEIM is an extension of that specific framework, a direct comparison is necessary to isolate and quantify the specific value-add of the authors' new contributions, such as the simultaneous pairing and the time-variant search window. Furthermore, Table 1 identifies several other sophisticated methods (e.g., DMCA or modified baseflow filters) that are not addressed in the evaluation; including at least one other modern automated method would provide a more rigorous validation of RVEIM's "transferability" and "robustness" across Australia's diverse hydro-climates.
  We agree that a broader benchmarking would strengthen the evaluation of *RVEIM*. In the original manuscript, we selected the *local maxima* method as a widely used and representative benchmark. However, we acknowledge that additional comparisons can provide a more comprehensive assessment of robustness and transferability.
  Following the reviewer's suggestion, we have included an additional benchmarking to strengthen the evaluation of our proposed approach. Fischer et al. (2021) and Giani et al. (2022) developed two relevant but distinct frameworks. Fischer et al. (2021) identify runoff events and subsequently pair them with rainfall events in a separate step, whereas RVEIM performs event identification and pairing simultaneously. In contrast, the Detrending Moving-average Cross-correlation Analysis (DMCA) method developed by Giani et al. (2022) has a similar level of complexity to *RVEIM* as it has only two parameters (*rain_min*, and *max_window*) and simultaneously detects and pairs rainfall-runoff events. Hence, we decided to include DMCA as additional benchmarking method to compare our results with for the 8 representative

catchments. We have included the relevant results in the revised Section 3.2, 3.3, and 4.2. Specifically, we demonstrated:

1. DMCA generally shows uncertainty comparable to RVEIM across most catchments and event characteristics, and both methods are substantially less sensitive to parameter changes than local maxima. However, DMCA exhibits greater variability in some cases, like the mean length of paired runoff event for catchment A0020101, whereas RVEIM maintains more consistently low uncertainty across catchments (Figure 6).
2. DMCA produces the lowest proportion of rainfall-runoff events with RCs > 1, with values close to zero across the representative catchments. RVEIM also maintains a very low proportion of RCs greater than 1 (generally below 4%). These results show that both DMCA and RVEIM provide physically plausible event identification, although DMCA performs slightly better according to this specific criterion (Figure 7).
3. Both DMCA and RVEIM produce considerably narrower uncertainty band of ECDF of RCs than local maxima, indicating lower uncertainty in parameter value change. Nevertheless, their median ECDFs differ substantially in some wetter catchments. DMCA identifies a large concentration of rainfall-runoff events with RCs ≈ 0, whereas RVEIM identifies far fewer rainfall-runoff events with RC ≈ 0. Therefore, RVEIM represents these catchments as having a greater potential to convert event rainfall into runoff. Given the relatively wet climatic setting of these catchments, the representation produced by RVEIM appears more hydrologically plausible (Figure 8).

Overall, the comparison shows that DMCA performs well in terms of robustness against parameter value and the avoidance of physically implausible RCs; however, RVEIM provides a more plausible representation of runoff generation behaviour in wetter catchments.

- In Equation 2, the search window length is defined using best_lag, which is calculated as the average time between the start of a rainfall event (as shown in a hyetograph) and the start of a runoff event (as shown in a hydrograph). However, it is not clearly explained how this parameter is calculated in practice. In particular, in the procedure used to identify the start times of rainfall and runoff events, as illustrated in Figure 4, best_lag is not shown as how it is calculated.

  Thank you for this helpful comment. We have revised lines 165–170 as follows:

  > *"3. Cross-Correlation. A runoff event is paired to the corresponding rainfall event based on a time-varying search window controlled by two parameters (as shown in Eq. 2). The first parameter, which is constant across rainfall events, is best_lag. This parameter is defined as the lag between rainfall and streamflow that maximizes the cross-correlation between these variables (and is calculated prior to rainfall event detection). Physically, best_lag represents the average response time between rainfall and runoff. The second parameter controlling the length of the search window is l_rainfall, which corresponds to the duration of the rainfall event. Therefore, this parameter varies from one rainfall event to another."*

  As discussed above, *best_lag* is estimated before identifying rainfall events. Therefore, the procedure to estimate *best_lag* is not explicitly shown in Fig. 4, since the figure focuses on the runoff event identification procedure after rainfall events are defined.

Minor Comments:

- The subplot labeling is inconsistent across figures (e.g., use of "a" vs. "(a)" in Figures 6 and 8). Please standardize. Also, their font size is not consistent, e.g. Figure 9.

Thank you for this comment. We have standardized subplot labeling across all figures and ensured consistent font sizes throughout all figures to improve clarity and visual consistency.

Reviewer 2

This manuscript provides a very useful contribution by introducing a novel, physically based and parsimonious event identification method that exhibits excellent robustness. This offers a standardised and transferable approach for event identification that forms the starting point of much hydrological analysis. However, some aspects of the methodology would benefit from further elucidation.

We sincerely thank the reviewer for the positive assessment of our work and for the constructive and insightful comments provided. We particularly appreciate the recognition of the methodological contribution and robustness of the proposed approach. Below, we provide a detailed point-by-point response to each comment.

Major Comments:

- The algorithm flow chart in Figure 2 is a very useful inclusion that I commend, however, please consider the following:
    1. Consider spacing out the boxes within the "Streamflow statistics" window to allow better interpretation of the flow arrows.
    2. The left arrow from the first diamond "n_exceed >= len_rain & length(QD[idx_candidate]) >0" is misleading as it says, "We have no streamflow event" yet it flows to "Calculate streamflow event characteristics".
    3. For both diamonds it is unclear under what condition you should follow which arrow.

Thank you for these helpful suggestions regarding the flowchart in Figure 2.

1. We agree that the spacing within the "Streamflow statistics" section can be improved. In the revised figure, we have increased the spacing between elements and reorganize the layout so that the pathways to the green boxes are vertically aligned, improving readability and flow interpretation.
2. We acknowledge that the left arrow from the first decision diamond may be misleading. Our intention was to represent cases where rainfall does not generate a detectable runoff event. In such cases, we still consider this as part of the rainfall-runoff process, with resulting event characteristics (runoff volume, duration, and runoff coefficient) equal to zero. To avoid confusion, we have added a label to the left-hand arrow in the flowchart to clarify that it represents rainfall events that do not generate a detectable runoff event.
3. We agree that the decision logic within the diamonds was not sufficiently clear. In the revised flowchart, we have explicitly labelled each branch using "True" and "False" conditions. This allows readers to clearly understand under which conditions each pathway is followed and improves the overall interpretability of the flowchart.

- Please clarify the basis for the *3 scaling factor in equation 4.
  Thank you for this comment. This choice of *3 scaling factor is consistent with robust approaches based on median absolute deviation (MAD), where thresholds in the range of 2–3 are commonly used (Voloh et al., 2020). In this study, we adopt a factor of 3 as a conservative threshold to ensure that only substantial changes which exceed the normal range of streamflow variability are identified, while avoiding sensitivity to minor fluctuations.
  The scaling factor of 3 is also consistent with commonly adopted outlier detection methods which work based on standard deviation (SD). In such methods, values beyond ±3 standard deviations are typically considered rare and indicative of significant deviations from typical

behaviour (Fahrnberger, 2019; Santacroce, 2020). We have added a sentence in lines 188–189 as follows:

> *"… baseflow-dominant condition. The multiplier of 3 (in Eq. 4) was adopted because this threshold is commonly used in outlier detection procedures. The variance of baseflow values …"*

- Can you please add further explanation as to why n_exceed >= len_rain is a necessary condition. Are there any situations, such as on a small catchment, where this may not be the case? Does this add any uncertainty? I am not suggesting any further analysis is required, rather just some further discussion. For example, my understanding is that "best_lag" will vary based on the catchment and this should ensure that on average n_exceed >= len_rain. Some words to this effect would be beneficial, I think.

  Thank you for this insightful comment. We have added a sentence in lines 207–209 to make this clear:

  > *"… search window. This condition reflects the expectation that a rainfall event leading to runoff should produce a sufficiently sustained streamflow response, rather than a short-lived or spurious fluctuation which might be due to instrumental errors."*

  The condition n_exceed ≥ len_rain is introduced to ensure that the identified streamflow response is sufficiently sustained relative to the duration of the corresponding rainfall event. Conceptually, this reflects the expectation that a rainfall event leading to runoff should generate a streamflow signal that persists for at least a comparable duration, due to processes such as storage, routing, and recession, rather than a short-lived or spurious fluctuation (e.g., due in instrumental error etc.)

  Empirical evidence supports the expectation that runoff duration is typically equal to or longer than the duration of the corresponding rainfall event, even in relatively small catchments. For example, in a small headwater catchment in Hong Kong, Zhang et al. (2023) showed that rainfall duration ranged from 24–662 min (mean 108 min), while runoff recession alone ranged from 73–591 min. This reflects the persistence of streamflow due to routing and recession processes following rainfall. Therefore, the adopted condition represents a reasonable and physically consistent expectation for rainfall-runoff events.

  This constraint on *len_rain* is not determined by "*best_lag*", which is estimated separately from the rainfall-runoff relationship for each catchment and shows the typical timing of the runoff response.

- Line 254-255: I'm concerned about the lack of sensitivity testing for alpha in the local maxima event identification. Whilst I appreciate that it would unlikely change the results significantly or the conclusion of the study, I think for completeness alpha should either be sampled for both or neither algorithm, as it will have an impact on event identification.

  Thank you for this important comment. We have included the *alpha* parameter sensitivity for local maxima and updated figures and analysis throughout the manuscript (Table 3, Fig 6, 7, and 8). Overall, the inclusion of sensitivity to changes in *alpha* did not alter the sensitivity results of the *local maxima* or the comparative conclusion, with RVEIM continuing to exhibit consistently lower uncertainty.

- Could you please clarify the ranges of parameters for Local maxima in Table 3, particularly for delta.y? Often the delta.y parameter is specified as a proportion of the peak to account for different catchment scales, however, here it is flow value. Given the range of catchment sizes investigated, i.e. 102-105, is it possible that the uncertainty of the local maxima algorithm is inflated by sampling across an implausible range for some catchments? I also note that in the author's cited paper from 2025, in Table 1 the parament range for delta.y in the local maxima algorithm is the proportion 0.9-0.7 whilst the delta.y for a different algorithm, local minima, is 0-30 which is what is used in this manuscript. Can you please clarify this?

Thank you for this careful and important observation. We acknowledge that there was an error in reporting the range of the *Delta.y* parameter in Table 3. We have fixed this issue in Table 3 of the revised manuscript.

The correct range of *Delta.y* is between -0.9 and -0.7 (Mohammadpour Khoie et al., 2025), which represent a proportional decrease of 0.9 to 0.7 from the local maxima. Importantly, this error was limited to the reporting in Table 3 and did not affect the implementation of the *local maxima* algorithm in our analysis. The parameter sampling was conducted using the correct proportional definition, and therefore the reported uncertainty is not inflated by an implausible parameter range.

- Looking at the results in Figure 5 is it possible that RVEIM is truncating the start of events by starting at i, where QDi is first positive, which is resulting in the shorter event lengths and lower volumes in Figure 6 b) & c), and typically lower RC values in Figure 7)? Could you consider starting at i-1 instead?

  Thank you for this insightful comment. We acknowledge that, as illustrated in Fig. 5, some runoff events may appear visually truncated at their start. This is primarily a consequence of discretizing continuous streamflow signals into discrete time steps, rather than a limitation in the event identification concept itself. Shifting the starting point one time step backward only marginally affects event characteristics (As shown in Fig. R1), typically resulting in a one-step increase in event duration and a slight increase in event runoff volume, without altering the overall patterns or key conclusions of the study. To assess the impact of this choice, we have repeated the analysis for representative catchments by shifting the starting index of runoff events one time step backward. These changes do not alter the overall patterns or key conclusions of the study.

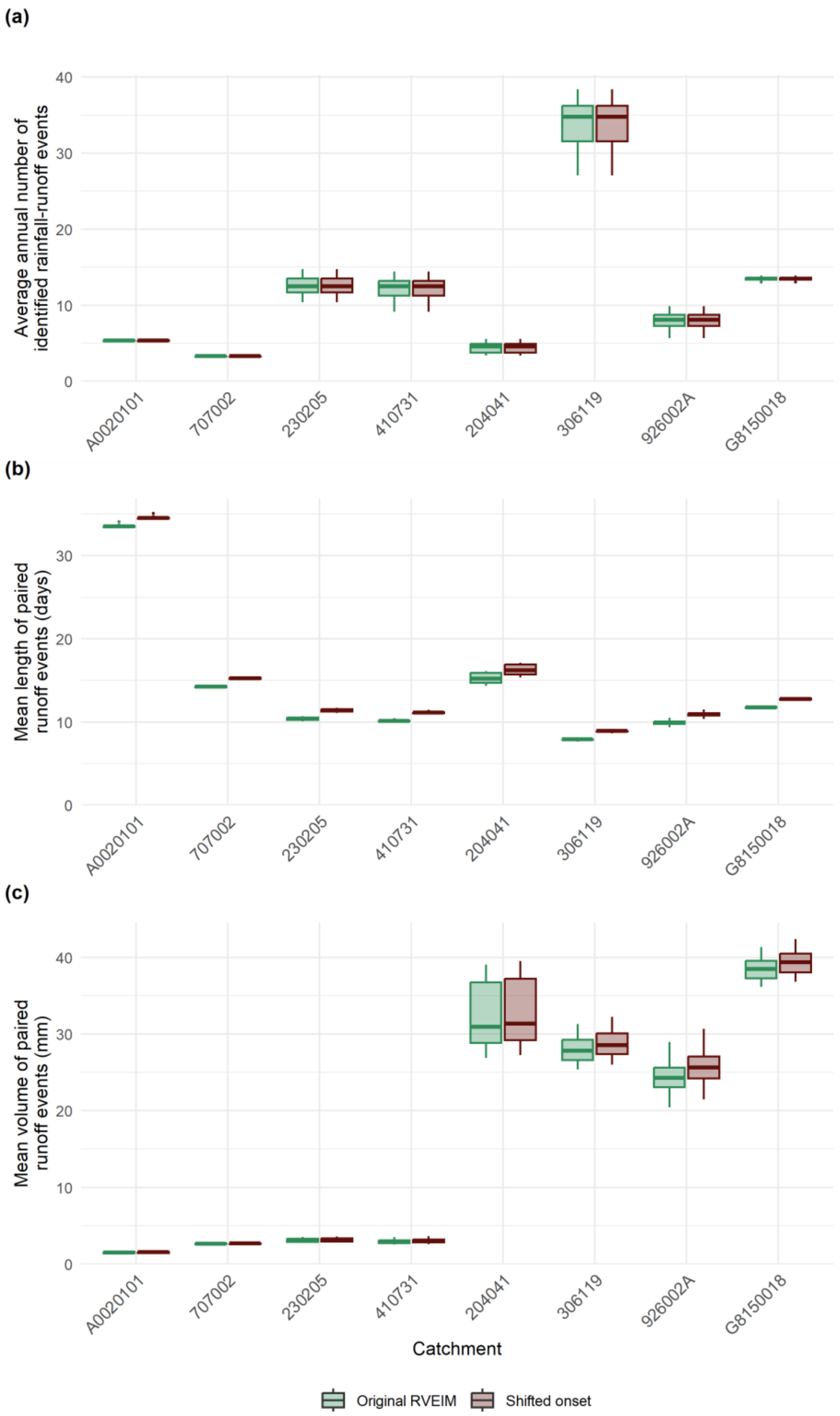

Fig. R1. Comparison of characteristics of rainfall-runoff events identified by RVEIM and an onset shifted one day earlier

We have clarified this by adding relevant discussion in lines 316–320 as:

> *"Although the same runoff event identification rules are applied consistently to all events, some events may appear visually truncated, such as the third and fourth runoff events in Fig. 5 (shown in green). This mainly reflects the temporal discretization involved in delineating runoff events from daily data. Shifting the runoff event start one time step earlier could make the runoff events appear more visually complete, but it produces only negligible changes in event characteristics and may incorrectly double-*

count time steps by creating overlap between the beginning of one runoff event and the end of the preceding event."

- I think the results shown in figure 5.2 should be moved out of the discussion and up into the results.

  We agree that the results shown in Fig. 11 (the figure presented in Section 5.2 of the original version, which we assume you are referring to) are presenting additional results. However, as this figure provides supplementary analysis supporting the main findings rather than a core result, we have moved Fig. 11 to the appendix A (now as Fig. A5) to improve the flow and readability of the main Discussion section.

- It would be worth commenting in the discussion whether RVEIM would generalise to use on subdaily data.

  Thank you for this valuable suggestion. We have added a discussion in the revised manuscript to clarify the potential applicability and limitations of *RVEIM* for sub-daily data. Please see lines 524–528:

  > *"RVEIM could, in principle, be applied to sub-daily data because its identification rules are based on physically meaningful changes in streamflow. However, because runoff event detection is initiated from independently identified rainfall events, the definition and consistent delineation of rainfall events at sub-daily resolutions would be critical. Although the underlying rules should remain physically applicable, the greater variability and noise at finer temporal resolutions may require adjustments to parameter definitions and thresholds."*

Minor Comments:

- Please review the manuscript and ensure correct spelling and grammar is used throughout.

  We have carefully reviewed the entire manuscript and corrected grammatical errors, typos, and wording where necessary to improve clarity and readability.

- Please rephrase lines 98-100 "We demonstrated that the proposed rainfall-runoff event identification method leads to noticeably lower percentage of RCs better reflecting physical processes" as it is hard to interpret.

  We have revised the sentence as (lines 97–99):

  > *"We show that the proposed rainfall-runoff event identification method maintains a low proportion of physically implausible RCs, while producing substantially lower uncertainty across parameter choices, indicating greater methodological robustness."*

- Please check that the correct conversion of units has been undertaken for Mean annual streamflow in Table 2.

  Thank you for your careful attention. We have revised the table caption as follows to add clarity:

  > *"Table 2. Hydroclimate and catchment-specific characteristics of selected catchments. Mean annual streamflow is expressed as an equivalent depth (mm/day) obtained by dividing gauged flow in ML/day by catchment area."*

  In this study, streamflow is originally measured in ML/day. However, in Table 2, for the purpose of comparison across catchments, we converted mean annual streamflow to an equivalent depth (mm) by normalizing by catchment area. This allows direct comparison with mean annual rainfall and provides a more intuitive interpretation of the proportion of rainfall contributing to runoff across different catchments.

References

Fahrnberger, G. (2019). Outlier removal for the reliable condition monitoring of telecommunication services. 2019 20th International Conference on Parallel and Distributed Computing, Applications and Technologies (PDCAT),

Fischer, S., Schumann, A., & Bühler, P. (2021). A statistics-based automated flood event separation. *Journal of Hydrology X*, *10*, 100070.

Giani, G., Tarasova, L., Woods, R. A., & Rico-Ramirez, M. A. (2022). An objective time-series-analysis method for rainfall-runoff event identification. *Water Resources Research*, *58*(2), e2021WR031283.

Mohammadpour Khoie, M. M., Guo, D., & Wasko, C. (2025). Improving the consistency of hydrologic event identification. *Environmental Modelling & Software*, *191*, 106521. https://doi.org/https://doi.org/10.1016/j.envsoft.2025.106521

Santacroce, M. (2020). SCREEN: Stata module to quickly identify possible outliers based on the interquartile range, percentile or standard deviation.

Voloh, B., Watson, M. R., König, S., & Womelsdorf, T. (2020). MAD saccade: Statistically robust saccade threshold estimation via the median absolute deviation. *Journal of eye movement research*, *12*(8), 65.

Zhang, L., Xu, Q., Wang, K., Chen, G., Luo, Z., Li, X., & Guo, B. (2023). Event-scaled hydrological response of a headwater catchment in Hong Kong [Original Research]. *Frontiers in Environmental Science*, *Volume 11 - 2023*. https://doi.org/10.3389/fenvs.2023.1218239