Reply to RC2 - Anonymous Referee #2

The manuscript presents an hourly, at 2 km resolution blended rainfall dataset (BRAIN) for Australia, derived from satellite, radar, and gauge observations. Given the current lack of a national hourly gridded rainfall product, this dataset addresses a clear need. The manuscript is generally well written, and the figures are clearly presented.

Response: Thank you for your positive assessment of our work and for recognising the significance of developing a national hourly, 2 km blended rainfall dataset for Australia.

However, the methodology used in the study, including the Optimal Interpolation (OI) and Bias Correction, was developed and applied in other studies for the same purpose of generating blended rainfall datasets from satellite, radar, and gauge data (e.g., Yu et al. 2020 (blended data in the UK), Xia et al. 2025 (blended data in China)). There is no significant improvement in this study when the method was applied to Australia.

Response: Thank you for this comment. We agree that the general methodological concepts of bias correction and optimal interpolation have been adopted in previous multi-source rainfall blending studies, including Yu et al. (2020) and Xia et al. (2025). Our intention was not to claim that bias correction or optimal interpolation itself is newly developed in this study.

In addition, the specific implementation in this study differs from previous applications. Yu et al. (2020) used PDF matching for bias correction and applied optimal interpolation to blend three data sources over the relatively smaller UK domain. Xia et al. (2025) first used a machine-learning method to combine satellite and radar data into a blended background field, followed by a two-source optimal interpolation procedure. In contrast, our study uses a Kalman-filter-based bias correction approach and a simplified three-source optimal interpolation framework. The simplification, including the nearest-radar approximation, is important for applying the method efficiently over the large domains, such as Australia.

In addition to these implementation differences, the contribution of this work lies in adapting and implementing the multi-source blending framework for national-scale, hourly, 2 km rainfall analysis across Australia, where the observing network is highly heterogeneous and where no national hourly gridded rainfall product is currently available. The Australian context presents specific practical challenges, including uneven gauge density, spatially variable radar coverage, extensive data-sparse regions, and the need for an efficient framework suitable for operational implementation. We believe this study also provides a strong continental-scale case study of multi-source precipitation blending that may serve as a useful reference for similar large-domain operational applications.

To address this comment, we will revise the manuscript to more clearly distinguish the contributions of this study from previous work. We will cite the relevant previous studies more clearly, including Yu et al. (2020) and Xia et al. (2025), and clarify that the contribution of this work lies in the implementation, adaptation, and evaluation of a computationally efficient national-scale multi-source precipitation blending framework for Australia, rather than the introduction of a new blending concept.

In addition, the use of inverse distance weighting (IDW) interpolation as a benchmark raises some concerns. Describing IDW as "the operational method in Australia for estimating hourly rainfall" may be somewhat misleading. IDW is a relatively simple interpolation technique, typically used for observing purposes rather than for producing authoritative rainfall datasets. Established Australian products such as AGCD, SILO, and ANUClimate employ more sophisticated interpolation methods, albeit primarily for daily timescales. As such, comparisons against IDW may not provide a sufficiently robust benchmark. The conclusion that BRAIN outperforms gauge data using the IDW interpolation is not convincing and can mislead users of the dataset.

Response: Thank you for this comment. Our intention was not to establish IDW as the primary benchmark, but rather to use it as a simple gauge-only reference to illustrate the added value of incorporating multiple observation sources. We also agree that established Australian rainfall products, such as AGCD, SILO, and ANUClimate, use more advanced interpolation methods. However, these products are designed primarily for daily or longer timescales and are not directly comparable to a near-real-time national hourly rainfall analysis. We also acknowledge that comparisons against a simple IDW interpolation alone are insufficient to demonstrate the superiority of the proposed framework. Accordingly, we will moderate our conclusions and avoid making broad statements that BRAIN outperforms gauge-based interpolation in general.

To address this comment, we will revise the manuscript by:

- Revising the description of IDW to clarify that it is used only as a simple gauge-only reference.
- Reducing the emphasis on the comparison with IDW and presenting it as an illustrative example instead of the primary benchmark.
- Moderating the conclusions throughout the manuscript to avoid overstating the advantages of BRAIN based solely on the comparison with IDW.
- Expanding the discussion of the limitations of IDW and the differences between IDW and established Australian rainfall products, including AGCD, SILO, and ANUClimate.

Unfortunately, I suggest rejecting the manuscript for publication in HESS at this stage. I hope that the authors will revise the method and validation processes after the trial. While the dataset has potential value, I suggest submitting the manuscript to another dataset journal.

Response: Thank you for your review and comments. We will thoroughly revise the manuscript by addressing all of the comments raised, including expanding the validation, improving the discussion of uncertainties, and clarifying the methodology.

We respectfully believe that this study is appropriate for Hydrology and Earth System Sciences. Although this study adopts the general concepts of bias correction and optimal interpolation, the contribution of this work lies in the implementation and evaluation of a computationally efficient national-scale multi-source precipitation blending framework for Australia, addressing the country's unique observational and operational challenges. In addition, this work can serve as an additional useful reference for other similar studies, which is also why research in this area continues to be actively developed and published.

In addition, once the long-term dataset has been completed, we plan to prepare a separate data paper for submission to an appropriate dataset journal.

Major comments

1. The introduction, section 2 and discussion sections would benefit from a more thorough review and discussion of potential errors and artifacts in the satellite and radar datasets. I suggest adding a few paragraphs to that outline the limitations across all three data sources used in the blending process, the issue that these errors may propagate into the final blended product, and the point-area spatial mismatch, even after the bias correction (Ochoa-Rodriguez et al, 2019).

Response: Thank you for this valuable suggestion. We agree that a more comprehensive discussion of the uncertainties and limitations associated with the input datasets is necessary. We also agree that errors in the satellite, radar, and gauge observations may propagate into the final blended product, and that issues such as the point-to-area spatial mismatch remain important considerations, even after bias correction.

To address this comment, we plan to make the following revisions in the manuscript:

- Introduction: Expand the literature review to provide a more balanced discussion of the strengths and limitations of the three primary precipitation data sources (satellite, radar, and rain gauges), together with the associated challenges for multi-source precipitation blending.
- Data section: Add more detailed descriptions of the limitations and uncertainty characteristics of each dataset used in this study, including the Himawari satellite precipitation estimates, Rainfields radar product, and sub-daily gauge observations.
- Results and Discussion: Include a dedicated discussion of the potential errors and artefacts associated with each input data source, how these uncertainties may propagate through the statistical interpolation framework, and the implications for the final blended precipitation estimates. We will also discuss the point-to-area spatial mismatch and other related issues (Ochoa-Rodriguez et al., 2019).

- Uncertainty analysis: We will also consider adding uncertainty-related diagnostics, either in the main manuscript or the Supplementary Material, to provide readers with additional information on the relative contributions of different observation sources and the confidence of the blended precipitation estimates.

2. IDW should not be used an operational benchmark for hourly gauge rainfall data in the validation. In the paper, the authors claim that "the gauge-based interpolation approach currently used in flood operations". In practice, IDW is a simplistic method, typically suitable for preliminary visualisation of rainfall patterns rather than for operational flood applications. To my knowledge, IDW is not used to generate regional or national daily or sub-daily rainfall products. In Australia, widely used datasets such as AGCD, SILO, and ANUClimate rely on more advanced interpolation techniques. At the hourly scale, more robust approaches have also been explored—for example, thin-plate spline interpolation at 1 km resolution (Nguyen et al., 2026). Given this context, using IDW as the primary benchmark is not appropriate. Comparing the BRAIN dataset against such a basic method does not provide a rigorous assessment of performance, and the conclusion that the blended product outperforms gauge-based interpolation is therefore not sufficiently justified. his framing may mislead readers, particularly for sensitive applications such as flood modelling, where reliable rainfall estimates are critical.

Response: Thank you for this comment. We agree that our original framing could be interpreted as suggesting that IDW represents the current state of the art for operational rainfall analysis, which was not our intention.

Our motivation for including IDW was not to establish it as the primary benchmark, but rather to provide a simple gauge-only reference for demonstrating the potential benefits of incorporating multiple observation sources. According to the literature (Malone, 1999; Malone and Carroll, 2022), IDW has been used in operational flood forecasting applications in Australian Bureau of Meteorology because of its computational efficiency and ability to provide timely rainfall estimates for hydrological modelling.

We also agree that widely used Australian gridded rainfall products, such as AGCD, SILO, and ANUClimate, employ more sophisticated interpolation methods. These products benefit from relatively dense daily gauge networks and are designed for daily or monthly rainfall analyses rather than near-real-time hourly applications. Similarly, we agree that more advanced gauge-only interpolation methods, such as thin-plate spline interpolation (e.g., Nguyen et al., 2026), provide a more robust reference than IDW, although they also involve substantially higher computational costs.

More importantly, our objective is not to develop an improved gauge-only interpolation product. Rather, the purpose of BRAIN is to address the limitations of gauge-only approaches under sparse observation conditions by integrating complementary information from satellite, radar, and gauge observations within a statistical interpolation framework.

To address this comment, we plan to make the following revisions:

- Revise the manuscript to clarify the purpose of including IDW and explicitly state that it is intended only as a simple gauge-only reference rather than the primary benchmark.
- Moderate our wording throughout the manuscript to avoid implying that BRAIN has been comprehensively demonstrated to outperform operational gauge-only interpolation methods.
- Reduce the emphasis on the IDW comparison in the case study and present it only as an illustrative example rather than a central conclusion.
- Expand the discussion of existing gauge-only interpolation approaches, including more advanced methods such as thin-plate spline interpolation, and clarify their advantages, limitations, and applicability for hourly rainfall analysis.
- Clarify that a more comprehensive evaluation of the BRAIN framework, including comparisons with additional rainfall products and interpolation methods, is part of our ongoing development of the long-term BRAIN dataset and will be presented in future work.

References:

- Malone, T. and Carroll, D.: Guidelines for URBS routing parameters, in: Hydrology & Water Resources Symposium 2022, The Past, the Present, the Future, Brisbane, 145–155, https://search.informit.org/doi/10.3316/informit.905985293458115, 2022.
- Malone, T.: Using URBS for Real Time Flood Modelling, in: Joint Congress: 25th Hydrology & Water Resources Symposium, 2nd International Conference on Water Resources & Environment Research, Handbook and Proceedings, 603–608, https://search.informit.org/doi/10.3316/informit.705872257387556, 1999.

3. What interpolation distance was used in the IDW interpolation (apologies if this was specified and I overlooked it)? Figures 7c and 7f suggest that the IDW-based interpolation of hourly gauge data performs poorly, even when compared with AGCD. In particular, the method appears to over-smooth the observations and can produce unrealistically broad rainfall patterns. For example, while AGCD (Fig. 7e) indicates no rainfall within the highlighted red-box area below, the IDW result shows a substantial area with 6-20 mm/day. This discrepancy further highlights the limitations of IDW and reinforces that it is not a reliable benchmark for validation.

Response: Thank you for this helpful comment and for carefully examining the comparison figures. We apologise that the implementation details of the IDW interpolation were not described clearly in the original manuscript.

In our implementation, no search radius or distance threshold was imposed for the IDW interpolation. This was done to ensure that rainfall estimates could be generated for every grid cell, as applying a distance limit would leave some regions without interpolated values because of the sparse distribution of hourly gauges.

We would also like to clarify that the rainfall fields shown in Figure 7 represent 24-hour accumulated rainfall, obtained by summing the hourly analysis results. These accumulated fields were included to facilitate comparison with the operational daily AGCD product.

Regarding the discrepancy, we agree that it reflects the limitations of IDW as a gauge-only interpolation method. It is also important to note that the three products are generated from different observation sources:

- AGCD is produced using daily rain gauge observations and an advanced interpolation framework.
- IDW is generated solely from the available hourly rain gauge observations.
- BRAIN integrates complementary information from satellite, radar, and hourly gauge observations through a statistical interpolation framework.

Consequently, differences between AGCD and IDW do not necessarily arise solely from the interpolation algorithm, but also from differences in the observation datasets used to generate the rainfall analyses.

To address this comment, we plan to make the following revisions:

- Expand the methodology section to provide a complete description of the IDW implementation, including the weighting function and interpolation settings.
- Clarify in the figure caption and main text that Figure 7 shows 24-hour accumulated rainfall derived from hourly analyses and explain the purpose of comparing these fields with the operational AGCD product.
- Revise the discussion to better explain the differences between AGCD, IDW, and BRAIN, including their respective input observations and methodological characteristics.
- As noted in our response to the previous comment, reduce the emphasis on the IDW comparison and present Figure 7 as an illustrative case study rather than as the primary evidence supporting the performance of the proposed framework.

4. Please justify the influence radius of 200 km in the OI. Is there any sensitivity analysis on that distance? It is unlikely that the rainfall value at one point in the grid is correlated with a point 200 km away. In other studies, the distance used is around 100 km. Figure 4 also shows that the error correlation is less than 0.1 if the distance is more than 100 km. Similarly, why is the calibration radius in R1 expanded to 300 km compared to 150 km in R0? Increasing the interpolation distance can generate the issue of over-smoothing the blended dataset.

Response: Thank you for this insightful comment. We also appreciate your careful observation that the estimated error correlation in Figure 4 decreases to below 0.1 at distances beyond approximately 100 km.

We agree that the choice of the influence radius should be better justified in the manuscript. It is important to note that the influence radius in the statistical interpolation framework defines the maximum search distance for identifying candidate observations rather than implying that distant observations have a strong influence on the analysis. In the statistical interpolation, observations located farther away receive progressively smaller weights according to the estimated error correlation function. Therefore, although observations up to 200 km may be included in the analysis, those beyond approximately 100 km contribute very little because of their weak error correlation.

We have also performed sensitivity analyses using different influence radius. The results indicate only negligible differences in the blended precipitation fields, suggesting that the analysis is relatively insensitive to the exact choice of the search radius. This is consistent with the rapidly decaying error correlation shown in Figure 4, which naturally limits the contribution of distant observations.

Regarding the larger spatial calibration window used in R1 (300 km) compared with R0 (150 km), the purpose is different from that of the statistical interpolation influence radius. The expanded spatial window in R1 is used during the bias-correction stage to increase the number of gauge observations available for estimating the bias, particularly in radar covered regions between 150 km and 300 km. It is not directly related to the spatial weighting applied during the statistical interpolation. We agree that this distinction was not sufficiently clear in the manuscript and should be explained more explicitly.

To address this comment, we plan to make the following revisions:

- Add a more comprehensive theoretical explanation of the influence radius used in the OI framework, clarifying that it defines the maximum search distance for candidate observations rather than the effective radius of influence.
- Expand the discussion on the potential effects of search radius and explain why this effect is limited in our framework because distant observations receive very small statistical weights.
- Add some explanations of our sensitivity analysis to demonstrate that varying the influence radius has only a negligible impact on the blended precipitation estimates.
- Clarify the different purposes of the OI influence radius and the larger spatial window used in R1, and explain why the latter is beneficial for bias correction.

5. Because the distribution of rain gauges and radar coverage in Australia is denser in the coastal areas, I suggest separating the interpolation and validation for regions with high and low rain-gauge density, and targeting different interpolation weights for gauged data in each region (i.e., put higher weights on gauge measurements in the areas with dense rain gauges). For example, Figures 7a and 7b show an issue of mismatch between the BR-SRG and AGCD datasets, in the area with dense daily rain gauges. This may reflect artifacts in the radar and/or satellite inputs, which are not sufficiently discussed in the manuscript. In this case, I believe the AGCD dataset is more reliable because of the dense distribution of rain gauges in the area to detect rainfall on the ground.

Response: Thank you for this valuable suggestion. We think that evaluating the performance of the blending framework separately under different observation densities would be helpful, as it provides a more informative assessment of its strengths and limitations.

We also understand your suggestion that gauge observations should receive higher weights in regions with dense rain-gauge networks. This is, in fact, an inherent characteristic of the statistical interpolation framework used in BRAIN. The observation weights are not prescribed according to predefined regions; instead, they are determined automatically from the estimated error covariance structure and the spatial distribution of the available observations. Consequently, where rain gauges are dense and reliable, the analysis naturally places greater weight on the gauge observations, while in data-sparse regions, the analysis relies more heavily on the background field and other available observation sources.

Regarding the differences between BR-SRG and AGCD shown in Figures 7a and 7b, we agree that these discrepancies deserve further discussion. It is also important to note that the two products are generated using different observation networks. AGCD incorporates a much larger network of daily rain gauges, whereas BR-SRG uses only the hourly rain gauges available within our framework. Therefore, some of the differences may arise from differences in the input observations rather than deficiencies in the blending methodology itself. We also acknowledge that artefacts in the satellite and radar observations may propagate into the blended product and should be discussed more thoroughly.

To address this comment, we plan to make the following revisions:

- Separate the validation results for regions with relatively high and low rain-gauge density to better demonstrate the performance of the proposed framework under different observation conditions.
- Expand the discussion of the differences between BR-SRG and AGCD, including the influence of their different input observation networks and methodologies.
- Add a more comprehensive discussion of the potential artefacts and uncertainties associated with the satellite, radar, and blended precipitation products, particularly in regions where discrepancies with AGCD are observed.
- Clarify in the methodology that the statistical interpolation framework automatically adjusts the relative contribution of each observation according to the estimated error covariance and observation configuration, rather than assigning different interpolation weights to predefined regions.

6. The hourly rainfall measurements contain many zeros, so even though the RMSE values seems to be low, the dataset can have an issue of significantly underestimating high rainfall values. Figure7 shows that the BR-SRG can underestimate the maximum rainfall substantially compared to the hourly and daily observed data (as noted in the figure caption). This is a critical limitation, as accurate representation of extreme rainfall is essential for applications such as storm and flood analysis. I recommend extending the validation to explicitly assess performance across different rainfall intensities (e.g., low, moderate, and

high thresholds). In addition, presenting time series comparisons (or RMSE time series) at selected key locations would provide a more robust and informative evaluation of the dataset's performance, particularly during high-impact events.

Response: Thank you for this valuable suggestion. We agree that hourly rainfall datasets contain a large proportion of zero-rainfall samples, which can mask deficiencies in representing moderate and heavy rainfall events when evaluated using aggregate metrics such as RMSE. We also agree that the accurate representation of heavy rainfall is important for applications such as storm analysis and flood forecasting.

The underestimation of peak rainfall shown in Figure 7 is an important aspect that warrants further investigation and discussion. We will expand the manuscript to examine the potential factors contributing to this behaviour.

To address this comment, we plan to make the following revisions:

- Clarify the limitations of relying solely on overall RMSE and consider incorporating additional evaluations across different rainfall intensity ranges to provide a more comprehensive assessment of model performance.
- Explore the inclusion of event-based evaluations, including time series comparisons at representative locations, to better demonstrate the temporal evolution of rainfall during selected high-impact events.
- Expand the discussion of the underestimation of heavy rainfall and extreme precipitation, including the potential contributing factors and the implications for hydrological applications such as flood forecasting.

Detailed comments

Lines 85-88: Please add references for stating that IDW was used in flood modelling and flood forecasting. From my experience in flood modelling, the rainfall dataset generated using the IDW interpolation is very poor and should not be used as an input in flood models.

Response: Thank you for this helpful comment. We agree that the statement should be better supported by appropriate references.

Our intention was not to imply that IDW represents the state-of-the-art approach for operational rainfall analysis or that it is universally adopted for flood modelling. Rather, we intended to indicate that simple interpolation methods have been used in operational flood forecasting applications because of their computational efficiency and ease of implementation.

To address this comment, we will revise the manuscript by adding appropriate references to support this statement, including:

- Malone, T. and Carroll, D.: Guidelines for URBS routing parameters, in: Hydrology & Water Resources Symposium 2022, The Past, the Present, the Future, Brisbane, 145–155, https://search.informit.org/doi/10.3316/informit.905985293458115, 2022.

- Malone, T.: Using URBS for Real Time Flood Modelling, in: Joint Congress: 25th Hydrology & Water Resources Symposium, 2nd International Conference on Water Resources & Environment Research, Handbook and Proceedings, 603–608, https://search.informit.org/doi/10.3316/informit.705872257387556, 1999.

We will also revise the wording to avoid implying that IDW is the preferred or state-of-the-art method for operational rainfall analysis, and instead clarify that it has been adopted in operational applications while acknowledging its well-recognised limitations compared with more advanced interpolation methods.

Lines 125-127: This validation method is not new, and it has been applied in other studies (Yu et al. 2020; Xia et al. 2025), so I don't think it is a contribution of the paper.

Response: Thank you for this helpful comment. We agree that this validation method is not new and has been applied in previous studies, including Yu et al. (2020) and Xia et al. (2025). It was not intended to be presented as a contribution of this paper. Rather, our intention was to summarise the findings of the validation as a conclusion of the study.

To address this comment, we plan to make the following revisions in the manuscript:

- Revise the relevant text to avoid implying that the validation methodology is a contribution or novelty of this study.
- Clarify that the statement is intended as a conclusion drawn from the validation results rather than a methodological contribution.
- Cite the relevant previous studies, including Yu et al. (2020) and Xia et al. (2025), where appropriate.

Section 2.1: Please add more details about the quality control of gauge data (what values are considered extremely high and why).

Response: Thank you for this helpful comment. We agree that the quality control procedures for the gauge observations should be described in greater detail.

The quality control of the hourly gauge data was a substantial component of this project. Our original intention was to describe the complete quality control framework in a separate methodological paper, which is currently in preparation. Consequently, only a brief description was included in the present manuscript.

To address this comment, we will expand Section 2.1 by providing additional details of the quality control procedures, including the criteria used to identify and remove unrealistic or extremely high rainfall values, together with the rationale for the selected thresholds. We will also add a supplementary section describing the quality control workflow in more detail, including additional information on the screening procedures and threshold selection, to improve the transparency and reproducibility of the dataset. We will clarify that a more comprehensive description of the quality control framework will be presented in the dedicated methodological paper.

Lines 170-172: What is the source and the reliability of the additional independent hourly data? Why are they not included in the interpolation and only used for extra validation?

Response: Thank you for this important question. We agree that the source of these additional hourly observations and the rationale for using them only for validation should be explained more clearly.

These additional hourly gauge observations were obtained from third-party providers, including local councils, water utilities, and other water management organisations. They were made available through ongoing data-sharing agreements, and the database has continued to expand throughout the project.

At the time when the bias correction and blending analyses were performed, these additional observations were not yet available. Therefore, they were not included in the bias correction or statistical interpolation, both of which are computationally intensive processes. After the analyses had been completed, these additional datasets became available and were subsequently collated. Rather than re-running the entire processing workflow, we used these newly acquired observations as an independent validation dataset to provide an additional assessment of the performance of the BRAIN framework. This also offers a more independent evaluation, as these observations were not used during product generation.

To address this comment, we will revise the manuscript by:

- Providing a clearer description of the sources and characteristics of the additional hourly gauge observations.
- Explaining why these datasets were not incorporated into the bias correction and blending processes, and why they were instead used as an independent validation dataset.
- Clarifying that these additional observations will be incorporated into future operational production of the BRAIN product as they become available.

Section 2.2: Since when has radar rainfall data been available in Australia? Please justify whether increasing the calibration distance from 150 km to 300 km to include more gauges makes sense, or if the improvement is mainly from increasing the time window (15-min to 1-hour).

Response: Thank you for these insightful questions.

Regarding the availability of radar rainfall data in Australia, weather radar observations have been available since the late 1990s. The network initially covered only major population centres, such as Sydney and Melbourne, and has since been progressively expanded to provide near-national coverage of approximately 72% of the Australian land area. We agree that this background information would help readers better understand the temporal availability and spatial coverage of radar data.

Regarding the calibration distance and temporal window, the operational radar calibration currently used by the Bureau typically considers gauges within a 150 km radius and a 15-minute temporal window. In our study, we increased both the spatial window (to 300 km) and the temporal window (to 1 hour) to increase the number of gauges available for calibration, particularly in regions between 150 km and 300 km.

The larger spatial window allows gauges located between 150 km and 300 km from the radar to contribute to the calibration, while the longer temporal window enables the inclusion of gauges with longer reporting latencies. Together, these two modifications increase the number of available gauge observations and improve the robustness of the calibration process. We acknowledge that the relative contributions of the larger spatial radius and the longer temporal window have not been separately quantified in this study, and we will clarify this in the manuscript.

To address this comment, we will:

- Add background information describing the development and spatial coverage of the Australian weather radar network.
- Expand the description of the radar calibration methodology, including the rationale for increasing both the spatial window and the temporal window.
- Clarify that both modifications were introduced to increase the number of gauges available for calibration, while acknowledging that their individual contributions to the improvement were not separately evaluated in this study.

Section 3.2: The method mentioned in this section is under cited. Equations 1-5 were developed in other studies. Please add references to the original paper that developed the OI method (Gardin L., 1965), and other studies had already used the method to interpolate blended rainfall datasets. In this section, the radar-to-radar and radar-gauge correlations are ignored. However, according to Seo et al. (2011), the covariance of radar-gauge can be considerable at large scales due to the significant variability. Please justify.

Response: Thank you for these comments and helpful references.

Regarding the references, we agree that this section is under-cited. We will cite the original reference describing the OI method, including Gandin (1965), to acknowledge the development of the methodology. As this classic reference is relatively old and is not readily accessible in a standard journal format, we will also cite more recent and widely available studies that describe the methodology and its application to precipitation analysis and multi-source rainfall blending.

Regarding the omission of radar-to-radar and radar-gauge error correlations, we agree that these correlations can be important under certain conditions, as discussed by Seo et al. (2011). In this study, however, we deliberately adopted a simplified error covariance structure to achieve a practical balance between physical realism and computational efficiency for national-scale, near-real-time precipitation analysis.

Specifically, our framework adopts a nearest-radar approximation, whereby only the nearest radar grid cell is selected as the radar observation for each target grid cell. Consequently, each analysis location contains at most one radar observation, eliminating the need to explicitly model radar-to-radar error correlations. Similarly, we ignore radar-gauge cross-correlations to maintain a computationally efficient covariance structure suitable for operational implementation over Australia's large spatial domain. We acknowledge that this is a simplifying assumption and that incorporating cross-source error correlations may further improve the analysis in future developments.

To address this comment, we will:

- Add the original references for the OI methodology, including Gandin (1965), together with additional references to previous OI-based precipitation blending studies.
- Clarify that the equations presented in this section are based on the established OI framework rather than being newly developed in this study.
- Expand the discussion of the assumptions underlying the error covariance model, including the rationale for neglecting radar-to-radar and radar-gauge error correlations.
- Explain that the nearest-radar approximation substantially simplifies the covariance structure by limiting each analysis location to a single radar observation, thereby improving computational efficiency and making the framework suitable for large-scale operational applications.

Figure 7: The “x” symbol is in red, and it is really hard to see in the figure. Please change the shape and colour of the symbol.

Response: Thank you for this helpful suggestion. We agree that the current red "×" symbol is difficult to distinguish against the rainfall colour scale and does not effectively highlight the intended locations.

We will revise Figure 7 to improve its readability by replacing the current marker with a more visible symbol (e.g., a rectangle or another clearly distinguishable marker) and selecting a colour that provides better contrast with the background. This will make the highlighted areas easier for readers to identify.

References

Nguyen, C., Vaze, J., Mateo, C. M. R., Hutchinson, M. F., & Teng, J. (2026). Elevation-dependent spatial interpolation of hourly rainfall for accurate flood inundation modelling. Hydrology and Earth System Sciences, 30(1), 45-66.

Ochoa-Rodriguez, S., Wang, L. P., Willems, P., & Onof, C. (2019). A review of radar-rain gauge data merging methods and their potential for urban hydrological applications. Water Resources Research, 55(8), 6356-6391.

Seo, B. C., & Krajewski, W. F. (2011). Investigation of the scale-dependent variability of radar-rainfall and rain gauge error covariance. Advances in Water Resources, 34(1), 152-163.

Xia, H., & Wang, K. (2025). Hourly, kilometer-scale precipitation merged from rain gauge, ground-based radar and satellites over East Asia: methods, evaluation and applications. Journal of Hydrology, 134148.

Yu, J., Li, X. F., Lewis, E., Blenkinsop, S., & Fowler, H. J. (2020). UKGrsHP: A UK high-resolution gauge–radar–satellite merged hourly precipitation analysis dataset. Climate Dynamics, 54(5), 2919-2940.