Reply to RC1 - Anonymous Referee #1

Overview

This manuscript describes the development of a new operational hourly, km scale precipitation analysis for the entire Australian continent, titled 'BRAIN'. The authors adapt an existing method - Optimal Interpolation - to blend gauged precipitation, a satellite precipitation estimate and precipitation estimates from radar. The use of a robust existing method in OI is in my view a sensible choice for an operational system, and the authors make well-judged simplifications to enable its deployment at continental scale (e.g. the use of single radar grid cells to estimate error characteristics). The authors use stringent verification measures, including strict cross-validation and comparisons against a range of existing products that are in wide use in Australia. These are high hurdles to clear, yet the authors are able to demonstrate that BRAIN clearly outperforms existing alternatives. In addition, the manuscript is concise and clear, and in general a pleasure to read. No continentally consistent precipitation product at this temporal or spatial resolution currently exists for Australia, and when BRAIN is operationalised it promises to be a game-changing innovation for a wide range of scientific, industrial and environmental applications. I essentially recommend that the manuscript be published as is, subject to technical corrections.

Response: Thank you for your thorough and highly positive evaluation of our manuscript. We greatly appreciate your recognition of the significance of this work and are encouraged by your assessment that BRAIN has the potential to become a game-changing precipitation product for Australia. We also appreciate your positive comments on our methodological choices, the evaluation strategy, and the potential value of the product for a wide range of scientific, operational, and environmental applications.

I leave it to the authors to determine which of my recommendations they wish to act on, of which I will note only one here: I thought too much emphasis was put on the performance of BRAIN shown in Figure 7, which shows only two isolated events. I prefer the authors to refer to their robust assessments of performance in Figures 5, 6 and 8 (and in the very useful Supplementary materials) than on Figure 7.

I congratulate the authors on what I believe will be a highly impactful product.

Response: Thank you for this very useful suggestion and for your encouraging comments on the BRAIN product. We fully agree that Figure 7 should be presented as an illustrative case study rather than the primary evidence of the framework's performance.

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

- Rewrite the discussion of Figure 7 to present it as an illustrative example of the product's behaviour during two contrasting events.

- Reduce the emphasis on Figure 7 throughout the manuscript and instead place greater weight on the results presented in Figures 5, 6, and 8, as well as the extensive Supplementary Material.

Specific comments

Introduction is lucid, economical and comprehensive. It clearly outlines 1) the promise of this product, 2) the challenges in establishing a sub-daily rainfall product in Australia 3) the range of methods available 4) a clear justification for the authors' use of optimal interpolation for their BRAIN product and 5) how the authors will adapt OI for their product.

Response: Thank you for your positive assessment of the Introduction section.

242 "the "true value"" suggest "the "true value", T,"

Response: Thank you for this suggestion. In the revised manuscript, we will introduce the notation at the first occurrence of the term.

L243 "the "true value"" - it's not clear from the equations below or the prose how the true values (T_k or T_i) are defined; accordingly the equations for errors are not explicit. Please describe how the 'true value' is estimated or defined (in prose is fine). I appreciate that OI methods are well known, but for those of us (like me) who are not very familiar with them, it would be nice to have this spelled out.

Response: Thank you for pointing this out. We agree that explicitly defining the true value is important for readers who are less familiar with optimal interpolation.

To clarify, the terms T_k and T_i represent the theoretical true values at the target grid cell k and at the observation location i, respectively. These true values are not directly known in practice. They are introduced only as conceptual variables for the purpose of deriving the optimal interpolation equations and formulating the error structures.

In the revised manuscript, we will add a clear prose definition at the first occurrence of the term.

L251 "in normalised form" should this be "in standardised form"? Normalisation can imply a normalising transformation.

Response: Thank you very much for pointing this out. We agree that the term “normalised form” could be misinterpreted as implying a normalisation transformation, which was not the intention. Therefore, we will revise the text accordingly, changing "normalised form" to "standardised form" throughout the relevant sections of the manuscript.

L255-267 "This system of linear equations is solved to determine the optimal weights w_i, which are used to obtain the best precipitation estimate at each target grid cell by optimally combining satellite, radar, and gauge data." The equations above use a single term (O) to describe observations from gauges or radar, and weights vary with i (location), not data source. What happens when both radar and gauges are available? I assume these have quite different error characeristics, so I couldn't follow how they can be all thrown in together. Oh - never mind, this is explained below (Lines 271-317). It might be good to either avoid the statement "are used to obtain the best precipitation estimate at each target grid cell by optimally combining satellite, radar, and gauge data" at this point, or foreshadow the coming explanation of how you actually combine different observational data sources below, or both. The authors can add a version of this statement at the end of Section 3.2 if they wish.

Response: Thank you for this helpful comment. We agree that the statement at this point may be premature, as the detailed treatment of different observation sources and their error characteristics is explained later in the manuscript.

To improve clarity and avoid confusion, we will revise the text to describe the estimation process more generally, and we will add a forward reference indicating that the handling of different observation sources (radar, gauge, satellite) and their associated error characteristics are introduced in the following sections (specifically Section 3.4). This will better guide the reader through the logical flow of the methodology.

L272 "while background-to-observation error correlations are assumed to be zero, as different sensor systems are considered independent" no change here, merely a comment: this is interesting, as it's possible to envisage cases where these are not independent (though I think the use of satellite as background vs gauges/radar means the authors are reasonably safe in this assumption).

Response: Thank you for this thoughtful comment. Actually the complete independence between different data sources is difficult to guarantee in practice. For example, some satellite retrieval algorithms incorporate gauge observations during calibration and/or bias correction, and radar products may also undergo gauge-based adjustments. Consequently, some degree of dependency may exist between the datasets.

L285 "only the nearest radar grid cell to each target grid cell is retained as a supplementary observation" Might be good to (very briefly) note how 'nearest' is defined here: is it possible to have more than one 'nearest' cell? Never mind, this is also covered below!

Response: Thank you for this comment and question. We agree that a brief clarification at this point improve readability.

To address this, we will revise the text to explicitly indicate that the nearest radar grid cell is identified based on the minimum Euclidean distance to the target grid cell. In addition, we will add a brief explanation of how cases involving multiple equidistant nearest grid cells are

handled—specifically, when multiple grid cells have the same minimum distance, the grid cell with the lower error variance is retained.

L321 "except for gauges located outside the radar range where no radar information is available" I think this is saying that gauges outside the radar range are used to correct satellite data. Is this correct? (And if not, why not?)

Response: Thank you for pointing this out. We apologise for the unclear wording.

Our intention was not to indicate that gauges outside radar coverage were excluded from satellite bias correction. Rather, satellite and radar bias correction used different gauge sets:

- Satellite bias correction used all available gauges.
- Radar bias correction only used gauges located within radar coverage, because radar values cannot be extracted at locations outside the radar range.

To avoid confusion, we will revise the manuscript to explicitly state this distinction.

L335 "For gauged locations, this task is relatively straightforward, as gauge observations can be used as ground truth to estimate the error characteristics of the background and observational inputs" Excepting that gauges are point data and satellite/radar are areal averages, which will be smoother in time and space. How did the authors account for areal averaging to compute errors against gauges?

Response: Thank you for this insightful comment. We agree that representativeness differences exist between point-based gauge observations and grid-based satellite or radar estimates, as gridded products represent areal averages and are generally smoother in space and time.

Given the relatively high spatial resolution of the analysis grid (2 km), we assumed that these representativeness differences are comparatively small and can be neglected for practical purposes. In addition, the same assumption was applied consistently to both satellite and radar products. Therefore, the estimated error characteristics should be interpreted primarily as relative measures of the uncertainties among different data sources within the framework, rather than absolute estimates of their true errors.

We will revise the manuscript to explicitly state this assumption and its implications.

L352 "Specifically, the Australian domain was divided into tropical, subtropical, and temperate subregions." It would be nice to see these regions on a map; could simply be added to Figure 3a or similar.

Response: We thank you for this constructive suggestion. We agree that visualising the three subregions on a map would enhance clarity and help readers better interpret the regional performance results presented later in the manuscript.

We will add the delineation of the tropical, subtropical, and temperate subregions to Figure 3a (or an appropriate figure) in the revised manuscript.

Evaluation strategies: outperforming both IDW and AGCD is a high bar to clear; I commend the authors for this kind of rigour.

Response: Thank you for your positive comment. We appreciate your recognition of the evaluation strategy adopted in this study.

L393 "The spatial match between the gridded data and gauge locations was established using nearest interpolation." I'm not sure what was being interpolated here - please clarify.

Response: Thank you for this question. We agree that the original wording was ambiguous. The intention was to indicate that values from the gridded precipitation products were extracted at gauge locations using nearest-neighbour interpolation for comparison with gauge observations. To avoid confusion, we will revise the manuscript to clarify this point.

L407 "direct comparison with global precipitation products would be neither fully fair nor particularly informative." Fair enough - I assume the authors are suggesting BRAIN would substantially outperform these global products. As the verification in this study is already very rigorous, and you can't do everything, this is fine. I think it would nonetheless be interesting to compare BRAIN to these global products in a future study if possible, as it would reinforce confidence in BRAIN for Australian users.

Response: Thank you for this valuable suggestion. We agree that a comparison with global precipitation products would provide useful additional context and could further strengthen confidence in BRAIN for Australian applications. We are currently working towards the development of a long-term, operational-quality dataset based on the proposed framework. Once this dataset becomes available, a more systematic and comprehensive comparison with global products will become both feasible and more meaningful.

L547 "BR-SRG clearly outperforms AGCD and IDW in representing fine-scale spaital structure and intensity". I understand the need to show example predictions to illustrate how BRAIN functions in comparison to other products. I do not think it's appropriate to use these figures to describe relative performance, in particular because e.g. BRAIN does not outperform AGCD in out-of-radar regions on average according to Figure 6. Suggest this

sentence be reworded to "BR-SRG shows fine-scale structure and intensities not present in AGCD and IDW."

Response: Thank you for this helpful comment. We agree that the example figures are intended to illustrate spatial characteristics rather than provide evidence of overall performance superiority.

We will therefore revise the text to avoid overinterpreting the visual comparison and instead focus on the differences in spatial structure represented by the products

L550 Figure 7. A few things to improve about this figure: 1) it would be nice to put the dates for the two rows in the y-axis label of the left-most panel. 2) I could not see the 'x' symbols described in the text; please choose a symbol that is legible. 3) I do not think it's good practice to present additional results on rainfall totals at gauges in the caption. This information could be presented in small text boxes within each panel; alternatively an additional panel could be added to each row with a bar chart showing the values for each product.

Response: Thank you very much for these constructive suggestions to improve Figure 7. We will address each point as below:

1. Dates in y-axis label: We agree this would improve readability. We will add the dates (e.g., "12 February 2020" and "12 March 2020") to the y-axis labels of the left-most panels for each row.
2. Legibility of 'x' symbols: We apologise for this oversight. We will replace the 'x' symbols with a more legible marker (e.g., a triangle).
3. Rainfall totals at gauges: We agree that presenting this information in the caption is not ideal. We will add small text boxes within each panel showing the gauge totals.

We will revise the figure accordingly in the revised manuscript.

L561 "underestimates peaks"; as per the previous comment, this does not reflect e.g. bias statistics shown in Figures 5 & 6. It's my view that readers are drawn to examples like these, and they may use them to infer that these results are generally true, rather than refer to the much more robust statistics presented in Figs 5 & 6. I think that all that can be said is that BRAIN shows much greater spatial detail than the other products. This detail is supported by the verification metrics shown in Figures 5 & 6.

Response: Thank you very much for this comment. We fully agree that case example plots should not be used to make general claims about performance, and that readers may place undue weight on such visual comparisons. We will revise the text to focus solely on the spatial detail captured by BRAIN, while explicitly referring to the quantitative verification results in Figures 5 and 6 for evidence of overall performance.

L565 "AGCD (Fig. 7e) fails to capture the peak at Cape Tribulation". 1) International readers may not know where Cape Tribuation is and 2) The totals by AGCD on and off the coast around Cape Tribulation are higher than either BR-SRG or IDW, so I do not think this comment accurate reflects what's shown in the figure. Again, I suggest avoiding making inferences about performance from figure 7. Figure 7 is useful in showing (1) the additional spatial detail shown by BRAIN over both AGCD and IDW and (2) that the spatial detail is consistent across the continent, while for IDW it is not. In my view conclusions on performance should be drawn only from the robust assessments shown in Figs 5 & 6.

Response: Thank you for this helpful suggestion. We agree with both points raised. First, the reference to Cape Tribulation may not be familiar to international readers. Second, we agree that Figure 7 should not be used to draw general conclusions regarding the relative performance of the different products. Instead, its purpose is to illustrate the spatial characteristics of the rainfall analyses, while the quantitative performance should be assessed using the comprehensive validation results presented in Figures 5, 6, and 8.

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

- Remove the location-specific reference to Cape Tribulation and revise the description so that it is more accessible to an international readership.
- Revise the discussion of Figure 7 to focus on the differences in spatial structure represented by the products, rather than making inferences about their relative performance.
- Refer readers more explicitly to the quantitative evaluation results in Figures 5, 6, and 8 as the primary evidence supporting the performance of BRAIN product.

L616 "Future development of BRAIN will focus on improving input data quality, refining error characterisation, and enhancing spatial and temporal resolution." (1) It would be nice to give an indication of how computationally demanding a re-estimation of the OI is when new predictor products are used or existing predictor products change. (2) Will the use of e.g. additional satellite products violate the assumption of error independence between products that is crucial to the simplifications pre

Response: Thank you for these forward-looking questions. We agree that both points are relevant to readers interested in the operational implementation and future development of BRAIN.

Regarding the computational demand, the error characterisation is performed offline using historical data and is not recomputed in real time. Once the error statistics have been estimated, the hourly OI analysis is computationally efficient, and the entire continental-scale analysis can be completed within a few minutes on a standard computing node. If new predictor products are introduced or existing products are updated, only the offline error characterisation needs to be repeated.

Regarding the assumption of error independence, we agree that introducing additional satellite products may increase dependencies between data sources because they often share common retrieval algorithms, input observations, or bias-correction procedures. As with many operational blending frameworks, this represents a simplifying assumption, and its practical validity should ultimately be assessed through comprehensive validation.

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

- Add a brief discussion of the computational cost of re-estimating the error statistics when new predictor products are introduced.
- Clarify that the error characterisation is performed offline, whereas the operational OI analysis is computationally efficient.
- Expand the discussion of the independence assumption, acknowledging that additional predictor products may introduce dependencies between data sources and discussing the implications for future developments of the BRAIN framework.

Typos etc.

L39 "Satellites-based" should be "Satellite-based"

Response: Thank you. Will be corrected.

L50 "collocation-based approach provides" should be "co-location-based approaches provide"

Response: Thank you. But “triple collocation” is a correct term in the literature. For example, see the title from Pan et al. (2015).

Pan, M., Fisher, C.K., Chaney, N.W., Zhan, W., Crow, W.T., Aires, F., Entekhabi, D., Wood, E.F., 2015. **Triple collocation**: Beyond three estimates and separation of structural/non-structural errors. Remote Sensing of Environment 171, 299–310. https://doi.org/10.1016/j.rse.2015.10.028

L51 Suggest a sentence break before 'However', and suggest "However, both performance-based and co-location-based approaches typically..."

Response: Thank you for your suggestion. Will be revised.

L58 Suggest a sentence break before 'While', and suggest 'While offering high flexibility, ML-based approaches have...'

Response: Thank you for your suggestion. Will be revised.

L59 "In addition, all above" should be "In addition, all the above"

Response: Thank you for your suggestion. Will be revised.

L75 "data source used for background field" should be "data source used for the background field"

Response: Thank you for your suggestion. Will be revised.

L251 "By expanding and expressing the above equation" suggest "By expanding and expressing Equation 4"

Response: Thank you for your suggestion. Will be revised.

L264 "is adjusted weight accounting" should be "is the adjusted weight accounting"

Response: Thank you for your suggestion. Will be revised.

L337 suggest deleting "due to the absence of gauge observations" - it's redundant.

Response: Thank you for your suggestion. Will be deleted.

L394 "For the trialled implementation" suggest "For the trial implementation"

Response: Thank you for your suggestion. Will be revised.

L414 "mean bias error" suggest deleting "error" as it's redundant

Response: Thank you for your suggestion. Will be deleted.

L604 "providing a consistent and flexible rainfall analysis capability" Suggest deleting "capability"

Response: Thank you for your suggestion. Will be deleted.

L607 "for gridded radar source" suggest deleting "source"

Response: Thank you for your suggestion. Will be deleted.