Responses to reviewers

An intercomparison of four gridded precipitation products over Europe using an extension of the three-cornered-hat method, by Llorenç Lledó, Thomas Haiden, and Matthieu Chevallier, submitted to HESS.

Reviewer 1

The revised paper is an improvement over an already excellent paper and is acceptable for publication. The authors have responded adequately to my first review and the paper is clearer as a result. I have only a few very minor suggestions for the final version. I do not need to see it again unless there are specific questions or concerns.

Minor comments:

1. The notation $\binom{N}{2}$ was unfamiliar to me. The factorial form would be more understandable to all readers.

The notation $\binom{N}{2}$ denotes the binomial coefficient N over 2 or N choose 2 and, as far as the authors know, is a standard mathematical notation in the field of combinatorics, and more compact than the expanded form with three factorial terms. Unless the journal editor has a clear requirement, authors would prefer to keep this compact form.

2. I am not sure Eq. 4 is necessary. Also, the { in Eqs. (3)-(5) are not necessary.

We have included those brackets to denote a system of equations. This is also standard mathematical notation as far as authors know. Unless the journal editor has a clear requirement, authors would prefer to keep these brackets.

3. In line 178, the two datasets for which the error covariances are not assumed zero and instead calculated (IMERG and Opera) should be define explicitly. You could write:

Therefore, we have decided to explicitly compute the error covariance between IMERG and Opera.

Thanks for this suggestion. We have added an additional sentence for each bullet point where we explicitly write the corresponding error covariance term, and whether we set it to zero or we have computed it.

4. Shouldn't the Author Contributions, Competing interests, and Acknowledgements come before the Appendix, right after Code and data availability?

We have employed the Latex journal template provided by the journal, and this is the order that appears on compilation. If this is an issue, I would expect the copyediting to fix it.

Responses to reviewers

An intercomparison of four gridded precipitation products over Europe using an extension of the three-cornered-hat method, by Llorenç Lledó, Thomas Haiden, and Matthieu Chevallier, submitted to HESS.

Reviewer 2

General comment

This study examines the estimation of random error variance in four gridded precipitation datasets derived from various sources using a triangulation method known as the Four Cornered Hat (4CH). The research focuses on Europe (within the overlapping coverage area of the precipitation products) and analyzes the error variance patterns across the different datasets.

The authors have improved the logic and structure of their article while positively addressing my main concern regarding the innovation of the study: their extension of the 3CH method to the 4CH method is built on the work of Sjoberg et al. (2021). But they used additional datasets to compute two of the error covariance terms that must have been neglected in 3CH framework, which is the novel aspect of their work. However, I do not understand why the authors chose to highlight discussion on number that might be relevant for guidance on precipitation forecast verification but are not evaluated or validated, instead of better emphasizing the novelty of their methodological approach. That said, the paper, in its new structure, is acceptable for publication.

Nevertheless, I offer here some responses and clarifications on the authors' answers that might be considered to further improve the paper.

Comment

1) Regarding the evaluation, I acknowledge that the authors present valuable points showing that neither a) in situ network data nor b) E-OBS gauge density information is available for validating their error variance estimates. The authors must be aware that these two kinds of independent validation datasets were given as examples. It is the authors' responsibility to find a way to validate their analysis results to some extent. If this is not possible, I suggest explicitly discussing in the conclusion that validating such numbers is challenging given the current state of the art.

Authors do not want to set aside an independent dataset for verification (which would already imply a judgment of the verification data quality) and prefer to treat all data available with the 4CH method directly. With this, we have been able to show that radar estimates are sometimes (e.g. in Germany) of better quality than station-derived estimates. We have added a sentence in the conclusions stating that the observational error variances obtained are only accurate if the assumptions of the 4CH analysis are fulfilled:

"All the conclusions above are based on estimates of error variance obtained with an extension of the three-cornered-hat method, which in turn relies on a set of

error covariance assumptions. Those assumptions were carefully selected based on physical considerations of the different observational systems but have not been independently validated. The error variance estimates can differ from the true error variances whenever the error orthogonality assumptions are not fulfilled."

2) Regarding my comment suggesting masking (in grey) pixels where one of the datasets has negative variance in the 4CH method, the authors state in their response: "The numbers, albeit not exact, are still indicative of the product quality." From my point of view, deriving negative variance for at least one dataset indicates that the 4CH hypotheses do not hold at these particular locations. Since the variance (and covariance) are co-estimated dependently, I would not trust any of these estimates where one of them is incorrect. How can negative variance be linked to product quality? I suggest adding to the manuscript a short message noting that if one of the error estimates is non-physical, the other five estimates at the same location must be treated with caution.
We have added a sentence in the conclusions stating that the observational error variances obtained are only accurate if the assumptions of the 4CH analysis are fulfilled (see above). Authors believe those assumptions are strong and valid for most of the grid points.

3) Regarding my comment on the disappearance of quality issues observed for the OPERA dataset in the winter/summer stratification, the authors claim that this is a great illustration of the power of the 3CH. I respectfully disagree with this point because, in the conclusion, the authors' final guidance for choosing the verification dataset relies on the non-stratified variance estimates. In the authors' logic OPERA is less useful all year long due to only "a few outliers only present on some days in April or October", which does not seem accurate to me. The fact that a few outliers strongly influence the overall variance estimate is one of the limit of the current 4CH approach. On the contrary, I suggest the authors emphasize the power of their stratification to move from "static" error variance estimates toward state-dependent error variance estimates. Note that such an approach has been recently introduced in the framework of the Triple Collocation (Pellet et al. 2022).
Authors do not think that OPERA is less useful all year long. Indeed, the authors final guidance in the conclusions (line 240) already reads: *"The OPERA data is a higher resolution product, and although it might be biased with respect to in-situ measurements, with a bit of additional quality control it might be the best way to go for verifying high-resolution simulations over Europe".*

References :

• Sjoberg, J. P., Anthes, R. A., and Rieckh, T.: The Three-Cornered Hat Method for Estimating Error Variances of Three or More Atmospheric Datasets. Part I: Overview and Evaluation, Journal of Atmospheric and Oceanic Technology, 38, 555–572, https://doi.org/10.1175/JTECH-380D-19-0217.1, 2021.
• V. Pellet, "A Binned Triple Collocation for Estimating Regime-Dependent Uncertainties of Precipitation," in IEEE Geoscience and Remote Sensing Letters, vol. 19, pp. 1-5, 2022, Art no. 1005405, doi: 10.1109/LGRS.2022.3162703