

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

This study explores the estimation of random error variance in four gridded precipitation datasets sourced from various natural sources using the triangulation method called the Four Cornered Hat (4CH). In my assessment, the paper falls short of the standard expected by the Journal of HESS and would benefit significantly from a major revision. My main concern is the ambiguity regarding whether the variant of the 3CH, the 4CH, is an innovative aspect of this study. My full review is included as a Supplement.

[We want to thank the reviewer for taking the time to read the manuscript and suggest aspects that require improvement. In the revised manuscript we have clarified the novelty of our method by putting it in context with existing multiple-collocation techniques.](#)

General comment

This study explores the estimation of random error variance in four gridded precipitation datasets sourced from various natural sources using the triangulation method called the Four Cornered Hat (4CH). The investigation focuses on Europe (over the intersection coverage of the precipitation products) and analyses the error variance patterns across different datasets.

However, the overall outcome of the study is somewhat modest, resembling more of an experimental report rather than a comprehensive scientific paper. One notable concern is the ambiguity regarding whether the variant of the 3CH, the 4CH, represents an innovative aspect of this study. Furthermore, the introduction of the framework and methodology is brief and suffer from a lack of clarity, and the evaluation of the estimated variances is missing. In my assessment, the paper falls short of the standard expected by the Journal of HESS and would benefit significantly from a major revision. Below, I have listed my comments, which I believe will contribute to the preparation of an improved version of the manuscript.

[We agree that the methods section of the original manuscript lacked clarity and context and we have worked to improve this aspect in the revised manuscript. Regarding the modesty of the results, not for being concise a study should be dismissed of its value. In this case, the results we present are very relevant for guidance on forecast verification of precipitation. The fact that conclusions can be extracted from a limited number of backed assumptions is a strength rather than a weakness. We have been very careful not to run too many analyses based on different assumptions and then try to reconcile the results. Instead, we believe the best we can do is select a careful set of assumptions and discuss the results as the best estimates we can obtain from the data at hand.](#)

Major comments

1) The model error hypothesis within the framework of the 3CH should be introduced more effectively. Specifically, the estimation of random error variance assumes additive noise ε_A (along with some unquantifiable bias b_A) in the estimate A concerning the truth t :

$$A = t + b_A + \varepsilon_A,$$

While this may seem self-evident, it's worth noting that this is not always the case, especially considering the further extension of triangulation techniques, such as Triple Collocation, where uncertainty estimation is developed for multiplicative error (Alemohammad et al., 2015). I believe this section would benefit from a clear introduction of the formalism utilized by the 3CH, including the notion and notation of the truth t , the estimates A , B , C , and D , as well as their respective errors, before presenting Equation 1.

- Alemohammad, S. H., McColl, K. A., Konings, A. G., Entekhabi, D., and Stoffelen, A.: Characterization of precipitation product errors across the United States using multiplicative triple collocation, *Hydrol. Earth Syst. Sci.*, 19, 3489–3503, <https://doi.org/10.5194/hess-19-3489-2015>, 2015.

We agree that nomenclature was not clear enough in the first manuscript, and the methodology lacked clarity. The only error hypothesis we use is that errors are defined additively, by setting $E_A = A - t$ as the error of product A . This is the usual hypothesis in forecast verification. We have included this information in the revised manuscript, and we have enhanced the notation of sections 3.3 and 3.4 and explained in more detail the equations. The relationship of the 3CH method with the triple collocation method is fully described by Sjoberg et al. 2021 and there is no need to repeat those considerations here. The 3CH methodology is well aligned with our hypothesis and hence suitable here.

2) From my review, it remains unclear whether the variant 4CH introduced in the submitted manuscript represents a novelty.

The main point of the manuscript is to compare the uncertainties of four different precipitation datasets, and this is done by using an extension of the 3CH method, the 4CH method. Sjoberg et al. (2021) already introduced the notion of the N-CH analysis, but in that case, they use the additional datasets to compute multiple 3CH estimates from triplets. As far as the authors know, using the additional dataset to explicitly compute some of the 3CH covariance terms is novel, but it is not the main point of the paper. We do not believe this is a substantial modification of the original method that deserves special attention. We simply adapted the existing tools to the specific characteristics of the available datasets.

a) If indeed it is a novel contribution, it is imperative to underscore this aspect within the manuscript. This could be achieved by delineating the limitations of the 3CH in the methods section, thereby emphasizing the innovation brought forth by the 4CH. Additionally, the title and abstract should be modified to highlight this advancement compared to existing state-of-the-art techniques. It's noteworthy to mention that extensions beyond three datasets have previously been explored within the framework of Triple Collocation (Pan et al., 2015; Gruber et al., 2016; Vogelzang, 2021). Therefore, the authors should reference these prior works while describing the 4CH approach.

We agree with these suggestions. We have divided section 3.3 into one part for the 3CH and one part presenting the 4CH extension in the revised manuscript. We have also modified the title, and included a paragraph where we discuss the novelty of our method in relation to the existing literature. We have been cautious of not mixing 3CH and TC in the methodology section, which have different assumptions, but we have otherwise cited these studies because the rationale is similar.

b) If the variant 4CH has already been documented in the literature, this information must be explicitly stated, and the relevant references should be cited.

- Gruber, A., Su, C.-H., Crow, W. T., Zwieback, S., Dorigo, W. A., & Wagner, W. (2016a). Estimating error cross-correlations in soil moisture data sets using extended collocation analysis. *Journal of Geophysical Research: Atmospheres*, 121, 1208–1219.
<https://doi.org/10.1002/2015JD024027>
- Pan, M., Fisher, C. K., Chaney, N. W., Zhan, W., Crow, W. T., Aires, F., et al. (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>
- Vogelzang, J., & Stoffelen, A. (2021). Quadruple collocation analysis of in-situ, scatterometer, and NWP winds. *Journal of Geophysical Research: Oceans*, 126, e2021JC017189.
<https://doi.org/10.1029/2021JC017189>

Thanks for providing these interesting references. They all refer to TC and not to 3CH (albeit Pan et al, indeed present the 3CH equations under the TC name). Indeed Vogelzang et al. use a very similar concept to what we present, but with TC. As far as the authors know, this is the first time the idea is used for the 3CH. We have included these references and also described the relationship with our extension at the end of the new 4CH sub-section.

3) The error covariance estimate (as illustrated in Figure A.2) should be incorporated into the main text of the article and deeper analyzed. For example, comparing it with the variance estimates (figure 2) could demonstrate to what extent neglecting these quantities in 3CH methods is accurate. Since these covariance estimates are among the primary outputs of the 4CH method, their inclusion and examination could significantly enhance the understanding and interpretation of the study's findings.

We have moved this figure to the main text and added a description of it. Also, in reaction to a suggestion from reviewer 1, we have included error correlations in the appendix, which might be easier to interpret in some cases. Indeed, the areas with negative variances are collocated with large error covariances, and we discuss this in the text.

4) 3CH and 4CH method could be compared at least in the appendix or in the main text.

We have included a brief discussion of the differences between the 3CH and 4CH results in the appendix. We do not see large differences, except a slightly lower error variance for IMERG with the assumption of uncorrelated errors. We believe the assumptions for the 4CH analysis are better than those for the 3CH and therefore those results should be superior.

5) The study lacks in terms of evaluation. While I understand that the triangulation technique assumes the unavailability of ground truth, there are strategies that could be employed. For instance, using some gauge stations (not interpolated) from dense networks (over France and Germany) :

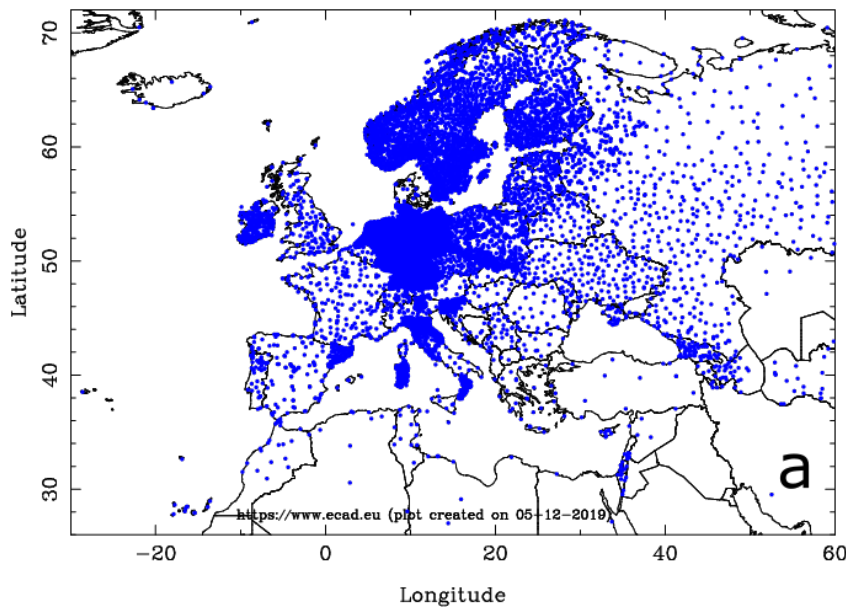
E-OBS is very efficient in making use of as many stations as possible from European high-density networks, and there is no such independent set available for verification, as far as the authors know. KNMI states in the E-OBS product documentation that: “For a considerable number of countries the number of stations used is the complete national network and therefore much more dense than the station network that is routinely shared among NMHSs (which is the basis of other gridded datasets). The density of stations gradually increases through collaborations with NMHSs within European research contracts.”

a) The author could analyze the impact of adding another dataset (e.g., EOBS) in the estimation of uncertainties (error variance), comparing results from Figure A3 and Figure 2. This comparative analysis would provide valuable insights into the effectiveness of incorporating additional datasets within the 4CH framework.

Please see our answer to point 4) above. The increase in the variance estimates not only comes from having more datasets but also from having better assumptions on the error correlations.

b) The author could evaluate the error variance pattern of EOBS with respect to the gauge density to investigate how its error variance is link to the interpolation itself. Such an evaluation, albeit limited, would enhance the robustness of the proposed analysis.

We do not have access to the underlying station density of E-OBS, unfortunately. The product documentation (<https://confluence.ecmwf.int/display/CKB/E-OBS+daily+gridded+observations+for+Europe+from+1950+to+present%3A+Product+user+guide#EOBSdailygriddedobservationsforEuropefrom1950topresent:Productuserguide-Theunderlyingstationdataset>) shows this plot for v20 of the product, but we use v26 so probably the situation changed in some countries.



6) In Figure 5, the quality issues observed for the OPERA dataset, as seen previously in Figure 2, almost disappeared in the winter/summer stratification. The author suggests that a plausible explanation for this phenomenon is the presence of outliers that are only present on certain days of April or October (therefore excluded in both composites), which exert a strong influence on the overall result. I recommend that the author thoroughly investigate this point and, if necessary, consider filtering out these outliers before estimating the error variance.

We believe it is a great illustration of the power of the 3CH method, and hence decided to leave this visible. It is not the intent of this work to curate specific datasets (i.e. apply additional quality controls, post-process, calibrate) but rather estimate their usefulness for forecast verification.

7) In Figures 4, 5, and A3, the gray area representing negative variance for one dataset should be masked for all the other three datasets in the figures. Furthermore, this discrepancy should be thoroughly discussed in the text as it underscores potential shortcomings in the 4CH approach.

Masking the grid-points in which one estimate is negative would hide which of the datasets received the negative value. The numbers, albeit not exact, are still indicative of the product quality. These negative values were already discussed in section 4 of the original manuscript. The negative variances can only be seen over a small portion of Europe and are due to the fact the error correlation assumptions are not correct in those grid points. In the revised manuscript we have backed this sentence by the inclusion of references discussing when those negative values occur, (in particular, those are described in section 4e of Sjöberg et al (2021) and section 2.5 of Pan et al. (2015)).

Other comments

1) Title: The title should introduce the notion of "variant 4CH" rather than the classical 3CH.

We have modified the title to highlight that we use an extension of the 3CH method.

2) In the article structure, Section 3.3 Method could be separated from Section 3.1 Data to balance the length of each section.

We have separated the 3CH and 4CH description into two subsections, resulting in better balanced sections.

3) Notation used in different equations should be harmonized.

Agreed and changed. This is in line with comments from reviewer 1 also.

4) Figures 2, 5, A2, and A3 show the boundaries of Europe and Africa outside the frame of the figures when the PDF is read locally.

We have updated the software stack and now used a more modern version of python, xarray and matplotlib to recreate the figures. However, we could not check this particular issue because the figures look good in Adobe Acrobat, and we could not reproduce the problem.

6) L112: “In particular, the OPERA radar data has been interpolated from a finer grid using a conservative interpolation”. Please add the interpolation scheme that has been used here?

The OPERA data has been upscaled with the gdalwarp utility, and in particular the “average” method, which according to the documentation (<https://gdal.org/programs/gdalwarp.html>) “computes the weighted average of all non-NODATA contributing pixels”. We have included this information in the revised manuscript in the Data collocation section.