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

This is an excellent paper that applies the three-cornered hat (3CH) and four-cornered hat (4CH) methods to estimate the uncertainties (random error variances) of four precipitation datasets. It is acceptable for publication after the authors consider some relatively minor changes that would improve the clarity of the paper. My full review with suggested edits is included as a Supplement.
We want to thank the reviewer for the positive appreciation of this manuscript, and for taking the time to read it and suggest specific improvements.

This is an excellent paper that applies the three-cornered hat (3CH) and four-cornered hat (4CH) methods to estimate the uncertainties (random error variances) of four precipitation datasets. It is acceptable for publication after the authors consider some relatively minor changes that would improve the clarity of the paper.
A novel aspect of the paper is to use the 4CH method to compute two of the error covariance terms that must be neglected in 3CH error variance estimates. Sjoberg et al. (2021) show that N datasets lead to (N - 1)(N - 2)/2 unique error variance estimates for each dataset when using the 3CH method. Furthermore, under the assumption of zero error covariances among all the datasets, a single error estimate for each dataset can be computed using all N datasets simultaneously (N-cornered hat method, a generalization of the 3CH method), which is identical to the mean of all the individual estimates. This paper goes further and shows that the N-cornered hat method is potentially more powerful than multiple applications of the 3CH method to various combinations of N datasets. It allows the estimation of some of the error covariance terms, which can be chosen arbitrarily, but based on insights of which pairs of datasets are likely to be most independent (have smallest error covariances). The authors illustrate this with four different precipitation datasets.
Thanks for providing this explanation of how N datasets were employed in Sjoeberg et al 2021. We have included this information in a new paragraph at the end of section 3.4, where we describe the novelty of our extension of the 3CH.

It is important when describing the 3CH and 4CH to be very clear in describing the methods, notation, and equations, because the equations can quickly become confusing, especially to those unfamiliar with these relatively new methods. My major suggestion is to improve the description of the 3CH and 4CH method in Section 3.3. I found the discussion a bit unclear, in part due to the notation in Equations (1)-(3):

1. There are currently two ways of writing the variance of the difference between two datasets. For example, in (1) var(A-B) is used, while in (2) and (3) vA-B is used. Please use one or the other for consistency.
2. Also, vA-t is not defined, although one can assume it is the error variance of A. I suggest writing it as something like var(Aerr) or var(Aϵ), where Aerr or ϵ denotes random errors.
3. More seriously, the covariance terms in (2) and (3) are not defined and could be misinterpreted. For example, cAtBt should be the covariance of the errors of A and B, but it could be misinterpreted as the covariance between the true values of A and B. Write something like cov(AϵBϵ) or cov(AerrBerr).
4. Finally, in Eqs. (2) and (3), each equation actually represents two equations. For example, the first equation represents the following two equations:
   VA-B = VAerr + VBerr – 2cov(AerrBerr)
   VA-B @ VAerr + VBerr
   The second equation is only an approximation, and equality holds only when the error covariance term is zero.

We agree that the notation was not carefully defined in the first manuscript. We have now defined E_A=A-t as the additive error of product A, and explicitly written all variance and covariance as Var(x) or Cov(x,y). We have also stated equation (1) as two separate equations and stated the system of equations (2) as two separate systems of equations, adding more clarity to the discussion. Equation (3) has been moved to a new subsection to present the four-cornered-hat methodology extension we employ.

In addition to the notation issues, I suggest a rewriting of lines 142-150 (I did not change the notation in this suggested revision):

To that end, if we have a third independent observation system C, we can repeat the above steps with differences between A and C, and B and C to obtain a system of three equations and six unknowns (the error variances of the three datasets and three error covariance terms):

vA−B = vA−t +vB−t −2cAtBt
vA−C = vA−t +vC−t −2cAtCt
vB−C = vB−t +vC−t −2cBtCt

If the three error covariance terms are small and can be neglected, the equations for the estimates of the true error variances can be solved from the remaining terms, which can be computed from the three collocated datasets.

In our case we have identified four observational systems, therefore we can constrain the computation a bit more. The four-cornered hat (4CH) method extends the system of equations with three more equations, (six in total, one for each pair of products), with 10 unknowns (the error variances of the four datasets and six error covariance terms (Eq. (3):

vA−B = vA−t +vB−t −2cAtBt

$vA-C = vA-t + vC-t - 2cAtCt$
$vA-D = vA-t + vD-t - 2cAtDt$
$vB-C = vB-t + vC-t - 2cBtCt$
$vB-D = vB-t + vD-t - 2cBtDt$
$vC-D = vC-t + vD-t - 2cCtDt$

If any four of the error covariance terms are assumed zero, we can solve for the error variance of the four datasets and the remaining two error covariance terms. Assessing which two covariance terms should be computed is a science-informed but subjective matter. The independence.....
We have followed this rationale in the reorganization of section 3.3. In response to reviewer 2 we have divided the section in two, with the 4CH method being presented in new section 3.4. Thanks for taking the time to suggest a clearer way to present the equations.

Minor comments:

1. The paper uses “data” as a singular noun. I realize many people use it this way, but technically “data” refers to more than one datum, and hence is plural. Consider changing all the “data is” to “data are”.
   Unless the editor has a strong opinion about this, we would prefer to use this generally accepted singular usage.
2. Fig. 2—What is the circular feature over central Spain in three of the panels? And one panel in Fig. 4. A comment about this in the caption would be useful.
   This pattern is related to quality issues in the OPERA dataset, specifically the weather radar in Madrid has some strong outliers in specific dates. Those are discussed in line 215 and also later in 236 of the annotated manuscript. 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 its usefulness for forecast verification. We have added a note on the figure caption, as suggested.
3. Line 175—It appears that the largest values of the differences between IMERG and ERA5 are along the coasts of western Europe and northern Italy (2-5 mm/day), not over the oceans where the largest magnitudes are 0.5-1.0 mm/day.
   We realize that this sentence was difficult to follow. We wanted to describe the behavior of the products over the ocean. We have changed the sentence to: “The largest differences can be seen over the ocean, where OPERA is drier than IMERG and ERA5 by 2 to 5 mm, and IMERG is wetter than ERA5 by 1 to 2 mm”.
4. Both E-OBS and EOBS are used for the rain gauge dataset; please use one of these throughout. EOBS is used in many of the figures and in Line 180. E-OBS is used in many other places.
   Thanks for pointing this out. We have corrected EOBS to E-OBS everywhere, except in the figures that show differences of products, where we wrote E_OBS to differentiate clearly the two products being subtracted.

5.
   A. Line 183-It might be better to change the wording from “worst” and “best” to the less pejorative “highest” and “lowest” e.g. ranking the products from highest to lowest.
   B. Line 185-Same as A-highest instead of worst.
   C. Line 191-Change “better” to “lower”
   D. Line 245-Change “best to worst” to “lowest to highest.”

   Agreed and changed. We also changed the legend of fig. A1 accordingly.
6. Line 195—Sjoberg et al. (2021) discuss in some detail how negative error variances can be obtained and could be included here as a reference. Line 249 as well.
   We have included a reference to section 4e of Sjoberg et al. (2021), and also a reference to section 2.5 of Pan et al. (2015) which also contains a good theoretical basis for understanding when negative error variances happen.
7. Are the units of Figs. 2, 4 and 5 mm^2 or $(mm/day)^2$?
   Total precipitation is measured in mm, and its variance in mm^2. Figure 2 showed mm/day which was erroneous and has been corrected. We also corrected the sentences of section 4 where this figure is discussed.
8. Line 230-You could compute the error covariance of ERA5 and Opera in this case using the 4CH method, but it is not necessary.
   We agree that it is not impossible to find plausible assumptions for running the 4CH method over the US, but we haven’t explored that avenue in depth.
9. Fig. A1 is outside Appendix A.
   This was handled by the journal latex template. I would expect the copy-edition to fix that. Anyways, after revising the manuscript the figure is placed in the right place.
10. The error covariances shown in Fig. A2 are interesting, but it would also be interesting, and easier to interpret, if the corresponding error correlation map were shown
    We have produced error correlation maps, and we present them in the appendix. The covariances have been moved to the main text as a response to a suggestion by reviewer 2. One particular problem with the error correlations, though, is that computing them from estimated covariance and variances does not guarantee that the result is in the [-1,1] range, and the occurrence of spurious zero or negative variances leads to undefined values. This has been mentioned in the Appendix.