# Interactive comment on "Characterization of precipitation product errors across the US using multiplicative Triple Collocation" *by* S. H. Alemohammad et al.

W.T. Crow (Referee)

wade.crow@ars.usda.gov

The paper describes the application of a modified triple collocation approach to the problem of evaluating large-scale precipitation data sets. The proposed modification allows for the more appropriate treatment of precipitation errors as multiplicative in nature. Issues surrounding the potential impact of error cross-correlation are examined via the decomposition of triple collocation error results over heavily-gauged reference sites.

Overall, this is a high-quality paper on a topic of significant interest. The application

C812

of a log-transform to deal with multiplicative precipitation errors is a very nice methodological extension and clearly superior to the existing approach (of naively assuming that all errors are additive in nature in order to shoehorn them into a TC framework).

I also appreciated the effort to explicitly examine the role of error cross-correlation on TC-derived error estimates (in Section 5). However, the one important thing I felt was missing was a re-examination of results in Section 4 based on the (non-trivial) impacts of cross-correlated errors (isolated in Section 5). For example, a key result in Section 4 is the relative lack of accuracy for the GPI precipitation product. However, in Section 5 (and the supplementary materials) we also see that GPI is relatively more independent (i.e. contains less error cross-correlation) than the other precipitation data sets. Given that TC will penalize GPI for this lack of dependence. . .does this mean that the analysis in Section 4 is truly even-handed? Is GPI being unfairly penalized due to being truly independent from the other products – as opposed to it being FAIRLY penalized for its weak relationship with "true" precipitation? So basically, I'd like a little bit of guidance about how the conclusions presented in Section 4 should be re-examined given the cross-correlation issues presented in Section 5. Should the reader really trust that relative rankings presented in Section 4?

I understand that this is a generic problem with any TC analysis; however, I think there are a couple of things that the authors could do to better address this point. First, they could examine whether or not the relative rankings that they derive using TC (at the 6 reference pixel sites examined in Section 5) accurately reflect the rankings they achieve when comparing all the products against the high-quality rain gauge observations acquired at each sites. If TC can successfully replicate the gauge-based rank correlation analysis at these 6 sites – that would be good evidence that the spatially-distributed TC results in Section 4 are robust in a relative sense (despite the known bias issues associated with the neglect of error cross-correlation).

Another step that could be taken would be replace the TMPA 3B42 dataset with its "real-time" (RT) equivalent (TMPA 3B42RT) which is not gauge-corrected. This tran-

sition would make the "TRMM" precipitation product relatively more independent from the NEXRAD and GPCP datasets (which also have a gauge-correction component). Therefore, this transition towards greater error independence should lead to an increase in TC-derived error for the NEXRAD and GPCP products and a decrease in error for GPI (when considered as part of triplet that includes GPI) . The size of this increase (or decrease) could be used of an indication of how serious the cross-correlation problem is across the entire study domain. Does it – for example – significantly close the gap between GPI and NEXRAD TC results over the eastern part of the study area?

Therefore, prior to publication, I would strongly recommend that the authors address this issue in some manner. At the very least, add 2-3 sentences describing the consequences of the analysis in Section 5 on earlier results in Section 4.

Some other minor suggestions:

1) Page 2536, Line 9-11: Clarify what exactly is meant by "homogeneous"? You mean homogeneous in a statistical climate sense...correct?

2) Page 2535, Line 20-22: I don't see how zeros would violate the assumption of error independence...however I can see how they would cause fatal problems in a log-transform analysis. The authors might want to re-write this sentence (or improve its clarity). Also, what about extremely low rainfall values (right at the edge of numerical precision)... can they skew results conducted in log-transform space?

3) Page 2541, Line 1-4: I had to read this sentence several times to follow it...I'd recommend re-writing to clarify its meaning (e.g. be a bit more specific...representativeness error in what?....and cross-correlation in errors between what and what?).

Wade Crow