Comment on hess-2021-262
Anonymous Referee #3

Referee comment on "Citizen rain gauges improve hourly radar rainfall bias correction using a two-step Kalman filter" by Punpim Puttaraks Mapiam et al., Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2021-262-RC3, 2021

General comments:

The sparsity of the conventional rain gauge network is a limiting factor for radar rainfall bias correction. Citizen rain gauges offer an opportunity to provide additional information at higher spatial resolution. This paper performs radar rainfall bias adjustment using two sources of rainfall information: observations measured by TMD (Thailand Meteorological Department) and daily citizen rainfall observations. The radar rainfall bias correction factor was sequentially updated based on the TMD data and downscaled hourly citizen data via a two-step Kalman filter. The results showed that citizen rain gauges improved the performance of radar rainfall bias adjustment especially for small ranges from the center of these gauges.

My reading through this manuscript suggests that the following three issues should be dealt with or stressed more clearly in the paper:

1) The citizen rain gauges are very important in the context of this paper. However, the relevant introduction is too simple. What kind of equipment are they? How uncertain are the measurements compared to the official TMD data? Please add more information on them.

2) The strategy used to downscale the daily measured citizen rainfall observations into hourly temporal resolution is essential in this context. However, there were no references on this topic. Four relatively simple downscaling strategies were tested, and the results showed that the one that utilized gauge-respective radar rainfall patterns performed the best. Yet I do think there is room for improvement.

3) I am concerned about the strategy used for applying the proposed approach. There is a
spatial separation of the gauges used for performing the 1st step of the Kalman filter (KF) and the gauges used for performing the 2nd step of the KF. The 1st step used 14 TMD gauges that are within 100 km radius of the center of Tubma basin, from which only 1 is inside the basin, whereas the 2nd step used 16 citizen rain gauges that are within the basin. The 1st step was applied under the assumption that the radar rainfall bias correction factor is relatively stable in space. However, there were signs of spatial instability, e.g., the downscaling strategy $G_{\{MP\}}$ (where the hourly rainfall patterns of the 14 TMD gauges were averaged and used for downscaling) had the worst performance, as shown in Fig. 6(c). More obvious signs were shown in Figs. 8 and 9. Hence, I strongly recommend the authors improve the strategy for applying the approach.

Specific comments / technical corrections:

1. Introduction

The description related to Citizen rain gauges/rainfall observations is too simple. Whereas, as indicated by the title "Citizen rain gauge improves hourly radar rainfall bias correction using a two-step Kalman filter", the readership might expect more information on the citizen rain gauges/rainfall observations from the introduction. Besides, it seems that the downscaling strategy is very important in this context. It might be necessary to give a brief review of the downscaling methods in scientific literature.

2. Study Area and Data

(Pg 5, L.125-126) I'm not clear with the sentence "A rain gauge with more than 80% of the dataset below the threshold was excluded from the analysis." Please explain.

(Pg 5, L.128-129) It is necessary to show the 14 rain gauges in Fig. 1 as well because the 14 rain gauges were intensively utilized in the following sections.

(Pg.5, L. 131-136) Concerning the low-cost citizen rain gauges, are they tipping buckets? What kind of equipment are they? Were the measurements compared with the TMD rain gauges (As shown in Fig. 1, there are 2 TMD gauges very close to them)? If so, how uncertain were the data? Or, in the other case, were the quality checks performed in an intra-group manner. It is hard to tell from the current description. Please make it clear.

3. Methods
Concerning the downscaling method "G_{Tubma}", instead of using only one TMD station within the basin (I noticed there is another one very close to the basin as shown in the small figure in Fig. 1), why not use this one as well. Besides, how variable are the hourly rainfall patterns across the basin. A comparison of the patterns from these two stations might provide useful information.

Please keep the term "KF-TMD" consistent with that shown in Table 2.

4. Results and discussion

There might be a mistake. The R_P-based result has a sharp mean cumulative fraction curve and large variance in the downscaled hourly data if one observes the box plots, whereas the other has a flat mean cumulative fraction curve and small variance in the downscaled hourly data. In other words, these two are different in both respects. The problem comes down to how the variability is defined. Please explain.

I am concerned about the validation scheme referred to as "KF -TMD-D", especially about the separation of the gauges for performing KF/MFB and gauges for validation purposes. The bias correction was made for a large area with a radius of 100 km, whereas the validation was performed on a much smaller basin. I am afraid the separation makes the validation results less representative. Perhaps LOOCV for the 14 TMD gauge (as used in the "KF -TMD-H" strategy) is more persuading. If the purpose of using "KF -TMD-D" is to validate the bias correction performance within the basin, the authors should specify it explicitly (perhaps, also in Sect. 3.4). Anyhow, the separation mentioned above could be a bit problematic.

"While there is a modest improvement in mean RMSE, the upper 75%-ile RMSE is reduced from about 6 mm/h to 3.5 mm/h. Mean MBE is changed from 0.1 to -0.15 mm/h." I found it hard to follow here. Please use the terms presented in Fig. 7(b) to refer to the results.

Please correct "MFB-TMB" and "KF-TMB" to "MFB-TMD" and "KF-TMD", respectively.

I think the description of the results/figures could be organized in a more readable way. For example, there is a direct shift from the description of Fig. 7 to that of Fig. 8 (Pg.19, L. 440-444) without telling the readers that the content is related to Fig. 8, and there is a separation in the description of Fig. 7 (Pg. 19, L. 433- and Pg. 20, L. 451-).
I am not clear with the sentence "Figure 9 illustrates that hourly rainfall distribution patterns of TMD rain gauges in the 40-90 km range, influenced mainly by the southwest monsoon, appear to be more similar to the mean citizen rain gauge data than the range beyond 40 km." Please explain.