

## Review

Title: Citizen rain gauge improves hourly radar rainfall bias correction using a two-step Kalman filter

Author(s): Punpim Puttaraksa Mapiam et al.

MS No.: hess-2021-262

MS Type: Research article

## Major comments

- In Chapter 2.3.1 you describe that all TMD rain gauges „with more than 80% of the dataset below the threshold was excluded“ (P5L125). Since the threshold of 0.5 is equal to the tipping-bucket resolution, you excluded all station with an hourly p0 (probability of value zero in the dataset) above 0.8. On P16L363 you mention that „the others having a period of heavy rainfall around 4-5 hours a day“ which leads to a p0 of around 0.8. This applied filter excluded approximately 55% of the stations which is a lot due to the fact that the data is provided by the Thai Meteorological Department. Can you please provide a figure showing all TMD gauges on the one side and the used gauges on the other side?
- Was there a specific reason why the validation was done with two different data sets? This means that a comparison between the daily and the hourly data set is not meaningful.
- Was there a specific reason why the boundary between "near" stations and the "far" stations was chosen at 40km in Case 4? I think that the boundary could have been anywhere outside the Tubma basin and there would have been a significant change in the results.

## Minor comments

- Title: „Citizen rain gauge improves ...“ → Citizen rain gauges improve ...“
- Title, P1L11, P1L15, P1L20, P2L55,...: Kalman filter, Kalman Filter, ... please be consistent in using capital letters or not
- P2L32: „the A and b parameters“ → „the parameters A and b“
- P2L43, P2L44: „... the reference A parameters (...) to the A parameters for sub-daily resolutions.“ → please reformulate
- P2L61 and P2L62: „They found ...“ → please reformulate
- P3L66: „...typically provided at daily scale“ → Please define citizen rain observation more precisely so that it is clear that this refers to soda bottles, for example.
- P3L75: 101°17'51" → : 101°17'51"E
- P3L75: „of approximately 197km<sup>2</sup>“ → „of 197km<sup>2</sup>“ or “of approximately 200km<sup>2</sup>”
- P3L83, P3L84: “240 km x 240 km (...) 0.6x0.6km” → Please stay with one style
- P3L84: “... spatial resolution and 6-min temporal resolution” → „... spatial and 6-min temporal resolution“
- P3L90: „3 datasets“ → „three datasets“
- P3L90-92: What happened between 2014 and 2019?
- P3L121: Is it “Thai Meteorological Department” or “Thailand Meteorological Department”?
- P4L102 and L103: “1-hour” → “hourly”
- P4L103: “A parameters” → “parameters A”
- P4L104: “b exponent” → “exponent b”

- P4L106: “b parameter” → “parameter b”
- P4L108: „mean absolut error“ → „MAE“ (defined on P4L104)
- P4L110: Since  $N_{G,t}$  seems not be equal to N over the entire period T, the Equation is incorrect. Which equation was used in your calculations?  

$$\rightarrow MAE = \frac{1}{T} \sum_{t=1}^T \frac{1}{N_{G,t}} \sum_{i=1}^{N_{G,t}} |G_{i,t} - R_{i,t}|$$
- P4L122: „have tipping-bucket sizes of 0.5mm” → something like “have a resolution of 0.5mm”
- P4L130: “in the 197km<sup>2</sup> Tubma basin” → „in the Tubma basin“
- P4L133: 1 gauge/15km<sup>2</sup>: How did you calculate that? One TMD station + 16 citizen rain gauges = 17 stations. 197km<sup>2</sup> / 17 stations → 1 gauge/~12km<sup>2</sup>
- P6Table1: „Code description“ → „Code Description“
- P7L158: Is MFB the abbreviation for „Mean field bias“ or „Mean fiel bias adjustment”? If the latter, I would suggest MFBA, since MFB is a common abbreviation for mean field bias. If the former, use this abbreviation (P7L162, P7L168).
- P7L163f: „...Smith and Krajewski (1991), Anagnostou et al. (1998), and Seo et al. (1999), Chumchean et al. (2006), Kim and Yoo, (2014), Shi et al. (2018).“ → „...Smith and Krajewski (1991), Anagnostou et al. (1998), Seo et al. (1999), Chumchean et al. (2006), Kim and Yoo, (2014), and Shi et al. (2018).“
- P7L177: “The radar bias at time t ...” → The radar bias at time t ...”)
  - P10L254: „... day i ...” → „... day i ...“
- P11Figure3: „Is hourly TMD data at hour t available?” → „Is hourly TMD data y at hour t available?”
- P11Figure3: „Is hourly citizen rain gauge data at hour t available?” → „Is hourly citizen rain gauge data z at hour t available?”
- P12L287: „... hour t ...” → „... hour t ...“
- P12L290: „... time t ...” → „... time t ...“
- P12L295: „Kalman Filter“ → „KF“ From here on, no more explicit mention for missing abbreviations. Please check independently in the following.
- P13L303: „1 TMD“ → „one TMD“
- P13L306: „randomly”: Was the LOOCV done for all 16 citizen rain gauges or did you randomly sample 16 times? If the former, it is not really randomly.
- P13L306: „3 different techniques, and 1 rain gauge“ → „three different techniques, and one rain gauge“
- P13L310: „... gauge i ...” → „... gauge i ...“
- P14L324: see P13L306
- P14L325: „1 TMD“ → „one TMD“
- P14L327f: „fourteen TMD“ → „14 TMD“
- P14L338: „(leave 1 TMD out)” → „(leave one TMD out)”
- P14L345 and L347: “Kalman Filter” → “KF”
- P15L347: “r<sub>1</sub> parameter” → “parameter r<sub>1</sub>”
- P15L350: “over the same time-series period” → “over the same period”
- P16L372: “figure 5” → “Fig.5”
- P17L391: “observations” → “observations”
- P17L393: “are based on 4” → “are based on four”
- Figure 6 and 7: Since RMSE and MBE have different limits (0 to infinity vs. -infinity to infinity), it does not make sense to put both assessment measures on one graph.

- P18L414: “respectively)” → “respectively).”
- P18L432: “Figure 7 (b) and Fig. 7 (c)” → “Figure 7 (b) and (c)”
- P18L435: see P18L432
- Figure 9: Please increase the fontsize and add a grid.
- P22L490: “(August-October, 2019)” → “(August-October 2019)”
- P22L495 and L497: “Kalman filter” → “KF” or „Kalman Filter“