Author response to referee #3 (Søren Thorndahl)

HESS-2021-105

Ruben Imhoff

Ruben.Imhoff@deltares.nl

May 21, 2021

Dear reviewer,

We would like to thank you for your interest in our work and the enthusiastic reaction to our manuscript. With four constructive and elaborate reviews, we think we are very well served by our reviewers. Your comments have been valuable and have helped us to improve the manuscript.

Below, we give a response to the given suggestions. We have placed the reviewer's comments in black font and below that, our response in blue font for clarity.

Sincerely,

Ruben Imhoff, Claudia Brauer, Klaas-Jan van Heeringen, Hidde Leijnse, Aart Overeem, Albrecht Weerts and Remko Uijlenhoet

General comments

1) In another study that the authors also refer (Schleiss 2020), we have seen significant differences between radar products based on single radars versus products based on composites of multiple radars. The latter being more reliable especially for the estimation of high-intensity rainfall. Can you maybe comment on how radar data from the two radars that you apply in the study are merged and how this relates to the larger biases as a function of distance from the radar. If there are significant range effects, I guess that this will be present both in Ra and Ru and therefore not necessarily lead to bias differences. Is there differences in the bias estimations depending on whether the location in question is covered by one or two radars.

   That finding in Schleiss et al. (2020) is indeed a very interesting one. As mentioned in Sec. 2.1, the data of both radars in this study were merged using range-weighted compositing. This is further described in Overeem et al. (2009). In short, the weight a radar gets for grid cell (*i*,*j*) depends on the range from that radar (in km):

$$w(r) = \begin{cases} (\frac{r}{70})^2, & if\ r \leq 70 \\ 1 - \left(\frac{r-70}{200-70}\right)^2, & if\ r > 70\ and\ r < 200 \\ 0, & if\ r \geq 200 \end{cases}$$

   Here, 200 km is the maximum used range of the weather radars (note that 320 km is used operationally nowadays) and 70 km is seen as the range where the weight is maximum (note that the distance between both radars is approximately 100 km, approximately 120 km between the new radars in Herwijnen and Den Helder).

   The subsequent bias in this merged product is indeed influenced by the presence of multiple (two) radars in the composite. See for example also Fig. 9 in Holleman (2007), which shows that the bias is lowest in a close region around a line between both operational radars (three notes: the QPE product was slightly different then and based on a pseudo-CAPPI at 800 m instead of 1500 m now; the weighing function was slightly different in Holleman, 2007, and the color scheme for the biases is not ideal to distinguish between low ranges of biases, e.g. 0.0 to 0.6). The bias increases further away from this line, with highest errors in the East, South East and South West of the country. This corresponds also to our findings with the remaining bias in these regions after the MFB-adjustment has been applied.

   R_A is also spatially corrected with the daily manual rain gauge observations, which should correct for a large part of the range effect. So, we expect that the remaining bias is low and shows up when the unadjusted radar QPE is compared to R_A.

2) The idea of a climatological bias factor works if the drop size distribution and the Z-R relationship do not change with season. It is especially important here to distinguish between convective and stratiform precipitation. In NL the winter is probably dominated by statiform precip., but summer precipitation is probably a combination of both…which would lead to more uncertainty in estimation of summer rainfall. Can you maybe provide an insight into the variability of Fclim depending on season and year. Like in figure 4 (a) where you show MFB for 2018, but Fclim for all years with confidence bands or monthly boxplots describing the variability for each month and year.

This is a good point, the error in the summer precipitation estimation depends on the type of rainfall for a given event. CARROTS should partly compensate for differences in the Z-R relationship over the year, but for individual events this can indeed still go wrong. However, Reviewer #1 also suggested (to comment on) a dBZ-dependent correction factor. This is outside the scope of our current work, but we think it is worth looking into as a follow up. We actually performed a small analysis of the relationship between the drop size distribution and the correction factor (prior to submitting this manuscript), but this gave no clear relationship so far. However, we would like to explore this further in future work.

As F_{clim} is derived as an average over 10 years of radar data, we cannot show the annual variability of the factor. However, the sensitivity to leaving individual years out of the derivation is visualized as the vertical bars in Fig. 4c. An indication of the variability from season to season is present in Fig. 4a where a clear seasonality is present. To get an idea of the variability of the unadjusted radar QPE (R_U) quality over time, it is more insightful to plot the difference between the reference and R_U when both are accumulated over a 31-day window (the used moving window in this study).

Specific comments and technical corrections

3) In figure 7 (a) is the bias correction factor derived for all years or only 2015? Since the idea of Carrot is to use the climatological average, I would prefer to see data for the whole period 2009-2018.

We have derived the factor for all years, just like the other results. As the factor does not change over the years (it is based on the mean over the ten years), we only show one year. In line with Fig. 6, we also only show the hydrological model simulations for the year 2015. We have chosen to show one year in order to be able to see individual discharge peaks, which would become more difficult when showing ten full years. We propose to state this more explicitly in the caption to avoid any confusion.

4) To be consistent please indicate on figures 3, 5 and 7 that the “bias correction factor” corresponds to Fclim.

Thanks for suggesting this, that would make the results indeed more consistent. We will apply it to the indicated figures.

References

Holleman, I.: Bias adjustment and long-term verification of radar-based precipitation estimates, Meteorological Applications, 14, 195–203, https://doi.org/10.1002/met.22, 2007.

Overeem, A., Holleman, I., and Buishand, A.: Derivation of a 10-year radar-based climatology of rainfall, Journal of Applied Meteorology and Climatology, 48, 1448–1463, https://doi.org/10.1175/2009JAMC1954.1, 2009.

Schleiss, M., Olsson, J., Berg, P., Niemi, T., Kokkonen, T., Thorndahl, S., Nielsen, R., Ellerbæk Nielsen, J., Bozhinova, D., and Pulkkinen, S.: The accuracy of weather radar in heavy rain: a comparative study for Denmark, the Netherlands, Finland and Sweden, Hydrology and Earth System Sciences, 24, 3157–3188, https://doi.org/10.5194/hess-24-3157-2020, 2020.