Author's replies to the referee's comments to manuscript AMT-2021-34

Please see the file with track changes to check the edited lines.

The original referee's comments are written in bold and the author's replies are written in regular font.

Anonymous Referee #2, revision 3

The authors have responded to all my comments and adjusted the manuscript accordingly. There are some minor details left that should be corrected, mainly regarding how the wet-dry classification results are discussed. See my comments below.

We gratefully thank the Referee for the constructive comments and recommendations. We carefully checked the quality of the added text.

L22: write "…by a delay…"

Reply:

L22: "…This decrease was caused by a delay in data delivery…"

L244: Since the authors acknowledge that they have no good explanation for the importance of WD_p2, they should add this information to the text. I suggest to write that there are some uncertainties (which I do not know of, but the authors should) regarding how the importance of the parameters are derived and that there is no clear explanation for the importance of WD_p2.

Reply:

We added a sentence in 2.3 section in order to clarify how the importance of parameters are derived. Likely, using the right terminology let the things clearer.

L169-170 "…Thus, local sensitivities (i.e. a partial effects) get integrated o a global sensitivity measure…"

Moreover,

L241-244 "…It is important to highlight that the max(P_{min}) is only computed if at least a minimum number of hours (defined by WD_{p1}) of data are available; otherwise it is nor computed and no rainfall intensities are retrieved. Certainly, this results in uncertainties about the computation of sensitivity when rainfall is not retrieved. Thus, a clearer explanation for the highest importance of WD_{p2} is not pointed.

Fig. 2: The figure caption should explain what the number in the upper-right diagonal are. From the plot in the last revision it is clear to me that the first pair defines the range of the colours and I guess that the number behind is the correlation of the points within that range. But this is not clear to me, and might be even less clear to a reader that has not seen the previous version of this plot.

Reply:

Thank you so much for this observation, we corrected the figure caption accordingly.

L335: Better write "However, taking these parameters into account in the optimization..."

Reply:
L325: "...However, taking these parameters into account in the optimization..."

L346: "...for classified rainy events rightly..." sounds strange. If I understand correctly you mean the true positive rate. Please reformulate.

Reply:

L340:"...However, the Sensitivity metric shows that the calibrated parameters are worst for classified the true positive rate..."

L348 and following: In my opinion you do not have to use such negative formulations ("did not generalise at all"). It would suffice to point out that the optimization might not have generalized well enough. As you write in the text below the different fractions of wet events in the calibration and validation data set might also have contributed to this decrease of the MCC.

Reply:

L340-342:"... This occurred because the optimization might not have generalized well enough the wet-dry classification process..."

L357-360: Polz et al 2020 have not "reached" a result which is similar to, as you write, "approximately 50% of the rainy events are classified as dry". Their Fig 9 shows a much lower number of false-negatives compared to the true positives. Also Fig A1 clearly shows that there is approx. a 0.25 to 0.75 spilt between true-positives and false-negatives. Furthermore, I find these two new sentence a bit confusing because the first one seems to talk about false-negatives and the sec-

ond one about false-positives. Maybe this is only because the wording is not precise. I suggest to reformulate this new part slightly.

Reply:

L354-355: "…Using a convolutional neural networks for classifying wet-dry periods, Polz et al. (2020) found a proportion of approximately 25% for false positives …"

L374-377: The logic of this new sentence is not correct. One cannot directly conclude that an overestimation (i.e. a high PBIAS) is caused by false-positives. Even if you have no false-positives and no false-negatives there could be bias, e.g. due to wrong WAA compensation or other effect. In my first two reviews I concluded that the false-positives have a strong effect on PBIAS because PBIAS changes dramatically if you do not include the false-positives in the analysis, as can be seen in Table 5. for the case "Reference > 0" which removes all cases where CMLs estimate rainfall but the reference is dry. Hence, I suggest to update this sentence and state that the results from Table 5 (as described above) indicate a strong effect of the false-positives on PBIAS. The fact that Polz et al 2020 do not see such a strong effect might be due to a lower number of false-positives in their analysis or due to a different distribution of false-positives. In their Fig 9 it is clear that the occurrence of false-positives has a peak at fairly small rain rates (0.1 mm/h) whereas the true-positives have the peak at 1 mm/h.

Reply:

L385-387: "…Indeed, the false positives presented a strong effect on PBIAS, when they were removed (i.e., "Reference > 0") PBIAS changed significantly. As for Polz et al. (2020) a different behavior was observed, maybe due to a lower number of false positives or due to a different distribution of false positives…"