Response to Referee Comment 1 on 'Impact of bias nonstationarity on the performance of uni- and multivariate bias-adjusting methods' by J. Van de Velde et al.

Bastien François (Referee)

We would like to thank Bastien François for his final comments on this paper. Below, we give a comment-by-comment response.

General Comments

Comment: In the third draft, the authors of "Impact of bias nonstationarity on the performance of uni- and multivariate bias-adjusting methods: a case study on data from Uccle, Belgium" took into account all my comments in a satisfactory way, including 1) the improvement of the design experiment to compare the different bias correction methods properly, 2) the description of Rank Resampling for Distributions and Dependences, 3) the improvement of section 4.1 "Bias change". These modifications improve the manuscript. I appreciate the work done by the authors to modify the initial draft.

However, the second reviewer mentioned the issue that the evaluation has been done only for one cell and a single model. I agree with the second reviewer that this issue is key to (potentially) reach a firm conclusion on the performance of univariate and multivariate bias correction methods, which is not the case in the present study. Indeed, drawing conclusions on the performance of these statistical methods using only one grid point seems to be subject to much uncertainty.
From my point of view, this point is so important and should be considered in a future version, to provide clearer conclusions to end-users.

Response: We thank the reviewer for his appreciation of our work. In addition to the better framing of the choices that was added in the previous review round, some ideas on how to assess the bias nonstationarity on a continental scale have been added.

Specific Comments

Comment: Figure 1: the panel (f) is missing.

Response: We thank the reviewer for noticing this. The figure has been updated.

Comment: After the different modifications during the peer-reviewing process, I found the final submission very detailed, which is quite interesting, but makes it sometimes difficult to read. From my personal point of view, an effort to be concise must be made, so that readers do not have problems following the manuscript. In particular, are there strong differences of conclusions between 4.2-4.4? Would gathering the results and summarizing concisely the conclusions obtained possible (and preferable)?

Response: We thank the reviewer for indicating lack of clarity. To improve this, the last paragraph of each subsection has been expanded to summarize the results for the corresponding variable or measures. In addition, the second paragraph of the Discussions and conclusions section has been extended to give a global summary of the results.

Comment: I often found it hard to link the information given by the subsection 4.1 Bias Change, and the other sections 4.2-4.7 giving the results for the performance of the methods while this link is the core of the study.

Response: This information was given throughout the text, but to further improve clarity, the extended paragraphs mentioned in the response to the previous comment also allowed for a better discussion of the link between the R index values and the performance of the methods.

Response to Referee Comment 2 on 'Impact of bias nonstationarity on the performance of uni- and multivariate bias-adjusting methods' by J. Van de Velde et al.

Anonymous referee

We would like to thank the referee for the additional comments on this paper. Below, we give a comment-by-comment response.

General Comments

Comment: The reviewer appreciates the work made by the authors to improve the manuscript, this includes additional framing of the results and the work as a case study of value, with much improved justification of the choice and discussion of the limitations. The additional results tables provide much needed clarity to the results. My only general comment would be to urge the authors to limit their use of long paragraphs where important ideas can get lost to the reader (two examples of this are included in the specific comments). Some minor comments follow.

Response: We thank the reviewer for the appreciation of our work. In addition to the paragraphs mentioned in the specific comments, the comments by Referee #1 have also led to more clarity and some improved paragraphs.

Specific Comments

Comment: L 19-20: To which adjustment do you refer here?

Response: This refers to the adjustment (by all methods) in the validation period, in comparison with the adjustment in the calibration method. The sentence has been updated.

Comment: Table 3: P5 is NAN in all seasons, why? Should it remain in the table?

Response: The R index cannot be calculated for P5, as all values of P5 refer to dry days with zero precipitation and no bias. This is not referred to anymore in the text, and P5 is not used in the evaluation. Hence, we followed the suggestion to remove P5 from the table.

Comment: L 405: Consider separating table 3 and table 4 results into two paragraphs.

Response: The discussion of these tables is now split into two paragraphs.

Comment: Figure 1: Panel f is missing.

Response: We thank the reviewer for noticing this. The figure has been updated

Comment: L 437: Why can't these percentiles be plotted, is it by choice? If this is the case, a reference in the text justifying the decision to limit the figure's range would be useful.

Response: This is indeed by choice. The RB_MB or RB_O values of these percentiles are (much) larger than 1, and thus plotting them wouldn't allow an easy comparison between either calibration and validation period or variables. This clarification has also been added to the text.

Comment: Figure 6 caption: R2D2 tag is missing next to (f).

Response: We thank the reviewer for noticing this, the caption has been updated accordingly.

Comment: L 528: It could be worth discussing that there is a seasonal component to the effect of ETP, could winter biases be more relevant for flood modelling than summer biases in ETP?

Response: This is indeed relevant, as causes and types of floods differ depending on the season. The difference between pluvial floods in summer and fluvial floods in winter, and how the PDM is primarily meant to capture the latter type of floods, is now discussed in the text. In addition, the text has been expanded to more extensively discuss the nonstationarity propagation of evaporation, as the biases are largest in spring, whereas the impact of soil moisture on floods is largest in winter.

Comment: Table 8: mQDM, Winter, Cal: value is 1%, should this be 100%?

Response: This should indeed be 100%, and has been updated accordingly.

Comment: L 627-660: This is a very long paragraph which can be split to ease reading (e.g., along sentences starting with First, Second, Third…).

Response: We thank the reviewer for providing this perspective. As guidance in the form of First, Second… was already present, the paragraph has now been split in several smaller paragraphs.

Comment: L 648: Related to the above, the sentence starting here is important and should be highlighted as an important result (perhaps by starting a new paragraph?)

Response: In addition to splitting the paragraph based on the four observations, the last part of the former paragraph, starting at line 648, has now been made a separate paragraph.