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 time spent reviewing our manuscript. Below, we give a comment-by-comment response.

General Comments

Comment: The authors did not explicitly present the study area of this paper. After looking for a while, I'm surprised to find that only one station and one grid cell were used (lines 115-143). In addition, there is only one GCM-RCM combination was used, and no information on its spatial resolution. As such, the results of this study are subjected to large uncertainty.

Response: We believe our research design is still informative, as it is too date scarcely addressed and can provide a base-line for more extended follow-up studies. Yet we simultaneously agree on the limitation of our design, and will ensure these are properly stated both the Introduction and the Discussion.

Comment: The manuscript is lengthy and hard to follow. I think some information does not need to be presented in detailed in the manuscript. For example, the introduction of the bias correction methods (section 3.1 to 3.3, almost 14 pages of the main body). All these information can be found in literatures.

Response: We prefer not to completely remove the technical information, but move it to the Appendix instead, as suggested by Referee #1. The reasons for not removing the information are threefold. First, not every original paper is easy to follow without understanding additional technical details. In our paper, we tried to find a balance between mathematical background and the essence of the paper. Second, we have made some small changes to implementations and these are easier to understand with the full background. Third, the characteristics of the multivariate methods are of great importance to the results and to give the reader an easy reference, we wanted to have the most important information close by.

Comment: Although the authors have stated in line 527 "As the effect on discharge is the overarching goal of this paper", I think the information on the hydrological model and hydrological simulation this quite poor in the manuscript. Firstly, I did not find how the hydrological model performed in the study area (e.g. Nash value or some other criteria). Although the goal of this study is to compare the difference between the univariate method and multivariate method in the hydrology perspective, the hydrological model should be well calibrated. Secondly, as is stated that the PDM model was not calibrated in Uccle but Grote Nete watershed, please give the evidence to show that it is feasible to drive the PDM using the climate data in Uccle.

Response: Some details on the calibration performance will be added. Besides, the Grote Nete watershed is roughly 50 km from the Uccle station and the effect of topography on weather is negligible. Hence, the assumption that the Uccle data can be used for the watershed is acceptable, as discussed in the referenced and other earlier papers.

Comment: The authors used scatter plots throughout the whole paper. It is difficult to see how many dots are located within the 0-1 square or to do the comparison (e.g. Figure 6). I suggest that maybe the authors can use some more quantitative metrics and figures to show the results.

Response: The scatter plots allowed for the inclusion of multiple indices or interesting percentiles of the variables' distribution in one plot and are hence practical for conveying information. However,

we understand that using only the RBmb and RBo values might impede some readers in quickly grasping the results. As such, we will also include other measures, such as Perkins Skill Score and ETCCDI indices, wherever applicable. Yet, not every metric is as practical. For example, while the RMSE provides general information on the fit, it hides the detailed information on the percentiles. Nonetheless, by balancing both new and standard metrics, we aim to provide a good overview.

Comment: Figure 3 to figure 9, the authors did not mention which period (calibration or validation) and which month the results are based, or is it based on the whole year? In this case, readers cannot understand and assess the results. For example, for the correlations between two variables (e.g. P and T), the correlation coefficients are quite different for each month (e.g. summer and winter), therefore, at least, they should be evaluated separately for each month.

Response: These figures were based on the validation period, with data for the whole year. All figures will be updated to clarify the time frame. Besides, as suggested, we have calculated monthly and seasonal data. This indeed allows for a better assessment of the inter-seasonal differences, which provides interesting insights. To provide a balanced overview of the results and keep a proper length of the paper, we prefer the seasonal evaluation. All results will be rewritten to reflect the updated validation.

Comment: Table 1, I am wondering whether the Spearman correlation coefficients and lag-0 cross-correlation between two variables reflect the same thing.

Response: The Spearman correlation is based on the rank of the value, the crosscorrelation is based on its effective value; these two values thus do not reflect the same thing. This can also be seen in Figure 7, where the RBmb and RBo values for the Spearman correlation and the always differ by some extent.

Comment: In general, I found that many expressions in the current manuscript are not very accurate. For example, in line 568 the authors write "A surprising result for P is the high RBMB value for P99.5 for MRQNBC", but I found in figure 3 that the corresponding value for MRQNBC is quite small. Therefore, I'm quite confused by many expressions in the current manuscript.

Response: This was an incorrect statement and we thank the referee for pointing this out. After revising the manuscript, we will make sure to properly proofread the manuscript, so there won't be any incorrect and confusing expressions left.