Richard Chandler - 'Comment on egusphere-2023-1481' - October 2nd, 2023

Comments by the reviewer in blue italics - responses by the authors in black

We thank Richard Chandler for his comments which have strongly improved the manuscript. We provide detailed responses below.

This paper makes a welcome contribution to the sometimes murky world of statistical bias correction, by providing a publicly available software tool that allows users easily to assess the effects / unintended consequences of different "correction" methods. It will be interesting to see whether this makes a substantial change to current practice.

I have just three comments:

1. *If I understand correctly, the tool allows users to assess the effect of bias correction methods on a limited number of threshold-based indices. Some of the visualisations are linked to specific metrics (e.g. the cumulative distribution functions for the spell lengths in Figures 4 and 5). Others, such as the boxplots, are completely generic however. I wonder how easy it would be to link to, say, the xclim library (https://xclim.readthedocs.io/en/stable/) which defines a whole range of other climate indices? If you could import those index definitions and provide some core visualisations - such as boxplots - for them, then ibicus would become a really powerful tool.*

Response: This is correct with one exception: the threshold-based indices within ibicus are meant to be generic, meaning that users can define their metrics and use the full range of generic visualisations on them (ranging from the CDF plots to boxplots and analysis of trend modification). The pointer to the xclim library however is very helpful. In order to minimise dependencies, we decided that explicitly including xclim with the ibicus package is not useful. However, it is easy to export the output of ibicus, compute xclim indices on that and visualise the result of that using xclim. We will include an example of that in the next software release.

2. *Section 4.2.4 claims that bias adjustment changes the uncertainty in an ensemble. Although this kind of claim is commonly made, it is not true: the uncertainty is what it is, and it doesn't change just by massaging the data. Bias adjustment changes the variation which is a symptom of the underlying uncertainty, but that's not the same as changing the uncertainty itself! This is connected to my final point, which is ...*

Response: Thank you for this comment. To address it, as well as the comments of Jorn Van de Velde, we have expanded our discussion of this point and adjusted the text as follows:

Original text: “4.2.4 Evaluation of the uncertainty in the climate model ensemble before and after bias adjustment “

Modified text: “4.2.4 Evaluation of the variation in the climate model ensemble before and after bias adjustment “

Original text: Figure 8 shows that the climate model ensemble spread of the trend in mean seasonal precipitation is modified when applying bias adjustment. This means that not only the trend but also the range of uncertainty and possible worst-case scenarios analysed in impact studies depend on the bias adjustment method used to pre-process the climate model. As shown in the previous sections, the ’best’ bias adjustment method for a given use case depends on the variable, region and impact variable studied. The result shown in figure 8 demonstrates that bias adjustment can add an additional source of uncertainty if the method is applied blindly and not evaluated properly. Interestingly the uncertainty range is not necessarily narrowed as has been postulated by some authors (Ehret et al., 2012), but even extended and shifted in some cases.

Modified text: Figure 8 shows that the climate model ensemble spread of the trend of mean seasonal precipitation is modified in different ways by different bias adjustment methods which is in line with previous findings in the literature (Maraun and Widman 2018, Lafferty et al. 2023). Interestingly the variation (often interpreted as the uncertainty range) is not necessarily narrowed as has been postulated by some authors (Ehret et al., 2012), but even extended and shifted in some cases. From this finding, it follows that the range of uncertainty and possible worst-case scenarios analysed in subsequent impact studies might depend on the bias adjustment method used to pre-process the climate model.

The interpretation of this shift in uncertainty is related to the previously discussed questions on trend preservation, namely whether the change in the climate model trend through a statistical bias adjustment method is justified or not. This issue was mentioned by Maraun and Widmann (2018), who discuss that a minimum requirement to justify a change in the uncertainty spread through bias adjustment should be a critical evaluation of the validity of the results and the assumptions of the underlying statistical model. Given the finding in the previous section, namely that the best bias adjustment method depends on the variable, region and impact variable studied, it follows that indiscriminately applying a bias adjustment method across regions and variables without evaluation can shift the spread of the results of subsequent impact studies in a non-justified manner.

3. Users may feel as though you've provided a tool that is specifically designed to pull the rug out from under their feet, in the sense that it will almost certainly reveal that there are problems with their chosen bias adjustment method. I am personally rather supportive of any contribution that demonstrates the problems of bias adjustment, but it would perhaps be helpful to provide some constructive suggestions for how to proceed if your software reveals major problems. One such alternative, for example, is to postprocess the entire ensemble within a statistical framework that acknowledges the discrepancies between climate models and the real world, and that aims to derive defensible uncertainty assessments for the real world on the basis of all the available information. There is a fair bit of literature on this: I have made a limited contribution myself, but other authors include Michael Goldstein, Jonathan Rougier, Phil Sansom, Christoph Buser and Claudia Tebaldi (the list goes on!).

We thank Richard Chandler for the pointer to the literature on post-processing entire ensembles. We have expanded our discussion section quite substantially overall and have added a reference to this strand of literature. However, we felt that a detailed discussion of alternative approaches would be outside the scope of this particular GMD submission.