Responses to Anonymous Referee #1

We thank the Anonymous Referee #1 for their time and effort to review our manuscript, which helps to improve the paper. We provide a detailed response to the Referee's comments.

Below, reviewer comments are marked in red.
Responses to the comments are marked in blue.
Changes that have been made in the manuscript are marked in *italic*.

1. Line 35-40: This is a good and important discussion on key limitations of bias correction. However, the methodology used in this study did not addresses these challenges if your BC approach did help to "address the origin of model errors", that would be a significant contribution to the community. As it stands, the manuscript does not seem to offer solutions for tackling the root causes of model errors.

Thank you for raising this important point. We agree that addressing the origin of model errors would indeed represent a significant advance in the field. However, this is beyond the scope of our study, which focuses on providing a high-quality, bias-corrected dataset using established methods to support impact assessments, rather than addressing the underlying causes of these errors.

We refer to the manuscript types specified by ESSD (https://www.earth-system-science-data.net/about/manuscript_types.html). According to the journal's guidelines, data description papers in ESSD emphasize the quality, accessibility, and usability of datasets. Although demonstrating data quality is essential, extensive interpretations of data, as expected in a research article, fall outside the scope of this journal. ESSD specifically focuses on publishing high-quality research datasets to promote data reuse, rather than introducing novel methodologies or performing comprehensive comparisons.

Our study aligns with these goals by applying established bias-correction methods to improve the usability of climate projections for impact assessments. To further support the value of our work, we note that one of our datasets, available at https://zenodo.org/records/6337381, has already been downloaded 245 times as of 05/11/2024, demonstrating its significant contribution to the community.

2. Line 66: "simple BC methods used in CHESS-SCAPE": what methods they use?

We appreciate your request for clarification. In fact, an explanation is provided in Lines 60-65 of the manuscript. The CHESS-SCAPE dataset applies simple linear scaling, with additive corrections for air temperature and multiplicative corrections for precipitation. Seasonal offsets are used for temperature, while scaling factors are applied for precipitation. Specifically, temperature offsets are calculated as the difference between CHESS-SCAPE and CHESS-met observations for each season and grid cell, and precipitation scaling factors are the ratio of CHESS-SCAPE to CHESS-met precipitation. We have revised the manuscript to clarify this method.

However, the simple linear scaling BC method used in the CHESS-SCAPE does not take into account changes in higher-order moments than the mean.

3. The data you corrected is from a perturbed physics ensemble, which is designed to both explore the influence of parameter variations on simulations and to reduce uncertainty. It would be highly valuable to examine how bias correction affects the ensemble results. For example, does using a single observation dataset to correct multiple ensemble runs impact the ensemble's representation of uncertainty? Additionally, it would be useful to see a discussion on the performance of individual ensemble members before and after bias correction. Such an analysis could provide insights into how bias correction interacts with model parameterizations and the ensemble spread, and provide physical interpretation of your results.

We appreciate this suggestion. In fact, we have already considered the effects of bias-correction on the ensemble spread, as seen in Fig. 4 showing the spread of monthly P/T before and after BC and figures for standard deviation ('sd' rows in Fig. 5, 6, 11, and 12) of reference period error and projected changes of a number of P/T indices. The individual ensemble members' performance and projected changes are also shown in these figures, showing how the bias-correction influences them. It is important to note that there are 12 sets of perturbed physics ensemble (PPE) parameters drawn from distributions of 47 parameters (as described in Section 2.1.1). This sparse coverage does not allow for deriving physical explanations of model behaviour or performance related to these perturbations, and the UKCP18-PPE-ensemble was not constructed with this purpose in mind.

4. The overestimation of dry-day frequency is a common issue in climate models, often linked to the drizzle effect. I recommend addressing the drizzle effect first, as this might improve the accuracy of subsequent bias corrections and separate this systematic bias from others.

Thank you for raising this point. Your comment might be referring to overestimation of wet-day frequency and hence the drizzle effect. The bias of dry-day frequency is already provided in Fig. 1 along with the text in Section 3.1.1 (Lines 231-240). Figure 1 shows the ensemble mean errors of the raw UKCP18-RCM projections in the dry-day frequency, mean daily precipitation and the Q95 of precipitation in the reference period, expressed as a percentage of the observed value. In general, the frequency of dry days in UKCP18-RCM is too low (and therefore the wet-day frequency is too high), particularly in the winter and in regions of higher elevation. In summer, the dry-day frequency bias is very small for most of England. The precipitation mean and Q95 are strongly overestimated across the UK in winter, although in highly elevated areas this bias is smaller or even reversed in sign (especially for Q95). In summer, however, the mean and Q95 biases show a strong spatial variability, with underestimations toward the south and at high elevation levels, and a wet bias in the north of the UK. These seasonal bias differences result in an annual bias of too few dry days almost everywhere, too wet mean precipitation in most regions, and more mixed wet and dry Q95 relative biases.

The QM and trend-preserving BC methods employed in this study address the drizzle effect (overestimation of wet-day frequency) by matching the quantiles or re-distributing the precipitation values to align with the observed. In other words, both BC methods adjust the distribution of those lower-quantile values (light precipitation events) to align with the observed. The sections 2.4.1 and 2.4.2 already explained the methods and how the distribution is adjusted, therefore no further change is made in the manuscript.