Responses to Anonymous Referee#2

We thank Anonymous Referee #2 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, referee's 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. This paper evaluates two relatively simple statistical methods for correcting climate model output directly. The results indicate that bias correction techniques can mitigate biases related to the indices utilized in this study. Nonetheless, my principal concern is the paper's originality. Numerous studies have already been published that compare various statistical and machine learning methods across different domains and temporal contexts, ranging from historical data to future projections. Consequently, I do not believe this paper contributes significantly new information relative to existing literature, particularly given that available published data could offer additional insights. While the journal is dedicated to data publication, it is essential for this paper to either introduce a novel methodology or provide a comprehensive analysis of each method, highlighting the advantages and disadvantages of their application in specific circumstances. Merely comparing simple methods that produce comparable results for future projections may not provide substantial value.

Thank you for raising this important concern. We acknowledge that introducing a novel methodology would represent a significant contribution to 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 introducing a new bias-correction method.

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.

Finally, we would like to point out that even though the two bias-correction techniques are not novel, their application to the entire UKCP18 regional dataset (whole of UK, 12 members, 1981 to 2080, and three variables, including PET) is novel. The process of narrowing down the selection of the two bias-correction methods, preprocessing the data, generating the bias-adjusted data, and evaluating the results for broader interests took about a year. This substantial effort ensures that research on climate change impacts in the UK can progress without duplicating efforts to bias-correct data that has already been carefully processed. The reviewer writes "Merely comparing simple methods that produce comparable results for future projections may not provide substantial value". We believe instead that these datasets do provide substantial value, and based on the fact that one of our datasets (https://zenodo.org/records/6337381) has been downloaded 246 times as of 19/11/2024, we would like to argue that the community sees this value too.

2. The manuscript lacks a comprehensive explanation for the preference of two simpler methods over more sophisticated approaches. The selection of the degree of bias correction should depend on the specific application, as advanced techniques have been shown to produce greater improvements, particularly in the context of extreme events. While quantile mapping and simple climatological mean correction have demonstrated their advantages in preventing excessive correction based on observational data, they may still permit the persistence of biases related to low-frequency variability, which can complicate the direct correction of surface variables in climate model outputs.
Numerous studies have employed sophisticated bias correction across various time scales to adjust the outputs of GCM and RCM or the boundary conditions for RCM inputs, with the objective of enhancing the accuracy of simulations for extreme and compound events. It would be advantageous to incorporate additional references that pertain to bias correction prior to addressing the limitations of existing studies. Furthermore, it is essential to provide detailed information and explanations regarding using simpler methods in the introduction or conclusion to ensure a thorough understanding of the techniques ranging from simple to sophisticated methods.

- Correcting outputs
Wood AW, Leung LR, Sridhar V, Lettenmaier D (2004) Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs. Clim Change 62:189–216
Cannon, A. J. (2018). Multivariate quantile mapping bias correction: an N-dimensional probability density function transform for climate model simulations of multiple variables. Climate dynamics, 50(1), 31-49.

- Correcting RCM input full variable fields
Bruyere CL, Done JM, Holland GJ, Fredrick S (2014) Bias corrections of global models for regional climate simulations of high-impact weather. Clim Dyn 43:1847–1856
Kim, Y, Evans, JP, Sharma, A (2023). Can Sub‐Daily Multivariate Bias Correction of Regional Climate Model Boundary Conditions Improve Simulation of the Diurnal Precipitation Cycle?. Geophysical Research Letters, 50(22), p.e2023GL104442.

- Software for correcting climate model variables
Cannon, A.J. (2016). Multivariate bias correction of climate model output: Matching marginal distributions and intervariable dependence structure. Journal of Climate, 29(19), pp.7045-7064.
Kim, Y., Evans, J. P., & Sharma, A. (2023). A software for correcting systematic biases in RCM input boundary conditions. Environmental Modelling & Software, 168, 105799.

We agree that the selection of the bias-correction method should ideally be tailored to each specific impact study and the types of hydrological or meteorological extreme events being studied. However, barriers such as time/budget constraints, lack of knowledge of the available methods and the evolving field of bias adjustment, lack of validation/testing in impact modelling studies, and/or lack of available software implementing these innovative specialized bias-correction methods, can stand in the way of their uptake in many impact modelling studies. We make available two sets of bias-adjusted datasets for the community and provide results of a multi-metric evaluation, but as we emphasize in our conclusions section, in the end it is up to the user to decide which bias-correction method or bias-corrected dataset fits their specific purpose. We actively encourage potential users to perform their own evaluations in lines 431-432. (The significant added value of our work has been discussed in our response to major comment 1).

With the established quantile mapping method and the novel quantile mapping-based trend preserving bias-correction method developed for the 3rd phase of the ISIMIP-project, we believe we strike a good balance between simple and more sophisticated methods, as these (1) correct higher order moments (which commonly applied simpler methods don't do), and (2) have indeed proven themselves and have been put to the test in hydrological impact modelling studies.

Thank you for raising the concern of lower-frequency variability. To provide prospective users with insight on the degree to which multi-day metrics are corrected, we addressed this in the evaluation of the bias-adjusted datasets by looking at the errors and projected changes in the longest annual/seasonal streaks of consecutive wet days, consecutive dry days, and maximum 5-day total rainfall. As we discussed in L295-301, Fig. 5 shows that these are indeed more challenging to correct, but they are nevertheless improved to a great degree by both bias-adjustment methods.

We agree to discuss the suggested additional literature in the introduction section.

3. The authors have undertaken corrections for three variables: precipitation, temperature, and potential evapotranspiration (PET). It would be beneficial to clarify the methodology employed for correcting PET. Did the authors correct PET derived from the raw variables directly, or did they utilize the adjusted variables in the PET calculations? As indicated by the authors, PET is influenced by several variables generated from the climate model, including specific humidity, pressure, and temperature, which can be modified by adjusting surface variables interconnected with these factors. Correcting these variables statistically, without accounting for the physical relationships among them, may lead to inconsistencies and produce unrealistic results.

Thank you for the comment. PET was calculated from the uncorrected (i.e. "raw") model variables (specific humidity, atmospheric pressure, net downwelling longwave and shortwave radiation, 10m wind speed and surface air temperature) and then the PET was bias-corrected against PET calculated from observed variables. One reason we did not do the alternative (i.e. bias correct the individual variables and then calculate PET from the bias-corrected individual variables) is that the PET might still show biases compared with observation-based PET because, as the referee notes, there may be inconsistencies between the individual variables in the model simulation. Taking this approach might then require a further bias-correction step to correct for remaining biases in the calculated PET, which our approach avoids. We thank the referee for highlighting this question and we have added the following sentence to the end of section 2.2 (potential evapotranspiration) to make it clear to the reader:

PET was calculated from the uncorrected model variables and then it was bias-corrected against PET calculated from observed variables (i.e. CHESS-PE). This was preferred to the alternative approach of bias correcting the individual variables and then calculating PET from the bias-corrected individual variables because that PET might still show biases compared with observation-based PET, requiring a further bias-correction step.

Specific comments:

L66. "the simple methods …" It would be beneficial to provide more details about what these methods entail.

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.

L68. "the quantile mapping (QM) method outperforms … the standards deviation and percentiles." Quantile mapping (QM) can outperform simple mean correction for variance since the latter does not address the standard deviation or percentiles. I recommend incorporating more details about bias correction techniques, including methods like simple mean and standard deviation correction. This will help justify the use of empirical QM when publishing datasets for broader applications.

Thank you for your comment. Several studies have conducted similar comparisons of bias-correction methods. We have revised the manuscript as follows:

Several studies have compared different BC methods including linear scaling, delta change, local intensity scaling, variance scaling and QM. These studies consistently find that QM outperforms other BC methods in effectively correcting higher-order statistical properties, such as standard deviation and percentiles (Azmat et al., 2018; Teutschbein and Seibert, 2012; Worku et al., 2020). Specifically, empirical QM approaches, such as those applied by Fang et al. (2015) and shown by Enayati et al. (2021) to be effective for correcting precipitation and temperature biases, highlight the strengths of nonparametric techniques.

L82. “the trend-preserving BC method.” It is essential to justify the trend-preserving method for future projections, as climate models also contain biases in trends.
We employed two methods. Both correct distribution biases: one does not explicitly consider trends in the projections, while the other one explicitly considers them and is designed to preserve them. Our results show that the projected changes are essentially insensitive to these two different approaches. This is illustrated by the future changes in the raw data and in the two bias-corrected datasets being overall similar (Fig. 9 and 10). As the referee mentions, the differences between these maps are “minimal”. Importantly, they are much smaller than differences between members (see e.g. Fig. 12). In other words, larger uncertainties are introduced by running an ensemble than by treating future trends differently. In addition, the Met Office performed an evaluation of long-term drift in the UKCP18 global and regional PPE simulations. Optionally, we can also discuss this report in the "data" section to address this comment.

Figures 9 and 10. I recommend modifying the figures to use a bias map instead of presenting each one individually, as the differences are minimal.
Thank you for your comment. While we recognize that the differences between the raw and bias-corrected data in Fig. 9 and Fig. 10 are small, these figures serve two purposes, which are more efficiently fulfilled by the current representation. First, they demonstrate that the projected changes are retained after bias-correction. Second, they are crucial for illustrating the spatial distribution of temperature and precipitation changes within the UKCP18 model. This spatial representation provides valuable insights into how these projected changes across different regions, and the minor distortions of the spatial coherence of projections by the quantile mapping method in the context of the magnitude of the projected changes (e.g. for summer temperature in Fig. 10), which would not be as clear in a bias map. Therefore, we believe that maintaining the current figure more effectively communicates both the preservation of projected changes and their regional variability.

References

Azmat, M., Qamar, M. U., Huggel, C., and Hussain, E.: Future climate and cryosphere impacts on the hydrology of a scarcely gauged catchment on the Jhelum river basin, Northern Pakistan, Science of The Total Environment, 639, 961–976, https://doi.org/10.1016/j.scitotenv.2018.05.206, 2018.

Enayati, M., Bozorg-Haddad, O., Bazrafshan, J., Hejabi, S., and Chu, X.: Bias correction capabilities of quantile mapping methods for rainfall and temperature variables, Journal of Water and Climate Change, 12, 401–419, https://doi.org/10.2166/wcc.2020.261, 2021.

Fang, G. H., Yang, J., Chen, Y. N., and Zammit, C.: Comparing bias correction methods in downscaling meteorological variables for a hydrologic impact study in an arid area in China, Hydrology and Earth System Sciences, 19, 2547–2559, https://doi.org/10.5194/hess-19-2547-2015, 2015.

Teutschbein, C. and Seibert, J.: Bias correction of regional climate model simulations for hydrological climate-change impact studies: Review and evaluation of different methods, Journal of Hydrology, 456–457, 12–29, https://doi.org/10.1016/j.jhydrol.2012.05.052, 2012.

Worku, G., Teferi, E., Bantider, A., and Dile, Y. T.: Statistical bias correction of regional climate model simulations for climate change projection in the Jemma sub-basin, upper Blue Nile Basin of Ethiopia, Theor Appl Climatol, 139, 1569–1588, https://doi.org/10.1007/s00704-019-03053-x, 2020.