Responses to Anonymous Referee #1

We thank the Anonymous Referee #1 for their time and effort to review our manuscript, which helped to further improve the quality of 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*.

General Comments:

1. The paper describes the method used to empirically quantile map hourly rainfall and daily temperature from the UKCP18 CPM model using observed datasets. The article is well presented and very clearly described, enabling readers to understand the underlying methods and reasoning with relative ease.
   We appreciate the referee's positive assessment of our manuscript.

2. Some additional content would improve the justification of the method – specifically the impact of using the bias controlled dataset relative to the original. This update should be relatively simple and not significantly add to the length of the paper. The paper also alludes to river catchments – it should be stated why this is as it is otherwise out of context.
   Thank you for this suggestion. We agree that the manuscript should more clearly justify the impact of using bias-corrected data relative to the original (raw) CPM outputs and clarify why catchments are introduced. We have strengthened the Discussion to better justify the use of the bias-adjusted dataset relative to the raw CPM output. We now emphasise that the raw UKCP18-CPM precipitation and temperature data show systematic biases across multiple temporal and spatial scales, which could affect downstream impact assessments if used directly. We have also added references to previous studies to place these biases in a broader context.
   The second paragraph in Discussion (lines 413-422) has been revised to: *"Before bias correction, the UKCP18-CPM simulations show considerable biases in both precipitation and temperature across various temporal and spatial scales. In general, UKCP18-CPM simulations show wet precipitation biases (especially in winter) and cool biases in temperature. These broad patterns are consistent with previous evaluations of UKCP18-RCM simulations (Reyniers et al., 2025) and the UKCP18-CPM science report (Kendon et al., 2019b). These residual biases are consistent with the broader CPM literature, which shows that although kilometre-scale models can improve the representation of precipitation compared with convection-parameterised RCMs, systematic biases and model uncertainty may remain, particularly for sub-daily precipitation and across different regional settings (Ban et al., 2021; Correa-Sánchez et al., 2025; Soares et al., 2024). After bias correction, both precipitation and temperature show substantially improved agreement with observations at monthly and*

diurnal timescales (Figs. 4 and 5). This suggests that bias correction provides a more reliable basis for downstream impact assessments than raw UKCP18-CPM output."

The opening sentence of the last paragraph in the Discussion (lines 451–452) has been revised to: *"This high-resolution bias-corrected dataset offers enhanced reliability, making it suitable for a wide range of climate change impact simulations and studies."*

Regarding catchments, we agree this required clearer framing. We have revised Section 2.1 to clarify that catchments are used only to define the analysis mask (i.e., to select which grid-cells are included). The bias correction itself is applied at the grid-cell level, and no hydrological modelling is performed in this study. This clarification is now given in the revised opening paragraph of Section 2.1. We have also revised Figure 1 to show the processed 1 km grid cells (displayed by their centre points) and removed the gauges, as they were not necessary for explaining how the analysis mask was defined.

The paragraph in Section 2.1 has been replaced with: *"This study focuses on England (Fig. 1), with the primary aim of assessing the bias-correction method. Catchment boundaries are obtained from 249 catchments in the National River Flow Archive (NRFA) and are used only to define the analysis mask (i.e., to select the 1 km grid cells processed). No hydrological modelling or catchment-scale simulation is performed. The processed HadUK-Grid cells within this mask are displayed in Fig. 1 by their centre points. Given computational and storage constraints, restricting the analysis to grid cells within these catchments provides a practical and geographically coherent subset for the study. The catchments were selected based on the availability of complete gauged records for December 1990 to November 2000, overlapping with both the CEH-GEAR1hr observational dataset and the UKCP18-CPM baseline period (see Sect. 2.2.2; hereafter referred to as the reference period), and on minimal anthropogenic influence. Bias correction was applied independently at the grid-cell level to UKCP18-CPM precipitation and temperature for 62,488 1 km grid cells. Across the processed grid cells, the mean annual precipitation is approximately 2.64 mm d^{-1}, and the mean annual temperature is around 9.3 °C, based on the CEH-GEAR1hr and HadUK-Grid datasets, respectively."*

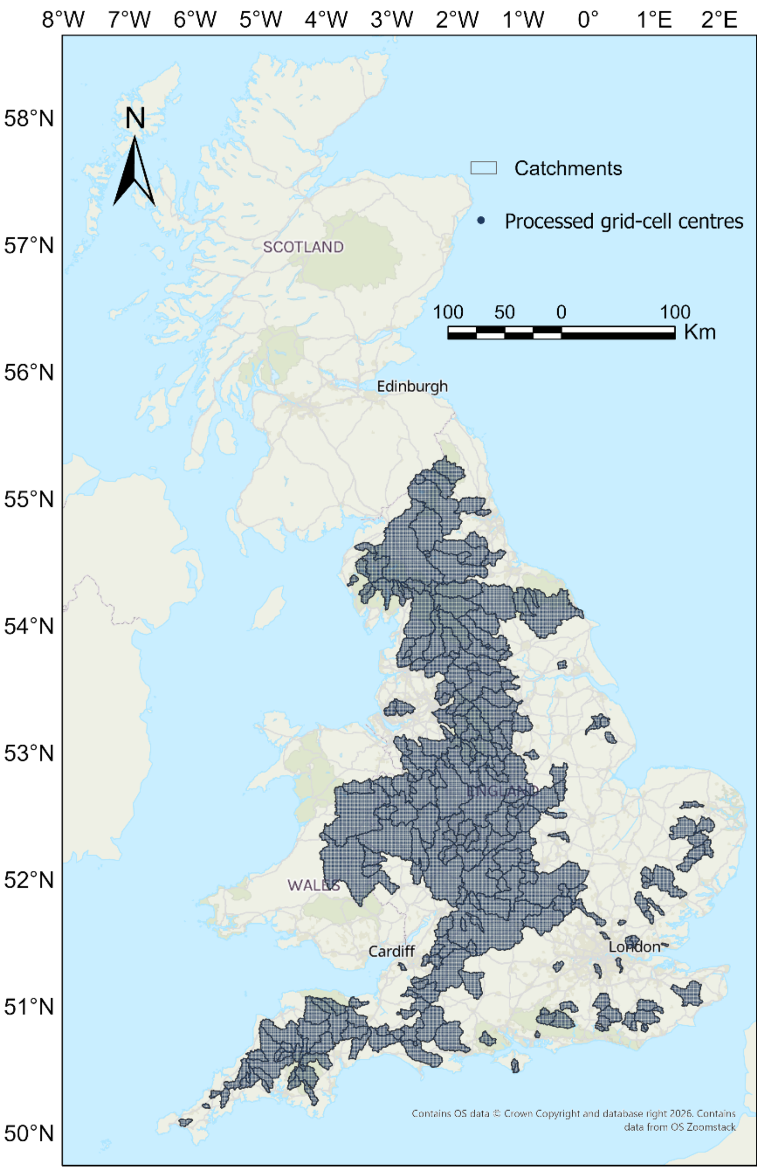

Figure 1: Study domain in England. Catchment boundaries (249 catchments) are used only to define the analysis mask for grid-cell processing. The processed 1 km HadUK-Grid cells are displayed by their centre points.

3. The purpose of the paper should also be stated more clearly from the abstract and introduction. Is the paper there to develop a method, to build a dataset for others to use, to demonstrate the inadequacy of the data in its raw form, or to demonstrate the importance of quantile mapping. Whichever the motivation, some additional content is required – either clearly stating that this approach is new and how others can use it, stating within the article where the data can be found and how it can be used, clearly stating that the raw data should not be used for the authors purpose, or by demonstrating more directly the impact of the BC on the future datasets.
Thank you for the comment. We agree that the purpose of the paper and the justification for using the bias-corrected dataset needed to be stated more clearly. We have therefore revised the Abstract (including the purpose sentence and final sentence) and added a

clearer statement of scope and aim in the Introduction, as detailed in our response to the first Specific Comment below.

Regarding the bias-correction method, we have expanded the Discussion to clarify this bias-correction method, how we use it, and how it may be used by others.

The following paragraph has been added after the first paragraph in Discussion (line 413): *"We applied empirical QM to daily mean temperature to address seasonal biases and distributional differences using a reproducible procedure. It is widely used because it can correct not only mean biases but also biases in variability and quantiles (Fang et al., 2015; Themeßl et al., 2011; Wilcke et al., 2013). For hourly precipitation, we used DBC by hour of day, because applying a single mapping across all hours can overlook systematic diurnal-cycle biases in sub-daily precipitation. Figure 5 supports this choice, showing that the bias-corrected ensemble more closely reproduces the observed hour-of-day pattern. This hour-of-day treatment is intended to reduce discrepancies that matter for sub-daily impact applications, consistent with the broader argument that correcting the diurnal cycle can improve sub-daily bias-correction strategies (Faghih et al., 2022). Therefore, we recommend using DBC for sub-daily variables. However, in this study we applied a 3 h moving window in the DBC. By incorporating hourly data from the previous and following hours, this may have contributed to some of the remaining discrepancies. It may be worth exploring whether omitting the moving window gives better results when the training period sample size is sufficiently large."*

We have also clarified in the Data availability section that the bias-corrected dataset can be accessed from the existing Zenodo archive (line 489).

In addition, we revised the Discussion (lines 413–422) to make clearer that the raw UKCP18-CPM outputs exhibit substantial biases across multiple temporal and spatial scales and are therefore not suitable for our intended impact-assessment purpose without bias correction. The revised text also explains that the bias-corrected data provide a more reliable basis for downstream impact assessments than the raw CPM output. This is addressed in our response to the second General Comment above.

There are several additional comments/queries/suggestions, but all are minor edits and do not require any amount of additional work. While not ground breaking (or intending to be), this is a good clear paper of direct interest to those in this field and looking to use bias correction.

Specific Comments

Abstract – is the purpose of this paper to simply present a dataset? Or to present a method? Or to present findings on the biases? Unclear. Also worth stating around line 91. The paper needs a very minor reframing of its purpose (maybe only a sentence or two) and a bit of additional material to complete that purpose. See general comment.

Thanks for the comment. The main aim of this paper is to present and assess a bias-correction method for UKCP18-CPM hourly precipitation and daily temperature, and to quantify how closely the bias-adjusted UKCP18-CPM simulations align with observations relative to the raw simulations. The bias results are included to motivate the need for correction and to assess the method performance, rather than to present a dataset as the main outcome. We have revised the Abstract and added a clear purpose sentence in the Introduction (around line 91).

Line 11 in the Abstract has been revised to: *"In this study, we apply and evaluate a quantile mapping (QM) bias correction method for UKCP18-CPM hourly precipitation (with diurnal correction) and daily temperature over England. We quantify how closely the bias corrected simulations align with observations relative to the raw simulations."*

The last sentence of the Abstract (line 19) has been revised to: *"Overall, the applied bias-correction method brings UKCP18-CPM simulations into closer agreement with observations for both mean behaviour and extremes, providing a more reliable basis for high-resolution impact modelling and assessments that require hourly precipitation forcing."*

The following sentence has been added at the beginning of the last paragraph of the Introduction (around line 91): *"The purpose of this paper is to present and assess a bias-correction method for UKCP18-CPM hourly precipitation and to quantify how closely the bias-adjusted UKCP18-CPM simulations align with observations relative to the raw simulations. We bias-corrected and evaluated the hourly precipitation…"*

Line 48 – '…the ensemble generally underestimates values across the mean, cold, and hot tails of the distribution' – is that correct? This sounds like it just underestimates everything. (Though I suppose it is possible that it overestimates the shoulders and underestimates the middle/ends).

Thanks for the comment. The sentence could be misleading in suggesting that temperature is underestimated at all parts of the distribution. Our intention was to summarise the findings of Reyniers et al. (2025) that, at the annual scale, UKCP18-RCM shows an overall cold bias (as reflected by the mean and the indices used to represent the cold and warm ends of the distribution), while the sign and magnitude of the biases vary by season and region. We have therefore rephrased and condensed the sentence to avoid implying a systematic underestimation everywhere and to provide a concise annual/winter/summer summary.

The sentence around line 50 has been revised to: *For temperature, the ensemble shows an overall cold bias at the annual scale in the mean and extremes. In winter, biases vary spatially and suggest overestimated variability, whereas in summer temperatures are more consistently underestimated across the UK (Reyniers et al., 2025).*

Line 76 – what distributions? Perhaps a timescale or few extra words would help. E.g. 'distribution of rainfall within the period'.

Thanks for the comment. We have clarified that the limitation refers to the distribution of daily precipitation values within each month.

The sentence around line 76 has been revised to: *However, this monthly scaling applies a uniform adjustment to all daily precipitation values within each month. It corrects monthly means but does not explicitly adjust the within-month distribution (e.g., quantiles and extremes).*

Line 87 – What is the temporal resolution of the '1 km HadUK-Grid dataset'?

The HadUK-Grid temperature product used here is at a daily resolution.

The sentence around line 87 has been revised to: *Therefore, bias correction was performed on the UKCP18-CPM daily temperature using the Met Office's 1 km daily HadUK-Grid dataset (Hollis et al., 2019; Met Office et al., 2022).*

Section 2.1 is the first mention of catchments and hydrology – until this point I thought that this was a climate paper, but maybe it will become a hydrology paper. Introduce the use of this data being for hydrological purposes earlier in the paper (and abstract?) so that this is not out of place. Looking forward there does not seem to be future mention of hydrology/rivers – why was this process only done for some river catchments – would it not have been simpler, and more useful to others, to have conducted the processing nationally rather than for such vocation specific areas? Justify/explain this or consider presenting data for whole of England. If the purpose of this paper is to provide a dataset then it would probably be more useful to process all of England and to present it alongside a table of which catchments have missing data.

Thank you for this comment. The main aim of this paper is to present and assess a bias-correction method for UKCP18-CPM hourly precipitation (and daily temperature), and to quantify how much it improves the raw CPM simulations against observations. We agree that it should be introduced earlier in Abstract and Introduction.

Catchments are used only to define a hydrologically relevant analysis domain (i.e., which grid cells are included), while the bias correction itself is applied at the grid-cell level and no hydrological modelling is performed in this study. Given computational and storage constraints, restricting the analysis to grid cells within some catchments provides a practical and hydrologically meaningful subset for the study. In this paper we focus on England and four members to provide a clear benchmark for the method, and the work is directly extensible to additional members and a larger domain.

The last sentence of the Abstract (line 19) has been revised to: *"Overall, the applied bias-correction method brings UKCP18-CPM simulations into closer agreement with observations for both mean behaviour and extremes, providing a more reliable basis for high-resolution impact modelling and assessments that require hourly precipitation forcing."*

The following sentence in Lines 83-85 has been revised to: *" The CEH-GEAR1hr dataset (Lewis et al., 2022) integrates data from over 1,900 quality-controlled rainfall gauges over Great Britain, providing highly accurate precipitation measurements and enabling more robust sub-daily evaluation and bias correction of precipitation timing and intensity."*

The following sentence has been added at the beginning of the last paragraph of the Introduction (around line 91): *"The purpose of this paper is to present and assess a bias-correction method for UKCP18-CPM hourly precipitation and to quantify how closely the bias-adjusted UKCP18-CPM simulations align with observations relative to the raw simulations. We bias-corrected and evaluated the hourly precipitation…"*

The paragraph in Section 2.1 has been replaced with: *"This study focuses on England (Fig. 1), with the primary aim of assessing the bias-correction method. Catchment boundaries are obtained from 249 catchments in the National River Flow Archive (NRFA) and are used only to define the analysis mask (i.e., to select the 1 km grid cells processed). No hydrological modelling or catchment-scale simulation is performed. The processed HadUK-Grid cells within this mask are displayed in Fig. 1 by their centre points. Given computational and storage constraints, restricting the analysis to grid cells within these catchments provides a practical and geographically coherent subset for the study. The catchments were selected based on the availability of complete gauged records for December 1990 to November 2000, overlapping with both the CEH-GEAR1hr observational dataset and the UKCP18-CPM baseline period (see Sect. 2.2.2; hereafter referred to as the reference period), and on minimal anthropogenic influence. Bias correction was applied independently at the grid-cell level to UKCP18-CPM* *precipitation and temperature for 62,488 1 km grid cells. Across the processed grid cells, the mean annual precipitation is approximately 2.64 mm d^{-1}, and the mean annual temperature is around 9.3 °C, based on the CEH-GEAR1hr and HadUK-Grid datasets, respectively."*

Section 2.1 – Please also explain why only England (if there is a reason). Consider whether to say England and Wales as you have 'English' catchments that cover much of Wales.

Thank you for the comment. We acknowledge that a small number of the selected catchments extend into Wales. However, their gauging stations are located in England, and these cross-border catchments were therefore retained in the analysis mask. We continue to describe the study domain as England for simplicity, while recognising that a small number of selected catchments extend across the England–Wales border. England was chosen because the primary aim of the paper is to present and assess the bias-correction method rather than to produce a national dataset, and limiting the domain also helps keep the computational demand and storage requirements manageable.

Line 107 – CAMELS-GB is mentioned? Are you taking other parameters from this other than catchment outlines? If not then NRFA is perhaps a more appropriate reference for these, or consider stating '… catchment boundaries taken from the CAMELS-GB dataset…').

Thank you for the advice. We only use CAMELS-GB to obtain catchment boundaries and do not use any additional catchment attributes or parameters from the dataset. We therefore removed the mention of CAMELS-GB. We have revised the text around line 107 accordingly and added the NRFA.

The sentence in lines 106-107 has been replaced with: *"This study focuses on England (Fig. 1), with the primary aim of assessing the bias-correction method. Catchment boundaries are obtained from 249 catchments in the National River Flow Archive (NRFA) and are used only to define the analysis mask, i.e., to identify the 1 km grid cells included in the analysis."*

Figure 1 - Would be interesting to also see which catchments were removed due to issues with either human or rainfall data or temperature data, but perhaps surplus.

Thank you for this suggestion. We agree that showing the catchments excluded during the selection process could provide additional context. However, the main purpose of Figure 1 is to present the final study domain and analysis mask used in this work, and adding excluded catchments would make the figure considerably more crowded and less clear. Moreover, exclusions were based on several considerations, including the availability of complete gauged records during the reference period and the degree of anthropogenic influence, and these are more appropriately described in the text (lines 107–110) than represented in a single map. Information on catchment characteristics and anthropogenic influence is also available through the NRFA. For clarity and focus, we have therefore kept Figure 1 as a presentation of the final selected catchments only.

Line 127 – what strand are the CPMs? Or is this a different release/dataset?

Thank you for the comment. The CPM projections used here are the UKCP18 Local (2.2 km) convection-permitting projections (released as an additional UKCP18 product rather than one of the three land strands). The 12 CPM simulations are driven by (nested within) the 12-

member UKCP18 Regional (12 km) Strand 3 RCM ensemble. We have clarified this in line 127.
The sentence in line 127 has been revised to: *"The UKCP18 convection-permitting model (UKCP18-CPM) projections are the UKCP18 Local (2.2 km) product (Kendon et al., 2019b), providing high-resolution climate information at ~2.2 km resolution and driven by the UKCP18 Regional (12 km) Strand 3 ensemble. "*

Line 224 – 'As Reiter et al. (2018) pointed out, QM can correct the distribution'… Is this just for a certain period? I.e. '…annual QM can correct…'. Or 'QM, taking the period as a whole, can correct…'
Thank you for the comment. In Reiter et al. (2018), "QM can correct the complete distribution" refers to applying QM to the calibration period as a whole (i.e., without sub-annual subsampling), which improves the overall distribution but does not correct errors in the annual cycle.
The sentence in line 224 has been changed to: *"As Reiter et al. (2018) pointed out, when applied to the full calibration period as a whole, QM can correct the overall distribution, but it does not correct errors in the annual cycle."*

Section 2.3.2 – Presumably you calculate the two CDFs using the baseline period for 1990-2000 and then apply the mapping function to the whole period of the modelled data? So that the baseline periods have the same relative and absolute distributions and the 2000+ period has the same relative values (i.e. distribution). If so, state that the CDFs are made using only the 1990-2000 period and that mapping is applied to the whole period.
Thank you for the comment. Yes, the empirical CDFs are estimated using the baseline period only, and the same mapping function is then applied to the full model period.
The sentence in line 222 has been revised to: *"In this study, the empirical CDFs for observations and raw CPM are constructed using the baseline period, and the resulting mapping function is applied to the full CPM time series (3 20-year time slices)."*

Does using DQM have a significant effect?
Yes. Diurnal quantile mapping (DQM) has a clear effect on hourly precipitation because it corrects hour-of-day dependent biases and improves the simulated diurnal cycle. This is consistent with previous studies showing that applying bias correction separately for each hour of the day improves the representation of the precipitation diurnal cycle (Faghih et al., 2022). In our results, the bias-corrected ensemble follows the observed diurnal pattern much more closely than the raw UKCP18-CPM output (Figure 5).

Line 255 – not sure I agree with '…although the spatial differentiation between north and south is relatively moderate on an annual basis.' There seems to be a clear difference in north and south biases. Is relatively moderate high or low?
Thank you for the comment. We agree that "relatively moderate" was unclear. We have revised the sentence to avoid this ambiguous wording and to describe the north–south contrast more directly.
The sentence in line 254 has been revised to: *"From a spatial perspective, the annual bias (ANN) shows a consistent pattern across the four ensemble members, with wetter biases in the*

north and drier biases in the south. This north–south contrast is clearer in seasonal results than in the annual mean."

Line 278 – Unsure whether treating EM08 as a separate, instead of part of the same continuous range is odd…

Thank you for the comment. We agree that treating EM08 as a separate case is unnecessary.

The sentence in Lines 277–279 has been revised to: "*For the annual mean temperature (ANN), the ensemble members show a generally cool bias across England relative to HadUK-Grid, with mean biases ranging from -0.87 °C (EM04) to +0.02 °C (EM08). Three out of the four members have negative mean biases.*"

Figure 5 – Where are the thin yellow lines on the plot? These are presumably buried beneath the thicker yellow line. Would it be simpler to remove these and explain that a single line represents all EMs? If not then please state whether the thin yellows are beneath the red of the thick yellow.

Thanks for the advice. We agree that the thin yellow lines are not distinguishable because the individual ensemble member curves overlap closely and are effectively hidden by the thicker ensemble-mean line.

To improve clarity, the figure caption has been updated to: *"Figure 5: Diurnal cycle of mean hourly precipitation before bias correction (blue lines: UKCP18-CPM ensemble) and after bias correction (yellow lines: UKCP18-CPM ensemble BC) for the reference period. The CEH-GEAR1hr (observations) are shown as a red solid line with circles. The ensemble mean before bias correction is shown as a darker blue line (UKCP18-CPM ensemble mean), and the bias-corrected ensemble mean is shown as a darker yellow line (UKCP18-CPM ensemble mean BC). The thin yellow lines (individual members) are overlain by the thick yellow line (ensemble mean), as the diurnal cycles are similar across the four members."*

Line 323 – You say that the 3hr window is used to remove unrealisable observations – is this due to a data accuracy/ error? If so, does the large number of values used in each hourly calculation (10 years x 30 days?) not drown out the impact of these naturally?

Thank you for the comment. Because precipitation is intermittent, the sample within any single hour can be limited. We therefore used a 3 h moving window to increase the sample pool and make the training more robust. This approach helps to stabilise the estimation of the hour-of-day empirical CDFs and the resulting mapping and reduces noisy hour-to-hour variability in the derived correction, particularly for the upper tail.

The sentence in line 323 has been replaced by: "*These discrepancies are mainly related to the 3 h moving window used in the DBC process, which pools data from the target hour and its neighbouring hours to stabilise the hour-of-day correction."*

Figure 6 - What causes banding in Observations? Limit of resolution in recording measurements?

Thank you for the comment. The apparent banding in the observations is likely related to the discrete nature of the gridded hourly precipitation values and the use of a percentile-based metric, rather than to the bias-correction itself. Because precipitation is intermittent, many hourly values are repeated or clustered at a limited set of values, and the estimated P_{95} can therefore also take repeated values across many grid cells. This effect may be further enhanced by the finite precision or rounding of the stored values. As a result, many cells share the same observation value on the x-axis while differing in their raw UKCP18-CPM or bias-corrected values on the y-axis, which appears as vertical banding in the scatter plots.

Line 358 – This 'almost perfect' match should be treated as a mathematical function as a result of the method of mapping the modelled to the observed, rather than a success or proof that it is working well. Similar comment for line 361 – need to be careful using the word 'correct' – it should be used to mean 'adjust' rather than 'right'.

Thank you for the comment. We have removed the relevant wording to avoid over-interpreting. We have also replaced "correct" with "adjust" where appropriate in the text, to avoid implying that the outputs are "right".

The sentences around lines 357–362 have been revised to: "*After bias correction, the fitted slope and R^2 values for all four ensemble members are close to 1 in the reference period. Figure 8 complements this analysis by displaying violin plots of the distributions of T_{95} and T_5 for the observations (black), raw UKCP18-CPM (light pink), and bias-adjusted UKCP18-CPM (dark pink). The plots for T_{95} (left side) and T_5 (right side) 360 clearly show that bias correction not only adjusts the central tendencies of the temperature extremes but also adjusts their distribution.*"

Figures 10 and 11 – Colour scale is difficult to use – figures are interpreted as 'red' – consider adding another colour into the colour bar to differentiate between the temperature increases. Or consider cropping colours to 0-10 degrees if there is no cooling.

Thank you for the comment. We revised Figures 10 and 11 by limiting the colour bar range to 0–10 °C, because the future changes show warming only (no cooling). In addition, we added Figure 12 as a heatmap to show the projected changes in the 2030s and 2070s for each ensemble member, so that the values can be read more clearly than from the maps alone. The updated figures are as follows.

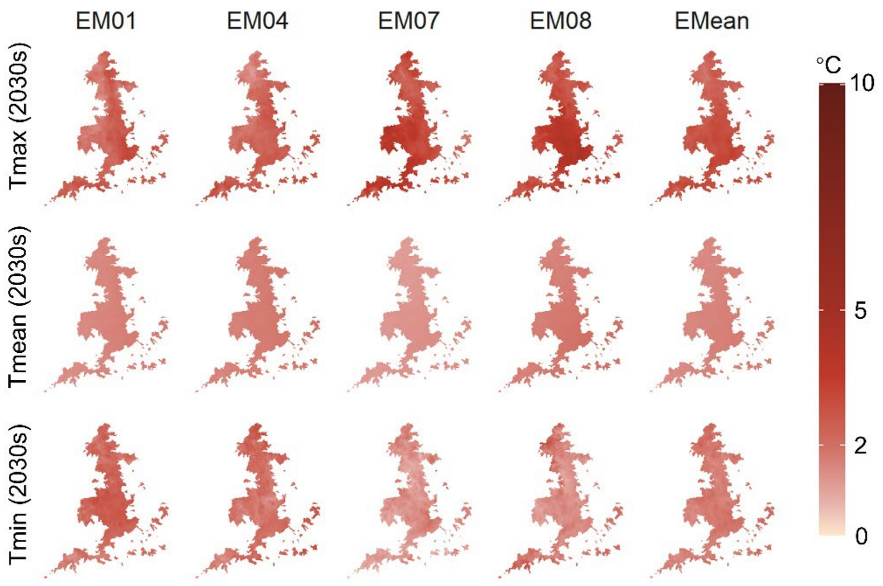

Figure 10: Projected changes (°C) in the annual maximum (Tmax), annual mean (Tmean), and annual minimum (Tmin) of daily mean temperature from bias-corrected UKCP18-CPM

simulations. Results are shown for the 2030s relative to the baseline period, for each processed grid cell within the analysis mask. Columns show individual ensemble members (EM01, EM04, EM07, and EM08) and the ensemble mean (EMean).

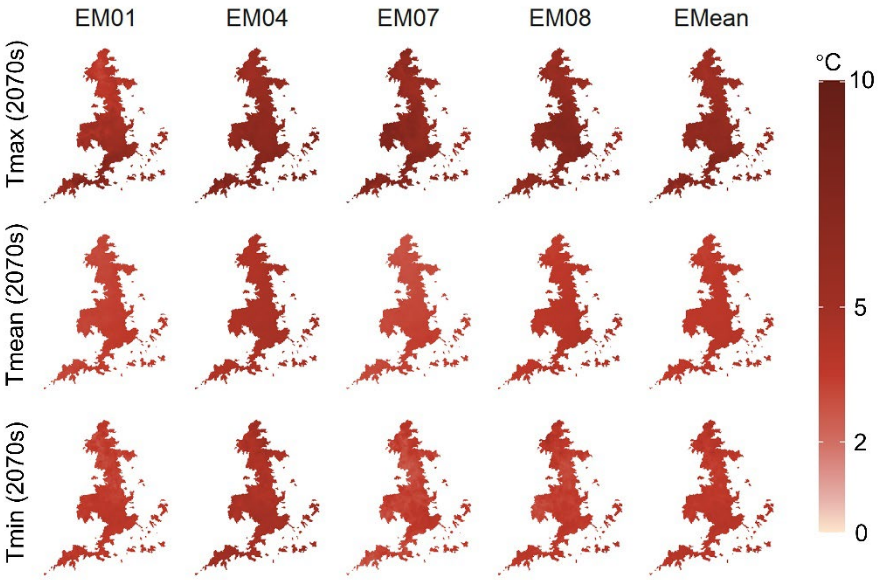

Figure 11: Projected changes (°C) in the annual maximum (Tmax), annual mean (Tmean), and annual minimum (Tmin) of daily mean temperature from bias-corrected UKCP18-CPM simulations. Results are shown for the 2070s relative to the baseline period, for each processed grid cell within the analysis mask. Columns show individual ensemble members (EM01, EM04, EM07, and EM08) and the ensemble mean (EMean).

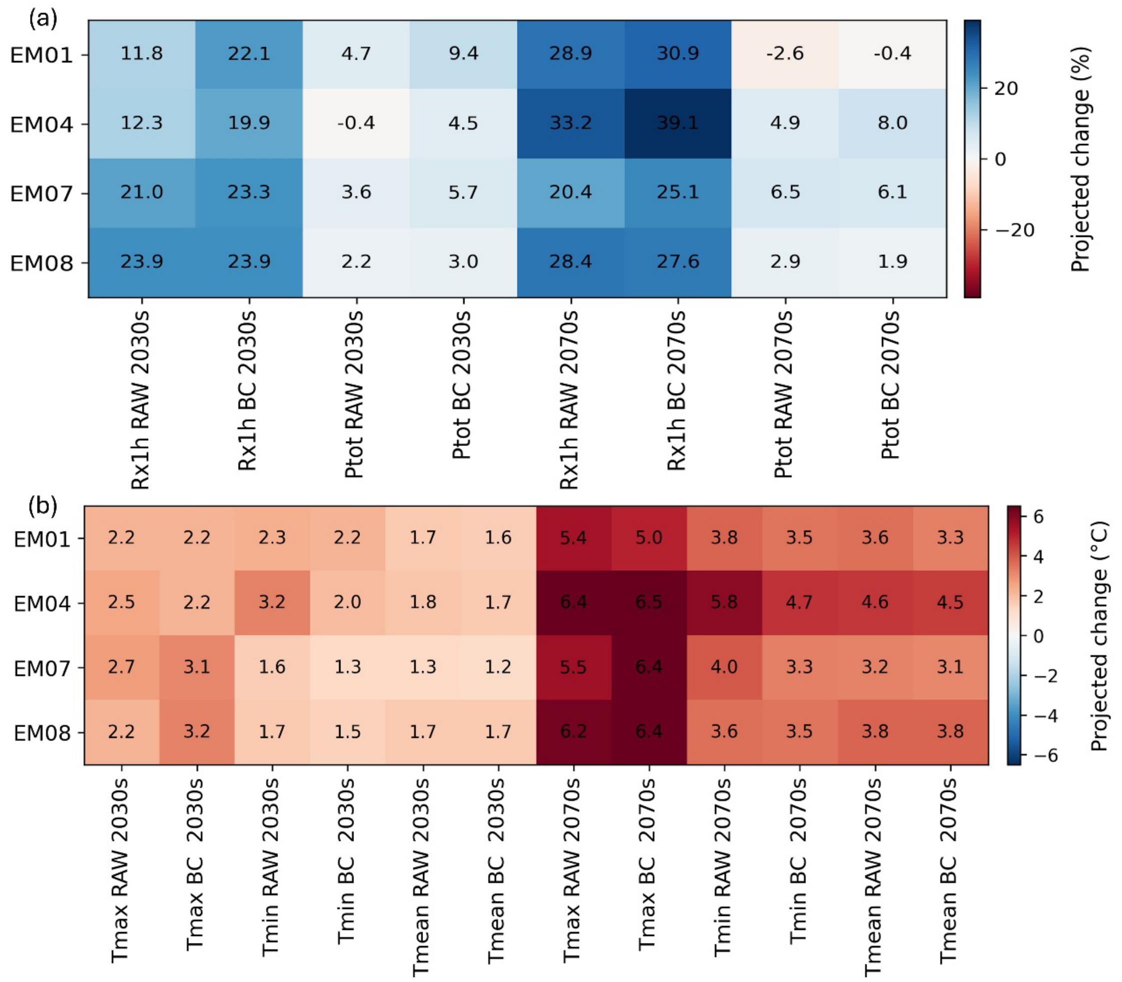

Figure 12: Projected changes in (a) annual maximum 1 h precipitation (Rx1h) and annual total precipitation (Ptot), and (b) annual maximum (Tmax), annual mean (Tmean), and annual minimum (Tmin) of daily mean temperature for UKCP18-CPM simulations before (RAW) and after bias correction (BC). Each value is the spatial average over the processed grid cells, shown for the 2030s and 2070s relative to the baseline period.

Line 469 – If there was a seasonal pattern in temperature bias then please include that or state otherwise earlier in the text.

Thanks for the comment. There is a seasonal pattern in temperature bias. We have added a brief statement at the beginning of the temperature bias in Results to summarise the main differences between DJF and JJA.

The sentence in lines 468-469 has been revised to: "*Temperature biases show a clear seasonal dependence, with winter (DJF) exhibiting larger spatial contrasts and summer (JJA) showing a more consistent cool bias across all members. On an annual mean basis, the temperature bias ranged from −0.87 °C (EM04) to +0.02 °C (EM08), with three of the four members showing a cool bias.*"

Conclusion point 2 – as mentioned earlier – the bias corrected data may be 'better' and more fit for purpose than the original, however you are treading the line of making it sounds as if the alignment of BC data to the observations is proof that this is the case, when this is simply a direct result of remapping the UKCP18 data. I.e. Doubling a UKCP18 temperature 5°C and matching the observation of 10°C does not prove that the bias controlled data is better, only that the process is working as it should. Consider rephrasing and the appropriateness of words/phrases like 'marked improvement' and 'corrected' rather than 'bias corrected'. Perhaps 'After bias correction, both precipitation and temperature simulations aligned well with observational data. Monthly precipitation and temperature biases were substantially reduced, with bias corrected outputs closely following observed monthly patterns. The diurnal cycle of bias corrected precipitation captured the variability and magnitude of observed precipitation across most hours of the day. The 95th percentile (P95) of hourly precipitation and temperature extremes (T95) matched observational datasets after bias correction demonstrating that the method had been correctly applied across the dataset, including the tails.'

Thank you for the comment. We agree and have revised Conclusion point 2 accordingly. The text has been rephrased to avoid overstating the close agreement between the bias-adjusted data and the observations as proof of improved performance, and to describe it more appropriately as the expected outcome of the bias-correction procedure. We have also replaced potentially overstated wording where appropriate.

The revised Conclusion point 2 now reads: "*After bias correction, both precipitation and temperature simulations aligned well with observational data. Monthly precipitation and temperature biases were substantially reduced, with bias corrected outputs closely following observed monthly patterns. The diurnal cycle of bias corrected precipitation captured the variability and magnitude of observed precipitation across most hours of the day. The 95th percentile (P_{95}) of hourly precipitation and temperature extremes (T_{95}) showed markedly improved agreement with observations after bias correction for both mean behaviour and high-percentile extremes (P95 and T95).*"

The following sentences have been added to line 61 in Introduction: "*In this paper, we use the term bias correction (BC) following common usage in the climate impact literature. However, these methods do not remove model biases in a physical sense; rather, they statistically adjust model outputs to better match an observational reference over a calibration period. We adopt this interpretation when discussing our results.*"

Discussion – If possible, I would like to see a little more detail on how this compares with other bias corrections, even if these are the RCMs rather than CPMs. The content on comparison is limited other than saying that there is general agreement with Robinson (2023) and Kendon (2019b). It would be interesting to hear whether the similarities are in the patterns, or the values themselves and if there are any differences and why this may be. That said, it is difficult to compare between datasets and so this may be difficult to do in depth or in a fully quantitative manner.

Thank you for this helpful suggestion. We agree that the previous Discussion provided only a limited qualitative comparison with earlier bias-correction studies and did not clearly indicate whether the agreement referred mainly to spatial patterns or to the magnitude of the results. We have expanded the Discussion to place our results in a broader context by comparing the main precipitation and temperature bias characteristics, as well as the projected future changes, with findings from previous studies. In particular, we now make clearer where the agreement is mainly in the overall patterns and direction of change and provide additional supporting references from related studies.

The paragraph in Discussion (lines 413-422) has been revised to: *"Before bias correction, the UKCP18-CPM simulations show considerable biases in both precipitation and temperature across various temporal and spatial scales. In general, UKCP18-CPM simulations show wet precipitation biases (especially in winter) and cool biases in temperature. These broad patterns are consistent with previous evaluations of UKCP18-RCM simulations (Reyniers et al., 2025) and the UKCP18-CPM science report (Kendon et al., 2019b). These residual biases are consistent with the broader CPM literature, which shows that although kilometre-scale models can improve the representation of precipitation compared with convection-parameterised RCMs, systematic biases and model uncertainty may remain, particularly for sub-daily precipitation and across different regional settings (Ban et al., 2021; Correa-Sánchez et al., 2025; Soares et al., 2024). After bias correction, both precipitation and temperature show substantially improved agreement with observations at monthly and diurnal timescales (Figs. 4 and 5). This suggests that bias correction provides a more reliable basis for downstream impact assessments than raw UKCP18-CPM output."*
The revised paragraph in lines 423-434 reads as follows: *"Future changes were also analysed for both precipitation and temperature in Figs. 9–12. In general, the bias-corrected projections indicate a clear intensification of annual maximum 1 h precipitation (Rx1h) by both the 2030s and 2070s, while changes in annual total precipitation (Ptot) remain smaller and spatially variable (Figs. 9 and 12). This pattern (i.e., stronger changes in short-duration extremes than in totals) is broadly consistent with CPM-based studies that report amplified changes for sub-daily extremes and a strong dependence on dynamical factors and regional setting (Dallan et al., 2024; Pichelli et al., 2021). The projected temperature changes are consistent with findings from the UKCP18-CPM science report (Kendon et al., 2019b), which indicated that mean temperature is expected to increase across all regions and seasons. Our findings also align well with existing literature, such as Robinson et al. (2023), who reported similar projected changes for UKCP18-RCM. This suggests that the bias-corrected projections remain broadly consistent with previously reported climate change signals.*"

Discussion – The paper ticks off all of the expected content apart from a comparison with unbiased corrected futures. In my view this would demonstrate the need / impact of the bias correction and demonstrate to others that they need to employ (or at least consider) methods such as this before using the raw data. This would also make the paper more applicable/relevant and increase its impact. The difference in raw/BC data is discussed with regard to the reference period, however this is the simple part as, simply due to the method, we expect the BC data to align with the observations. I would like to see a few plots of future rainfall/temperature with and without BC. Perhaps an annual timeseries for the 2035 or 2075 showing the difference, or a seasonal plot of the impact as this has clear consequences for the hydrological/agricultureal uses mentioned in the paper. Seasonal differences were discussed in the text in relation to biases in the reference period, but a direct future example would be more impactful. The referenced Robinson et al. (2023) shows seasonal breakdowns, clearly demonstrating the seasonality of the impact of bias control. Another approach could be to add a column to the mapped figures with EMean Raw, showing the un-bias corrected data.
Thank you for the suggestion. We agree that it is useful to compare future changes from the raw CPM simulations with those from the bias-corrected simulations, to illustrate the impact of bias correction on projected changes. We have therefore added Figure 12 (heatmaps) to

summarise projected changes for both raw and bias-corrected simulations in the 2030s and 2070s.

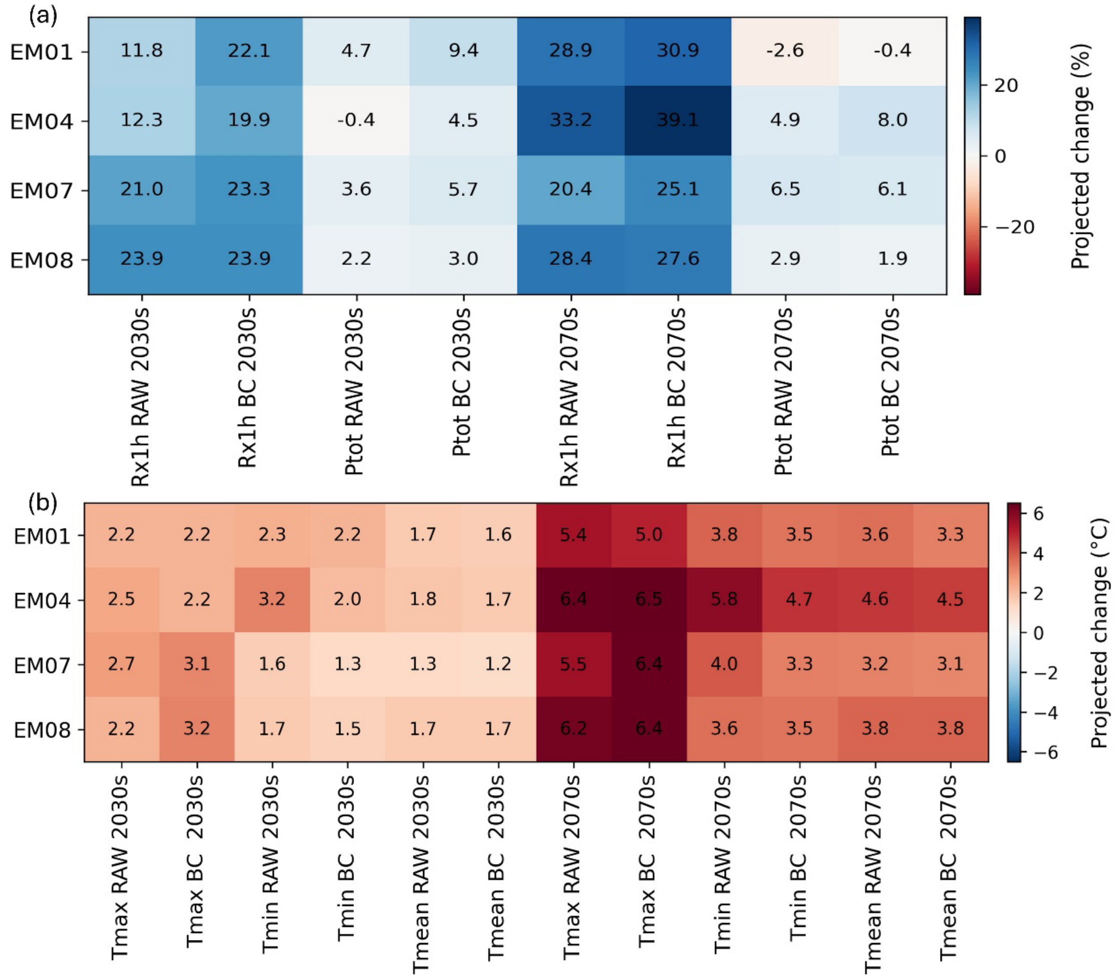

Figure 12: Projected changes in (a) annual maximum 1 h precipitation (Rx1h) and annual total precipitation (Ptot), and (b) annual maximum (Tmax), annual mean (Tmean), and annual minimum (Tmin) of daily mean temperature for UKCP18-CPM simulations before (RAW) and after bias correction (BC). Each value is the spatial average over the processed grid cells, shown for the 2030s and 2070s relative to the baseline period.

We revised the Discussion section on future changes (lines 423–434): *"Future changes were also analysed for both precipitation and temperature in Figs. 9–12. In general, the bias-corrected projections indicate a clear intensification of annual maximum 1 h precipitation (Rx1h) by both the 2030s and 2070s, while changes in annual total precipitation (Ptot) remain smaller and spatially variable (Figs. 9 and 12). This pattern (i.e., stronger changes in short-duration extremes than in totals) is broadly consistent with CPM-based studies that report amplified changes for sub-daily extremes and a strong dependence on dynamical factors and regional setting (Dallan et al., 2024; Pichelli et al., 2021). The projected temperature changes are consistent with findings from the UKCP18-CPM science report (Kendon et al., 2019b), which indicated that mean temperature is expected to increase across all regions and seasons. Our findings also align well with existing literature, such as Robinson et al. (2023), who* *reported similar projected changes for UKCP18-RCM. This suggests that the bias-corrected projections remain broadly consistent with previously reported climate change signals."*

Discussion on the limitations of BC in general would be benificial if possible. I.e. does bias correction detract from the point of using a climate model in the first place?
Thank you for this helpful suggestion. We do not consider bias correction to detract from the purpose of using a climate model. Rather, it is a statistical post-processing step that can improve the usability of climate model outputs for impact applications by reducing systematic biases relative to an observational reference. We also now talked more about the BC method we use.
The following paragraph has been added after the first paragraph in Discussion (line 413):
"We applied empirical QM to daily mean temperature to address seasonal biases and distributional differences using a reproducible procedure. It is widely used because it can correct not only mean biases but also biases in variability and quantiles (Fang et al., 2015; Themeßl et al., 2011; Wilcke et al., 2013). For hourly precipitation, we used DBC by hour of day, because applying a single mapping across all hours can overlook systematic diurnal-cycle biases in sub-daily precipitation. Figure 5 supports this choice, showing that the bias-corrected ensemble more closely reproduces the observed hour-of-day pattern. This hour-of-day treatment is intended to reduce discrepancies that matter for sub-daily impact applications, consistent with the broader argument that correcting the diurnal cycle can improve sub-daily bias-correction strategies (Faghih et al., 2022). Therefore, we recommend using DBC for sub-daily variables. However, in this study we applied a 3 h moving window in the DBC. By incorporating hourly data from the previous and following hours, this may have contributed to some of the remaining discrepancies. It may be worth exploring whether omitting the moving window gives better results when the training period sample size is sufficiently large."

Technical Corrections:

- Line 33 – '…provide the…'
  Corrected.
  "The UK Climate Projections 2018 (UKCP18) provide the latest generation of national climate projections for climate impact studies in the UK (Murphy et al., 2018)."

- Line 42 – What has been widely applied? CMPs (presumably not UKCP18-CPM)
  Thank you for the comment. We agree that the subject was unclear. We have revised the text to make clear that the sentence refers to convection-permitting models (CPMs) in general, rather than specifically to UKCP18-CPM.
  The sentences in lines 39-42 have been revised to "*Convection-permitting models (CPMs) have become a valuable tool in short-range weather forecasting due to their enhanced ability to represent convective processes in detail, improve forecast accuracy, and capture localised high-impact rainfall that coarser models often miss (e.g., Done et al., 2004; Lean et al., 2008; Roberts and Lean, 2008; Weisman et al., 2008; Weusthoff et al., 2010). CPMs have also been widely applied in Europe (Berthou et al., 2020; Pichelli et al., 2021), Asia (Murata et al., 2017; Yun et al., 2020), and Africa (Kendon et al., 2019a; Maurer et al., 2017), as well as other regions (Trapp et al., 2011), to*

enhance the representation of convective processes and improve the accuracy of climate projections."

- Line 54 – 'while in summer it shows spatial variability, being too wet in the north and too dry in the south'. Would delete 'shows spatial variability', as that is good, being too wet/dry is the issue.
  Revised as suggested.
  Line 54: *"…while in summer it is too wet in the north and too dry in the south (Kendon et al., 2021)"*

- Line 75 – 'applies' -> 'applied' … 'cannot' -> 'could not'
  Corrected.
  "However, the scaling method applied a uniform adjustment across all precipitation values and could not account for the distribution."

- Line 78 – Comma before 'which'.
  Corrected.
  "As a result, it may not accurately correct biases in extreme precipitation events. It also fails to correct the diurnal cycle of precipitation, which is a key feature of high-resolution climate models (Bannister et al., 2019; Scaff et al., 2019)."

- Line 83 – Incorrect '.' before integrates.
  Removed the '.'.

- Line 88 – Faghih reference seems slightly out of place – perhaps better with mention of diurnal cycle in paragraph above.
  Thank you for the comment. We moved this sentence to line 78 and revised it as follows: *"This is consistent with Faghih et al. (2022), who showed that accounting for the diurnal cycle in bias correction can improve the representation of sub-daily precipitation and temperature."*

- Line 96 – is the driest member EM01 or EM04? Unclear from list arrangement.
  Clarified the member classification.
  Line 96: *"…including the default configuration EM01 (also the driest member), the wettest members (EM04 and EM07), and a moderate member (EM08)."*

- Line 108 – Confusingly long bracketed sentence, consider making it its own sentence. Split into two sentences for clarity.
  The sentence has been revised to: *"The catchments were selected based on the availability of complete gauged records for December 1990 to November 2000, overlapping with both the CEH-GEAR1hr observational dataset and the UKCP18-CPM baseline period (see Sect. 2.2.2; hereafter referred to as the reference period), and on minimal anthropogenic influence."*

- Line 109 – Typo in 'datasetshereafter'.

We revised the sentence, as described in our response to the comment on manuscript line 108 under Technical Corrections above.

- Line 112 – Unsure of purpose of giving mean annual rainfall/temp.
  Thank you for the comment. These values were included to provide a brief climatic context for the selected grid cells, giving readers a general indication of the mean hydroclimatic conditions of the analysis domain.

- Figure 1 – Reference data sources and resolution for elevation and catchment outlines. Consider rounding Elevation upper bound to 1000m.
  Thank you for the comment. As the elevation map is not relevant to this study, it was not used to define either the analysis mask or the catchment selection, and we have therefore removed it. We have also added the processed 1 km grid cells (shown by their centre points) used in this study and removed the gauges, as they were not necessary for explaining how the analysis mask was defined. The sentence describing the Figure 1 in Section 2.1 has also been revised, as explained in our response to the second General Comments above.

- Line 122 - Duplicated? 'Strand 3 is a new perturbed parameter ensemble (PPE) of regional climate model (RCM) projections. It introduces a new ensemble consisting of 12 RCM simulations, which form a PPE of RCM variants derived from 12 out of the 15 GC3.05-PPE simulations (Murphy et al., 2018).' à Strand 3 is a new perturbed parameter ensemble (PPE) of 12 regional climate model (RCM) projections derived from 12 of the 15 GC3.05-PPE simulations (Murphy et al., 2018).
  Thank you for the comment. We agree that the original wording was repetitive and have revised the sentence for clarity.
  The text at line 122 now reads: *"Strand 3 is a perturbed parameter ensemble (PPE) of 12 regional climate model (RCM) projections derived from 12 of the 15 GC3.05-PPE simulations (Murphy et al., 2018)."*

- Line 163 – Sect. should presumably be 'Section' as in the text, not brackets.
  Thank you for this comment. We have checked the HESS style guidance, which states that "Sect." should be used in running text and followed by a number unless it appears at the beginning of a sentence. We have therefore retained "Sect. 2.2.1" here to remain consistent with the journal style.

- Line 174 – are the sizes correct? 576 GB of inputs to only 57 GB of outputs seems unlikely…
  Thank you for the comment. 576 GB inputs refer to the 1 km UKCP18-CPM data obtained by using the Nearest Neighbour Interpolation used. The result is a compressed version. I have removed this part of the content, as the level of detail was excessive for the main text. We have therefore substantially shortened this paragraph, kept only a brief justification and removed the specific runtime and storage numbers.
  The sentences in lines 169-176 have been replaced with: *"The UKCP18-CPM consists of 12 projections driven by the 12 km RCM ensemble. Due to computational*

and storage constraints, it was not practical to bias-correct the full CPM-PPE at 1 km hourly resolution. We therefore selected a subset of ensemble members for bias correction and implemented a parallel workflow on the University of East Anglia's high-performance computing (HPC) system to complete the processing within a reasonable time. Four ensemble members were selected to represent the broad range of possible precipitation outcomes within the ensemble."

- Line 246 – Should say Figure 2, not 2a. Should say '…and the change in the number of events exceeding in the UKCP18 data relative to the CEH-GEAR data'.
  Thank you for the comment. We describe figure 2a and figure 2b separately.
  The text now reads: "*Figure 2a shows the spatial distribution of relative biases (%) in mean hourly precipitation in the raw UKCP18-CPM simulations compared with CEH-GEAR1hr for the reference period. Figure 2b shows the change in the number of hourly rainfall events exceeding 20 mm h^{-1} in UKCP18-CPM relative to CEH-GEAR1hr over the reference period. The 20 mm h^{-1} threshold is commonly used as an indicator of potential flash flood producing rainfall in the UK (e.g., Kendon et al., 2023)."*

- Figure 2 may benefit from actual figure titles rather than equations. I.e. Percentage change in precipitation bias (%).
  Thank you for the suggestion. We have removed the equations from the figure and revised the caption accordingly.
  The caption of Figure 2 now reads: *"Figure 2: Spatial distribution of (a) relative biases (%) in mean hourly precipitation between UKCP18-CPM and CEH-GEAR1hr and (b) differences in the number of events exceeding 20 mm h^{-1} between UKCP18-CPM and CEH-GEAR1hr during the reference period across England. Rows represent annual (ANN), winter (DJF: December, January, February), and summer (JJA: June, July, August) results, while columns show individual ensemble members (EM01, EM04, EM07, EM08) and the ensemble mean (EMean).*

- Figure 2b – The duration over which the sum is calculated should be emphasised – you do state that this is for the reference period, but I missed this on first reading and it could be taken as annual values in this period by a reader flicking through. Consider stating '…GEAR1hr over the 10-year reference period…'.
  Line 246 has been revised to: *"Figure 2a shows the spatial distribution of relative biases (%) in mean hourly precipitation in the raw UKCP18-CPM simulations compared with CEH-GEAR1hr for the reference period. Figure 2b shows the change in the number of hourly rainfall events exceeding 20 mm h^{-1} in UKCP18-CPM relative to CEH-GEAR1hr over the reference period. The 20 mm h^{-1} threshold is commonly used as an indicator of potential flash flood producing rainfall in the UK (e.g., Kendon et al., 2023)."*

- Figure 2 – A thin GB outline would make the figure geography clearer, if it does not make the maps to cluttered.
  Thank you for the suggestion. We agree that a Great Britain outline could provide additional geographical context. However, as Figure 1 already shows the study domain

within the wider GB context, we felt that adding a GB outline to every panel in Figure 2 would make the figure more cluttered while providing only limited additional information. We have therefore kept the current layout for clarity.

- Figure 3 – Same comments as for figure 2.
  Thank you for the suggestion. We have responded in the same way as for Figure 2 above.

- I do not think there should be a space between a temperature and the degrees symbol.
  Thank you for the comment. We have checked the formatting of temperature units and retained spacing between the number and °C in accordance with the journal style.

- Line 321 – Should 'darker' be 'thicker'? They look to be the same colour, but different thicknesses.
  Thank you for the comment. We agree that "thicker" is the more appropriate description here, and we have revised the wording accordingly.
  It now reads: *"… (thicker yellow line) closely aligning…"*.

- Figure 6 – Backwards order of data – I would present chronologically (black, light blue, dark blue from left to right). Minor point. Would also mean that order in figure matches order in caption. And order in Fig 8.
  Thank you for the suggestion. We have reordered the violin plots in the second row of Figure 6 to follow the same order as the legend (black, light blue, dark blue).

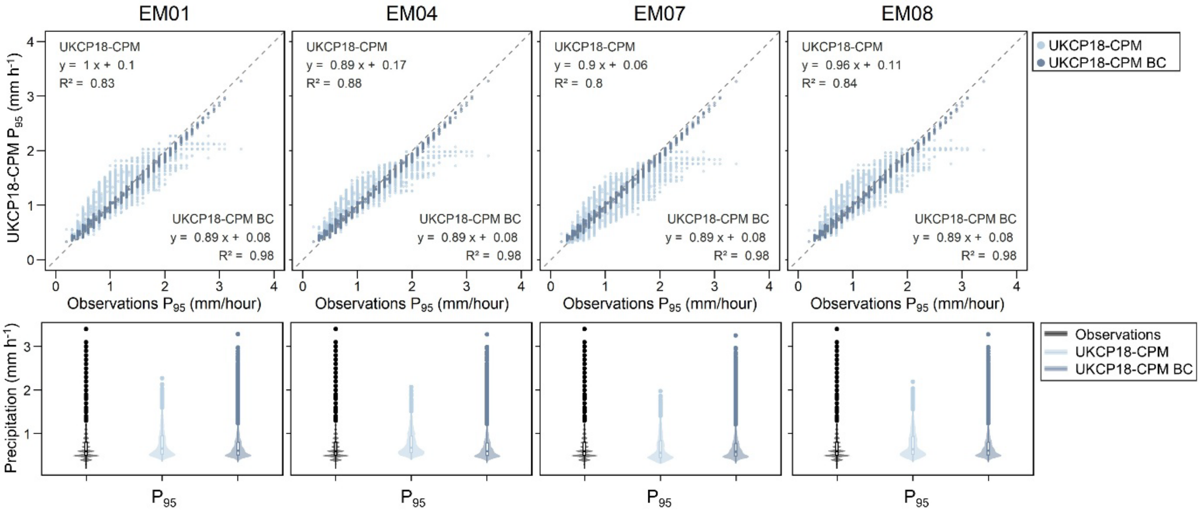

Figure 6: Comparison of 95th percentile (P_{95}) hourly precipitation values for the reference period. The top row shows scatter plots of UKCP18-CPM P_{95} values for each ensemble member, plotted against CEH-GEAR1hr observations for each 1 km grid, with raw (light blue) and bias-corrected (dark blue) data. The bottom row presents violin plots of P_{95} values for observations (black), raw UKCP18-CPM (light blue), and bias-corrected (dark blue) data.

- Line 350 – Does 5th percentile use all values beneath this threshold, or just those on it? I was presuming that 95th percentile used all values on and above that threshold…

Here, the 5th and 95th percentiles refer to the percentile values themselves, used as summary measures of the lower and upper tails of the distribution, rather than to all values below or above those thresholds. We added a clarifying sentence at the first mention of the percentile.

A sentence has been added to line 334: *"Here, the percentile refers to the percentile value itself (i.e., the threshold value), rather than to all values above or below that threshold."*

- Figure 7/8 – 'bias-corrected UKCP18-CPM BC' does not need 'BC' to match other figures.

  Thank you for this suggestion. We have updated the caption of both Figures 7 and 8.

  Now the caption of Figure 7 reads: *"Figure 7: The 95th (T_{95}, top row) and 5th (T_5, bottom row) percentile values of daily mean temperature for each ensemble member, plotted against HadUK-Grid (observations). Raw UKCP18-CPM (light pink) and bias-corrected UKCP18-CPM (dark pink) values are compared for the reference period."*

  Now the caption of Figure 8 reads: "*Figure 8: Violin plots of T_{95} and T_5 values for observations (HadUK-Grid, black), raw UKCP18-CPM (light pink), and bias-corrected UKCP18-CPM (dark pink) for each ensemble member during the reference period."*

- Figure 9 – would be clearer to mention Annual Total and Annual Maximum in caption so that it is clear what abbreviations mean in plot.

  Thank you for this suggestion. We agree that the caption should explain these abbreviations more clearly. To avoid confusion between annual maximum 1 h precipitation and annual maximum temperature, we have standardised these terms throughout the manuscript. We now use annual maximum 1 h precipitation (Rx1h), annual total precipitation (Ptot), annual maximum temperature (Tmax), annual mean temperature (Tmean), and annual minimum temperature (Tmin). All related abbreviations in the figures, captions, and main text have been updated accordingly.

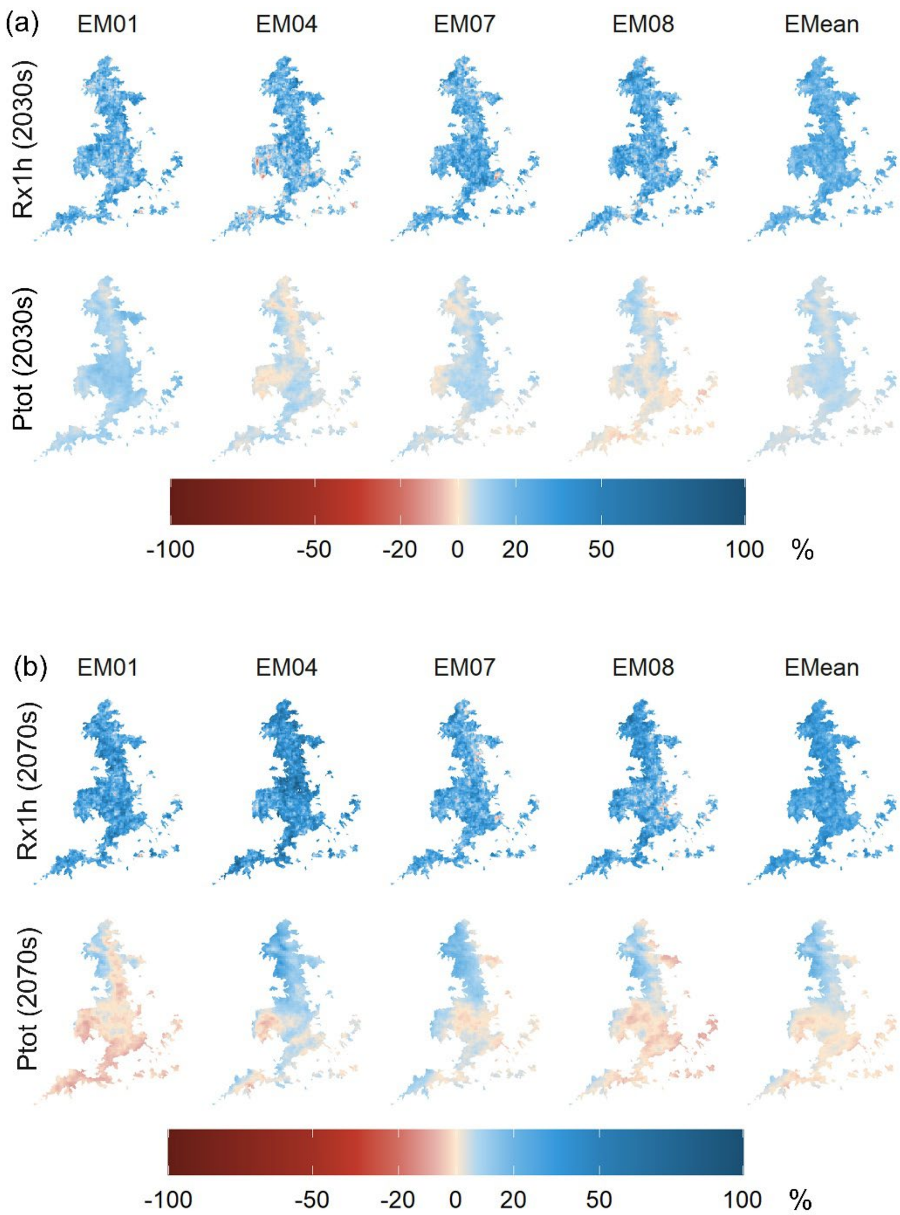

Figure 9: Projected percentage changes in annual maximum 1 h precipitation (Rx1h) and annual total precipitation (Ptot) from bias-corrected UKCP18-CPM simulations for the (a) 2030s and (b) 2070s, compared with the baseline period.

- Line 378 – 'mostly positive' sounds like a good thing, when it is meant as an increase.
  Thank you for the suggestion.
  The sentence now reads: *"Ptot also increases, but the changes are less pronounced, with values of 3% (EM08) and 9.4% (EM01)."*

- Line 397 – 'minimum,by' has missing space.
  Corrected typo, and now it reads: *"…i.e., maximum, mean, and minimum, by the 2070s."*

- Line 454 – Future research… does not read correctly. Remove 'that' from 'that would'?
  Thank you for the comment. We have revised the sentence accordingly.

Now it reads: *"Future research could extend this analysis to other regions and incorporate additional climate models, providing a broader understanding of the impacts of climate change and helping to assess the applicability of these methods in different contexts."*

- Line 477 – 'increasesin' has missing space.
  Corrected typo. *"…increases in both precipitation and temperature."*

References

Ban, N., Caillaud, C., Coppola, E., Pichelli, E., Sobolowski, S., Adinolfi, M., Ahrens, B., Alias, A., Anders, I., Bastin, S., Belušić, D., Berthou, S., Brisson, E., Cardoso, R. M., Chan, S. C., Christensen, O. B., Fernández, J., Fita, L., Frisius, T., Gašparac, G., Giorgi, F., Goergen, K., Haugen, J. E., Hodnebrog, Ø., Kartsios, S., Katragkou, E., Kendon, E. J., Keuler, K., Lavin-Gullon, A., Lenderink, G., Leutwyler, D., Lorenz, T., Maraun, D., Mercogliano, P., Milovac, J., Panitz, H.-J., Raffa, M., Remedio, A. R., Schär, C., Soares, P. M. M., Srnec, L., Steensen, B. M., Stocchi, P., Tölle, M. H., Truhetz, H., Vergara-Temprado, J., de Vries, H., Warrach-Sagi, K., Wulfmeyer, V., and Zander, M. J.: The first multi-model ensemble of regional climate simulations at kilometer-scale resolution, part I: evaluation of precipitation, Clim Dyn, 57, 275–302, https://doi.org/10.1007/s00382-021-05708-w, 2021.

Bannister, D., Orr, A., Jain, S. K., Holman, I. P., Momblanch, A., Phillips, T., Adeloye, A. J., Snapir, B., Waine, T. W., Hosking, J. S., and Allen-Sader, C.: Bias Correction of High-Resolution Regional Climate Model Precipitation Output Gives the Best Estimates of Precipitation in Himalayan Catchments, Journal of Geophysical Research: Atmospheres, 124, 14220–14239, https://doi.org/10.1029/2019JD030804, 2019.

Berthou, S., Kendon, E. J., Chan, S. C., Ban, N., Leutwyler, D., Schär, C., and Fosser, G.: Pan-European climate at convection-permitting scale: a model intercomparison study, Clim Dyn, 55, 35–59, https://doi.org/10.1007/s00382-018-4114-6, 2020.

Correa-Sánchez, N., Dallan, E., Marra, F., Fosser, G., and Borga, M.: Orographic control on bias and uncertainty in extreme sub-daily precipitation simulations from a convection-permitting ensemble, Journal of Hydrology, 659, 133324, https://doi.org/10.1016/j.jhydrol.2025.133324, 2025.

Dallan, E., Marra, F., Fosser, G., Marani, M., and Borga, M.: Dynamical Factors Heavily Modulate the Future Increase of Sub-Daily Extreme Precipitation in the Alpine-Mediterranean Region, Earth's Future, 12, e2024EF005185, https://doi.org/10.1029/2024EF005185, 2024.

Done, J., Davis, C. A., and Weisman, M.: The next generation of NWP: explicit forecasts of convection using the weather research and forecasting (WRF) model, Atmosph. Sci. Lett., 5, 110–117, https://doi.org/10.1002/asl.72, 2004.

Faghih, M., Brissette, F., and Sabeti, P.: Impact of correcting sub-daily climate model biases for hydrological studies, Hydrology and Earth System Sciences, 26, 1545–1563, https://doi.org/10.5194/hess-26-1545-2022, 2022.

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.

Hollis, D., McCarthy, M., Kendon, M., Legg, T., and Simpson, I.: HadUK-Grid—A new UK dataset of gridded climate observations, Geoscience Data Journal, 6, 151–159, https://doi.org/10.1002/gdj3.78, 2019.

Kendon, E. J., Stratton, R. A., Tucker, S., Marsham, J. H., Berthou, S., Rowell, D. P., and Senior, C. A.: Enhanced future changes in wet and dry extremes over Africa at convection-permitting scale, Nat Commun, 10, 1794, https://doi.org/10.1038/s41467-019-09776-9, 2019a.

Kendon, E. J., Fosser, G., Murphy, J., Chan, S., Clark, R., Harris, G., Lock, A., Lowe, J., Martin, G., Pirret, J., Roberts, N., Sanderson, M., and Tucker, S.: UKCP Convection-permitting model projections: Science report, 2019b.

Kendon, E. J., Prein, A. F., Senior, C. A., and Stirling, A.: Challenges and Outlook for Convection-Permitting Climate Modelling, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 379, 20190547, https://doi.org/10.1098/rsta.2019.0547, 2021.

Kendon, E. J., Fischer, E. M., and Short, C. J.: Variability conceals emerging trend in 100yr projections of UK local hourly rainfall extremes, Nat Commun, 14, 1133, https://doi.org/10.1038/s41467-023-36499-9, 2023.

Lean, H. W., Clark, P. A., Dixon, M., Roberts, N. M., Fitch, A., Forbes, R., and Halliwell, C.: Characteristics of High-Resolution Versions of the Met Office Unified Model for Forecasting Convection over the United Kingdom, Mon. Wea. Rev., 136, 3408–3424, https://doi.org/10.1175/2008MWR2332.1, 2008.

Lewis, E., Quinn, N., Blenkinsop, S., Fowler, H. J., Freer, J., Tanguy, M., Hitt, O., Coxon, G., Bates, P., Woods, R., Fry, M., Chevuturi, A., Swain, O., and White, S. M.: Gridded estimates of hourly areal rainfall for Great Britain 1990-2016 [CEH-GEAR1hr] v2, 2022.

Maurer, V., Bischoff-Gauß, I., Kalthoff, N., Gantner, L., Roca, R., and Panitz, H.-J.: Initiation of deep convection in the Sahel in a convection-permitting climate simulation for northern Africa, Quarterly Journal of the Royal Meteorological Society, 143, 806–816, https://doi.org/10.1002/qj.2966, 2017.

Met Office, Hollis, D., McCarthy, M., Kendon, M., and Legg, T.: HadUK-Grid Gridded Climate Observations on a 1km grid over the UK, v1.1.0.0 (1836-2021), https://doi.org/10.5285/BBCA3267DC7D4219AF484976734C9527, 2022.

Murata, A., Sasaki, H., Kawase, H., and Nosaka, M.: Evaluation of precipitation over an oceanic region of Japan in convection-permitting regional climate model simulations, Clim Dyn, 48, 1779–1792, https://doi.org/10.1007/s00382-016-3172-x, 2017.

Murphy, J. M., Harris, G. R., Sexton, D. M. H., Kendon, E. J., Bett, P. E., Clark, R. T., Eagle, K. E., Fosser, G., Fung, F., Lowe, J. A., McDonald, R. E., McInnes, R. N., McSweeney, C. F., Mitchell, J. F. B., Rostron, J. W., Thornton, H. E., Tucker, S., and Yamazaki, K.: UKCP18 Land Projections: Science Report, Met Office Hadley Centre, UK, 2018.

Pichelli, E., Coppola, E., Sobolowski, S., Ban, N., Giorgi, F., Stocchi, P., Alias, A., Belušić, D., Berthou, S., Caillaud, C., Cardoso, R. M., Chan, S., Christensen, O. B., Dobler, A., de Vries, H., Goergen, K., Kendon, E. J., Keuler, K., Lenderink, G., Lorenz, T., Mishra, A. N., Panitz, H.-J., Schär, C., Soares, P. M. M., Truhetz, H., and Vergara-Temprado, J.: The first multi-model ensemble of regional climate simulations at kilometer-scale resolution part 2: historical and future simulations of precipitation, Clim Dyn, 56, 3581–3602, https://doi.org/10.1007/s00382-021-05657-4, 2021.

Reiter, P., Gutjahr, O., Schefczyk, L., Heinemann, G., and Casper, M.: Does applying quantile mapping to subsamples improve the bias correction of daily precipitation?, International Journal of Climatology, 38, 1623–1633, https://doi.org/10.1002/joc.5283, 2018.

Reyniers, N., Zha, Q., Addor, N., Osborn, T. J., Forstenhäusler, N., and He, Y.: Two sets of bias-corrected regional UK Climate Projections 2018 (UKCP18) of temperature, precipitation and potential evapotranspiration for Great Britain, Earth System Science Data, 17, 2113–2133, https://doi.org/10.5194/essd-17-2113-2025, 2025.

Roberts, N. M. and Lean, H. W.: Scale-Selective Verification of Rainfall Accumulations from High-Resolution Forecasts of Convective Events, Mon. Wea. Rev., 136, 78–97, https://doi.org/10.1175/2007MWR2123.1, 2008.

Robinson, E. L., Huntingford, C., Semeena, V. S., and Bullock, J. M.: CHESS-SCAPE: high-resolution future projections of multiple climate scenarios for the United Kingdom derived from downscaled United Kingdom Climate Projections 2018 regional climate model output, Earth System Science Data, 15, 5371–5401, https://doi.org/10.5194/essd-15-5371-2023, 2023.

Scaff, L., Prein, A. F., Li, Y., Liu, C., Rasmussen, R., and Ikeda, K.: Simulating the convective precipitation diurnal cycle in North America's current and future climate, Clim Dyn, 55, 369–382, https://doi.org/10.1007/s00382-019-04754-9, 2019.

Soares, P. M. M., Careto, J. A. M., Cardoso, R. M., Goergen, K., Katragkou, E., Sobolowski, S., Coppola, E., Ban, N., Belušić, D., Berthou, S., Caillaud, C., Dobler, A., Hodnebrog, Ø., Kartsios, S., Lenderink, G., Lorenz, T., Milovac, J., Feldmann, H., Pichelli, E., Truhetz, H., Demory, M. E., de Vries, H., Warrach-Sagi, K., Keuler, K., Raffa, M., Tölle, M., Sieck, K., and Bastin, S.: The added value of km-scale simulations to describe temperature over complex orography: the CORDEX FPS-Convection multi-model ensemble runs over the Alps, Clim Dyn, 62, 4491–4514, https://doi.org/10.1007/s00382-022-06593-7, 2024.

Themeßl, M. J., Gobiet, A., and Leuprecht, A.: Empirical-statistical downscaling and error correction of daily precipitation from regional climate models, International Journal of Climatology, 31, 1530–1544, https://doi.org/10.1002/joc.2168, 2011.

Trapp, R. J., Robinson, E. D., Baldwin, M. E., Diffenbaugh, N. S., and Schwedler, B. R. J.: Regional climate of hazardous convective weather through high-resolution dynamical downscaling, Clim Dyn, 37, 677–688, https://doi.org/10.1007/s00382-010-0826-y, 2011.

Weisman, M. L., Davis, C., Wang, W., Manning, K. W., and Klemp, J. B.: Experiences with 0–36-h Explicit Convective Forecasts with the WRF-ARW Model, Wea. Forecasting, 23, 407–437, https://doi.org/10.1175/2007WAF2007005.1, 2008.

Weusthoff, T., Ament, F., Arpagaus, M., and Rotach, M. W.: Assessing the Benefits of Convection-Permitting Models by Neighborhood Verification: Examples from MAP D-PHASE, Mon. Wea. Rev., 138, 3418–3433, https://doi.org/10.1175/2010MWR3380.1, 2010.

Wilcke, R. A. I., Mendlik, T., and Gobiet, A.: Multi-Variable Error Correction of Regional Climate Models, Climatic Change, 120, 871–887, https://doi.org/10.1007/s10584-013-0845-x, 2013.

Yun, Y., Liu, C., Luo, Y., Liang, X., Huang, L., Chen, F., and Rasmmusen, R.: Convection-permitting regional climate simulation of warm-season precipitation over Eastern China, Clim Dyn, 54, 1469–1489, https://doi.org/10.1007/s00382-019-05070-y, 2020.

Responses to Anonymous Referee #2

We thank the Anonymous Referee #2 for their time and effort to review our manuscript, which helped to further increase the quality of the paper. We provide a detailed response to the Referee's comments.

Below, referee 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*.

The paper proposes the bias correction (BC) of precipitation and temperature for 4 members from the convection-permitting simulations over UK (UKCP18-CPM), based on gridded observation products. Quantification of biases are presented for average and "extreme" magnitudes, at annual and seasonal scale, before and after correction, evidencing over- or underestimation depending on the members, the variable considered, the region in uk. The future change in precipitation and temperature is then presented, based on the bias corrected simulations, finding a general increase extreme precipitation (Annual Maxima) but moderate change on annual totals, and increase in temperatures.

The paper is generally well written with clear figures, presenting a topic (bias correction of climate models for precipitation and temperature) which is of interest for the hydrological community. Anyway, I think there are some major weakness that need to be addressed before publication, particularly on the novelty, the analyzed domain, the future change from raw and corrected simulations, the discussion section.
We thank the referee for the careful reading of the manuscript and for the constructive summary of the study. We appreciate the positive comments on the overall presentation. We also acknowledge the referee's major concerns regarding novelty, the analysed domain, the comparison between raw and bias-corrected future simulations, and the Discussion section, and we address these points in detail below.

I list below my major concerns, and then bullet points on more specific/minor comments.

General Comments:
1) the novelty of this work should be better highlighted, in the introduction and discussion. Considering that methodology is based on already existing approaches, author should state more clearly why this study is relevant.
We thank the referee for the comment. We agree that quantile mapping (QM) and diurnal bias correction are existing approaches. We have revised Introduction, Discussion and Conclusion to explain that our novelty is not the development of a new bias-correction method, but the systematic testing and evaluation of diurnal QM for sub-daily precipitation in a domain where such applications are still rarely reported, particularly for convection-permitting climate model output. We now emphasise that the key contribution of the study is to provide evidence on the added value of incorporating hour-of-day structure in QM when correcting hourly precipitation, assessed using metrics relevant to sub-daily impacts (diurnal cycle behaviour and high-

percentile rainfall). These revisions clarify why the study is relevant despite building on existing approaches.

In Introduction, we add the following text between Line 90 and 91: "*Building on the findings of Faghih et al. (2022), we address a remaining gap: diurnal-cycle-aware bias correction is still rarely implemented and systematically evaluated for sub-daily precipitation in convection-permitting climate simulations, despite its importance for impact models that are sensitive to rainfall timing and intensity. The contribution of this paper is therefore application-driven evaluation in a CPM sub-daily setting, providing evidence on whether incorporating diurnal structure within quantile mapping improves (i) conventional climatological statistics and (ii) more challenging characteristics, including the diurnal cycle and sub-daily extremes.*"

In Discussion, we revised the first paragraph in Lines 407-412: "*This study provides a practical and reproducible bias correction approach for UKCP18-CPM hourly precipitation and daily mean temperature over England and evaluates how the adjusted outputs compare with observational reference datasets across mean behaviour, diurnal characteristics, and extremes. A set of bias-corrected 1 km UKCP18-CPM temperature and precipitation projections for England is provided, using empirical quantile mapping (QM) for daily mean temperature and diurnal bias correction (DBC) for hourly precipitation. The methodological contribution is not the development of a new bias-correction method, but the transparent implementation and assessment of established quantile-mapping methods for a large CPM dataset at hourly resolution. This is relevant for impact applications where modelled rainfall timing and intensity can strongly influence simulated responses, especially for fast-response impact applications (e.g., short-duration rainfall-driven flooding). These results add evidence on the value of diurnal-cycle bias correction for sub-daily precipitation in convection-permitting simulations, a setting where such evaluations remain relatively limited.*"

In Conclusion, we add a sentence in Line 465: "*Although the bias-correction approaches applied here are established, this study provides new evidence of their performance for sub-daily precipitation in convection-permitting UK climate projections, including improvements in diurnal-cycle behaviour and sub-daily extremes relevant to impact modelling.*"

2) I find not clear motivation for limiting the analysis based on catchments (any hydrological modelling is applied here) instead of whole UK, which I believe could be more relevant. I strongly suggest to expand the study domain, and completely change section 2.1 and figure 1 (why flow gauges are shown?)

Thank you for the comment. In this study, catchments are used only to define the analysis mask (i.e., which 1 km grid cells are processed), while the bias correction is applied at the grid-cell level and no hydrological modelling is performed. We restricted the processing to this mask because bias correcting the full UK domain and all CPM members at 1 km hourly resolution is computationally and storage intensive, and the main aim of the paper is to present and evaluate the bias correction method rather than to deliver a national scale dataset. We agree that showing flow-gauge locations was misleading, and we have updated Figure 1 accordingly.

Section 2.1 has been revised to clarify that catchments are used only to define the analysis mask and that all bias correction is performed at the grid-cell level (no hydrological simulation). Figure 1 has been updated by removing the flow-gauge locations and instead showing the processed grid cells used for bias correction. The caption has been revised accordingly.

The paragraph in Section 2.1 has been replaced with: *"This study focuses on England (Fig. 1), with the primary aim of assessing the bias-correction method. Catchment boundaries are obtained from 249 catchments in the National River Flow Archive (NRFA) and are used only to define the analysis mask (i.e., to select the 1 km grid cells processed). No hydrological modelling or catchment-scale simulation is performed. The processed HadUK-Grid cells within this mask are displayed in Fig. 1 by their centre points. Given computational and storage constraints, restricting the analysis to grid cells within these catchments provides a practical and geographically coherent subset for the study. The catchments were selected based on the availability of complete gauged records for December 1990 to November 2000, overlapping with both the CEH-GEAR1hr observational dataset and the UKCP18-CPM baseline period (see Sect. 2.2.2; hereafter referred to as the reference period), and on minimal anthropogenic influence. Bias correction was applied independently at the grid-cell level to UKCP18-CPM precipitation and temperature for 62,488 1 km grid cells. Across the processed grid cells, the mean annual precipitation is approximately 2.64 mm* d^{-1}*, and the mean annual temperature is around 9.3 °C, based on the CEH-GEAR1hr and HadUK-Grid datasets, respectively."*

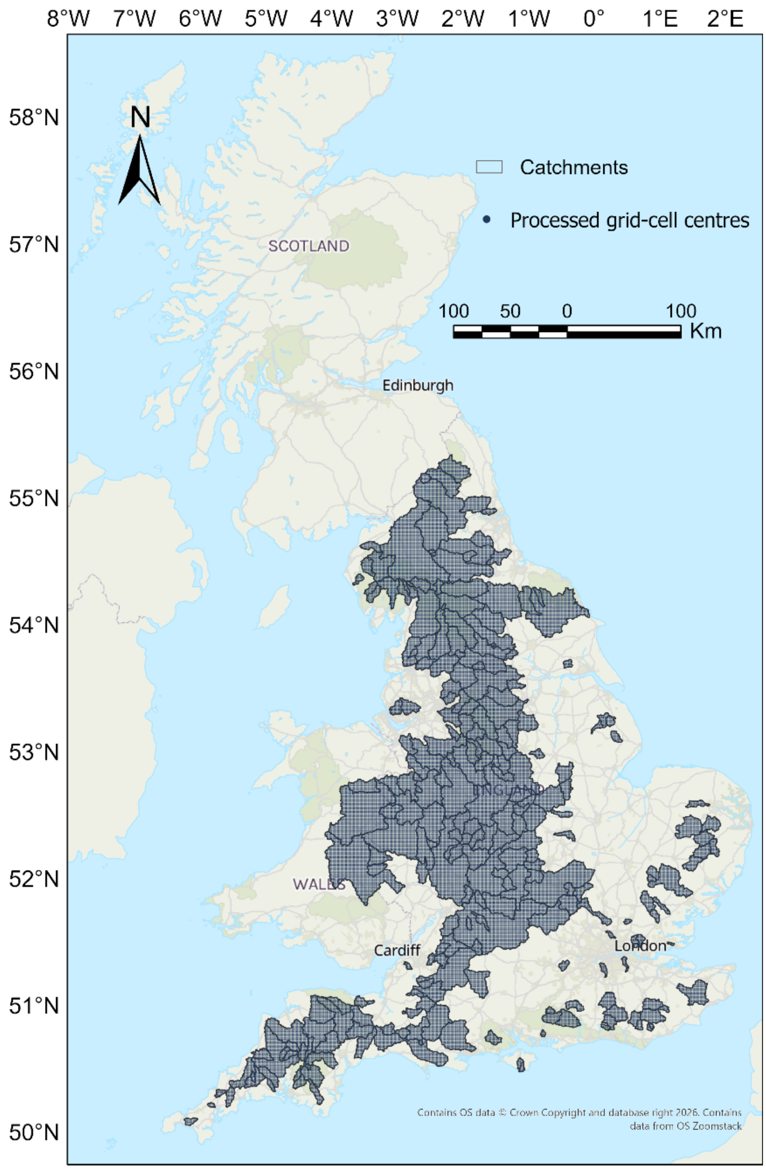

Figure 1: Study domain in England. Catchment boundaries (249 catchments) are used only to define the analysis mask for grid-cell processing. The processed 1 km HadUK-Grid cells are displayed by their centre points.

3) I suggest to compare also future changes based on raw simulations with those from bias corrected ones, to show/discuss the impact of bias correction on projected changes

Thank you for the suggestion. To facilitate the comparison between RAW and bias-corrected (BC) projections, we have added Figure 12 (heatmaps) in Section 3.3, which provides a clear numerical summary of the projected changes in the 2030s and 2070s for both RAW and BC simulations. This figure supports the projected change values reported in the text, which are not easy to read directly from the spatial maps in Figures 9, 10 and11.

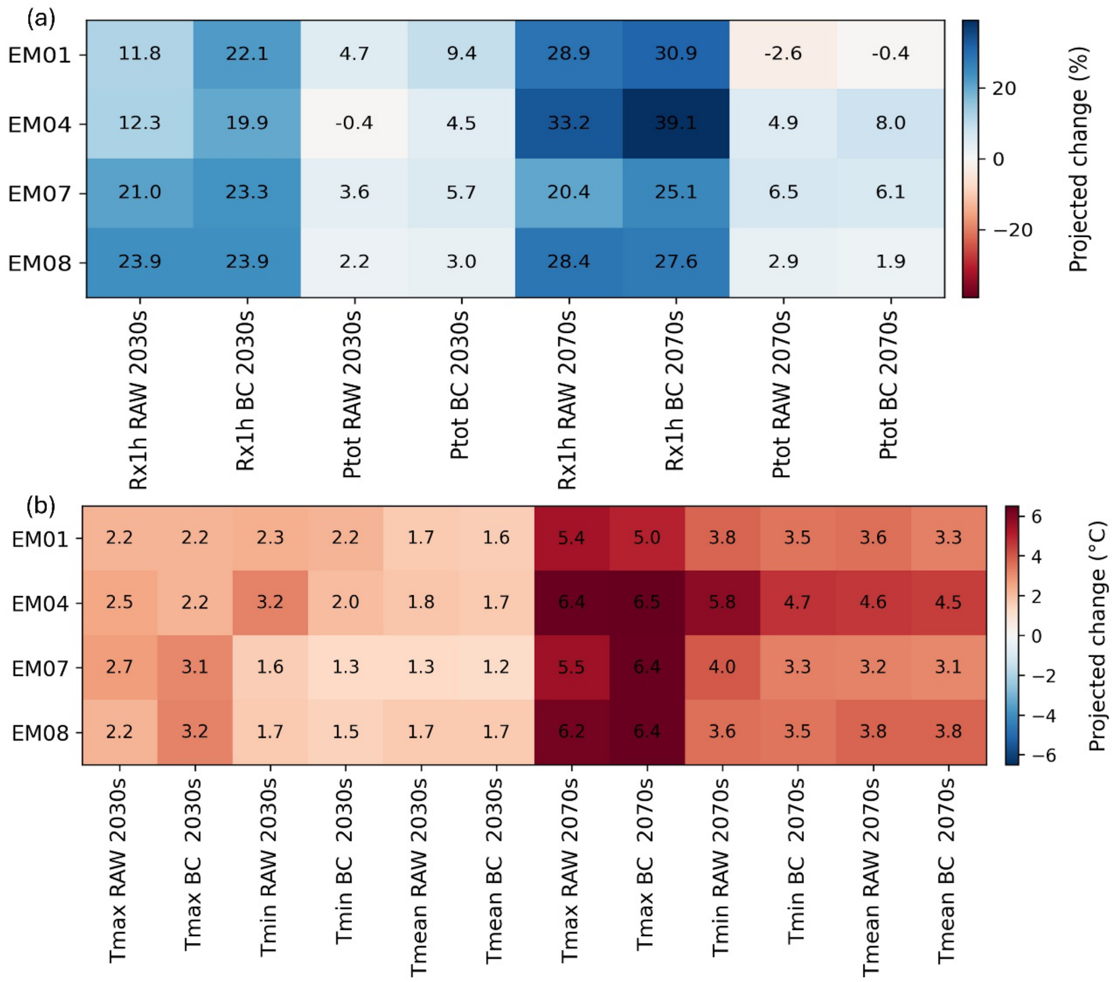

Figure 12: Projected changes in (a) annual maximum 1 h precipitation (Rx1h) and annual total precipitation (Ptot), and (b) annual maximum (Tmax), annual mean (Tmean), and annual minimum (Tmin) of daily mean temperature for UKCP18-CPM simulations before (RAW) and after bias correction (BC). Each value is the spatial average over the processed grid cells, shown for the 2030s and 2070s relative to the baseline period.

4) No clear why bias on extremes is shown on P95, and changes are shown for Annual Maxima. I suggest to also show biases on AM (at least in supplementary).

Thank you for this comment. Figures 6 and 9 were intended for different purposes. Figure 6 uses P_{95} to evaluate how bias correction improves agreement with observations for high-end precipitation during the reference period, whereas Figure 9 shows projected future changes in annual maximum 1 h precipitation (Rx1h) and annual total precipitation (Ptot). The two figures were designed to address different questions rather than to use identical metrics.

5) Discussion is more a summary of results (all lines from 407 to 434!), with no explanation/interpretation of your findings (bias and change) and very poor comparison with other works (just a few sentences based on studies using the same models). And on the choice/impact of the specific BC method with respect of others. Some example references for biases: https://doi.org/10.1016/j.jhydrol.2025.133324; https://doi.org/10.1007/s00382-021-05708-w; https://doi.org/10.1007/s00382-022-06593-7; for future changes: https://doi.org/10.1029/2024EF005185; https://doi.org/10.1007/s00382-021-05657-4; ….
Thank you for this comment. We agree that the original Discussion relied too much on a summary of results and did not provide enough interpretation or engagement with the wider literature. We have therefore revised this section substantially. We have reduced the descriptive repetition of the results, expanded the comparison with previous studies, and clarified the discussion of the bias-correction methods used in this study.
In addition, we added a new paragraph in the Discussion (line 413) to clarify the selected bias-correction methods and to discuss their potential limitations, as follows: *"We applied empirical QM to daily mean temperature to address seasonal biases and distributional differences using a reproducible procedure. It is widely used because it can correct not only mean biases but also biases in variability and quantiles (Fang et al., 2015; Themeßl et al., 2011; Wilcke et al., 2013). For hourly precipitation, we used DBC by hour of day, because applying a single mapping across all hours can overlook systematic diurnal-cycle biases in sub-daily precipitation. Figure 5 supports this choice, showing that the bias-corrected ensemble more closely reproduces the observed hour-of-day pattern. This hour-of-day treatment is intended to reduce discrepancies that matter for sub-daily impact applications, consistent with the broader argument that correcting the diurnal cycle can improve sub-daily bias-correction strategies (Faghih et al., 2022). Therefore, we recommend using DBC for sub-daily variables. However, in this study we applied a 3 h moving window in the DBC. By incorporating hourly data from the previous and following hours, this may have contributed to some of the remaining discrepancies. It may be worth exploring whether omitting the moving window gives better results when the training period sample size is sufficiently large*."
The revised Discussion text (lines 413–434) now reads as follows: "*Before bias correction, the UKCP18-CPM simulations show considerable biases in both precipitation and temperature across various temporal and spatial scales. In general, UKCP18-CPM simulations show wet precipitation biases (especially in winter) and cool biases in temperature. These broad patterns are consistent with previous evaluations of UKCP18-RCM simulations (Reyniers et al., 2025) and the UKCP18-CPM science report (Kendon et al., 2019b). These residual biases are consistent with the broader CPM literature, which shows that although kilometre-scale models can improve the representation of precipitation compared with convection-parameterised RCMs, systematic biases and model uncertainty may remain, particularly for sub-daily precipitation and across different regional settings (Ban et al., 2021; Correa-Sánchez et al., 2025; Soares et al., 2024). After bias correction, both precipitation and temperature show*

substantially improved agreement with observations at monthly and diurnal timescales (Figs. 4 and 5). This suggests that bias correction provides a more reliable basis for downstream impact assessments than raw UKCP18-CPM output.
Future changes were also analysed for both precipitation and temperature in Figs. 9–12. In general, the bias-corrected projections indicate a clear intensification of annual maximum 1 h precipitation (Rx1h) by both the 2030s and 2070s, while changes in annual total precipitation (Ptot) remain smaller and spatially variable (Figs. 9 and 12). This pattern (i.e., stronger changes in short-duration extremes than in totals) is broadly consistent with CPM-based studies that report amplified changes for sub-daily extremes and a strong dependence on dynamical factors and regional setting (Dallan et al., 2024; Pichelli et al., 2021). The projected temperature changes are consistent with findings from the UKCP18-CPM science report (Kendon et al., 2019b), which indicated that mean temperature is expected to increase across all regions and seasons. Our findings also align well with existing literature, such as Robinson et al. (2023), who reported similar projected changes for UKCP18-RCM. This suggests that the bias-corrected projections remain broadly consistent with previously reported climate change signals."

Specific/minor Comments:

line 39-45: maybe better after line 32, where the different models (GCM – RCM) are presented

Thank you for this suggestion. The first paragraph of the Introduction is intended to set out the general progression from coarse-resolution GCMs to finer-resolution RCMs, thereby establishing the motivation for using higher-resolution climate models. The second paragraph then narrows the focus to the UK context, introducing UKCP18 and, within that framework, UKCP18-CPM as the specific convection-permitting model dataset used in this study. We agree, however, that introducing other convection-permitting models immediately after UKCP18-CPM could create some ambiguity, partly because the abbreviation "CPM" was used too broadly in the original text. We have therefore rewritten this part of the Introduction to make the subject more explicit by referring clearly to convection-permitting models in general, while keeping UKCP18-CPM as the specific dataset used in this study.

It now reads: "*Convection-permitting models (CPMs) have become a valuable tool in short-range weather forecasting due to their enhanced ability to represent convective processes in detail, improve forecast accuracy, and capture localised high-impact rainfall that coarser models often miss (e.g., Done et al., 2004; Lean et al., 2008; Roberts and Lean, 2008; Weisman et al., 2008; Weusthoff et al., 2010). CPMs have also been widely applied in Europe (Berthou et al., 2020; Pichelli et al., 2021), Asia (Murata et al., 2017; Yun et al., 2020), and Africa (Kendon et al., 2019a; Maurer et al., 2017), as well as other regions (Trapp et al., 2011), to enhance the representation of convective processes and improve the accuracy of climate* projections."

line 90: no diurnal cycle for temperature if you sue daily values!

Thank you for pointing this out. The original sentence was intended to cite the general finding of Faghih et al. (2022) that correcting the diurnal cycle can be beneficial for sub-daily precipitation and temperature. In our study we use daily temperature, so there is no temperature diurnal cycle to correct. To avoid any ambiguity, we have removed temperature from this sentence and now refer only to the diurnal-cycle correction for sub-daily precipitation.

The sentence in lines 88-90 has been revised to: *Faghih et al. (2022) compared two bias-corrected time series using a multivariate quantile mapping method, both with and without correction of the diurnal cycle, and found that bias correction of the diurnal cycle for sub-daily precipitation is preferable.*

from line 161: In this section I find some repetitions of information contained in 2.2.1. Consider a careful re-reading for optimizing

Thank you for the suggestion. We agree that some background information about the UKCP18 ensemble was repeated in Section 2.2.1 and Section 2.3 (from line 161). To improve readability, we have streamlined the text by keeping the dataset description in Section 2.2.1 and shortening the opening of Section 2.3.1 to avoid repeating information already provided.

The Sentences in lines 163-175 have been revised to: "*The UKCP18-CPM consists of 12 projections driven by the 12 km RCM ensemble. Due to computational and storage constraints, it was not feasible to bias-correct the full CPM-PPE at 1 km hourly resolution. We therefore selected a subset of ensemble members for bias correction and implemented a parallel workflow on the University of East Anglia's high-performance computing (HPC) system to complete the processing within a reasonable time. Four ensemble members were selected to represent the broad range of possible precipitation outcomes within the ensemble.*"

lines 169-174. No useful to have all these details on computational cost.

Thank you for this suggestion. We included the computational and storage information to justify why bias-correcting the full UKCP18 CPM-PPE (12 members) was not practical. However, we agree that the level of detail was excessive for the main text. We have therefore substantially shortened this paragraph, kept only a brief justification and removed the specific runtime and storage numbers.

The text in lines 169-176 have been replaced with: "*The UKCP18-CPM consists of 12 projections driven by the 12 km RCM ensemble. Due to computational and storage constraints, it was not practical to bias-correct the full CPM-PPE at 1 km hourly resolution. We therefore selected a subset of ensemble members for bias correction and implemented a parallel workflow on the University of East Anglia's high-performance computing (HPC) system to complete the processing within a reasonable time. Four ensemble members were selected to represent the broad range of possible precipitation outcomes within the ensemble.*"

Line 230-233: redundant sentences.

Thank you for this comment. **This sentences in lines 229- 234 have been revised to:** "*For hourly precipitation, a single correction applied to all hours can miss the diurnal cycle. This may influence the reliability of climate model outputs, particularly in regional impact assessments involving small catchments that exhibit rapid hydrological responses to precipitation (Ban et al., 2014; Dai et al., 1999).*"

Line 237: no clear to me why a 3h moving window; precipitation is intermittent and variable in time, not with "continuous variations" as for temperature

Because precipitation is intermittent, the sample within any single hour can be limited. We therefore used a 3 h moving window to increase the sample pool and make the training more robust. This approach helps to stabilise the estimation of the hour-of-day empirical CDFs and the resulting mapping and reduces noisy hour-to-hour variability in the derived correction, particularly for the upper tail.

Line 237: 24 unique correction factors … per month?

Yes, the 24 unique correction factors are calculated per month. As stated in Line 235, we first split the data into 12 monthly groups, and the data for each month was then further divided into 24 hourly groups. Therefore, we derived 24 hour-specific corrections for each month.

Line 246; "precipitation biases" … on which prec. amount? Seasonal total?

Thank you for the comment. In Figure 2a, "precipitation biases (%)" refer to the relative bias of mean hourly precipitation (mm h^{-1}), calculated separately for ANN, DJF and JJA over the reference period. It is therefore based on mean precipitation intensity, not seasonal totals. We have clarified this in the text.

The sentence in line 246 has been changed to: "*Figure 2a shows the spatial distribution of relative biases (%) in mean hourly precipitation in the raw UKCP18-CPM simulations compared with CEH-GEAR1hr for the reference period. Figure 2b shows the change in the number of hourly rainfall events exceeding 20 mm h^{-1} in UKCP18-CPM relative to CEH-GEAR1hr over the reference period. The 20 mm h^{-1} threshold is commonly used as an indicator of potential flash flood producing rainfall in the UK (e.g., Kendon et al., 2023).*"

Line 262: already said at line 259-260

Agreed. We removed the repeated sentence in lines 259-260.

Figure 2b: maybe better %bias also for event number; add mean as done for panel b. I suggest to add also a metric for the ranges in the domain (e.g. st.dev or iqr) for all panels. I suggest to use different colors than red/blue for the color bar, because it is confounding to have then red/blue in figure 3 for opposite biases

Thank you for the suggestions. For Figure 2b, we did not add a mean value, because the spatial pattern includes both positive and negative differences and the domain average would be strongly affected by domain-mean cancellation. A single mean value would therefore be of limited interpretive value and could be misleading. In this figure the main purpose is to show the spatial pattern of the biases across England, and the maps themselves already convey the spatial heterogeneity and range of values within the domain. We therefore chose not to add an additional summary metric such as standard deviation or interquartile range.

We also keep Figure 2b as an absolute difference in event counts, because ⩾20 mm h^{-1} events are rare and percentage values can become unstable when observed counts are very small. The purpose of Figure 2b is to provide a visual check of whether UKCP18-CPM reproduces the occurrence of very intense hourly rainfall events (particularly in winter), whereas the main evaluation of precipitation extremes and the bias-correction performance is presented in Figure 6 (P_{95}).

Regarding the colour bar, we keep the current colour scales because they follow an intuitive physical interpretation: positive temperature biases are shown in red (warmer) in Figure 3, whereas positive precipitation biases are shown in blue (wetter) in Figure 2. To avoid any confusion, we have clarified the sign conventions in the captions of Figures 2 and 3.

The caption of Figure 2 has been changed to: "*Figure 2: Spatial distribution of (a) biases (%) of mean precipitation between UKCP18-CPM and CEH-GEAR1hr and (b) differences in the number of events exceeding 20 mm h^{-1} between UKCP18-CPM and CEH-GEAR1hr for the reference period over England. Rows represent annual (ANN), winter (DJF: December, January, February), and summer (JJA: June, July, August) biases, while columns show individual ensemble members (EM01, EM04, EM07, EM08) and the ensemble mean (EMean). Blue indicates positive (wetter) biases and red indicates negative (drier) biases.*"
The caption of Figure 3 has been revised to: "*Figure 3: Temperature biases (°C) in UKCP18-CPM for the reference period over England. 295 The panels show biases for each ensemble member (EM01, EM04, EM07, and EM08) and the ensemble mean (EMean). Rows represent different time scales: annual (ANN), winter (DJF: December, January, February), and summer (JJA: June, July, August). Red indicates positive (warmer) biases and blue indicates negative (cooler) biases.*"

Line 278-279: why to put EM08 separated? Just mention in the sentence before that the range is -0.87 +0.02, with 3 out of 4 models with negative mean bias.
Thank you for this comment. We agree with this.
The sentence in Lines 277–279 has been revised to: "*For the annual mean temperature (ANN), the ensemble members show a generally cool bias across England relative to HadUK-Grid, with mean biases ranging from -0.87 °C (EM04) to +0.02 °C (EM08). Three out of the four members have negative mean biases.*"

Line 282: "more pronounced" …based on mean values, this is no true for 3 models …
Thank you for the comment. We agree that "more pronounced" could cause misunderstanding. What we intended to describe was the clearer contrast between members in DJF. We have therefore revised this sentence accordingly.
The sentences in lines 282-284 have been changed to: "*During winter (DJF), the temperature biases vary more clearly across ensemble members. EM04 exhibits a cold bias of -0.85 °C, whereas EM08 shows a warm bias of 0.46 °C, indicating marked inter-member differences in winter temperature bias. The spatial distribution of winter…*"

Lin 291-294: merge sentences, expressing same concepts
Thank you for the comment.
The sentences in Lines 291–294 have been merged to avoid repetition and now read: *"The spatial pattern during JJA shows a clear north–south gradient, with stronger cooling biases in northern England and smaller biases in the south, a pattern that is consistent across all ensemble members."*

Lin 324: "unrealistic fluctuations" … precipitation is intermittent !
Thank you for the comment. Our previous wording was inaccurate. Because precipitation is intermittent, the sample within any single hour can be limited. We therefore used a 3 h moving window to increase the sample pool and make the training more robust. This approach helps to stabilise the estimation of the hour-of-day empirical CDFs and the resulting mapping and reduces noisy hour-to-hour variability in the derived correction, particularly for the upper tail.

The sentence in line 323 has been replaced by: *"These discrepancies are mainly related to the 3 h moving window used in the DBC process, which pools data from the target hour and its neighbouring hours to stabilise the hour-of-day correction."*

Figure 5: average hourly precipitation?

We clarified in the caption that Figure 5 shows the mean hourly precipitation for each hour of the day.

The caption of Figure 5 has been revised to: "*Figure 5: Diurnal cycle of mean hourly precipitation before bias correction (blue lines: UKCP18-CPM ensemble) and after bias correction (yellow lines: UKCP18-CPM ensemble BC) for the reference period. The CEH-GEAR1hr (observations) are shown as a red solid line with circles. The ensemble mean before bias correction is shown as a darker blue line (UKCP18-CPM ensemble mean), and the bias-corrected ensemble mean is shown as a darker yellow line (UKCP18-CPM ensemble mean BC). The thin yellow lines (individual members) are largely overlain by the thick yellow line (ensemble mean), as the diurnal cycles are very similar across the four members.*"

From line 332: I suggest to also report the %bias for the 95th percentiles as made in the previous section

Thank you for the suggestion. We focus on scatter plots with fitted slopes and R^2 because they directly assess grid-cell agreement across the full study domain and are less sensitive to cancellation effects that can occur when summarising bias with a single study domain mean percentage. We therefore retain the current presentation.

Figure 6 second row: Logical order of violin plot is like the legend: obs-raw-corrected

Thank you for the suggestion. We have reordered the violin plots in the second row of Figure 6 to follow the same order as the legend (obs-raw-corrected).

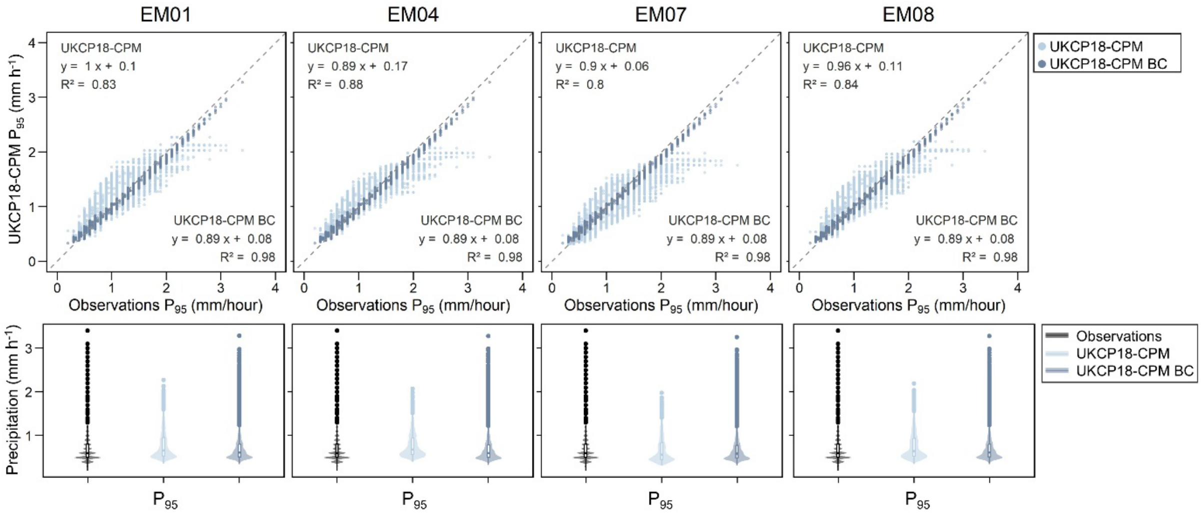

Figure 6: Comparison of 95th percentile (P_{95}) hourly precipitation values for the reference period. The top row shows scatter plots of UKCP18-CPM P_{95} values for each ensemble member, plotted against CEH-GEAR1hr observations for each 1 km grid, with raw (light blue) and bias-corrected (dark blue) data. The bottom row presents violin plots of P_{95} values for observations (black), raw UKCP18-CPM (light blue), and bias-corrected (dark blue) data.

Line 350-354: I suggest to shorten

The text in Lines 350-354 has been shortened to: "*Temperature extremes were evaluated by comparing the 95th (T_{95}) and 5th (T_5) percentiles of daily mean temperature in UKCP18-CPM with HadUK-Grid for the reference period. Figure 7 shows grid-cell scatter plots (T_{95} in the top row and T_5 in the bottom row) for each ensemble member, while Fig. 8 summarises the distributions using violin plots.*

For both T_{95} and T_5, the raw model outputs (light pink) show a fair level of correlation with the observations, with gradient values…"

Figure 8: consider to add as 3rd row in figure 7, as figure 6 (considering also the very short description of this figure)

Thank you for this suggestion. We considered merging Figure 8 into Figure 7 (as a third row), similar to Figure 6. However, Figure 7 already contains eight scatter panels (T_{95} and T_5 for four ensemble members) with regression information, and adding an additional row would make the figure overly dense and reduce readability. We therefore retain Figure 8 as a separate figure to clearly present the distributional information (violin plots). In addition, we have merged the two paragraphs in Lines 353–362 into a single paragraph, so that Figures 7 and 8 are discussed together in the text, which better reflects their close relationship in the analysis.

Figure 9-10-11: add mean change and a metric of range, in each panel; color bar for temperature doesn't allow to distinguish different changes

Thank you for this comment. In Figures 10 and 11, we have improved readability by limiting the colourbar range to 0–10 °C, since the projected changes show warming only, and updated the caption. In addition, we added Figure 12 (heatmaps) to report the study domain mean projected changes for the 2030s and 2070s across ensemble members. We did not add range metrics, as Figure 12 provides a concise numerical summary of the mean changes, while spatial variability is shown by the maps (Figures 10 and 11).

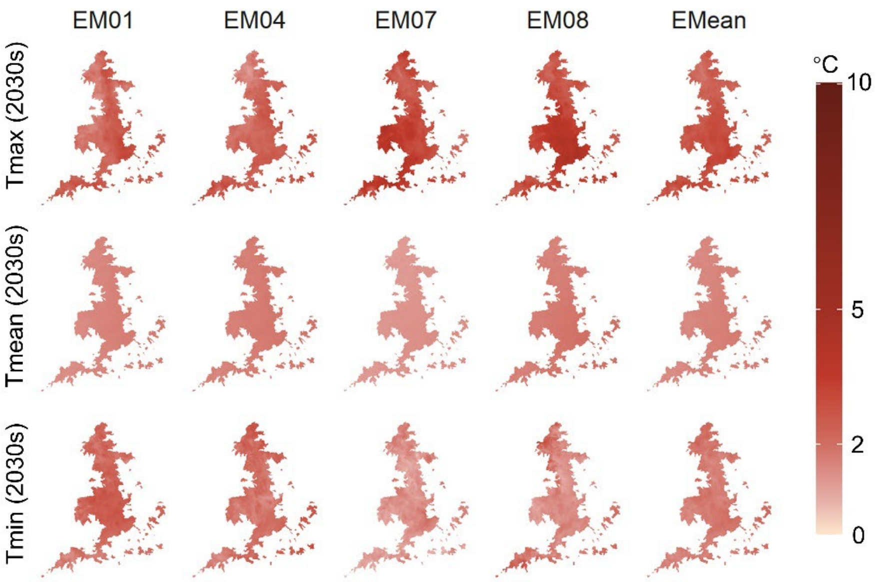

Figure 10: Projected changes in temperature (°C) for the 2030s (Dec 2020–Nov 2040) relative to the baseline period from the bias-corrected UKCP18-CPM simulations. Panels show (top to bottom) changes in the annual maximum (Tmax), annual mean (Tmean) and annual minimum

(Tmin) of daily mean temperature, for each ensemble member (EM01, EM04, EM07, EM08) and the ensemble mean (EMean).

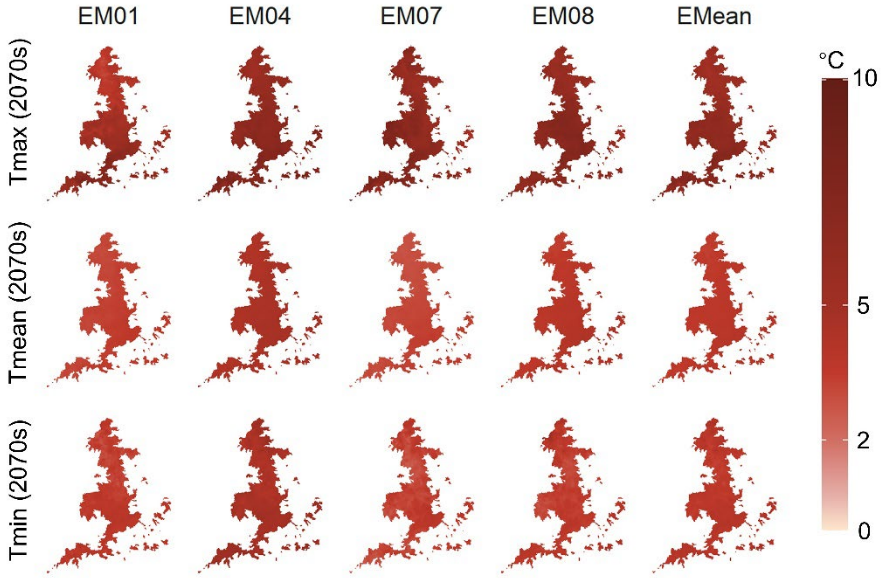

Figure 11: Projected changes in temperature (°C) for the 2070s (Dec 2060–Nov 2080) relative to the baseline period from the bias-corrected UKCP18-CPM simulations. Panels show (top to bottom) changes in the annual maximum (Tmax), annual mean (Tmean) and annual minimum (Tmin) of daily mean temperature, for each ensemble member (EM01, EM04, EM07, EM08) and the ensemble mean (EMean).

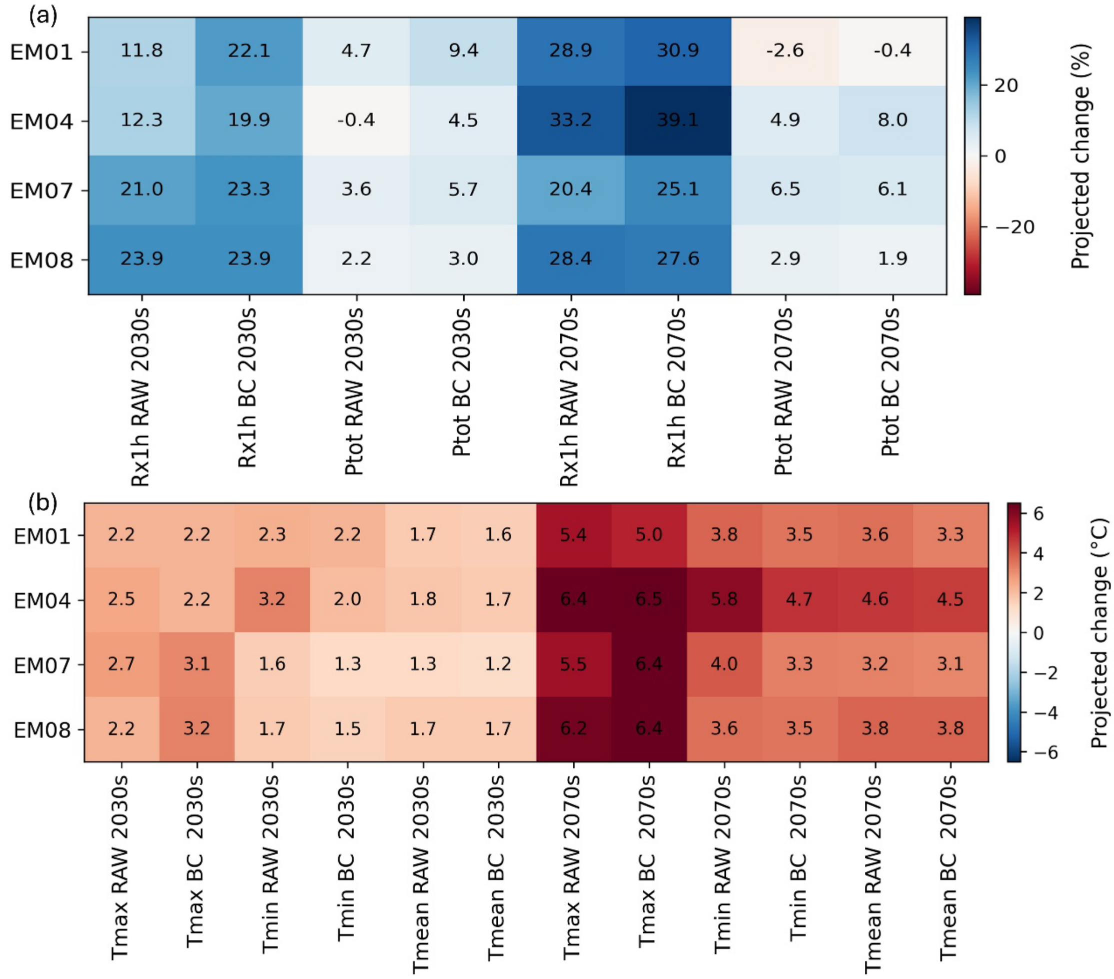

Figure 12: Projected changes in (a) annual maximum 1 h precipitation (Rx1h) and annual total precipitation (Ptot), and (b) annual maximum (Tmax), annual mean (Tmean), and annual minimum (Tmin) of daily mean temperature for UKCP18-CPM simulations before (RAW) and after bias correction (BC). Each value is the spatial average over the processed grid cells, shown for the 2030s and 2070s relative to the baseline period.

Line 453-439: already said previously in the paper.

Thank you for the suggestion. We assume you are referring to lines 435–439. We agree that this point was repetitive, as the limitation related to computational resources had already been mentioned in the previous paragraph. We have therefore revised this part of the Discussion by moving the relevant sentence into the paragraph in lines 435–441 and reorganising it accordingly.

It now reads (lines 422–434): *"Future changes were also analysed for both precipitation and temperature in Figs. 9–12. In general, the bias-corrected projections indicate a clear intensification of annual maximum 1 h precipitation (Rx1h) by both the 2030s and 2070s, while changes in annual total precipitation (Ptot) remain smaller and spatially variable (Figs. 9 and 12). This pattern (i.e., stronger changes in short-duration extremes than in totals) is broadly consistent with CPM-based studies that report amplified changes for sub-daily extremes and a strong dependence on dynamical factors and regional setting (Dallan et al., 2024; Pichelli et al., 2021). The projected temperature changes are consistent with findings from the UKCP18-*

CPM science report (Kendon et al., 2019b), which indicated that mean temperature is expected to increase across all regions and seasons. Our findings also align well with existing literature, such as Robinson et al. (2023), who reported similar projected changes for UKCP18-RCM. This suggests that the bias-corrected projections remain broadly consistent with previously reported climate change signals."

Lines 435–441: "*Due to limitations in computational resources and time, we focused our analysis on four ensemble members: EM01, EM04, EM07, and EM08. These were selected to represent a diverse range of climate outcomes, from the driest to the wettest scenarios, allowing the bias-corrected dataset to effectively capture the range of possible climate responses in England. This sub-ensemble was chosen to balance computational efficiency with representativeness while ensuring that the selection captures a wide range of precipitation responses. The reduced number of ensemble members may influence the ensemble mean and spread compared to using all 12 members. Future studies could address this by expanding the bias correction to include more ensemble members and broader spatial coverage. Additionally, the integration of multivariate bias correction methods (Cannon, 2018; Faghih et al., 2022) could offer the advantage of preserving inter-variable dependencies, ensuring that precipitation and temperature are corrected consistently without disrupting their natural relationship.*"

Useless table A! with information of catchments of any interest in this study.

Thank you for the comment. The list of NRFA catchment IDs is provided in the Appendix for transparency and reproducibility, because the selected catchments define the analysis mask (i.e., which 1 km grid cells are included). This table also allows readers to identify and cross-reference the included catchments with NRFA records if they wish to use or interpret the dataset for specific catchments. To make the purpose clearer, we have added one sentence in Appendix A introducing Table A1 (and we keep the table in the Appendix rather than the main text).

The sentence in line 106 has been changed to: *"Appendix A. Table A1 lists the NRFA IDs of the 249 catchments used to define the analysis mask in this study."*

References

Ban, N., Schmidli, J., and Schär, C.: Evaluation of the convection-resolving regional climate modeling approach in decade-long simulations, Journal of Geophysical Research: Atmospheres, 119, 7889–7907, https://doi.org/10.1002/2014JD021478, 2014.

Ban, N., Caillaud, C., Coppola, E., Pichelli, E., Sobolowski, S., Adinolfi, M., Ahrens, B., Alias, A., Anders, I., Bastin, S., Belušić, D., Berthou, S., Brisson, E., Cardoso, R. M., Chan, S. C., Christensen, O. B., Fernández, J., Fita, L., Frisius, T., Gašparac, G., Giorgi, F., Goergen, K., Haugen, J. E., Hodnebrog, Ø., Kartsios, S., Katragkou, E., Kendon, E. J., Keuler, K., Lavin-Gullon, A., Lenderink, G., Leutwyler, D., Lorenz, T., Maraun, D., Mercogliano, P., Milovac, J., Panitz, H.-J., Raffa, M., Remedio, A. R., Schär, C., Soares, P. M. M., Srnec, L., Steensen, B. M., Stocchi, P., Tölle, M. H., Truhetz, H., Vergara-Temprado, J., de Vries, H., Warrach-Sagi, K., Wulfmeyer, V., and Zander, M. J.: The first multi-model ensemble of regional climate simulations at kilometer-scale resolution, part I: evaluation of precipitation, Clim Dyn, 57, 275–302, https://doi.org/10.1007/s00382-021-05708-w, 2021.

Berthou, S., Kendon, E. J., Chan, S. C., Ban, N., Leutwyler, D., Schär, C., and Fosser, G.: Pan-European climate at convection-permitting scale: a model intercomparison study, Clim Dyn, 55, 35–59, https://doi.org/10.1007/s00382-018-4114-6, 2020.

Cannon, A. J.: Multivariate Quantile Mapping Bias Correction: An N-dimensional Probability Density Function Transform for Climate Model Simulations of Multiple Variables, Climate Dynamics, 50, 31–49, https://doi.org/10.1007/s00382-017-3580-6, 2018.

Correa-Sánchez, N., Dallan, E., Marra, F., Fosser, G., and Borga, M.: Orographic control on bias and uncertainty in extreme sub-daily precipitation simulations from a convection-permitting ensemble, Journal of Hydrology, 659, 133324, https://doi.org/10.1016/j.jhydrol.2025.133324, 2025.

Dai, A., Giorgi, F., and Trenberth, K. E.: Observed and model-simulated diurnal cycles of precipitation over the contiguous United States, J. Geophys. Res., 104, 6377–6402, https://doi.org/10.1029/98JD02720, 1999.

Dallan, E., Marra, F., Fosser, G., Marani, M., and Borga, M.: Dynamical Factors Heavily Modulate the Future Increase of Sub-Daily Extreme Precipitation in the Alpine-Mediterranean Region, Earth's Future, 12, e2024EF005185, https://doi.org/10.1029/2024EF005185, 2024.

Done, J., Davis, C. A., and Weisman, M.: The next generation of NWP: explicit forecasts of convection using the weather research and forecasting (WRF) model, Atmosph. Sci. Lett., 5, 110–117, https://doi.org/10.1002/asl.72, 2004.

Faghih, M., Brissette, F., and Sabeti, P.: Impact of correcting sub-daily climate model biases for hydrological studies, Hydrology and Earth System Sciences, 26, 1545–1563, https://doi.org/10.5194/hess-26-1545-2022, 2022.

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.

Kendon, E. J., Stratton, R. A., Tucker, S., Marsham, J. H., Berthou, S., Rowell, D. P., and Senior, C. A.: Enhanced future changes in wet and dry extremes over Africa at convection-permitting scale, Nat Commun, 10, 1794, https://doi.org/10.1038/s41467-019-09776-9, 2019a.

Kendon, E. J., Fosser, G., Murphy, J., Chan, S., Clark, R., Harris, G., Lock, A., Lowe, J., Martin, G., Pirret, J., Roberts, N., Sanderson, M., and Tucker, S.: UKCP Convection-permitting model projections: Science report, 2019b.

Kendon, E. J., Fischer, E. M., and Short, C. J.: Variability conceals emerging trend in 100yr projections of UK local hourly rainfall extremes, Nat Commun, 14, 1133, https://doi.org/10.1038/s41467-023-36499-9, 2023.

Lean, H. W., Clark, P. A., Dixon, M., Roberts, N. M., Fitch, A., Forbes, R., and Halliwell, C.: Characteristics of High-Resolution Versions of the Met Office Unified Model for Forecasting Convection over the United Kingdom, Mon. Wea. Rev., 136, 3408–3424, https://doi.org/10.1175/2008MWR2332.1, 2008.

Maurer, V., Bischoff-Gauß, I., Kalthoff, N., Gantner, L., Roca, R., and Panitz, H.-J.: Initiation of deep convection in the Sahel in a convection-permitting climate simulation for northern Africa, Quarterly Journal of the Royal Meteorological Society, 143, 806–816, https://doi.org/10.1002/qj.2966, 2017.

Murata, A., Sasaki, H., Kawase, H., and Nosaka, M.: Evaluation of precipitation over an oceanic region of Japan in convection-permitting regional climate model simulations, Clim Dyn, 48, 1779–1792, https://doi.org/10.1007/s00382-016-3172-x, 2017.

Pichelli, E., Coppola, E., Sobolowski, S., Ban, N., Giorgi, F., Stocchi, P., Alias, A., Belušić, D., Berthou, S., Caillaud, C., Cardoso, R. M., Chan, S., Christensen, O. B., Dobler, A., de Vries, H., Goergen, K., Kendon, E. J., Keuler, K., Lenderink, G., Lorenz, T., Mishra, A. N., Panitz, H.-J., Schär, C., Soares, P. M. M., Truhetz, H., and Vergara-Temprado, J.: The first multi-model ensemble of regional climate simulations at kilometer-scale resolution part 2: historical and future simulations of precipitation, Clim Dyn, 56, 3581–3602, https://doi.org/10.1007/s00382-021-05657-4, 2021.

Reyniers, N., Zha, Q., Addor, N., Osborn, T. J., Forstenhäusler, N., and He, Y.: Two sets of bias-corrected regional UK Climate Projections 2018 (UKCP18) of temperature, precipitation and potential evapotranspiration for Great Britain, Earth System Science Data, 17, 2113–2133, https://doi.org/10.5194/essd-17-2113-2025, 2025.

Roberts, N. M. and Lean, H. W.: Scale-Selective Verification of Rainfall Accumulations from High-Resolution Forecasts of Convective Events, Mon. Wea. Rev., 136, 78–97, https://doi.org/10.1175/2007MWR2123.1, 2008.

Robinson, E. L., Huntingford, C., Semeena, V. S., and Bullock, J. M.: CHESS-SCAPE: high-resolution future projections of multiple climate scenarios for the United Kingdom derived from downscaled United Kingdom Climate Projections 2018 regional climate model output, Earth System Science Data, 15, 5371–5401, https://doi.org/10.5194/essd-15-5371-2023, 2023.

Soares, P. M. M., Careto, J. A. M., Cardoso, R. M., Goergen, K., Katragkou, E., Sobolowski, S., Coppola, E., Ban, N., Belušić, D., Berthou, S., Caillaud, C., Dobler, A., Hodnebrog, Ø., Kartsios, S., Lenderink, G., Lorenz, T., Milovac, J., Feldmann, H., Pichelli, E., Truhetz, H., Demory, M. E., de Vries, H., Warrach-Sagi, K., Keuler, K., Raffa, M., Tölle, M., Sieck, K., and Bastin, S.: The added value of km-scale simulations to describe temperature over complex orography: the CORDEX FPS-Convection multi-model ensemble runs over the Alps, Clim Dyn, 62, 4491–4514, https://doi.org/10.1007/s00382-022-06593-7, 2024.

Themeßl, M. J., Gobiet, A., and Leuprecht, A.: Empirical-statistical downscaling and error correction of daily precipitation from regional climate models, International Journal of Climatology, 31, 1530–1544, https://doi.org/10.1002/joc.2168, 2011.

Trapp, R. J., Robinson, E. D., Baldwin, M. E., Diffenbaugh, N. S., and Schwedler, B. R. J.: Regional climate of hazardous convective weather through high-resolution dynamical downscaling, Clim Dyn, 37, 677–688, https://doi.org/10.1007/s00382-010-0826-y, 2011.

Weisman, M. L., Davis, C., Wang, W., Manning, K. W., and Klemp, J. B.: Experiences with 0–36-h Explicit Convective Forecasts with the WRF-ARW Model, Wea. Forecasting, 23, 407–437, https://doi.org/10.1175/2007WAF2007005.1, 2008.

Weusthoff, T., Ament, F., Arpagaus, M., and Rotach, M. W.: Assessing the Benefits of Convection-Permitting Models by Neighborhood Verification: Examples from MAP D-PHASE, Mon. Wea. Rev., 138, 3418–3433, https://doi.org/10.1175/2010MWR3380.1, 2010.

Wilcke, R. A. I., Mendlik, T., and Gobiet, A.: Multi-Variable Error Correction of Regional Climate Models, Climatic Change, 120, 871–887, https://doi.org/10.1007/s10584-013-0845-x, 2013.

Yun, Y., Liu, C., Luo, Y., Liang, X., Huang, L., Chen, F., and Rasmmusen, R.: Convection-permitting regional climate simulation of warm-season precipitation over Eastern China, Clim Dyn, 54, 1469–1489, https://doi.org/10.1007/s00382-019-05070-y, 2020.