

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 that 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 method testing in a new domain, 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 small and fast-responding catchments. 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 correction 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 identify the 1 km grid cells included in the analysis. 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 hydrologically meaningful 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 UKCPI8-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 UKCPI8-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 $9.3 \text{ }^\circ\text{C}$, based on the CEH-GEAR1hr and HadUK-Grid datasets, respectively.”

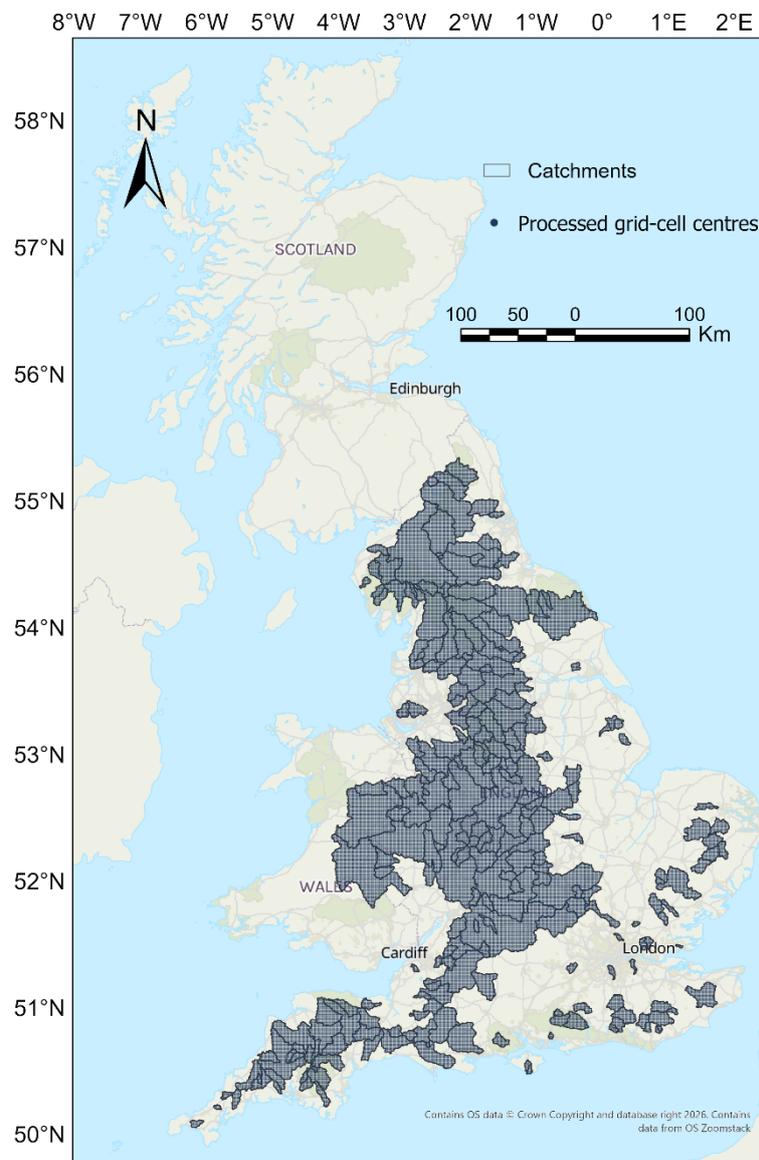


Figure 1: Study domain in England. Catchment boundaries (249 catchments from NRFA) are used to define the analysis mask. The processed 1 km HadUK-Grid cells are displayed by their centre points.

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 and 11.

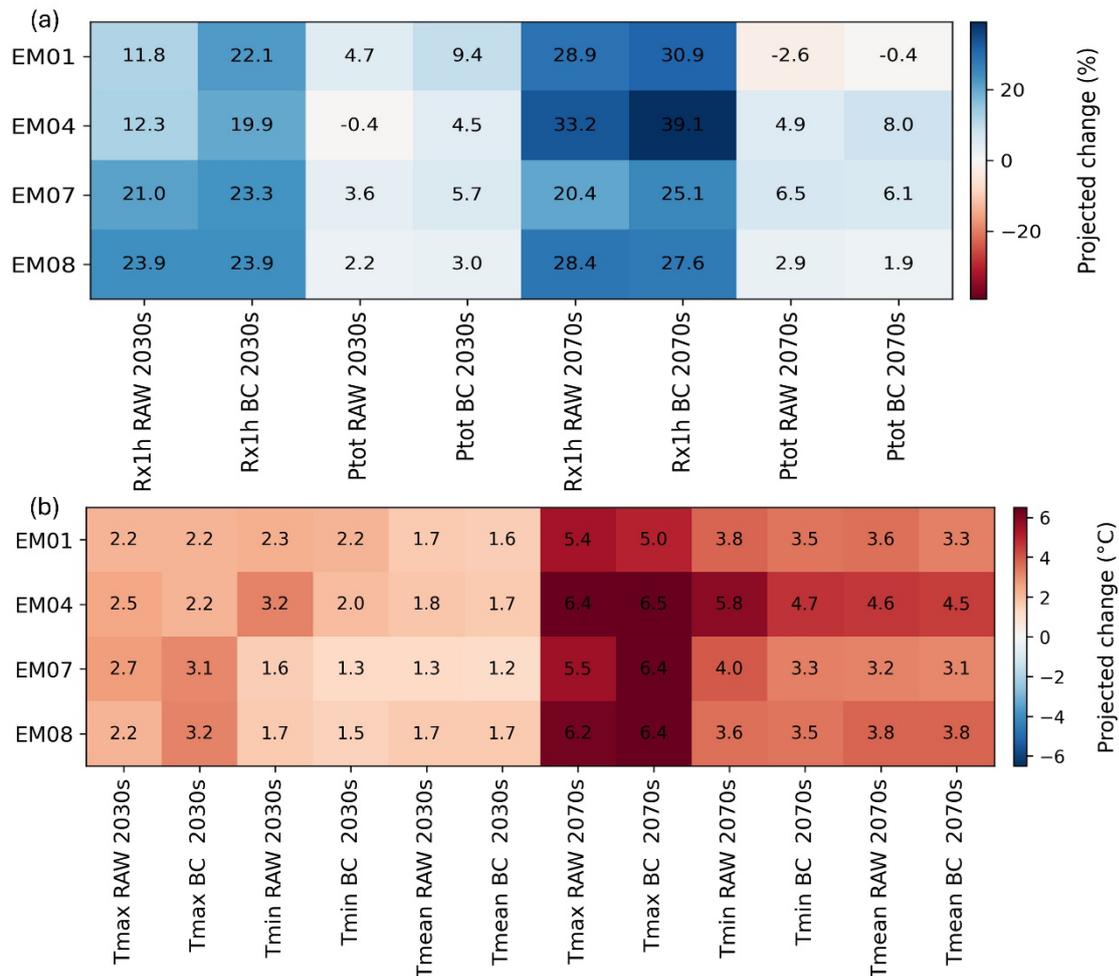


Figure 12: Projected changes in (a) annual maximum 1 h precipitation (Rx1h) and annual total precipitation (Ptot), and (b) annual maximum (Tmax), annual mean (Tmean), and annual minimum (Tmin) of daily mean temperature for UKCP18-CPM simulations before (RAW) and after bias correction (BC). Each value 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 localized 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 use 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: *Faghieh 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 line 163-175 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 $\geq 20 \text{ mm h}^{-1}$ 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 (P₉₅).

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), with three out of 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, a clear north–south gradient is evident, 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 CHE-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).

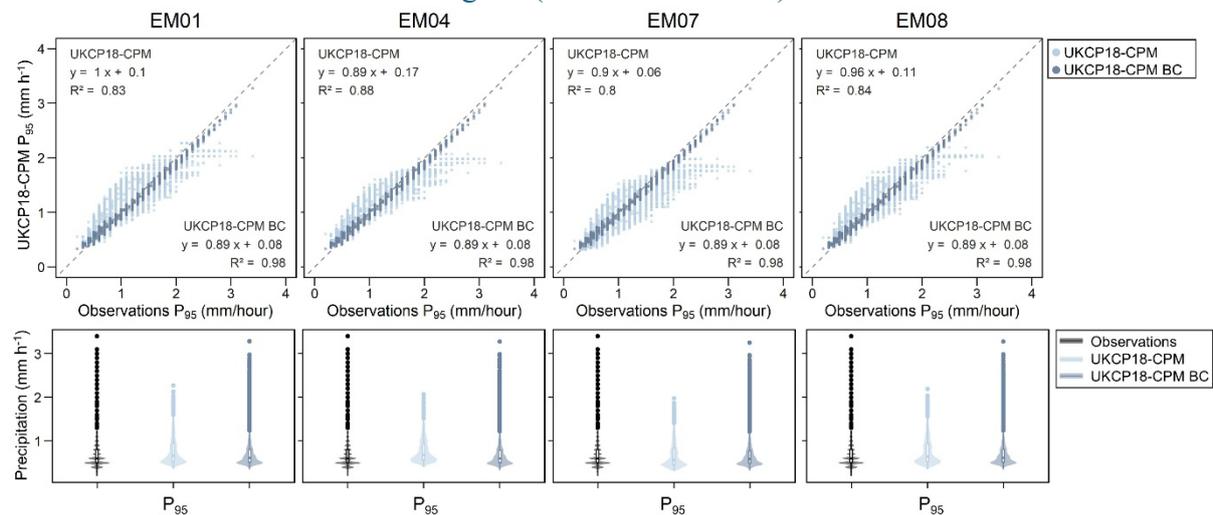


Figure 6: Comparison of 95th percentile (P₉₅) hourly precipitation values for the reference period. The top row shows scatter plots of UKCP18-CPM P₉₅ values for each ensemble member, plotted against CEH-GEAR1hr observations for each 1 km grid, with raw (light blue) and bias-corrected (dark blue) data. The bottom row presents violin plots of P₉₅ 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).

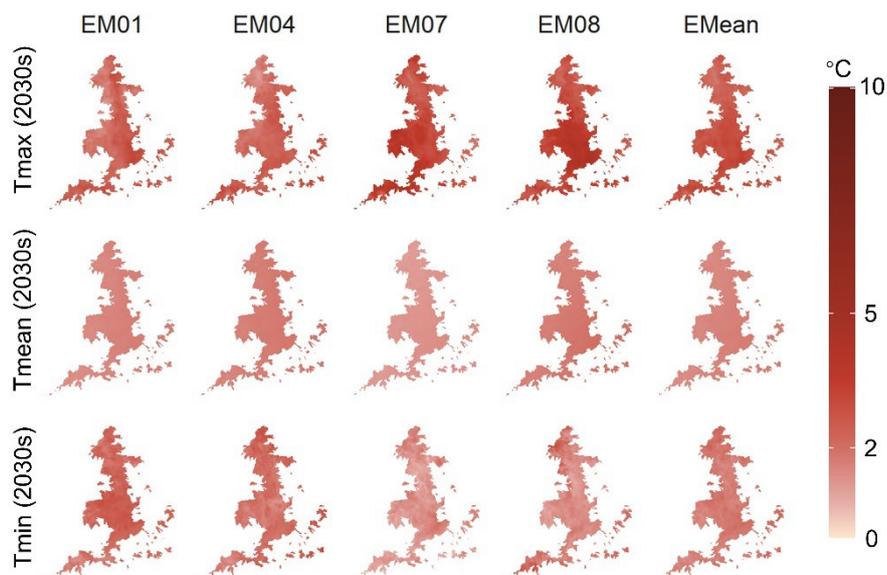


Figure 10: Projected changes in temperature (°C) for the 2030s (Dec 2020–Nov 2040) relative to the baseline period from the bias-corrected UKCP18-CPM simulations. Panels show (top to bottom) changes in the annual maximum (T_{max}), annual mean (T_{mean}) and annual minimum

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

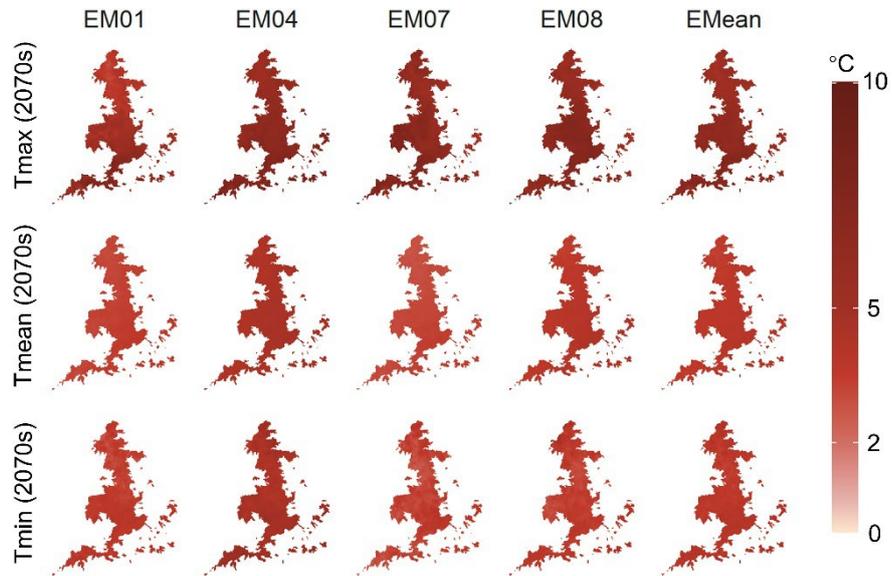


Figure 11: Projected changes in temperature (°C) for the 2070s (Dec 2060–Nov 2080) relative to the baseline period from the bias-corrected UKCP18-CPM simulations. Panels show (top to bottom) changes in the annual maximum (Tmax), annual mean (Tmean) and annual minimum (Tmin) of daily mean temperature,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

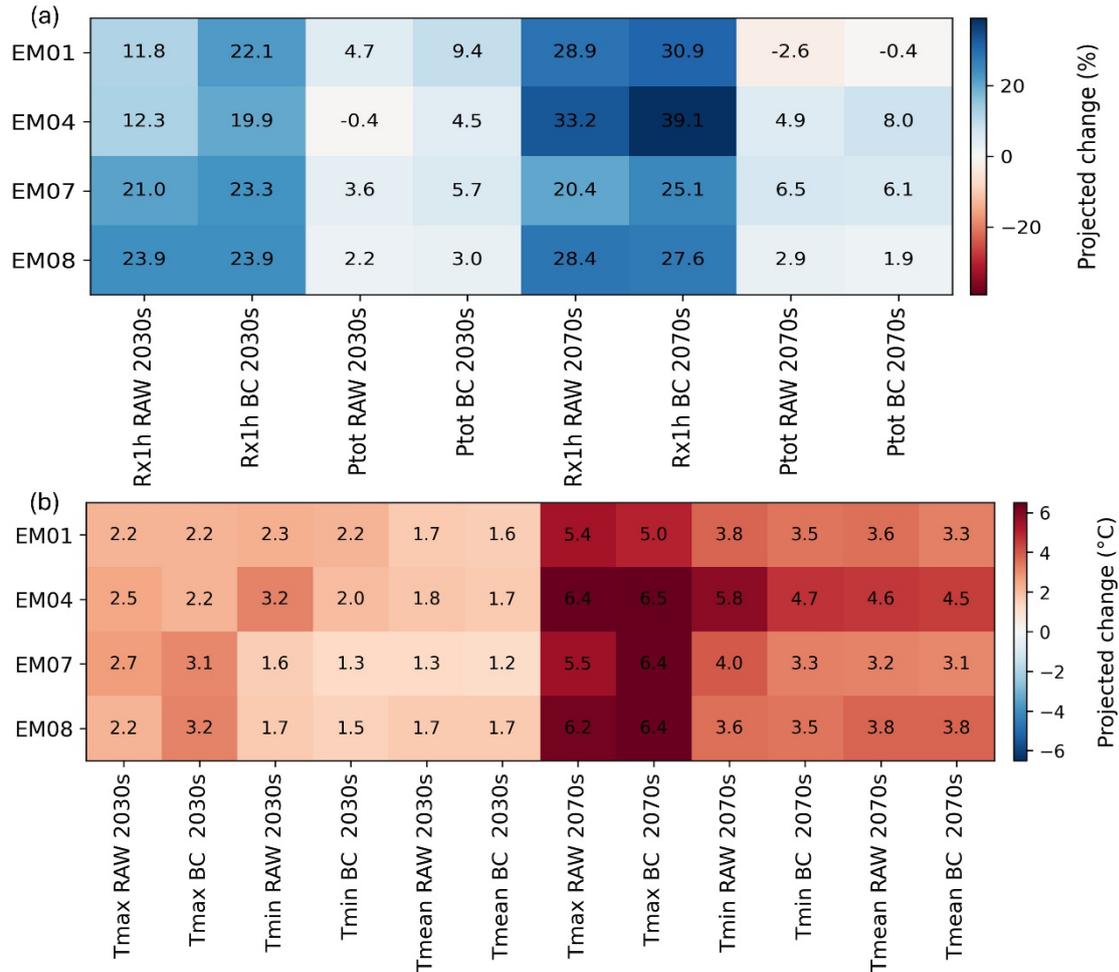


Figure 12: Projected changes in (a) annual maximum 1 h precipitation (Rx1h) and annual total precipitation (Ptot), and (b) annual maximum (Tmax), annual mean (Tmean), and annual minimum (Tmin) of daily mean temperature for UKCP18-CPM simulations before (RAW) and after bias correction (BC). Each value 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., Forstnhäusler, 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.: CHES-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 Rasmussen, 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.