Author's Response to the Referees

Manuscript ID: https://doi.org/10.5194/egusphere-2025-1112

Title: *Bias Correcting Regional Scale Earth System Model Projections: Novel Approach using Empirical Mode Decomposition*

We thank both the reviewers for their careful and constructive feedback. Point-by-point replies were originally uploaded as separate attachments (https://egusphere.copernicus.org/#AC4 and https://egusphere.copernicus.org/#AC5) in our earlier interactive comments. For completeness, we now merge those two rebuttal letters into a single, consolidated author-response file as requested. Reviewer remarks are set in **bold**; our responses follow each remark. When necessary, we use *italics* to indicate direct quotes from the manuscript, blue to mark text revised in response to comments from Referee 1, and green for text revised in response to comments from Referee 2.

Response to Reviewer 1

Thank you very much for your thorough and constructive review. We greatly appreciate the time and effort you invested in evaluating our work. Below we address each of your comments in turn. Reviewer remarks are set in **bold**; our responses follow each remark. When necessary, we use *italics* to demarcate quotes directly from the manuscript and *blue* marks revised text. We have also prepared a revised manuscript that addresses all of your comments and would be happy to provide it if needed.

This paper proposes a complex new way to do bias correction as a function of time scale. They claim that this technique is useful for impact analysis and improves on older simpler methods. But the paper does not provide any examples of this. It would greatly benefit from a couple examples where the new technique provides significantly improved understanding of the effects of climate change on a particular impact.

1. **Why does the paper use such old simulations to demonstrate the technique?**

   Response: While we acknowledge that this dataset was developed approximately a decade ago, it remains one of the high-resolution regional climate datasets available for North America and comparable to others (e.g., CORDEX; https://na-cordex.org/). It continues to be widely used by stakeholders and researchers, as demonstrated by its integration into publicly available platforms such as climRR (e.g., https://climrr.anl.gov/).

   Importantly, our goal in this study was to develop and demonstrate a new methodology for bias correction, rather than to evaluate future projections or explore new regional climate scenarios. The dataset we used provides a robust and validated testbed for this methodological development. In this context, the age of the dataset does not compromise the utility or rigor of our demonstration.

   That said, we fully recognize the importance of applying the method to more recent high-resolution simulations. Our team is currently developing a new 4 km convection-permitting regional climate dataset, and we intend to apply the proposed method to that product in future work. These simulations are undergoing evaluations for physical consistency, performance and other model evaluation criteria before we attempt any bias correction (Akinsanola et al., 2024).

2. **The paper seems to ignore the diurnal cycle. If you are going to bias correct based on time scales, this needs to be included. The daily maximum and minimum temperatures are very important for many impacts. Just using daily means is not sufficient for temperature.**

   Response: We appreciate the reviewer's insightful comment regarding the importance of the diurnal cycle. In this study, we focused on bias correction of daily mean temperature as a first step, largely due to limitations in the availability of spatially and temporally high-resolution observational datasets that cover the entire CONUS domain (e.g., Livneh, PRISM). These datasets

do not include sub-daily or hourly observations, which limits the feasibility of bias correction at finer temporal scales. Since our model domain spans the entire CONUS, it is important to use observational data that provides full spatial coverage, and currently, such datasets are only available at daily resolution.

With respect to daily maximum and minimum temperatures, we agree that applying bias correction to these variables is important for generating a more comprehensive and impact-relevant dataset. In this study, our primary objective was to demonstrate the methodological advantages of incorporating time-step-specific bias correction. For clarity and focus, we limited our main experiments to daily mean temperature. Nonetheless, our approach is directly applicable to the daily minimum and maximum temperature fields and users can extend the method to those variables without modifications beyond training separate models.

3. **And why is bi-weekly the shortest timescale? Many impacts depend on daily or subdaily timescales. Some extremes, which are very important, occur on these short time scales, including heatwaves and floods.**

Response: We appreciate this observation. Our framework employs empirical-mode decomposition (EMD) to split each temperature series into intrinsic mode functions (IMFs) of progressively lower frequency (see Fig. 1). We then form physically interpretable bands by summing subsets of these IMFs.

- The "bi-weekly band" aggregates IMFs with periods starting from a few days up to 14 days. However, as the observation data is only available at the daily level, a direct hourly (or finer) bias-correction step cannot be performed under the current data constraints.
- The "seasonal band" aggregates IMFs with periods between 14 days and 90 days, and so on for longer scales.

We chose the 14-day cutoff because it cleanly separates high-frequency "weather" variability from subseasonal fluctuations while leaving enough IMFs in each band for different timescale representations. The method itself is flexible: if higher-resolution observations become available, additional finer bands (e.g., hourly) can be introduced with the first few IMFs without altering the core algorithm. We have added the following clarification to the revised manuscript (Section 2.4.1, lines 180: *The bi-weekly band aggregates all IMFs with periods shorter than 14 days—thereby encapsulating the entire sub-daily to bi-weekly spectrum—while longer bands are formed by summing progressively lower-frequency IMFs.*):

4. **Figure 1 has several problems. It is colored, but there is no information about what the colors mean. The caption that a colorbar is not needed is absolutely wrong. What is the variable and how was it calculated? Also, there are no x- and y-axes. And the edges of the boxes are not discernible. There needs to be a table with each region and its location with the lat and lon of each edge.**

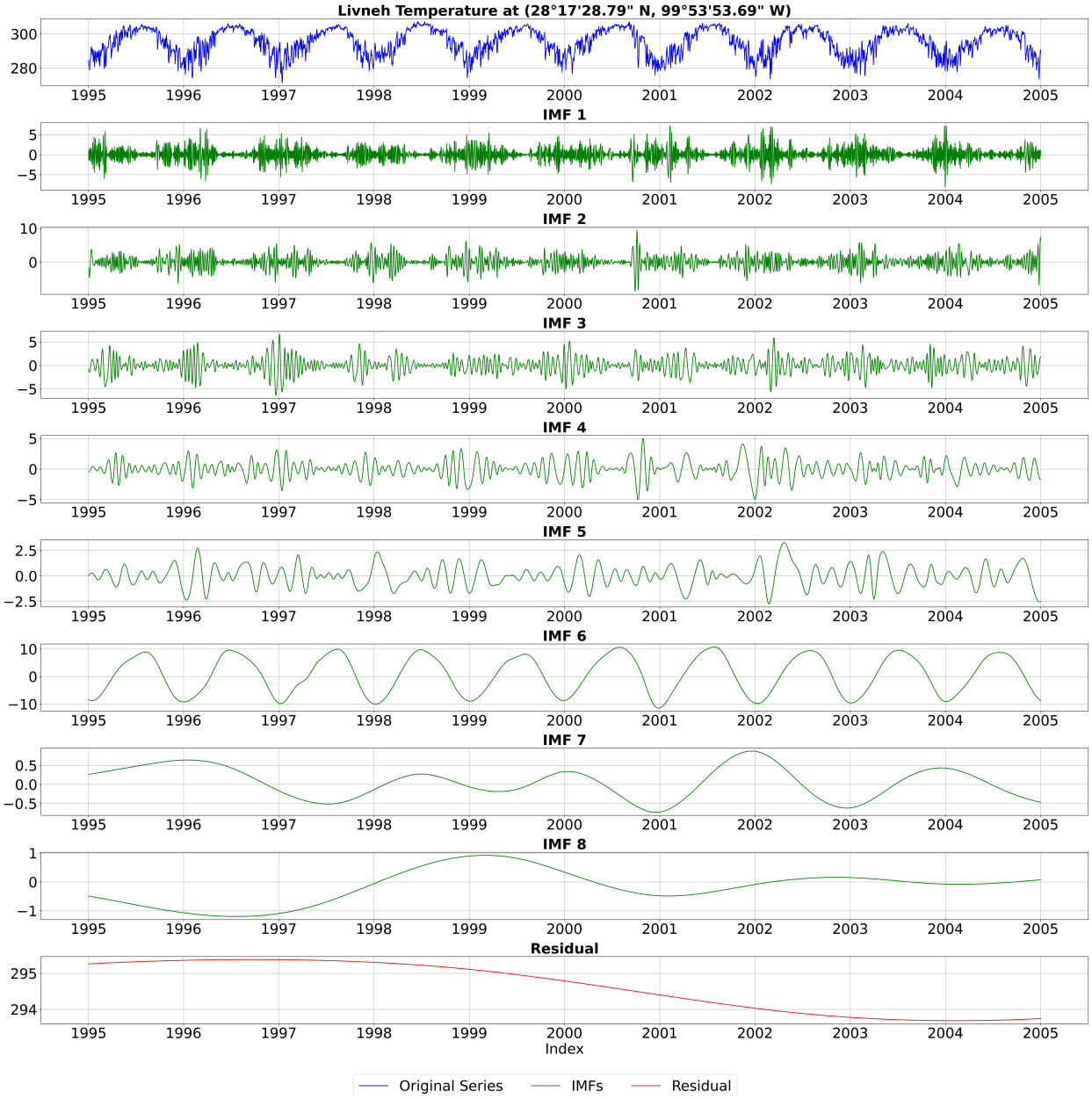

Figure 1: Empirical Mode Decomposition (EMD) of a representative daily temperature time series from Livneh observations (1995–2004). The top panel shows the original temperature signal. The following eight panels display the Intrinsic Mode Functions (IMFs), representing oscillatory components with progressively lower frequencies. The bottom panel shows the residual trend, capturing the long-term monotonic variation.

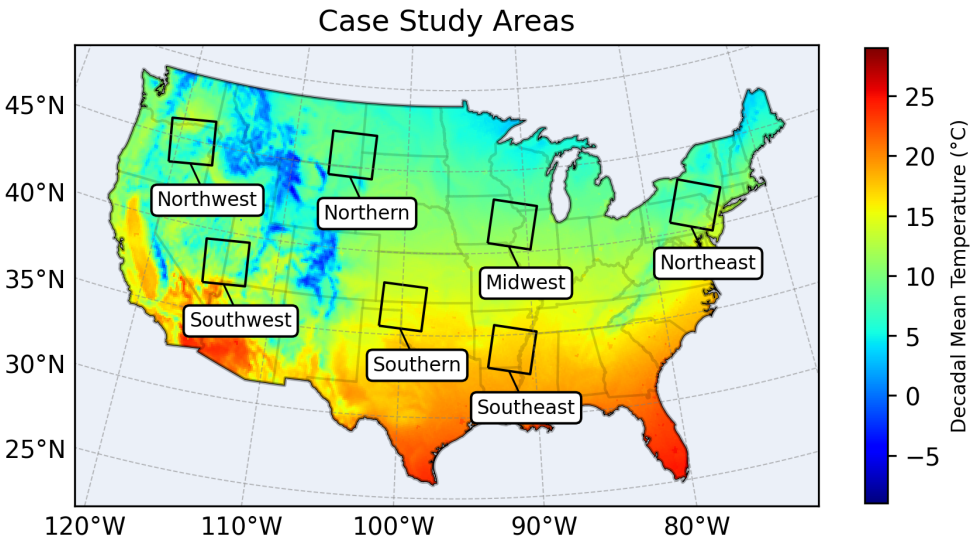

Figure 2: *Map of case study regions selected for evaluating the bias correction methods. Boundaries are overlaid on the average 1995-2004 mean temperature field from WRF-CCSM, illustrating the diverse range of temperature regimes captured by the case study areas.*

Response: We thank the reviewer for noting the problems in Figure 1. The figure has been updated by adding a color bar, labeling both axes, and overlaying grid lines to aid interpretation. Boundary vertices of each case study region were corrected to reflect the exact polygons used in the analysis, making each region visually distinct. The caption has been rewritten accordingly. The updated figure and caption is shown in Figure 11.

To ensure reproducibility, the appendix now includes a table that lists, for each case study region, the latitude/longitude coordinates of the bounding-box vertices together with their absolute array indices. The added table is shown in Table 1. The following text was added to the manuscript body to reflect this change: Lines 114: *The seven areas are geographically defined in Appendix A.*

5. **The paper has very many acronyms, and learning all of them makes the paper hard to read. And some of them are not defined at all. If the terms are not used multiple times, don't use acronyms and just write them out. And it would help to include an appendix with a list of the acronyms and their definitions.**

Response: Thank you for the suggestion. In the revised manuscript, we have included the following comprehensive table (Table 4 in the appendix that explains both the domain-specific acronyms as well as the mathematical notations and abbreviations used throughout the paper.

Region	$y_{\min}$	$y_{\max}$	$x_{\min}$	$x_{\max}$	UL Corner	UR Corner	LR Corner	LL Corner
Midwest	163	187	384	408	40.6233°N, 93.2427°W	40.1100°N, 89.8016°W	42.7208°N, 89.0243°W	43.2588°N, 92.6135°W
Northeast	190	214	482	506	40.6657°N, 78.3030°W	39.6782°N, 75.0555°W	42.1325°N, 73.6529°W	43.1648°N, 77.0141°W
Northern	189	213	292	316	44.4211°N, 106.7944°W	44.3514°N, 103.0425°W	47.0303°N, 102.8481°W	47.1034°N, 106.7843°W
Northwest	187	211	204	228	43.3865°N, 120.3843°W	43.7721°N, 116.7218°W	46.4336°N, 117.1932°W	46.0293°N, 121.0272°W
Southeast	95	119	394	418	33.0354°N, 93.4738°W	32.5457°N, 90.4126°W	35.1159°N, 89.7489°W	35.6294°N, 92.9269°W
Southern	110	134	330	354	35.4968°N, 101.6792°W	35.2683°N, 98.4510°W	37.9028°N, 98.0920°W	38.1425°N, 101.4546°W
Southwest	123	147	229	253	36.7126°N, 115.5168°W	36.9397°N, 112.2155°W	39.5947°N, 112.4409°W	39.3564°N, 115.8829°W

Table 1: *Bounding-box definitions for each case-study region. Columns $y_{\min}$, $y_{\max}$, $x_{\min}$, and $x_{\max}$ list the 0-based Python array indices that isolate the region within the WRF domain supplied with the dataset linked in the data availability statement; the remaining columns give the decimal-degree latitudes and longitudes of the four bounding-box corners: upper left (UL), upper right (UR), lower right (LR), and lower left (LL).*

Table 2: Acronyms and Symbols used in this study

Acronym / Symbol	Full Form	Brief Description (incl. equations)
CMIP6	Coupled Model Intercomparison Project Phase 6	Multi-model ensemble of coordinated global climate simulations.
GCM	Global Climate Model	Dynamical model representing physical processes of the climate system on a global grid.
RCM	Regional Climate Model	Higher-resolution model nested within a GCM to resolve regional detail.
BC	Bias Correction	Statistical adjustment applied to model output to align it with observations.
QM	Quantile Mapping	Bias-correction technique that remaps model quantiles to observed quantiles.
CDF	Cumulative Distribution Function	$F_X(x) = \Pr[X \leq x]$ for a random variable X.
QDM	Quantile Delta Mapping	Bias-correction method that preserves the modeled change signal while correcting quantiles.
EMD	Empirical Mode Decomposition	Data-adaptive decomposition that yields oscillatory components called IMFs.

Continued on next page

Table 2 (continued)

Acronym / Symbol	Full Form	Brief Description (incl. equations)
EEMD	Ensemble Empirical Mode Decomposition	Noise-assisted EMD variant that improves mode separation.
WRF–CCSM	Weather Research and Forecasting–Community Climate System Model	Dynamical downscaling chain coupling WRF with CCSM boundary fields.
IMF	Intrinsic Mode Function	Oscillatory component extracted by EMD, each with well-behaved local extrema.
EMDBC	EMD-based Bias Correction	Bias-correction framework that operates on time-scale-specific IMFs before reconstruction.
W_p (WD)	Wasserstein Distance	$W_p(P,Q) = \left(\inf_{\gamma\in\Gamma(P,Q)} \int_{\mathcal{X}\times\mathcal{X}} \|x - y\|^p \, \mathrm{d}\gamma(x,y)\right)^{1/p}$; where $\Gamma(P,Q)$ denotes the set of all couplings with marginals P and Q, commonly $p = 1$.
MSE	Mean Squared Error	$\mathrm{MSE}(y,\hat{y}) = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$; average squared deviation between predictions and observations.
FFT	Fast Fourier Transform	Algorithm that computes the discrete Fourier transform in $O(n \log n)$ operations.

6. **What about spatial scale? Usually bias correction is paired with downscaling. Are they independent? Does it matter in what order they are done?**

Response: The bias correction in our analysis has been applied after the regional downscaling step. We acknowledge that the order could influence the results, and in this case, the dynamical downscaling was performed first, followed by bias correction to adjust the downscaled output. We have conducted numerical experiments with bias corrected input fields (Wang and Kotamarthi, 2015). This process didn't always reduce the bias, particularly in near surface values. As a result, we have made the decision of conducting our downscaling simulations with inputs that are not bias corrected. In general, bias correction is often integrated into statistical downscaling, in the case of dynamical downscaling, it is commonly applied after the simulation. This approach directly addresses biases in the final model output at its resolved scale when using dynamical

downscaling. We acknowledge that there is no one method that is vastly superior to the other but this is the process we adopted.

Dynamical downscaling involves running a high-resolution regional model driven by a coarser GCM. While this method adds fine-scale physical processes and resolves local properties (e.g., terrain, land type, etc.), the RCM itself can introduce own systematic biases through its parameterized physics (e.g., microphysics, PBL, convective parameterizations). These RCM-specific biases are not necessarily present in the driving GCM and cannot be fully rectified by simply bias correcting the GCM input.

Given this, applying bias correction after downscaling directly targets biases inherited from the driving GCM as well as those inherent to the 12 km RCM output. This is crucial because these RCM-specific biases are distinct from the driving GCM's biases and significantly impact the regional-scale climate information relevant for many applications.

We will clarify these distinctions and the rationale for our sequencing within the revised manuscript to enhance its clarity and provide further context on our methodological choices.

7. **Figure 2 has too many panels with tiny font that is illegible. I can't read the subpanels in the first row. The text in the last row is too tiny. And what is this time series? Please explain where the data for the observations and model came from.**

Response: We thank the reviewer for noting the problems in Figure 2. The figure has been updated for improved readability by (1) increasing font sizes and (2) separating the "IMF" inset plot from row 1 into its own row. Additionally, the caption was updated to clarify the time series and data source. We would like to note that Figure 2 serves as a schematic of the EMDBC workflow, so specifics about the underlying data are not essential for illustrating the method. However, we agree that adding some additional details about where the data for the observation and model come from improves the clarity of the figure. We have added additional details in the caption to address this. The updated figure and caption is shown in Figure 11.

8. **Line 203 says, "At longer timescales (seasonal or annual), biases often manifest in more systematic patterns that persist across multiple years." but the authors provide no evidence for this. Why do you think the climate system would behave this way? Please provide references.**

Response: Thank you for flagging this point. In our experiments we consistently find that short-term IMFs ($<$ 14 days, i.e., the bi-weekly band) oscillate rapidly around zero, whereas seasonal and annual IMFs drift smoothly and keep the same sign for months to years—clear evidence of a persistent, systematic bias. To understand the temporal dependency across different timescale representations of our daily data, we calculated autocorrelation at different lags from 1-365. This analysis shows that the biweekly bias consistently exhibits low autocorrelation, confirming its weak temporal dependence and justifying a non-parametric treatment, whereas

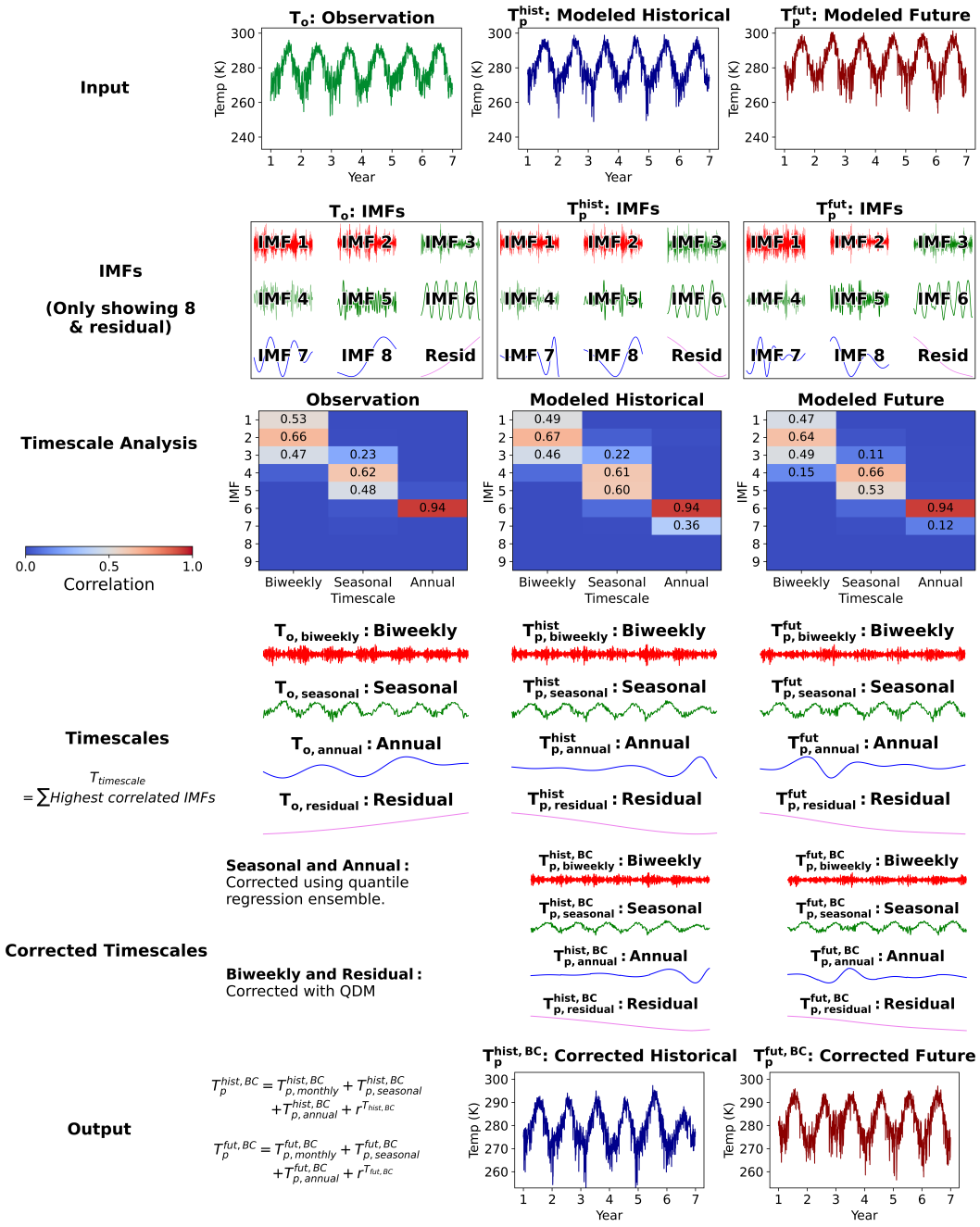

Figure 3: *Timescale-wise bias correction framework using EMDBC. Three inputs are required: temperature timeseries from observation, modeled historical, and modeled future datasets. Here, to demonstrate EMDBC, timeseries data are extracted from Livneh (T_o), WRF-CCSM historical (T_p^{hist}), and WRF-CCSM mid-century (T_p^{fut}) at an arbitrary location. The input temperature series are decomposed into IMFs using EEMD. IMFs are then classified into predefined timescales: biweekly, seasonal, and annual. Bias correction is applied using QDM for biweekly timescale and residuals, and quantile regression for seasonal or annual timescales. Finally, the corrected timescales are summed to reconstruct the bias corrected temperature series.*

for longer timescales (seasonal and annual), the autocorrelation metrics are much higher. Figure 20 shows this behaviour at a representative grid cell.

This observed behavior, where longer-term IMFs exhibit greater persistence and systematicity, while shorter-term IMFs are less autocorrelated and more stochastic, is consistent with established understanding of climate model uncertainties and errors. For example, Hawkins and Sutton (2009, 2012) demonstrate that while internal climate variability dominates uncertainties in near-term projections (reflecting a more stochastic and less predictable component), model uncertainty becomes increasingly dominant over longer timescales, indicative of persistent, systematic differences in model representations.

Internal variability (e.g., weather, chaotic atmospheric fluctuations) operates on short timescale and tends to average out over longer periods, reducing its relative contribution to bias. As lead time increases the random fluctuations from internal dynamics become less dominant. In contrast, what becomes prominent is the underlying, systematic bias that arises directly from the model's fixed structural shortcomings in representing physical processes at longer timescales. These are the persistent errors built into the model's physics and parameterizations. Because the source of these biases remains persistent, their effect accumulates and persists over longer periods. The autocorrelation of these biases is high because the reason for the bias doesn't change from one season or year to the next. This fundamental difference in bias characteristics at varying timescales directly informs and justifies our multi-timescale bias correction approach.

We have added this explanation at line 203 of the revised manuscript: *With daily-resolution data, longer timescales like seasons or years appear more structured, while shorter timescales show less pattern.*

9. **The paper uses "validation" in many places where it should be "evaluation." Please correct these. Validation means that you already know the results are valid.**

Response: In our manuscript we use "validation" in its standard machine-learning sense, not as a synonym for "evaluation of something already proven correct." Specifically:

- Training vs. validation splits – We divided the historical record into two non-overlapping windows: 1995 – 1999 for model development (training) and 2000 – 2004 for validation. The latter window is completely unseen during the fitting, so the performance metrics computed there estimate the generalization error of the model.
- Evaluation – Metrics calculated on the training window (or when we do not know the observed series, e.g., the mid and late century projections) are referred to as evaluations, because the data were used in optimization and the results primarily reflect in-sample fit.

Thus, wherever the manuscript discusses out-of-sample performance on the 2000 – 2004 window, we intentionally use "validation." We have double-checked the text to ensure the term "evaluation" is reserved for in-sample diagnostics, while "validation" appears only for results on the held-out split.

10. **Figures 3 and 4 are missing units. And what is Wasserstein?**

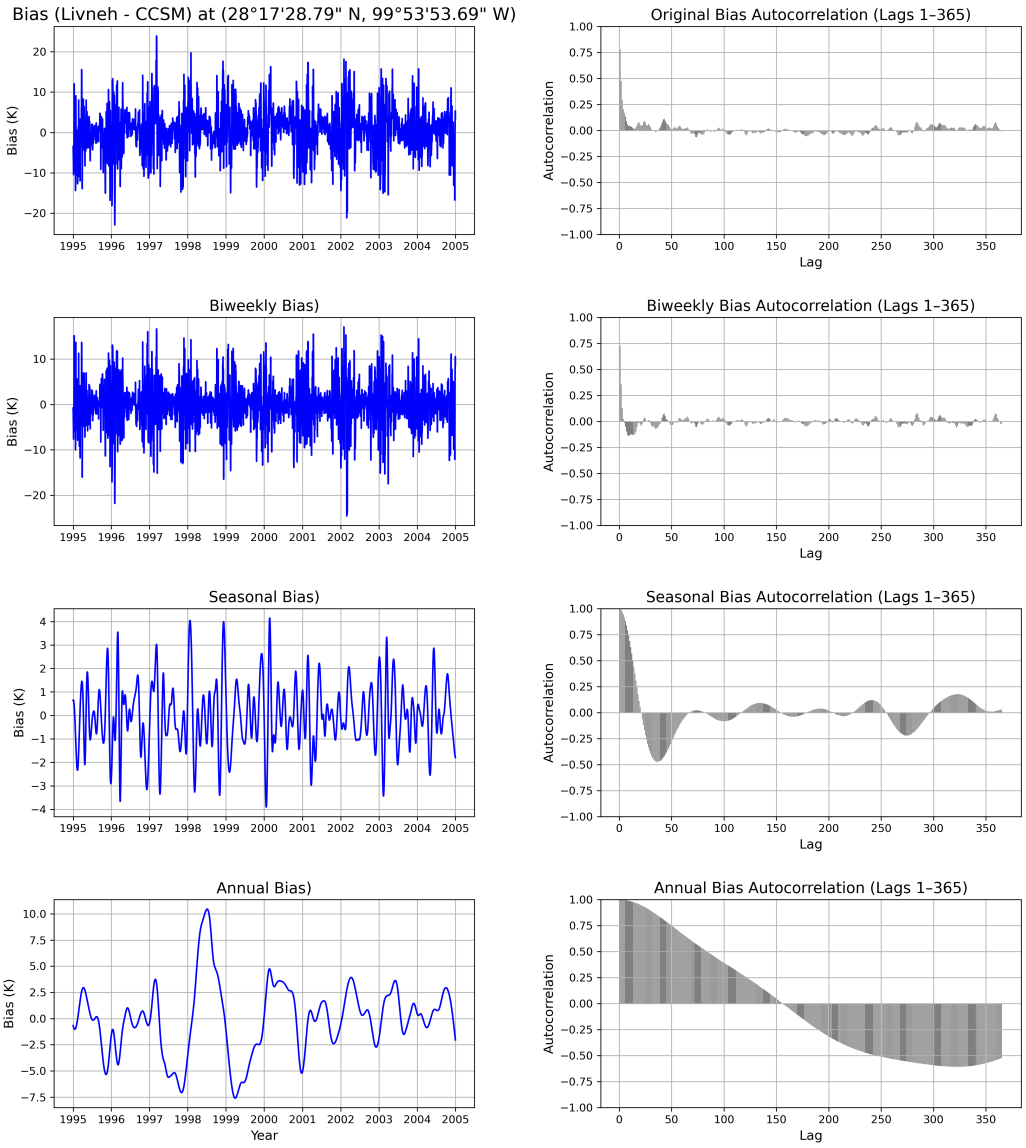

Figure 4: Bias decomposition by time scale for a representative grid cell. The first plot, labeled "Original," shows the daily bias (Livneh – CCSM). Empirical Mode Decomposition (EMD) is used to group intrinsic mode functions (IMFs) into three aggregated bands: biweekly, seasonal, and annual. Autocorrelation values at different lags are shown in the second column to quantify temporal dependence for multiple timescales extracted from daily samples. Notably, the biweekly bias exhibits much lower autocorrelation, indicating weaker time-dependency compared to the seasonal and annual bands.

Response: Thank you for noticing these omissions. We have updated Figures 3 and 4 to include the appropriate units on all axes and color bars.

Regarding "Wasserstein," we are referring to the first-order Wasserstein distance (also called the Earth-Mover's Distance), a metric that quantifies the minimal "cost" of transforming one probability distribution into another. It is well suited for comparing full distributions rather than individual moments and is therefore a natural choice for our quantile-based bias-correction evaluation. A more detailed description can be found at Panaretos and Zemel (2019). The following brief definition and formula have been added to the Appendix, and the term is now introduced at first mention in the main text:
WD is defined as a distance between two probability measures P and Q on a metric space $(\mathcal{X}, \|\cdot\|)$ by

$$W_p(P,Q) = \left(\inf_{\gamma \in \Gamma(P,Q)} \int_{\mathcal{X}\times\mathcal{X}} \|x-y\|^p \, \mathrm{d}\gamma(x,y) \right)^{1/p}, \tag{1}$$

where $\Gamma(P,Q)$ denotes the set of all couplings with marginals P and Q; throughout this study we use the common choice $p = 1$ (Panaretos and Zemel, 2019).

11. **The first three panels in Fig. 5 just show lots of colored lines, and it is impossible to compare them. The last row only has 4 lines, but the colors are similar and again it is hard to figure out which line is which. And what is MSE and what are its units? Why are the last three panels centered on 0? Are these anomalies? With respect to what?**

Response:

(a) Improved visual clarity: We have reorganised Figure 22 into two side-by-side columns. *Column 1* retains the full 1995–2004 record for context, while *Column 2* zooms into 1999–2001 so individual trajectories can be inspected easily. Livneh and CCSM are plotted with solid lines, and the two bias-corrected series—EMDBC and QDM—are plotted with dashed lines, making each data product visually distinct.

(b) Mean-squared error (MSE): Because individual curves remain hard to separate in the full-record view, we quantify performance with the mean-squared error between each series and Livneh. The caption now supplies the definition and units:

$$\mathrm{MSE}(x,y) \;=\; \frac{1}{N}\sum_{t=1}^{N}\big[x(t)-y(t)\big]^2,$$

where x is the bias-corrected (or raw) model series, y is Livneh, and N is the number of days; While QDM matches EMDBC at the native daily scale, EMDBC achieves lower MSE at seasonal and annual scales—evidence that it better preserves large-scale temperature behaviour.

(c) Why the lower-row curves are centred on zero: The bottom three panels display the bi-weekly, seasonal, and annual reconstructions obtained by summing the corresponding in-

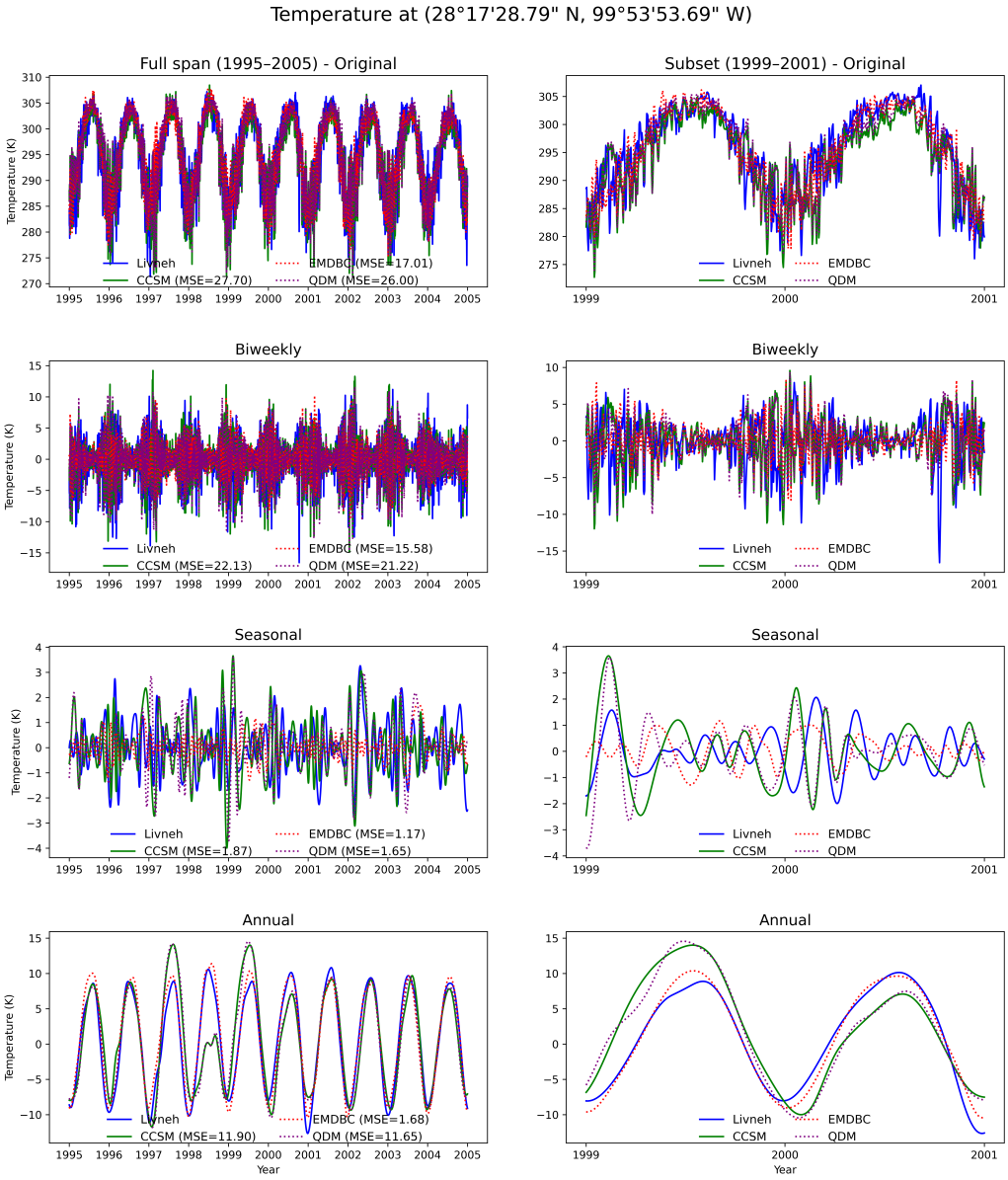

Figure 5: Bias-correction comparison across multiple time-scales at a representative grid cell. Column 1 shows the full 1995–2004 record, while Column 2 zooms into 1999–2001 for clarity. Solid lines correspond to the observed Livneh series (blue) and the raw CCSM projection (green); dashed lines show the bias-corrected outputs from EMDBC (red) and QDM (purple). Each row presents the original daily series and its bi-weekly, seasonal, and annual components, obtained by aggregating intrinsic mode functions as described in Section 2.4. The numbers at right report the mean-squared error (MSE, $^{\circ}K^2$) between each series and Livneh. While QDM matches EMDBC at the native daily scale, EMDBC yields consistently lower MSE at the bi-weekly, seasonal, and annual bands, indicating superior preservation of large-scale temperature variability.

trinsic mode functions (IMFs). Each IMF is zero-mean by construction, so their aggregates are also zero-mean; the curves are therefore centred on 0 and do *not* represent anomalies relative to an external baseline.

12. **For the example in Fig. 5, the largest differences in the annual temperature are just for three years. Why, and how does this skew the average scores?**

Response: For this sample location, the largest deviations at the annual scale occur in three years. However, they still raise the mean-squared error (MSE) across every time scale—exactly the kind of behaviour a time-scale-aware bias-correction scheme should fix. EMDBC lowers the error in each of those high-bias years and keeps errors small in the remaining years, thereby reducing the overall MSE. By contrast, QDM leaves larger residual errors at the annual scale, so its average MSE remains higher.

More broadly, a successful bias-correction should reduce error *consistently across all time scales*. EMDBC is designed for that purpose:

- It maintains the observed structure in the bi-weekly, seasonal, and annual bands (Figure 22, lower rows).
- Consequently, it also achieves a lower MSE than QDM at the original daily scale (Figure 22, top row).

Thus the few extreme-error years do not "skew" the evaluation; instead, they provide a stringent test, underscoring EMDBC's ability to correct large, infrequent biases without sacrificing day-to-day fidelity.

13. **Why are some plots semi-annual and others annual?**

Response: Thank you for pointing this out. We agree that presenting some plots at a semi-annual scale while others are shown annually could introduce representational inconsistency. Our original intention was to demonstrate that the proposed timescale-aware bias correction method is effective across different temporal scales. However, for consistency and clarity, we have revised the manuscript to replace the semi-annual plot with an annual version. The updated figure is shown below as Figure 6.

14. **In Fig. 6 in the last row, I can't tell any differences in the distributions. Are the differences really significant?**

Response: The purpose of the bottom row in Fig. 6 is to demonstrate that both methods preserve the overall distribution of temperatures. Because QDM is designed to map model quantiles directly onto observed quantiles, its post-correction curve almost perfectly overlays the Livneh distribution—hence the lines look identical. EMDBC achieves the same distributional match while simultaneously lowering the mean bias (top row).

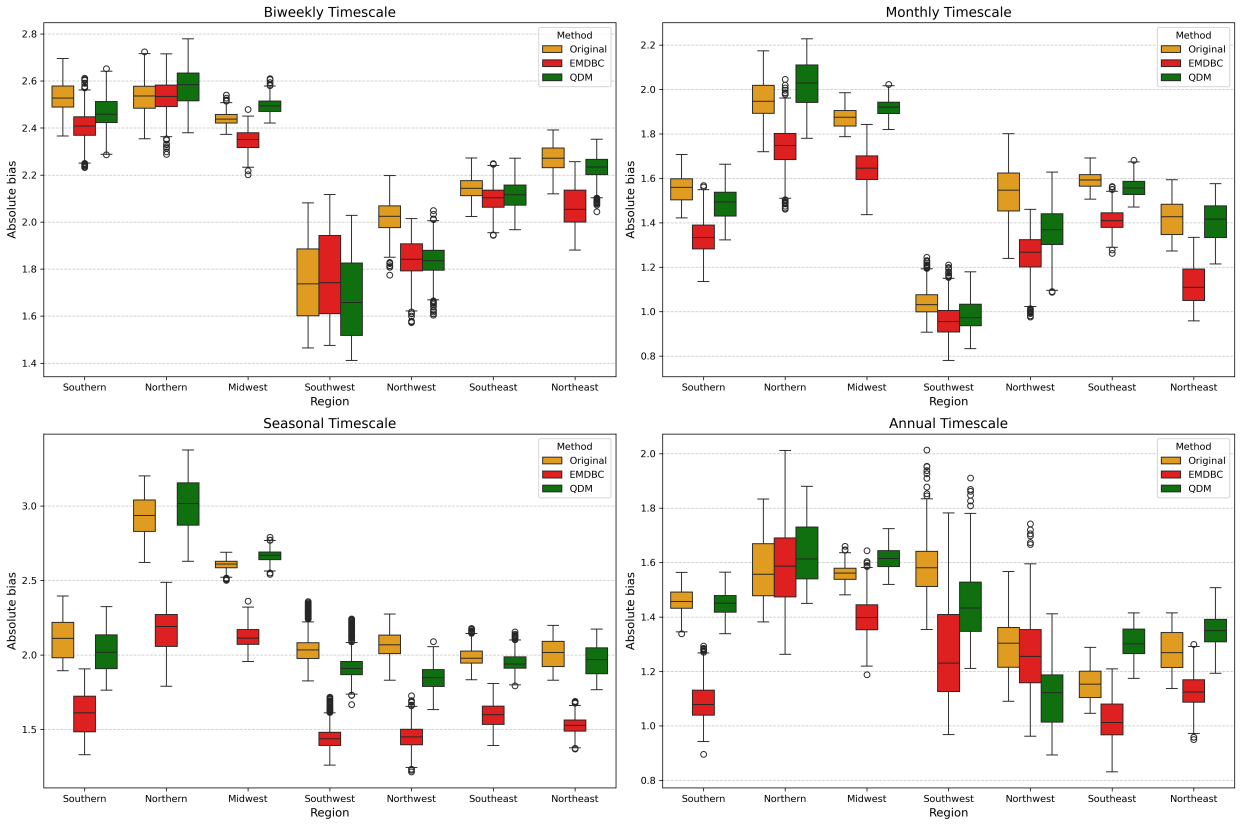

Figure 6: The timescale wise average absolute bias per subregion on the validation dataset.

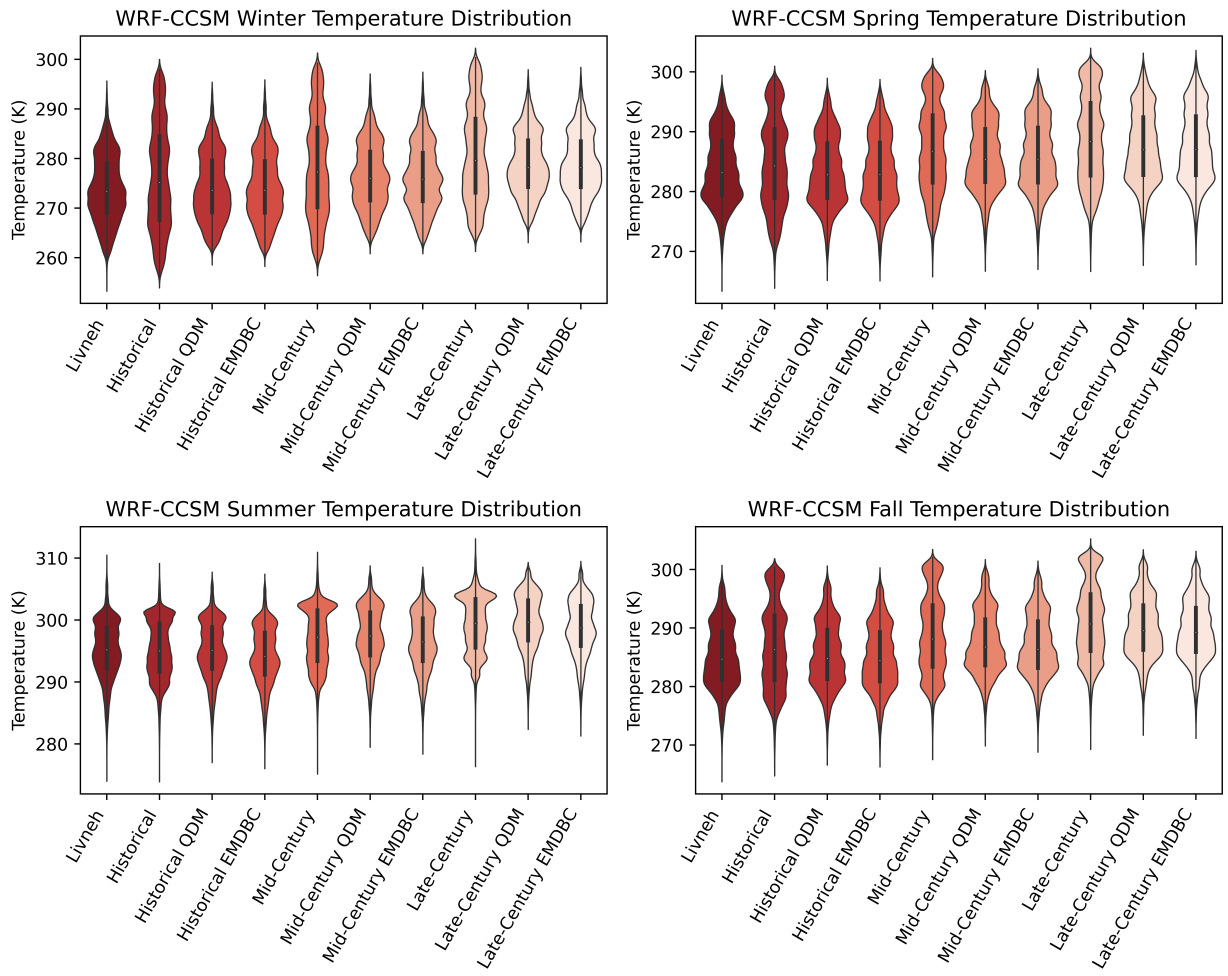

Figure 7: *The mean daily temperature by season (winter, spring, summer, fall) across Livneh (1995-2004) and WRF-CCSM historical (1995-2004), mid-century (2045-2054), and late-century (2085-2094) timeframes before and after bias correction. Results for QDM and EMDBC are included.*

15. **Figs. 7-9: There is not enough information to understand what is plotted? What are the exact time periods? What are the sources of the data?**

Response: We acknowledge that the captions for Figures 7–9 did not explicitly repeat the time windows and data sources, which may impede readers who consult the figures before the Methods section. We have therefore revised each caption to include: (1) the exact observational and model data sets and (2) the temporal periods used.

No changes were made to the plots themselves; the added detail simply mirrors the information already provided in Section 2.1 ("Data"). We believe this clarification makes the figures fully self-contained. The updated figure captions are shown in Figures 16, 17, and 18, with revisions shown in blue.

16. **Also, any response has to address the 30 comments in the attached annotated**

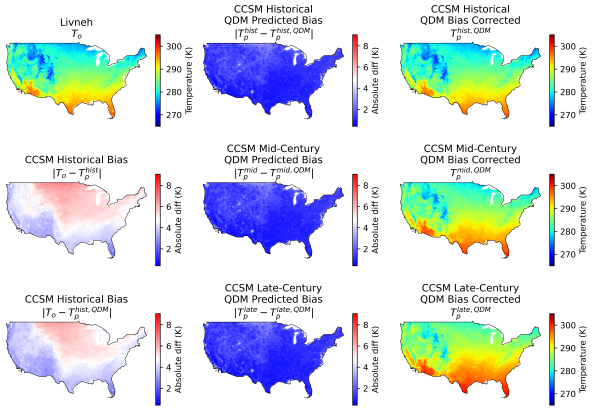

Figure 8: *Temperature and temperature bias comparisons over CONUS before and after applying QDM.* ***Left:*** *Observed temperature (Livneh, 1995-2004), WRF-CCSM historical (1995-2004) average daily absolute bias, and QDM-corrected WRF-CCSM historical average daily absolute bias.* ***Middle:*** *Magnitude of QDM correction in historical, mid-century (2045-2054), and late-century (2085-2004) timeframes.* ***Right:*** *QDM-corrected temperatures for WRF-CCSM historical, mid-century, and late-century periods.*

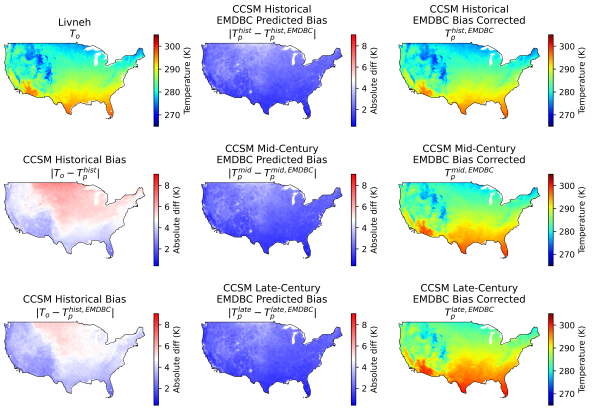

Figure 9: *Temperature and temperature bias comparisons over CONUS before and after applying EMDBC.* ***Left:*** *Observed temperature (Livneh, 1995-2004), WRF-CCSM historical (1995-2004) average daily absolute bias, and EMDBC-corrected WRF-CCSM historical average daily absolute bias.* ***Middle:*** *Magnitude of EMDBC correction in historical, mid-century (2045-2054), and late-century (2085-2004) timeframes.* ***Right:*** *EMDBC-corrected temperatures for WRF-CCSM historical, mid-century, and late-century periods.*

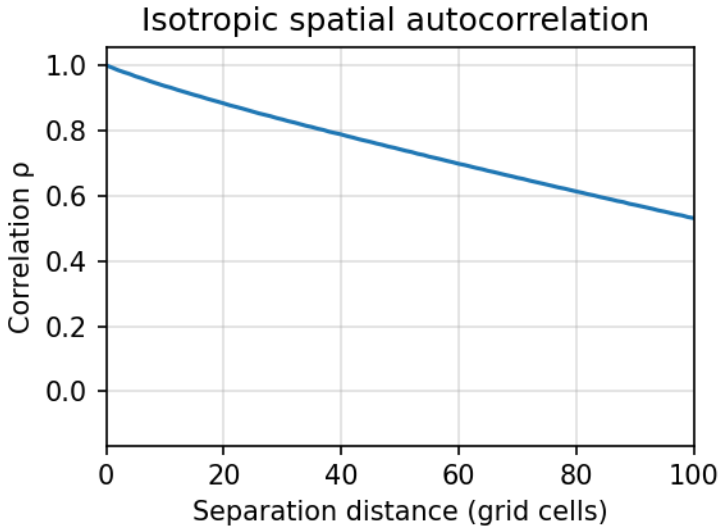

Figure 10: Isotropic spatial autocorrelation of daily surface temperature from the CCSM model at a representative time point. For each distance r, we select all pairs of grid points exactly r cells apart, align their anomaly values into two vectors, and compute a single Pearson correlation between those vectors. The strong correlation at short distances and its gradual decay with increasing separation illustrate the spatial coherence characteristic of temperature fields.

manuscript

Response:

(a) Comment on: *" Since temperature typically exhibits strong spatial coherence"*>> On what time and space scale? Show data to prove this.

The statement "Since temperature typically exhibits strong spatial coherence" is grounded in well-established physical understanding: surface temperature is influenced by spatially continuous factors such as elevation, vegetation, soil moisture, and land use, all of which shape the local surface energy balance. As a result, neighboring regions tend to exhibit similar temperatures, leading to spatially coherent temperature fields. This understanding is also supported by recent literature; for example, Kunz and Laepple (2024) describe a "large-amplitude short-distance component" in the spatial correlation function of temperature fields, indicating strong local spatial correlation.

As an empirical validation that the CCSM model realistically captures this expected spatial behavior, we have added a spatial autocorrelation analysis using a representative snapshot of the model's surface temperature field. Specifically, we compute the isotropic spatial autocorrelation, which measures how correlated the temperature is between locations separated by a given distance, averaged over all directions and presented in Figure 21. The line shows how correlated two grid cells are at increasing radius: for every possible pair of cells, we calculated their temperature correlation, grouped the pairs by distance, averaged

those correlations, and plotted the result. Because the correlation stays well above 0.8 for up to 40 grid cells away (12*40=480 km), it justifies our use of local smoothing.

(b) All the other typos are corrected in the revised manuscript.

Table 3: Reviewer highlights and responses

Highlighted text	**Reviewer comment**	**Our response**
Systems	System	Typo corrected.
systems	system	Typo corrected.
(EMD)	Don't define acronyms you don't then use	Eliminated the acronym from the abstract.
systems	system	Typo corrected.
Abstract	What about spatial scale? Usually bias correction is paired with downscaling. Are they independent? Does it matter in what order they are done?	Discussed in our response to 6.
CMIP6	define	Defined on line 15 and acronym defined in appendix acronym table.
(NA-CORDEX)	(NA-CORDEX)	Although the acronym is used only at its first mention, we retain it because most domain scientists know this dataset exclusively by that shorthand.
QM	The QM	Typo corrected.
reflects	reflect	Typo corrected.
validated	evaluated	Corrected
EEMD	define	Defined on line 69 and acronym defined in appendix acronym table.
systems	system	Typo corrected.
moodeled	moodeled	Typo corrected.
s1	s-1	We thank the reviewer for pointing out this typo. Though our manuscript source code had the proper notation, it was not being rendered properly in the PDF document. The text was corrected: Line 91: *... using a nudging coefficient of* $3 \times 10^{-5} s^{-1}$
leverage	???	We acknowledge the awkward word choice. We replaced "leverage" with "used."

Continued on next page

Table 3 – continued from previous page

Highlighted text	Reviewer comment	Our response
Daily mean temperature	What about the diurnal cycle? Why are you ignoring it? Certainly daily max and min T are important variables for impact analysis.	Discussed in our response to 2.
is	are ["data" is plural]	Typo corrected.
projected onto	What does this mean? How are the different grids combined?	We acknowledge the inadequate explanation regarding how the Livneh dataset was remapped onto the WRF-CCSM grid. We revised the text appropriately (our revisions in blue): Lines 102-103: *Similarly, the 1/16 degree Linveh data is remapped onto the 12-km simulation mesh used by WRF-CCSM using the bilinear interpolation operator provided in the Climate Data Operator software version 2.5.2 (Schulzweida, 2023) to match the spatial scales.*
25x25	What does this mean? What are the units?	25x25 refers to computation cells used in the WRF model simulation. We have clarified this in the manuscript: (Lines 112-113): *For a comprehensive spatial evaluation, we randomly selected seven areas, each measuring 25×25 grid cells (300 km × 300 km), from major subregions across the continental United States, shown in Figure 1, ensuring a diverse set of conditions.*
Map	Need axes showing lat and lon, especially for the regions.	Discussed in our response to 4.
(no colorbar needed	Absolutely wrong. What are the colors? And for what specific period?	Discussed in our response to 4.
CDF	What is this? It needs to be defined	Defined on line 46 and acronym defined in appendix acronym table.
validation	validation	Discussed in our response to 9.
same as above	evaluation	Discussed in our response to 9.

Continued on next page

Table 3 – continued from previous page

Highlighted text	**Reviewer comment**	**Our response**
strong spatial coherence	On what time and space scales? Show the data to prove this.	Discussed in our response to 16(a).
absolute temperature bias	Over what time period? And what time scale? And what are the units?	Updated the Figure 3 caption with the relevant timescale as *.. the original daily timescale* and the unit as *in K*
Wasserstein distance	What is this? And what are the units?	Updated Lines 260-263 with the definition and add in the appendix acronym table
doesn't	does not	Corrected
2.1.Figure	2.1. Figure	Corrected
Wasserstein distance	What is this?	Updated Lines 260-263 with the definition and add in the appendix acronym table
FFT	???	Defined on line 266 and acronym defined in appendix acronym table.

We trust that our revisions and clarifications satisfactorily address your concerns and significantly improve the manuscript. Thank you again for helping us strengthen this work.

Response to Reviewer 2

Thank you very much for your thorough and constructive review. We greatly appreciate the time and effort you invested in evaluating our work. Below we address each of your comments in turn. Reviewer remarks are set in **bold**; our responses follow each remark. When necessary, we use *italics* to indicate direct quotes from the manuscript, blue to mark text revised in response to comments from previous reviewers, and green for text revised in response to comments from Referee 2. We have also prepared a revised manuscript that addresses all of your comments and would be happy to provide upon request.

The authors have combined various existing bias correcting methods of the literature to account for timescale-aware bias corrections. They compared a model to a set of observations on a historical period and propagated the bias correction to mid- and long-term predictions. For some timescales, this bias correction method improves upon existing methods, on one specific region studied here and for an atmospheric model. The paper reads well and has many figures illustrating the results.

However, the text and the figures should be improved to clarify the method, the results and their performance. Thank you.

Here are some general comments:

1. **Please clarify the novelty of the method: combining several methods (name them) into one framework?**

   Response: **Methodological novelty - timescale-aware bias correction:** Our key innovation is to make bias correction explicitly timescale-aware by separating a model signal into physically meaningful bands and correcting each band individually before reconstruction. This idea has only been hinted at in previous work (e.g., Haerter et al. (2011)) and, to our knowledge, has not been operationalized within a complete framework. The approach is agnostic to the downstream correction tool: any standard univariate scheme can be slotted in, and a multivariate extension is straightforward via multivariate EMD-type decompositions, which would preserve inter-variable dependencies in the bias-corrected outputs. This method is a significant advance because it accounts for bias in phenomena that operate at different time-scales (mesoscale, synoptic, seasonal, etc.). Doing so preserves the added value of the dynamical downscaling at the scales where it is accurate, while still correcting errors that originate from the driving model or from imperfect physics at other scales.

   Application novelty - continental-scale evaluation across multiple timescales We test the framework on WRF-CCSM temperature fields over North America and show that it delivers larger bias reductions and closer agreement to the observed temperature distribution than both raw model output and a state-of-the-art Quantile Delta Mapping (QDM) benchmark. These gains—quantified by lower mean-squared error and smaller Wasserstein distance—are consistent across multiple timescales (biweekly, monthly, seasonal, and half-annual), demonstrating the practical value of addressing biases at their native timescales.

2. **To improve the performance study it would be helpful to broaden the evaluation: why is a regional dataset enough (USA)? what about using a larger spatial extent (global instead of USA)? Can the method be compared to other methods (mentioned in the Introduction) to illustrate its performance? If only one model is used in the end, add name of the model in the title.**

   Response: **Regional Focus (CONUS):** We agree that a broader spatial evaluation is an important direction for future work. We focus here on the challenge of correcting bias in a dynamically downscaled Earth system model projections without losing the value added by dynamical downscaling. Thus in this study, we used a high-resolution observational dataset (i.e., Livneh: 1/16 degree spatial resolution) that only covers the CONUS domain and a high resolution dynamically downscaled model output. This dataset provides daily meteorological fields at 1-km resolution, which is critical for assessing the performance of our EMDBC method across multiple timescales. While a global-scale evaluation is valuable, it would require consistent, high-quality observational datasets at comparable resolution, which are not readily available globally. Therefore, our current focus on the CONUS region allows us to rigorously validate the method using reliable and spatially detailed observations.

   Comparison to Other Methods: While we acknowledge the importance of comparing to other bias correction methods, this study aims to introduce and validate a novel, timescale-aware approach. Due to space limitations and the scope of this work, we chose to focus on demonstrating EMDBC's performance against high-resolution observations.

   That said, to provide context and justification, we have compared EMDBC against one established method, QDM, and found that our approach performs comparably or better in representing temperature across timescales. Additionally, the observational comparisons already provide compelling evidence of the method's effectiveness. Nevertheless, we view direct inter-method comparison as an important area for future investigation and will consider this in subsequent studies.

   Model Name in Title: We appreciate the reviewer's feedback about including the model name (WRF-CCSM) in the manuscript title, since only one model is used. As the primary purpose of this paper is to present the novel EMDBC method and not to present results from WRF-CCSM, we have decided to leave the title as-is.

3. **When explaining the methods, one or several schematics would help − > with method is used for which time scale, how is the data used for the evaluation, etc. . .**

   Response: We agree that a visual workflow is essential.

   To that end, Figure 2 (12) has been redesigned as a six-row schematic that tracks the data from raw input to final evaluation:

Row 1 – Inputs Displays the observed Livneh series alongside the raw CCSM simulations for each grid cell.

Row 2 – EMD decomposition Shows the empirical-mode decomposition of each series into intrinsic mode functions (IMFs).

Row 3 – IMF–frequency correlation A heat-map links individual IMFs to their dominant frequencies, indicating which ones capture sub-daily, bi-weekly, seasonal, or annual variability.

Row 4 – Timescale construction IMFs with similar periods are aggregated to build four working bands: bi-weekly, seasonal, annual, and a residual high-frequency component.

Row 5 – Component-wise bias correction

- **Bi-weekly band and residual** — corrected with Quantile-Delta Mapping (QDM).
- **Seasonal and annual bands** — corrected with an ensemble quantile-regression method suited to slowly varying biases.

Row 6 – Reconstruction and evaluation The four corrected bands are summed to form the final EMDBC series.

This expanded schematic now makes explicit (i) which algorithm is applied to each timescale, (ii) how the corrected components are recombined, for clearer methodological presentation. We have also added a description of this figure at the beginning of Section 2.4:

> Building on EEMD, we introduce an *Empirical Mode Decomposition–based Bias Correction* (EMDBC) framework for rectifying model biases across multiple timescales. As sketched in Figure 2, EMDBC proceeds in three steps: (i) **timescale decomposition**—the daily Livneh and CCSM series (Row 1) are split via EEMD into four bands (Rows 2–4); (ii) **timescale-specific correction**—the residual and bi-weekly bands are adjusted with QDM, while the seasonal and annual bands use ensemble quantile regression (Row 5); and (iii) **reconstruction**—the corrected bands are recombined to yield the final series for evaluation (Row 6). We describe each of these steps in detail in the subsections that follow.

4. **The figures should be more understandable, some graphs are not readable, all plots should be commented. Please add a letter to each sub-figure and refer to the letter in the text.**

Response: Thank you for highlighting opportunities to improve the clarity of our figures. In response, we systematically reviewed each figure against the following criteria:

- Adding titled labels and lettering to all subplots where appropriate
- Enhancing diagrammatic figures to better guide readers
- Providing explanations for subplots within the captions
- Appropriately referencing subplots in the main text

To that end, individual updates to each figure are documented below.

- **Figure 1:** We added a title, color bar, axis labels, and grid lines to improve interpretability. Boundary vertices for each case study region were corrected to match the exact polygons used in the analysis, ensuring clear visual distinction. The caption was rewritten accordingly. The revised figure and caption are shown in Figure 11.
- **Figure 2:** We improved readability by increasing font sizes and separating the "IMF" inset plot into its own row. The caption was updated to clarify the time series and data source. As Figure 2 serves as a schematic of the EMDBC workflow, we opted against lettering every subplot, as this could distract readers. Instead, we labeled rows as Step 1, Step 2, and Step 3, consistent with the Methods section. Labels in the first column were refined to better align with the in-text explanation, and the caption now explicitly references each row. The updated figure and caption are shown in Figure 12.
- **Figure 3-4,6:** We individually lettered each subplot and mentioned them in the caption. Also, used color-blind compatible colors. The updated figures and captions are shown in Figures 13, 14, 15.
- **Figure 7:** We added a centered supertitle to reduce visual clutter, then individually lettered each subplot title to reference in the figure caption. The figure caption was revised to explicitly reference each subplot. The updated figure and caption are shown in Figure 16.
- **Figures 8 and 9:** We individually lettered each subplot title to reference in the figure caption and font sizes were increasd to improve readability. The figure caption was revised to explicitly reference each subplot. The updated figures and captions are shown in Figures 17 and 18

.

Here are some specific comments:

5. **L5: "Meteorological signal"− > "Atmospheric variables" ? Which variables / components of the earth system model? which time scales? Which area (USA)?**

Response: Thank you for noting that our abstract does not explicitly mention the meteorological variable used in our experimental validation. We will revise the abstract to clarify that our validation focuses on WRF-CCSM daily temperature data over historical (1995–2004), mid-century (2045–2054), and late-century (2085–2094) periods. The revised abstract text reads:

> *By decomposing meteorological variables into multiple oscillatory components and aggregating them to represent distinct timescales, we apply targeted corrections to each component, thereby preserving both short- and long-term structure in the data.*

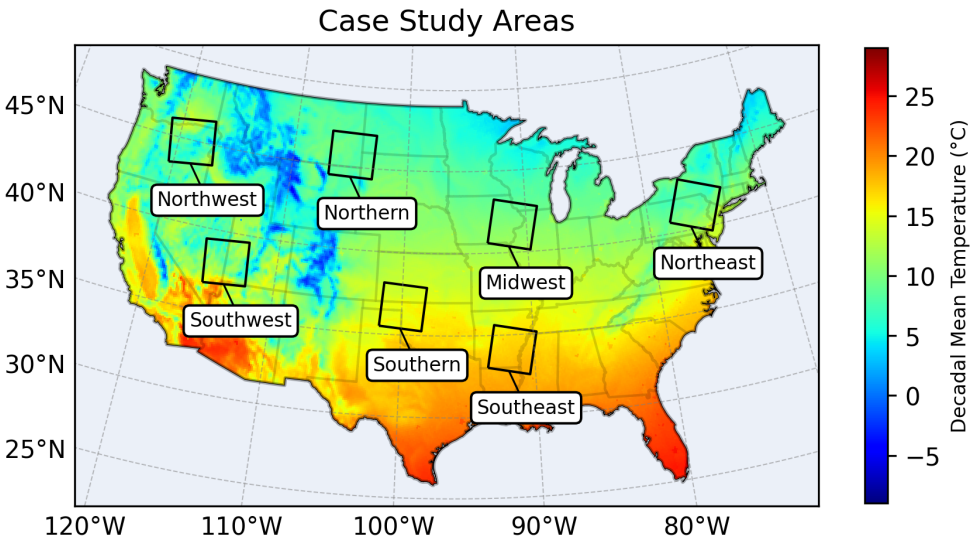

Figure 11: *Map of case study regions selected for evaluating the bias correction methods. Boundaries are overlaid on the average 1995-2004 mean temperature field from WRF-CCSM, illustrating the diverse range of temperature regimes captured by the case study areas.*

> *Experimental validations on WRF-CCSM daily temperature data across historical (1995–2004), mid-century (2045–2054), and late-century (2085–2094) periods demonstrate that this finer-grained method substantially improves upon existing bias-correction techniques such as quantile mapping.*

6. **L7: "Experimental validations demonstrate that this finer-grained method substantially improves upon existing bias-correction techniques such as quantile mapping": it was not demonstrated. "illustrates"? And add on which timescale it improves.**

Response: We will restructure the sentence for greater precision in the revised manuscript :

> *Experimental illustrations show that the timescale-aware EMDBC framework matches the performance of conventional quantile-delta mapping (QDM) at the native daily scale and achieves progressively larger bias reductions at bi-weekly, seasonal, and annual scales.*

The underlying reason is that QDM performs quantile matching only at the original daily resolution, whereas *EMDBC first decomposes each series into multiple timescales, then applies the most appropriate correction in each band* - QDM for the bi-weekly and residual components, and ensemble quantile regression for the seasonal and annual components, where time-dependent biases are stronger.

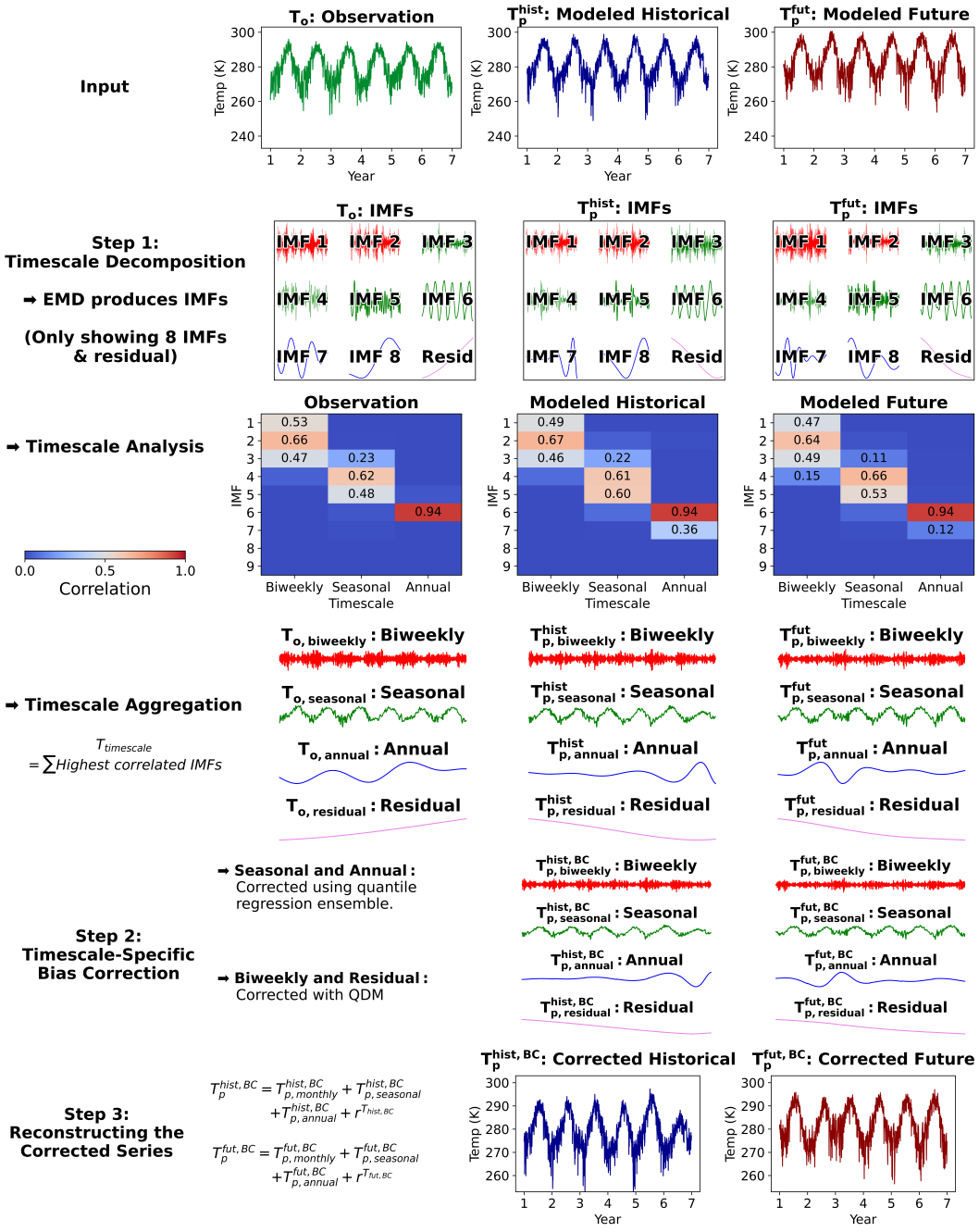

Figure 12: *Timescale-wise bias correction framework using EMDBC. Three inputs are required: temperature timeseries from observation, modeled historical, and modeled future datasets. Here, to demonstrate EMDBC, timeseries data are extracted from Livneh (T_o), WRF-CCSM historical (T_p^{hist}), and WRF-CCSM mid-century (T_p^{fut}) at an arbitrary location. The input temperature series are decomposed into IMFs using EEMD. IMFs are then classified into predefined timescales: biweekly, seasonal, and annual. Bias correction is applied using QDM for biweekly timescale and residuals, and quantile regression for seasonal or annual timescales. Finally, the corrected timescales are summed to reconstruct the bias corrected temperature series.*

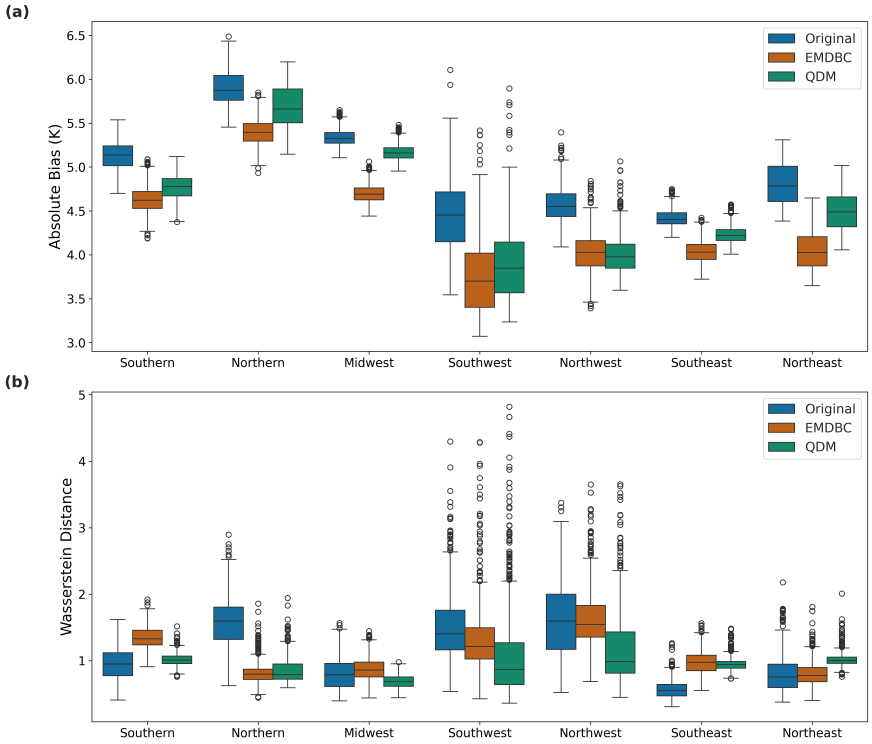

Figure 13: **Comparison of Absolute Biases and Wasserstein Distances Across Sub-Regions in the original daily timescale. (a):** Boxplots of the absolute temperature (in K) bias for the original (CCSM) and bias-corrected (EMDBC and QDM) simulations across sub-regions on the validation dataset. **(b):** Boxplots of the corresponding Wasserstein distances between the observed and modeled temperature distributions across sub-regions on the validation dataset.

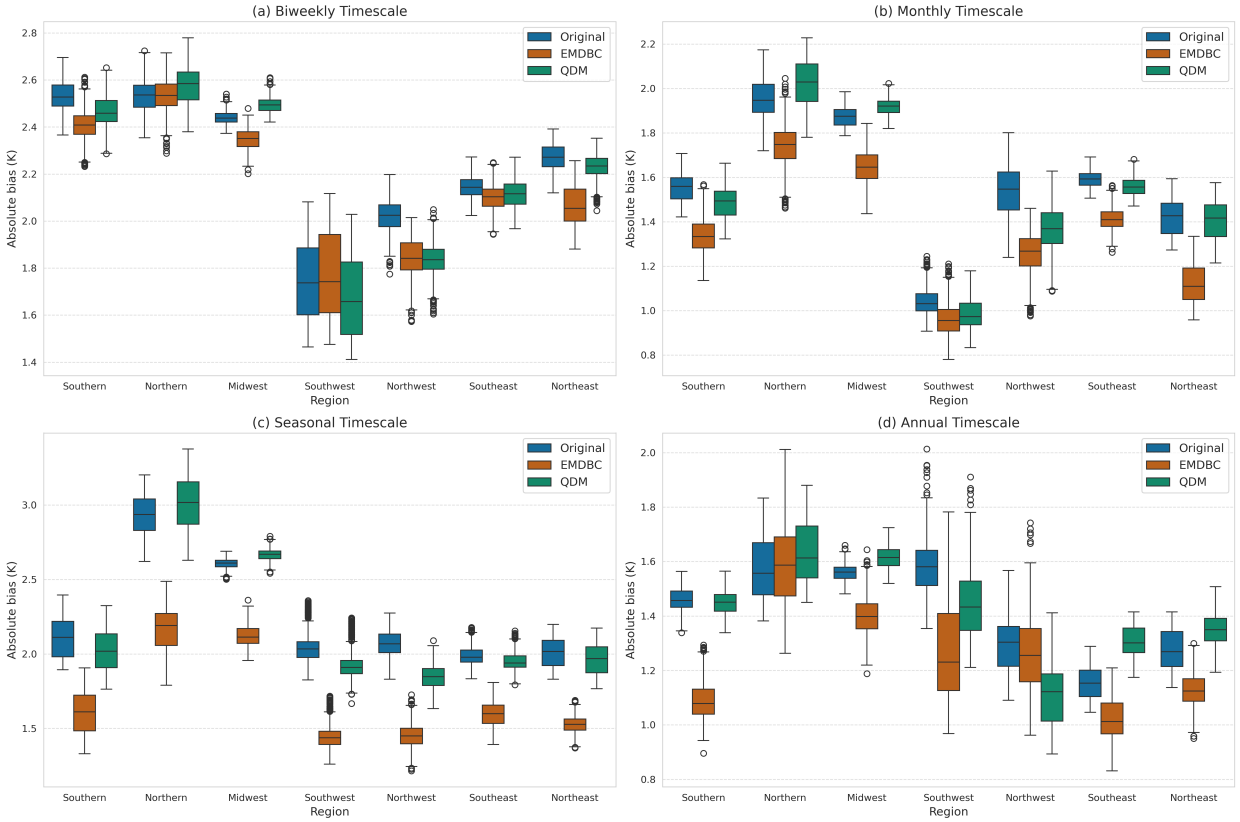

Figure 14: The timescale wise average absolute bias per subregion on the validation dataset. Included timescapes are (a) biweekly, (b) monthly, (c) seasonal, and (d) annual.

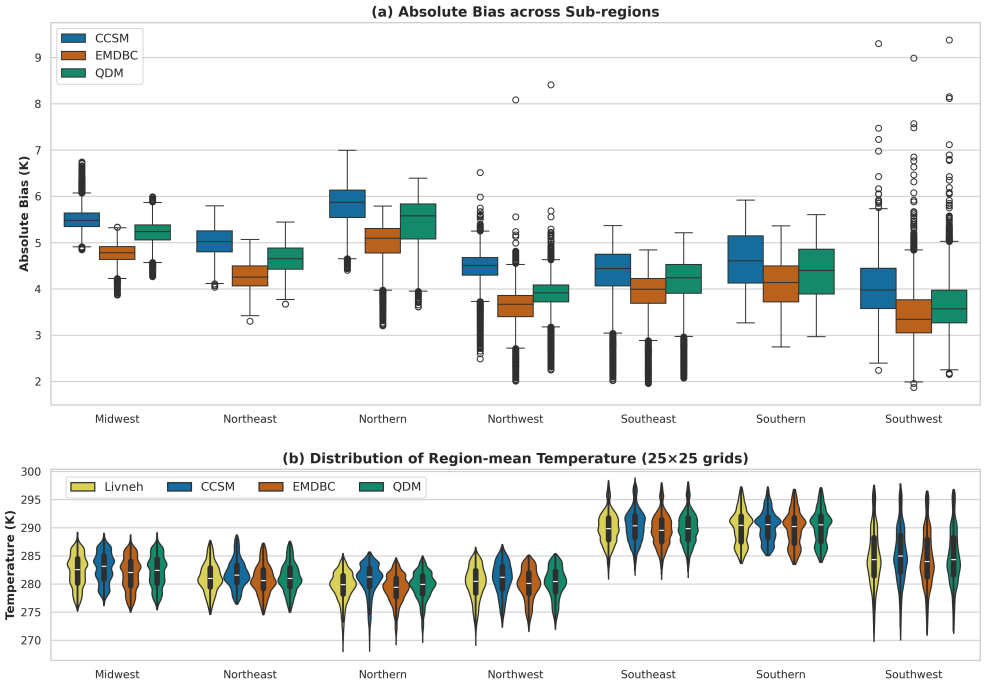

Figure 15: **Comparison of CCSM Temperature Biases and Temperature Distributions Across Sub-Regions. (a)** Boxplots of the absolute temperature bias before and after applying EMDBC and QDM corrections. **(b)** Violin plots showing the distribution of the average temperature for each sub-region.

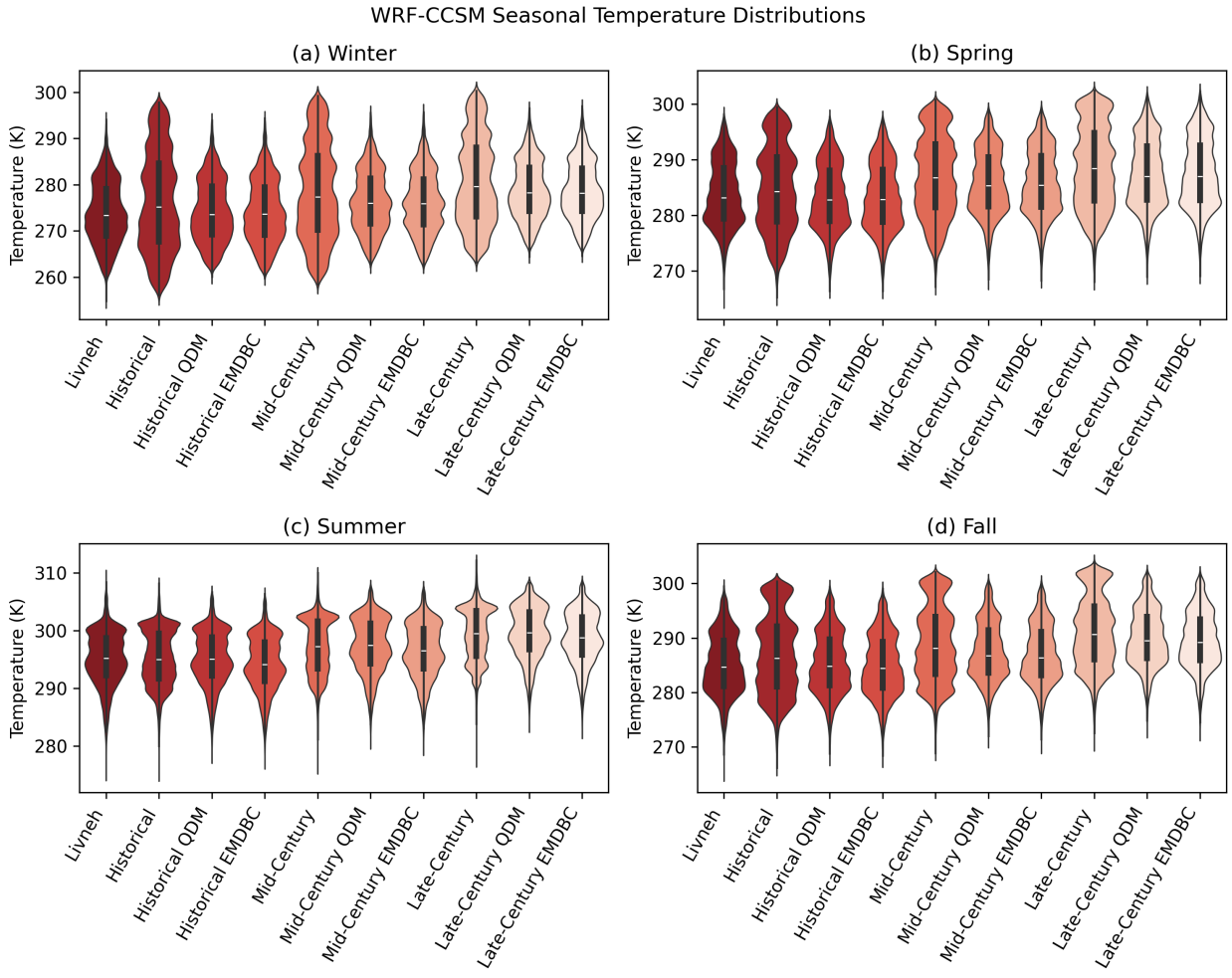

Figure 16: The mean daily temperature by season ((a) winter, (b) spring, (c) summer, (d) fall) across Livneh (1995-2004) and WRF-CCSM historical (1995-2004), mid-century (2045-2054), and late-century (2085-2094) timeframes before and after bias correction. Results for QDM and EMDBC are included. Violin plots displaying all timeframes on a common axis illustrate how both QDM and EMDBC preserve the shape of the observed spatial temperature distribution, while also showing the distribution's shift across centuries as projected by the WRF-CCSM model.

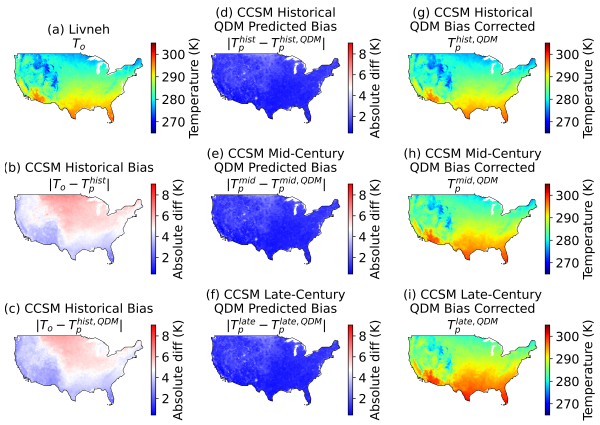

Figure 17: Temperature and temperature bias comparisons over CONUS before and after applying QDM. **Left:** (a) Observed temperature (Livneh, 1995-2004), (b) WRF-CCSM historical (1995-2004) average daily absolute bias, and (c) QDM-corrected WRF-CCSM historical average daily absolute bias. **Middle:** (d) Magnitude of QDM correction in historical, (e) mid-century (2045-2054), and (f) late-century (2085-2004) timeframes. **Right:** (g) QDM-corrected temperatures for WRF-CCSM historical, (h) mid-century, and (i) late-century periods.

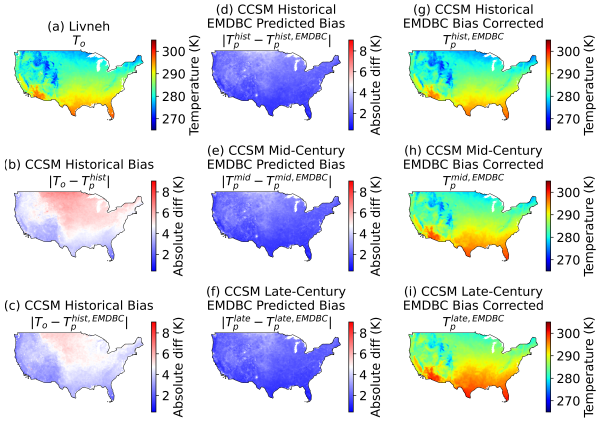

Figure 18: Temperature and temperature bias comparisons over CONUS before and after applying EMDBC. **Left:** (a) Observed temperature (Livneh, 1995-2004), (b) WRF-CCSM historical (1995-2004) average daily absolute bias, and (c) EMDBC-corrected WRF-CCSM historical average daily absolute bias. **Middle:** (d) Magnitude of EMDBC correction in historical, (e) mid-century (2045-2054), and (f) late-century (2085-2004) timeframes. **Right:** (g) EMDBC-corrected temperatures for WRF-CCSM historical, (h) mid-century, and (i) late-century periods.

Because of this timescale-aware correction, EMDBC **consistently reduces bias across all sub-regions even at the original daily input scale** (see absolute-bias maps in Figures 3 and 6). Moreover, Figure 4 demonstrates that the bias reduction magnifies as the aggregation scale increases from bi-weekly to seasonal and annual bands. These results collectively substantiate the claimed improvement.

7. **L8-10: add the limitations of the method (see conclusion).**

Response: We agree that outlining the limitations of EMDBC is important and have discussed them explicitly in the *Conclusion* (Section 5, lines 319-325). There we note, for example, the method's reliance on empirical-mode decomposition (EMD)—which can be computationally demanding and is not yet supported by a full theoretical framework—and suggest future work on alternative, lighter-weight timescale decompositions. To preserve the abstract's focus and brevity, however, we have not duplicated the full list of caveats in Lines 8–10 of the abstract. We believe this keeps the abstract concise while directing interested readers to a dedicated limitations paragraph in the main text, as customary in the literature.

Introduction

8. **L 21. "Unlike statistical downscaling": first define statistical downscaling.**

Response: Thank you for pointing out that we do not apprioriately introduce statistical downscaling. Statistical downscaling typically uses empirical relationships between large-scale climate model outputs and local observations to infer fine-scale climate information. We have added this clarification:

> *Unlike statistical downscaling which relies on drawing empirical relationships between large-scale Earth system models and local observations to infer fine-scale meteorological information, dynamical downscaling can simulate a range of physical processes and their interactions within the Earth system, producing a comprehensive set of dynamically consistent high-resolution meteorological variables.*

9. **L 25. "Region-level modeling": North America Mearns et al. 2012, North American component of the Coordinated regional downscaling experiment NACORDEY, Mearns et al, 2017 − > add other examples on other regions and other authors?**

Response: We agree that this section is improved by providing additional examples and references. Citations have been added to the revised manuscript and the updated text is provided below:

> *Various region-level modeling and assessment initiatives have adopted this approach, including the North American Regional Climate Change Assessment Program (Mearns*

et al., 2012), the North American component of the Coordinated Regional Downscaling Experiment (NA-CORDEX) (Mearns et al., 2017), targeted evaluations for Tasmania (Corney et al., 2013), the central United States (Bukovsky and Karoly, 2011), the European Coordinated Regional Downscaling Experiment (EURO-CORDEX) ensemble over Europe (Jacob et al., 2014), and the Coordinated Regional Downscaling Experiment over Africa (CORDEX-Africa) initiative (Nikulin et al., 2012), demonstrating the enhanced capability to capture fine-scale features and provide more realistic, detailed projections at regional and local scales.

10. **L 29-31: "Despite these improvements... from forcing data and inherent systematic errors such as those ...." − > add references for the source of the RCMs biases**

Response: We agree that adding references to support the claim that bias is driven by several factors improves this statement. Citations have been added to the revised manuscript and the updated text is provided below:

Despite these improvements, RCMs continue to face challenges with biases arising from both their forcing data and inherent systematic errors, such as those related to model resolution (Christensen et al., 2008), simplified physical parameterizations (Misra, 2007; Bukovsky and Karoly, 2011; Jacob et al., 2014), and incomplete understanding of the Earth system (Christensen et al., 2008), all of which degrade the downscaled simulations.

11. **L47. QM method: transfer function based on the quantile distribution, daily values − > not necessarily daily?**

Response: We agree that quantile mapping (QM) is not intrinsically limited to daily data; it can be applied at any temporal resolution for which paired model–observation series exist. In our study the finest common resolution is daily, because the Livneh reference data are available only as daily aggregates. We will revise Line 47 to make this explicit:

In the QM method, a transfer function is created by matching model-simulated and observed quantiles at their common temporal resolution (daily in this study) during a reference period; this function is then applied to future model simulations.

12. **L 51. QM improve model accuracy for both mean and extreme events "(Wood, 2002; Wood et al., 2004; Boé et al., 2007; Piani et al., 2009, 2010; Ashfaq et al., 2010; Teutschbein and Seibert, 2012; Gudmundsson, 2012)" : split the references between the type of application**

Response: Thank you for the suggestion. This clarification has been incorporated into the revised manuscript:

> *Previous studies have shown that QM effectively removes biases, improving model accuracy for both mean values (Wood, 2002; Wood et al., 2004; Boé et al., 2007; Piani et al., 2009) and extreme events (Piani et al., 2010; Ashfaq et al., 2010; Teutschbein and Seibert, 2012; Gudmundsson, 2012).*

13. **L61: "generally parallels that of quantile-based techniques and does not address the core challenge of biases that occur across multiple distinct timescales" – > "generally" suggests it is well known in the literature – > add other references. What does Dhawan et al 2024 show?**

Response: Our intent was to convey that, in many comparative studies, machine-learning (ML) bias-correction schemes perform *similarly* to classical quantile-based methods when both are applied at the same native timescale (daily in our case). We will revise Lines 60–61 to include additional supporting references and to clarify the role of Dhawan et al. (2024):

> *Several ML-based bias-correction schemes have been proposed (e.g., Sarhadi et al. (2016); Miftahurrohmah et al. (2024); Das et al. (2022); Feng et al. (2024)); however, comprehensive intercomparisons such as Dhawan et al. (2024) show that their daily-scale performance is broadly comparable to that of quantile-based approaches like QDM and that none addresses biases occurring across multiple distinct timescales.*

Clarification on Dhawan et al. (2024): Dhawan et al. (2024) conducted a large-sample benchmark using ERA5 pseudo-observations, testing a range of statistical and ML bias-correction algorithms. Their results indicate that—at the daily resolution—*quantile-delta mapping (QDM) achieves the highest overall skill for temperature*, while ML methods offer no systematic improvement. This finding underpins our choice of QDM as the state-of-the-art baseline and motivates the need for the timescale-aware EMDBC framework.

14. **L62. Biases different at daily monthly seasonal annual scales – > different in what sense? not multi annual? Or decadal?**

Response: In this context "different" refers to the fact that the *form, magnitude, and physical source* of model bias change with the temporal scale considered. Short-period fluctuations (hours to days) are governed by weather processes such as convection and frontal passages, whereas seasonal to annual variability reflects changes in radiation balance, soil-moisture feedbacks, or large-scale circulation. As noted by Haerter et al. (2011), lumping all timescales into one aggregate correction mixes these distinct signals and can obscure how a bias will propagate into future-scenario projections. They therefore advocated separating the

record into individual bands and applying a cascade of bias corrections—an idea we implement here via EMD.

Our analysis spans daily up to an "annual" band that, by construction, aggregates all low-frequency IMFs whose dominant periods exceed one year. Consequently, any multi-annual variability that is resolvable within the 1995–2004 Livneh record is already embedded in this band. True decadal signals, however, cannot be isolated with only ten years of data; capturing and correcting such lower-frequency biases would require a substantially longer observational series and is therefore left as future work.

15. **L67 "Empirical mode decomposition-based bias correction (EMDBC)": leveraging the adaptive nature of Empirical Mode Decomposition EMD and its ensemble variant EEMD: explain on which timescale it is used and on which timescale QDM is used**

Response: EMDBC first applies EEMD to the full daily series, producing intrinsic mode functions (IMFs) plus a residual. These IMFs are then aggregated into three bands. Each band (and the residual) is bias-corrected with the method most appropriate to its frequency content:

Timescale (IMF aggregation)	**Bias-correction method**
Bi-weekly (hours–14 days)	QDM
Seasonal (14 days–∼3 months)	Ensemble quantile regression
Annual ($>$ ∼3 months, incl. multi-annual)	Ensemble quantile regression
Residual	QDM

Thus, EEMD is used solely for the decomposition step, while QDM corrects the high-frequency residual and bi-weekly bands, and ensemble quantile regression corrects the lower-frequency seasonal and annual bands.

16. **L 73: section 3: demonstrates EMDBC effectiveness-¿ "demonstrates" is too strong here**

Response: We will adopt a more neutral phrasing in the revised manuscript to reflect an appropriate tone.

Methods:

17. **L 80: why is this dataset used?**

Response: We use the 1995–2004 WRF-CCSM regional simulation because it supplies a publicly accessible, high-resolution (≈ 12 km) North-American temperature field that has already been thoroughly vetted in the literature. The product is comparable in spatial detail

and quality to NA-CORDEX and is hosted on the ClimRR platform, which ensures straightforward community access. This well-validated data set gives us a solid testbed for developing and illustrating the EMDBC bias-correction method without introducing additional uncertainty from a less-studied model archive.

18. **L 78. observed and modeled temperature – > detail what is used for what?**

Response: Livneh serves as the observational reference against which biases are quantified and corrected, while WRF–CCSM provides the model fields to be bias-corrected. In the manuscript, Section 2.1, Lines 80–96 describe each dataset, its spatial resolution, temporal coverage, and specific role in our analysis.

19. **L 80: WRF-CCSM: explain acronym. Only atmosphere or coupled ?**

Response: In the manuscript, Line 82-84 describes the WRF-CCSM as follows:

> *These projections, called WRF-CCSM, are generated by dynamically downscaling the Community Climate System Model version 4 (CCSM4) using the Weather Research and Forecasting (WRF) model version 3.3.1 (Skamarock et al., 2008).*

CCSM4 is a coupled global climate model incorporating an ocean component, whereas the regional WRF simulations we use are atmosphere-only. WRF uses the boundary conditions and prescribed sea-surface temperature fields supplied by CCSM4, and couples internally to the Noah land-surface model. The physics schemes listeed in Lines 85-88 apply within this WRF configuration and are provided below for reference:

> *The model uses the Grell-Devenyi convective parametrization (Grell and Dévényi, 2002), the Yonsei University planetary boundary layer scheme (Noh et al., 2003), the Noah land surface model (Chen and Dudhia, 2001), the longwave and shortwave radiative schemes of the Rapid Radiation Transfer Model for GCM (Iacono et al., 2008), and the Morrison microphysics scheme (Morrison et al., 2009).*

20. **L 81 "moodeled"**

Response: We have corrected the typo in the revised manuscript.

21. **L 85. "RCP 8.5 scenario is used": explain acronym + add reference**

Response: Thank you for pointing out that we do not introduce this acronym appropriately. Because many readers recognize the pathway as "RCP 8.5," we retain the abbreviation as well. We have revised the line to read:

For future periods (mid- and late-century), we use the Representative Concentration Pathway 8.5 (RCP 8.5) scenario, which corresponds to a high greenhouse gas concentration trajectory, reaching approximately 8.5 W/m^2 of radiative forcing by 2100 (Riahi et al., 2011).

22. **L 90: "$3 \times 10s^1$" $->$ why s^1?**

Response: We thank the reviewer for pointing out this typo. Though our manuscript source code had the proper notation, it was not being rendered properly in the PDF document. The text was corrected:

...using a nudging coefficient of $3 \times 10^{-5}s^{-1}$

23. **L 93: why is this dataset used?**

Response: We discussed this above in comment 17.

24. **L 97 "leverage observation data"-¿ in which way?**

Response: By "leverage" we mean that the Livneh observations from 1995–2004 will serve as the reference against which the WRF-CCSM model projections will be bias-corrected. We have revised the text to use the word "use" instead of "leverage."

25. **L 98: typo: "is used to THE learn"**

Response: We will correct the typo in the revised manuscript.

26. **L99: daily mean temp data calculated from the 3h outputs of WRF CCSM to match the temporal resolution of the observed Livneh data $->$ upscaling**

Response: WRF–CCSM provides 3-hourly temperatures, whereas Livneh is available only as daily aggregates. We therefore average the 3-hourly model values to obtain daily means, ensuring both data sets share the same temporal resolution before bias correction. This step is a temporal aggregation, which can be termed as upscaling to daily resolution.

27. **L 102: Statistical framework is then applied to identify and learn the systematic biases in the simulation data. $->$ what statistical framework? which simulation data, WRF-CCSM or Livneh?**

Response: Here "statistical framework" refers to the bias-correction procedure itself: we estimate a transfer function that matches the empirical distribution of the WRF–CCSM simulated series to that of the Livneh observations over the historical period, and then apply this learned function to adjust the future WRF–CCSM projections. To avoid ambiguity, we have replaced the original sentence with:

> *We estimate a transfer function that aligns the empirical distribution of the WRF–CCSM daily series with the corresponding Livneh observations for 1995–2004, and then apply this function to correct the future WRF–CCSM projections.*

28. **L 104 "the model generates bias-corrected future predictions that scale more closely with observational data"– > which observations are available for future predictions ? or maybe the authors refer to the next paragraph ? This paragraph should probably be merged with the next paragraph for clarity. Not clear how Livneh data is used: the data is spilt into 2 parts?**

> *By correcting the learned biases, the model generates bias-corrected future predictions that retain the projected distribution shift while better matching the observed distributional shape, thereby reducing systematic and known biases in the model output.*

Response: Future observations are unavailable, so a bias-correction transfer function learned from a historical calibration period can only be applied to future model projections; its real-world performance at that time cannot be verified. To gauge how well each method would behave in such a setting, we carry out a split-sample experiment: the Livneh–WRF-CCSM record is divided into a calibration period (1995–1999) for estimating the transfer function and a validation period (2000–2004) that serves as a pseudo-future for benchmarking the corrected output.

We have kept the two paragraphs separate—one describing the standard bias-correction workflow, the other describing our split-sample validation—to preserve this distinction while clarifying the text as shown above.

29. **L 116: QM and QDM: add a Graphic illustrating the concepts of BQM and QDM (many examples in the literature, e.g. HESS - Peer review - Precipitation ensembles conforming to natural variations derived from a regional climate model using a new bias correction scheme + https://doi.org/10.1007/s40641-016-0050-x – > maybe also add these two references)**

Response: We appreciate the reviewer's suggestion. Because the core objective of the manuscript is to introduce the *Empirical Mode Decomposition-based Bias Correction* (EMDBC) framework, we aim to keep the figures focused on elements that are unique to EMDBC. Quantile-based approaches such as BQM and QDM serve here only as interchangeable bias-correction operators within EMDBC rather than as contributions in their own right. To

balance clarity with concision, we have therefore: (i) ensured the explanatory text provides a self-contained summary of BQM and QDM; and (ii) added the reviewer's recommended citations. This directs interested readers to authoritative information while allowing the manuscript to remain focused on the EMDBC paradigm. Here is an excerpt of the revised section:

- ***Basic Quantile Mapping (QM):*** *This approach directly corrects model outputs by aligning their quantile functions to that of the observed data(Tong et al., 2021; Kim et al., 2016) ...*
- ***Quantile Delta Mapping (QDM):*** *QDM extends QM by accounting for shifts between the historical and future model distributions (Tong et al., 2021; Maraun, 2016)...*

30. **L 120: F(T) and CDF: give the formula**

Response: The cumulative distribution function (CDF) of a random variable X is defined by

$$F_X(x) = P\big(X \le x\big), \qquad x \in \mathbb{R},$$

while its empirical counterpart based on n observations $\{X_i\}_{i=1}^{n}$ is

$$\widehat{F}_n(x) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{\{X_i \le x\}}.$$

(The abbreviation "CDF" is also listed in the acronym table in Table 4, and will be added in the appendix of the revised manuscript.)

Table 4: Acronyms and Symbols used in this study

Acronym / Symbol	Full Form	Brief Description (incl. equations)
CMIP6	Coupled Model Intercomparison Project Phase 6	Multi-model ensemble of coordinated global climate simulations.
GCM	Global Climate Model	Dynamical model representing physical processes of the climate system on a global grid.
RCM	Regional Climate Model	Higher-resolution model nested within a GCM to resolve regional detail.
BC	Bias Correction	Statistical adjustment applied to model output to align it with observations.
QM	Quantile Mapping	Bias-correction technique that remaps model quantiles to observed quantiles.

Continued on next page

Table 4 (continued)

Acronym / Symbol	Full Form	Brief Description (incl. equations)
CDF	Cumulative Distribution Function	$F_X(x) = \Pr[X \leq x]$ for a random variable X.
QDM	Quantile Delta Mapping	Bias-correction method that preserves the modeled change signal while correcting quantiles.
EMD	Empirical Mode Decomposition	Data-adaptive decomposition that yields oscillatory components called IMFs.
EEMD	Ensemble Empirical Mode Decomposition	Noise-assisted EMD variant that improves mode separation.
WRF–CCSM	Weather Research and Forecasting–Community Climate System Model	Dynamical downscaling chain coupling WRF with CCSM boundary fields.
IMF	Intrinsic Mode Function	Oscillatory component extracted by EMD, each with well-behaved local extrema.
EMDBC	EMD-based Bias Correction	Bias-correction framework that operates on time-scale-specific IMFs before reconstruction.
W_p (WD)	Wasserstein Distance	$W_p(P,Q) = \left(\inf_{\gamma\in\Gamma(P,Q)} \int_{\mathcal{X}\times\mathcal{X}} \|x - y\|^p \, \mathrm{d}\gamma(x,y)\right)^{1/p}$; where $\Gamma(P,Q)$ denotes the set of all couplings with marginals P and Q, commonly $p = 1$.
MSE	Mean Squared Error	$\mathrm{MSE}(y,\hat{y}) = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$; average squared deviation between predictions and observations.
FFT	Fast Fourier Transform	Algorithm that computes the discrete Fourier transform in $O(n \log n)$ operations.

31. **L 122: which model outputs?**

Response: Throughout the manuscript, "model output" refers specifically to the WRF–CCSM regional climate simulation that we seek to bias-correct, while "observations" denote the Livneh gridded temperature data. To make this explicit, we have revised Line 97 as follows:

To bias-correct the future WRF–CCSM projections, we use the 1995–2004 Livneh observations as the calibration reference.

32. **L123 and 131: Eq 1 and Eq2: "p" refers to what?**

Response: We use the subscripts "p" and "o" consistently to distinguish model projections from observations:

$$\begin{aligned} T_o &= \text{observed (Livneh) temperature,} \\ T_p^{\text{hist}} &= \text{historical WRF–CCSM simulation,} \\ T_p^{\text{fut}} &= \text{future WRF–CCSM projection,} \\ T_p^{\text{fut, BC}} &= \text{bias-corrected future projection.} \end{aligned}$$

In the manuscript, Lines 119-120 define these notations.

33. **L 135: "Nonparametric empirical CDFs are commonly used for flexibility, although parametric and semiparametric distributions can also be employed" – > which formula for parameter (Semi-) distributions? which adjustable parameters?**

Response: The specific form and adjustable parameters of a parametric or semiparametric CDF for QM depend on the variable being bias-corrected and the aspect of its distribution that is most important. For example, precipitation studies often fit a Gamma or mixed-exponential distribution for the bulk of the data and an extreme-value tail (e.g., Generalized Pareto) to capture rare events. Temperature, by contrast, is usually well handled by the empirical CDF, which is why we employ the non-parametric QDM in this paper. We mention parametric and semiparametric options only to alert readers who may wish to adapt EMDBC to other variables; the full methodological details can be found in Gudmundsson et al. (2012) and in additional sources such as Rajulapati and Papalexiou (2023), which we will cite in the revised manuscript.

34. **L 143: "meteorological time series" is it meteorological modelling or climate modelling ? is it "meteorological variables" in a "climate model"**

Response: Thank you for pointing this out. To avoid ambiguity, we have replaced "meteorological time-series" with "time-series of meteorological variables produced by Earth system models" in the revised manuscript.

35. **L 146: "future projections" – > is it fitted only to climate projections or also to short term weather forecasts?**

Response: Thank you for this question. Our study is focused on future climate projections, particularly analyzing long-term signals from Earth system models. However, this method is not limited to a specific scale (e.g., long-term projections). It is a general correction workflow coupled with EMD that can also be applied to high-frequency data such as weather forecasts.

That said, applying EMDBC to short-term forecasts may require different parameter choices or bias correction methods depending on the temporal resolution, noise characteristics, and overall objective. This could be an interesting direction for future work.

36. **L 153: "Can suffer mode mixing" – > explain more how this occurs**

Response: "Mode mixing" is an artifact of EMD where one IMF includes oscillations from very different frequency ranges or a single physical mode is split across several IMFs. This happens when the original signal contains intermittent or abrupt changes, which causes the spline envelopes used in EMD to misassign local extrema. The issue is well known in the EMD literature, and the ensemble EMD method proposed by Wu and Huang (2009) offers a practical remedy. EEMD adds small independent white–noise realizations to the signal, performs EMD on each noisy copy, and then averages the resulting IMFs. The added noise encourages the envelopes to sample the time–frequency space more uniformly, while the ensemble average removes the noise itself, which significantly reduces mode mixing. This point is discussed in detail at Line 153 and in the opening paragraph of Appendix A, where we also provide several supporting references:

> The performance of the proposed EMDBC framework depends on the quality and separation of the IMFs generated during the decomposition process. A common challenge in EMD methods is *mode-mixing*, where oscillatory modes of different frequencies are entangled within a single IMF, reducing interpretability and effectiveness (Tang et al., 2012). While the Ensemble EMD (EEMD) approach (Wu and Huang, 2009) mitigates mode-mixing by introducing random noise, it does not fully eliminate the issue. Several alternative strategies have been proposed to ensure distinct frequency bands for IMFs (Tang et al., 2012; Fosso and Molinas, 2018), but none has proven universally robust.

37. **L 155. Add other references using EEMD in Climate sciences: e.g.**

Investigating monthly precipitation variability using a multiscale approach based on ensemble empirical mode decomposition — Paddy and Water Environment

Identification of relationships between climate indices and long-term precipitation in South Korea using ensemble empirical mode decomposition - ScienceDirect,

The multi-timescale temporal patterns and dynamics of land surface temperature using Ensemble Empirical Mode Decomposition - ScienceDirect,

A time series processing tool to extract climate-driven interannual vegetation dynamics using Ensemble Empirical Mode Decomposition (EEMD) - ScienceDirect

Response: Thank you for suggesting these references. They demonstrate the practical value and broad use of the EEMD approach. We will revise Line 155 to read:

EEMD has been successfully incorporated in several recent studies, for example, Alizadeh et al. (2019); Kim et al. (2018); Liu et al. (2019); Hawinkel et al. (2015)

38. **L160 Explain why the noise helps**

Response: We discussed this above in comment 36.

39. **L161. All the operation can be done with the Python package PyEMD?**

Response: Yes. Once a time series is provided, the open-source PyEMD package can run the full EEMD procedure and return the complete set of IMFs together with the monotonic residual. Users may also adjust noise amplitude, ensemble size, and other EEMD parameters within the package if needed.

40. **L 171. "Total number of extracted IMFs m^s" $->$ why "^{s}"?**

Response: The superscript s indicates that the total number of IMFs depends on the specific time series being decomposed. We first define

$$m^s = \text{number of IMFs extracted from series } s.$$

Accordingly, Eq. (6) uses notations like m^{T_o} to denote the number of IMFs obtained from the observed temperature series T_o. This notation distinguishes IMF counts for different data sets without introducing separate symbols for each one.

41. **L 174. "To address this, we implement an additional hyperparameter-tuning step that reinforces distinct frequency separation and minimizes overlap among IMFs"$->$ give type of procedure used in a few words ?**

Response: To enforce clear separation among IMFs, we add an iterative step: after each EEMD run, we estimate the dominant frequency of every IMF, impose frequency-spacing constraints to ensure minimal overlap, and rerun EEMD with adjusted settings until those constraints are met. The precise criteria are documented in Appendix A. We will add the following brief sentence in the revised manuscript to describe this procedure: —

After each EEMD pass, we evaluate the peak frequency of every IMF, impose spacing constraints to minimize overlap, and iterate the decomposition with adjusted parameters until those constraints are satisfied (see Appendix A for details).

42. **Eq 6: write the formula showing how it aggregates: To=...**

Response: We agree that Eq. 6 should explicitly show how the IMFs are summed into the three working time-scale bands and then recombined. In the revised manuscript, we will update Eq. 6:

$$T_{o,\text{biweekly}} = \sum_{j=1}^{[\tau_1 m^{T_o}]} s_j^{T_o}, \quad T_{o,\text{seasonal}} = \sum_{j=[\tau_1 m^{T_o}]+1}^{[\tau_2 m^{T_o}]} s_j^{T_o}, \quad T_{o,\text{annual}} = \sum_{j=[\tau_2 m^{T_o}]+1}^{m^{T_o}} s_j^{T_o}, \tag{2}$$

$$T_o = T_{o,\text{biweekly}} + T_{o,\text{seasonal}} + T_{o,\text{annual}} + r^{T_o} \tag{3}$$

43. **Eq 6: what is "tau 1 m T0"? "tau 2 m T0"?**

Response: In Eq.∼(6), the expressions $\tau_1 m^{T_o}$ and $\tau_2 m^{T_o}$ denote the index cut-points that partition the m^{T_o} intrinsic mode functions (IMFs) of the observed series T_o into three frequency bands. Here $0 < \tau_1 < \tau_2 < 1$, and the integer parts $\lfloor \tau_1 m^{T_o} \rfloor$ and $\lfloor \tau_2 m^{T_o} \rfloor$ mark the boundaries between the bi-weekly, seasonal, and annual groups.

We obtain τ_1 and τ_2 through the following procedure (described in the paragraph immediately after Eq. 6 in the manuscript):

(a) Band-pass filter the daily time series to isolate the bi-weekly, seasonal, and annual frequency ranges.

(b) Compute the correlation of every IMF with each band-pass-filtered signal.

(c) Select τ_1 and τ_2 so that the IMFs most strongly correlated with each range are placed in the corresponding group, minimizing overlap between bands.

This procedure yields distinct IMF groups for the bi-weekly, seasonal, and annual timescales, which are then bias-corrected separately in the subsequent steps of EMDBC.

44. **L 177 and L 181: give the values of the thresholds τ_1 and τ_2. How are they estimated?**

Response: The thresholds τ_1 and τ_2 are determined *data-adaptively* for each individual time series, so no fixed numerical values apply. As detailed in our reply to Comment #43—and in the paragraph immediately after Eq. 6—they are obtained by correlating every IMF with band-pass-filtered versions of the series (bi-weekly, seasonal, annual) and then selecting the cut-points that best separate these frequency ranges.

45. **L 181: "selecting τ_1 and τ_2 such that the IMFs most closely matching each frequency range are grouped together." − > give the formula to select the values**

Response: For each IMF $s_j^{T_o}$ we compute its Pearson correlation with the band-pass-filtered series representing the three target frequency ranges, denoted B_1 (bi-weekly), B_2 (seasonal), and B_3 (annual). Let

$$r_{j,k} = \mathrm{corr}\big(s_j^{T_o},\ B_k\big), \qquad k \in \{1, 2, 3\}.$$

Each IMF is assigned to the band for which the correlation is maximal, $\arg\max_k r_{j,k}$. Let m^{T_o} be the total number of IMFs for T_o. We define the cut-points

$$\tau_1 = \frac{\max\{\, j : \arg\max_k r_{j,k} = 1\}}{m^{T_o}}, \qquad \tau_2 = \frac{\max\{\, j : \arg\max_k r_{j,k} \le 2\}}{m^{T_o}},$$

so $\lfloor \tau_1 m^{T_o} \rfloor$ and $\lfloor \tau_2 m^{T_o} \rfloor$ are the last indices assigned to the bi-weekly and seasonal groups, respectively. This procedure groups the IMFs that are most strongly correlated with each frequency band into three contiguous, non-overlapping sets. The same correlation-based scheme is applied to the IMFs of T_p^{hist} and T_p^{fut} to construct their corresponding time-scale bands.

46. **L 179: maybe change the order in the sentence: "In this study, we use the butter function available in scipy (Virtanen et al., 2020) to perform bandpass filtering of the original signal, isolating the frequencies associated with each timescale" − > "In this study, we perform bandpass filtering of the original signal, isolating the frequencies associated with each timescale using the butter function available in scipy (Virtanen et al., 2020)."**

Response: Agreed. We have adopted the suggested phrasing and will revise the sentence accordingly.

Appendix A

47. **L 345: What is the relative change in frequency between consecutive IMFs?: there is a gap or no gap between IMFs?**

Response:

Response: Ideally consecutive IMFs should exhibit a clear frequency gap; otherwise mode mixing can occur. We monitor this by computing the relative change in dominant frequency

$$\frac{\Delta f_{\max}^{(j)}}{f_{\max}^{(j)}} = \frac{f_{\max}^{(j-1)} - f_{\max}^{(j)}}{f_{\max}^{(j)}},$$

where $f_{\max}^{(j)}$ is the peak frequency of the j-th IMF. We require this ratio to lie within a user-defined band $0 < \delta_{\min} < \frac{\Delta f_{\max}^{(j)}}{f_{\max}^{(j)}} < \delta_{\max}$ for every pair of adjacent IMFs. If any pair falls outside the band we rerun the EEMD step with adjusted noise amplitude and ensemble size until all IMFs satisfy the criterion, thereby enforcing a gradual decrease in frequency.

The thresholds $\delta_{\min}$ and $\delta_{\max}$ are hyper-parameters tuned by split-sample cross-validation: we divide the historical period into two halves (as in our main validation setup), generate IMFs on the training and validation split, and select the thresholds that yield both stable IMF separation and good bias-correction skill on the second half. In our experiments $\delta_{\min} = 0.2$ and $\delta_{\max} = 0.8$ provided reliable results across all grid cells.

48. **L 355: Show the case of 2 IMFs: values of f0, f1, f2**

Response: As the number of generated IMFs is data-adaptive and depends on each input time-series, here we present the related frequencies and relative changes for the Livneh observed series and CCSM historical series for a representative grid cell ($2817'28.79"N, 9953'53.69"W$) in the following table:

(a) Livneh IMF dominant frequencies

IMF #	Dominant Freq	$\Delta f / f_{\text{prev}}$
1	0.18685	—
2	0.07753	0.585
3	0.04904	0.367
4	0.02356	0.520
5	0.01096	0.535
6	0.00274	0.750
7	0.00137	0.500
8	0.00055	0.600

(b) CCSM IMF dominant frequencies

IMF #	Dominant Freq	$\Delta f / f_{\text{prev}}$
1	0.12356	—
2	0.09151	0.259
3	0.05863	0.359
4	0.01973	0.664
5	0.00603	0.694
6	0.00274	0.545
7	0.00164	0.400
8	0.00082	0.500
9	0.00027	0.667

Table 5: Dominant frequencies and relative changes for Livneh and CCSM IMFs.

49. **L 355: Tuned through cross-validation? With which data?**

Response: Discussed above in Comment 47.

50. **L 356: yielded satisfactory results: which criteria is used?**

Response: Discussed above in Comment 47.

51. **In Algo 1: Recompute f(j)max – > how are the frequencies updated?**

Response: If the frequency-spacing constraints are violated, the current set of IMFs is discarded and the EEMD procedure is rerun; no incremental adjustment of individual modes is made. After each new decomposition, we recompute every IMF's dominant frequency $f_{\max}^{(j)}$ from its amplitude spectrum obtained via the fast Fourier transform (FFT; Rockmore (2000)). This regenerate–retest loop continues until all IMFs satisfy the prescribed spacing criteria.

52. **A schematic of the algorithm would help.**

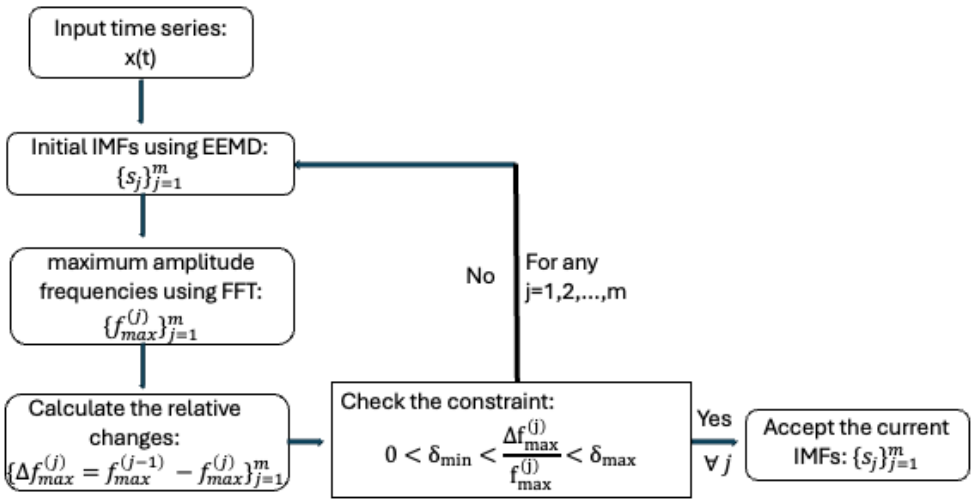

Figure 19: Iterative constriction of IMFs maintaining a gradual decrease in mode frequency for EMDBC

Response: For clarity, we have included here the schematic Figure 19 of the iterative approach for IMF generation and included it in this rebuttal. The manuscript already contains Algorithm 1, which describes the same sequence of steps in pseudocode; adding the figure alongside the algorithm would therefore be redundant. If the reviewer believes the schematic conveys the procedure more effectively, we are happy to replace Algorithm 1 with the figure in the revised manuscript.

53. **Fig 1: how are regions selected? Add color bar or remove colors.**

Response: We thank the reviewer for pointing out the issues in Figure 1. The suggested figure revisions were made and documented in our response above to Comment 4.

To select regions, we randomly sample a 25-cell × 25-cell area from each of the 7 US regions that the Fifth National Climate Assessment explores over the continental United States. This sampling is performed iteratively until the entire sampled area falls within the continental United States. We have added this explanation to the manuscript text:

> *For a comprehensive spatial evaluation, we randomly selected seven areas, each measuring 25×25 grid cells (300 km × 300 km), from major subregions defined in the Fifth National Climate Assessment (USGCRP, 2023)across the continental United States, shown in Figure 1, ensuring a diverse set of conditions.*

54. **Fig 2: Difficult to understand which are the important information on the plots**

Response: We appreciate the feedback. To improve the clarity and visual priority of Figure 2, we have directly labeled the relevant rows as Step 1, Step 2, and Step 3, consistent with the Methods section of our manuscript. We have also inserted ASCII arrows to convey

the font-type hierarchy and better guide the reader across the figure. The updated figure is included here as Figure 12. The in-text citation of this figure has likewise also been revised and documented in our response to Comment 4.

55. **Fig. 2: Why not compared directly the observation VS the corrected on the same plots? and the input VS corrected output on the same plots? and the Timescales VS Corrected Timescales on the same plot.**

Response: We appreciate the reviewer's suggestion. The role of Figure 2 is to provide a graphical schematic of the EMDBC workflow. Because each original signal (observations, historical simulations, future simulations) produces several derived series, the figure focuses on showing how intrinsic mode functions are grouped by timescale, bias corrected, and recombined to form the final output. With this in mind, we agree with the reviewer about the importance of graphical and quantitative comparisons (e.g., inputs versus outputs, observations versus outputs, timescale-to-timescale evaluations). As such, validated benchmarks of the bias-corrected outputs that use these criteria are included in Section 3.1, "Validation results."

56. **Fig 2: A schematic of the method would help understand the successive plots**

Response: Thank you for the recommendation. We have addressed this comment in our response to Comment 3.

57. **L 186 "the nature of the biases can vary greatly depending on whether we are dealing with short-term fluctuations (e.g., biweekly scales) or longer-term patterns (e.g., seasonal or annual). " $->$ indicate how it changes**

Response: Our analysis shows a clear scale dependence in model bias. Short-term IMFs ($<$ 14 days, the bi-weekly band) fluctuate rapidly about zero, whereas the seasonal and annual IMFs drift smoothly and retain the same sign for months to years, signalling a persistent structural bias. To quantify this behaviour we computed autocorrelations at lags 1–365 days. The bi-weekly bias displays consistently low autocorrelation—evidence of weak temporal dependence that favours a non-parametric correction—while the seasonal and annual bands show much higher values. Figure 20 illustrates these patterns for a representative grid cell.

This scale-dependent behaviour aligns with the broader literature. Internal variability dominates near-term projections, producing stochastic, low-persistence errors, while structural model shortcomings become the chief source of bias at seasonal to multi-annual horizons (Hawkins and Sutton, 2009, 2012). Because these structural errors stem from fixed deficiencies in model physics, they persist from one season or year to the next and exhibit high autocorrelation. Recognising this contrast between short- and long-term bias underpins our decision to apply a distribution-based method (QDM) to the high-frequency band and a regression-based method to the more systematic seasonal and annual bands.

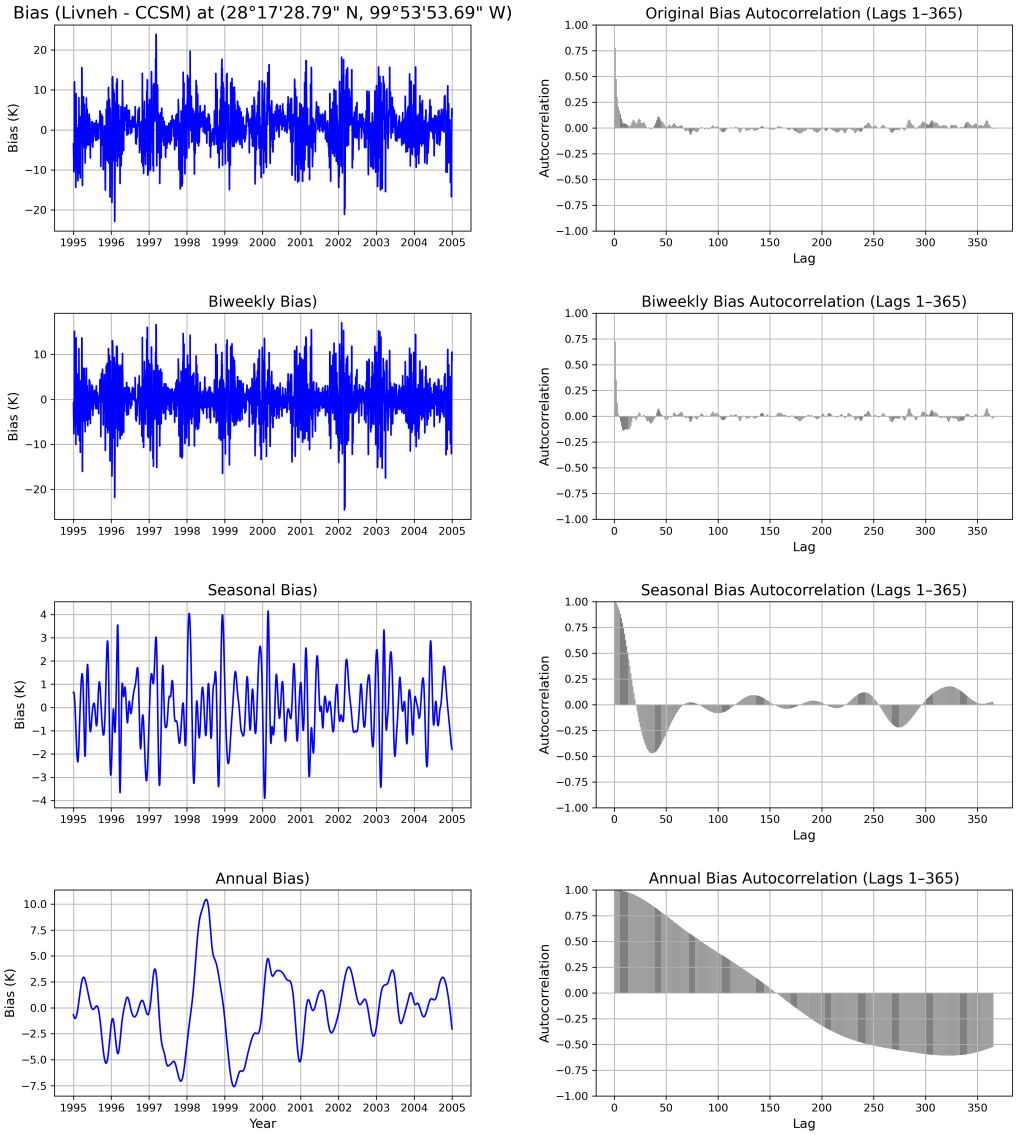

Figure 20: Bias decomposition by time scale for a representative grid cell. The first panel ("Original") shows the raw daily bias (Livneh – CCSM). Using EMD, intrinsic mode functions are grouped into bi-weekly, seasonal and annual bands; the second column reports lagged autocorrelations for each band. The bi-weekly component has markedly lower autocorrelation than the seasonal and annual components, indicating weaker persistence.

58. **L 191" At the biweekly scale, signals often exhibit substantial variability and frequent extremes, yet show little in the way of stable temporal patterns that persist across years" : reference ?**

Response: Discussed above in Comment 57.

59. **L 192- 193"Because a more complex regression approach is unlikely to provide significant benefits at this resolution, we use the QDM to correct these components": reference?**

Response: At the shorter timescales, the bias shows little temporal persistence (see Figure 20); covariates such as day-of-year therefore add almost no explanatory power. In this low-dependence setting, regression terms collapse to the mean of the target variable, offering no advantage over a purely distributional adjustment. Several comparison studies reach the same conclusion for sub-monthly temperature: non-parametric quantile-based methods already achieve the minimum error attainable at that resolution, while adding time-dependent predictors yields negligible improvement (e.g., Gudmundsson et al., 2012; Cannon et al., 2015). For this reason, we correct the high-frequency bands with QDM, which directly aligns the empirical quantiles of the model with those of the observations and has proven robust in the daily-scale temperature literature.

60. **L 202: Were other methods tested to confirm the hypothesis that QDM is the best here?**

Response: Discussed above in Comment 58.

61. **L 203 "At longer timescales (seasonal or annual), biases often manifest in more systematic patterns that persist across multiple years." − > reference?**

Response: Discussed above in Comment 57.

62. **L 237: Here a schematic showing the successive steps would be very helpful**

Response: Thank you for the recommendation. We have addressed this comment in our response to Comment 3.

63. **L 237: Was the QDM tested also on these timescale to confirm the hypothesis that the used method is better than the QDM on these timescales?**

Response: Yes. We applied QDM to each timescale and found that, owing to the strong temporal dependence in the seasonal and annual bands (see above Comments 57–59), the regression-based approach outperforms QDM at those scales.

64. **L 237: When is EEMD applied?**

Response: EEMD is applied at the very start of the EMDBC workflow. It decomposes each input temperature series into its IMFs before any bias-correction step is performed (Step 1, see Figure 12).

2.5 Visualization:

65. **L 239: Give the numbers of the figures here − > is a sub section (2.5) needed just for 1 sentence?**

Response: Agreed. We will delete subsection 2.5 for clarity in the revised manuscript.

Results

66. **- Fig3: in the caption: explain what values are represented in the box VS lines VS dotes**

Response: Figure 3 displays one boxplot per subregion showing the spatial distribution of the absolute biases (Top) and Wasserstein distances (Bottom) for each subregion. For each plot, the boxes span the interquartile range (25-75 percentiles) of the metric across all grid cells in that region; the horizontal line inside each box marks the median; whiskers extend to 1.5 × IQR; and individual dots represent grid cell outliers beyond the whiskers. The upper panel uses this format for absolute bias, and the lower panel does the same for the Wasserstein distance.

67. **- L 242: "we apply a spatial smoothing procedure to the bias corrected daily temperature fields" can you justify why?**

Response: Temperature fields are known to be highly spatially coherent because surface temperature is controlled by smoothly varying factors—elevation, vegetation, soil moisture, land use—that modulate the surface energy balance. Consequently, adjacent areas tend to share similar temperatures. Recent work supports this physical picture: Kunz and Laepple (2024) identify a "large-amplitude, short-distance component" in temperature spatial-correlation functions, confirming strong local correlations.

As an empirical validation that the CCSM model realistically captures this expected spatial behavior, we have added a spatial autocorrelation analysis using a representative snapshot of the model's surface temperature field. Specifically, we compute the isotropic spatial autocorrelation, which measures how correlated the temperature is between locations separated by a given distance, averaged over all directions and presented in Figure 21. The line shows how correlated two grid cells are at increasing radius: for every possible pair of cells, we calculated their temperature correlation, grouped the pairs by distance, averaged those correlations, and plotted the result. Because the correlation stays well above 0.8 for up to 40 grid cells away (12*40=480 km), it justifies our use of local smoothing.

68. **- L 253: "Section 2.1.Figure 3" need a space before "Figure"**

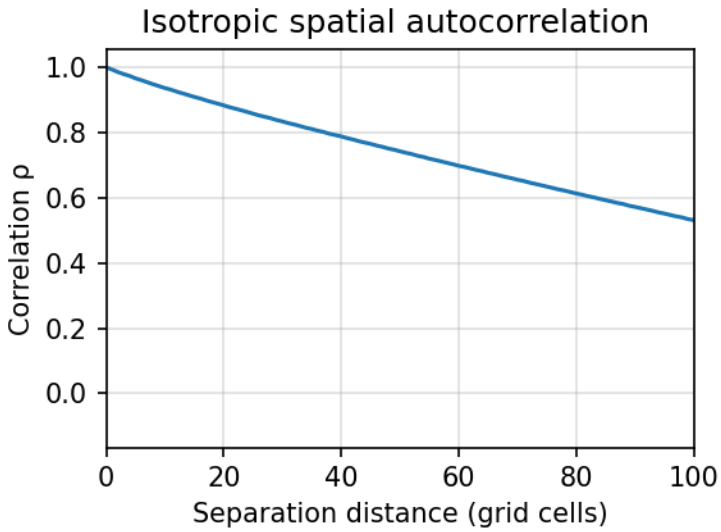

Figure 21: Isotropic spatial autocorrelation of daily surface temperature from the CCSM model at a representative time point. For each distance r, we select all pairs of grid points exactly r cells apart, align their anomaly values into two vectors, and compute a single Pearson correlation between those vectors. The strong correlation at short distances and its gradual decay with increasing separation illustrate the spatial coherence characteristic of temperature fields.

Response: We will correct it in the revised manuscript.

69. **- L 257: give the definition (and/or a reference) for the WD**

Response: The first-order Wasserstein distance (also called the Earth-Mover's Distance) is a metric that quantifies the minimal "cost" of transforming one probability distribution into another. It is well suited for comparing full distributions by the distance between them rather than individual moments and is therefore a natural choice for our quantile-based bias-correction evaluation. A more detailed description can be found at Panaretos and Zemel (2019). The following brief definition and formula will be added to the Appendix, and the term will be introduced at first mention in the main text:

WD is defined as a distance between two probability measures P and Q on a metric space $(\mathcal{X}, \|\cdot\|)$ by

$$W_p(P,Q) = \left(\inf_{\gamma \in \Gamma(P,Q)} \int_{\mathcal{X}\times\mathcal{X}} \|x-y\|^p \, \mathrm{d}\gamma(x,y) \right)^{1/p}, \tag{4}$$

where $\Gamma(P,Q)$ denotes the set of all couplings with marginals P and Q; throughout this study we use the common choice $p = 1$ (Panaretos and Zemel, 2019).

70. **- L.258: Fig 3 top: comment on why larger bias for some regions (S, N, Midwest)**

and larger uncertainties for SW, NW?

Response: The variations in bias magnitude and variability reflect the distinct geophysical and climatological characteristics of each sub-region. Our focus is to show that, regardless of those inherent differences, EMDBC reduces bias systematically across all regions.

71. **- Fig 3 bottom: why similar WD for all region although biases are bigger for some regions (top)? I would expect a larger WD for the regions with higher biases?**

Response: The WD assesses the similarity of entire distributions, whereas the absolute-bias metric in the top panel reflects average daily errors. A region can exhibit a noticeable daily bias yet still have a model distribution whose overall shape closely matches the observations. For example, because QDM explicitly aligns model quantiles with observed quantiles, it yields a small WD even if the resulting series retains some residual daily bias. EMDBC, in turn, lowers the daily bias across all regions (top panel, Figure 3) while preserving the distributional match (consistent lower WD compared to the original CCSM model projection).

72. **- Fig 4 and L273: with Northern and Midwest regions have larger biases for the GDM corrected datasets than for the not corrected datasets? $->$ should not be used then here?**

Response: Figure 4 shows bias for each EMD timescale separately, not the overall daily bias. In the North and Midwest, the high-frequency (bi-weekly) timescale is particularly noisy; after correction, its bias can appear slightly larger at that specific scale. The seasonal and annual timescales in those regions, however, exhibit substantial bias reductions. When these timescales and the residual are recombined, the net daily bias is lower, as demonstrated in the top panel of Figure 3. Therefore, a small increase in one narrow timescale does not contradict the overall improvement of the bias-corrected series.

73. **- Figure 5: add letters for sub plots (also in other figures):**

Response: Thank you for the suggestion. We have already restructured Figure 5 for greater clarity (see our response to Comment 74 for Figure 22). Given the new side-by-side layout and the descriptive caption, we believe subplot letters may no longer be necessary. However, if the reviewer prefers, we will gladly add alphabetical labels to each row in the revised manuscript.

74. **- Figure 5: difficult to see something. In particular in the top 2 plots: we cannot see the curves. "While QDM achieves performance comparable to EMDBC at the daily (training) scale," $->$ how is this observed?**

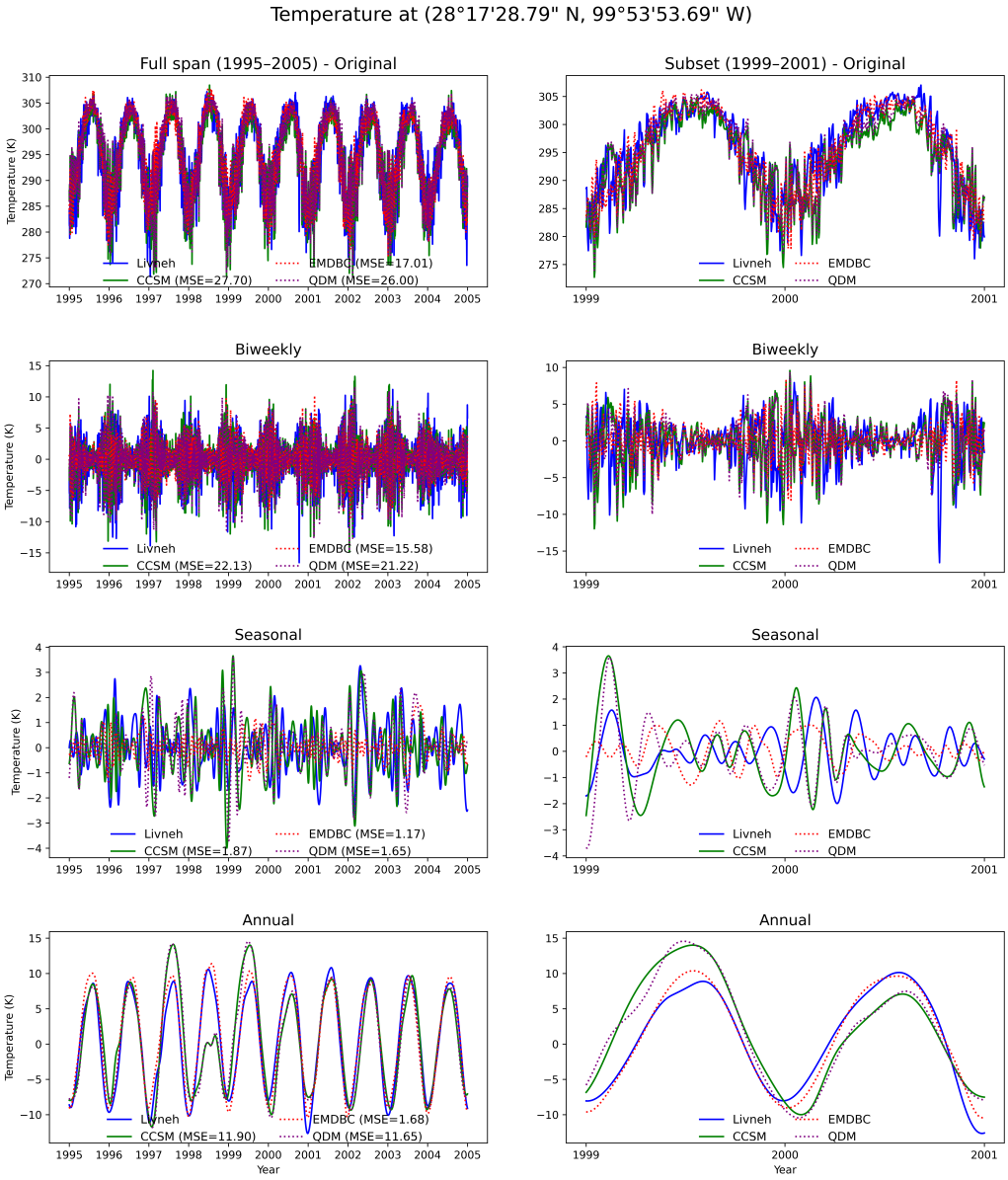

Figure 22: Bias-correction comparison across multiple time-scales at a representative grid cell. Column 1 shows the full 1995–2004 record, while Column 2 zooms into 1999–2001 for clarity. Solid lines correspond to the observed Livneh series (blue) and the raw CCSM projection (green); dashed lines show the bias-corrected outputs from EMDBC (red) and QDM (purple). Each row presents the original daily series and its bi-weekly, seasonal, and annual components, obtained by aggregating intrinsic mode functions as described in Section 2.4. The numbers at right report the mean-squared error (MSE, $^{\circ}K^2$) between each series and Livneh. While QDM matches EMDBC at the native daily scale, EMDBC yields consistently lower MSE at the bi-weekly, seasonal, and annual bands, indicating superior preservation of large-scale temperature variability.

Response: To improve readability and make the daily-scale comparison explicit, we have revised Figure 5 as follows:

(a) **Clearer layout.** The figure is now arranged in two side-by-side columns. Column 1 shows the full 1995–2004 record; Column 2 zooms into 1999–2001 so individual trajectories can be inspected. Livneh (blue) and raw CCSM (green) are plotted with solid lines, whereas the bias-corrected series—EMDBC (red) and QDM (purple)—are plotted with dashed lines. This colour/line-type combination allows each product to be distinguished even in the full-record view.

(b) **Quantitative metric.** Because overlapping curves remain hard to separate visually over ten years of daily data, we now report the *mean-squared error* (MSE) between each series and Livneh at the right-hand side of every row:

$$\mathrm{MSE}(x, y) = \frac{1}{N} \sum_{t=1}^{N} \big[x(t) - y(t)\big]^2,$$

with units $^{\circ}K^2$. These values show that QDM and EMDBC have nearly identical MSE at the native daily scale, confirming the statement in the text, while EMDBC yields lower MSE at the bi-weekly, seasonal, and annual timescales.

(c) **Caption updated.** The caption now explains the new layout, colour scheme, and the meaning of the MSE numbers.

75. **- Figure 5: in the legend add the "Reference datasets" to the "Livneh" legend. Change the Livneh curve to a dotted line.**

Response: We recognise the need to make Livneh's role as the reference/observed data set explicit. In the revised Figure 5, we now mention *"observed Livneh series"* to make it explicit. To maintain a clear distinction between solid (model projection or observational) and dashed (bias-corrected) series, we have kept Livneh as a solid blue line, while the model products remain dashed.

76. **- Figure 5: the biweekly plot is not commented in the caption**

Response: In the updated Figure 5 (now Fig.22) the caption explicitly describes the biweekly timescale plot, so this omission has been corrected.

77. **- Figure 6: explain what is plotted in the box plots (box, line, dotes) and in the violin plots (tail cutted? what is the black box and the white line?)**

Response: Figure 6 combines box–whisker plots (top row) with violin plots (right column) to convey both bias statistics and full temperature distributions:

- **Box–whisker panels (top row)**

  - The box spans the inter-quartile range (25th–75th percentiles) of the absolute daily bias across all grid cells in the sub-region.
  - The horizontal line inside the box marks the median bias.
  - Whiskers extend to 1.5 × IQR; points beyond the whiskers are plotted individually as outliers.
- **Violin panels (right column)**
  - The grey silhouette shows a kernel-density estimate of the region-averaged temperature distribution. Tails are truncated at the observed minimum and maximum values.
  - The black rectangle inside the violin denotes the inter-quartile range.
  - The white horizontal line indicates the median temperature.

78. **- Figure 6: not color-blind compatible?)**

Response: We agree. In the revised manuscript we will replace the current colour scheme in Figure 6 with a colour-blind–friendly palette to ensure accessibility.

79. **- L. 275 "These results demonstrate that EMDBC successfully preserves bias-corrected signals over a broad range of temporal frequencies" − > but it is a specific dataset, is it representative of other regions and other periods?**

Response: We chose the 1995–2004 WRF–CCSM regional simulation because it is a publicly accessible, high-resolution (≈ 12 km) North-American temperature data set that has been thoroughly evaluated in earlier studies and is comparable in quality to NA-CORDEX. It provides a reliable and well-documented testbed for demonstrating the EMDBC workflow without the added uncertainty of a less-studied archive. Future work will extend the method to other regions and periods to confirm its broader applicability.

80. **- 3.2 "Over full domain" which domain here? temporal spatial? which domain was used previously?**

Response: Here "full domain" refers to the entire spatial domain used in this study, namely the contiguous United States, rather than the seven 25 × 25 grid-cell sub-regions used in the validation analysis. The temporal period is unchanged: historical (1995–2004), mid-century (2045–2054), and late-century (2085–2094).

81. **- Figure 7: what should "Mid-century" violin plots be compared too? it is misleading to have them on the same plot at the historical data − > it should not be compared to it? Maybe bring the violins that should be compared closer to one another. E.g. 3 block: 1 historical bloc (Livneh + historical models) − > then a gap − > 1 block with 3 mid-century violins-¿ gap − > 1 bog with 3 Late-century violins**

Response: We thank the reviewer for the suggestion. In Figure 7 all time-period violins share a common axis so readers can see two things at once: (i) QDM and EMDBC preserve the overall shape of the observed historical temperature distribution, and (ii) both methods retain the projected mid-century and late-century shifts produced by WRF-CCSM. Plotting the timeframes on separate blocks would make those joint comparisons less immediate. To improve clarity we have revised the figure caption to include these points. The updated caption is included in Figure 16

82. **- L 285 "In each sub-region, the top panel compares the absolute temperature bias between the model projected and the observed series before and after correction with EMDBC and QDM, whereas the right panel shows the distribution of the average temperature" − > "the top panel"... "whereas the right panel" ? is it here "whereas the bottom panel"?**

Response: Thanks for checking this. We will correct this in the revised manuscript.

83. **- Fig 8: CONUS is not defined?**

Response: Thank you for pointing this out. We now define continental United States (CONUS) in the manuscript text before its use in Figure 8:

> *Turning next to broader spatial analyses, Figure 6 focuses on various sub-regions across continental United States (CONUS). In each sub-region, ...*

84. **- L 295: Fig 9 and Fig8 − > Fig 8 and Fig 9**

Response: Thank you for pointing this out, we have revised the text to read: "*Figures 8 and 9*"

85. **- Fig 8 and Fig9. give letters to the subplots**

Response: Thank you for the recommendation. We have revised Figures 8 and 9 to include lettered titles and documented these revisions in our response to Comment 4 above.

86. **- L 295: please comment each of the subplots or remove if not used in the text.**

Response: Thank you for pointing out that we do not explicitly reference each subplot in the manuscript text. All nine plots in both Figures 8 and 9 are important. This paragraph has been revised to directly mention each subplot by lettered subtitle:

> *Finally, Figures 8 and 9 show the average predicted daily bias ((d)−(f)) and the corresponding spatial maps (((g)−(i)); e.g., annual or multi-year averages) for the raw and bias-corrected WRF-CCSM outputs. For reference, average Livneh observation*

data is also plotted, along with the average WRF-CCSM historical bias before and after correction (((a)−(c))). Here, "predicted bias" refers to the difference between the modeled temperature and its bias-corrected counterpart. EMDBC generally applies a stronger correction than QDM, resulting in slightly cooler daily temperature fields and a more uniform reduction of bias across the domain. Although we cannot fully validate future-period corrections in the absence of observations, EMDBC's stronger alignment with historical data and its lower bias in validation sub-regions suggest it is well-equipped to handle changing conditions while preserving both short- and long-term temperature variability.

87. **- Fig 8 and 9: why northern America has large biases (left column)?**

Response: Thank you for this question. We have not conducted a specific analysis of the warm bias in northern North America. A posible contributor is the land–atmosphere interactions related to snow handling. Investigating these processes would require additional experiments outside the scope of the present methods paper, but we have noted this as an avenue for future work.

Conclusion:

88. **L303-304: add the spatial extent of the output on which it is applied, and the resolution.**

Response: Thanks for the suggestion. We will revise the first sentence of this section as:

This study proposes a new timescale-aware bias-correction methodology, EMDBC, and applies it to 12 km WRF-CCSM daily temperature simulations, covering historical (1995 − 2004), mid-century (2045 − 2054), and late-century (2085 − 2094) periods across the contiguous United States.

89. **L304: add "temporal" to "Temporal downscaled"**

Response: Thank you for the suggestion. Dynamical downscaling with WRF is done primarily to refine spatial resolution. To keep the terminology precise and consistent with its common usage in the literature we decided to retain the phrase *dynamical downscaling* without the qualifier "temporal." We have, however, clarified the sentence for readability in our response to Comment 88.

90. **L305: add the years to the periods: "historical 19..-19.., mid-century 20..-20.. and late-century 20..-20.."**

Response: Addressed in Comment 88.

91. **L312: "meteorological variables" − > atmospheric" ?**

Response: Thank you for the suggestion. We think "atmospheric variables" is the more appropriate term and have updated the manuscript.

92. **Code and reproducibility: The zenodo repository contains only 1 script for 1 figure that is not one of the figures of the paper. It would be nice to have the scripts to compute plots of the paper. Could the repository be added to a public github page ? The link given is not public: https://git.cels.anl.gov/jfeinstein/emdbc-paper (Feinstein, 2025).**

Response: Thank you for the suggestion. The repository is already available on our public GitHub page, as noted in an earlier thread on the discussion portal. The revised Code and Data Availability section now reads:

All Python scripts for the Empirical Mode Decomposition-based Bias Correction, the full-domain WRF-CCSM dataset used in this manuscript, and the validation areas mapping WRF-CCSM indices to 25×25 case study regions are available in a Zenodo repository at https://doi.org/10.5281/zenodo.15244202 (Ganguli et al. (2025). Livneh daily CONUS observational data (Livneh et al. (2013)), provided by NOAA Physical Sciences Laboratory (NOAA-PSL) in Boulder, Colorado, USA, are available at https://psl.noaa.gov/data/gridded/data.livneh.html (NOAA-PSL (2013)). For Livneh, daily mean temperatures are computed as the average of the daily minimum and maximum values. Finally, the Empirical Mode Decomposition-based Bias Correction code is also available in the EMDBC GitHub repository at https://github.com/jeremyfifty9/emdbc (Ganguli and Feinstein (2025)). Thank you.

We trust that our revisions and clarifications satisfactorily address your concerns and significantly improve the manuscript. Thank you again for helping us strengthen this work.

References

Akinsanola, A. A., Jung, C., Wang, J., and Kotamarthi, V. R. (2024). Evaluation of precipitation across the contiguous united states, alaska, and puerto rico in multi-decadal convection-permitting simulations. *Scientific Reports*, 14(1):1238.

Alizadeh, F., Roushangar, K., and Adamowski, J. (2019). Investigating monthly precipitation variability using a multiscale approach based on ensemble empirical mode decomposition. *Paddy and Water Environment*, 17:741–759.

Ashfaq, M., Bowling, L. C., Cherkauer, K., Pal, J. S., and Diffenbaugh, N. S. (2010). Influence of climate model biases and daily-scale temperature and precipitation events on hydrological impacts assessment: a case study of the united states. *Journal of Geophysical Research*, 115:D14.

Boé, J., Terray, L., Habets, F., and Martin, E. (2007). Statistical and dynamical downscaling of the seine basin climate for hydro-meteorological studies. *International Journal of Climatology*, 27:1643–1655.

Bukovsky, M. S. and Karoly, D. J. (2011). A regional modeling study of climate change impacts on warm-season precipitation in the central united states. *Journal of Climate*, 24(7):1985 – 2002.

Cannon, A. J., Sobie, S. R., and Murdock, T. Q. (2015). Bias correction of gcm precipitation by quantile mapping: how well do methods preserve changes in quantiles and extremes? *Journal of Climate*, 28(17):6938–6959.

Chen, F. and Dudhia, J. (2001). Coupling an advanced land surface–hydrology model with the penn state–ncar mm5 modeling system. part i: Model implementation and sensitivity. *Monthly Weather Review*, 129(4):569 – 585.

Christensen, J. H., Boberg, F., Christensen, O. B., and Lucas-Picher, P. (2008). On the need for bias correction of regional climate change projections of temperature and precipitation. *Geophysical Research Letters*, 35(20).

Corney, S., Grose, M., Bennett, J. C., White, C., Katzfey, J., McGregor, J., Holz, G., and Bindoff, N. L. (2013). Performance of downscaled regional climate simulations using a variable-resolution regional climate model: Tasmania as a test case. *Journal of Geophysical Research: Atmospheres*, 118(21):11,936–11,950.

Das, P., Zhang, Z., and Ren, H. (2022). Evaluation of four bias correction methods and random forest model for climate change projection in the mara river basin, east africa. *Journal of Water and Climate Change*, 13(4):1900–1919.

Dhawan, P., Dalla Torre, D., Niazkar, M., Kaffas, K., Larcher, M., Righetti, M., and Menapace, A. (2024). A comprehensive comparison of bias correction methods in climate model simulations: Application on era5-land across different temporal resolutions. *Heliyon*, 10(23):e40352.

Feng, S., Tan, Y., Kang, J., Zhong, Q., Li, Y., and Ding, R. (2024). Bias correction of tropical cyclone intensity for ensemble forecasts using the xgboost method. *Weather and Forecasting*, 39(2):323–332.

Fosso, O. B. and Molinas, M. (2018). Emd mode mixing separation of signals with close spectral proximity in smart grids. In *2018 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe)*, pages 1–6.

Ganguli, A. and Feinstein, J. (2025). EMDBC: Empirical mode decomposition-based bias correction.

Ganguli, A., Feinstein, J., Raji, I., Akinsanola, A., Aghili, C., Jung, C., Branham, J., Wall, T., Huang, W., and Kotamarthi, R. (2025). Full simulation data, validation indices, and frozen repository for the empirical mode decomposition based bias correction approach.

Grell, G. A. and Dévényi, D. (2002). A generalized approach to parameterizing convection combining ensemble and data assimilation techniques. *Geophysical Research Letters*, 29(14):38–1–38–4.

Gudmundsson, L. (2012). qmap: Statistical transformations for postprocessing climate model output. R package version 1.0–3. Available online at `https://cran.r-project.org/web/packages/qmap/`.

Gudmundsson, L., Bremnes, J. B., Haugen, J. E., and Engen-Skaugen, T. (2012). Downscaling rcm precipitation to the station scale using statistical transformations–a comparison of methods. *Hydrology and Earth System Sciences*, 16(9):3383–3390.

Haerter, J. O., Hagemann, S., Moseley, C., and Piani, C. (2011). Climate model bias correction and the role of timescales. *Hydrology and Earth System Sciences*, 15(3):1065–1079.

Hawinkel, P., Swinnen, E., Lhermitte, S., Verbist, B., Van Orshoven, J., and Muys, B. (2015). A time series processing tool to extract climate-driven interannual vegetation dynamics using ensemble empirical mode decomposition (eemd). *Remote Sensing of Environment*, 169:375–389.

Hawkins, E. and Sutton, R. (2009). The potential to narrow uncertainty in regional climate predictions. *Bulletin of the American Meteorological Society*, 90:1095–1107.

Hawkins, E. and Sutton, R. (2012). Time of emergence of climate signals. *Geophysical Research Letters*, 39:L01702.

Iacono, M. J., Delamere, J. S., Mlawer, E. J., Shephard, M. W., Clough, S. A., and Collins, W. D. (2008). Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models. *Journal of Geophysical Research*, 113:D13103.

Jacob, D., Petersen, J., Eggert, B., Alias, A., Christensen, O. B., Bouwer, L. M., Braun, A., Colette, A., Déqué, M., Georgievski, G., Georgopoulou, E., Gobiet, A., Menut, L., Nikulin, G., Haensler, A., Hempelmann, N., Jones, C., Keuler, K., Kovats, S., Kröner, N., Kotlarski, S., Kriegsmann, A., Martin, E., van Meijgaard, E., Moseley, C., Pfeifer, S., Preuschmann, S., Radermacher, C., Radtke, K., Rechid, D., Rounsevell, M., Samuelsson, P., Somot, S., Soussana, J.-F., Teichmann, C., Valentini, R., Vautard, R., Weber, B., and Yiou, P. (2014). Euro-cordex: new high-resolution climate change projections for european impact research. *Regional Environmental Change*, 14(2):563–578.

Kim, K. B., Kwon, H.-H., and Han, D. (2016). Precipitation ensembles conforming to natural variations derived from a regional climate model using a new bias correction scheme. *Hydrology and Earth System Sciences*, 20(5):2019–2034.

Kim, T., Shin, J.-Y., Kim, S., and Heo, J.-H. (2018). Identification of relationships between climate indices and long-term precipitation in south korea using ensemble empirical mode decomposition. *Journal of Hydrology*, 557:726–739.

Kunz, T. and Laepple, T. (2024). Effective spatial degrees of freedom of natural temperature variability as a function of frequency. *Journal of Climate*, 37(8):2505–2518.

Liu, H., Zhan, Q., Yang, C., and Wang, J. (2019). The multi-timescale temporal patterns and dynamics of land surface temperature using ensemble empirical mode decomposition. *Science of the Total Environment*, 652:243–255.

Livneh, B., Rosenberg, E. A., Lin, C., Nijssen, B., Mishra, V., Andreadis, K. M., Maurer, E. P., and Lettenmaier, D. P. (2013). A long-term hydrologically based dataset of land surface fluxes and states for the conterminous united states: Update and extensions. *Journal of Climate*, 26(23):9384 – 9392.

Maraun, D. (2016). Bias correcting climate change simulations — a critical review. *Current Climate Change Reports*, 2(4):211–220.

Mearns, L. O., Arritt, R., Biner, S., Bukovsky, M. S., McGinnis, S., Sain, S., Caya, D., Correia, J., Flory, D., Gutowski, W., Takle, E. S., Jones, R., Leung, R., Moufouma-Okia, W., McDaniel, L., Nunes, A. M. B., Qian, Y., Roads, J., Sloan, L., and Snyder, M. (2012). The north american regional climate change assessment program: Overview of phase i results. *Bulletin of the American Meteorological Society*, 93(9):1337 – 1362.

Mearns, L. O. et al. (2017). The na-cordex dataset, version 1.0. Accessed [date].

Miftahurrohmah, B., Kuswanto, H., Pambudi, D. S., Fauzi, F., and Atmaja, F. (2024). Assessment of the support vector regression and random forest algorithms in the bias correction process on temperatures. *Procedia Computer Science*, 234:637–644.

Misra, V. (2007). Addressing the issue of systematic errors in a regional climate model. *Journal of Climate*, 20(5):801 – 818.

Morrison, H., Thompson, G., and Tatarskii, V. (2009). Impact of cloud microphysics on the development of trailing stratiform precipitation in a simulated squall line: Comparison of one- and two-moment schemes. *Monthly Weather Review*, 137:991–1007.

Nikulin, G., Jones, C., Giorgi, F., Asrar, G., Büchner, M., Cerezo-Mota, R., Christensen, O. B., Déqué, M., Fernandez, J., Hänsler, A., van Meijgaard, E., Samuelsson, P., Sylla, M. B., and Sushama, L. (2012). Precipitation climatology in an ensemble of cordex-africa regional climate simulations. *Journal of Climate*, 25(18):6057 – 6078.

NOAA-PSL (2013). Livneh daily conus near-surface gridded meteorological and derived hydrometeorological data. `https://psl.noaa.gov/data/gridded/data.livneh.html`. Data provided by the NOAA PSL, Boulder, Colorado, USA. Accessed April 2025.

Noh, Y., Cheon, W. G., Hong, S. Y., and et al. (2003). Improvement of the K-profile model for the planetary boundary layer based on large eddy simulation data. *Boundary-Layer Meteorology*, 107:401–427.

Panaretos, V. M. and Zemel, Y. (2019). Statistical aspects of wasserstein distances. *Annual review of statistics and its application*, 6(1):405–431.

Piani, C., Haerter, J. O., and Coppola, E. (2009). Statistical bias correction for daily precipitation in regional climate models over europe. *Theoretical and Applied Climatology*, 99(1–2):187–192.

Piani, C., Weedon, G. P., Best, M., Gomes, S. M., Viterbo, P., Hagemann, S., and Haerter, J. O. (2010). Statistical bias correction of global simulated daily precipitation and temperature for the application of hydrological models. *Journal of Hydrology*, 395(3–4):199–215.

Rajulapati, C. R. and Papalexiou, S. M. (2023). Precipitation bias correction: A novel semi-parametric quantile mapping method. *Earth and Space Science*, 10(4):e2023EA002823. e2023EA002823 2023EA002823.

Riahi, K., Rao, S., Krey, V., Cho, C., Chirkov, V., Fischer, G., Kindermann, G., Nakicenovic, N., and Rafaj, P. (2011). Rcp 8.5—a scenario of comparatively high greenhouse gas emissions. *Climatic Change*, 109(1):33.

Rockmore, D. (2000). The fft: an algorithm the whole family can use. *Computing in Science Engineering*, 2(1):60–64.

Sarhadi, A., Burn, D. H., Johnson, F., Mehrotra, R., and Sharma, A. (2016). Water resources climate change projections using supervised nonlinear and multivariate soft computing techniques. *Journal of Hydrology*, 536:119–132.

Schulzweida, U. (2023). Cdo user guide.

Skamarock, W., Klemp, J., Dudhia, J., Gill, D. O., Barker, D., Duda, M. G., Huang, X.-Y., Wang, W., and Powers, J. G. (2008). A description of the advanced research WRF version 3. Technical report, University Corporation for Atmospheric Research.

Tang, B., Dong, S., and Song, T. (2012). Method for eliminating mode mixing of empirical mode decomposition based on the revised blind source separation. *Signal Processing*, 92(1):248–258.

Teutschbein, C. and Seibert, J. (2012). Bias correction of regional climate model simulations for hydrological climate-change impact studies: review and evaluation of different methods. *Journal of Hydrology*, 456–457:12–29.

Tong, Y., Gao, X. J., Han, Z. Y., Xu, Y., and Giorgi, F. (2021). Bias correction of temperature and precipitation over china for rcm simulations using the qm and qdm methods. *Climate Dynamics*, 57:1425–1443.

USGCRP (2023). Fifth national climate assessment. Technical report.

Wang, J. and Kotamarthi, V. R. (2015). High-resolution dynamically downscaled projections of precipitation in the mid and late 21st century over north america. *Earth's Future*, 3(7):268–288.

Wood, A. W. (2002). Long-range experimental hydrologic forecasting for the eastern united states. *Journal of Geophysical Research*, 107(D20).

Wood, A. W., Leung, L. R., Sridhar, V., and Lettenmaier, D. P. (2004). Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs. *Climatic Change*, 62(1):189–216.

Wu, Z. and Huang, N. E. (2009). Ensemble empirical mode decomposition: a noise-assisted data analysis method. *Advances in adaptive data analysis*, 1(01):1–41.