Evaluation of CMIP6 GCMs over the CONUS for downscaling studies

Moetasim Ashfaq*[1], Deeksha Rastogi[1], Joy Kitson[1,2], Muhammad Adnan Abid[3], Shih-Chieh Kao[1]
[1] Oak Ridge National Laboratory, Oak Ridge, TN, USA
[2] Department of Computer Science, University of Maryland, College Park, MD, USA
[3] Earth System Physics, Abdus Salam International Centre for Theoretical Physics, Trieste, Italy

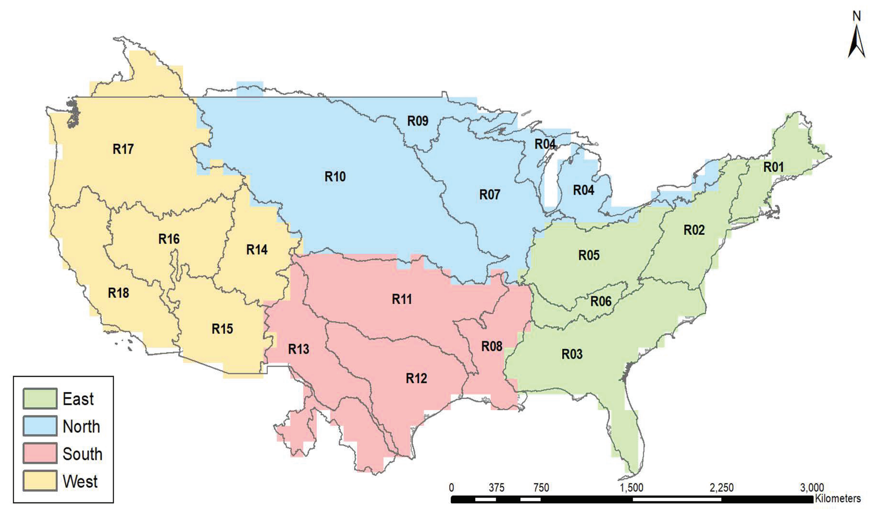
Abstract

Despite the necessity of Global Climate Models (GCMs) sub-selection in downscaling studies, an objective approach for their selection is currently lacking. Building on the previously established concepts in GCMs evaluation frameworks, we develop a weighted averaging technique to remove the redundancy in the evaluation criteria and rank 37 GCMs from the sixth phase of the Coupled Models Intercomparison Project over the contiguous United States. GCMs are rated based on their average performance across 66 evaluation measures in the historical period (1981–2014) after each metric is weighted between zero and one, depending on its uniqueness. The robustness of the outcome is tested by repeating the process with the empirical orthogonal function analysis in which each GCM is ranked based on its sum of distances from the reference in the principal component space. The two methodologies work in contrasting ways to remove the metrics redundancy but eventually develop similar GCMs rankings. A disparity in GCMs' behavior related to their sensitivity to the size of the evaluation suite is observed, highlighting the need for comprehensive multi-variable GCMs evaluation at varying timescales for determining their skillfulness over a region. The sub-selection goal is to use a representative set of skillful models over the region of interest without substantial overlap in their future climate responses and modeling errors in representing historical climate. Additional analyses of GCMs' independence and spread in their future projections provide the necessary information to objectively select GCMs while keeping all aspects of necessity in view.

Significance

The evaluation in this study is intended for downscaling studies where GCM sub-selection is necessary due to many unavoidable factors. We develop a weighted averaging technique for model evaluation that removes redundancy in the selected metrics. Additionally, we highlight the need for comprehensive multi-variable evaluation criteria at varying timescales to determine models' skillfulness over a region, their independence, and representativeness in capturing spread in future projections.

Study Area: The contiguous United States (CONUS) is divided into four Hydrological Unit Codes Level 2-based regions for the evaluation of GCMs.



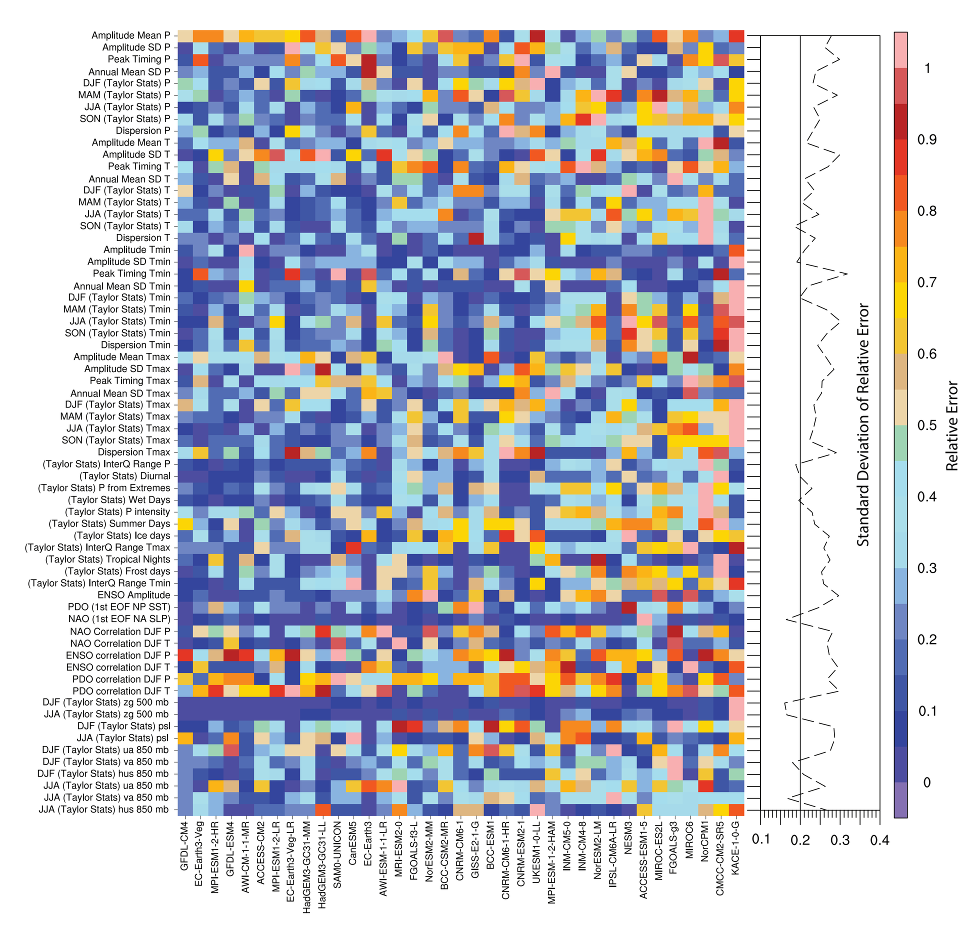
Relative Models Ranking

Models are ranked using two methodologies:
1) **Weighted Averaging**: Evaluation metrics are weighted based on their uniqueness so that highly correlated metrics are downweighed.

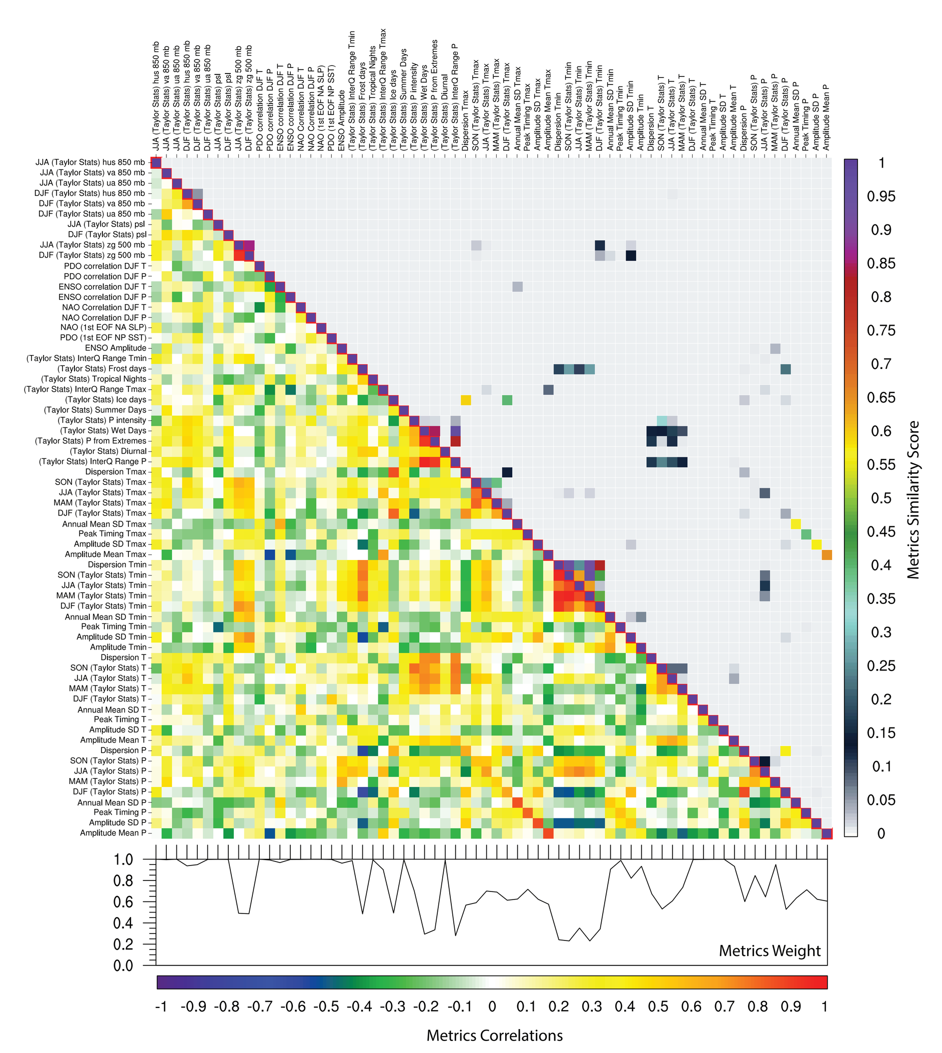
2) **EOF-based Strategy**: Accounts for the distance of each simulated metric from the reference in the PC space. Sum of Euclidean distances from the observations defines its rank.

1. GCMs Evaluation: Relative Error and Metrics Uniqueness

We analyze the performance of CMIP6 GCMs across sixty-six evaluation metrics . Many metrics exhibit considerable correlations. Models' relative ranking can be affected by the redundancy of information contained in these metrics. Weighted averaging based on metrics uniqueness solves this issue.



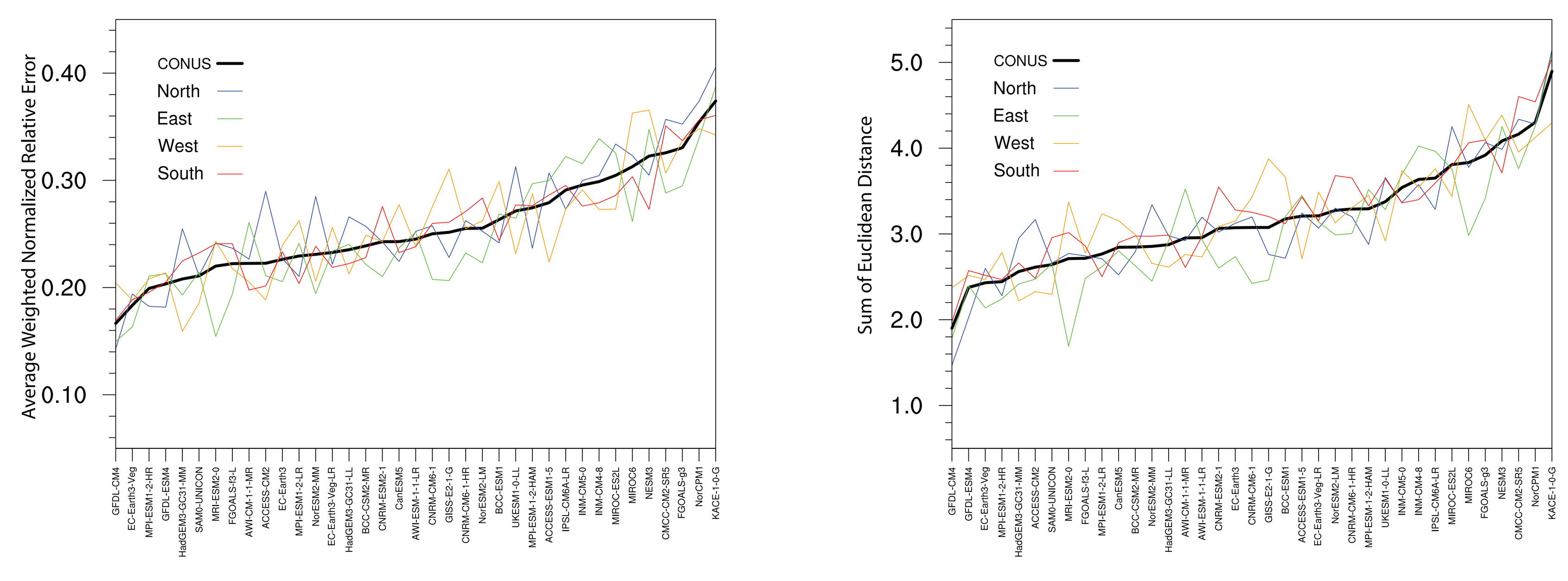
The unweighted relative errors of GCMs over the North (region). The left panel shows relative errors corresponding to each metric across all GCMs, and the line plot on the right shows the standard deviation of the relative error for each metric across all GCMs.



Metrics Uniqueness based on the similarity score. The correlation between the pairwise metrics (bottom triangle) and the corresponding similarity score (top triangle) over the North. Metrics with high correlations exhibit a high similarity score and are down-weighted. The line plot at the bottom shows the overall weight for each metric.

2. GCMs Ranking

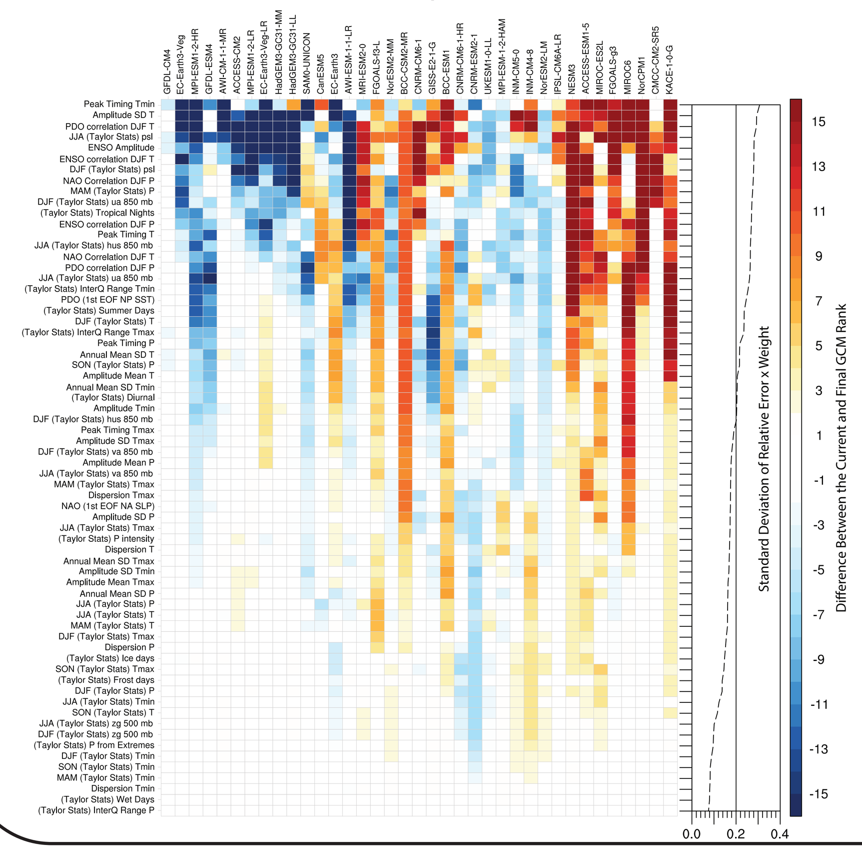
The regional and CONUS scale relative GCM rankings based on the two methodologies. The two approaches yield reasonably similar results at the CONUS scale.



The weighted averaging (left) and EOF-based Euclidian distances (right). The thin lines represent the models' relative ranking over four sub-regions, and the thick line represents the overall CONUS scale ranking.

3. Relative Importance of Individual Metrics

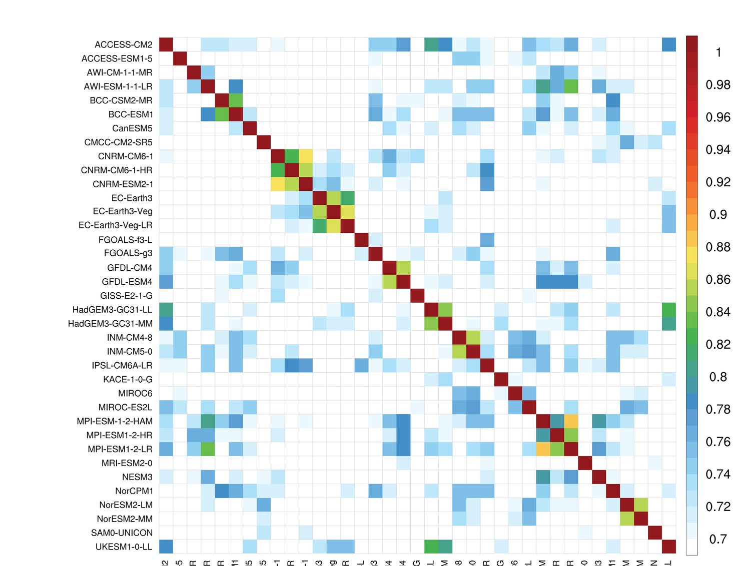
In the weighted averaging technique, the relative importance of an individual metric depends on two main factors: 1) the skill variation for that metric across the GCMs and 2) its weight or uniqueness. Substantial inconsistency exists in the performance of average-performing models. Therefore, using only a handful of metrics in an evaluation risks causing errors in the GCM selection process.

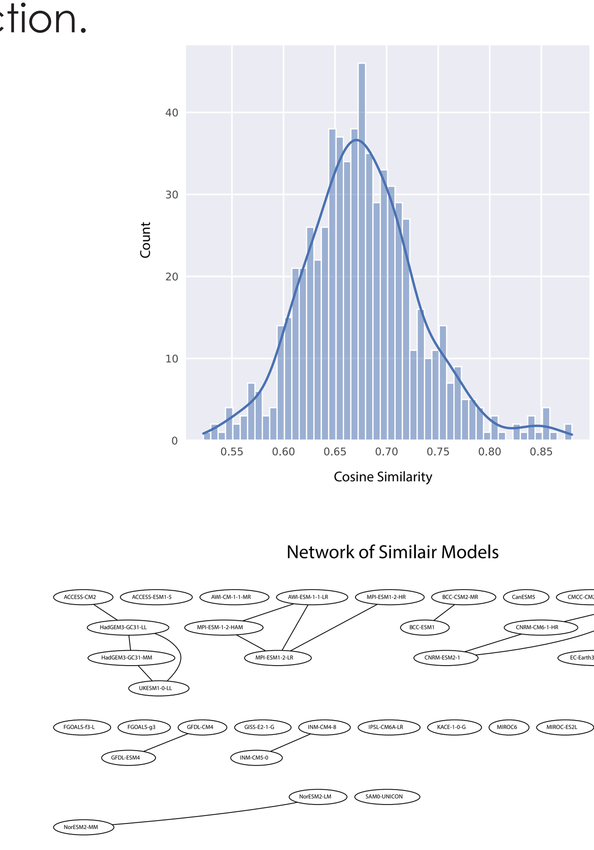


Difference between the current and final GCMs rank with the addition of evaluation metrics over the North region (left panel). Adding individual metrics in the weighted averaging follows the decreasingly ranked multiplicative product of their standard deviation and weights (line plot on the right), meaning that the ones with the higher magnitudes of this multiplicative product are considered first.

4. GCMs Independence

We use cosine similarity to quantify the independence or interdependence of GCMs. The cosine similarity of two vectors quantifies how close their directions are based on the cosine of the angle between them. The cosine similarity equals one when the two vectors point in the same direction, while it equals zero when the two vectors are orthogonal. Several models in the CMIP6 share modeling component and exhibit similar behavior. Therefore, model independence must be a consideration is sub-selection.

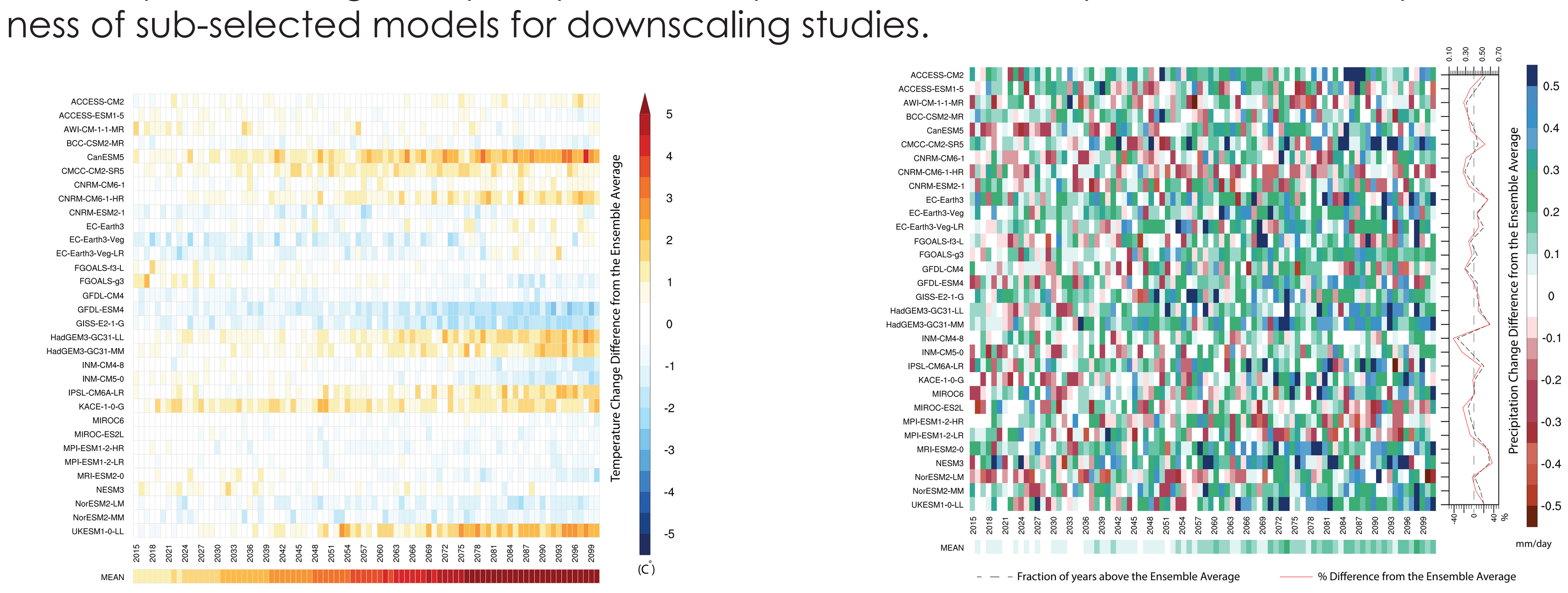




The cosine similarity score for each pair of GCMs using the weighted metrics data (left), its distribution across GCMs (top right) and the network of similar models with scores ≥ 0.8 (bottom right).

5. GCMs Regional Climate Sensitivity and Spread

Regional climate sensitivity is defined as GCMs simulated temperature changes over the region in future period of interest. Careful examination of regional climate sensitivity and GCMs spread in regional precipitation responses is necessary to ensure the representativeness of sub-selected models for downscaling studies.



Projected yearly change in temperature (left) and precipitation (right) over with reference to 1995–2014, as a difference from the ensemble mean (shown in the bottom row). The dotted line in the right line plot represents the fraction of years when the projected absolute change in an individual GCM is above the ensemble average. The red line in the right line plot represents the % difference between the ensemble average and each model in projected absolute precipitation changes over 2014–2100.

Ashfaq, M., Rastogi, D., Kitson, J., Abid, M. A., & Kao, S.-C. (2022). Evaluation of CMIP6 GCMs over the CONUS for downscaling studies. Journal of Geophysical Research: Atmospheres, 127, e2022JD036659. https://doi.org/10.1029/2022JD036659