

Earth's Future

RESEARCH ARTICLE

10.1029/2022EF002734

Key Points:

- Choice of observation used in the bias-correction of downscaled historical climate data can result in marked differences in the outcome
- Differences in future projections arising from the choice of downscaling techniques are more prominent than those from the observations
- A more comprehensive framework is needed to capture the full range of methodological uncertainties in climate data downscaling

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

D. Rastogi,
rastogi@ornl.gov

Citation:

Rastogi, D., Kao, S.-C., & Ashfaq, M. (2022). How may the choice of downscaling techniques and meteorological reference observations affect future hydroclimate projections? *Earth's Future*, 10, e2022EF002734. <https://doi.org/10.1029/2022EF002734>

Received 21 FEB 2022

Accepted 8 JUL 2022

© 2022 Oak Ridge National Laboratory, managed by UT- Battelle, LLC. This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs License](#), which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

How May the Choice of Downscaling Techniques and Meteorological Reference Observations Affect Future Hydroclimate Projections?

Deeksha Rastogi¹ , Shih-Chieh Kao² , and Moetasim Ashfaq¹ 

¹Oak Ridge National Laboratory, Computational Sciences and Engineering Division, Oak Ridge, TN, USA, ²Oak Ridge National Laboratory, Environmental Sciences Division, Oak Ridge, TN, USA

Abstract We present an intercomparison of a suite of high-resolution downscaled climate projections based on a six-member General Circulation Model (GCM) ensemble from Coupled Models Intercomparison Project (CMIP6). The CMIP6 GCMs have been downscaled using dynamical and statistical downscaling techniques based on two meteorological reference observations over the conterminous United States. We use the regional climate model, RegCM4, for dynamical downscaling, double bias correction constructed analogs method for statistical downscaling, and Daymet and Livneh datasets as the reference observations for statistical training and bias-correction. We evaluate the performances of downscaled data in both historical and future periods under the SSP585 scenario. While dynamical downscaling improves the simulation of some performance evaluation indices, it adds an extra bias in others, highlighting the need for statistical correction before its use in impact assessments. Downscaled datasets after bias-correction compare exceptionally well with observations. However, the choice of downscaling techniques and the underlying reference observations influence the hydroclimate characteristics of downscaled data. For instance, the statistical downscaling generally preserves the GCMs climate change signal but overestimates the frequency of hot extremes. Similarly, simulated future changes are sensitive to the choice of reference observations, particularly for precipitation extremes that exhibit a higher projected increase in the ensembles trained and/or corrected by Daymet than Livneh. Overall, these results demonstrate that multiple factors, including downscaling techniques and reference observations, can substantially influence the outcome of downscaled climate projections and stress the need for a comprehensive understanding of such method-based uncertainties.

1. Introduction

Investigation of climate change impacts at local to regional scales requires spatially resolved climate projections to sufficiently capture fine-scale responses and feedbacks (Ashfaq et al., 2016). The average horizontal grid spacing of General Circulation Models (GCMs) in the latest sixth Phase of Coupled Models Intercomparison Project (CMIP6) ensemble is still coarser than 1° (~110 km), which warrants the need for further spatial refinement before their use for climate change impact assessments (Maurer & Hidalgo, 2008). This limitation is currently overcome by the downscaling of GCMs using process-based dynamical downscaling (also known as regional climate modeling), empirical-based statistical downscaling, or a combination of both (Ashfaq et al., 2016; Maurer et al., 2013; Wood et al., 2004).

A regional climate model (RCM) is forced with prognostic variables from coarse resolution GCM simulations at its lateral boundaries in sub-daily time steps for dynamical downscaling. The design of regional model experiments allows RCMs to independently resolve fine-scale responses, such as those due to complex topography, coastal lines, and mesoscale convection within its domain, which are often poorly represented in the numerical solution of GCMs. Therefore, while computationally intensive and time-consuming, the dynamical downscaling approach enables the investigation of complex processes by providing a suite of physically consistent variables. However, due to the influence of numerical model bias, outputs of RCM (e.g., precipitation) should still be bias-corrected before their use in downstream modeling approaches, such as high-resolution hydrologic simulations (Potter et al., 2020; Hnilica et al., 2017; Addor & Seibert, 2014; Ashfaq et al., 2010).

On the other hand, statistical downscaling is based on the empirical relationships between the simulated and observed climate during the historical period. It is computationally efficient and can downscale large ensembles of GCMs representing different future climate scenarios. However, statistical downscaling is limited to only those

variables for which long-term observational records are available (Benestad, 2004). Moreover, it also intrinsically assumes that the empirical relationships between the simulated and observed climate remain stationary, which is likely invalid at higher radiative forcing levels in the future when Earth system dynamics is expected to be statistically different from the present day. The statistical downscaling also depends on the choice and length of reference data, and both factors may significantly impact the outcome.

While both approaches for the downscaling of GCMs have been in use for decades, their quantitative differences remain unclear. Most evaluations of downscaled hydroclimate projections are limited to the historical period and at a regional scale (Ayar et al., 2016; Bettolli et al., 2021; Manzanas et al., 2018; Nikulin et al., 2018; Roux et al., 2018; Yoon et al., 2012), as limited studies have evaluated similarities and/or disagreements in their downscaled climate change and subsequent climate change impacts (Jiang et al., 2018; Li et al., 2018; Mezghani et al., 2019). Since downscaled climate simulations data have been widely used in many hydroclimate risk assessments (Batibeniz et al., 2020; Naz et al., 2016; Pagan et al., 2016; Rastogi et al., 2020, 2019), the need for a thorough understanding of methodological uncertainties associated with the choices in the downscaling approaches cannot be overstated. Likewise, how does the choice of meteorological reference observations – used in the bias correction of dynamically downscaled projections or in establishing empirical relationships between GCM simulation and observations in statistical downscaling – impact the outcome is a factor which is often overlooked or disregarded. Therefore, not only is there a need for a more thorough comparison of different downscaling approaches, but there is also a need to understand the uncertainties related to the observational data in the downscaling process.

Owing to the need for comparative analyses between dynamical and statistical downscaling approaches and between downscaling using multiple observations, this study investigates how the choice of downscaling techniques, and meteorological reference observations, can affect future hydroclimate responses representing average and extreme states. We construct these comparisons using data from six CMIP6 GCM simulations that have been downscaled using a regional climate model, a statistical technique, and two observations over the conterminous United States (CONUS). While these downscaling techniques have been extensively applied and evaluated on previous generations of CMIPs (Ashfaq et al., 2016; Mearns et al., 2017; Pierce et al., 2015), the studies using the latest CMIP6 are just getting started. To the best of our knowledge, this is perhaps the first CONUS-scale intercomparison of multiple CMIP6-based downscaled climate projections in both historical and projected future periods that evaluates the differences arising from the choice of downscaling approaches and meteorologic observations. As previously noted, uncertainties in the regional downscaling of future climate projections can be detrimental in the long-term planning to cope with the impacts of climate change. For instance, underestimating regional-scale climate change responses can result in under-preparedness from a planning and mitigation perspective. Contrarily, overestimating these responses may cause overbudgeting to deal with the consequences. While we aim to evaluate the key differences arising from the downscaling choices, understanding the factors that drive these changes require distinctively different analytical frameworks beyond the scope of the study. Nevertheless, we hope that these comparative analyses should inform the effort to develop reliable approaches for future impact assessment and mitigation planning.

2. Data and Methodology

2.1. Data

Daily precipitation (P), maximum temperature (T_{\max}), and minimum temperature (T_{\min}) are obtained from the following two meteorologic reference observations:

1. Daymet: maintained by the Distributed Active Archive Center at Oak Ridge National Laboratory (ORNL DAAC). It provides daily gridded surface weather parameters over North America at a 1-km horizontal grid spacing from 1980 to 2019 (Thornton et al., 2021). For this study, we employ the latest Daymet V4, which includes observational timing correction for daily precipitation and maximum temperature, along with other enhancements.
2. Livneh: initially produced by the University of Colorado at Boulder (UCB), updated version available from the University of California Los Angeles. It provides daily gridded meteorological estimates over CONUS, Mexico, and the part of Canada from 1950 to 2018 at a $1/16^\circ$ (~ 6 km) horizontal grid spacing (Livneh

Table 1
List of CMIP6 GCMs Used for Downscaling

CMIP6 GCM name	Spatial resolution	Ensemble member	GCM institute	References
ACCESS-CM2	144 × 192	r1i1p1f1	The commonwealth Scientific and Industrial Research Organization, Australia	Dix et al. (2019)
BCC-CSM2-MR	160 × 320	r1i1p1f1	Beijing Climate Center	Wu et al. (2018)
CNRM-ESM2-1	256 × 128	r1i1p1f2	French Center National de la Recherche Scientifique	Seferian (2018)
MRI-ESM2-0	160 × 320	r1i1p1f1	Meteorological Research Institute Japan	Yukimoto et al. (2019)
MPI-ESM1-2-HR	192 × 384	r1i1p1f1	The German Climate Computing Center	von Storch et al. (2017)
NorESM2-MM	192 × 288	r1i1p1f1	Multi-institutional, coordinated climate research project in Norway	Bentsen et al. (2019)

Note. CMIP, Coupled Models Intercomparison Project; GCM, General Circulation Model.

et al., 2015). The new update also corrects a timing adjustment issue and improves the characterization of precipitation extremes (Pierce et al., 2021).

Given their differences in the original spatial resolution, both datasets were either aggregated or bilinearly interpolated to common 1° and 1/24° (~4 km) grids to support downscaling and evaluation. The 1° resolution grids are the common grids used to process all raw CMIP6 GCM outputs. The 1/24° is an intermediate resolution between Daymet and Livneh that we selected for evaluation and can also feed into an existing hydrologic model (Naz et al., 2016) in the following hydrologic modeling efforts. As reported by Henn et al. (2018), local differences in different gridded datasets exist, and the differences are not simply due to their different data resolution. The impacts of different meteorologic reference datasets on the resulting hydroclimate impact assessment are subjects in the subsequent study.

We chose Daymet and Livneh as meteorological reference observations since they have been widely used in many hydroclimate studies. While both datasets generally look similar at the regional scale, noticeable local differences exist (e.g., see Henn et al., 2018), and the differences are not simply due to their different spatial resolution. These local differences may pass onto the downscaled data products and affect the resulting impact assessment (e.g., see Alder & Hostetler, 2019). For instance, Livneh exhibits cold bias in T_{\min} over the western United States. These biases in T_{\min} are associated with the method and interpolation technique used when developing the gridded data set (Walton & Hall, 2018). When using Livneh as the meteorological reference observation, the downscaled data product will possess similar features that need to be better acknowledged.

We select 6 CMIP6 GCMs (Table 1) based on the availability of 6-hourly atmospheric data for dynamical downscaling and their historical performance over the CONUS which are detailed in Ashfaq et al., 2022. The data for selected GCMs is obtained from Earth System Grid Federation (ESGF) archives (<https://esgf-node.llnl.gov/search/cmip6>) for 1980–2014 in the historical period and 2015–2059 in the future period under the SSP585 scenario. Before statistical downscaling, all CMIP6 GCMs were also remapped to the common 1° × 1° latitude-longitude grids.

2.2. Methods

2.2.1. Dynamically Downscaling and Correction

We use the Regional Climate Model version 4 (RegCM4), developed and maintained by the Abdus Salam International Center for Theoretical Physics (Giorgi et al., 2012), to dynamically downscale the CMIP6 GCMs (outputs referred to as RCM_O). The RegCM4 is configured at 25 km horizontal grid spacing and 18 vertical levels over a domain that covers CONUS, parts of Mexico and Canada (Figure S1 in Supporting Information S1). It uses the hydrostatic dynamical core from the Fifth Generation Mesoscale Model (MM5) (Grell et al., 1994), the radiation package from the Community Climate Model version 3 (CCM3) (Kiehl et al., 1998), the Community Land Model (CLM) version 4.5 (Tawfik & Steiner, 2011), the Holtslag boundary layer package (Holtslag et al., 1990), the Subgrid Explicit Moisture Scheme (SUBEX) scheme of Pal et al. (2000) and the Tiedke cumulus convection parameterization (Tiedtke, 1989). RegCM4 and its earlier versions have been extensively used to downscale CMIP5 and earlier generations of GCMs over North America [Ashfaq et al., 2010, 2013 and 2016;

Diffenbaugh & Ashfaq, 2010]. In the current configuration, the RegCM4 grid was centered at 39.00°N and 98.00°W. It consisted of 157 points in the latitude direction and 227 points in the longitude direction, covering CONUS, parts of Mexico and Canada. Further, daily scale bias-correction (method described in the following section) was applied to P, T_{\max} and T_{\min} from each of the RegCM simulations using both Daymet (outputs referred to as RCM_{CD}) and Livneh (outputs referred to as RCM_{CL}) at 1/24° (~4 km) spatial grid.

2.2.2. Statistical Downscaling

For statistical downscaling, we apply the Double Bias Correction Constructed Analogs (DBCCA) method (Werner & Cannon, 2016) to daily scale P, T_{\max} and T_{\min} from the six selected CMIP6 GCMs using both Daymet (outputs referred to as DBCCA_D) and Livneh (outputs referred to as DBCCA_L) as the training datasets. DBCCA is based on the widely used Bias Correction Constructed Analogs (BCCA) method developed by Maurer et al. (2010), with another round of bias correction at the end to fix some known BCCA issues such as drizzling and residual biases. Werner and Cannon (2016) conducted a comprehensive evaluation of several statistical downscaling methods and reported DBCCA as one of the two best-performing methods in their case study. DBCCA includes four main steps, in which steps 1 and 2 are conducted at the 1° horizontal grid spacing, and steps 3 and 4 are performed at the 1/24° horizontal grid spacing.

1. *First Bias Correction* (at 1° grid resolution): The observations are aggregated to match the 1° latitude-longitude GCM grid. For each calendar day and each GCM grid, a 31-day (± 15 days) window is used to create the cumulative distributions across all years from observations, GCM baseline, and GCM future. Bias correction is then applied to the daily GCM output using a quantile mapping approach (Thrasher et al., 2012). A maximum of 150% future to baseline quantile ratio is enforced for precipitation to avoid unreasonably large adjustments (Gutman et al., 2014).
2. *Analogue Selection and Weight Determination* (at 1° grid resolution): For each GCM day, a 91-day (± 45 days) window is used to create a historic library across all years in observation. A total of 30 analogs from the historic library that best match the spatial pattern of the targeted GCM day are selected. The statistical regression technique is then used to determine the weights of each analogue (Maurer et al., 2010).
3. *Construct Analogue* (at 1/24° grid resolution): Assuming the same analogs and weights exist at both spatial resolutions, the previous step is reversed. The 30 analogs from observation at the 1/24° grid are first looked up. A linear combination of these 30 analogs with the identified weights is formed as the downscaled climate variable. Output at this step is the conventional BCCA.
4. *Second Bias Correction* (at 1/24° grid resolution): A second daily bias correction is applied to the 1/24° constructed analogs to generate the final results of DBCCA. In addition, to maintain a reasonable diurnal temperature range (DTR; Thrasher et al., 2012), we apply the second bias correction on T_{\max} and DTR. T_{\min} is then calculated based on the corrected T_{\max} and DTR ($T_{\min} = T_{\max} - \text{DTR}$). This is also the same step we used to bias-correct the RegCM outputs.

2.2.3. Analysis

We evaluate the performance of the six datasets: the original GCMs (GCM_O), RCM_O, RCM_{CD}, RCM_{CL}, DBCCA_D, and DBCCA_L in the historical period by using annual and seasonal averages of P, T_{\max} , and T_{\min} , annual 95th percentile of P (P_{95}) and T_{\max} (T_{95}), fifth percentile of T_{\min} (T_{05}), the annual number of rainy days with $P > 1$ mm (R_{day}), and frost days with $T_{\min} < 0$ (F_{day}). These metrics are also used to investigate differences in the projected future period changes across the six datasets. We also compare future period changes in the number of days with P above historical P_{95} , T_{\max} above historical T_{95} , and T_{\min} below historical T_{05} in these ensembles. Further, we compare the output from the downscaled dataset with the GCM_O. We spatially interpolate GCM_O and the corresponding RCM_O outputs to the common 1/24° (~4 km) spatial grid spacing for ease of comparison. The historical comparisons are based on several statistical measures, including the Pearson product-moment coefficient of linear correlation (also known as pattern correlation (PC)), the ratio of spatial standard deviation (ROS), and absolute bias in the annual and seasonal magnitudes of each metric.

Further, we present the historical comparison and future changes using spatial maps, heat maps, time series plots, and plots of probability density function (PDF). The analysis is presented for 1980–2019 reference and 2020–2059 projection periods. The reference period represents the most recent historical period, while the projection period represents the near-term future for which the downscaling has been conducted. Further, we chose the highest emission to evaluate hydroclimate changes under the potential worst-case scenario. The projected future changes

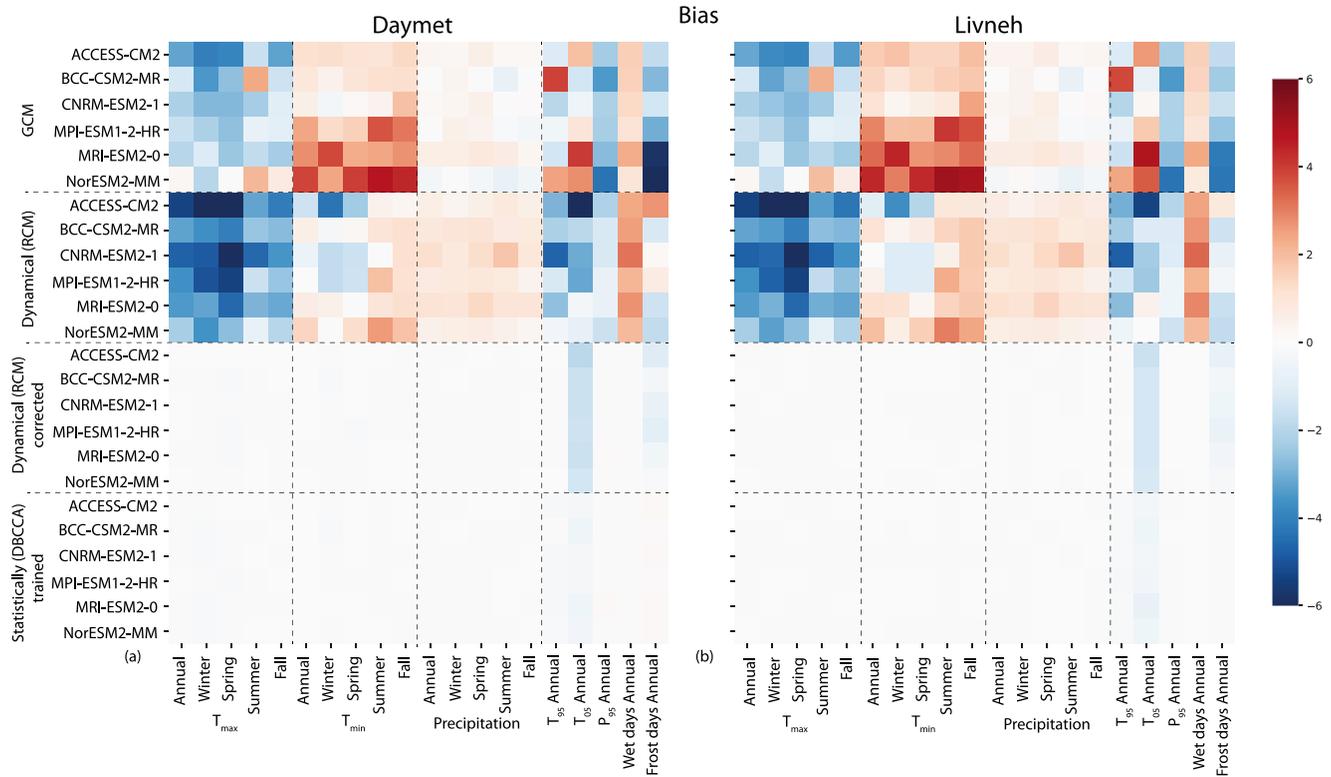


Figure 1. Heat maps showing CONUS scale biases in annual and seasonal scale daily maximum temperature (T_{\max}), daily minimum temperature (T_{\min}), and Precipitation (P) and annual magnitude of 95th percentile of T_{\max} (T_{95}), fifth percentile of T_{\min} (T_{05}), 95th percentile of P (P_{95}), number of wet days (R_{days}) and number of frost days (F_{days}) in the original GCMs (GCM_o), dynamically downscaled with RegCM4 (RCM_o), dynamically downscaled and corrected with Daymet (RCM_{CD}) and Livneh (RCM_{CL}) and statistically trained and corrected simulations with Daymet ($DBCCA_D$) and Livneh ($DBCCA_L$) as compared to (a) Daymet and (b) Livneh observations respectively. The simulations trained and/or corrected with an observation are compared with the reference observation. The biases in T_{\max} , T_{\min} , Precipitation, T_{95} , T_{05} , and P_{95} are presented in °C, °C, mm/day, °C, °C, and mm/day respectively whereas those in the number of wet days and number of frost days are normalized using the mean ensemble bias for presentation and comparison.

are presented as the average differences between 2020–2059 and 1980–2019. Statistical significance of future change is tested using Student's t -test at a 95% confidence interval. For the PDF plots, the difference for each year in the future period is calculated with respect to the average historical climatology for each grid point, and the plots are then created using all the grid points over CONUS. For P, T_{\min} and T_{\max} , monthly data is used, whereas for P_{95} , T_{95} , T_{05} , number of days above P_{95} and T_{95} , number of days below T_{05} , annual data is used to create the PDFs. Time series plots are plotted as anomalies with respect to 1981–2018. We use nine National Center for Environmental Information (NCEI) regions to perform regional-scale analysis following Ashfaq et al. (2016) that also use the NCEI climate regions to evaluate projected changes at a regional scale. The climate within each NCEI region exhibit relatively consistent characteristics, which enables us to compare these datasets at the sub-CONUS scale but over areas with uniform characteristics.

3. Results

3.1. Performance Evaluation of Historical Climate Ensembles

We assess the performance of historical ensembles by evaluating biases, pattern of correlation (PC) and ratio of standard deviation (ROS) of various metrics with respect to reference observations (Figures 1–3, Tables S1–S6 in Supporting Information S1). The following section provides details on these statistical measures.

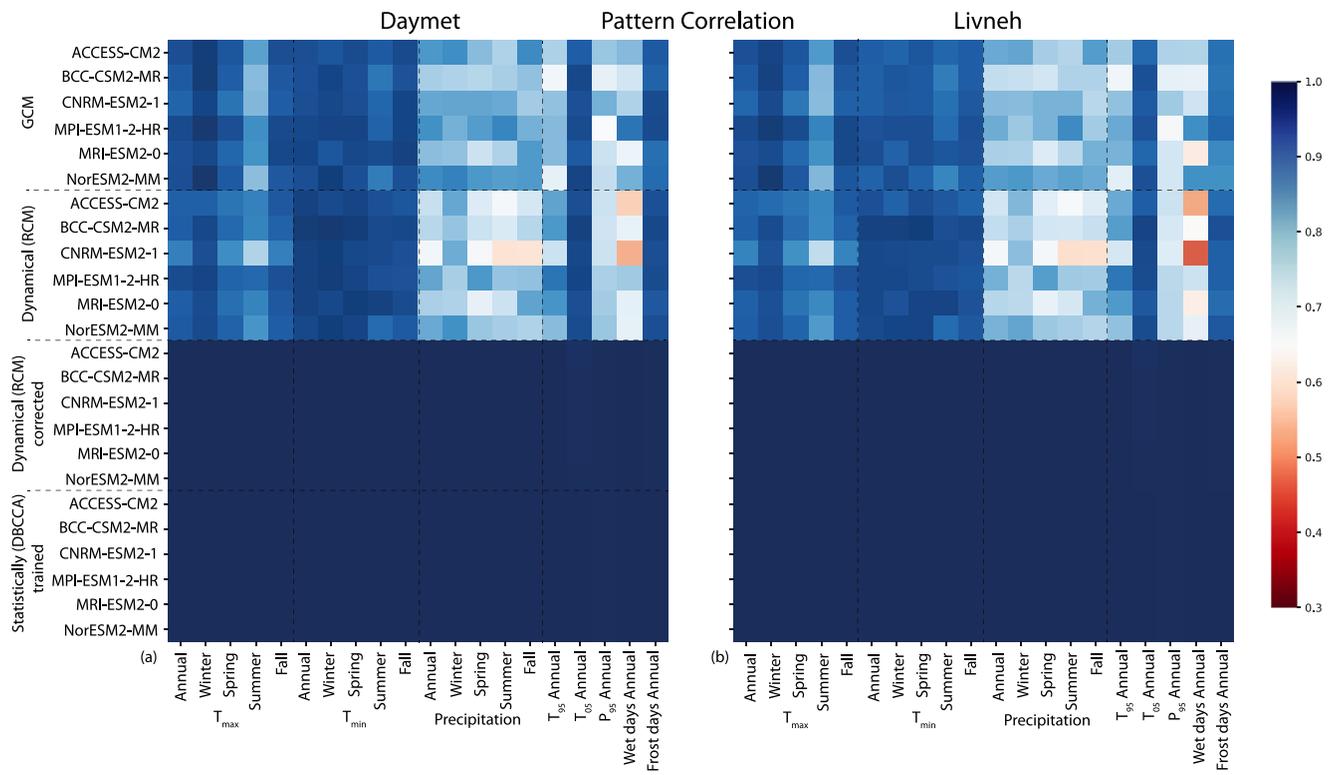


Figure 2. Same as Figure 1 but for pattern correlation.

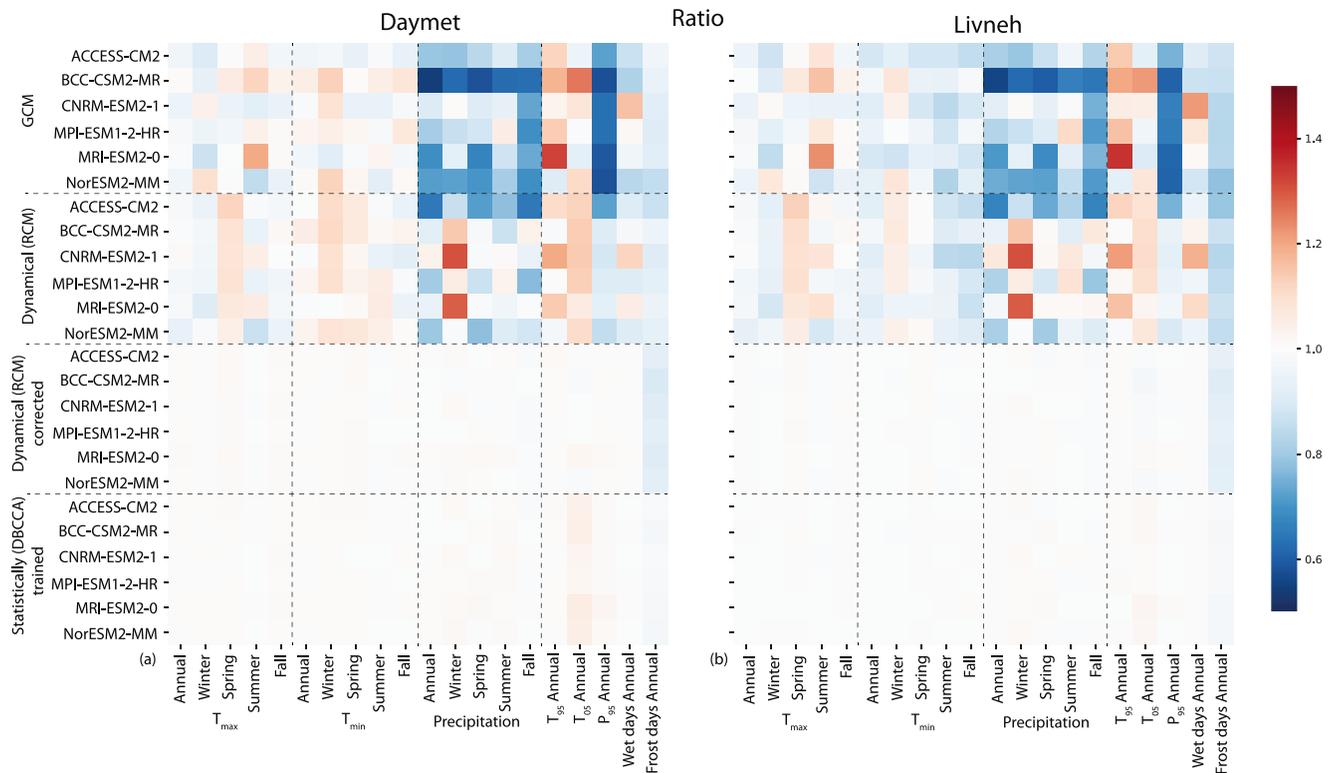


Figure 3. Same as Figure 1 but for ratio of standard deviation.

3.2. Biases

In terms of biases (Figure 1), our findings include:

1. The dynamically downscaled and corrected (RCM_{CD}, RCM_{CL}) and statistically downscaled (DBCCA_D, DBCCA_L) ensembles exhibit negligible biases at the CONUS scale in most cases. For example, the bias in P (± 0.04 mm/day), T_{\min} (-0.13 – 0.1 °C), T_{\max} (-0.15 – 0.1 °C), T_{95} (-0.25 – 0.02 °C), P_{95} (<0.4 mm/day) and R_{day} (<1 day) are small, where the values in parenthesis shows the range of CONUS-scale biases across ensemble members and the sign reflects the direction of the bias. The only exception where RCM_{CD} and RCM_{CL} exhibit higher biases than DBCCA_L and DBCCA_D is for T_{05} (up to -1.8 °C) and F_{day} (up to 8 days) (Tables S1 and S2 in Supporting Information S1).
2. In contrast, GCM_O and RCM_O ensembles, which represent uncorrected raw versions of the aforementioned datasets, exhibit relatively significant differences from their corrected counterparts across all evaluation metrics (Figure 1). For instance, most of the ensemble members display cold bias in T_{\max} over the higher elevations and parts of the upper Midwest, Southwest, South, and Southeast, which results in a CONUS-scale bias of up to -4 °C and -7.4 °C in GCM_O and RCM_O ensemble, respectively (Figure 1; Figure S2, Tables S1 and S2 in Supporting Information S1).
3. The cold biases in T_{\max} that also persist in T_{05} are partly a result of errors in the distribution of P in the GCM_O ensemble that are further exacerbated in the RCM_O ensemble after downscaling (Figure 1; Figure S3, S4 and Tables S1, S2 in Supporting Information S1).
4. The majority of GCM_O ensemble members exhibit strong wet precipitation bias except during summer and fall, when the nature of bias varies across models (Figure 1). The exacerbation of wet bias after dynamical downscaling is particularly noticeable over Northeast, parts of Southeast, and West (Figure S4 and S5 in Supporting Information S1). Interestingly, both GCM_O and RCM_O ensembles exhibit lower than observed magnitudes of P_{95} , which implies that the wet bias results from too many precipitation days (R_{day}) that lower the magnitudes of extremes by distributing annual precipitation over an excessive number of days. Nonetheless, the RCM_O ensemble reduces the underestimation bias in P_{95} over several hot spots of extreme precipitation, such as Southeast and West (Figure S6 in Supporting Information S1).
5. The RCM_O also displays improvements in the simulation of T_{\min} , T_{05} , and F_{days} (Figure 1; Figure S7–S9 in Supporting Information S1). For instance, the GCM_O ensemble exhibits substantial warm bias (up to 5 °C) in T_{\min} over parts of Southwest, South and Southeast US, Upper Midwest, and Northern Rockies (Figure S7 in Supporting Information S1). This warm bias in T_{\min} is reduced in the RCM_O ensemble, as it ranges between -1.75 to 2.6 °C (-3.7 – 3 °C) with respect to Daymet (Livneh). Most GCM_O also exhibit warm bias in simulating T_{05} , which improves after downscaling (Figure 1; Figure S8, Tables S1 and S2 in Supporting Information S1).
6. Overall, both ensembles (GCM_O and RCM_O) display higher biases in the simulated characteristics of T_{\max} and P. Nonetheless, the RCM_O ensemble shows better skill in the simulation of T_{\min} , F_{days} , T_{05} , and P_{95} as compared to the GCM_O ensemble over a majority of CONUS.

3.3. Pattern Correlation and Ratio of Standard Deviation

In terms of PC (Figure 2) and ROS (Figure 3), our findings include:

1. The dynamically downscaled and corrected (RCM_{CD}, RCM_{CL}) and statistically downscaled (DBCCA_D, DBCCA_L) ensembles have pattern correlation (PC) above 0.99 and a ratio of standard deviation (ROS) close to 1 for most of the indices (Tables S3–S6 in Supporting Information S1).
2. In the GCM_O and RCM_O ensembles, the PC for most of the metrics for T_{\max} and T_{\min} is above 0.9 except for the summer T_{\max} (0.75–0.92) and T_{95} (0.67–0.9) for both ensembles (Figure 2, Tables S3 and S4 in Supporting Information S1). In contrast, the ROS value depends on the reference observation datasets. It is close to 1 for a majority of T_{\max} and T_{\min} metrics with respect to Daymet, while it is <0.9 with respect to Livneh for several T_{\min} metrics (12/30) in the case of GCM_O ensembles and a few T_{\min} metrics (6/30) in the case of RCM_O ensembles (Tables S5 and S6 in Supporting Information S1). This discrepancy is associated with the T_{\min} biases in Livneh over the West-Southwest US (Walton & Hall, 2018).
3. For annual and seasonal P, P_{95} and R_{days} , PC ranges between 0.6 and 0.87 for both ensembles (Tables S3 and S4 in Supporting Information S1). The ROS value for P ranges between 0.6 and 1.4, with values <1 for most metrics/models in the GCM_O ensemble and close to or greater than 1 for the downscaled RCM_O ensemble.

(Figure 3). For P_{95} , the ROS value is ≥ 0.9 for 5/6 RCM_O simulations as compared to ≤ 0.75 for 6/6 GCM_O. The ROS value for wet days is between 0.9 and 1.18 for RCM_O and 0.82 and 1.22 for GCM_O ensemble members (Figure 3, Tables S5 and S6 in Supporting Information S1).

3.4. Projected Changes in Future Climate

For all six ensembles (GCM_O, RCM_O, RCM_{CL}, RCM_{CD}, DBCCA_L, DBCCA_D), we show the spatial comparison of changes in the characteristics of precipitation and temperature distributions, in addition to their summary statistics at a regional scale for nine NCEI regions (Figures 4–6; Figures S10–S13 and Tables S7–S17 in Supporting Information S1). The changes are based on the difference between the future period (2020–2059) from the reference period (1980–2019).

3.4.1. Changes in the Characteristics of T_{\max}

All six ensembles representing future climate change projections show a statistically significant increase in the magnitudes of T_{\max} and T_{95} over NCEI regions, with the strongest changes in the Upper Midwest and Northern Rockies regions (Figures 4 and 5; Figures S10, S16 and Tables S7, S8 in Supporting Information S1). Compared to the GCM_O ensemble, the corresponding statistically downscaled ensembles (DBCCA_L and DBCCA_D) project generally similar changes, while the dynamically downscaled RCM_O and its corresponding bias-corrected ensembles (RCM_{CL} and RCM_{CD}) project comparatively smaller future period changes in T_{\max} . The smallest (0.05°C) and most significant (up to 0.27°C) differences in projected changes in T_{\max} between GCM_O and RCM_O, RCM_{CL} and RCM_{CD} are over the West and the Northeast, respectively (Figure 4a). Similarly, the future period changes in T_{95} , projected by DBCCA_L and DBCCA_D, are comparable whereas those projected by RCM_O, RCM_{CL}, and RCM_{CD} are lower (by 0.27–0.7°C) than the GCM_O ensemble (Figure 4b). The six ensembles show marked differences in the future period changes in the frequency of T_{\max} extremes (number of days above the historical T_{95}). As compared to GCM_O, RCM_{CL} and RCM_{CD} project a similar increase, RCM_O projects a lower increase (by up to 8 days), whereas DBCCA_L and DBCCA_D project a much higher increase (by up to 11 days) in the frequency of T_{\max} extremes across most regions (Figures 4d and 5a–5f; Table S9 in Supporting Information S1). The differences in the frequencies of the T_{\max} extremes, projected by DBCCA_L and DBCCA_D, are higher during the last two decades of analyses (2040–2059) across CONUS (Figure S14g–S14o in Supporting Information S1).

3.4.2. Changes in the Characteristics of T_{\min}

All the six ensembles project statistically significant increase in T_{\min} and T_{05} and decrease in extremely cold days (number of days below historical period T_{05}) and number of frost days over all the regions (Figures 4e–4h; Figure S11, S12 and Tables S10–S13 in Supporting Information S1). When compared to the GCM_O, contrary to the future changes in T_{\max} , the dynamically downscaled RCM_O and its corresponding bias-corrected ensembles (RCM_{CL} and RCM_{CD}) project higher future period warming in T_{\min} over a majority of the regions with the largest differences (up to 0.23°C) in the West (Figure 4f; Table S10 in Supporting Information S1). These ensembles also projected higher increases in T_{05} than the GCM_O, with the largest differences in warming over the West and Southwest (Figure 4e, Table S11 in Supporting Information S1). On the other hand, similar to the future changes in T_{\max} , DBCCA_L and DBCCA_D exhibit future changes in T_{\min} comparable to GCM_O over most regions except West and Northwest, where differences between GCM_O and the corresponding statistically downscaled ensembles are higher (by 0.1–0.23°C) (Figure 4f, Table S10 in Supporting Information S1). The magnitude of future period changes in T_{05} is lower in DBCCA_L and DBCCA_D when compared with the GCM_O except over Northern Rockies and Upper Midwest (Figure 4e, Table S11 in Supporting Information S1). Further, there are also differences between DBCCA_L and DBCCA_D in the projected warming in T_{\min} and T_{05} over most regions. Lastly, the projected changes in the number of days below historical T_{05} are generally comparable across all the ensembles (Figure S12, Table S12 in Supporting Information S1). However, the projected decrease is higher in the number of frost days in RCM and its bias-corrected ensembles (RCM_O, RCM_{CL}, and RCM_{CD}) (by 1–5 days) than in statistically downscaled ensembles (DBCCA_L and DBCCA_D) (by up to 3 days) in comparison to the GCM_O (Figure 4h).

3.4.3. Changes in the Characteristics of P

All the ensembles project a significant increase in P, P_{95} , and extreme P (days > historical P_{95}) over the majority of CONUS, except over parts of South and Southwest in the GCM_O and the RCM_O ensembles (Figures 4i–4l and 6;

Projected Regional Future Changes

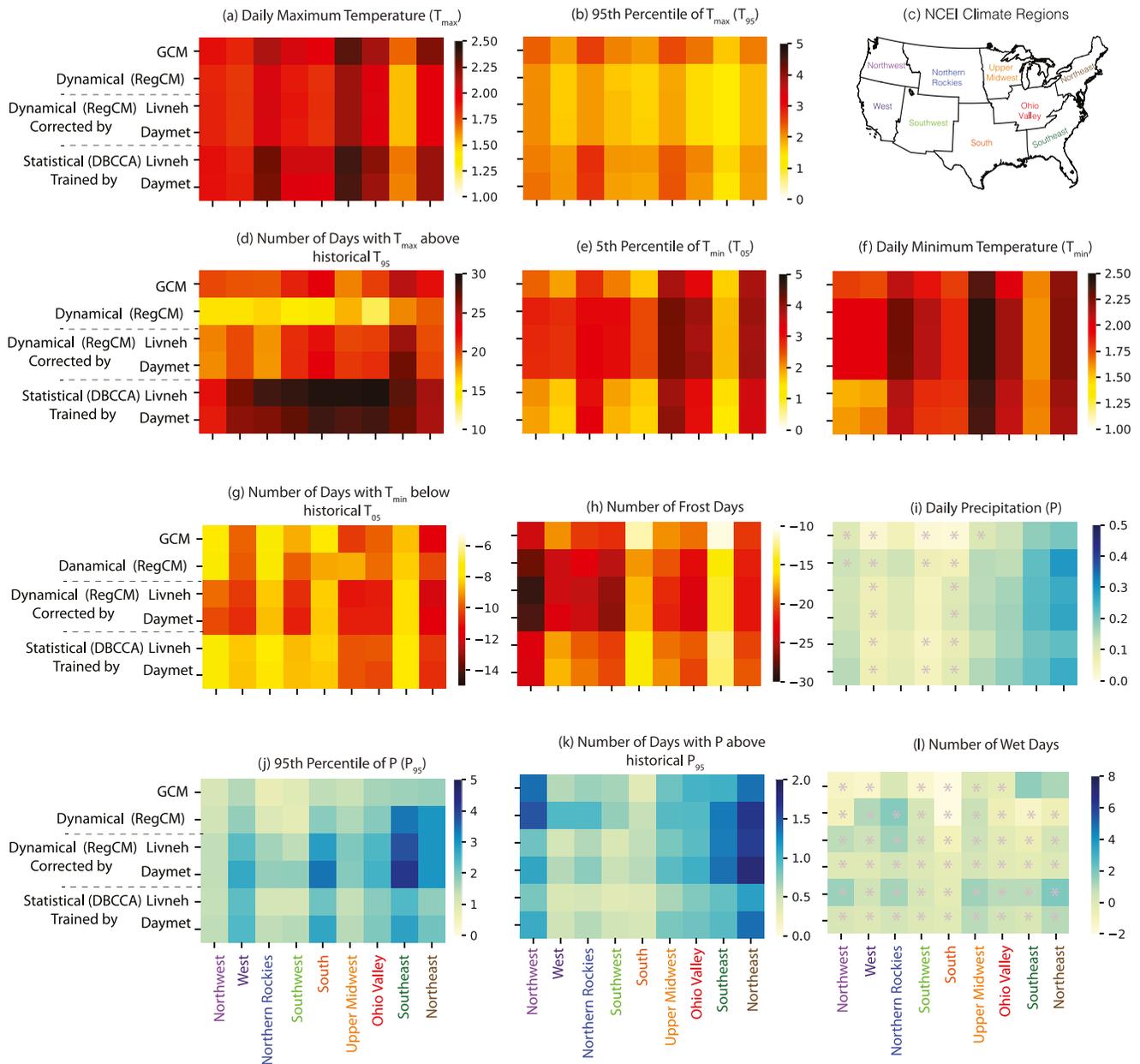


Figure 4. Heat maps showing projected future changes (2020–2059 minus 1980 to 2019) as simulated by GCM_O , RCM_O , RCM_{CL} , RCM_{CD} , $DBCCA_L$, and $DBCCA_D$ in (a) T_{max} (b) T_{95} (d) number of days with T_{max} above historical T_{95} (e) T_{05} (f) T_{min} (g) number of days with T_{min} below historical T_{05} (h) F_{days} (i) P (j) P_{95} (k) number of days with P above historical P_{95} (l) R_{days} across the nine National Center for Environmental Information (NCEI) climate regions shown in (c). The changes are significant for all the boxes at a 95% confidence interval except for those marked with *.

Figures S13, S15, Tables S14–S16 in Supporting Information S1). The future period increase is higher over the Southeast, the Northeast, and parts of the Western US. On the other hand, the projected changes in R_{days} are small and insignificant, implying that the future period increase in P results from an increase in precipitation intensity and P_{95} (Figures 4i–4l, Table S17 in Supporting Information S1). The GCM_O ensemble exhibits the smallest rise in P_{95} compared to the rest of the five ensembles (Figures 4j and 6). Furthermore, both the choice of downscaling techniques and the observations used for training and correction impact the projected future period changes in the magnitude of P_{95} . For instance, over the Southeast and the Northeast, RCM_{CL} and RCM_{CD} project a higher increase in P_{95} and the number of days above historical P_{95} compared to $DBCCA_L$ and $DBCCA_D$. Similarly,

Change in Number of Extreme Hot Days (days above T_{95}) (2020 to 2059 minus 1980 to 2019)

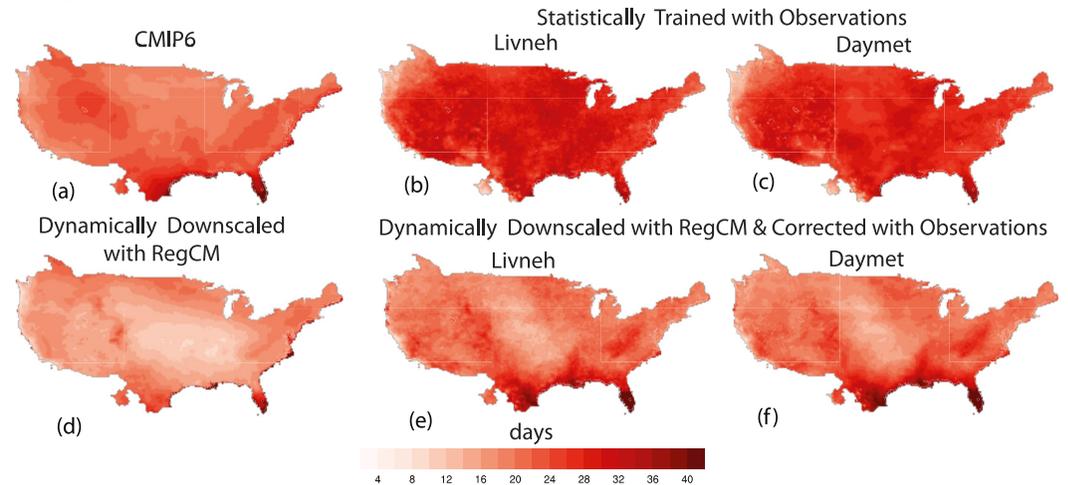


Figure 5. Spatial maps of projected changes (2020–2059 minus 1980 to 2019) in number of days above historical T_{95} in (a) GCM_O , (b) $DBCCA_L$, (c) $DBCCA_D$, (d) RCM_O (e) RCM_{CL} and (f) RCM_{CD} . Changes are significant across the map at a 95% confidence interval.

RCM_{CD} and $DBCCA_D$ project a higher increase in both metrics than RCM_{CL} and $DBCCA_L$ over the Southeast, Northeast, and West (Figure 6).

4. Summary and Discussion

We present a comprehensive comparison of an ensemble of dynamical and statistical downscaled CMIP6 GCMs in the historical and future periods to investigate how the choice of downscaling techniques and meteorological reference observations affect future hydroclimate projections. To the best of our knowledge, this is the first time that the CMIP6 GCMs have been used for downscaling over the CONUS with two very different downscaling methods. For dynamical downscaling, we employ the RegCM4 model, and the downscaled outputs are subsequently bias-corrected. For statistical downscaling, we apply the DBCCA technique to selected variables.

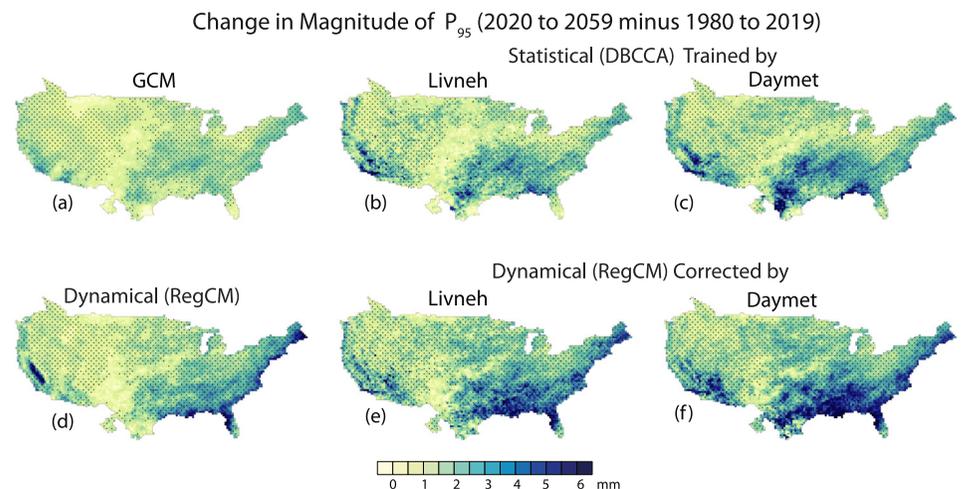


Figure 6. Spatial maps of projected changes (2020–2059 minus 1980 to 2019) in the magnitude of P_{95} in (a) GCM_O , (b) $DBCCA_L$, (c) $DBCCA_D$, (d) RCM_O (e) RCM_{CL} and (f) RCM_{CD} . Stippled area shows significant change at 95% confidence interval.

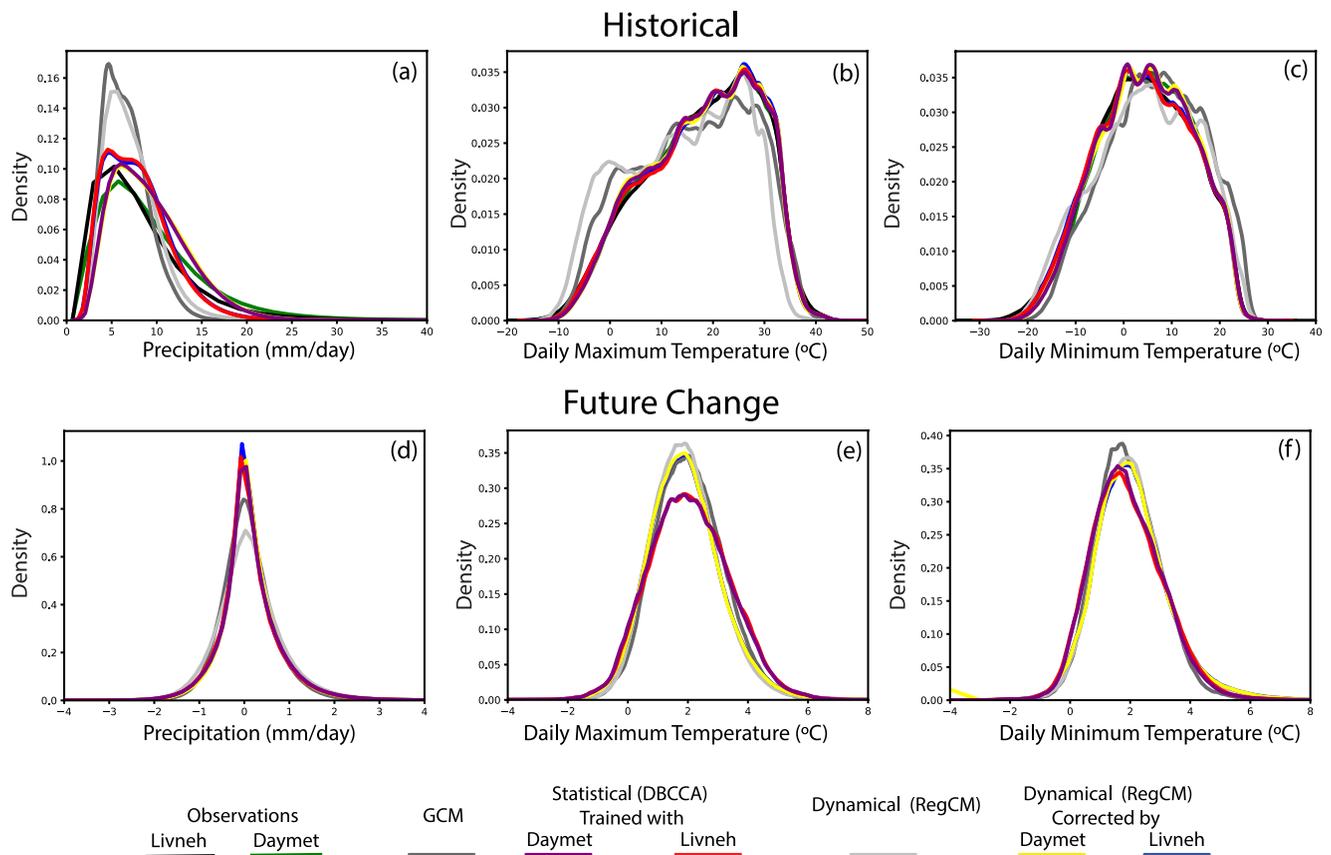


Figure 7. Probability density plots for historical period/future change in (a/d) precipitation (b/e) daily maximum temperature and (c/f) daily minimum temperature. The plots are created using monthly average values and all the grid points over the CONUS domain. The difference for each year in the future period is calculated with respect to the average historical climatology for each grid point.

Additionally, we use two high-resolution observations in the training and/or bias correction process to account for the uncertainties resulting from the choice of reference observations. Our main findings include:

4.1. Choice of Meteorological Reference Observation Used in the Bias-Correction of Downscaled Historical Climate Output Can Result in Marked Differences in the Outcome

Dynamical downscaling is unable to perform better than the driving GCMs consistently. As a result, the dynamically downscaled RCM (RCM_O) exhibits improvements over the driving GCMs (GCM_O) in several evaluation metrics but exacerbates biases in a few evaluation metrics. For instance, the RCM_O simulations exhibit cold biases than GCM_O in T_{max} . The attribution of such exacerbation is beyond the scope of current analyses because identifying physical processes that drive biases in RCM_O requires a distinctively different analytical framework beyond the scope of this study. Nonetheless, our analyses identify potential linkages between T_{max} errors are excessive P and high R_{days} in RCM_O , which provide the basis for more targeted rigorous investigations in future studies. Despite the biases in P, RCM_O simulations exhibit improvements in the distribution of precipitation extremes. The bias-corrected RCM_O simulations (RCM_{CL} and RCM_{CD}) and/or the statistical downscaling of GCM_O ($DBCCA_L$ and $DBCCA_D$) mostly remove their simulated errors in the historical climates, and these improvements are mostly insensitive to the choice of the downscaling technique.

The choice of reference observations yields some marked differences in the bias-corrected data. We evaluate these differences using the PDF plots (Figures 7a–7c; Figure S16a–S16c in Supporting Information S1), plots for cumulative density function (CDF) (Figure S16d–S16f in Supporting Information S1) and survival function (SF) (Figure S16g–S16i in Supporting Information S1) for the historical period. For instance, we find that precipitation extremes are underestimated in Livneh compared to Daymet, which is evident from the differences

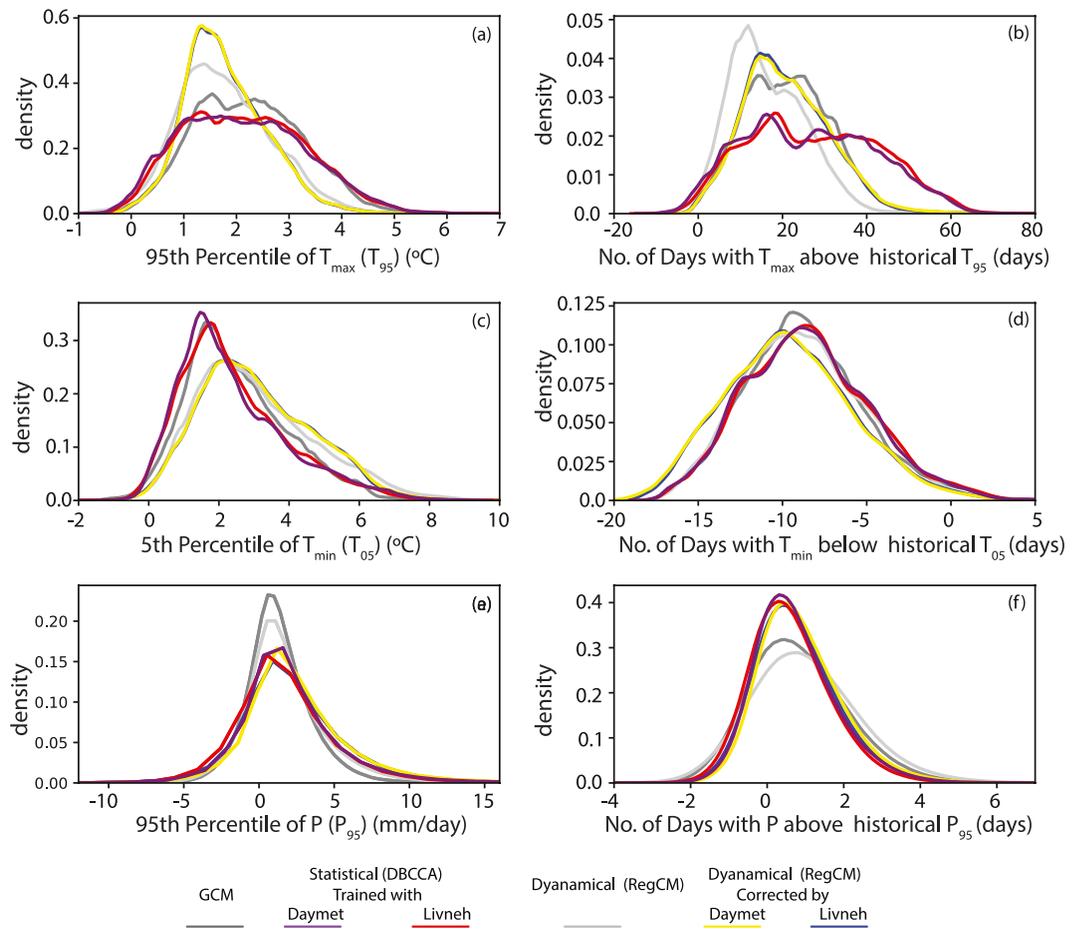


Figure 8. Probability density plots for future changes in (a) 95th percentile of T_{\max} (T_{95}) (b) the number of days above historical T_{95} (c) fifth percentile of T_{\min} (T_{05}) (d) number of days below historical T_{05} (e) 95th percentile of daily precipitation (P_{95}) (f) number of days above historical P_{95} . For each grid point of the CONUS domain, the differences are calculated for the annual average values for each year in the future period (2020–2059) with respect to the average historical 40-year climatology (1980–2019).

in the right tails of precipitation PDF plots as well as from the CDF and SF plots (Figure S16a, S16d and S16g in Supporting Information S1). Similarly, the differences in T_{\min} are evident from the disparities in left tails as well as the mode of the T_{\min} PDF plots as well as from the CDF and SF plots that show a larger number of negative T_{\min} values for Livneh than Daymet (Figure 7c, Figure S16c, S16f and S16i in Supporting Information S1). These differences in T_{\min} between observations are also visible in the spatial maps, especially over the western half of the US (Figure S7 in Supporting Information S1). Studies demonstrate that choice of interpolation technique in the development of the gridded datasets can partly be responsible for such disagreements (Walton & Hall, 2018). The downstream impact of these differences in observations is also visible between the distributions of Livneh (RCM_{CL} and DBCCA_L) versus Daymet (DBCCA_L and DBCCA_D) downscaled and/or corrected ensembles (Figures 7a–7c; Figure S16 in Supporting Information S1).

4.2. The Differences in Future Hydroclimate Projections Arising From the Choice of Downscaling Techniques Are More Prominent Than Those From the Reference Observations

This study has not addressed how historical period biases influence simulated future responses, as it requires more in-depth targeted and beyond the scope analyses. Nevertheless, we highlight the key differences in future hydroclimate projections that arise from the choice of downscaling technique and reference observation using PDF plots of changes in various metrics (Figures 7 and 8). The projected differences in the mean P , T_{\max} and T_{\min} are generally more uniform across the simulations (Figures 7d–7f). However, there are more considerable

differences in the simulation of extremes. For instance, RCM and its bias-corrected ensembles (RCM_O , RCM_{CL} , and RCM_{CD}) project comparatively lower warming in the magnitudes of T_{95} (Figure 8a). Similarly, statistically downscaled ensembles ($DBCCA_L$ and $DBCCA_D$) project significantly higher increases in the frequency of T_{95} extremes. These differences are evident from the elongated right tails and the differences in the peak of the PDF plots (Figures 8a and 8b). It is interesting to note that dynamically downscaled RCM (RCM_O) underestimates the frequency of T_{95} extremes, whereas its corresponding bias-corrected ensembles (RCM_{CL} and RCM_{CD}) project similar future changes when compared to those projected by driving GCMs (GCM_O) (Figure 8b). RCM and its bias-corrected ensembles (RCM_O , RCM_{CL} , and RCM_{CD}) overestimate, whereas statistically downscaled ensembles ($DBCCA_L$ and $DBCCA_D$) project similar change in the magnitude of T_{05} as the driving GCMs (GCM_O) (Figure 8c). RCM_{CL} and RCM_{CD} also simulate a more substantial decrease in T_{05} extremes, while RCM_O , $DBCCA_L$, and $DBCCA_D$ generally show comparable changes as GCM_O (Figure 8d). These differences are visible from the right and left shifts in the distributions of RCM, and its bias-corrected ensembles with respect to the driving GCMs and the corresponding statistically downscaled ensembles in Figures 8c and 8d, respectively. The distributions of the magnitude of change in P_{95} and the frequency of P_{95} extremes simulated by RCM_{CL} and RCM_{CD} show a slight shift to the right as compared to those by $DBCCA_L$ and $DBCCA_D$, reflecting more remarkable changes in the former (Figures 8e and 8f). These differences are also evident in the spatial plots that show a higher increase in the magnitude of P_{95} in RCM_{CL} and RCM_{CD} than in $DBCCA_L$ and $DBCCA_D$, especially over the southeastern US (Figure 6). The differences arising from the choice of observations, though less prominent, are evident in the simulations of the precipitation extremes. Daymet downscaled and/or corrected ensembles ($DBCCA_D$ and RCM_{CD}) project a higher number of P_{95} extremes as compared to Livneh-based ensembles ($DBCCA_L$ and RCM_{CL}) (Figures 6 and 8f).

The underlying causes of the disparities in future changes due to the choice of downscaling technique are difficult to systematically ascertain due to the inherently different nature of the two downscaling techniques. The projected changes in dynamical downscaling are governed by simulated physical responses, parameterizations' choice, and RCM's internal dynamics. On the other hand, the different response in the DBCCA technique could arise at different steps of the statistical downscaling procedure. Nevertheless, these analyses elucidate the key factors inducing uncertainties in the spatial downscaling, such as the downscaling techniques and the choice of observations used for training and/or correction of simulated data.

4.3. A More Comprehensive Framework Is Needed to Capture the Full Range of Methodological Uncertainties in Climate Downscaling

Different methodological choices in the downscaling techniques inevitably add a layer of uncertainty and complexity to the spatial refinement of projected future climate change, which is increasingly necessary to understand regional and local-scale climate change impacts. These include but are not limited to the choice of (a) driving GCM, (b) RCM used for dynamical downscaling, (c) statistical downscaling techniques, and (d) observations used for statistical training and/or correction. While the analyses described in this study provide helpful information regarding some of the sources of method-based uncertainties, the study is still limited in some aspects. The use of a single RCM, one statistical downscaling technique, two observational datasets, and a limited set of GCMs in the current study does not capture the full range of uncertainties associated with the methodological choices in the downscaling of GCMs. Therefore, further studies using multiple RCM simulations and statistical techniques, a larger suite of driving GCMs, and observational datasets are warranted to establish a more comprehensive and robust understanding of method-based uncertainties in this regard. In the future, we plan to extend current work by incorporating other statistical downscaling techniques and RCMs and by utilizing a more extensive set of driving GCMs. The output from the existing suite of downscaled data is currently used to run multiple hydrologic models, which should help to elucidate further the issues related to hierarchical uncertainty in downscaling procedures.

The high-resolution downscaled climate output analyzed in this study are being used in various regional scale hydro-climate risk assessment exercises including but not limited to evaluations of climate change impacts on hydrology, reservoir operations, hydropower generation, future energy demand and climate extreme events. The multi-model framework provides a new dimension for the impact assessment community to explore the additional uncertainties along the chain of hydroclimate modeling and analysis.

Data Availability Statement

All the CMIP6 GCMs data are publicly available (from <https://esgf-node.llnl.gov/projects/cmip6/>). The source code for RegCM4 is available online (from <https://github.com/ICTP/RegCM>). A non-github version cannot be provided since the code is maintained by a separate organization (International Centre for Theoretical Physics, Trieste, Italy). The analysis code used in this study is available in a repository at <https://doi.org/10.5281/zenodo.6784778>. The Daymet observations data (Thornton et al., 2021) are publicly available (from <https://daac.ornl.gov/DAYMET/>) and Livneh observations (Livneh et al., 2015) are publicly available (from <https://psl.noaa.gov/data/gridded/data.livneh.html>).

Acknowledgments

The authors thank Dr. David W. Pierce and Dr. Daniel R. Cayan from the University of California, Santa Cruz for their advice and help with Livneh data set. This study is supported by the US Department of Energy (DOE) Water Power Technologies Office as a part of the SECURE Water Act Section 9505 Assessment. Support for the climate simulations, data storage, and analysis is provided by the Oak Ridge Leadership Computing Facility, which is a DOE Office of Science User Facility. This manuscript has been authored by employees of UT-Battelle, LLC, under contract DEAC05-00OR22725 with the US Department of Energy (DOE). Accordingly, the publisher, by accepting the article for publication, acknowledges that the US government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this manuscript, or allow others to do so, for US government purposes. DOE will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan (<https://www.energy.gov/downloads/doe-public-access-plan>).

References

- Addor, N., & Seibert, J. (2014). Bias correction for hydrological impact studies—beyond the daily perspective. *Hydrological Processes*, 28(17), 4823–4828. <https://doi.org/10.1002/hyp.10238>
- Alder, J. R., & Hostetler, S. W. (2019). The dependence of hydroclimate projections in snow-dominated regions of the Western United States on the choice of statistically downscaled climate data. *Water Resources Research*, 55(3), 2279–2300. <https://doi.org/10.1029/2018WR023458>
- Ashfaq, M., Bowling, L. C., Cherkauer, K., Pal, J. S., & 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, D14116. <https://doi.org/10.1029/2009jd012965>
- Ashfaq, M., Ghosh, S., Kao, S. C., Bowling, L. C., Mote, P., Touma, D., et al. (2013). Near-term acceleration of hydroclimatic change in the western US. *Journal of Geophysical Research: Atmospheres*, 118, 10–676. <https://doi.org/10.1002/jgrd.50816>
- Ashfaq, M., Rastogi, D., Abid, M. A., & Kao, S. C. (2022). Evaluation of CMIP6 GCMs over the CONUS for downscaling studies.
- Ashfaq, M., Rastogi, D., Mei, R., Kao, S. C., Gangrade, S., Naz, B. S., & Touma, D. (2016). High-resolution ensemble projections of near-term regional climate over the continental United States. *Journal of Geophysical Research: Atmospheres*, 121, 9943–9963. <https://doi.org/10.1002/2016jd025285>
- Ayar, P. V., Vrac, M., Bastin, S., Carreau, J., Déqué, M., & Gallardo, C. (2016). Intercomparison of statistical and dynamical downscaling models under the EURO-and MED-CORDEX initiative framework: Present climate evaluations. *Climate Dynamics*, 46(3–4), 1301–1329. <https://doi.org/10.1007/s00382-015-2647-5>
- Batibenz, F., Ashfaq, M., Diffenbaugh, N. S., Key, K., Evans, K. J., Turuncoglu, U. U., & Önel, B. (2020). Doubling of US population exposure to climate extremes by 2050. *Earth's Future*, 8(4), e2019EF001421. <https://doi.org/10.1029/2019ef001421>
- Benestad, R. E. (2004). Empirical-statistical downscaling in climate modeling. *EOS, Transactions American Geophysical Union*, 85(42), 417–422. <https://doi.org/10.1029/2004eo420002>
- Bentsen, M., Olivière, D. J. L., Seland, Ø., Toniazzo, T., Gjermundsen, A., Graff, L. S., et al. (2019). *NCC NorESM2-MM model output prepared for CMIP6 CMIP historical*. Earth System Grid Federation. <https://doi.org/10.22033/ESGF/CMIP6.8040>
- Bettolli, M. L., Solman, S. A., Da Rocha, R. P., Llopart, M., Gutierrez, J. M., Fernández, J., et al. (2021). The CORDEX flagship pilot study in southeastern South America: A comparative study of statistical and dynamical downscaling models in simulating daily extreme precipitation events. *Climate Dynamics*, 56(5–6), 1–20. <https://doi.org/10.1007/s00382-020-05549-z>
- Diffenbaugh, N. S., & Ashfaq, M. (2010). Intensification of hot extremes in the United States. *Geophysical Research Letters*, 37(15), L15701. <https://doi.org/10.1029/2010gl043888>
- Dix, M., Bi, D., Dobrohotoff, P., Fiedler, R., Harman, I., Law, R., et al. (2019). *CSIRO-ARCCSS ACCESS-CM2 model output prepared for CMIP6 CMIP historical*. Earth System Grid Federation. <https://doi.org/10.22033/ESGF/CMIP6.4271>
- Giorgi, F., Coppola, E., Solmon, F., Mariotti, L., Sylla, M. B., Bi, X., et al. (2012). RegCM4: Model description and preliminary tests over multiple CORDEX domains. *Climate Research*, 52, 7–29. <https://doi.org/10.3354/cr01018>
- Grell, G. A., Dudhia, A. J., & Stauffer, D. R. (1994). *A description of the fifth-generation Penn State/NCAR mesoscale model (MM5)*. NCAR Tech. Note NCAR/TN-398+STR. Natl. Cent. for Atmos. Res.
- Gutmann, E., Pruitt, T., Clark, M. P., Brekke, L., Arnold, J. R., Raff, D. A., & Rasmussen, R. M. (2014). An intercomparison of statistical downscaling methods used for water resource assessments in the United States. *Water Resources Research*, 50(9), 7167–7186. <https://doi.org/10.1002/2014wr015559>
- Henn, B., Newman, A. J., Livneh, B., Daly, C., & Lundquist, J. D. (2018). An assessment of differences in gridded precipitation datasets in complex terrain. *Journal of Hydrology*, 556, 1205–1219. <https://doi.org/10.1016/j.jhydrol.2017.03.008>
- Hnilica, J., Hanel, M., & Puš, V. (2017). Multisite bias correction of precipitation data from regional climate models. *International Journal of Climatology*, 37(6), 2934–2946. <https://doi.org/10.1002/joc.4890>
- Holtzlag, A. A. M., Debruin, E. I. F., & Pan, H. L. (1990). A high-resolution air-mass transformation model for short-range weather forecasting. *Monthly Weather Review*, 118(8), 1561–1575. [https://doi.org/10.1175/1520-0493\(1990\)118<1561:Ahramt>2.0.Co;2](https://doi.org/10.1175/1520-0493(1990)118<1561:Ahramt>2.0.Co;2)
- Jiang, Y., Kim, J. B., Still, C. J., Kerns, B. K., Kline, J. D., & Cunningham, P. G. (2018). Inter-comparison of multiple statistically downscaled climate datasets for the Pacific Northwest, USA. *Scientific Data*, 5(1), 1–18. <https://doi.org/10.1038/sdata.2018.50>
- Kiehl, J. T., Hack, J. J., Bonan, G. B., Boville, B. A., Williamson, D. L., & Rasch, P. J. (1998). The national center for atmospheric research community climate model: CCM3. *Journal of Climate*, 11(6), 1131–1149. [https://doi.org/10.1175/1520-0442\(1998\)011<1131:Tncfar>2.0.Co;2](https://doi.org/10.1175/1520-0442(1998)011<1131:Tncfar>2.0.Co;2)
- Le Roux, R., Katurji, M., Zawar-Reza, P., Quénol, H., & Sturman, A. (2018). Comparison of statistical and dynamical downscaling results from the WRF model. *Environmental Modelling & Software*, 100, 67–73. <https://doi.org/10.1016/j.envsoft.2017.11.002>
- Li, D., Zou, L., & Zhou, T. (2018). Extreme climate event changes in China in the 1.5 and 2 C warmer climates: Results from statistical and dynamical downscaling. *Journal of Geophysical Research: Atmospheres*, 123(18), 10–215. <https://doi.org/10.1029/2018jd028835>
- Livneh, B., Bohn, T. J., Pierce, D. W., Munoz-Arriola, F., Nijssen, A., Vose, R., et al. (2015). A spatially comprehensive, hydrometeorological data set for Mexico, the US, and Southern Canada 1950–2013. *Scientific Data*, 2(1), 1–12. <https://doi.org/10.1038/sdata.2015.42>
- Manzanas, R., Gutiérrez, J. M., Fernández, J., Van Meijgaard, E., Calmanti, S., Magariño, M. E., et al. (2018). Dynamical and statistical downscaling of seasonal temperature forecasts in Europe: Added value for user applications. *Climate Services*, 9, 44–56. <https://doi.org/10.1016/j.cliser.2017.06.004>
- Maurer, E. P., Das, T., & Cayan, D. R. (2013). Errors in climate model daily precipitation and temperature output: Time invariance and implications for bias correction. *Hydrology and Earth System Sciences*, 17(6), 2147–2159. <https://doi.org/10.5194/hess-17-2147-2013>

- Maurer, E. P., & Hidalgo, H. G. (2008). Utility of daily vs. monthly large-scale climate data: An intercomparison of two statistical downscaling methods. *Hydrology and Earth System Sciences*, *12*(2), 551–563. <https://doi.org/10.5194/hess-12-551-2008>
- Maurer, E. P., Hidalgo, H. G., Das, T., Dettinger, M. D., & Cayan, D. R. (2010). The utility of daily large-scale climate data in the assessment of climate change impacts on daily streamflow in California. *Hydrology and Earth System Sciences*, *14*, 1125–1138. <https://doi.org/10.5194/hess-14-1125-2010>
- Mearns, L. O., McGinnis, S., Korytina, D., Arritt, R., Biner, S., Bukovsky, M., et al. (2017). *The NA-CORDEX dataset, version 1.0*. NCAR Climate Data Gateway. <https://doi.org/10.5065/D6SJ1JCH>
- Mezghani, A., Dobler, A., Benestad, R., Haugen, J. E., Parding, K. M., Piniewski, M., & Kundzewicz, Z. W. (2019). Subsampling impact on the climate change signal over Poland based on simulations from statistical and dynamical downscaling. *Journal of Applied Meteorology and Climatology*, *58*(5), 1061–1078. <https://doi.org/10.1175/jamc-d-18-0179.1>
- Naz, B. S., Kao, S. C., Ashfaq, M., Rastogi, D., Mei, R., & Bowling, L. C. (2016). Regional hydrologic response to climate change in the conterminous United States using high-resolution hydroclimate simulations. *Global and Planetary Change*, *143*, 100–117. <https://doi.org/10.1016/j.gloplacha.2016.06.003>
- Nikulin, G., Asharaf, S., Magariño, M. E., Calmanti, S., Cardoso, R. M., Bhend, J., et al. (2018). Dynamical and statistical downscaling of a global seasonal hindcast in eastern Africa. *Climate Services*, *9*, 72–85. <https://doi.org/10.1016/j.cliser.2017.11.003>
- Pagán, B. R., Ashfaq, M., Rastogi, D., Kendall, D. R., Kao, S. C., Naz, B. S., et al. (2016). Extreme hydrological changes in the southwestern US drive reductions in water supply to Southern California by mid century. *Environmental Research Letters*, *11*(9), 094026. <https://doi.org/10.1088/1748-9326/11/9/094026>
- Pal, J. S., Small, E. E., & Eltahir, E. A. B. (2000). Simulation of regional-scale water and energy budgets: Representation of subgrid cloud and precipitation processes within RegCM. *Journal of Geophysical Research*, *105*(D24), 29579–29594. <https://doi.org/10.1029/2000JD900415>
- Pierce, D. W., Cayan, D. R., Maurer, E. P., Abatzoglou, J. T., & Hegewisch, K. C. (2015). Improved bias correction techniques for hydrological simulations of climate change. *Journal of Hydrometeorology*, *16*(6), 2421–2442. <https://doi.org/10.1175/jhm-d-14-0236.1>
- Pierce, D. W., Su, L., Cayan, D. R., Risser, M. D., Livneh, B., & Lettenmaier, D. P. (2021). An extreme-preserving long-term gridded daily precipitation dataset for the conterminous United States. *Journal of Hydrometeorology*, *22*(7), 1883–1895. <https://doi.org/10.1175/jhm-d-20-0212.1>
- Potter, N. J., Chiew, F. H., Charles, S. P., Fu, G., Zheng, H., & Zhang, L. (2020). Bias in dynamically downscaled rainfall characteristics for hydroclimatic projections. *Hydrology and Earth System Sciences*, *24*(6), 2963–2979. <https://doi.org/10.5194/hess-24-2963-2020>
- Rastogi, D., Holladay, J. S., Evans, K. J., Preston, B. L., & Ashfaq, M. (2019). Shift in seasonal climate patterns likely to impact residential energy consumption in the United States. *Environmental Research Letters*, *14*(7), 074006. <https://doi.org/10.1088/1748-9326/ab22d2>
- Rastogi, D., Touma, D., Evans, K. J., & Ashfaq, M. (2020). Shift toward intense and widespread precipitation events over the United States by mid-21st century. *Geophysical Research Letters*, *47*(19), e2020GL089899. <https://doi.org/10.1029/2020gl089899>
- Seferian, R. (2018). *CNRM-CERFACS CNRM-ESM2-1 model output prepared for CMIP6 CMIP*. Earth System Grid Federation. <https://doi.org/10.22033/ESGF/CMIP6.1391>
- Tawfik, A. B., & Steiner, A. L. (2011). The role of soil ice in land-atmosphere coupling over the United States: A soil moisture-precipitation winter feedback mechanism. *Journal of Geophysical Research*, *116*, D02113. <https://doi.org/10.1029/2010JD014333>
- Thornton, M. M., Shrestha, R., Thornton, P. E., Kao, S., Wei, Y., & Wilson, B. E. (2021). *Daymet version 4 monthly latency: Daily surface weather data*. ORNL DAAC. <https://doi.org/10.3334/ORNLDAAC/1904>
- Thrasher, B., Maurer, E. P., McKellar, C., & Duffy, P. B. (2012). Bias correcting climate model simulated daily temperature extremes with quantile mapping. *Hydrology and Earth System Sciences*, *16*(9), 3309–3314. <https://doi.org/10.5194/hess-16-3309-2012>
- Tiedtke, M. (1989). A comprehensive mass flux scheme for cumulus parameterization in large-scale models. *Monthly Weather Review*, *117*(8), 1779–1800. [https://doi.org/10.1175/1520-0493\(1989\)117<1779:acmsf>2.0.co;2](https://doi.org/10.1175/1520-0493(1989)117<1779:acmsf>2.0.co;2)
- Von Storch, Jin Song, Putrasahan, D., Lohmann, K., Gutjahr, O., Jungclaus, J., et al. (2017). *MPI-M MPIESM1.2-HR model output prepared for CMIP6 HighResMIP*. Earth System Grid Federation. <https://doi.org/10.22033/ESGF/CMIP6.762>
- Walton, D., & Hall, A. (2018). An assessment of high-resolution gridded temperature datasets over California. *Journal of Climate*, *31*(10), 3789–3810. <https://doi.org/10.1175/jcli-d-17-0410.1>
- Werner, A. T., & Cannon, A. J. (2016). Hydrologic extremes—an intercomparison of multiple gridded statistical downscaling methods. *Hydrology and Earth System Sciences*, *20*(4), 1483–1508. <https://doi.org/10.5194/hess-20-1483-2016>
- Wood, A. W., Leung, L. R., Sridhar, V., & Lettenmaier, D. P. (2004). Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs. *Climatic Change*, *62*(1–3), 189–216. <https://doi.org/10.1023/b:clim.0000013685.99609.9e>
- Wu, T., Chu, M., Dong, M., Fang, Y., Jie, W., Li, J., et al. (2018). *BCC BCC-CSM2MR model output prepared for CMIP6 CMIP piControl*. Earth System Grid Federation. <https://doi.org/10.22033/ESGF/CMIP6.3016>
- Yoon, J. H., Ruby Leung, L., & Correia, J., Jr. (2012). Comparison of dynamically and statistically downscaled seasonal climate forecasts for the cold season over the United States. *Journal of Geophysical Research*, *117*, D21109. <https://doi.org/10.1029/2012jd017650>
- Yukimoto, S., Koshiro, T., Kawai, H., Oshima, N., Yoshida, K., Urakawa, S., et al. (2019). *MRI MRI-ESM2.0 model output prepared for CMIP6 CMIP*. Earth System Grid Federation. <https://doi.org/10.22033/ESGF/CMIP6.621>