Title
Understanding the Cascade: Removing GCM Biases Improves Dynamically Downscaled Climate Projections

Journal
Geophysical Research Letters, 51(9)

ISSN
0094-8276

Authors
Rahimi, Stefan
Huang, Lei
Norris, Jesse
et al.

Publication Date
2024-05-16

DOI
10.1029/2023gl106264

AGU ADVANCING EARTH AND SPACE SCIENCES

Geophysical Research Letters

RESEARCH LETTER
10.1029/2023GL106264

Understanding the Cascade: Removing GCM Biases Improves Dynamically Downscaled Climate Projections

Stefan Rahimi[1,2], Lei Huang[1], Jesse Norris[1], Alex Hall[1], Naomi Goldenson[1], Mark Risser[3], Daniel R. Feldman[3], Zachary J. Lebo[4], Eli Dennis[1], and Chad Thackeray[1]

[1]Department of Atmospheric and Oceanic Sciences, Center for Climate Science, University of California Los Angeles, Los Angeles, CA, USA, [2]Department of Atmospheric Science, University of Wyoming, Laramie, WY, USA, [3]Earth and Environmental Sciences Area, Lawrence Berkeley National Laboratory, Berkeley, CA, USA, [4]School of Meteorology, University of Oklahoma, Norman, OK, USA

Key Points:

- Bias correction of Global Climate Models (GCMs) reduces biases in downscaled mean precipitation, snow, and temperature across the western United States
- Cascading cold, thermodynamically unstable, and cyclonic vorticity biases from GCMs to regional climate models drive wet biases in dynamical downscaling
- CMIP6-wide GCM biases are similar suggesting that biases in dynamically downscaled precipitation and temperature can be anticipated

Supporting Information:

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

Correspondence to:

S. Rahimi,
srahimi@uwyo.edu

Citation:

Rahimi, S., Huang, L., Norris, J., Hall, A., Goldenson, N., Risser, M., et al. (2024). Understanding the cascade: Removing GCM biases improves dynamically downscaled climate projections. *Geophysical Research Letters*, *51*, e2023GL106264. https://doi.org/10.1029/2023GL106264

Received 7 SEP 2023
Accepted 30 MAR 2024

Author Contributions:

Conceptualization: Stefan Rahimi, Naomi Goldenson, Chad Thackeray
Data curation: Stefan Rahimi, Lei Huang
Formal analysis: Stefan Rahimi
Funding acquisition: Alex Hall
Investigation: Stefan Rahimi, Jesse Norris, Mark Risser
Methodology: Stefan Rahimi, Jesse Norris, Chad Thackeray
Project administration: Stefan Rahimi
Resources: Stefan Rahimi, Zachary J. Lebo
Software: Stefan Rahimi
Supervision: Alex Hall

Abstract

Polarization surrounding bias correction (BC) in creating climate projections arises from its lack of physicality. Here, we perform and analyze 18 dynamical downscaling simulations (with and without BC) to better understand the physical impacts of BC, applied before downscaling, on regional climate output across the western United States. Without BC, downscaled precipitation is systematically and unrealistically wet biased compared to a hierarchy of observationally based datasets over the 1980–2014 period due to cascading mean-state Global Climate Model (GCM) biases: (a) overly strong lower-tropospheric lapse rates (5 K/km), (b) overly cold (2 K) tropospheric temperatures, and (c) anomalous mid-tropospheric cyclonic vorticity advection. With BC, downscaled precipitation (snow) biases are virtually eliminated (halved). Identified GCM biases are common to the broader Coupled Model Intercomparison Project ensemble. Physical effects of BC on the quality of the regionalized projections, pending an evaluation of BC's distortion of the downscaled climate response, may motivate its broader application by dynamical downscalers.

Plain Language Summary

Global Climate Models (GCMs) are known to have biases that, when dynamically downscaled, damage the credibility of the e. A longstanding way around this problem is bias correction (BC) after downscaling, but this practice rarely involves physics and can mislead climate data users into overvaluing the quality of the downscaled data. Further, post-downscaling BC techniques can over correct the higher-order statistics, calling into question the faithful preservation of the original simulated signals. For the first time, we apply a minimally invasive BC procedure to a group of 9 GCMs in order to define physical relationships between mean GCM biases and their dynamically downscaled hydroclimate variables across the western United States. We find that native GCMs tend to exhibit surprisingly common mean biases that, when downscaled, effectuate an overly wet, cold, and snowy climate across the region.

1. Introduction

From water supply to electricity load forecasts, climate data users across the western US often rely on direct dynamically downscaled global climate model (GCM) projections of the 21st Century. This approach involves running a regional climate model (RCM) with boundary conditions provided from GCMs across a focused area of the planet (Bukovsky & Karoly, 2011; Giorgi et al., 1994; Huang et al., 2020, 2021; Rastogi et al., 2022; Wang & Kotamarthi, 2013, 2015; Xu et al., 2021). This domain restriction allows the RCM to operate with a more realistic representation of the physical processes within a GCM grid box that impact variables such as near-surface air temperature and precipitation. The resulting regional climate change projections inform new infrastructure design standards, allow for assessing changes in water vulnerability and drought, and provide crucial insight into human health risks on decision-relevant scales.

Although RCMs provide a more realistic representation of the local climate, they, like all downscaling methods, inherit the biases of the driving GCM when used for climate change projections (Maraun et al., 2017; Rastogi et al., 2022; Xu et al., 2019; Zhang et al., 2022). GCM bias propagation can lead to large biases and other artifacts that can easily mute the benefits of dynamical downscaling, namely increased resolution, thereby raising legitimate questions about the fidelity and trustworthiness of their future climate projections. To cope with this issue, modelers and the community of climate data users resort to bias correction (BC; properly referred to as bias adjustment per the most recent IPCC report).

Geophysical Research Letters
10.1029/2023GL106264

Validation: Stefan Rahimi, Lei Huang
Visualization: Stefan Rahimi
Writing – original draft: Stefan Rahimi, Jesse Norris, Chad Thackeray
Writing – review & editing: Stefan Rahimi, Jesse Norris, Chad Thackeray

In the context of dynamical downscaling, implementation of BC occurs either before or after running the RCM. For a priori (before downscaling) BC, biases in the GCM's mean-state and/or variability are quantified against reanalyzes and removed from the GCM output (Bruyère et al., 2014; Holland et al., 2010; Xu et al., 2021). Variable biases, specifically for those solved for in the GCM's dynamic core, are computed over a historical reference period to quantify bias shifts; these shifts are then applied to the entire time series to reduce downscaled biases in high-resolution RCM-simulated fields. The application of bias shifts, assumed to be time-invariant, breaks the physical consistency between variables in the GCM boundary conditions, but the RCM establishes its own internal relationships between variables in the downscaled outputs. More commonly applied however is a posteriori (after downscaling) BC, implemented on the variables of interest to climate data users (e.g., surface air temperature, precipitation, and surface winds). Often performed by quantile, this method may present challenges for researchers using the data for process studies, namely artifacts in downscaled signals which lack physical interpretability; these issues persist even in multivariate a posteriori BC routines (Cannon, 2018; Pierce et al., 2014). Further, a posteriori BC often assumes that GCM trends are preserved which is almost certainly not the case on decision-relevant scales. Like a priori BC, a posteriori BC assumes time-invariance in its application.

Despite downscaling type, BC routines cannot by themselves identify "absurdity" within the data being bias corrected, and it can easily distort future climate trends. Maraun et al. (2017) highlighted this distortion potential by bias correcting the historical jet stream, which unphysically transformed a GCM-predicted future-era poleward shift in storm tracks into an equatorward shift. The authors also highlighted the ease to which BC can be misused by bias correcting Southern Ocean surface air temperature to European precipitation. They effectively motivated the responsible usage of BC, an idea reinforced in the Intergovernmental Panel on Climate Change's Working Group 1 report (IPCC, 2021) and Goldenson et al. (2023).

Given the risk a posteriori BC routines may pose in inflating the fidelity of downscaled climate data, in addition to its ability to distort climate trends, and because dynamically downscaling GCMs without any type of BC may lead to extremely large biases at local scales, there is a need to test the effects of a priori BC on an ensemble of GCMs in order to provide the community with updated guidance on its consequences and applicability.

Below, we quantify the impacts of applying a simple a priori BC of the mean-state climate to a small ensemble of 9 carefully selected GCMs (See Sec. S1) from the 6th Coupled Model Intercomparison Project (CMIP6) across the western US during the recent historical period (1980–2014). The complex terrain and many micro-climates of the western US make GCMs particularly ill-equipped to represent its heterogeneous regional climate. Thus, it is a region where BC may be particularly impactful. We focus on instances where preventing the transmission of mean-state biases from the GCM to the RCM (termed here bias cascade) leads to physically explainable improvements in downscaled precipitation, snow, and temperature. The results here reinforce those in Maraun et al. (2017) and Goldenson et al. (2023) that BC should only be applied to GCMs which demonstrate relatively high skill in simulating the regional atmospheric circulation. However, specific to dynamical downscaling, we argue that BC should (a) yield physically interpretable impacts, and (b) be simple in that it prioritizes the correction of the lower order statistics. By adopting these BC guidelines, modelers will be left with less-absurd outputs for process studies, while climate data users needing additional layers of BC a posteriori for impact assessments may expect less distortion of the original RCM output.

2. Methods

Dynamical downscaling of GCMs to a 9-km grid is accomplished using the Weather Research and Forecasting (WRF; Skamarock et al., 2019) model from 1 September 1980 through 31 August 2014. Rahimi et al. (2023), which overviews the Western United States Dynamically Downscaled Dataset (WUS_D3), describes these GCMs in more detail. Spectral nudging is implemented on an intermediate 45-km grid (Figure S1 in Supporting Information S1) to prevent model drift for temperature and horizontal winds above the planetary boundary layer. These options, described in more detail in Text S2 of Supporting Information S1, were applied by Rahimi et al. (2022) when dynamically downscaling a 40-year reanalysis-driven simulation, the ECMWF's Reanalysis Version 5 (ERA5; Hersbach et al., 2020), which is used below for comparison purposes (WRF-ERA5). Internal to the WRF domain, transient local-scale aerosol, land-use,/land coverage, and greenhouse gas changes are not explicitly simulated, however the GCM boundary condition data implicitly contain these effects from far-field sources.

To explore the effects of BC in an ensemble sense, nine downscaled GCMs are selected from (Rahimi et al., 2023; Table S1 in Supporting Information S1). These GCMs were initially chosen for downscaling based on an extensive GCM evaluation process (Krantz et al., 2021; Simpson et al., 2020; and Text S1 in Supporting Information S1). For this study, nine identical WRF simulations are conducted, except that a gridpoint-specific ERA5-relative BC of each GCM's temperature, moisture, horizontal winds, surface pressure, and SSTs is performed.

BC is the subject of much debate in the regional downscaling community given that, if not applied carefully, a highly biased dataset may appear to be of high fidelity (e.g., Ehret et al., 2012; Maraun et al., 2017). Thus, we apply a minimally invasive monthly-varying BC of the mean-state GCM fields that preserves the regional synoptic dynamics following the methods of Bruyère et al. (2014). This method decomposes a GCM-simulated variable (x_{GCM}) into its mean-state historical climatological mean ($x_{\mathrm{GCM},0}$) plus the anomaly (x'_{GCM}):

$$x_{\mathrm{GCM}} = x_{\mathrm{GCM},0} + x'_{\mathrm{GCM}}. \tag{1}$$

Next, we do the same for a reference dataset, in this case, ERA5:

$$x_{\mathrm{ERA5}} = x_{\mathrm{ERA5},0} + x'_{\mathrm{ERA5}}. \tag{2}$$

We then define the mean-state bias as $\Delta = x_{\mathrm{GCM},0} - x_{\mathrm{ERA5},0}$, and subtract this from the full GCM signal to arrive at our bias corrected signal (x_{BC}) and use this signal to drive WRF:

$$x_{\mathrm{BC}} = x_{\mathrm{GCM},0} + x'_{\mathrm{GCM}} - \Delta, \tag{3}$$

$$x_{\mathrm{BC}} = x_{\mathrm{ERA5},0} + x'_{\mathrm{GCM}}. \tag{4}$$

BC shifts are computed monthly by differencing the 1980–2014 climatologies of the respective GCMs and ERA5 before upsampled to 6-hourly intervals via linear interpolation to match the boundary condition update frequency within WRF. While this BC technique preserves the long-term trends and variability of the original GCM data, it breaks the physical consistency between a GCM's dynamic fields, specifically winds and temperature. This is a prime reason for us bias correcting only the means versus a more aggressive multivariate or quantile-mapping-based approach. We argue here that the latter method may be more likely to lead to dynamical inconsistencies, violations of fundamental conservation laws, and manifest as unrealistic solutions in downscaling.

3. Results

3.1. BC Impacts on Mean Climate Across the Western US

We begin by examining the climatological (1981–2010) biases in ensemble mean temperature (Figures 1a and 1b) and precipitation (Figures 1c and 1d) relative to the Parameter-elevation Regressions on Independent Slopes Model (PRISM; Daly et al., 1994, 2008; PRISM Climate Group, 2014). When no BC is employed (henceforth labeled no-BC), simulations exhibit cold and wet biases across the western US (Figures 1a–1c), with biases particularly pronounced at high elevations (around −4 K and +4 mm d^{-1}). Even larger biases (up to −10 K and +10 mm d^{-1}) are found for individual experiments (Figures S2 and S3 in Supporting Information S1). As expected, downscaling with a priori BC (henceforth labeled w/BC) dramatically reduces these biases (Figures 1b–1d). This makes the biases of similar magnitude to WRF-ERA5.

Since precipitation and temperature biases are amplified at higher elevations, we examine how BC affects simulated snow water equivalent (SWE) and precipitation in the mountains. These biases are calculated on more than 700 high-elevation in situ SNOw TELemetry sites (Serreze et al., 1999; Figure S1 in Supporting Information S1, right). Eliminating bias cascade generally dries the downscaled precipitation and snow solutions in the ensemble mean, making the precipitation result almost precisely equal to WRF-ERA5 and SNOTEL (Figure 1e). SWE on the other hand is dry-biased (by 12%) when BC is implemented. We attribute this dry SWE bias to overly smoothed WRF topography due to the 9-km grid spacing, particularly across the southwestern US SNOTEL sites, which can affect simulated surface temperatures and subsequent rain-snow partitioning in the land-surface model (Jordan, 1991), as well as mute orographically-influenced precipitation. Meanwhile, no-BC experiments

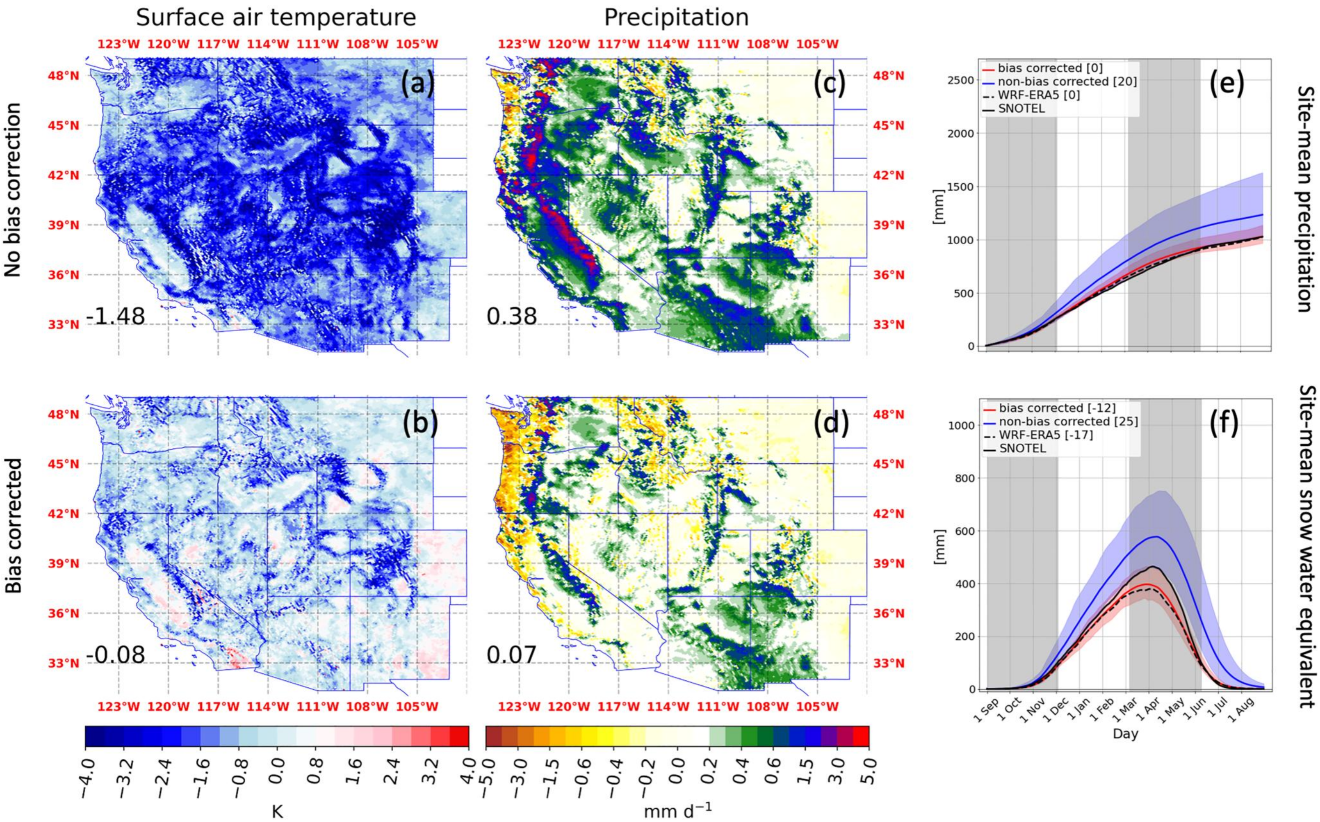

Figure 1. Relative to Parameter-elevation Regressions on Independent Slopes Model, ensemble-mean climatological (1981–2010) bias patterns in annual-mean (a and b) temperature and (c and d) precipitation, with 11-state-mean biases indicated in the lower-left of each subpanel. (e and f) Time series of cumulative annual precipitation and snow water equivalent, respectively [mm], across 703 western US SNOTEL sites. Shading denotes the range of bias corrected (red) and non-bias corrected (blue) experiments. Observations (black solid curve) and WRF-ERA5 (black dashed curve) are also presented.

overestimate precipitation and peak SWE by about 20% and 25%, respectively. Additionally, the range of no-BC precipitation and snow (blue shading) is much wider and wetter across the region than in the w/BC experiments (red shading), as downscaled precipitation (peak snow) in the former can be as much as 65% (80%) wetter than observations. Sub-regionally, no-BC simulations can be even more severely biased.

Sub-regionally, BC brings the largest improvements across the Pacific Northwest (PNW; Figure S4 in Supporting Information S1). Here, some no-BC experiments simulate precipitation (peak snow) that is 65% (100%) wetter than observations. A 31% ensemble-mean precipitation wet bias is reduced to only 8% when BC is employed. With BC across other subregions, regional peak SWE biases can be larger in magnitude than those without. In the Southwest (SW), the no-BC experiments have a 19% wet bias in the GCM mean, while the w/BC experiments have up to a 50% dry bias. Rahimi et al. (2022) found that decreasing WRF's grid spacing to 3-km greatly reduced this dry bias, hypothesizing that better-resolved topography leads to more accurate precipitation and snow over the Sierra Nevada. Other studies have noted the difficulties in simulating precipitation and SWE across this region (He et al., 2019). This result illustrates that the dry bias associated with the representation of orographic precipitation is largely canceled in the no-BC experiments by a tendency toward wet bias originating from cascading mean-state biases from GCMs.

3.2. Precipitation Extremes Across California

It is unsurprising that BC improves the downscaled means of temperature, precipitation, and SWE, while also leading to a reduced ensemble spread. However, it is less obvious that precipitation extremes should be improved, given that BC is only applied to the climatological monthly means of GCM variables. Figure 2 shows statistics across the downscaled GCM ensemble for annual maximum cumulative precipitation totals over five intervals (ranging from 1 to 30 days). These are calculated at each gridcell and averaged across California. BC yields a

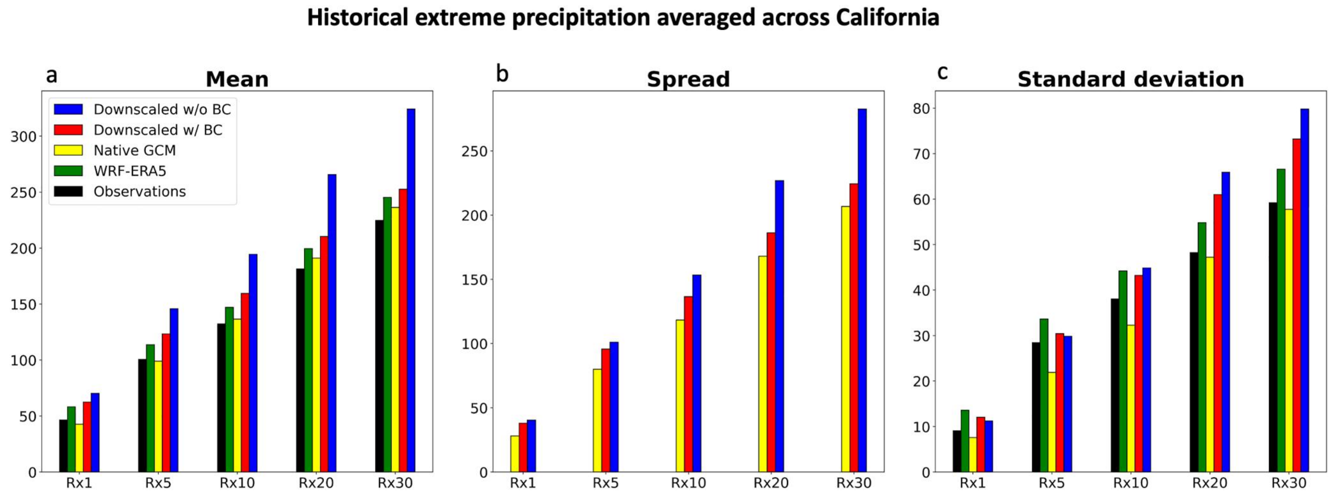

Figure 2. (a) Ensemble mean, (b) intermodel spread, and (c) the standard deviation of climatology-averaged maximum 30-day (Rx30), 20-day, 10-day, 5-day, and 1-day cumulative precipitation [mm] across California. Ensemble means and spreads are computed by considering the nine different Global Climate Models (GCMs) and their respective downscaled simulations, while the standard deviation is computed by taking the root of the mean variance computed for each GCM's time series.

better (drier) representation of extreme precipitation than no-BC simulations when compared to PRISM (Figure 2a). Ensemble-mean wet biases in extreme precipitation (~41%–49% for the various metrics) are reduced to between 16% and 31% when BC is implemented. Wet bias reductions are larger for longer intervals (e.g., rx30 day). Clearly, bias cascade is leading to a more biased representation of extreme precipitation averaged across California relative to native GCMs, but BC ameliorates this issue.

Regardless of bias cascade, dynamical downscaling as implemented here increases the spread (Figure 2b) and interannual variability in precipitation extremes, the latter being quantified by the standard deviation (Figure 2c), relative to parent GCMs in the ensemble mean. The interannual variability of Rx10 day, Rx20 day, and Rx30 day events in w/BC simulations is closer to observations (Figure 2c). However, w/BC and no-BC experiments simulate very similar Rx1 day and Rx5 day events, suggesting that the most extreme events are insensitive to BC. Meanwhile, the amplification of spread in no-BC experiments is unanimously reduced with BC (Figure 2b).

3.3. Connecting GCM Biases to RCM Output

Next, we connect the overly wet behavior of the no-BC experiments to mean-state GCM bias transmission. Since winds, moisture, temperature, and SSTs are all bias corrected, we hypothesize that mean-state biases in these fields are driving the downscaled precipitation biases. To test this hypothesis, Figure 3 shows the vertical profiles of wintertime climatological (1980–2014) biases in temperature, zonal wind, and meridional wind in the driving GCMs. These are averaged across the Northeast Pacific and western North America (20°–55°N and 140°–90°W) to represent the upstream and overhead atmospheric environment for systems bringing precipitation to the western US. These biases are almost identically preserved in WRF's intermediate 45-km grid solutions due to our spectral nudging configuration (Section 2). Moisture is addressed later since it is not spectrally nudged.

For all variables, the 9-GCM ensemble of dynamically downscaled GCMs is almost precisely equal to the broader CMIP6 ensemble (compare solid and dashed black curves in each panel), meaning that the selected GCMs are representative. In the 9-GCM mean, there is a near-surface warm bias and deep-layer cold bias aloft (Figure 3a). Thus, there is a lower-tropospheric (surface to 850 hPa) lapse rate bias of ~5 K km^{-1}. The tropospheric cold bias persists in other seasons (Figure S5 in Supporting Information S1) while the low-level instability bias is confined to winter, the wet season for California, Oregon, and Washington. Within the 9-GCM sub-ensemble, CESM2 is the least biased model in terms of column temperature, while CNRM-ESM2-1 and FGOALS-g3 are the most biased. Meanwhile, CESM2 (CNRM-ESM2-1 and FGOALS-g3) biases are smallest (largest) in terms of its dynamically downscaled precipitation (Figure S3 in Supporting Information S1 and Figure 4).

Biases in wintertime zonal wind (Figure 3b) are much larger than those of the meridional wind (Figure 3c), with a near-zero 9-GCM mean bias for the latter. For zonal wind, there is a deep-layer positive bias from about 900 hPa

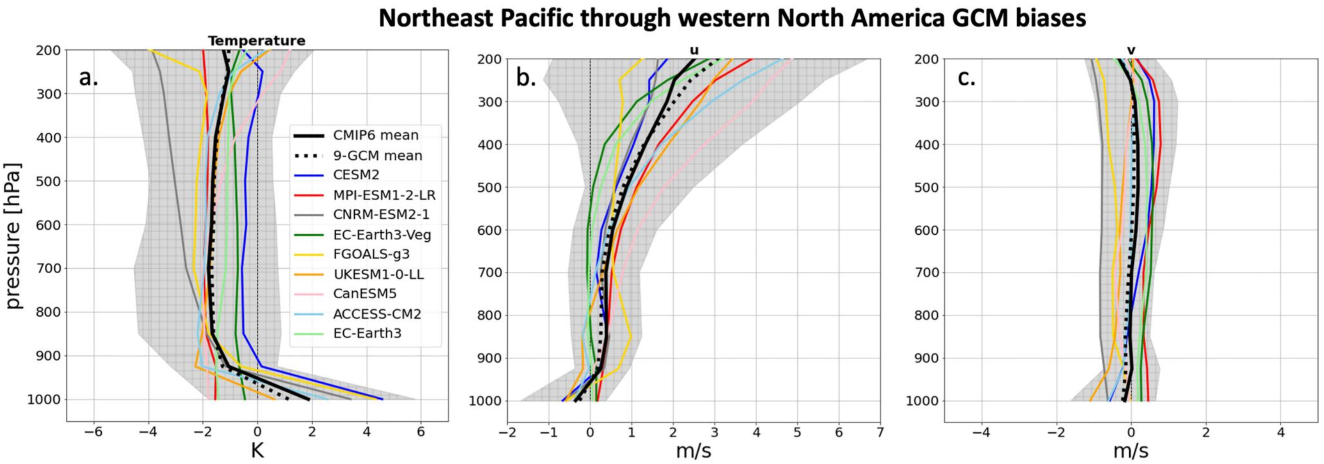

Figure 3. Vertical profiles of wintertime Global Climate Model (GCM) biases, computed against ERA5 from 1980 to 2014, across the Northeast Pacific and western US (20°–55°N and 140°–90°W) for (a) temperature [K], (b) zonal wind [m s^{-1}], and (c) meridional wind [m s^{-1}]. Colored curves are for GCMs that were dynamically downscaled, while the black solid (black dashed) curve represents the Coupled Model Intercomparison Project ensemble (downscaled sub-ensemble) mean (30 different GCMs, 1 realization each). The gray hatched region illustrates the bias spread from the 30 GCM ensemble.

upwards. Further, there is a zonal wind shear bias of +3 m s^{-1} over the depth of the troposphere. This bias is smaller in other seasons (Figure S5 in Supporting Information S1), but also present in springtime.

Next, we explore the connection between each GCM's mean-state bias and bias in dynamically downscaled precipitation bias in the no-BC experiments. This is done by regressing GCM biases for individual variables against dynamically downscaled annual precipitation biases across land regions of the western US (35°–49°N, 125°–105°W mean; Figure 4). GCM bias for a given variable is defined as the difference in the mean-states between the w/BC and no-BC experiments. This is a safe assumption for temperature, zonal wind, and meridional wind since spectral nudging almost precisely conserves the ERA5 and GCM mean states in the former and latter, respectively, as well as SSTs, which are prescribed in downscaling. Regarding moisture, because specific

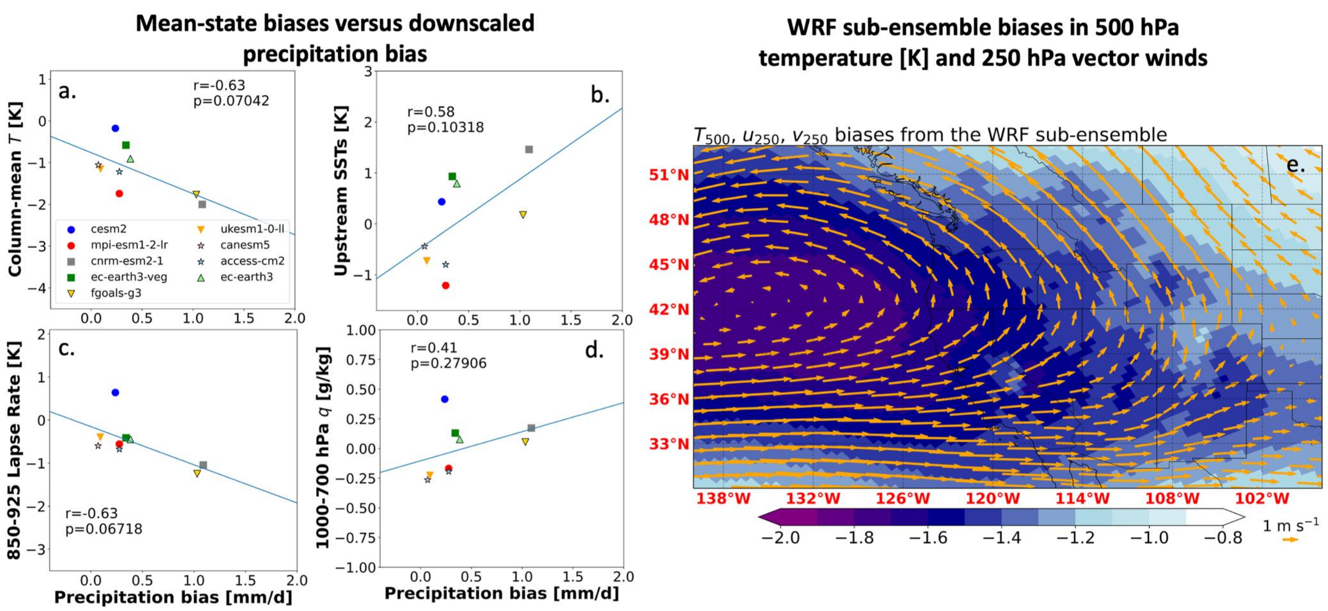

Figure 4. Subpanels (a)–(d) show climatological (1981–2014) mean-state biases in no-BC experiments regressed against dynamically downscaled annual-mean precipitation across the western US (land regions, 35°–49°N, 125°–105°W mean). Panel (a) is for 1,000–250 hPa temperature, (b) for upstream SSTs, (c) for low-level lapse rates, and (d) for lower-tropospheric specific humidity. Correlation coefficients and p-values for each regression are presented in each subpanel. Subpanel (d) shows mean-state biases Weather Research and Forecasting sub-ensemble's 500 hPa temperature and 250 hPa vector wind field.

humidity is un-nudged in the expansive WRF domain, the deviations of the GCMs' mean states from the ERA5 climatology cannot be regarded as inherited GCM biases since WRF integrates its own moisture budget. We thus refer to these moisture deviations as differences rather than biases.

Biases in tropospheric-mean temperature (Figure 4a), low-level (925–850 hPa) lapse rates (Figure 4c), and upstream SSTs (Figure 4b), as well as differences in specific humidity (Figure 4d) correlate positively with dynamically downscaled PRISM-relative precipitation biases. Mean-state biases in low-level lapse rates and column-mean temperature correlate most highly with precipitation biases, as evidenced by a Pearson correlation coefficient (r) of −0.63. Here, r is negative because lapse rate biases become more negative, and tropospheric-mean temperature biases become colder, as precipitation bias becomes more positive. Meanwhile, lower-tropospheric (1000–700 hPa) specific humidity differences are modestly correlated with precipitation bias ($r = 0.40$). Despite the correlation between specific humidity differences and downscaled precipitation bias, it is difficult to definitively assess the role of inherited moisture biases on downscaled precipitation since we are not forcing WRF to conserve GCM-simulated specific humidity. We do note that no-BC- and w/BC-simulated specific humidity values are similar in their respective ensemble means (Figure S6 in Supporting Information S1), and four of nine no-BC experiments are drier than the ERA5 climatology across the Northeastern Pacific and western US (Figure 4d) despite being PRISM-relative wet biased. Meanwhile, upstream SST biases also correlate modestly with precipitation biases ($r = 0.58$).

Several of the thermodynamic biases are dependent on each other. For instance, since the Northeast Pacific is a main moisture source for the western US, it is unsurprising to find that specific humidity differences strongly correlate with upstream SST biases ($r = 0.76$). MPI-ESM1-2-LR, ACCESS-CM2, UKESM1-0-LL, and CanESM5 have a dry bias in mean-state specific humidity relative to the ERA5 climatology. These simulations also have a cold upstream SST bias. Meanwhile CNRM-ESM2-1 and EC-Earth simulations have the warmest SST and highest specific humidity biases. In addition to having the warmest SST bias of any GCM, CNRM-ESM2-1 has the coldest tropospheric bias of any GCM, culminating in the coldest lapse rate bias of any GCM (Figure 4c) Also, tropospheric cold biases are highly correlated with lapse rate biases ($r = 0.85$).

Counterintuitively, biases in zonal and meridional wind speed and the vertical shear therein are negatively correlated with PRISM-relative biases in downscaled precipitation across the western US. The Pacific Northwest is the only subregion where a positive correlation is found (Figure S7 in Supporting Information S1). More generally however, wind biases in the sub-ensemble are tightly coupled to a −2 K minimum in the temperature bias field upstream of the western US via the thermal wind relationship (Figure 4e). This vorticity bias can vary in location and intensity amongst GCMs in the sub-ensemble (Figure S8 in Supporting Information S1) and in the broader 30-GCM CMIP6 ensemble (Figure S9 in Supporting Information S1).

4. Using Physics to Tie GCM Biases to Precipitation Biases

Explanations for why GCM bias cascade is driving the overly wet behavior in the no-BC experiments, and why the wet bias is removed after BC, are rooted in fundamental atmospheric physical concepts. First, the overly unstable lower troposphere in GCMs effectuates increased ascent through a stronger convective available potential energy (Markowski & Richardson, 2011). Second, the overly cold troposphere predisposes rising air parcels to reach saturation and precipitate more quickly through lowered saturation vapor pressures and increased potential for ice nucleation processes (Lamb & Verlinde, 2011; Yau & Rogers, 1996). Across the Pacific Northwest, overly strong zonal wind speed biases from GCMs favor stronger orographically driven ascent (Figure S7 in Supporting Information S1), which can directly enhance precipitation formation through enhanced moisture convergence. Finally, it is important to bias correct the geometry of the dynamic fields as evidenced by the cold bias minimum upstream of the western US, which is tightly coupled to a cyclonic upper-tropospheric vorticity bias. Residing within the mid-latitude westerlies, this feature is directly tied to downstream synoptic-scale ascending motion via the quasigeostrophic omega equation (Bluestein, 1992). Preventing bias cascade eliminates all discussed mean-state biases, subsequently drying the downscaled outputs, improving the representation of mean and extreme precipitation, temperature, and snowpack.

5. Conclusions

Here, we perform and analyze 18 dynamical downscaling simulations across the western United States to better understand the impacts of BC. We find that preventing GCM bias cascade contributes to a relatively minimally biased outcome compared to the case where GCM boundary conditions are not modified prior to downscaling. Mean-state BC also reduces the downscaled spread in precipitation and SWE across simulations forced by different GCMs. Different from post-downscaling BC techniques, the impact of BC is physically interpretable.

We assert that the mean states between GCMs and RCMs must be different for their precipitation yields to be similar over regions of complex terrain. This is because GCMs currently do not resolve the hypergradients in orography necessary to realistically trigger and sustain precipitation systems at sub-synoptic scales. When a biased GCM's mean-state atmosphere is inherited by a high-resolution RCM, which can more appropriately resolve these triggers, storms will be unrealistically predisposed to increased precipitation and storm longevity. The consequence of dynamically downscaling these biases (bias cascade) is an overly wet solution not only for precipitation means (Figure 1c) and extrema (Figure 2a), but also for state-dependent variables such as SWE (Figure 1f). In the end, the value of downscaling may be taken away by propagated GCM biases that interact with realistically resolved features. As applied here, the relatively noninvasive a priori BC of the GCMs' mean-states increases the physical realism of the dynamically downscaled product through physically interpretable mechanisms.

It is beyond the scope of this paper to assess why CMIP6 GCMs share these biases in temperature and winds across the region. However, it is possible that biases in the vertical profile of temperature are due to assumptions in computing vertical velocities within hydrostatic frameworks, as this process can be directly related to temperature via the adiabatic term in the thermodynamic energy equation. Further, known dry SWE biases in GCMs may be responsible for low-level warm biases presented above. More generally, biases in temperature can drive biases in winds and vice versa, and it is unclear if the former is driving the latter. What is clear however is that these biases are dynamically linked, and that gridpoint-specific bias corrections should be applied in dynamical downscaling versus a spatial averaging BC approach. Specifically, if a spatial mean BC method were applied above, the effects of GCMs' vorticity bias on downscaled precipitation would remain. This study also illustrates that, if spectral nudging is employed alongside BC, it should be implemented only for bias-corrected atmospheric variables, moisture aside. To not do so would break fundamental meteorological relationships (e.g., the thermal wind).

Dynamically downscaled GCMs are crucial to better understanding regional climate change, but biases in their historical model behavior are likely to continue into the future period. Moreover, their historical biases may be larger than their future change signals. We strongly urge dynamical downscalers to consider the notion of a priori BC (more appropriately referred to as bias adjustment), in addition to other dynamical downscaling techniques that account for GCM biases (such as the pseudo-global warming technique; Liu et al., 2017; Rasmussen et al., 2014; Schär et al., 1996). Despite the clear strengths of BC presented above, the community should prioritize exploring the physical impacts of a priori BC on future hydroclimate trends and variability as well as its introduced uncertainty.

Data Availability Statement

The versions of WRF used in this study, a sample Jupyter notebook, their attendant files, and the geography files are archived with Zenodo in an open DOI (https://zenodo.org/records/10871230) subject to a Creative Commons License version 4 (Rahimi & Huang, 2024b). Another DOI comprising the WUS-D3 is also available at https://zenodo.org/records/10635867 (Rahimi & Huang, 2024c). All downscaled data for the Western US Dynamically Downscaled Dataset (WUS-D3), including the full 6-hourly WRF datastream (Tier 1), hourly data for select land-surface variables (Tier 2), and a daily post-processed datastream (Tier 3) are located in an open-data bucket on Amazon S3 and in Rahimi and Huang (2024a; https://registry.opendata.aws/wrf-cmip6/). These data are completely open and free to the public. As recommended in the document, these data are most easily downloaded when using Amazon Web Service's (AWS') Command Line Interface (CLI) or with wget. An example is presented in the technical access and usage document. Individualized preprocessing codes were developed to create the intermediate binary files for each GCM before ingestion into WRF.

Acknowledgments
We acknowledge the supporting funding agencies supporting this project: the Department of Energy's HyperFACETSS (A23-1053-S003) project, the Strategic Environmental Research and Development Program Project, the California Energy Commission (EPC-20-006), and the University of California's Climate Ecosystems Future (LRF-18-542511) project. We also thank the computational support through the NCAR-Wyoming Supercomputing Center and the Computational and Information Systems Laboratory. We also acknowledge the TORNERDO consortium, as well as Cora for writing assistance.

References

Bluestein, H. B. (1992). *Synoptic-dynamic meteorology in midlatitudes: Volume I: Principles of kinematics and dynamics: Principles of kinematics and dynamics* (Vol. 1). Oxford University Press.

Bruyère, C. L., Done, J. M., Holland, G. J., & Fredrick, S. (2014). Bias corrections of global models for regional climate simulations of high-impact weather. *Climate Dynamics*, *43*(7), 1847–1856. https://doi.org/10.1007/s00382-013-2011-6

Bukovsky, M. S., & 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. https://doi.org/10.1175/2010JCLI3447.1

Cannon, A. J. (2018). Multivariate quantile mapping bias correction: An N-dimensional probability density function transform for climate model simulations of multiple variables. *Climate Dynamics*, *50*(1), 31–49. https://doi.org/10.1007/s00382-017-3580-6

Daly, C., Halbleib, M., Smith, J. I., Gibson, W. P., Doggett, M. K., Taylor, G. H., et al. (2008). Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. *International Journal of Climatology*, *28*(15), 2031–2064. https://doi.org/10.1002/joc.1688

Daly, C., Neilson, R. P., & Phillips, D. L. (1994). A statistical-topographic model for mapping climatological precipitation over mountainous terrain. *Journal of Applied Meteorology and Climatology*, *33*(2), 140–158. https://doi.org/10.1175/1520-0450(1994)033<0140:ASTMFM>2.0.CO;2

Ehret, U., Zehe, E., Wulfmeyer, V., Warrach-Sagi, K., & Liebert, J. (2012). HESS Opinions "Should we apply bias correction to global and regional climate model data?". *Hydrology and Earth System Sciences*, *16*(9), 3391–3404. https://doi.org/10.5194/hess-16-3391-2012

Giorgi, F., Brodeur, C. S., & Bates, G. T. (1994). Regional climate change scenarios over the United States produced with a nested regional climate model. *Journal of Climate*, *7*(3), 375–399. https://doi.org/10.1175/1520-0442(1994)007<0375:RCCSOT>2.0.CO;2

Goldenson, N., Leung, L. R., Mearns, L. O., Pierce, D. W., Reed, K. A., Simpson, I. R., et al. (2023). Use-inspired, process-oriented GCM selection: Prioritizing models for regional dynamical downscaling. *Bulletin of the American Meteorological Society*, *104*(9), E1619–E1629. https://doi.org/10.1175/BAMS-D-23-0100.1

He, C., Chen, F., Barlage, M., Liu, C., Newman, A., Tang, W., et al. (2019). Can convection-permitting modeling provide decent precipitation for offline high-resolution snowpack simulations over mountains? *Journal of Geophysical Research: Atmospheres*, *124*(23), 12631–12654. https://doi.org/10.1029/2019JD030823

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., et al. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*(730), 1999–2049. https://doi.org/10.1002/qj.3803

Holland, G., Done, J., Bruyere, C., Cooper, C. K., & Suzuki, A. (2010). Model investigations of the effects of climate variability and change on future Gulf of Mexico tropical cyclone activity. *Presented at the Offshore Technology Conference*. OnePetro. https://doi.org/10.4043/20690-MS

Huang, X., Swain, D. L., & Hall, A. D. (2021). Future precipitation increase from very high resolution ensemble downscaling of extreme atmospheric river storms in California. *Science Advances*, *6*(29), eaba1323. https://doi.org/10.1126/sciadv.aba1323

Huang, X., Swain, D. L., Walton, D. B., Stevenson, S., & Hall, A. D. (2020). Simulating and evaluating atmospheric river-induced precipitation extremes along the U.S. Pacific coast: Case studies from 1980–2017. *Journal of Geophysical Research: Atmospheres*, *125*(4), e2019JD031554. https://doi.org/10.1029/2019JD031554

IPCC. (2021). In V. Masson-Delmotte, P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, et al. (Eds.), *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (p. 2391). Cambridge University Press. https://doi.org/10.1017/9781009157896

Jordan, R. E. (1991). A one-dimensional temperature model for a snow cover: Technical documentation for SNTHERM.89. *This Digital Resource Was Created from Scans of the Print Resource*. Retrieved from https://erdc-library.erdc.dren.mil/jspui/handle/11681/11677

Krantz, W., Pierce, D. W., Goldenson, N., & Cayan, D. (2021). *Memorandum on evaluation global climate models for studying regional climate change in California*. California Energy Commission. Retrieved from https://www.energy.ca.gov/sites/default/files/2022-09/20220907_CDAWG_MemoEvaluating_GCMs_EPC-20-006_Nov2021-ADA.pdf

Lamb, D., & Verlinde, J. (2011). *Physics and chemistry of clouds*. Cambridge University Press.

Liu, C., Ikeda, K., Rasmussen, R., Barlage, M., Newman, A. J., Prein, A. F., et al. (2017). Continental-scale convection-permitting modeling of the current and future climate of North America. *Climate Dynamics*, *49*(1), 71–95. https://doi.org/10.1007/s00382-016-3327-9

Maraun, D., Shepherd, T. G., Widmann, M., Zappa, G., Walton, D., Gutiérrez, J. M., et al. (2017). Towards process-informed bias correction of climate change simulations. *Nature Climate Change*, *7*(11), 764–773. https://doi.org/10.1038/nclimate3418

Markowski, P., & Richardson, Y. (2011). *Mesoscale meteorology in midlatitudes*. John Wiley and Sons.

Pierce, D. W., Cayan, D. R., & Thrasher, B. L. (2014). Statistical downscaling using localized constructed analogs (LOCA). *Journal of Hydrometeorology*, *15*(6), 2558–2585. https://doi.org/10.1175/JHM-D-14-0082.1

PRISM Climate Group. (2014). Oregon State University. Retrieved from https://prism.oregonstate.edu

Rahimi, S., & Huang, L. (2024a). CMIP6 GCMs downscaled using WRF [Data]. *Amazon S3*. https://registry.opendata.aws/wrf-cmip6/

Rahimi, S., & Huang, L. (2024b). Jupyter Notebook and comprising data for GRL2023GL106264R: Understanding the Cascade: Removing GCM biases improves dynamically downscaled climate projections [Notebook, some data]. *Zeondo*. https://zenodo.org/records/10871230

Rahimi, S., & Huang, L. (2024c). Versions of WRF, Figure Jupyter notebooks, and geo files for WUS-D3 (version 1) [Software, a Notebook, some data]. *Zenodo*. https://zenodo.org/records/10286545

Rahimi, S., Huang, L., Norris, J., Hall, A., Goldenson, N., Krantz, W., et al. (2023). An overview of the western United States dynamically downscaled dataset (WUS-D3). *Geoscientific Model Development*, *17*(6), 2265–2286. https://doi.org/10.5194/gmd-17-2265-2024

Rahimi, S., Krantz, W., Lin, Y.-H., Bass, B., Goldenson, N., Hall, A., et al. (2022). Evaluation of a reanalysis-driven configuration of WRF4 over the western United States from 1980 to 2020. *Journal of Geophysical Research: Atmospheres*, *127*(4), e2021JD035699. https://doi.org/10.1029/2021JD035699

Rasmussen, R., Ikeda, K., Liu, C., Gochis, D., Clark, M., Dai, A., et al. (2014). Climate change impacts on the water balance of the Colorado headwaters: High-resolution regional climate model simulations. *Journal of Hydrometeorology*, *15*(3), 1091–1116. https://doi.org/10.1175/JHM-D-13-0118.1

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*(8), e2022EF002734. https://doi.org/10.1029/2022EF002734

Schär, C., Frei, C., Lüthi, D., & Davies, H. C. (1996). Surrogate climate-change scenarios for regional climate models. *Geophysical Research Letters*, *23*(6), 669–672. https://doi.org/10.1029/96GL00265

Serreze, M. C., Clark, M. P., Armstrong, R. L., McGinnis, D. A., & Pulwarty, R. S. (1999). Characteristics of the western United States snowpack from snowpack telemetry (SNOTEL) data. *Water Resources Research*, *35*(7), 2145–2160. https://doi.org/10.1029/1999WR900090

Simpson, I. R., Bacmeister, J., Neale, R. B., Hannay, C., Gettelman, A., Garcia, R. R., et al. (2020). An evaluation of the large-scale atmospheric circulation and its variability in CESM2 and other CMIP models. *Journal of Geophysical Research: Atmospheres*, *125*(13), e2020JD032835. https://doi.org/10.1029/2020JD032835

Skamarock, C., Klemp, B., Dudhia, J., Gill, O., Liu, Z., Berner, J., et al. (2019). A description of the advanced research WRF model version 4. https://doi.org/10.5065/1dfh-6p97

Wang, J., & Kotamarthi, V. R. (2013). Assessment of dynamical downscaling in near-surface fields with different spectral nudging approaches using the nested regional climate model (NRCM). *Journal of Applied Meteorology and Climatology*, *52*(7), 1576–1591. https://doi.org/10.1175/JAMC-D-12-0302.1

Wang, J., & 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. https://doi.org/10.1002/2015EF000304

Xu, Z., Han, Y., Tam, C.-Y., Yang, Z.-L., & Fu, C. (2021). Bias-corrected CMIP6 global dataset for dynamical downscaling of the historical and future climate (1979–2100). *Scientific Data*, *8*(1), 293. https://doi.org/10.1038/s41597-021-01079-3

Xu, Z., Han, Y., & Yang, Z. (2019). Dynamical downscaling of regional climate: A review of methods and limitations. *Science China Earth Sciences*, *62*(2), 365–375. https://doi.org/10.1007/s11430-018-9261-5

Yau, M. K., & Rogers, R. R. (1996). *A short course in cloud physics*. Elsevier.

Zhang, M.-Z., Xu, Z., Han, Y., & Guo, W. (2022). Evaluation of CMIP6 models toward dynamical downscaling over 14 CORDEX domains. *Climate Dynamics*, 1–15. https://doi.org/10.1007/s00382-022-06355-5

References From the Supporting Information

Bi, D., Dix, M., Marsland, S., O'Farrell, S., Sullivan, A., Bodman, R., et al. (2020). Configuration and spin-up of ACCESS-CM2, the new generation Australian community climate and earth system simulator coupled model. *Journal of Southern Hemisphere Earth Systems Science*, *70*(1), 225–251. https://doi.org/10.1071/ES19040

Chen, D., Norris, J., Thackeray, C., & Hall, A. (2022). Increasing precipitation whiplash in climate change hotspots. *Environmental Research Letters*, *17*(12), 124011. https://doi.org/10.1088/1748-9326/aca3b9

Danabasoglu, G., Lamarque, J.-F., Bacmeister, J., Bailey, D. A., DuVivier, A. K., Edwards, J., et al. (2020). The community earth system model version 2 (CESM2). *Journal of Advances in Modeling Earth Systems*, *12*(2), e2019MS001916. https://doi.org/10.1029/2019MS001916

Dennis, E. J., & Berbery, E. H. (2021). The role of soil texture in local land surface–atmosphere coupling and regional climate. *Journal of Hydrometeorology*, *22*(2), 313–330. https://doi.org/10.1175/JHM-D-20-0047.1

Döscher, R., Acosta, M., Alessandri, A., Anthoni, P., Arneth, A., Arsouze, T., et al. (2022). The EC-earth3 Earth system model for the climate model intercomparison project 6. *Geoscientific Model Development Discussions*, *2021*, 1–90. https://doi.org/10.5194/gmd-2020-446

Durack, P. J., Taylor, K. E., Eyring, V., Ames, S. K., Doutriaux, C., Hoang, T., et al. input4MIPs: Making [CMIP] model forcing more transparent. Lawrence Livermore National Laboratory (LLNL). https://doi.org/10.2172/1463030

Iacono, M. J., Delamere, J. S., Mlawer, E. J., Shephard, M. W., Clough, S. A., & Collins, W. D. (2008). Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models. *Journal of Geophysical Research*, *113*(D13), D13103. https://doi.org/10.1029/2008JD009944

Li, L., Yu, Y., Tang, Y., Lin, P., Xie, J., Song, M., et al. (2020). The flexible global ocean-atmosphere-land system model grid-point version 3 (FGOALS-g3): Description and evaluation. *Journal of Advances in Modeling Earth Systems*, *12*(9), e2019MS002012. https://doi.org/10.1029/2019MS002012

Morrison, H., & Milbrandt, J. (2015). Parameterization of cloud microphysics based on the prediction of bulk ice particle properties. Part I: Scheme description and idealized tests. *Journal of the Atmospheric Sciences*, *72*(1), 287–311. https://doi.org/10.1175/JAS-D-14-0065.1

Niu, G.-Y., Yang, Z.-L., Mitchell, K. E., Chen, F., Ek, M. B., Barlage, M., et al. (2011). The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements. *Journal of Geophysical Research*, *116*(D12), D12109. https://doi.org/10.1029/2010JD015139

Séférian, R., Nabat, P., Michou, M., Saint-Martin, D., Voldoire, A., Colin, J., et al. (2019). Evaluation of CNRM Earth system model, CNRM-ESM2-1: Role of Earth system processes in present-day and future climate. *Journal of Advances in Modeling Earth Systems*, *11*(12), 4182–4227. https://doi.org/10.1029/2019MS001791

Spero, T. L., Otte, M. J., Bowden, J. H., & Nolte, C. G. (2014). Improving the representation of clouds, radiation, and precipitation using spectral nudging in the Weather Research and Forecasting model. *Journal of Geophysical Research: Atmospheres*, *119*(20), 11682–11694. https://doi.org/10.1002/2014JD022173

Swart, N. C., Cole, J. N. S., Kharin, V. V., Lazare, M., Scinocca, J. F., Gillett, N. P., et al. (2019). The Canadian Earth system model version 5 (CanESM5.0.3). *Geoscientific Model Development*, *12*(11), 4823–4873. https://doi.org/10.5194/gmd-12-4823-2019

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:ACMFSF>2.0.CO;2

Zhang, C., Wang, Y., & Hamilton, K. (2011). Improved representation of boundary layer clouds over the southeast Pacific in ARW-WRF using a modified Tiedtke cumulus parameterization scheme. *Monthly Weather Review*, *139*(11), 3489–3513. https://doi.org/10.1175/MWR-D-10-05091.1