JGR Atmospheres

RESEARCH ARTICLE
10.1029/2023JD040591

An Evaluation of Dynamical Downscaling Methods Used to Project Regional Climate Change

Alex Hall[1], Stefan Rahimi[1], Jesse Norris[1], Nikolina Ban[2], Nicholas Siler[3], L. Ruby Leung[4], Paul Ullrich[5], Kevin A. Reed[6], Andreas F. Prein[7], and Yun Qian[4]

[1]Atmospheric and Oceanic Sciences Department, University of California Los Angeles, Los Angeles, CA, USA, [2]Department of Atmospheric and Cryospheric Sciences, University of Innsbruck, Innsbruck, Austria, [3]College of Earth, Ocean, and Atmospheric Sciences, Oregon State University, Corvallis, OR, USA, [4]Atmospheric, Climate, & Earth Sciences Division, Pacific Northwest National Laboratory, Richland, WA, USA, [5]Department of Land, Air and Water Resources, University of California Davis, Davis, CA, USA, [6]School of Marine and Atmospheric Sciences, Stony Brook University, Stony Brook, NY, USA, [7]National Center for Atmospheric Research, Boulder, CO, USA

Key Points:

- Pseudo global warming and direct dynamical downscaling produce generally similar regional projections over Western North America and Europe
- First-order moisture budget terms determining future hydroclimate change are similar between the two techniques
- Regions whose hydroclimates are dominated by higher-order moisture budget terms require evaluation prior to downscaling

Supporting Information:

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

Correspondence to:

A. Hall,
alexhall@atmos.ucla.edu

Citation:

Hall, A., Rahimi, S., Norris, J., Ban, N., Siler, N., Leung, L. R., et al. (2024). An evaluation of dynamical downscaling methods used to project regional climate change. *Journal of Geophysical Research: Atmospheres*, *129*, e2023JD040591. https://doi.org/10.1029/2023JD040591

Received 21 DEC 2023
Accepted 11 NOV 2024

Abstract In the past decade, dynamical downscaling using "pseudo-global-warming" (PGW) techniques has been applied frequently to project regional climate change. Such techniques generate signals by adding mean global climate model (GCM)-simulated climate change signals in temperature, moisture, and circulation to lateral and surface boundary conditions derived from reanalysis. An alternative to PGW is to downscale GCM data directly. This technique should be advantageous, especially for simulation of extremes, since it incorporates the GCM's full spectrum of changing synoptic-scale dynamics in the regional solution. Here, we test this assumption, by comparing simulations in Europe and Western North America. We find that for warming and changes in temperature extremes, PGW often produces similar results to direct downscaling in both regions. For mean and extreme precipitation changes, PGW generally also performs surprisingly well in many cases. Moisture budget analysis in the Western North America domain reveals why. Large fractions of the downscaled hydroclimate changes arise from mean changes in large-scale thermodynamics and circulation, that is, increases in temperature, moisture, and winds, included in PGW by design. The one component PGW may have difficulty with is the contribution from changes in synoptic-scale variability. When this component is large, PGW performance could be degraded. Global analysis of GCM data shows there are regions where it is large or dominant. Hence, our results provide a road map to identify, through GCM analyses, the circumstances when PGW would not be expected to accurately regionalize GCM climate signals.

Plain Language Summary Simulation of regional climate change requires regional climate models, given that global models cannot adequately represent local climate. Two techniques are typically employed for this purpose. In the first, known as direct downscaling, future climate simulated by a global model is fed to a regional model over the domain of interest. In the second, known as pseudo global warming (PGW), large-scale observations of current climate are perturbed by the global model's future changes, and then fed to the regional model. Direct downscaling appears advantageous because it is more physically consistent with the global model. We test this idea by comparing these two techniques over both Western North America and Europe. They produce surprisingly similar temperature and precipitation projections in both regions. The reason for this similarity is investigated by analyzing the model physics in Western North America. We show there is one component, representing changes to individual storms, that PGW may have difficulty representing. PGW succeeds because this component makes a small contribution in Western North America. But there are other regions where it makes a larger or dominant contribution. In such regions, regional modeling using PGW may produce errors in simulating regional climate change.

1. Introduction

With global warming on track to exceed 2°C by midcentury, there is an urgent need for policies that will help local communities and ecosystems adapt to its impacts. Unfortunately, many local climate change impacts remain poorly constrained, in part because the underlying processes operate at scales that are too small to be resolved by most state-of-the-art global climate models (GCMs) (e.g., Doblas-Reyes et al., 2021). The coarse resolution of GCMs poses a particular problem for hydrologic variables like precipitation and streamflow, which are strongly modulated by subgrid-scale processes like atmospheric convection and turbulence, and also by subgrid-scale

variations in vegetation and topography (e.g., Shepherd, 2019). To compensate for some of these limitations, regional assessments of climate change impacts often rely on simulations from regional climate models (RCMs), which can be run at much higher resolution than a typical GCM (Giorgi & Gutowski, 2015; Giorgi & Mearns, 1999; Gutowski et al., 2021, and many others).

The conventional approach to regional climate modeling is called "direct dynamical downscaling" (e.g., Giorgi, 2019). It involves running a high-resolution RCM (horizontal grid spacing smaller than ∼50 km) over a limited domain, with lateral boundary conditions taken from a GCM simulation. This approach can help to compensate for a low GCM resolution (horizontal grid spacing larger than ∼100 km), and seems straightforward: The GCM's weather and climate is simply regionalized using very similar numerical techniques to what was used to produce the GCM data itself, but at a higher resolution and over a limited area. However, the technique does not address two other problems associated with GCM simulations that are particular to regional climate science. First, GCM simulations of the historical period often exhibit substantial and unambiguous biases relative to observations. These biases are passed to the RCM through lateral and surface boundary conditions, distorting the RCM's characterization of both historical climate and possibly also the response to anthropogenic forcing (e.g., Flato et al., 2013). The bias can be large enough to make the regional data unusable for climate change impact analyses. Bias correction techniques can help mitigate this problem, but they do not entirely eliminate it (e.g., Ehret et al., 2012; Lafon et al., 2013; Teutschbein & Seibert, 2012) and may also distort climate change signals (Maraun et al., 2017). Second, GCM simulations, like the real climate, exhibit significant internal (i.e., unforced) variability. Internal variability is particularly large in magnitude at local and regional scales, where it can obscure the true response to anthropogenic forcing in RCM simulations, even over timescales of several decades (Deser et al., 2012; Lehner et al., 2020). This phenomenon makes it difficult to interpret simulated climate change signals as purely anthropogenic in origin, particularly at local and regional scales. The ambiguity can be greatly mitigated by dynamically downscaling a large ensemble of GCM simulations, but at a high and often prohibitive computational cost (Gutowski et al., 2021).

These limitations of direct dynamical downscaling have given rise to an alternative approach to regional climate modeling called "pseudo global warming" (PGW) experiments (Brogli et al., 2023; Hara et al., 2008; Ikeda et al., 2021; Kawase et al., 2009; Liu et al., 2011, 2017; Rasmussen et al., 2011; Schär et al., 1996; Ullrich et al., 2018; Xue et al., 2023). With the PGW method, a GCM is still used, but only to estimate the mean changes in large-scale variables such as winds, temperature, humidity, geopotential height, and sea-surface temperatures (SSTs) under global warming. These mean-state changes are then added as perturbations to a baseline climate (typically defined using historical reanalysis), and a high-resolution RCM is used to dynamically downscale both the baseline and perturbed climates. The difference between these two simulations reveals how the historical climate of a given region would have been different had it occurred with exactly the same large-scale variability, but with changes in the mean-state climate corresponding to what occurred in the GCM.

The PGW method has several advantages over direct dynamical downscaling. First, by using reanalysis for the lateral boundary conditions for the baseline simulations, PGW simulations have much smaller mean-state biases during the historical period. Second, because the same lateral boundary conditions are used in both the baseline and PGW simulations (differing only in the monthly mean perturbations), the differences between the two simulations are not obscured by internal variability, and thus represent a cleaner approximation of the forced global warming response (Lenderink et al., 2023). This advantage arises from the damping of internal variability in the GCM perturbations themselves, by averaging over long baseline and future periods, typically decades. The damping of internal variability is even greater if the averaging is done over multiple realizations of a large ensemble, or over an ensemble of multiple GCMs. This technique in turn eliminates the need for a large ensemble of regional simulations, and the associated computational cost. Finally, PGW is more effective than direct dynamical downscaling at connecting the impacts of climate change directly to climate events that human societies already encountered in the historical baseline climate. Certain extreme weather events—droughts, floods, blizzards, hurricanes, etc.—leave a strong impression on those who must deal with their effects. Exemplifying the "storyline" approach to regional climate prediction (Shepherd, 2019), PGW simulations present a vision of the future as an alternate history. PGW shows how these same extreme events might have been different, and in many cases more severe, had they occurred several decades later in a world significantly warmer than today's (Li et al., 2023; Patricola & Wehner, 2018; Prein et al., 2017; Ullrich et al., 2018; Wehner et al., 2019; Xue & Ullrich, 2021). PGW approaches have also been used to attribute the impact that climate change has had on recent extreme events by removing the historical large-scale climate change signal (Bercos-Hickey et al., 2022;

Lackmann, 2015; Patricola & Wehner, 2018; K. A. Reed et al., 2020; Kevin A. Reed et al., 2022). Finally, PGW can be useful for separating out the impacts of various elements of the GCM mean climate change (e.g., overall warming, lapse rate changes, and circulations changes) on regional climate signals (Kröner et al., 2017).

Despite these advantages, the PGW method also has a major limitation: because it only accounts for changes in mean-state variables, it neglects aspects of climate change captured by higher-order statistics. In general, it is clear that *anthropogenic changes in any variable at any particular moment in time are not necessarily the same as the mean changes in that variable*. For example, shifts in the tracks of midlatitude storms (e.g., Scheff & Frierson, 2012), or changes in vapor transport resulting from amplified warming at high latitudes and over land (Siler et al., 2019) are best characterized by changes in hourly to daily variability statistics rather than changes in the mean (Siler et al., 2023). While PGW simulations may be useful to understand how historic extreme events would have unfolded in a warmer, moister world, they typically cannot produce extreme events with unprecedented dynamics, and do not modify event sequencing.

Here we are particularly concerned with the impact of using PGW to produce anthropogenic signals in temperature and precipitation. Given the limitation of PGW noted above, the impact on temperature and precipitation extremes may be particularly large. For example, applying mean atmospheric changes in PGW may exclude the changes in atmospheric circulation anomalies associated with heat waves and cold air outbreaks. Such changes could influence simulated anthropogenic changes in temperature extremes. Precipitation may be even more problematic for PGW because it is the result of nonlinear atmospheric processes occurring only in specific phases of synoptic variability. Thus, it is not clear whether the mean large-scale changes in temperature, water vapor, circulation, etc. imposed in PGW will produce even the same mean precipitation changes as direct downscaling, let alone the same change in precipitation extremes. Direct downscaling, on the other hand, should in principle capture thermodynamic and dynamic changes that may be nonuniform across the temperature and precipitation variability distribution. While these limitations of PGW are widely acknowledged (e.g., Gutowski et al., 2021; Ikeda et al., 2021), especially for anthropogenic changes in hydroclimate, the errors they introduce in PGW projections of climate change remain largely unknown.

In this paper, we investigate differences in dynamically downscaled climate change signals derived from direct downscaling (DIRECT) and PGW downscaling techniques using two GCMs and two RCMs across two regions, the western United States and Europe. We seek to determine the circumstances under which PGW produces very similar outcomes to conventional dynamical downscaling and hence can be considered a valid downscaling technique. In certain cases, PGW may also be even superior due to its other advantages noted above. We examine geographical distributions of changes in temperature and precipitation, both mean and extreme. We also take a brief detour to examine snow changes in the western U.S., to illustrate a case where PGW may be more credible because of its more realistic baseline state. Since PGW's performance for changes in hydroclimate is particularly in doubt, we analyze the projected changes in precipitation further by performing a decomposition of the atmospheric moisture budget. We break down the precipitation change into a thermodynamic component, which we expect PGW to simulate well, and a dynamic component, where PGW's performance may be more in doubt. We further decompose these components into mean and eddy subcomponents, where the latter is particularly unlikely to be well produced by PGW (Siler et al., 2023). Our goal is to use this information to provide guidance about the credibility of climate change impact data produced with the PGW technique. We also use these results to provide guidelines as to those GCM analyses that should first be done to ensure PGW will produce credible results in a particular region. A critical and urgent motivation is the widespread and increasing use of the PGW method to provide high-resolution climate change data, without a deep understanding of the method's credibility.

2. Downscaling Techniques

Here, we describe the dynamical downscaling techniques used for the western U.S. (WUS) and European (EU) domains (shown in Figure S1 in Supporting Information S1). These methods are summarized in Table 1.

2.1. WUS Downscaling

For all dynamical downscaling across the WUS, we use the Weather Research and Forecasting Model version 4.1.3 (C. Skamarock et al., 2019; W. C. Skamarock & Klemp, 2008). The boundary conditions (GCM or reanalysis) are applied to a 45-km grid followed by downscaling to 9 km via one-way nesting. Boundary conditions are updated at 6-hourly intervals. Spectral nudging of horizontal winds, temperature, and geopotential is

Table 1
Dynamical Downscaling Technique Characteristics and Their Differences Between the WUS and EU Regions

	WUS	EU
General characteristics		
GCM	CESM2	MPI-ESM-LR
Variant	r1i1p1f1	r1i1p1
RCM	WRF	COSMO
Grid spacing	GCM → 45 km → 9 km	GCM → 12 km
Nudging	Spectral; T, u, and v on 45-km grid	None
Spin-up	One month for each year; years ran in parallel	5 years; time continuous simulations
Emissions scenario	SSP3-7.0	RCP8.5
DIRECT		
Historical years	1980–2000	1995–2005
Future years	2080–2100	2090–2099
PGW		
Simulated reanalysis years	1980–2000	1999–2008
Reanalysis product	ERA5	ERA-Interim
Periods used for deltas	2080–2100 minus 1980–2000	2070–2099 minus 1971–2000
Deltas computed	Monthly, with interpolation	Daily

employed above the PBL to prevent model drift in the larger 45-km domain. WRF configurations optimized for simulating climate across the WUS are used here following the results of Rahimi et al. (2022). In all WRF experiments, individual years are run in parallel to reduce integration times. Each individual year is initialized on 1 August and integrated through August 31 of the following year, with the first month excluded in the simulation output as a spin-up. Note that the land model was initialized with interpolated reanalysis or GCM data. The ramifications of this approach are touched upon by Rahimi et al. (2022), who found that soil moisture and temperatures, as well as surface air temperature and precipitation, were similar between simulations spun up for 1 month versus several years. In all WRF experiments, historical land use and land cover data are used, and the aerosol activation number in the microphysics package is unchanged between experiments.

The historical components of both the PGW and DIRECT experiments span the years 1980–1999. These are driven by ERA5 (Hersbach et al., 2020) and the r11i1p1f1 variant of the Community Earth System Model version 2 (CESM2; Danabasoglu et al., 2020), respectively. For the future DIRECT experiment, the same CESM2 variant's projection (r11i1p1f1) is used from 2080 to 2099. The associated emissions scenario is Shared Socioeconomic Pathways (SSP) 3–7.0, with an end-of-century radiative forcing of 7 W/m^2. For the future PGW experiment, monthly mean deltas are computed between the same CESM2 variant's historical (1980–2000) and future (2080–2100) climate representations. These deltas are imposed on ERA5 on the 15th of each month. On other days, the deltas are linearly interpolated to the given 6-hourly time step. In this way, there is no discontinuity in deltas between the end of 1 month and the beginning of the next. Deltas in surface pressure, mean sea-level pressure, 3-D temperature, zonal wind, meridional wind, geopotential height, SSTs, soil moisture, and soil temperature are applied. For the computation of the specific humidity deltas, a relative humidity conservation constraint is applied, such that the future period has the same time variation in relative humidity as the baseline. Note that because of this arrangement, the implied delta in specific humidity is time-varying in the boundary conditions.

2.2. EU Downscaling

For dynamical downscaling across the EU domain, we used the COSMO-crClim version of the limited-area model from the COnsortium for Small-scale MOdeling (COSMO; Baldauf et al., 2011; Rockel et al., 2008; www.cosmo-model.org). Unlike the WUS experiments, we downscale reanalysis and GCM outputs directly to a single 12-km grid without nudging. The historical simulations are driven by ERA-Interim (ERAI; Berrisford

et al., 2011) over 1999–2008. (Note this is a different reanalysis product from the one used for the WUS simulations.) For the historical portion of the DIRECT experiment, boundary conditions were derived from variant r1i1p1 from the Max Planck Institute's low-resolution Earth System Model (MPI-ESM-LR; Stevens et al., 2013) under CMIP5 historical emissions from 1995 to 2005.

For the future experiments, PGW deltas were computed by differencing the periods 1971–2000 and 2070–2099 from the same MPI-ESM-LR variant, representing a century of this model's climate change signals, applied to a reanalysis baseline corresponding to the beginning of the 21st century. The associated emissions scenario is the Representative Concentration Pathway (RCP) 8.5, with an approximate radiative forcing of 8.5 W/m^2 by the end of the century. Deltas were computed at daily intervals and then smoothed. Note that this technique differs slightly from that used to compute deltas for the WUS. As with the WUS, deltas in surface pressure, mean sea-level pressure, 3-D temperature, zonal wind, meridional wind, geopotential height, SSTs, soil moisture, and soil temperature are applied. For the computation of the specific humidity deltas, new specific humidity values are computed to match the change in temperature and relative humidity simulated by the GCM. Note that if relative humidity changes are small, in practice, this is similar to the WUS experiments. In contrast to the WUS WRF experiments, deltas were computed and applied daily. Meanwhile, direct downscaling of the same MPI-ESM-LR variant over 2090–2099 was conducted for the DIRECT experiment. An additional difference compared to the WU simulations is that all EU runs are continuous with an initial spin-up period of 5 years. Although the historical and future periods are slightly different for the PGW and DIRECT EU experiments, in both cases, approximately 100 years' worth of climate change from the MPI-ESM-LR GCM is downscaled. Thus, if PGW is giving the same outcomes as DIRECT, the climate change patterns in the two experiments ought to be very similar, even if their magnitudes differ slightly. Finally, we note that internal variability effects may be slightly larger for the EU DIRECT experiment than for its WUS counterpart, owing to the shorter simulation periods (one decade vs. two).

3. Temperature Changes

3.1. Mean Warming

We begin our comparison of PGW and DIRECT future climate signals by examining mean wintertime and summertime warming in the two domains (Figure 1). For reference, we also include the associated GCM warming in both domains. We also report spatial correlations between PGW and DIRECT change patterns (Table 2, first two rows). This is intended to be a simple similarity metric between the two patterns. There are many ways in which it can be misleading, including not accounting for differences in the magnitude of the fields, and lumping together varying degrees of correlation across multiple spatial scales (e.g., a correlation could arise from fine-scale or large-scale variability). Still, between visual inspection of the fields and the accompanying correlation, meaningful assessments of their similarity can be made.

In general, the PGW and DIRECT produce very similar warming patterns (spatial correlations >0.9). Moreover, both the PGW and DIRECT patterns generally exhibit the same distinct differences from the driving GCM warming pattern. This indicates that the same local processes shaping warming patterns are at work in the regional climate model, no matter whether it is run in PGW or DIRECT mode.

Evidence of the above statements can be seen in the warming patterns of the WUS. During winter, the GCM exhibits anomalous warming over a broad swath of the northern Rockies, likely due to snow albedo feedback. (Figure S2 in Supporting Information S1 shows the corresponding snow losses in the GCM, PGW, and DIRECT cases.) In the PGW and DIRECT simulations, this anomalous warming is largely suppressed, albeit with smaller pockets of enhanced warming clearly associated with snow loss in the regional simulations (Figure S2 in Supporting Information S1). The more realistic topography and hence snow distributions in the two regional simulations appear to be similarly correcting a snow albedo feedback error in the GCM due to an unrealistic snowpack distribution, an effect seen in other future regional simulations (Walton et al., 2017). We note that in the DIRECT experiment, this warming enhancement in snow loss areas is somewhat magnified relative to PGW. This can be traced to somewhat larger snow losses in the DIRECT experiment, likely due to unrealistically large snow cover in the historical component of the DIRECT simulation (Rahimi et al., 2024; Risser et al., 2024). We discuss this effect further in Section 4b. In the WUS summertime, DIRECT and PGW are also similar to one another (spatial correlation 0.97), and are similarly differentiated from the pattern of the driving GCM. However, there is again greater warming in the DIRECT experiment where there is greater snow loss, as in winter. For the EU domain, the wintertime warming is quite similar in GCM, PGW, and DIRECT (spatial correlation 0.9).

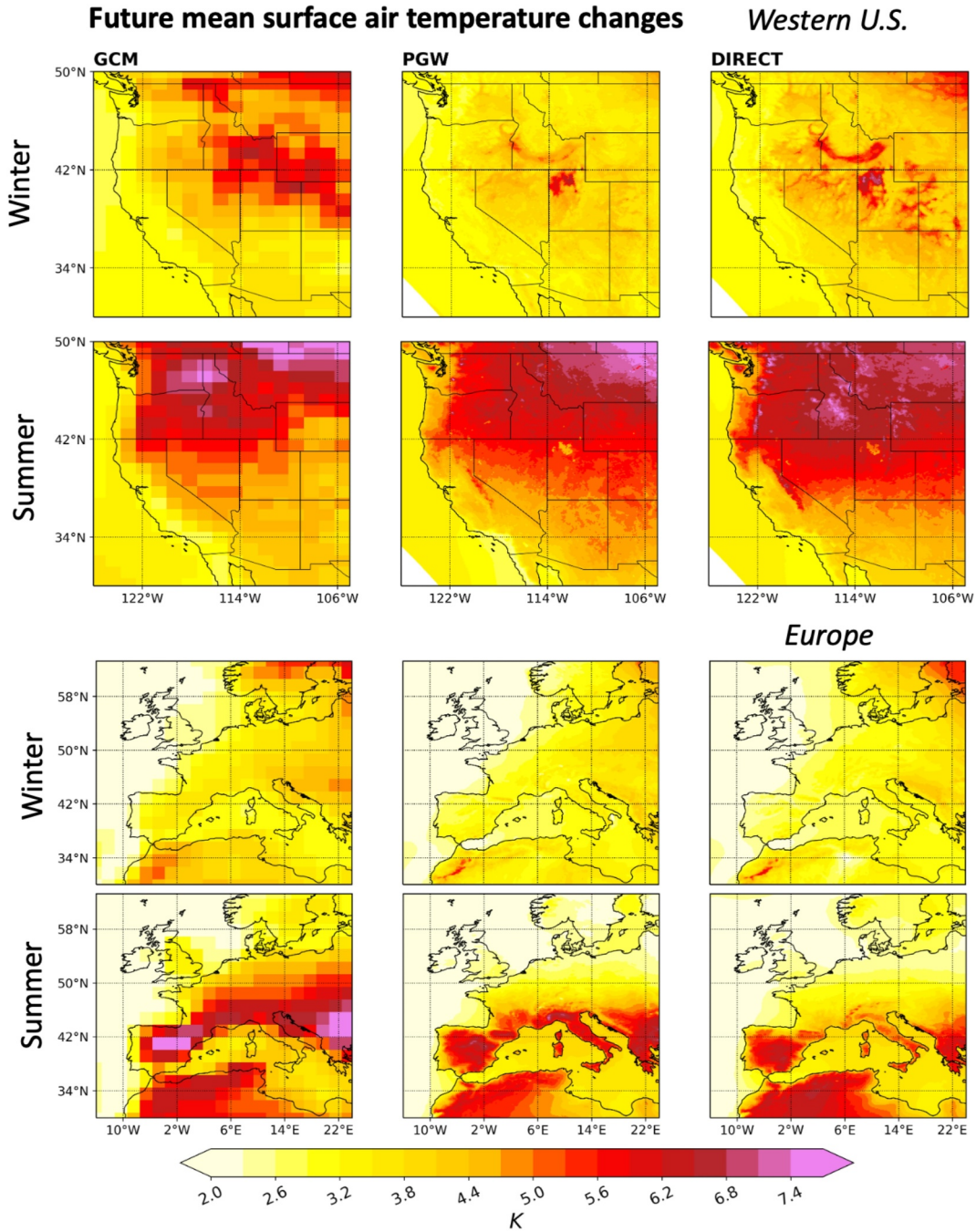

Figure 1. Future climatological changes in (top 6 panels) western U.S. and (bottom 6 panels) European experiments. Wintertime and summertime results are shown for each region. Results from the driving GCM are shown in the leftmost column, followed by the PGW results in the middle column. Finally, the DIRECT results are shown in the rightmost column. Units are degrees K. For reference, the corresponding 21st century global warming in CESM2 is about 3.2 K, while for MPI-ESM-LR, the figure is about 3.6 K. See Table 1 for further information about the historical and future time periods used to compute the climate change signals in the WUS and EU DIRECT and PGW experiments. For the WUS domain, the GCM change signals were computed for the same time periods as the DIRECT and PGW experiments. For the EU domain, the GCM change signals were computed for the same time period as the DIRECT experiment (2090–2099 minus 1995–2005). The spatial correlations between the PGW and DIRECT patterns are as follows: 0.92 (WUS DJF), 0.97 (WUS JJA), 0.90 (EU DJF), and 0.97 (EU JJA), also reported in Table 2.

Table 2
Spatial Correlations Between the PGW and DIRECT Patterns Shown in Figures 1, 2, 3, and 5

	WUS	EU
DJF ΔT	0.92	0.90
JJA ΔT	0.97	0.97
ΔT*at P*95	0.90	0.93
DJF ΔP	0.72	0.79
JJA ΔP	0.05	0.66
DJF ΔP *at P*95	0.84	0.79
JJA ΔP *at P*95	0.09	0.66

Examining the summertime EU warming, the PGW and DIRECT experiments exhibit greatly reduced warming in continental Europe north of the Alps, compared to the GCM. This could be due to systematic differences in the land-atmosphere feedbacks between the RCM and GCM. In any event, the PGW technique captures the essential features of the DIRECT warming pattern very well (spatial correlation 0.97).

3.2. Changes in Extreme Temperature

Next, we examine differences in future temperature changes on extremely hot days (Figure 2, spatial correlations between PGW and DIRECT in Table 2). We define an extremely hot day as the 95th percentile of the summertime temperature distribution. In the WUS, the warming on such days exhibits a broadly similar pattern in the PGW and DIRECT experiments (spatial correlation 0.9), with large temperature increases greater than 5°C on the interior side of the coastal ranges, and north of about 37°N. In these areas, however, there are also differences, with extreme temperatures being about 1°C warmer in the Northern Great Basin and about 1°C cooler in the Northern Great Plains in the DIRECT experiment compared to PGW. These differences could be partly due to anthropogenic circulation changes on extremely hot days incorporated in the DIRECT experiment, but not captured by the anthropogenic changes in the mean circulation imposed in the PGW experiment.

In the EU domain, the spatial pattern of the warming on extremely hot days is also broadly similar in PGW and DIRECT cases (spatial correlation 0.93), with greater increases over southern Europe and North Africa. But as with the WUS, the differences between PGW and DIRECT are locally significant. The increases in the DIRECT experiment are approximately 2°C less than PGW in southern and southeastern Europe. Again, this could indicate that in the DIRECT experiment, there are changes in synoptic circulation associated with heat waves that are not included in the PGW experiment. In this case, such circulation anomalies, which could include changes in

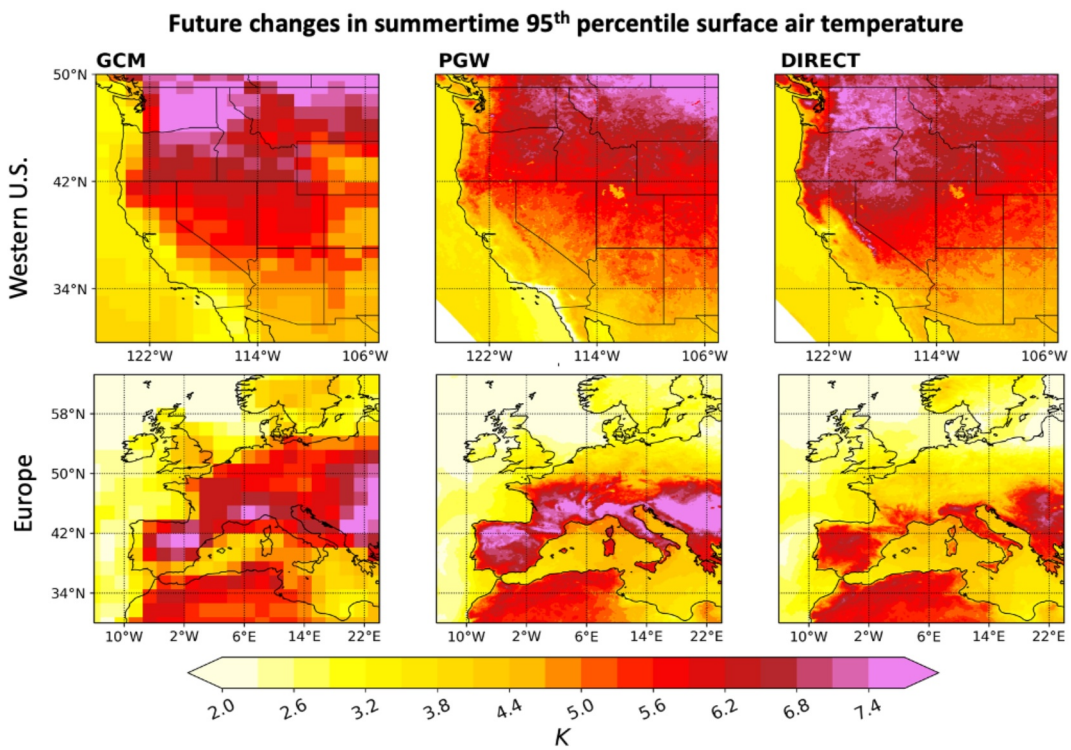

Figure 2. As in Figure 1, but for the 95th percentile of daily mean JJA surface air temperature. Units are degrees K. The spatial correlations between the PGW and DIRECT patterns are as follows: 0.90 (WUS) and 0.93 (EU), also reported in Table 2.

blocking circulation patterns or land/sea breeze systems, dampen extreme heat. These statements appear to be contradicted by the fact that the PGW pattern is more similar to the GCM pattern than that of the DIRECT experiment. (Presumably, the GCM has similar circulation changes during extreme heat as the DIRECT experiment, albeit at a coarser resolution.) However, the RCM has a different land surface model than the GCM, in addition to having finer resolution. Land-atmosphere feedbacks such as the latent and sensible heat flux damping response to temperature on extremely hot days are likely different in the RCM compared to the GCM. This makes it difficult to fully disentangle the reasons for similarities and differences between the GCM and RCM patterns. For example, the PGW experiment may be accidentally more similar to the GCM, because of the combination of the effects of the RCM land surface model and the absence of synoptic circulation changes. Finally, we note that land-atmosphere feedbacks may also depend on the baseline climate, which is somewhat different among GCM, DIRECT, and PGW cases. For example, temperature biases in the historical climate could result in different soil states and hence a different feedback response upon warming.

4. Changes in Hydroclimate

4.1. Mean Precipitation

Next, we turn to an examination of changes in mean precipitation (Figure 3, spatial correlations in Table 2). In the WUS wintertime, when most of the annual precipitation falls in the region, the PGW experiment produces a similar precipitation change as the DIRECT experiment, although the spatial correlation (0.72) is lower than that associated with warming patterns. A large increase is seen, inherited from the orographical-modulated GCM with downscaling, especially in the mountains adjacent to the Pacific coast. Similar orographic increases are also seen in the interior, being somewhat larger in DIRECT than in PGW. Despite a broad similarity (spatial correlation of 0.79), differences between PGW and DIRECT are also visible in the EU domain in winter. Although the general pattern of drying over the Mediterranean and wetting further north is apparent in both experiments, the wetting is significantly greater in the DIRECT case. In some locations, such as the Iberian Peninsula, differences are particularly pronounced. The magnitude of the wetting in the DIRECT experiment is generally much more similar to that of the driving GCM than PGW. One explanation for this is that aspects of the GCM's wintertime synoptic variability during precipitation events are inappropriately suppressed when the PGW technique is applied. Another explanation could be that internal variability is especially impacting EU precipitation signals, owing to the design of the EU experiments. The time periods used to produce the DIRECT and GCM change patterns in Figure 3 are the same (2090–2099 minus 1995–2005), in contrast to those using the PGW deltas (2070–2099 minus 1971–2000). Similar phasing of internal variability during the 2090–2099 and 1995–2005 periods could account for the greater similarity of the GCM and DIRECT patterns.

Meanwhile, in the WUS summertime, the PGW and DIRECT experiments mostly disagree (spatial correlation 0.05). In the DIRECT experiment, the change is characterized by a small wetting in the Great Basin, and drying further north, somewhat similar to the GCM. In the PGW counterpart, by contrast, the wetting in the Great Basin is largely absent, and significant drying is seen in the Arizona/New Mexico area, a sign of a simulated weakening North American monsoon. This change is not seen in the GCM, which shows weak wetting throughout the domain. The weakening of the North American monsoon in the PGW experiment must be due to an aspect of synoptic variability that is unrealistically omitted in the PGW method. Boos and Pascale (2021) suggest that North American monsoon precipitation is associated with convectively enhanced orographic rainfall due to deflections of the extratropical jet stream. To the extent that there is an anthropogenic change in this mechanism, the PGW experiment could be missing its high-frequency (submonthly) component. We return to explanations for the failure of PGW to reproduce the DIRECT experiment in the WUS summertime in Section 5.2. Finally, for the EU domain in summertime, the precipitation change exhibits some similarities between DIRECT and PGW (spatial correlation 0.66). Both experiments also show similar departures from the GCM pattern, indicating that higher resolution in both cases helps to generate the relevant local dynamics shaping patterns of regional precipitation change.

4.2. Snow

Here, we take a brief detour to illustrate how the greater realism of the historical climate in the PGW setup can actually yield more credible climate change signals. The case we examine is the snow change in the WUS. In the WUS wintertime, we have seen that PGW and DIRECT produce similar changes in temperature and precipitation.

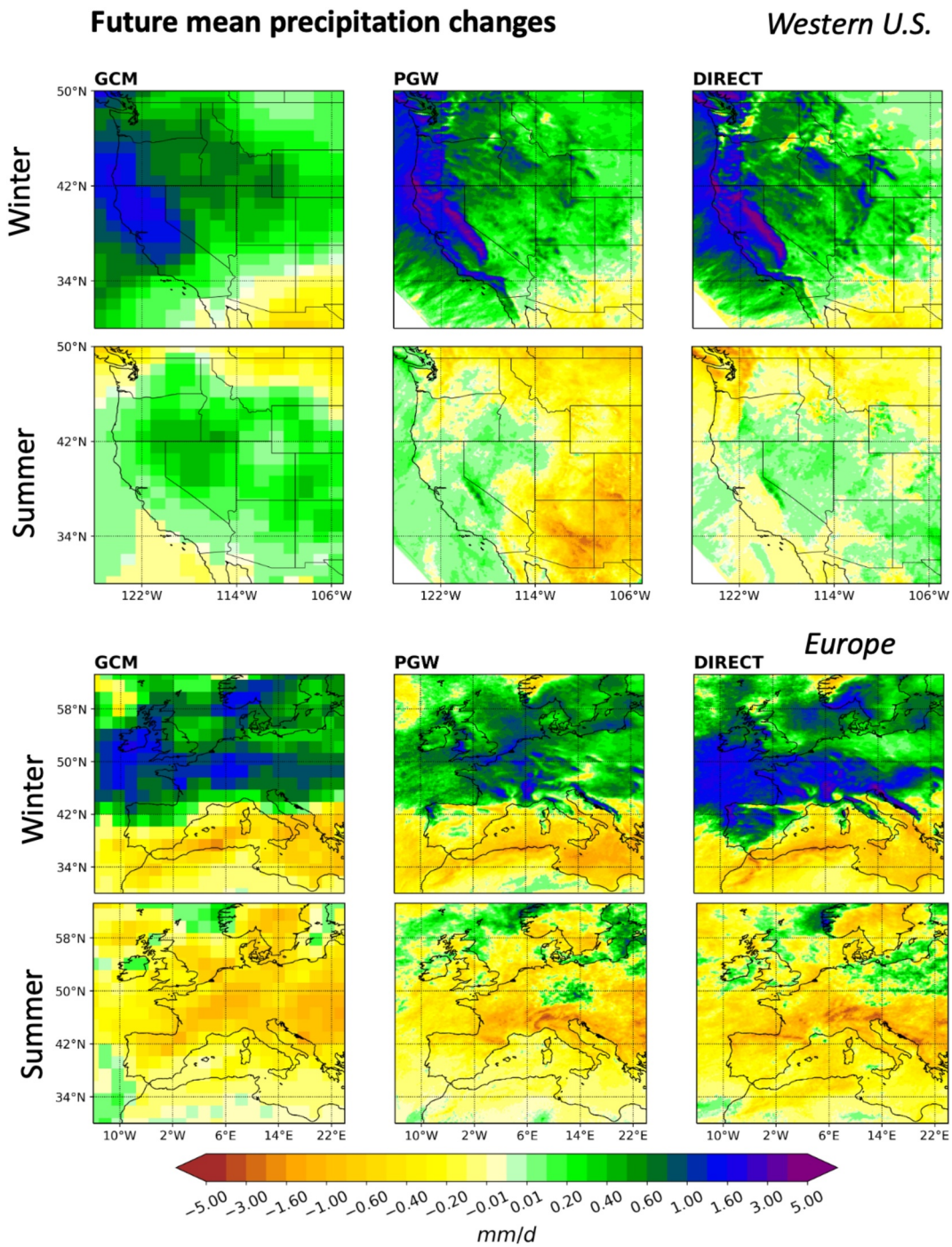

Figure 3. As in Figure 1, but for changes in climatological mean precipitation, including both rainfall and snowfall. Units are mm/d. The spatial correlations between the PGW and DIRECT patterns are as follows: 0.72 (WUS DJF), 0.05 (WUS JJA), 0.79 (EU DJF), and 0.66 (EU JJA), also reported in Table 2.

However, the two techniques produce very different outcomes for snow, a variable that is strongly dependent on temperature and precipitation. Figure 4 shows the seasonal climatologies of snow water equivalent (SWE) at over 700 locations where snow has been consistently observed in the WUS. We first compare the SWE values of the historical climate in the PGW and DIRECT experiments to the observed SWE values. The PGW baseline climate is a downscaling of renalysis. As expected, it agrees reasonably well with the observations, although its SWE is somewhat unrealistically low at the seasonal SWE peak. In contrast, the baseline climate of the DIRECT

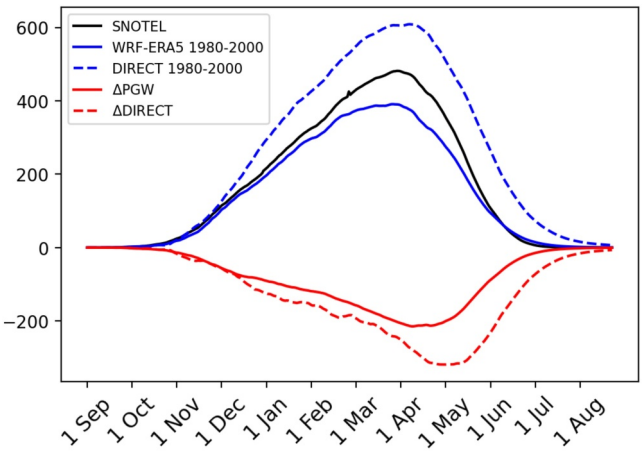

Figure 4. Snow water equivalent (SWE; mm) over the course of the WUS "water year," averaged at 703 western U.S. SNOw TELemetry (SNOTEL) locations, and anthropogenic changes therein. The black solid curve shows the seasonal climatology of the observed SNOTEL data, averaged over the period 1980–200. The blue dashed line shows the corresponding data from the historical period of the DIRECT experiment. The blue solid line corresponds to the values from the reanalysis-driven experiment that forms the historical baseline for the PGW experiment. Red curves show projected changes in the last 20 years of the 21st century relative to the 1980–2000 period for the PGW (solid) and DIRECT (dashed) cases.

experiment is a downscaling of the GCM historical climate. Because this particular GCM is wet-biased (Rahimi et al., 2024; Risser et al., 2024), the SWE values are unrealistically large, by as much as 30%. These biases have large consequences for the future SWE change (Figure 4; red lines). Because it starts out with so much more snow, the DIRECT experiment also loses more snow in the future than PGW. This difference between PGW and DIRECT can also be seen by closely comparing the geographical distributions of SWE loss in the PGW and DIRECT cases in Figure S2 in Supporting Information S1. In addition, the seasonal structure of the loss is different, with the date of maximum loss occurring nearly a month later in the DIRECT experiment. Since the changes in temperature and precipitation are so similar in the two experiments, the differences in the magnitude and timing of the SWE losses must be primarily due to the different SWE biases of the baseline climatology in the two experiments. Because the PGW baseline climate is less biased than that of the DIRECT experiment, its more modest and earlier-phased SWE losses are likely more physical than those of the DIRECT experiment. This illustrates a key advantage of PGW for simulation of variables whose behavior under climate change is sensitive to biases in the baseline climate.

The final variable we examine is the change in extreme precipitation in both domains, defined as the 95th percentile of daily mean precipitation occurring in wintertime and summertime (Figure 5, spatial correlations in Table 2). The changes are broadly consistent with the mean precipitation change (Figure 3): PGW reproduces the large increase in wintertime extreme precipitation seen in the DIRECT experiment in the mountain ranges adjacent to the Pacific, and the smaller increases elsewhere (spatial correlation 0.84). A modest difference between PGW and DIRECT is seen around the state of Washington, where the extremes increase in PGW and are generally drier in DIRECT. Meanwhile, PGW produces a muted increase in wintertime precipitation extremes in the midlatitudes of the EU domain compared to DIRECT, albeit with a similar spatial pattern (spatial correlation 0.79). PGW largely fails to capture the summertime extreme precipitation changes in WUS (spatial correlation 0.09), showing a large drying in the North American monsoon region which is absent in DIRECT. This is very similar to the PGW experiment's excessive mean drying in this subregion (Figure 3). Below, in Section 5.2, we return to the reasons for this discrepancy. PGW does a modest job capturing the change in JJA extreme precipitation in the EU case (spatial correlation 0.66).

5. Moisture Budget Analyses

In general, the PGW method captures the changes in means and extremes of temperature and precipitation reasonably well. However, there are exceptions to this general statement, with pattern similarities for mean and extreme precipitation being generally lower than for temperature. Here, we seek to understand why PGW reproduces the changes seen in the DIRECT method so well, and to the extent the two methods give different answers, to understand why. We focus on precipitation changes, because of the lower agreement between the two methods for precipitation, and because even mean precipitation changes are tightly associated with only particular phases of historical and future synoptic variability. This is precisely the type of variable that ought to challenge the PGW method. We use the WUS domain as our test bed. Our goal is to illustrate how the mean and extreme precipitation differences can be diagnosed using a moisture budget decomposition technique. Note that this thermodynamic/dynamic decomposition approach is similar in spirit to previous work focused on European precipitation (de Vries et al., 2022).

5.1. Methods

Here, we describe the moisture budget decomposition technique, which we evaluate using 6-hr data for the WUS domain (model-level data were not available for the European domain, leading to large errors in the moisture balance). The steady-state moisture budget under climate change is as follows:

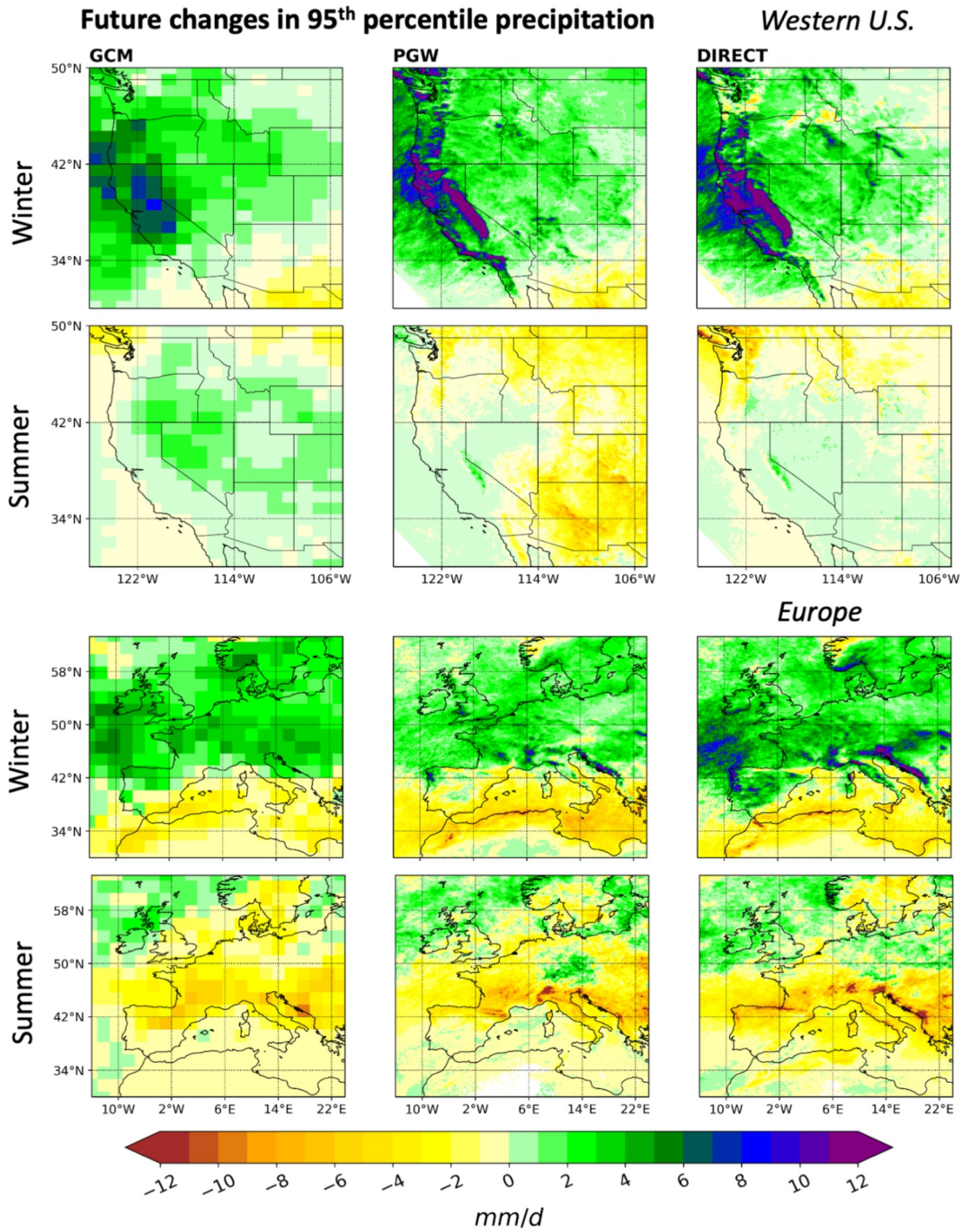

Figure 5. As in Figure 1, but for the 95th percentile of daily mean precipitation, which includes both rainfall and snowfall. Units are mm/d. The spatial correlations between the PGW and DIRECT patterns are as follows: 0.84 (WUS DJF), 0.79 (EU DJF), 0.09 (WUS JJA), and 0.66 (EU JJA), also reported in Table 2.

$$\Delta\overline{\text{P-E}} = -\nabla \cdot \left[\overline{\{q_w \mathbf{v}_w\}} - \overline{\{q\mathbf{v}\}}\right] \quad (1)$$

where q and $\mathbf{v}$ denote specific humidity and the horizontal wind vector (u; v), respectively; { }denotes a mass-weighted vertical integral, $\frac{1}{g}\int_0^{p_s} dp$; the overbar denotes the average of a given term or expression across a

given population of events (mean climate or extremes; annual or seasonal); the w subscript denotes a given term evaluated in the future warmer climate (thus, its absence means that the given term is evaluated in the historical climate); Δ denotes the change in the given term from historical to future climates; g denotes the gravitational constant; p denotes pressure; and p_s denotes surface pressure.

WRF uses an Arakawa C-grid, that is, with u and v defined at the east/west and north/south interfaces of grid points, respectively. Thus, to evaluate q and p at the same locations as u or v, we average q and p across adjacent grid points, for example, $q_{i,u} = (q_i + q_{i-1})/2$ and $q_{j,v} = (q_j + q_{j-1})/2$, where i,u denotes the i th x coordinate on the u grid and j,v denotes the j th y coordinate on the v grid. We then discretize horizontal divergence, $\nabla \cdot X_{i,j}$, for some vector $X = (U;V)$, as $\frac{U_{i+1,u} - U_{i,u} + V_{j+1,v} - V_{j,v}}{\delta x}$. Here $\delta x = 9{,}000$ m is the grid spacing. In this way, the variables are interpolated back to the center of grid cells, that is, at the same location as P and E. Meanwhile, WRF uses a terrain-following vertical coordinate (with 39 levels used in these simulations), with pressure output on model levels. Thus, a vertical integral of some expression X is calculated as $\frac{1}{g}\sum_k \frac{X_k + X_{k+1}}{2}\delta p_k$, where $\delta p_k = p_k - p_{k+1}$. In this way, X is averaged between the k th and $(k+1)$ th levels and multiplied by the corresponding pressure difference.

Typically, the right-hand side of Equation 1 is decomposed into thermodynamic, dynamic, and eddy terms (e.g., Seager et al., 2010). Here, we employ an updated method that allows thermodynamic and dynamic terms to be partitioned into mean and eddy components (Siler et al., 2023). This partitioning allows us to evaluate hypotheses relating to the performance of PGW and DIRECT techniques: Since the PGW method explicitly incorporates GCM changes in both the mean thermodynamic (e.g., $\overline{T}$ and $\overline{q}$) and mean dynamic (e.g., $\overline{u}$ and $\overline{v}$) terms, it seems reasonable to expect it to capture the mean thermodynamic and dynamic components of the moisture budget change. Meanwhile, because the thermodynamic and dynamic eddy components of the GCM driving data are not included in PGW, it seems reasonable that PGW would not faithfully capture those changes. Instead, PGW is likely to simply rescale the historical eddies by the future mean deltas. We note, however, that PGW may end up being able to capture these changes through simulated fine-resolution internal dynamics, for example, by the innermost domain of the regional model.

The updated decomposition method consists of first defining total thermodynamic and dynamic terms:

$$\Delta\overline{\mathrm{P\text{-}E}}_{\mathrm{q}} = -\nabla \cdot \overline{\{\mathbf{v}\Delta q\}}$$

$$\Delta\overline{\mathrm{P\text{-}E}}_{\mathrm{v}} = -\nabla \cdot \left[\overline{\{(q_w - \Delta q)\mathbf{v}_w\}} - \overline{\{q\mathbf{v}\}}\right] \quad (2)$$

These thermodynamic and dynamic terms comprise the total change to moisture flux convergence (right-hand side of Equation 1), excluding second-order terms, an approximation that we evaluate.

In these terms, q, u, and v are all sorted in ascending order of q prior to averaging over the given population of events. Thus, Δq denotes the change in a given percentile of q. And, for example, to calculate $v\,\Delta q$, u and v are evaluated at the given percentile of q. Conventional thermodynamic and dynamic decompositions only represent changes in the means ($\Delta\overline{q}$, $\Delta\overline{u}$, and $\Delta\overline{v}$). This method allows the thermodynamic and dynamic terms to encompass all changes to moisture and winds, respectively, whether changes to the means ($\Delta\overline{q}$, $\Delta\overline{u}$, $\Delta\overline{v}$) or eddies ($\Delta q'$, $\Delta u'$, $\Delta v'$). (Here, the prime denotes the deviation at a given time step from the mean of the population, represented by the overbar.) The asymmetry between the thermodynamic and dynamical terms arises from the fact that all variables are indexed on q, and not on u and v.

We then define mean thermodynamic (changes to $\overline{q}$) and dynamic terms (changes to $\overline{u}$, $\overline{v}$):

$$\Delta\overline{P - E}_{\overline{q}} = -\nabla \cdot \overline{\{\mathbf{v}\Delta\overline{q}\}}$$

$$\Delta\overline{P - E}_{\overline{\mathbf{v}}} = -\nabla \cdot \overline{\{q\Delta\overline{\mathbf{v}}\}} \quad (3)$$

And, finally, the thermodynamic eddy term (changes to q') and dynamic eddy term (changes to u', v') are calculated as the residuals between the total and mean terms:

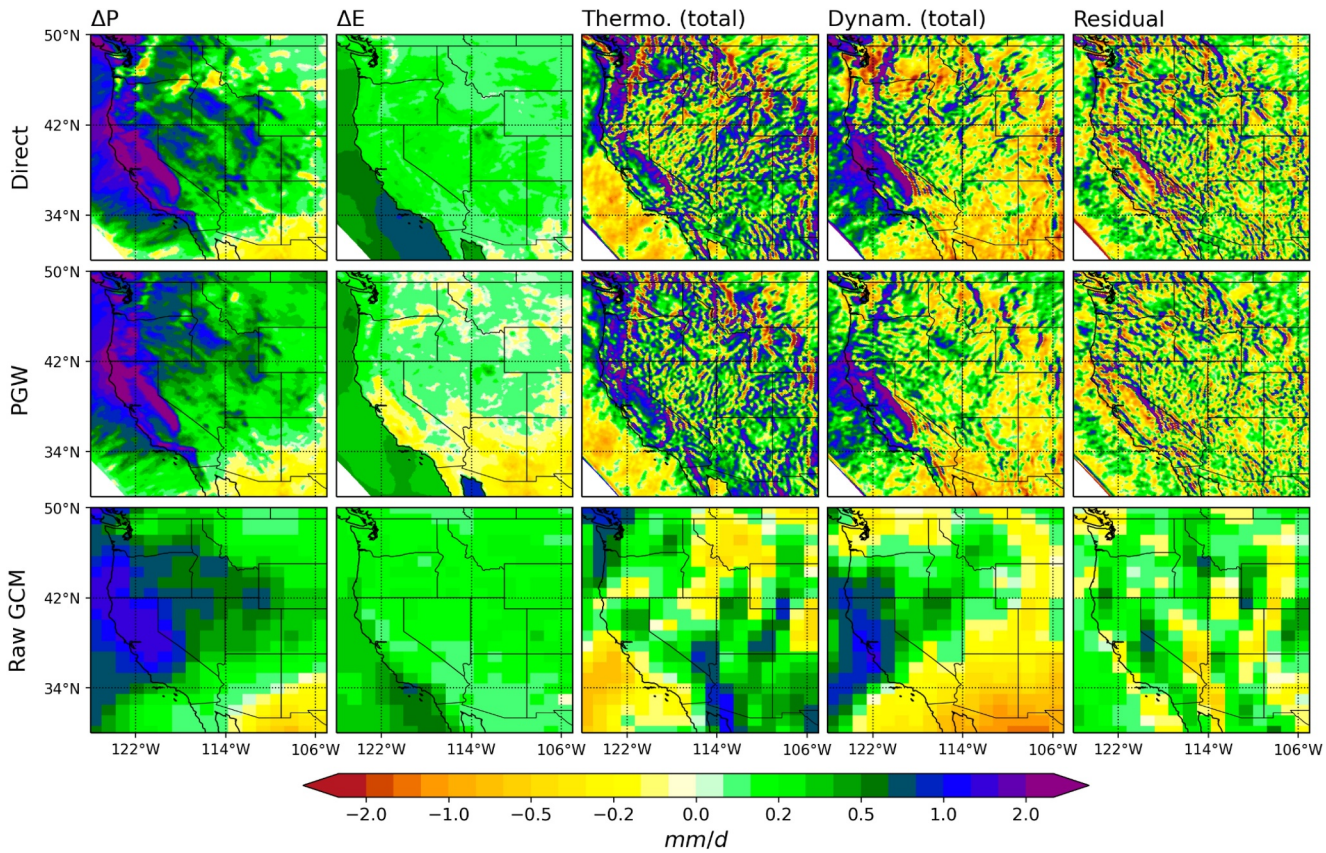

Figure 6. Moisture budget analysis for changes to WUS DJF mean precipitation, shown for (top row) DIRECT, (middle row) PGW, and (bottom row) raw GCM data. The first column shows the change in precipitation. The 2nd column shows the change in evaporation. The 3rd column shows the total thermodynamic component of the change, $\Delta\overline{P-E}_q$, while the 4th shows the total dynamical component, $\Delta\overline{P-E}_{\mathbf{v}}$. The residual, shown in the 5th column, is calculated as $\Delta P - \Delta E - \Delta\overline{P-E}_q - \Delta\overline{P-E}_{\mathbf{v}}$. Units are mm/d. The spatial correlations between the PGW and DIRECT patterns are as follows: 0.80 (thermodynamic), 0.74 (dynamical), and 0.94 (residual), and are also shown in Table 3.

$$\Delta\overline{P-E}_{q'} = \Delta\overline{P-E}_q - \Delta\overline{P-E}_{\bar{q}}$$
$$\Delta\overline{P-E}_{\mathbf{v}'} = \Delta\overline{P-E}_{\mathbf{v}} - \Delta\overline{P-E}_{\bar{\mathbf{v}}} \quad (4)$$

See Siler et al. (2023) for more details of the derivation and justification of these terms.

5.2. Decomposition Applied to WUS Changes in Mean Precipitation

We first focus on the moisture budget decomposition to the WUS mean precipitation change in DJF, the heart of the wet season in most of the region. Figure 6 shows the precipitation change patterns we are attempting to diagnose (the leftmost column, also shown in Figure 3), together with associated contributions from various elements of the moisture budget. The values seen in the residual term (the rightmost column of Figure 6) are generally smaller than the thermodynamic and dynamic terms, although locally they can be significant. The residual term contains higher-order components of the moisture budget. In addition, it contains differences between the numerical methods of the atmospheric models and those of moisture budget decomposition (Seager & Henderson, 2013). The first row of Table 3 shows the spatial correlations between PGW and DIRECT versions of the various moisture budget elements.

Recall from the discussion of Figure 3 in Section 4a, the PGW technique captures the change in mean wintertime precipitation in the DIRECT experiment with reasonable accuracy. The moisture budget decomposition reveals that the thermodynamic and dynamic components of this change are each individually also reproduced quite well

Table 3
Spatial Correlations Between the PGW and DIRECT Patterns Shown in Figures 6, 7, and 10

	Thermodynamic	Dynamic	Residual
DJF ΔP (Figure 6)	0.80	0.74	0.94
JJA ΔP (Figure 7)	0.83	0.73	0.86
ΔP95 (Figure 10)	0.79	0.63	0.74

with PGW, with reasonably high spatial correlations between PGW and DIRECT patterns (Table 3, first row). The mostly positive thermodynamic term over land reflects that almost all of the domain is a region of moisture convergence in the wintertime in the historical climate. Thus, greater moisture levels in a warmer climate produce an amplification of that moisture convergence. Likewise, the particularly large values over mountains, such as the Sierra Nevada and Cascades, reflect the fact that in the historical climate, significant moisture converges during orographic uplift of moist Pacific air masses encountering the continent. In a warmer world, with even moister air masses being uplifted over intense topography, moisture convergence intensifies further. These processes play out in a very similar way in both PGW and DIRECT simulations.

The dynamical term is also similar between DIRECT and PGW. There is a prominent positive contribution to the precipitation increase from dynamics over the Sierra Nevada, California coastal ranges, and Cascades that is likely due to a systematic intensification and shift in the jet stream between the latitudes of Baja California and Oregon (See Figure S3 in Supporting Information S1). A zone of positive values is seen in roughly the same location in the GCM's dynamical component. Note that the effects of such changes may end up being incorporated into PGW experiments through the imposition of the mean circulation change (an example of which can be seen in Figure S3 in Supporting Information S1). The PGW method fails to capture some features in its dynamic component. For example, the positive dynamical term over the Sierra Nevada is greater in DIRECT. And there is dynamical drying over the Pacific Northwest in DIRECT, that is not captured in PGW.

We also briefly discuss the high degree of similarity between the patterns of the PGW and DIRECT residual terms in Figure 6, expressed in their high spatial correlation of 0.94 (Table 3). As noted above, the residual term not only represents changes in the higher-order components of the moisture budget but also includes differences in numerical methods of the moisture budget decomposition and WRF. Since we do not expect PGW to capture the higher-order components of the moisture budget, the similarity between the two residual patterns is strongly reminiscent of structural numerical differences common to the two cases. These numerical differences would likely be where moisture and pressure gradients are largest, that is, over steep topography. Indeed, these are precisely the locations where the same prominent anomalies in the residual term are found in the PGW and DIRECT patterns (e.g., the steep faces of the Sierra Nevada range). Thus, in this case, we do not take the similarity of the two residual patterns to be evidence that PGW is successful in simulating higher-order components of the moisture budget. Rather it is the smallness of this term, together with the successful reproduction of the thermodynamic and dynamic terms, that account for the success of PGW in reproducing WUS DJF precipitation change.

We next show the same analysis for the summertime case (Figure 7, and 2nd row of Table 3). This figure is instructive because it illustrates why there is such a large difference between PGW and DIRECT in summertime ΔP, particularly over the North American monsoon subregion (Figure 3). In this subregion, extra drying is seen in the PGW case in the dynamic term. In addition, the moisture budget residual behaves very differently from the wintertime case. It is very large relative to ΔP, and is different in the DIRECT and PGW cases. This difference is clearly dominating similarities in this term that might arise from common numerical aspects of the decomposition. It is an indicator that the second-order changes ($-\nabla \cdot \overline{\{\Delta q \Delta \mathbf{v}\}}$) in the moisture budget are in fact different between DIRECT and PGW. Thus, in contrast to wintertime, there are low correlations between winds and moisture, preventing a simple decomposition between precipitation changes due to moisture versus winds. Instead, ΔP results from changes due to the synoptic-scale covariance of moisture and winds. We do not expect PGW to be able to represent this component because it applies mean time-invariant deltas to individual variables. Note that this residual term is also large in the GCM, indicating that PGW's failure could be predictable from GCM output. Given the importance of changes in synoptic-scale variability to the changes in mean summertime precipitation, it is reasonable to expect the PGW experiment to similarly fail to capture changes in extreme summertime precipitation in this subregion (Figure 5), which only arises from particular types of synoptic-scale disturbances.

Next, we partition the thermodynamic and dynamic terms further, into mean and eddy components for the DJF case (Equation 3, Figure 8, first row of Table 4). The purpose here is to more precisely isolate those components we do not expect the PGW method to accurately reproduce. As expected, the mean thermodynamic term is similar between the two experiments (spatial correlation 0.88). This term represents the historical PDFs of u and v scaled

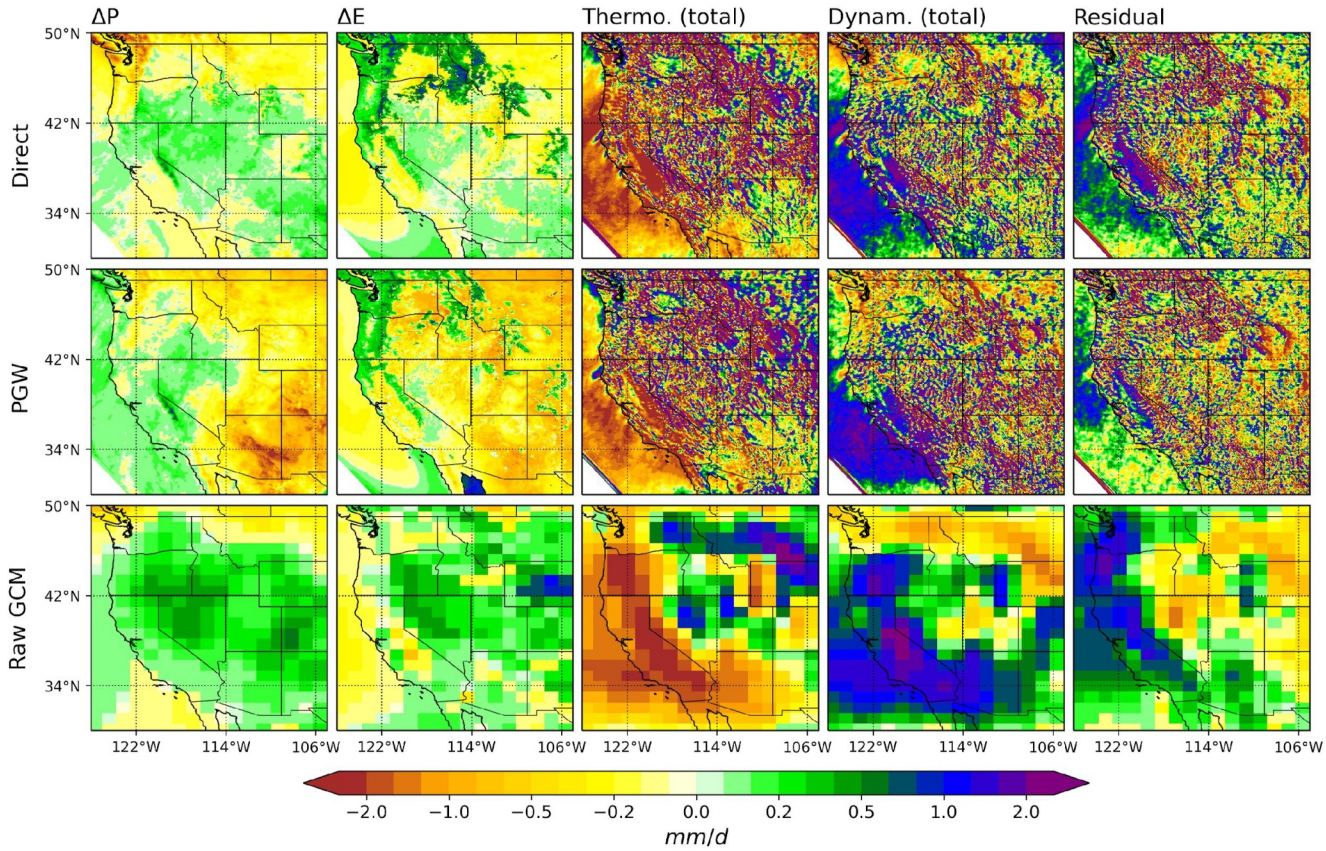

Figure 7. As in Figure 6, but for changes in WUS JJA precipitation. The spatial correlations between the PGW and DIRECT patterns are as follows: 0.83 (thermodynamic), 0.73 (dynamical), and 0.86 (residual), and are also shown in Table 3.

by $\Delta\overline{q}$, which is likely to be very similar in PGW and DIRECT cases because of the relative humidity conservation assumption in the PGW experiment. This term is generally positive, consistent with the WUS, especially the northwestern portion, being a zone of moisture convergence that simply intensifies in a warmer world due to the q increase. The mean dynamic term is also extremely similar (spatial correlation 0.85), although there are greater increases, in particular over California, in DIRECT. It may not be surprising that PGW captures this term well, because it represents the moisture budget change due to the historical q variability distribution and the mean future circulation change ($\Delta\overline{u}$ and $\Delta\overline{v}$). The mean circulation change (i.e., jet stream shift) is the same in DIRECT and PGW. Thus, any differences between the two simulations reflect differences in the historical q variability distribution between the GCM and reanalysis, which are probably small. In fact, since the reanalysis probably has the more realistic q variability distribution, it is possible that the PGW rendition of this component is actually more credible.

We next turn to the thermodynamic eddy term, which represents the contribution of historical synoptic-scale disturbances, but with $\Delta q'$ instead of q' at each time step. In the case of the WUS wintertime, the synoptic-scale disturbances are primarily atmospheric rivers arriving from the North Pacific. The geographical pattern associated with this term in the DIRECT experiment corresponds with the orographic precipitation pattern of extratropical cyclones. It reflects the strong convergence of moisture on the windward sides of mountain ranges along the Pacific coast during these synoptic disturbances, which increases significantly in the future due to elevated moisture levels. We do not necessarily expect the PGW method to perform well in reproducing this term, because it does not explicitly include any information about future specific humidity anomalies in synoptic disturbances in the calculation of boundary conditions. Yet, the PGW method performs surprisingly well in reproducing the pattern of the DIRECT experiment (spatial correlation 0.93). This demonstrates that the PGW experiment actually has similar specific humidity increases in future atmospheric rivers as the DIRECT experiment. This is consistent with Siler et al. (2023), who point out that most of the change in q' under warming is a

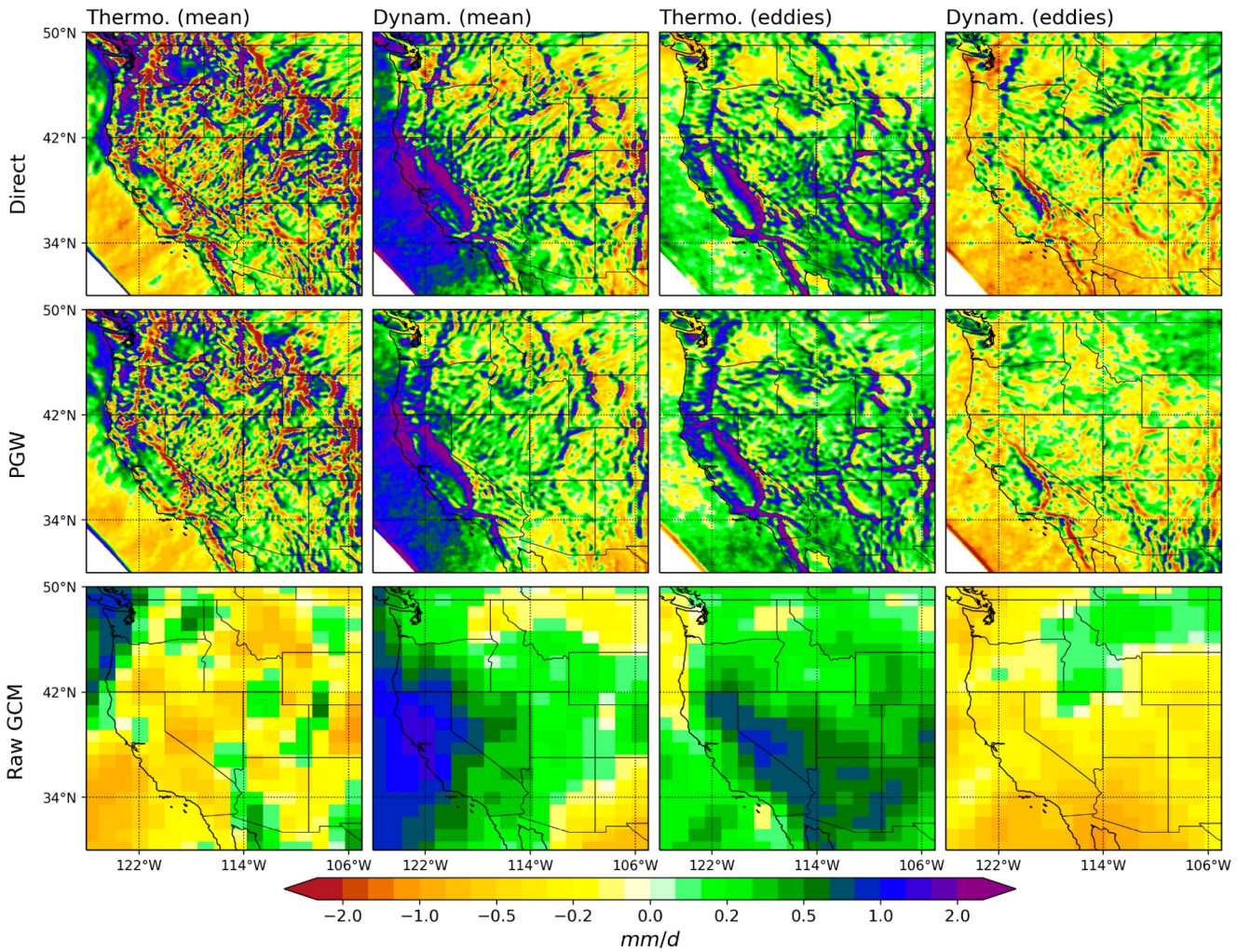

Figure 8. As in Figure 6, except further decomposing the thermodynamic and dynamic components of the DJF mean precipitation change moisture budget into those portions attributable to changes in $\overline{q}$ (thermo mean, 1st column), $\overline{u}$ and $\overline{v}$ (dynam mean, 2nd column), q' (thermo eddies, 3rd column), and u' and v' (dynam eddies, 4th column). The spatial correlations between the PGW and DIRECT patterns are as follows: 0.88 (thermodynamic mean), 0.85 (dynamical mean), 0.93 (thermodynamic eddies), and 0.72 (dynamic eddies), also shown in Table 4.

simple consequence of Clausius-Clapeyron scaling, with little contribution from changes in relative humidity or temperature variance, the two components of the thermodynamic response that PGW neglects. For example, when a historical atmospheric river from reanalysis is simulated in the warmer world of the future PGW experiment, the specific humidity is increased by a much larger amount than would be expected from a change in the climatological mean specific humidity, because of the requirement that relative humidity would be conserved. This confirms that relative humidity conservation is an excellent approximation of atmospheric behavior in the DIRECT experiment. It also underscores the importance of conserving relative humidity, rather than specific humidity, in computation of boundary conditions in the PGW method.

Table 4
Spatial Correlations Between the PGW and DIRECT Patterns Shown in Figures 8 and 9

	Thermo (mean)	Dynam (mean)	Thermo (eddies)	Dynam (eddies)
DJF ΔP (Figure 8)	0.88	0.85	0.93	0.72
JJA ΔP (Figure 9)	0.86	0.83	0.74	0.70

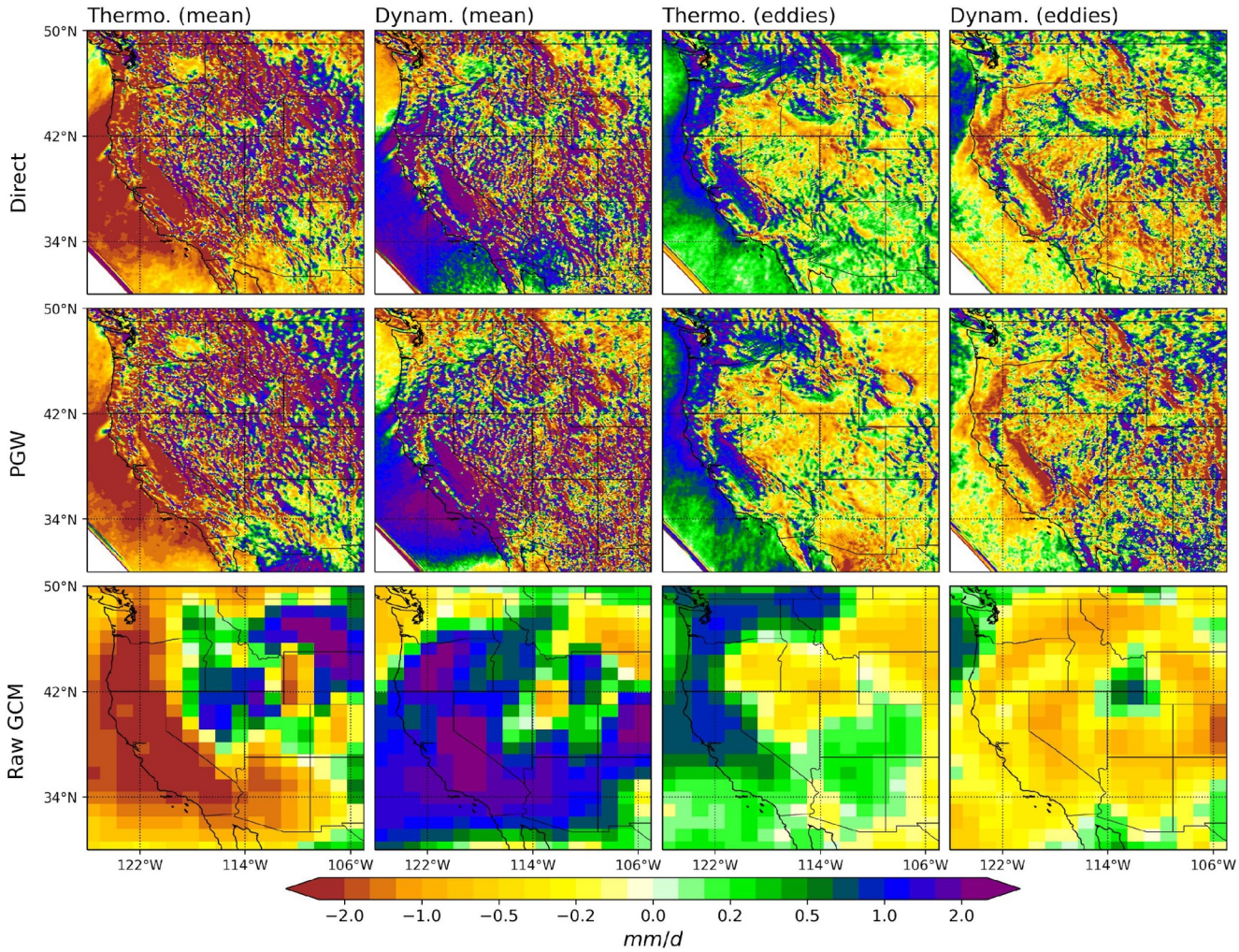

Figure 9. As in Figure 8, but for changes in WUS JJA precipitation. The spatial correlations between the PGW and DIRECT patterns are as follows: 0.86 (thermodynamic mean), 0.83 (dynamical mean), 0.74 (thermodynamic eddies), and 0.70 (dynamic eddies), also shown in Table 4.

Finally, we turn to the dynamic eddy term. This term represents the contribution from synoptic eddies but assuming future circulation changes, that is, $\Delta u'$ and $\Delta v'$, instead of u' and v'. Our baseline hypothesis is that PGW might have difficulty capturing this term because it only includes changes in mean circulation in the calculation of boundary conditions. Indeed the spatial correlation of the two patterns is lowest for this term (0.72), but still PGW appears to have some skill here. In the DIRECT experiment, this term shows wetting over the northeastern part of the domain, drying over the southern part of the domain, and drying over the Pacific Northwest, with strong anomalies associated with mountain ranges also seen. The PGW experiment generally shows a qualitatively similar pattern, albeit with reduced magnitudes. Our interpretation of these results is that the circulation anomalies during synoptic-scale precipitation events are large in magnitude compared to the circulation anomalies when precipitation is not occurring. Thus, the mean circulation change (i.e., what is imposed in the PGW experiment) is an attenuated version of the circulation changes during precipitation events. It is unclear whether this relationship between mean circulation changes and the circulation changes during precipitation events would hold in other regions. In any event, this term makes a small contribution to ΔP (cf. Figure 6), and the assumption of constant Δu and Δv in PGW is fairly inconsequential, at least in wintertime over the western US.

Briefly applying the same analysis to JJA (Figure 9, 2nd row of Table 4), we see that as for winter, all terms exhibit similarities, but with the dynamical eddy term being least similar (spatial correlation 0.70). The thermodynamic eddy term is also associated with a relatively low spatial correlation (0.74). The relatively low correlations arise in part from pronounced differences between the two experiments in the North American

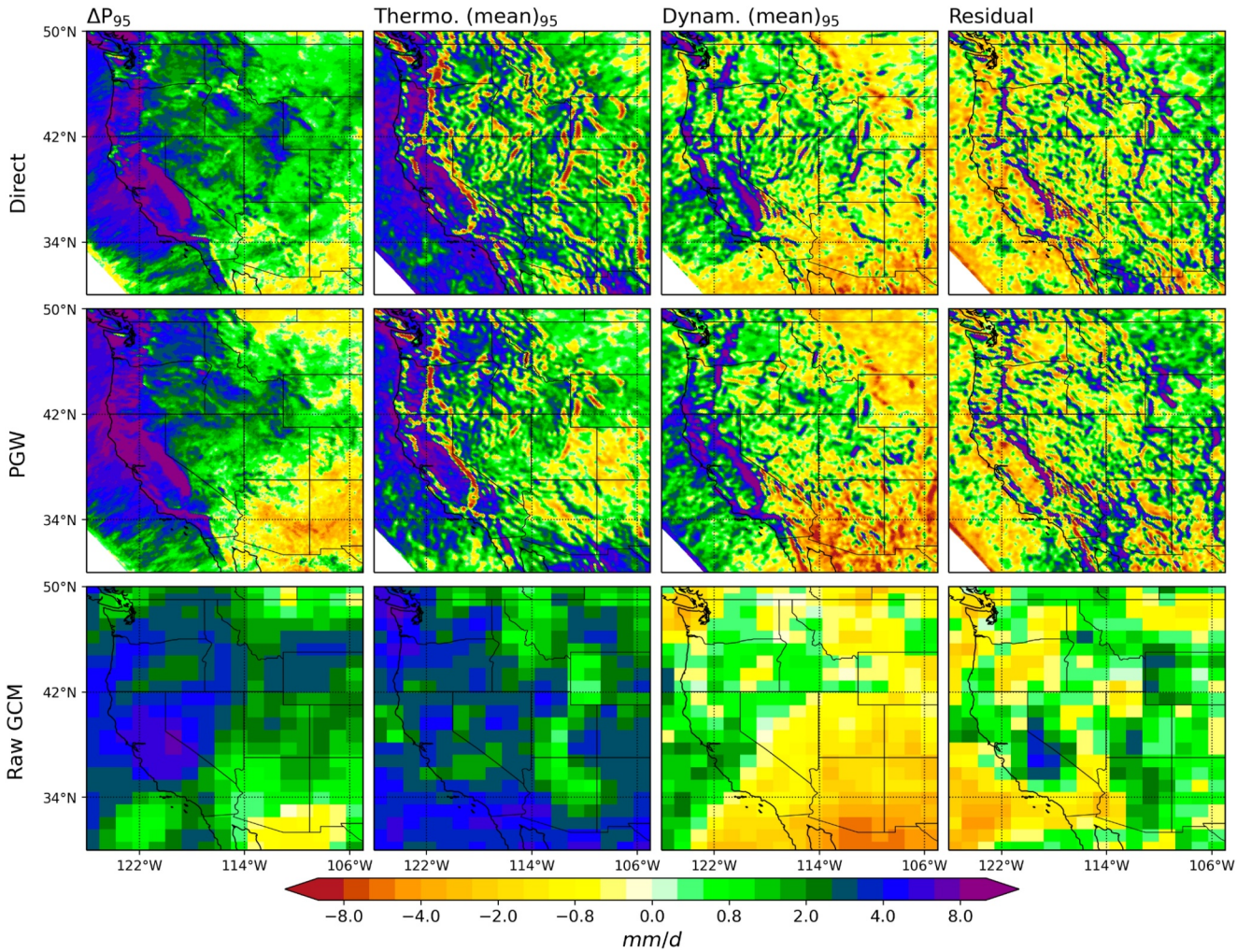

Figure 10. As in Figure 6, but based on the full calendar year and only sampling days when precipitation exceeds the 95th percentile. Based on this subsample, ΔE and the eddy terms are negligible and not plotted. The residual is calculated as $\Delta P - \Delta E - \Delta \overline{P} \cdot \overline{E_q} - \Delta \overline{P} - \overline{E_{\vec{v}}}$. The spatial correlations between the PGW and DIRECT patterns are as follows: 0.79 (thermodynamic), 0.63 (dynamical), and 0.74 residual, also shown in the 3rd row of Table 3.

monsoon region, with more drying in the PGW case. Thus, PGW may be less credible in this subregion in summertime due to its inability to represent changes in synoptic-scale moisture anomalies, and in the PDFs of u and v. Note that this lack of credibility may also stem from an inability to simulate second-order moisture budget terms, as discussed above.

5.3. Decomposition Applied to WUS Changes in Extreme Precipitation

Here, we decompose the moisture budget for the changes to WUS extreme precipitation, ΔP_{95}, in the two experiments and the GCM (Figure 10 and 3rd row of Table 3). This calculation is much simpler for the mean changes for two reasons. First, the evaporation change term is negligible relative to the change in extreme precipitation. Second, the eddy terms within extreme events are also negligible. This reflects that unlike for the mean precipitation changes, Δq, Δu, and Δv can all be assumed to be constant within the subset of extreme events, as found by Norris et al. (2019). Therefore, in this figure, we only plot ΔP_{95} along with the mean thermodynamic and dynamic terms. The residual is what is unaccounted for by these terms, and is relatively small, further confirming that the evaporation and eddy terms are negligible in this context. Here, the mean thermodynamic and dynamic terms represent changes to $\Delta \overline{q}$, $\Delta \overline{u}$, and $\Delta \overline{v}$, where the overbar denotes the average over greater-than-95th-percentile precipitation events.

As discussed previously in Section 4c, PGW captures the change in wet season (DJF) extreme precipitation well. In DIRECT, PGW, and GCM cases, the change in the mean thermodynamic term is the dominant contribution. In the raw GCM, the thermodynamic term is universally positive throughout the domain. PGW is very similar to DIRECT in this term. In both RCM experiments (and unlike in the GCM), the most positive values are overwhelmingly seen in the mountain ranges adjacent to the Pacific (i.e., coastal ranges and Sierra Nevada). This reflects the large increases in extreme orographic precipitation associated with atmospheric rivers, due to an increase in atmospheric moisture in a warmer world. Apparently, the relative humidity conservation requirement in the boundary conditions for the PGW experiment is an excellent approximation for the DIRECT experiment's simulated moisture increase during extreme atmospheric river events. The GCM produces these thermodynamic effects associated with moister atmospheric rivers too, but the associated precipitation increases are unrealistically distributed over poorly resolved topography, highlighting the need for downscaling. A small area of disagreement between the two RCM experiments are the negative values over the interior southwest in PGW that are not seen in DIRECT. This likely reflects disagreements between the GCM and reanalysis in whether in DJF, this subregion is moisture convergent or divergent, conditioned on extremes.

The excellent performance of the PGW experiment in reproducing the thermodynamic component of precipitation extremes may be physically reasonable in light of PGW's relative humidity requirement. However, it is not obvious that PGW should capture the dynamic term well. The climatological mean changes in winds (i.e., what is imposed in PGW) are not necessarily the same as the mean changes in winds when extreme precipitation occurs. And yet, they appear to be similar enough that the PGW method qualitatively reproduces the dynamic term. In both cases, enhanced westerly winds stretching from Baja California northward to much of the rest of the northwestern third of the domain can be inferred from an intensification of orographic precipitation associated with intense atmospheric rivers. (Again, Figure S3 in Supporting Information S1 provides evidence of this change in the GCM.). The mean dynamic term of the moisture budget decomposition of the mean precipitation change in Figure 6 (2nd column, bottom row) shows a qualitatively similar pattern. This is consistent with the idea that the climatological depiction of the GCM's circulation change is a reasonably good predictor of the circulation change during extreme atmospheric river events. Finally, we note the representation of the dynamic term is not perfect in PGW. The dynamic term is negative in much of Washington state in the DIRECT experiment, while it is largely positive in PGW. This largely accounts for the error in the PGW experiment in this subregion noted above in Section 4c. Also, there is a much more negative dynamic term in Arizona and New Mexico in PGW, which unlike in DIRECT, results in a decrease in extreme precipitation.

Overall, the above moisture budget analyses demonstrate that the WUS is a region well suited to performing PGW in winter (when most precipitation falls). In the case of the change in mean precipitation, the only term that is noticeably different between DIRECT and PGW is the dynamical eddies term. But even with this term, PGW is not without skill, and moreover, it makes a small contribution to ΔP. In the case of extreme precipitation, PGW reliably reproduces the intensification seen in the DIRECT experiment, due to similarity with DIRECT in both the thermodynamic and dynamic terms.

6. Discussion and Conclusions

In this study, we analyzed regional end-of-century temperature and precipitation projections using direct dynamical downscaling (DIRECT) and PGW techniques. Our goal was to understand how well the PGW technique can reproduce direct downscaling projections. In the DIRECT case, the boundary conditions are taken directly from GCM historical simulations and future projections. In PGW, by contrast, reanalysis is downscaled to produce the baseline climate, and the future climate is then produced by downscaling reanalysis perturbed by mean climate changes in a GCM. The variability sequencing in the PGW baseline and future experiments is identical. Therefore, errors in PGW could arise, primarily from GCM-projected changes that depend on the phase of weather and climate variability at any given time. At the same time, the PGW method has a more realistic historical climate and performs better for changes in variables that are sensitive to the realism of the mean state. We analyzed DIRECT and PGW simulations over both a WUS and an EU domain. The projections over the two domains were driven by the CESM2 and MPI-ESM-LR GCMs, respectively, downscaled by WRF and COSMO.

Both PGW and DIRECT simulations show very similar warming patterns in both domains, characterized by enhancement over high elevations. This partly reflects snow albedo feedback processes not properly represented in the parent GCMs. In the WUS domain, this effect is enhanced in DIRECT over PGW, due to unrealistically

large snow cover in the historical climate in the WRF downscaling of CESM2, leading to unrealistically high snow loss in the future, hence unrealistically large snow albedo feedback. Increases in extreme temperature are generally similar between the two techniques. An exception here is that PGW produces increases in temperature extremes over southern Europe that are noticeably larger than in the DIRECT case. This likely indicates that in the DIRECT experiment, there are changes in synoptic circulation associated with heat waves not included in the PGW experiment. PGW-generated precipitation projections are also broadly similar to those produced with the DIRECT method. Exceptions include less wintertime wetting in the EU domain in the PGW experiment, and a zone of summertime drying in the North American monsoon region in the WUS summertime, and also in the PGW case. Extreme precipitation projections are highly similar between DIRECT and PGW in the WUS in the winter, but are very different in summer. They are mostly similar in the EU domain in both seasons. Despite the similarities between PGW and DIRECT in mean temperature and precipitation signals, the two techniques produce very different outcomes for snow. In the WUS, the DIRECT experiment produces a much larger future snow loss than PGW. This is because it starts out with a snowpack that is clearly unrealistically large, due to the wet bias of the driving GCM. This illustrates a key advantage of PGW for simulation of variables, like snow, whose behavior under climate change is sensitive to biases in the baseline climate.

A moisture budget decomposition was performed over the WUS to shed light on why PGW performs as well as it does in reproducing mean precipitation changes during what is generally the region's wettest season (DJF). Because of its design, we expect the PGW technique to reproduce the changes associated with the mean change in water vapor (mean thermodynamic component), as well as those associated with the mean change in circulation (mean dynamic component), and confirm that this is the case. We further show that the PGW technique also captures the changes associated with water vapor anomalies during precipitation events (eddy thermodynamic component) with a high degree of fidelity. This success of PGW is likely due to the relative humidity conservation constraint in the boundary conditions imposed in the PGW experiment, underscoring the importance of this aspect of the PGW experimental design. We do not expect PGW to capture the changes associated with the future circulation anomalies (eddy dynamic component), and yet it does, albeit in a muted way. The only way this could happen is if the circulation changes during precipitation events happen to be similar to the mean circulation changes imposed in PGW (i.e., a jet stream shift). Even with PGW's limited success in capturing this term, it is also a small contributor to the overall mean precipitation change. During extreme precipitation events, PGW likewise captures the thermodynamic effect of increasing atmospheric water, indicating that its relative humidity constraint is an excellent approximation of the effect of increasing water vapor in the DIRECT experiment. While the thermodynamic contribution dominates the extreme precipitation change, the dynamical component is not negligible and is somewhat captured in the PGW experiment. Again, we reason that this modest success is due to a qualitative agreement between the circulation changes during precipitation extremes, and the mean circulation changes imposed in the PGW experiment. We also discuss the moisture budget analyses for the WUS in JJA, demonstrating that PGW fails to reproduce the DIRECT results in the North American monsoon subregion because the residual term in the moisture budget analysis is so large and because of large contributions to the change signal from the dynamic eddy term.

The discussion of the moisture budget decomposition leads us to suspect that there may be regions, seasons, and models where the dynamical eddy and residual terms, representing changes in synoptic-scale variability, make a much larger contribution to the mean precipitation change than in the wintertime WUS. In these cases, caution may be warranted in using the PGW method. In the top two panels of Figure 11, the magnitude of the sum of the dynamical eddy and residual terms in the moisture budget analysis is shown globally in the raw CESM2 projections, as a fraction of $\Delta\overline{P}$ in DJF and JJA. The red to black areas of these two panels show where these terms together are a large or dominant contributor to the mean precipitation change. In the wintertime WUS, they are not a large contributor, except over the far southwestern portion of the domain, consistent with the GCM panels of Figures 5 and 6, and the associated downscaled PGW and DIRECT analyses. However, during the summertime WUS, they are dominant. Circling back to the WUS JJA mean precipitation change in Figure 3 and the moisture budget analyses shown in Figures 7 and 9 , we see that this is indeed a season when PGW shows significant disagreement with DIRECT in the changes in the North American monsoon, and where the residual and dynamical eddy terms are large. Based on this evidence, it is clear that in even the driving GCM, there are large synoptic-scale circulation changes when summertime precipitation occurs that are significantly different from the summertime mean circulation changes imposed in the PGW experiment.

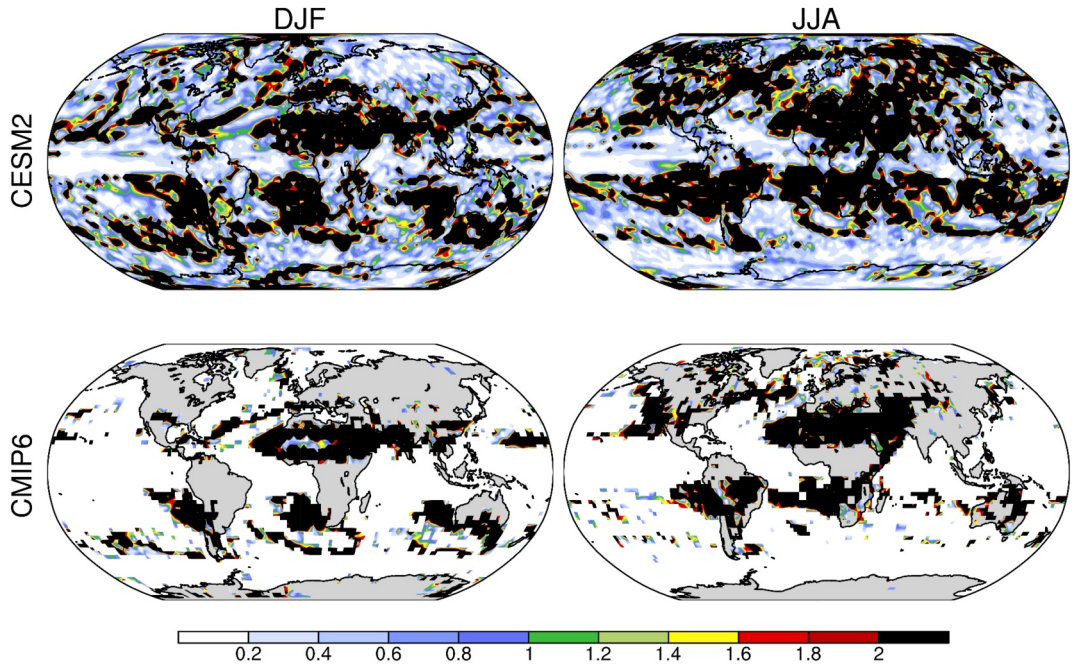

Figure 11. The absolute value of the sum of the dynamical eddy term, $\Delta\overline{P - E_{\psi}}$, and the moisture budget residual as fraction of ΔP, that is, $\left|\frac{R+\Delta\overline{P-E_{\psi}}}{\Delta P}\right|$. The residual is defined as in Figure 6 and encompasses second-order changes plus numerical errors. Evaluated for (left) DJF and (right) JJA; based just on (top) the CESM2 variant used to downscale WRF over the WUS; (bottom) the mean of one SSP3-7.0 projection of 8 CMIP6 GCMs (CanESM5, CESM2, EC-Earth3, FGOALS-g3, IPSL-CM6A-LR, MIROC6, MPI-ESM1-2-LR, and NorESM2-LM). Values are masked where fewer than ¾ of GCMs agree that the given fraction is <0.5. The periods are the same as the WRF simulations (top), but 1984–2014 (historical), and (2071–2100) for the CMIP6 mean. The limited number of GCMs reflects the large data storage and analysis required for moisture budget analysis on 3-D 6-hr data, but comprises most GCMs for which the required data were available.

The fact that the CESM2 moisture budget analysis could have been used to predict PGW's successes and failures suggests that the technique could be applied more broadly to make general predictions about the regional applicability and credibility of PGW. In the bottom two panels of Figure 11, we show the regions and seasons where the sum of the dynamical eddy and residual terms comprises a large fraction of the precipitation change in a CMIP6 ensemble of 8 GCMs. Regions that stand out include the following: the entire desert complex of the Sahara and Middle East, in both seasons, central Western North America during JJA, India and Southeast Asia during DJF, central South America and parts of the Congo during JJA, and Southern Hemisphere Mediterranean climates during DJF (Peru/Chile, South Africa, and Australia). We recommend that similar moisture budget analysis be done in any region prior to applying the PGW technique. But we highlight these regions as places where PGW could give particularly misleading results. Often they are arid or semi-arid regions, with small but potentially still consequential precipitation change signals.

Our study is limited by the fact that we only explicitly compared PGW and DIRECT techniques for two mid-latitude regions. But Figure 11 also illustrates how the GCM moisture budget decomposition can be used to predict when the PGW method might go awry in other places. In both seasons, very small dynamical eddy contributions occur in the deep tropics. In DJF, however, dynamical eddies dominate in southwestern North America, southern Europe/North Africa, and South Asia. Note that Figure 11 portrays CESM2 data rather than the MPI-ESM-LR used in the EU experiments discussed above, and thus, this statement about southern Europe may not apply to MPI-ESM-LR. In JJA, PGW might be problematic for this particular GCM throughout much of the U.S., the Mediterranean, the Middle East, South Africa, Western Australia, and the southern half of South America. Some of these regions are arid or semiarid, with associated precipitation change signals that are small in absolute terms. Thus, for water resource applications, the application of PGW may not produce significant errors. From an ecosystem perspective, however, the relative precipitation changes may be large and consequential, and

Table 5
Summary of the Strengths and Weaknesses of PGW versus Direct Downscaling Techniques in Handling Various Downscaling Challenges

Downscaling challenge	PGW	DIRECT
Historical climate biases	Historical climate is minimally biased	Historical climate inherits GCM biases, with unphysical consequences for some variables, for example, snow
Computational expense	Relatively low (only one baseline simulation needed for ensembles, and usually a time slice approach is taken for future)	Higher, since all GCM ensemble members require their own baseline simulation, and generally, a continuous time series is downscaled
Internal variability effects	Somewhat suppressed, if averaging period or ensemble size for computation of deltas is long enough	The RCM inherits the internal variability of the GCM. Forced climate change signals can be difficult to differentiate from internal variability.
Need for the "storyline" approach	More effective, if the impact of climate change on historical events is to be emphasized	More effective, if "no analog" futures are to be emphasized, since the GCM's future variability is directly downscaled
Need for credible simulation of changes in extremes	Can be credible, provided it is demonstrated that the "dynamical eddy" term is small in the driving GCM data	Can be credible, provided a reliable driving GCM

PGW should be used with caution. This highlights the need to consider the human or natural system for which climate change impacts are being simulated, when choosing between PGW and direct downscaling. In addition to the regions where the dynamical eddy term is large, a PGW approach in which no changes in relative humidity is assumed may also be problematic over those land areas where convection is a dominant precipitation-generating mechanism. In these situations, local evaporation can be an important moisture source for precipitation. Future GCM simulations often show a robust relative humidity reduction in such situations, which in turn affects the convective precipitation signal. We encourage further experimentation in the community with comparisons of PGW and DIRECT techniques, in other regions, with other GCMs, and with other GCM/RCM pairs, to test the assertions made here.

We conclude that we can understand why PGW does surprisingly well in reproducing the climatic changes of direct downscaling in our two test regions. There are also cases where PGW produces more credible projections, for example, when more realistic representations of snow cover in the historical climate lead to more credible snow loss, snow albedo feedback, and warming distributions. This highlights the utility of hybrid techniques where the GCM data is mean-state-bias-corrected prior to downscaling, generating a more realistic historical climate, while preserving the GCM anomaly fields (e.g., Bruyere et al. Rahimi et al., 2024). Our analyses also underscore the need for caution in applying the PGW method. While we do not have reasons to discount its temperature projections, there are regions and seasons where it could yield misleading hydroclimate projections. These regions and seasons can and should be identified through moisture budget analyses of GCM data being considered for RCM boundary conditions. To synthesize this discussion of PGW and DIRECT techniques, as well as the introductory discussion in Section 1, we provide in Table 5 a summary of the strengths and weaknesses of the two techniques insofar as they related to key contemporary downscaling challenges.

Data Availability Statement

All downscaled data, including the full 6-hourly WRF data stream, hourly data for select land surface variables, and a daily post-processed data stream are located in the following open-data bucket on Amazon S3: s3://wrf-cmip6-noversioning/at https://registry.opendata.aws/wrf-cmip6/. These data are completely open and free to the public. We have also developed a technical access and usage document that details these three data tiers, which can be found at this link. As recommended in the document, these data are most easily downloaded when using Amazon Web Services' (AWS') command line interface or with wget.

Acknowledgments
Work performed by PAU is done under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344. We would also like to acknowledge funding supporting for this work from the Department of Energy, Office of Science HyperFACETS project (award DE-SC0016605). N.B. acknowledges the Partnership for Advanced Computing in Europe (PRACE) for awarded access to Piz Daint at ETH Zurich/Swiss National Supercomputing Centre (CSCS, Switzerland).

References

Baldauf, M., Seifert, A., Förstner, J., Majewski, D., Raschendorfer, M., & Reinhardt, T. (2011). Operational convective-scale numerical weather prediction with the COSMO model: Description and sensitivities. *Monthly Weather Review*, *139*(12), 3887–3905. https://doi.org/10.1175/MWR-D-10-05013.1

Bercos-Hickey, E., O'Brien, T. A., Wehner, M. F., Zhang, L., Patricola, C. M., Huang, H., & Risser, M. D. (2022). Anthropogenic contributions to the 2021 Pacific Northwest heatwave. *Geophysical Research Letters*, *49*(23), e2022GL099396. https://doi.org/10.1029/2022GL099396

Berrisford, P., Dee, D. P., Poli, P., Brugge, R., Fielding, M., Fuentes, M., et al. (2011). The ERA-Interim archive Version 2.0.

Boos, W. R., & Pascale, S. (2021). Mechanical forcing of the North American monsoon by orography. *Nature*, *599*(7886), 611–615. https://doi.org/10.1038/s41586-021-03978-2

Brogli, R., Heim, C., Mensch, J., Sørland, S. L., & Schär, C. (2023). The pseudo-global-warming (PGW) approach: Methodology, software package PGW4ERA5 v1.1, validation, and sensitivity analyses. *Geoscientific Model Development*, *16*(3), 907–926. https://doi.org/10.5194/gmd-16-907-2023

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

Deser, C., Knutti, R., Solomon, S., & Phillips, A. S. (2012). Communication of the role of natural variability in future North American climate. *Nature Climate Change*, *2*(11), 775–779. https://doi.org/10.1038/nclimate1562

de Vries, H., Lenderink, G., van der Wiel, K., & van Meijgaard, E. (2022). Quantifying the role of the large-scale circulation on European summer precipitation change. *Climate Dynamics*, *59*(9–10), 2871–2886. https://doi.org/10.1007/s00382-022-06250-z

Doblas-Reyes, F. J., Sorensson, A. A., Almazroui, M., Dosio, A., Gutowski, W. J., Haarsma, R., et al. (2021). *Linking global to regional climate change*. In V. Masson-Delmotte, P. Zhai, A. Pirani, S. L. Connors, C. Pean, S. Berger, et al. (Eds.). Cambridge University Press. Retrieved from https://centaur.reading.ac.uk/99896/

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

Flato, G., Marotzke, J., Abiodun, B., Braconnot, P., Chou, S. C., Collins, W., et al. (2013). Evaluation of climate models. In T. F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Boschung, et al. (Eds.), *Climate change 2013: The physical science basis. Contribution of working group I to the fifth assessment report of the intergovernmental panel on climate change*. Cambridge University Press.

Giorgi, F. (2019). Thirty years of regional climate modeling: Where are we and where are we going next? - Giorgi - 2019. *Journal of Geophysical Research: Atmospheres - Wiley Online Library. Journal of Geophysical Research: Atmospheres*, *14*(11), 5696–5723. https://doi.org/10.1029/2018JD030094

Giorgi, F., & Gutowski, W. J. (2015). Regional dynamical downscaling and the CORDEX initiative. *Annual Review of Environment and Resources*, *40*(1), 467–490. https://doi.org/10.1146/annurev-environ-102014-021217

Giorgi, F., & Mearns, L. O. (1999). Introduction to special section: Regional climate modeling revisited. *Journal of Geophysical Research*, *104*(D6), 6335–6352. https://doi.org/10.1029/98JD02072

Gutowski, J., Arritt, R. W., Hall, A. D., Leung, L.-Y., Mearns, L. O., & Pryor, S. C. (2021). *A hierarchical evaluation framework for assessing climate simulations relevant to the energy-water-land nexus (Final Report) (No. DOE-IOWASTATE-SC0016438)*. Iowa State University. https://doi.org/10.2172/1828286

Hara, M., Yoshikane, T., Kawase, H., & Kimura, F. (2008). Estimation of the impact of global warming on snow depth in Japan by the pseudo-global-warming method. *Hydrological Research Letters*, *2*, 61–64. https://doi.org/10.3178/hrl.2.61

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

Ikeda, K., Rasmussen, R., Liu, C., Newman, A., Chen, F., Barlage, M., et al. (2021). Snowfall and snowpack in the western U.S. As captured by convection permitting climate simulations: Current climate and pseudo global warming future climate. *Climate Dynamics*, *57*(7), 2191–2215. https://doi.org/10.1007/s00382-021-05805-w

Kawase, H., Yoshikane, T., Hara, M., Kimura, F., Yasunari, T., Ailikun, B., et al. (2009). Intermodel variability of future changes in the Baiu rainband estimated by the pseudo global warming downscaling method. *Journal of Geophysical Research*, *114*(D24), D24110. https://doi.org/10.1029/2009JD011803

Kröner, N., Kotlarski, S., Fischer, E., Lüthi, D., Zubler, E., & Schär, C. (2017). Separating climate change signals into thermodynamic, lapse-rate and circulation effects: Theory and application to the European summer climate. *Climate Dynamics*, *48*(9–10), 3425–3440. https://doi.org/10.1007/s00382-016-3276-3

Lackmann, G. M. (2015). Hurricane sandy before 1900 and after 2100. *Bulletin of the American Meteorological Society*, *96*(4), 547–560. https://doi.org/10.1175/BAMS-D-14-00123.1

Lafon, T., Dadson, S., Buys, G., & Prudhomme, C. (2013). Bias correction of daily precipitation simulated by a regional climate model: A comparison of methods. *International Journal of Climatology*, *33*(6), 1367–1381. https://doi.org/10.1002/joc.3518

Lehner, F., Deser, C., Maher, N., Marotzke, J., Fischer, E. M., Brunner, L., et al. (2020). Partitioning climate projection uncertainty with multiple large ensembles and CMIP5/6. *Earth System Dynamics*, *11*(2), 491–508. https://doi.org/10.5194/esd-11-491-2020

Lenderink, G., de Vries, H., van Meijgaard, E., van der Wiel, K., & Selten, F. (2023). A perfect model study on the reliability of the added small-scale information in regional climate change projections. *Climate Dynamics*, *60*(9–10), 2563–2579. https://doi.org/10.1007/s00382-022-06451-6

Li, J., Qian, Y., Leung, L. R., Chen, X., Yang, Z., & Feng, Z. (2023). Potential weakening of the June 2012 North American derecho under future warming conditions. *Journal of Geophysical Research: Atmospheres*, *128*(2), e2022JD037494. https://doi.org/10.1029/2022JD037494

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

Liu, C., Ikeda, K., Thompson, G., Rasmussen, R., & Dudhia, J. (2011). High-resolution simulations of wintertime precipitation in the Colorado headwaters region: Sensitivity to physics parameterizations. *Monthly Weather Review*, *139*(11), 3533–3553. https://doi.org/10.1175/MWR-D-11-00009.1

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

Norris, J., Chen, G., & Neelin, J. D. (2019). Thermodynamic versus dynamic controls on extreme precipitation in a warming climate from the community Earth system model large ensemble. *Journal of Climate*, *32*(4), 1025–1045. https://doi.org/10.1175/JCLI-D-18-0302.1

Patricola, C. M., & Wehner, M. F. (2018). Anthropogenic influences on major tropical cyclone events. *Nature*, *563*(7731), 339–346. https://doi.org/10.1038/s41586-018-0673-2

Prein, A. F., Rasmussen, R. M., Ikeda, K., Liu, C., Clark, M. P., & Holland, G. J. (2017). The future intensification of hourly precipitation extremes. *Nature Climate Change*, *7*(1), 48–52. https://doi.org/10.1038/nclimate3168

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*(9), e2023GL106264. https://doi.org/10.1029/2023GL106264

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., Liu, C., Ikeda, K., Gochis, D., Yates, D., Chen, F., et al. (2011). High-resolution coupled climate runoff simulations of seasonal snowfall over Colorado: A process study of current and warmer climate. *Journal of Climate*, *24*(12), 3015–3048. https://doi.org/10.1175/2010JCLI3985.1

Reed, K. A., Stansfield, A. M., Wehner, M. F., & Zarzycki, C. M. (2020). Forecasted attribution of the human influence on Hurricane Florence. *Science Advances*, *6*(1), eaaw9253. https://doi.org/10.1126/sciadv.aaw9253

Reed, K. A., Wehner, M. F., & Zarzycki, C. M. (2022). Attribution of 2020 hurricane season extreme rainfall to human-induced climate change. *Nature Communications*, *13*(1), 1905. https://doi.org/10.1038/s41467-022-29379-1

Risser, M. D., Rahimi, S., Goldenson, N., Hall, A., Lebo, Z. J., & Feldman, D. R. (2024). Is bias correction in dynamical downscaling defensible? *Geophysical Research Letters*, *51*(10), e2023GL105979. https://doi.org/10.1029/2023GL105979

Rockel, B., Will, A., & Hense, A. (2008). The regional climate model COSMO-CLM (CCLM). *Meteorologische Zeitschrift - METEOROL Z*, *17*(4), 347–348. https://doi.org/10.1127/0941-2948/2008/0309

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

Scheff, J., & Frierson, D. M. W. (2012). Robust future precipitation declines in CMIP5 largely reflect the poleward expansion of model subtropical dry zones. *Geophysical Research Letters*, *39*(18). https://doi.org/10.1029/2012GL052910

Seager, R., & Henderson, N. (2013). Diagnostic computation of moisture budgets in the ERA-interim reanalysis with reference to analysis of CMIP-archived atmospheric model data. *Journal of Climate*, *26*(20), 7876–7901. https://doi.org/10.1175/JCLI-D-13-00018.1

Seager, R., Naik, N., & Vecchi, G. A. (2010). Thermodynamic and dynamic mechanisms for large-scale changes in the hydrological cycle in response to global warming. *Journal of Climate*, *23*(17), 4651–4668. https://doi.org/10.1175/2010JCLI3655.1

Shepherd, T. G. (2019). Storyline approach to the construction of regional climate change information. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *475*(2225), 20190013. https://doi.org/10.1098/rspa.2019.0013

Siler, N., Bonan, D. B., & Donohoe, A. (2023). Diagnosing mechanisms of hydrologic change under global warming in the CESM1 Large Ensemble. *Journal of Climate*, *1*(aop), 8243–8257. https://doi.org/10.1175/JCLI-D-23-0086.1

Siler, N., Roe, G. H., Armour, K. C., & Feldl, N. (2019). Revisiting the surface-energy-flux perspective on the sensitivity of global precipitation to climate change. *Climate Dynamics*, *52*(7), 3983–3995. https://doi.org/10.1007/s00382-018-4359-0

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

Skamarock, W. C., & Klemp, J. B. (2008). A time-split nonhydrostatic atmospheric model for weather research and forecasting applications. *Journal of Computational Physics*, *227*(7), 3465–3485. https://doi.org/10.1016/j.jcp.2007.01.037

Stevens, B., Giorgetta, M., Esch, M., Mauritsen, T., Crueger, T., Rast, S., et al. (2013). Atmospheric component of the MPI-M Earth system model: ECHAM6. *Journal of Advances in Modeling Earth Systems*, *5*(2), 146–172. https://doi.org/10.1002/jame.20015

Teutschbein, C., & Seibert, J. (2012). Bias correction of regional climate model simulations for hydrological climate-change impact studies: Review and evaluation of different methods. *Journal of Hydrology*, *456–457*, 12–29. https://doi.org/10.1016/j.jhydrol.2012.05.052

Ullrich, P. A., Xu, Z., Rhoades, A. m., Dettinger, M. d., Mount, J. f., Jones, A. d., & Vahmani, P. (2018). California's drought of the future: A midcentury recreation of the exceptional conditions of 2012–2017. *Earth's Future*, *6*(11), 1568–1587. https://doi.org/10.1029/2018EF001007

Walton, D. B., Hall, A., Berg, N., Schwartz, M., & Sun, F. (2017). Incorporating snow albedo feedback into downscaled temperature and snow cover projections for California's Sierra Nevada. *Journal of Climate*, *30*(4), 1417–1438. https://doi.org/10.1175/JCLI-D-16-0168.1

Wehner, M. F., Zarzycki, C., & Patricola, C. (2019). Estimating the human influence on tropical cyclone intensity as the climate changes. In J. M. Collins & K. Walsh (Eds.), *Hurricane risk* (pp. 235–260). Springer International Publishing. https://doi.org/10.1007/978-3-030-02402-4_12

Xue, Z., & Ullrich, P. (2021). A comprehensive intermediate-term drought evaluation system and evaluation of climate data products over the conterminous United States. *Journal of Hydrometeorology*, *22*(9), 2311–2337. https://doi.org/10.1175/JHM-D-20-0314.1

Xue, Z., Ullrich, P., & Leung, L.-Y. R. (2023). Sensitivity of the pseudo-global warming method under flood conditions: A case study from the northeastern US. *Hydrology and Earth System Sciences*, *27*(9), 1909–1927. https://doi.org/10.5194/hess-27-1909-2023