AMERICAN METEOROLOGICAL SOCIETY

Journal of Climate

EARLY ONLINE RELEASE

The DOI for this manuscript is doi: 10.1175/JCLI-D-15-0439.1

The influence of climate model biases on projections of aridity and drought

Darren L. Ficklin[1], John T. Abatzoglou[2], Scott M. Robeson[1], Anna Dufficy[3]

[1]701. E. Kirkwood Ave., Department of Geography, Indiana University, Bloomington, IN 47405

[2]Department of Geography, University of Idaho, Moscow, ID 83844-3021

[3]Geological Sciences, Indiana University, Bloomington, IN 47405

*corresponding author. Address: [1]Department of Geography, Indiana University, 701. E. Kirkwood Ave., Bloomington, IN 47405. Phone: 812-856-5047. Email: dficklin@indiana.edu

Abstract

Global Climate Models (GCMs) have biases when simulating historical climate conditions, which in turn have implications for estimating the hydrological impacts of climate change. This study examines the differences in projected changes of aridity (defined as precipitation over potential evapotranspiration, or P/PET)) and the Palmer Drought Severity Index (PDSI) between raw and bias-corrected GCM output for the continental United States (CONUS). For historical simulations (1950-1979) the raw GCM ensemble median has a positive precipitation bias (+24%) and negative PET bias (-7%) compared to the bias-corrected output when averaged over CONUS with the most acute biases over the interior western United States. While both raw and bias-corrected GCM ensembles project more aridity (lower P/PET) for CONUS in the late 21st century (2070-2099), relative enhancements in aridity were found for bias corrected data compared to the raw GCM ensemble owing to positive precipitation and negative PET biases in the raw GCM ensemble. However, the bias-corrected GCM ensemble projects less acute decreases in summer PDSI for the southwest United States compared to the raw GCM ensemble (1 to 2 PDSI units higher), stemming from biases in precipitation amount and seasonality in the raw GCM ensemble. Compared to the raw GCM ensemble, bias-corrected GCM inputs not only correct for systematic errors but also can produce high-resolution projections that are useful for impact analyses. Therefore, changes in hydroclimate metrics often appear considerably different in bias-corrected output compared to raw GCM output.

1. Introduction

With current and impending increases in air temperature and changes in precipitation around the world, water scarcity and security is becoming one of the most important topics of the 21st century. Issues of aridity and drought – which can directly cause water scarcity and food insecurity – depend on whether there is a water surplus or deficit at a given location, which is determined by the balance between precipitation and potential evapotranspiration (PET). With changes in climate, reduced precipitation amounts can induce short-term droughts that then are more readily reinforced by higher PET and temperature (Diffenbaugh et al. 2015). In addition, systematic changes in precipitation and PET at a given location may be part of a longer term shift to a more arid climate (Cook et al. 2014). Globally, about one-third of land area is currently designated as arid, with an expansion projected as a result of anthropogenic climate change (Feng and Fu 2013; Huang et al., 2015; Scheff and Frierson 2014, 2015).

A long history of droughts has been documented across the United States, both in the instrumental and paleoclimatic records (e.g., Cook et al. 2004; Herweijer and Seager 2008; Woodhouse and Overpeck 1998). However, numerous widespread severe droughts have been observed across the United States since the late 20th century, including a persistent drought across the southwestern United States (Woodhouse et al. 2010), intense but short-lived droughts across Texas and the southern Great Plains (Seager and Hoerling 2014) and the midwestern United States (Hoerling et al. 2013), and one of the most acute multi-year droughts in California in over a millennia (Griffin and Anchukaitis 2014). Additionally, drought occurrences for a large portion of the United States have been found to be increasing since the late 20th century (Ficklin et al. 2015). It is likely that such changes are a combination of anthropogenic forcing and natural variability, with a number of recent studies suggesting that enhanced greenhouse gas concentrations are

increasing the likelihood of such droughts, particularly ones associated with anomalously high PET (e.g., Diffenbaugh et al. 2015; Rupp et al. 2012). It is therefore imperative to understand changes in drought intensity, duration, and frequency – as well as overall shifts in aridity – with projected changes in climate.

Primarily through increases in temperature and vapor pressure deficit, PET is projected to increase (Ficklin et al. 2015; Sheffield et al. 2012; Vicente-Serrano et al. 2014). Regional changes in precipitation and its seasonality may offset, partially offset, or compound increases in PET. Several recent studies have examined such changes in precipitation and PET in the context of drought (Ault et al. 2014; Cook et al. 2015; Dai 2013; Sheffield and Wood 2008) and aridity (Feng and Fu 2013; Girvetz and Zganjar 2014; Scheff and Frierson 2014, 2015; Seager and Vecchi 2010; Seager et al. 2007). There is a general concurrence among these studies showing an overall increase in aridity and drought intensities under projected changes in climate. However, it is important to note that Johnson and Sharma (2010) show that PET could decrease with climate change due to decreases in wind speed.

A majority of prior analyses of changes in drought and aridity have used direct (or raw) output from Global Climate Models (GCMs) or coarse resolution GCM bias-corrected estimates, which have documented biases in simulating regional climate (e.g., Maurer and Hidalgo 2008; Polade et al. 2013; Rupp et al. 2013). Such biases include differences in the magnitude, seasonality, and variability of model output relative to observations. For example, Scheff and Frierson (2015) showed substantial intermodel biases in precipitation and PET across large geographic regions of the globe with individual model biases in precipitation typically opposing PET biases in sign and magnitude. Biases in GCM simulations run for contemporary conditions should translate into similar biases in projections of drought or aridity under climate change scenarios. At the local

scale, such biases can substantially alter the magnitude and sometimes the sign of change for annual moisture related variables (e.g., Christensen et al. 2008; Hagemann et al. 2011; Johnson and Sharma 2015; Maurer and Pierce 2014). Downscaling and bias-correction methods have been used to correct some GCM biases and translate the coarser resolution model output to localized scales (e.g., Fowler et al. 2007). The abundance of approaches to downscale GCM data confounds matters as dynamical downscaling typically allows for the persistence of systemic biases (e.g., Glotter et al. 2014) while statistical downscaling approaches are unable to account for sub-regional changes in climate that arise through physical processes at sub-GCM grid scales. Despite the limitations of statistical downscaling approaches, they remain the preferred approach for climate impact assessment using ensembles of GCMs.

There are still differences of opinion regarding whether direct or bias-corrected model output should be used in climate change assessments. While the usage of direct GCM or RCM output ensures internal consistencies across variables, substantial biases in resultant variables may render direct output unrealistic and ultimately unsuitable for subsequent modeling. For example, GCMs fail to resolve topographic features and thus may poorly model precipitation across complex terrain (Sheffield et al., 2013b). Likewise, model biases in precipitation seasonality can be problematic for realistic projections of regional water resources. Collectively, failure to account for GCM or RCM biases may result in gross errors for both historical and future climate simulations (e.g., Muerth et al., 2013). By contrast, statistical methods inherent in empirical bias correction often lack a physical basis (e.g., Ehret et al., 2012) and may obscure uncertainties inherent with downscaling itself. However, while the climate modeling community and computing resources continue to advance toward higher resolution GCMs, statistical bias correction currently is required to make climate projections usable for impact assessment.

To our knowledge, the effect of precipitation and PET bias from raw GCMs on drought has yet to be explored. Recent work by Huang et al., (2015) found that bias-correcting GCMs from the fifth Coupled Model Inter-Comparison Project (CMIP5) led to global increases in the aridity index (precipitation/PET) compared to using direct GCM output. Johnson and Sharma (2015) and Nasrollahi et al. (2015) assessed the performance of GCMs at simulating meteorological droughts against historical observations/downscaled GCM datasets, finding that raw GCM output tend to overestimate drought conditions based on the Standard Precipitation Index. However, these studies concentrated only on precipitation, neglecting the role of PET. In the interest of assessing and interpreting the impact of bias-corrected GCM output on projected changes in drought and aridity, we focus on the continental United States where a large number of weather stations and high-resolution gridded datasets provide the basis for downscaling and bias-correcting raw GCM output to local or regional scales. We concentrate on how bias-correcting raw GCM output can modify projections of precipitation and PET and subsequent changes in aridity and drought occurrences. Specifically, our research questions are:

[1] How do projected changes in precipitation and PET differ between raw and bias-corrected GCM output?

[2] How does the bias in [1] affect the precipitation-PET imbalance?

[3] How does the bias in [1] affect projections of aridity (precipitation/PET) and drought?

We expect the results to this study to be useful for researchers that have used or intend to use raw GCM output for aridity or drought projections. Additionally, this work also has important implications for water resource managers and planners.

2. Methods and Materials

2.1 Raw and bias-corrected/downscaled Global Climate Model data

Data from 17 CMIP5 GCMs that had daily output of 2-m maximum and minimum temperature, specific humidity, 10-m wind velocity, downward shortwave radiation and precipitation at the surface were acquired for both historical forcing experiments (1950-2005) and future forcing experiments (2006-2099) using Representative Concentration Pathway 8.5 (Table 1). Raw GCM output was aggregated to a common 2°×2° degree resolution grid using bilinear interpolation. Bias correction and statistical downscaling of these variables for all models was performed using the Multivariate Adaptive Constructed Analogs (MACA) approach (Abatzoglou and Brown 2012) to a reference, observed 4-km resolution surface meteorological dataset (Abatzoglou 2013). The observed dataset of Abatzoglou (2013) incorporated 4-km monthly output from PRISM (Parameterized Regression on Independent Slopes Model, Daly et al., 2008) with the higher temporal resolution output from the North American Land Data Assimilation System - Phase (NLDAS2). Variables of 10-m wind and solar radiation were derived exclusively from NLDAS2/North American Regional Reanalysis.

Raw daily GCM output was bias corrected prior to downscaling and a joint-bias correction was used for temperature and precipitation outputs (Li et al., 2014). MACA uses a joint constructed analogs approach for downscaling daily temperature (maximum and minimum air temperatures), wind velocity (u and v components) and regular constructed analogs for precipitation, solar radiation and specific humidity. A total of 30 analogs spanning the continental United States were used. To account for novel analogs, the residual (GCM target minus constructed analog) was bi-linearly interpolated to the final pattern. Finally, a secondary bias correction was applied to the end product to ensure adherence with the training data. We note that there are several statistical downscaling methods and uncertainty arising from the choice of downscaling can color results (e.g., Fowler et al., 2007). Likewise, differences in observed data that climate model is trained to

exist, although are often of much lower magnitude than biases between GCM output and observations (e.g., Rupp et al., 2013).

The bias correction procedures used in MACA allow for projected changes in the distribution of a given variable (e.g., precipitation) as simulated by the GCM at daily timescales to be entrained into the downscaled products. For example, projected changes in the 99th percentile distribution of precipitation in the raw GCM output are mirrored in the downscaled output. The MACA approach was further adjusted to ensure the preservation of trends from raw GCM output to downscaled fields at monthly timescales (Pierce et al., 2015). Daily output from both raw GCM and downscaled data were temporally aggregated to monthly timescales for subsequent analysis. The bias-corrected/downscaled GCM data can be obtained at: http://maca.northwestknowledge.net/

2.2 Potential evapotranspiration

Potential evapotranspiration for this work was estimated using the United Nations Food and Agriculture Organization Penman-Monteith method (Allen et al. 1998). Net radiation is estimated using the approach of Allen et al. (1998), whereby both downwelling and upwelling components of shortwave and longwave radiation are incorporated. Albedo from Barkstrom (1984) was used and the surface resistance was not modified from the reference crop. Full documentation and code of this procedure can be found in Ficklin et al. (2015).

Although approaches have been adopted to estimate the reduction in water demand for reference crops with elevated CO_2 (Roderick et al., 2015), the magnitude of changes in water use efficiency are variable across biomes and hence this factor was not included in this work. It is important to note that incorporating the response of vegetation to elevated CO_2 (or removing the assumption of fixed surface resistance) will result in increased vegetative water use efficiency,

resulting in changes in evapotranspiration, soil moisture, and surface water runoff (Ficklin et al., 2009; Donohue et al., 2013; Roderick et al., 2015). This study, however, examines how biases in raw GCM output, when compared to bias-corrected GCM output, will affect future projections of aridity and drought. While GCMs are able to incorporate the effects of CO_2 on precipitation and ET via increased water-use efficiency of vegetation (and therefore soil moisture), the degree to which vegetation is modeled in GCMs varies and, as shown here (and elsewhere), often uses biased values of precipitation and temperature to determine the impacts of the vegetation response.

2.3 Definition of Aridity and Drought

Two metrics of aridity and drought were used in subsequent analyses: a measure of short-term soil moisture imbalance -- Palmer Drought Severity Index (PDSI; (Palmer 1965)) -- and a measure of long-term soil moisture imbalance -- the Aridity Index (AI). PDSI is a drought index that characterizes the cumulative departure of local mean soil moisture conditions based on a simplified water balance calculation. Soil available water capacity used in the PDSI calculations was extracted from the State Soil Geographic Database (STATSGO) for every downscaled grid node. Although PDSI has several limitations (e.g., Alley 1984; Karl and Knight 1985; Werick et al. 1994), it is commonly used throughout the world to quantify observed drought and drought projections (e.g., Dai 2011). The PDSI calibration time period is the historical forcing period of 1950-2005 for each GCM and all deviations are compared to this time periods within the internal PDSI calculations.

To assess long-term soil moisture imbalances, we used the aridity index (AI) defined as the ratio of precipitation and PET (AI = P/PET). AI is widely used (Girvetz and Zganjar 2014; Hulme 1996) and is a useful parameter in terrestrial biome classifications and climatic suitability

for agriculture. Lower values represent more arid moisture regimes, while larger values represent more humid moisture regimes.

2.4 Statistical analyses

For this work, we relate the raw GCM ensemble to the bias-corrected/downscaled GCM ensemble because the bias-corrected/downscaled GCM ensemble has been trained against an observational dataset. The bias-corrected/downscaled GCM data (precipitation, inputs for PET calculation, AI, and PDSI) were upscaled to the same resolution as the raw GCM data by taking the average of all downscaled cells within a raw GCM grid element (2°×2°) and is henceforth referred to as "bias-corrected data." Statistical analyses were performed on the historical (1950-1979) and late-21st century (2070-2099) time periods for the GCM ensemble median (e.g., median of all annual AI values for 2070-2099). The comparison between the observed climate and bias-corrected climate should by definition be very similar when using the quantile bias correction approach (Abatzoglou and Brown, 2012).

To assess the variability of GCM projections around the GCM ensemble median, the interquartile range (IQR) of the GCM ensemble was calculated. Deficit months are defined as months where monthly PET exceeds monthly precipitation. The deficit months for each year are then summed. Raw and bias-corrected PDSI values are only presented for the late-21st century time period because the PDSI values are relative to the raw and bias-corrected historical time period, respectively, and thus PDSI for the historical time period would, on average, be near zero. Additionally, for consistency and comparison with other studies (e.g., Cook et al., 2015; Williams et al., 2015), the average PDSI during June, July, and August (JJA) is used in our analysis. It is important to note the biases for grid elements on the coast that contain large parts of the ocean may induce large biases from averaging a smaller number of bias-corrected and downscaled grid points.

2.5 Evaluation of GCM biases

Numerous approaches exist for evaluating biases in GCM simulation of regional climate. We focus on two aspects of biases that impact resulting fields of precipitation and PET: (i) biases in annual precipitation, annual PET, and the variables that contribute to annual PET and (ii) biases in the seasonality of precipitation. In both cases we only consider biases from the climatology and not those in interannual variability (although it is important to note that unrepresentative interannual variability is another potential bias in AI and PDSI). Raw GCM biases in annual and seasonal (e.g., DJF) GCM outputs for the historical time period (1950-1979) were compared to those from bias-corrected GCM data. We evaluate biases in spatially-weighted absolute and relative (percent changes) depending on the climate output. Biases are presented as the bias of the GCM raw or bias-corrected ensemble median unless otherwise noted.

3. Results

3.1 Differences in precipitation and potential evapotranspiration between raw and bias-corrected GCMs

Annual mean precipitation from the raw and bias-corrected GCM outputs are shown in Figure 1 for the historical and late 21st century time periods. Additionally, percent differences between the raw GCM and bias-corrected GCM ensembles for precipitation are shown for both time periods as well as the differences in precipitation between the end of the 21st century and historical time period for raw and bias-corrected GCM ensembles. Spatially-averaged over all grid cells of the continental United States, the raw GCM ensemble shows a +24% and +22% percent bias from the bias-corrected data for the historical and projected time periods, respectively (i.e., there is an overall "wet bias" in the median of the raw GCM ensemble). The wet bias in the raw GCM output is particularly acute across much of the interior western United States with numerous grid cells receiving over twice the precipitation as bias-corrected output. The northeastern United

States also shows wet biases, though not as extreme as the western United States (~15%). Less than 40% of grid cells across the continental United States show a dry bias. These dry biases are mostly concentrated in the coastal region of the western United States and the southern United States, centered on Arkansas and the Gulf Coast. Finally, around a third of all grid cells show minor bias (+/- 10%) in precipitation.

Projected changes in annual precipitation adhere to the same general latitudinal dipole pattern when comparing raw GCM output to bias-corrected GCM output, with increases in precipitation in the north and decreased precipitation in the south (Figure 1). Averaged over the continental United States, the bias-corrected GCM ensemble projections result in a slightly larger mean increase in annual precipitation (2% increase) than the raw GCM ensemble projections (0.75% increase). This difference between projected and historical precipitation is most apparent for the western United States, where the bias-corrected GCM ensemble leads to large increases in precipitation for the northern half of the western United States (~15-20% average increase compared to the historical time period), as compared to the raw GCM ensemble (5-15%) that shows a dipole pattern with increases north of 40N and decreases south of 35N. Both the raw and bias-corrected GCM ensembles project decreases in precipitation for the southern United States equator-ward of 35°N (southern California to Florida). Conversely, both the raw and bias-corrected GCM ensemble projections show comparable increases in precipitation of approximately 5 – 20% for the Midwest and northeastern United States. Examining the difference between the raw and bias-corrected GCM ensembles (lower-right panel in Figure 1) shows that the raw GCM projections lead to an overall drier western United States than does the bias-corrected GCMs by up to 15%, while the raw GCMs project an increase in precipitation from the central Midwest to Gulf Coast region.

The seasonal distribution of precipitation is an important element that determines the hydroclimatology of a region. Climate projections show distinct seasonal patterns of precipitation change across North America (Sheffield et al. 2013a; Sheffield et al. 2013b). Figure 2 shows the percent of annual precipitation for each season (DJF = winter; MAM = spring; JJA = summer; SON = fall) during 1950-1979 for the raw and bias-corrected GCM ensemble medians, as well as the difference between the two ensembles. Although summaries of changes in precipitation can be presented in absolute numbers, biases in precipitation magnitude often necessitate examining changes in precipitation as a percent of historical values.

Several important seasonal differences can be found in the raw GCM ensemble – bias-corrected GCM ensemble for the historical time period. The raw GCM ensemble overestimates the fraction of annual precipitation occurring in winter for a large portion of the United States and underestimates the fraction of annual precipitation occurring in summer in many of these same regions. The same overestimation and underestimation occurs for the spring and fall seasons. The winter overestimation can be found for much of the central United States, a portion of the western United States (Nevada and Utah), southern Texas, and southern Florida where, on average, 10-20% more of the annual precipitation occurs during winter compared to the bias-corrected GCM ensemble. The overestimation (~5 – 15% compared to the bias-corrected GCM ensemble) of the percent of annual precipitation in spring shifts eastward into New England and the Great Lakes region as well as across the southwestern United States. For the summer and fall seasons, the raw GCM ensemble underestimates precipitation compared to the bias-corrected GCM ensemble by approximately 3% overall, but with large spatial variation. This is especially the case for the central United States, stretching from Texas to North Dakota as well as along the Gulf Coast from Florida to Texas, during the summer, where the percent of annual precipitation during JJA is 10-25% less

than the bias-corrected GCM ensemble. Finally, a negative bias in the percent of annual precipitation occurring during autumn was seen for much of the United States, with large biases in the eastern United States and southern Texas.

Analogous to Figure 1, percent biases and percent changes from the historical to projected time periods for annual PET are shown in Figure 3. For a large portion of the continental United States (~75%), annual PET for the raw GCM ensemble is lower than bias-corrected GCM ensemble PET annual values with an average bias of -10 cm/year (-7% difference). Negative biases in PET derived from raw GCM output are pronounced for much of the western and eastern United States (-20 to -50 cm/year), whereas a positive bias in PET is apparent across the central United States (+10 to +30 cm/year).

The biases in the PET are a result of biases in the underlying variables used to derive PET. Figure 4 shows historical ensemble medians of average annual temperature, wind, specific humidity, and solar radiation, all of which are inputs for the Penman-Monteith PET calculations. For the average annual temperature, the raw GCM ensemble has a cold bias for the western United States (~2-4 ºC colder than the bias-corrected GCM ensemble) and a warm bias of 1-3 ºC for the central United States and 2-4 ºC for the coastal eastern United States. The spatial pattern of the average annual temperature biases was highly spatially correlated (0.64) to the PET biases, suggesting that temperature biases may have a large control on the PET bias. Much of the central to western United States and coastal eastern United States has a positive bias in wind of approximately 1-4 m/s, while much of the eastern United States has a neutral or negative wind bias (although wind biases did not correlate well with the PET biases; -0.02 correlation). Much of the United States has a positive bias in specific humidity, ranging from 0.0005 to 0.002 kg/kg. However, the specific humidity biases presented here are relatively small compared to the typical

values of specific humidity. Even so, there is a negative correlation (-0.48) of specific humidity biases to PET biases. Lastly, raw GCM ensemble solar radiation exhibited a positive bias of the United States with an average of 2.8 W/m^2. As with specific humidity, this overall bias is relatively small, but extreme positive biases in the Midwestern and western United States and extreme negative biases in the Pacific Northwest occur within a few grid cells. The only regions where negative biases occur are in the Pacific Northwest and along the western and eastern coasts. Solar radiation biases are moderately positively correlated (0.38) to PET biases.

Both the raw and bias-corrected GCMs project increases in PET by the late 21st century across the entire United States (Figure 3). The spatial patterns between the raw and bias-corrected GCM ensembles are similar (large increases in the central United States). The bias-corrected and raw GCM ensembles project similar increases in PET for the continental United States (~34 cm/year increase; 29% increase; Figure 3). A large portion of the United States PET projections are lower for the raw GCM ensemble (10-20 cm/year lower) compared to the bias-corrected GCM. The only extensive regions where PET is higher for the raw GCM ensemble are the southwestern and central United States.

3.2 Comparison of deficit months between raw and bias-corrected GCMs

The comparison of the historical time period between the raw and bias-corrected GCM ensembles indicates that the raw GCM ensemble underestimates the number of dry months (PET > precipitation) compared to the bias-corrected GCM ensemble by, on average, 1 month for the continental United States (Figure 5). For the southwestern United States, the bias-corrected GCM ensemble indicates that 11-12 months of the year will be dry, while the raw GCM ensemble indicates 10-11 deficit months.

For the projected time period, the changes in precipitation (Figure 1) and PET (Figure 3) cause the average number of deficit months for the raw and bias-corrected GCMs to increase by 0.88 months and 1.11 months, respectively. Deficit months calculated from the bias-corrected GCMs indicate that much of United States will have more than 8 deficit months out of the year, while for the raw GCMs there are still regions where the number of deficit months is less than 6. Increases in precipitation (Figure 1) are not large enough to overcome the increase in PET (Figure 3), resulting in projections of overall increases in deficit months for the entire continental United States.

3.3 Comparison of aridity between raw and bias-corrected GCMs

The bias-corrected GCM ensemble leads to more arid conditions (defined as a lower AI) for the historical and projected time periods (Figure 6). The raw and bias-corrected GCM ensembles are in general agreement of the AI values for much of the central United States; however, there the bias-corrected ensemble leads to more arid conditions from Louisiana to Iowa. There are large differences in AI for the eastern and western United States between the raw and bias-corrected GCM ensembles consistent with the aforementioned biases in PET and precipitation. For both regions, the historical and projected AI values for the raw GCM ensemble estimates much more humid conditions (larger AI values) than does the bias-corrected GCM ensemble. The raw GCM ensemble projects larger decreases in AI.

3.4 Comparison of PDSI between raw and bias-corrected GCMs

Drought occurrence and magnitude, defined by JJA PDSI, are projected to increase for the entire United States for both the raw and bias-corrected GCM ensembles (Figure 7). The raw and bias-corrected GCMs, however, have large differences (weighted mean absolute difference of 0.48) and produce a different spatial pattern of PDSI (spatially-weighted correlation of just 0.38).

For much of the western United States and southern Mississippi valley, the JJA PDSI from the raw GCM ensemble is more negative (drier) than the bias-corrected GCM ensemble on the order of 1 to 2 PDSI units.

To assess the spread of PDSI projections across the raw and bias-corrected GCM ensembles, the number of GCMs in the ensemble that result in a late 21st ensemble average JJA PDSI of at least -1 (mild drought), -2 (moderate drought), and -3 (severe drought) were tallied (Figure 8). The GCMs from the raw and bias-corrected ensemble general concur on mild drought conditions in the late 21st century (mean absolute error = 1.3; and spatially-weighted correlation = 0.6). The only major region of difference is the southwestern United States, where more GCMs (4-8) from the raw ensemble project mild drought conditions than does the bias-corrected GCM ensemble. The largest differences occur for moderate and severe drought conditions. For moderate drought conditions, the raw GCMs are in near concurrence (>15 GCMs) that the late 21st century time period will to more arid conditions for south-central United States and intermountain western United States. For the bias-corrected GCMs, there is still a concurrence that these regions will transition to more arid climates by the end of the 21st century. However, the number of GCMs that agree on the drought occurrences are lower. Another discrepancy is that the bias-corrected GCMs agree on moderate drought conditions for the upper Midwest (>10 GCMs), while moderate drought conditions are only projected for approximately half of the raw GCMs for this same region. Lastly, projections from both GCM ensembles suggest less than moderate drought conditions (< -2 JJA PDSI) for the California region, with the bias-corrected GCM ensemble encompassing a larger area. For severe drought conditions, the raw and bias-corrected GCM ensembles do not concur on severe drought conditions for both coasts of the United States. However, there are several regions in the central United States where both ensemble project, on average, severe drought conditions

during the late 21st century. Differences between the raw and bias-corrected ensembles for severe drought conditions are spatially similar to the moderate drought conditions.

3.5 Variation around the GCM ensemble median

Apart from the count of GCMs that agree on a mild, moderate, and severe drought for the late-21st century, the results presented in this study are for the ensemble medians. While the GCM ensemble median is resistant to outliers (i.e., unrepresentative GCMs) and, therefore, a reliable indicator of spatial variability and temporal changes, the variability around the GCM median can reveal the overall agreement between GCMs. Figure 9 presents the interquartile range of projected changes in annual precipitation, PET, AI, and JJA PDSI around the ensemble median for the raw and downscaled GCMs for the end of the 21st century time period. For precipitation, the average interquartile range is 12 cm for the raw GCM ensemble and 10 cm for the downscaled GCM ensemble. However, Figure 9 indicates that the spatial pattern of inter-GCM variability differs between the downscaled GCM ensemble and the raw GCM ensemble. For the downscaled GCM ensemble, the largest precipitation variability across GCMs occurs in the southeastern and northwestern United States, while a large portion of the western United States has low variability on the order of 6 cm. The variability around the raw GCM ensemble median is highest in the Pacific northwestern (20 cm) and southeastern United States (16 cm), though the variability is noticeably higher than the downscaled ensemble throughout the United States. For PET, the variability around the raw GCM ensemble median is much higher than the downscaled GCM variability. For a large portion of the United States, the variability of PET around the raw GCM ensemble median is greater than 20 cm, a value that is not exceeded for the downscaled GCM ensemble. Owning to the limited variability around the bias-corrected GCM ensemble, the AI (precipitation/PET) variability is much lower than the raw GCM ensemble. For JJA PDSI, there is

high variability for much of the United States, with an average JJA PDSI IQR of 3.8 PDSI units for both ensembles. Interestingly, JJA PDSI for both the raw and downscaled GCMs have a high variability around the upper Midwest of the United States. High variability with the raw GCMs encompasses a larger area in the upper Midwest, encroaching into parts of the intermountain western United States. It is important to note that the JJA PDSI IQR range spans both dry (negative PDSI) and wet conditions (positive PDSI).

4. Discussion and Summary

GCMs can exhibit significant biases when simulating the observed climate due to their spatial resolution and simplified parameterizations of important climatic processes (Randall et al. 2007). Even so, there is still confidence that GCMs provide useful estimates of future climatic changes (Randall et al. 2007). Given the results presented here, however, we recommend that raw GCM outputs be used with caution for climatic and hydrological impact studies due to their coarse resolution and precipitation and PET biases. To link output from GCMs to a local or regional scale, downscaling and bias-correction is needed. We show that using raw GCM data with biases in precipitation and the climatic inputs that determine PET can lead to inaccurate projections of aridity and drought for the continental United States. Both projections, however, do not include the role of increased vegetation water use efficiency in an elevated CO_2 environment.

For the GCM ensembles used in this work, the raw GCM ensemble had a positive (wet) precipitation bias compared to the bias-corrected GCM ensemble for much of the United States. This wet bias has been found in several recent studies (Mehran et al. 2014; Rupp et al. 2013; Sheffield et al. 2013b). A recent study by Nasrollahi et al. (2015) indicates annual precipitation biases in CMIP5 model simulations result in discrepancies in observed wetting and drying trends,

445 leading to the conclusion that much work needs to be done in simulating the daily and monthly
446 intensities and frequencies of regional precipitation.

447 The annual precipitation biases assessed in this study are rooted in the biases in seasonal
448 precipitation. We find that the seasonal differences were largest for the winter and summer time
449 periods. The raw GCM ensemble overestimates winter precipitation for a large portion of the
450 United States, while, for these same regions, the summer precipitation is underestimated. This is
451 likely a consequence of the raw GCMs doing a poor job in simulating the northward extent and
452 precipitation amount of the North American monsoon (Geil et al. 2013) and poor representation
453 of convective precipitation in the interior of the United States (Mehran et al. 2014). Additionally,
454 the winter biases in the western United States are likely due to orographic precipitation not being
455 fully resolved in the raw GCMs (Sheffield et al. 2013b).

456 To our knowledge, a detailed assessment of the PET bias that results from bias in PET
457 inputs has yet to be performed. We show that PET estimates from the bias-corrected GCM
458 ensemble lead to higher PET than does the raw GCM ensemble. Scheff and Frierson (2015) found
459 that precipitation biases were anti-phase to PET biases and our results confirm this (Figures 1 and
460 3). Examining the individual components of the Penman-Monteith inputs suggests that bias in
461 temperature is highly correlated with bias in PET from the raw GCM ensemble, which is physically
462 consistent. Saturation vapor pressure is exponentially related to temperature and therefore a bias
463 in temperature (particularly when temperature is high) can lead to even larger biases in the vapor
464 pressure deficit and the slope of the saturation vapor curve (Δ in Penman-Monteith PET). Biases
465 in wind were not correlated with PET biases. In the Penman-Monteith equation, wind is directly
466 but non-linearly related to PET, so once wind speed surpasses a relatively low value (~2-3 m/s),
467 PET responds weakly to wind speed. Nearly the entire United States is above this threshold for

both ensembles (Figure 4) and therefore any bias in wind will likely not result in much PET bias as long as the wind speeds do not become extremely low. It is important to note that the GCMs do not show declines in wind speeds on par with observations, although some GCMs do show a reduction by mid-century (Kumar et al., 2015). The biases in specific humidity and solar radiation were highly correlated with PET bias, but the magnitudes of their biases are relatively low.

The number of deficit months (PET > precipitation) for both ensembles indicates that the raw GCM ensemble underestimates the number of deficit months compared to the bias-corrected GCM ensemble, especially for the southwestern United States. Even with precipitation and PET biases, the raw and bias-corrected GCM ensembles project a drier future, albeit with different magnitudes. Biases in the raw GCM ensemble produce an overall lower number of deficit months (PET > precipitation) than the bias-corrected GCM ensemble largely as a result of higher PET estimates coupled with lower precipitation values from the bias-corrected ensemble. This monthly imbalance of PET and precipitation leads to larger areas and shifts in aridity for the bias-corrected GCM ensemble. While the comparison of AI shifts between raw and bias-corrected GCM ensemble is a novel result, previous studies on projected aridity changes largely agree with our findings and show an overall increase in arid lands will occur (Feng and Fu 2013; Overpeck and Udall 2010; Seager et al. 2007). However, these studies and our study do not consider the role that elevated CO_2 may play in increasing vegetation water use efficiency, which has recently been studied in Donohue et al. (2013) and Roderick et al. (2015). Their conclusion is that a warmer atmosphere does not necessarily lead to more arid conditions when additional effects of elevated CO_2 are considered. It is also important to note, however, that the use of AI may be too simplified to make strong conclusions on shifts of aridity. While the AI has been extensively used throughout the literature, recent work by Girvetz and Zganjar (2014) suggest that calculating AI only using

491 months where there is a precipitation deficit may be more useful than using all months. This
492 adjustment would correct for a bias in the original AI towards changes in precipitation that occur
493 during already wet months, resulting in a dampened precipitation signal during the deficit months
494 when summed at the annual time step (Chou et al. 2013). Using more advanced aridity metrics
495 such as the cumulative moisture surplus and cumulative moisture deficit may provide stronger
496 insight into the differences between the raw and downcaled GCM ensembles and is perhaps a
497 future avenue of study.

498 While the examination of long-term precipitation deficits (aridity) are important, short-
499 term precipitation deficits (drought; PDSI) can also give insight on the differences between the
500 raw and bias-corrected GCM ensemble, even if averaged over a 30-year time period. PDSI
501 calculates a simplified water balance at the monthly time step, and therefore month-to-month
502 variability in soil moisture (and the precipitation/PET balance) is captured. It is important to note
503 that future projections of soil moisture in the Intergovernmental Panel on Climate Change
504 Assessment Report 5 (IPCC AR5) indicate drier soil moistures in the United States, but the lower
505 values (<2 mm in the top 10 cm in the soil compared to historical averages) do not drastically
506 deviate from historical soil moisture (IPCC, 2013). This result compared to our result suggests two
507 very different futures of soil moisture and is likely due to the differences in soil moisture
508 estimation. A future avenue of study would be to bias-correct soil moisture output directly from
509 GCMs and then examine differences in projections.

510 We show that drought occurrences and magnitudes are projected to increase for the entire
511 United States for JJA, similar to other studies using different GCMs and time period (e.g., Cook
512 et al. 2015; Zhao and Dai 2015). There are some locations, mainly located on the eastern and
513 western United States coasts, where average JJA PDSI for the future time periods is projected to

be less than -2, indicating a transition to a more arid climate that is akin to a "moderate drought" under current climatic conditions. However, based on Figures 7 and 8, there are individual GCMs that project even more extreme (positive and negative) PDSI values at the end of the 21st century.

Compared to the raw GCM ensemble, we find higher PDSI values (decreased occurrence and magnitude of drought) and greater concurrence between GCMs for the southwestern United States for the bias-corrected GCM ensemble. As several recent studies have examined this area of the United States using paleoclimate reconstructions and GCM projections (Ault et al. 2014; Cook et al. 2015; Cook et al. 2004), it is important to note the large wet bias in raw GCM output in winter precipitation and biases in precipitation seasonality across this region. Coupling a constant or decrease in projected precipitation from the raw GCM ensemble with an increase in projected PET will result in increases in drought. Therefore, for the western United States, using a raw GCM ensemble may lead to drier conditions than would be expected if the GCM biases were bias-corrected. For other regions such as the Upper Midwest, the bias-corrected GCMs project larger increases in drought compared to the raw GCMs, primarily due to the positive precipitation bias in the raw GCMs in the seasons preceding the JJA time period.

It is important to note and discuss the different time scales used in PDSI and AI. AI does not include any precipitation and PET seasonality – it is simply the ratio of the two variables for the entire year. PDSI, on the hand, explicitly has a temporal (autoregressive) component that has a memory of up to 12-18 months (Cook et al., 1999; Vincente-Serrano et al., 2010). Not only does PDSI have a long memory, but moisture surpluses and deficits for individual months can have a long-lasting effect on drought conditions, which is not necessarily true with AI. Additionally, PDSI values are in reference to its mean and variance during the PDSI calibration period, and therefore changes in PDSI reflect changes in standardized units and not absolute changes. Therefore, any

changes in PDSI are relative to the historical time period, which can be quite different between the raw and bias-corrected GCMs. Changes in PDSI (for both ensembles) may also be a result of a signal-to-noise issues such that if model variability is low, a moderate change in aridity can produce a large decrease in PDSI. Lastly, the raw GCM ensemble has much more variability around the GCM ensemble median than the bias-corrected GCM ensemble for all climatological components, especially for PET, which is largely due to the raw GCMs not being constrained by an observed dataset.

Using both raw and bias-corrected GCM ensembles, our results point towards a drier future in terms of aridity and drought. Even with an overall increase in dry conditions, substantial regional differences exist between raw and bias-corrected GCM projections. These differences are particularly large in the southwestern and midwestern United States, which has important implications for water resources management and agriculture.

Acknowledgements

We acknowledge the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP, and we thank the climate modeling groups (listed in Table 1 of this paper) for producing and making available their model output. For CMIP, the US Department of Energy's Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals. The data used in this work are freely available at http://maca.northwestknowledge.net/.

References cited

Abatzoglou, J. T., 2013: Development of gridded surface meteorological data for ecological applications and modelling. *International Journal of Climatology*, **33,** 121-131.

Abatzoglou, J. T., and T. J. Brown, 2012: A comparison of statistical downscaling methods suited for wildfire applications. *International Journal of Climatology*, **32,** 772-780.

Allen, R. G., L. S. Pereira, D. Raes, and M. Smith, 1998: Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56. *FAO, Rome*, **300,** 6541.

564 Alley, W. M., 1984: The Palmer Drought Severity Index: Limitations and Assumptions. *Journal*
565 *of Climate and Applied Meteorology*, **23,** 1100-1109.
566 Ault, T. R., J. E. Cole, J. T. Overpeck, G. T. Pederson, and D. M. Meko, 2014: Assessing the Risk
567 of Persistent Drought Using Climate Model Simulations and Paleoclimate Data. *Journal*
568 *of Climate*, **27,** 7529-7549.
569 Barkstrom, B. R., 1984: The Earth Radiation Budget Experiment (ERBE). *Bulletin of the American*
570 *Meteorological Society*, **65,** 1170-1185.
571 Chou, C., J. C. H. Chiang, C.-W. Lan, C.-H. Chung, Y.-C. Liao, and C.-J. Lee, 2013: Increase in
572 the range between wet and dry season precipitation. *Nature Geosci*, **6,** 263-267.
573 Christensen, J. H., F. Boberg, O. B. Christensen, and P. Lucas-Picher, 2008: On the need for bias
574 correction of regional climate change projections of temperature and precipitation.
575 *Geophysical Research Letters*, **35,** L20709.
576 Cook, B., J. Smerdon, R. Seager, and S. Coats, 2014: Global warming and 21st century drying.
577 *Clim Dyn*, **43,** 2607-2627.
578 Cook, B. I., T. R. Ault, and J. E. Smerdon, 2015: Unprecedented 21st century drought risk in the
579 American Southwest and Central Plains. **1,** e1400082.
580 Cook E R, D.M. Meko, and M.K. Cleaveland, 1999: Drought Reconstructions for the Continental
581 United States. *Journal of Climat,e* **12,** 1145-62
582 Cook, E. R., C. A. Woodhouse, C. M. Eakin, D. M. Meko, and D. W. Stahle, 2004: Long-Term
583 Aridity Changes in the Western United States. *Science*, **306,** 1015-1018.
584 Dai, A., 2011: Characteristics and trends in various forms of the Palmer Drought Severity Index
585 during 1900–2008. *Journal of Geophysical Research: Atmospheres*, **116,** D12115.
586 ——, 2013: Increasing drought under global warming in observations and models. *Nature Clim.*
587 *Change*, **3,** 52-58.
588 Daly, C., M. Halbleib, J.I. Smith, W.P. Gibson, M.K. Doggett, G.H. Taylor, J. Curtis, and P.P
589 Pasteris, 2008: Physiographically sensitivity mapping of climatological temperature and
590 precipitation across the conterminous United States. *International Journal of Climatology*,
591 **28,** 2031-2064
592 Diffenbaugh, N. S., D. L. Swain, and D. Touma, 2015: Anthropogenic warming has increased
593 drought risk in California. *Proceedings of the National Academy of Sciences*, **112,** 3931-
594 3936.
595 Donohue, R.J., M.L., Roderick, T.R. McVicar, and G.D. Farquhar, 2013: Impact of CO_2
596 fertilization on maximum foliage cover across the globe's warm, arid environments.
597 *Geophysical Research Letters*, **40,** 3031-3035.
598 Ehret, U., E. Zehe, V. Wulfmeyer, K. Warrach-Sagi, and J. Liebert, 2012: Should we apply bias
599 correction to global and regional climate model data? *Hydrology and Earth System*
600 *Sciences*, **16,** 3391-3404.
601
602 Feng, S., and Q. Fu, 2013: Expansion of global drylands under a warming climate. *Atmos. Chem.*
603 *Phys.*, **13,** 10081-10094.
604 Ficklin, D.L., Y. Luo, E. Luedeling, and M. Zhang, 2009: Climate change sensitivity assessment
605 of a highly agricultural watershed using SWAT. *Journal of Hydrology*, **374**, 16-29.
606 Ficklin, D.L, S.L. Letsinger, H. Gholizadeh, and J. T. Maxwell, 2015. Incorporation of the
607 Penman–Monteith potential evapotranspiration method into a Palmer Drought Severity
608 Index Tool. *Computers and Geosciences*, **85,** 136-141**.**

Ficklin, D. L., J. T. Maxwell, S. Letsinger, L., and H. Gholizadeh, 2015: A climatic deconstruction of recent drought trends in the United States. *Environmental Research Letters*, **10,** 044009.

Fowler, H. J., S. Blenkinsop, and C. Tebaldi, 2007: Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling. *International Journal of Climatology*, **27,** 1547–1578, DOI: 1510.1002/joc.1556.

Geil, K. L., Y. L. Serra, and X. Zeng, 2013: Assessment of CMIP5 Model Simulations of the North American Monsoon System. *Journal of Climate*, **26,** 8787-8801.

Girvetz, E., and C. Zganjar, 2014: Dissecting indices of aridity for assessing the impacts of global climate change. *Climatic Change*, **126,** 469-483.

Glotter, M., J. Elliott, D. McInerney, N. Best, I. Foster, and E. J. Moyer, 2014: Evaluating the utility of dynamical downscaling in agricultural impacts projections. *Proceedings of the National Academy of Sciences*, **111,** 8776-8781.

Griffin, D., and K. J. Anchukaitis, 2014: How unusual is the 2012–2014 California drought? *Geophysical Research Letters*, **41,** 9017-9023.

Hagemann, S., C. Chen, J. O. Haerter, J. Heinke, D. Gerten, and C. Piani, 2011: Impact of a Statistical Bias Correction on the Projected Hydrological Changes Obtained from Three GCMs and Two Hydrology Models. *Journal of Hydrometeorology*, **12,** 556-578.

Herweijer, C., and R. Seager, 2008: The global footprint of persistent extra-tropical drought in the instrumental era. *International Journal of Climatology*, **28,** 1761-1774.

Hoerling, M., and Coauthors, 2013: Causes and Predictability of the 2012 Great Plains Drought. *Bulletin of the American Meteorological Society*, **95,** 269-282.

Huang, J., H. Yu, X. Guan, G. Wang, R, and R. Guo, 2015: Accelerated dryland expansion under climate change. *Nature Climate Change*, doi:10.1038/nclimate2837

Hulme, M., 1996: Recent Climatic Change in the World's Drylands. *Geophysical Research Letters*, **23,** 61-64.

IPCC, 2013: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1535 pp, doi:10.1017/CBO9781107415324.

Johnson, F., Sharma, A., 2010: A comparison of Australian open water body evaporation trends for current and future climates estimated from Class A evaporation pans and general circulation models. *Journal of Hydromet.*, **11**, 105-121.

Johnson, F., and A. Sharma, 2015: What are the impacts of bias correction on future drought projections? *Journal of Hydrology*, **525,** 472-485.

Karl, T., and R. W. Knight, 1985: *Atlas of monthly Palmer hydrological drought indices (1931-1983) for the contiguous United States.* National Climatic Data Center.

Kumar, D., V. Mishra, and A. R. Ganguly, 2015. Evaluating wind extremes in CMIP5 climate models. *Climate Dynamics*, **45**, 441-453.

Li, C., E. Sinha, D. E. Horton, N. S. Diffenbaugh, and A. M. Michalak (2014), Joint bias correction of temperature and precipitation in climate model simulations, J. Geophys. Res. Atmos., 119, 13,153–13,162, doi:10.1002/2014JD022514.

Maurer, E. P., and H. G. Hidalgo, 2008: Utility of daily vs. monthly large-scale climate data: an intercomparison of two statistical downscaling methods. *Hydrol. Earth Syst. Sci.*, **12,** 551-563.

Maurer, E. P., and D. W. Pierce, 2014: Bias correction can modify climate model simulated precipitation changes without adverse effect on the ensemble mean. *Hydrol. Earth Syst. Sci.*, **18,** 915-925.

Mehran, A., A. AghaKouchak, and T. J. Phillips, 2014: Evaluation of CMIP5 continental precipitation simulations relative to satellite-based gauge-adjusted observations. *Journal of Geophysical Research: Atmospheres*, **119,** 1695-1707.

Muerth, M.J., B. G. St-Denis, S. Ricard, J.A. Velázquez, J. Schmid, M. Minville, D. Caya, D. Chaumont, R. Ludwig, and R. Turcotte, 2013: On the need for bias correction in regional climate scenarios to assess climate change impacts on river runoff. *Hydrology and Earth System Sciences*, **17,** 1189-1204.

Nasrollahi, N., and Coauthors, 2015: How well do CMIP5 climate simulations replicate historical trends and patterns of meteorological droughts? *Water Resources Research*, **51,** 2847-2864.

Overpeck, J., and B. Udall, 2010: Dry Times Ahead. *Science*, **328,** 1642-1643.

Palmer, W., 1965: Meteorological Drought. U. S. Dept. of Commerce Weather Bureau. Research Paper 45..

Pierce, D.W., D. R. Cayan, E.P. Maurer, J.T. Abatzoglou, and K.C. Hegewisch, 2015: Improved bias correction techniques for hydrological simulations of climate change. *Journal of Hydrometeorology*, doi: http://dx.doi.org/10.1175/JHM-D-14-0236.1, in press.

Polade, S. D., A. Gershunov, D. R. Cayan, M. D. Dettinger, and D. W. Pierce, 2013: Natural climate variability and teleconnections to precipitation over the Pacific-North American region in CMIP3 and CMIP5 models. *Geophysical Research Letters*, **40**, 2296-2301.

Randall, D. A., and Coauthors, 2007: Climate Models and Their Evaluation. *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment, Report of the Intergovernmental Panel on Climate Change*, S. Solomon, and Coauthors, Eds., Cambridge University Press.

Roderick, M.L., P. Greve, and G.D. Farquhar, 2015: On the assessment of aridity with changes in atmospheric CO2. *Water Resources Research*, **51,** 5450 - 5463.

Rupp, D. E., J. T. Abatzoglou, K. C. Hegewisch, and P. W. Mote, 2013: Evaluation of CMIP5 20th century climate simulations for the Pacific Northwest USA. *Journal of Geophysical Research: Atmospheres*, **118,** 2013JD020085.

Rupp, D. E., P. Mote, N. Massey, C. J. Rye, R. Jones, and M. R. Allen, 2012: Did human influence on climate make the 2011 Texas drought more probable? *Bulletin of the American Meteorological Society* **93,** 1041-1067.

Scheff, J., and D. M. W. Frierson, 2014: Scaling Potential Evapotranspiration with Greenhouse Warming. *Journal of Climate*, **27,** 1539-1558.

——, 2015: Terrestrial aridity and its response to greenhouse warming across CMIP5 climate models. *Journal of Climate*. doi:10.1175/JCLI-D-14-00480.1, in press.

Seager, R., and G. A. Vecchi, 2010: Greenhouse warming and the 21st century hydroclimate of southwestern North America. *Proceedings of the National Academy of Sciences*, **107,** 21277-21282.

Seager, R., and M. Hoerling, 2014: Atmosphere and Ocean Origins of North American Droughts. *Journal of Climate*, **27,** 4581-4606.

Seager, R., and Coauthors, 2007: Model Projections of an Imminent Transition to a More Arid Climate in Southwestern North America. *Science*, **316,** 1181-1184.

699 Sheffield, J., and E. F. Wood, 2008: Global Trends and Variability in Soil Moisture and Drought
700 Characteristics, 1950–2000, from Observation-Driven Simulations of the Terrestrial
701 Hydrologic Cycle. *Journal of Climate*, **21,** 432-458.
702 Sheffield, J., E. F. Wood, and M. L. Roderick, 2012: Little change in global drought over the past
703 60 years. *Nature*, **491,** 435-438.
704 Sheffield, J., and Coauthors, 2013a: North American Climate in CMIP5 Experiments. Part II:
705 Evaluation of Historical Simulations of Intraseasonal to Decadal Variability. *Journal of*
706 *Climate*, **26,** 9247-9290.
707 Sheffield, J., and Coauthors, 2013b: North American Climate in CMIP5 Experiments. Part I:
708 Evaluation of Historical Simulations of Continental and Regional Climatology. *Journal of*
709 *Climate*, **26,** 9209-9245.
710 Vincente-Serrano, S.M., S. Beguería, and J.I. López-Moreno, 2010: A Multiscalar Drought Index
711 Sensitive to Global Warming: The Standardized Precipitation Evapotranspiration Index
712 *Journal of Climate*, **23** 1696-718.
713 Vicente-Serrano, S. M., and Coauthors, 2014: Evidence of increasing drought severity caused by
714 temperature rise in southern Europe. *Environmental Research Letters*, **9,** 044001.
715 Werick, W. J., G. E. Willeke, N. B. Guttman, J. R. M. Hosking, and J. R. Wallis, 1994: National
716 drought atlas developed. *Eos, Transactions American Geophysical Union*, **75,** 89-90.
717 Williams, A.P., R. Seager, J.T. Abatzoglou, B.I. Cook, J.E. Smerdon and E.R. Cook, 2015:
718 Contribution of anthropogenic warming to California drought during 2012-2012.
719 *Geophysical Research Letters*, **42,** 6819-6828.
720 Woodhouse, C. A., and J. T. Overpeck, 1998: 2000 Years of Drought Variability in the Central
721 United States. *Bulletin of the American Meteorological Society*, **79,** 2693-2714.
722 Woodhouse, C. A., D. M. Meko, G. M. MacDonald, D. W. Stahle, and E. R. Cook, 2010: A 1,200-
723 year perspective of 21st century drought in southwestern North America. *Proceedings of*
724 *the National Academy of Sciences*, **107,** 21283-21288.
725 Zhao, T., and A. Dai, 2015: The Magnitude and Causes of Global Drought Changes in the Twenty-
726 First Century under a Low–Moderate Emissions Scenario. *Journal of Climate*, **28,** 4490-
727 4512.
728
729 ***Figure and Table Captions***

730 **Table 1.** General Circulation Models used in this study for Representative Concentration

731 Pathway 8.5

732

733 **Figure 1**. Annual mean precipitation, differences, and biases between the raw and bias-corrected

734 GCM ensemble (rows) and projections for the 1950-1979 and 2070-2099 time periods (columns).

735 Differences in projections are shown in the lower-right panel.

Figure 2. Percent of annual total precipitation for 1950-1979 by climatological season for the raw and bias-corrected GCM ensemble as well as the difference.

Figure 3. Biases in annual mean potential evapotranspiration (PET) between the raw and bias-corrected GCM ensemble (rows) and PET projections for the 1950-1979 and 2070-2099 time periods (columns). Differences in projections are shown in the lower-right panel.

Figure 4. Biases in the climatic inputs for the Penman-Monteith annual mean potential evapotranspiration (PET) between the raw and bias-corrected GCM ensemble and PET projections for the 1950-1979 time period.

Figure 5. The number of annual deficit months (PET > P) for the raw and bias-corrected GCM ensembles (rows) for the 1950-1979 and 2070-2099 time periods (columns). Differences in projections are shown in the lower-right panel.

Figure 6. Aridity Index (precipitation/potential evapotranspiration) for the raw and bias-corrected GCM ensembles (rows) for the 1950-1979 and 2070-2099 time periods (columns). Differences in projections are shown in the lower-right panel.

Figure 7. Average June, July, and August PDSI for the 2070-2099 time period for the raw and bias-corrected GCM ensembles as well as the difference.

Figure 8. The number of raw and bias-corrected GCMs that agree on a 2070-2099 average of a PDSI value of less than -1 (mild drought), -2 (moderate drought), and -3 (severe drought).

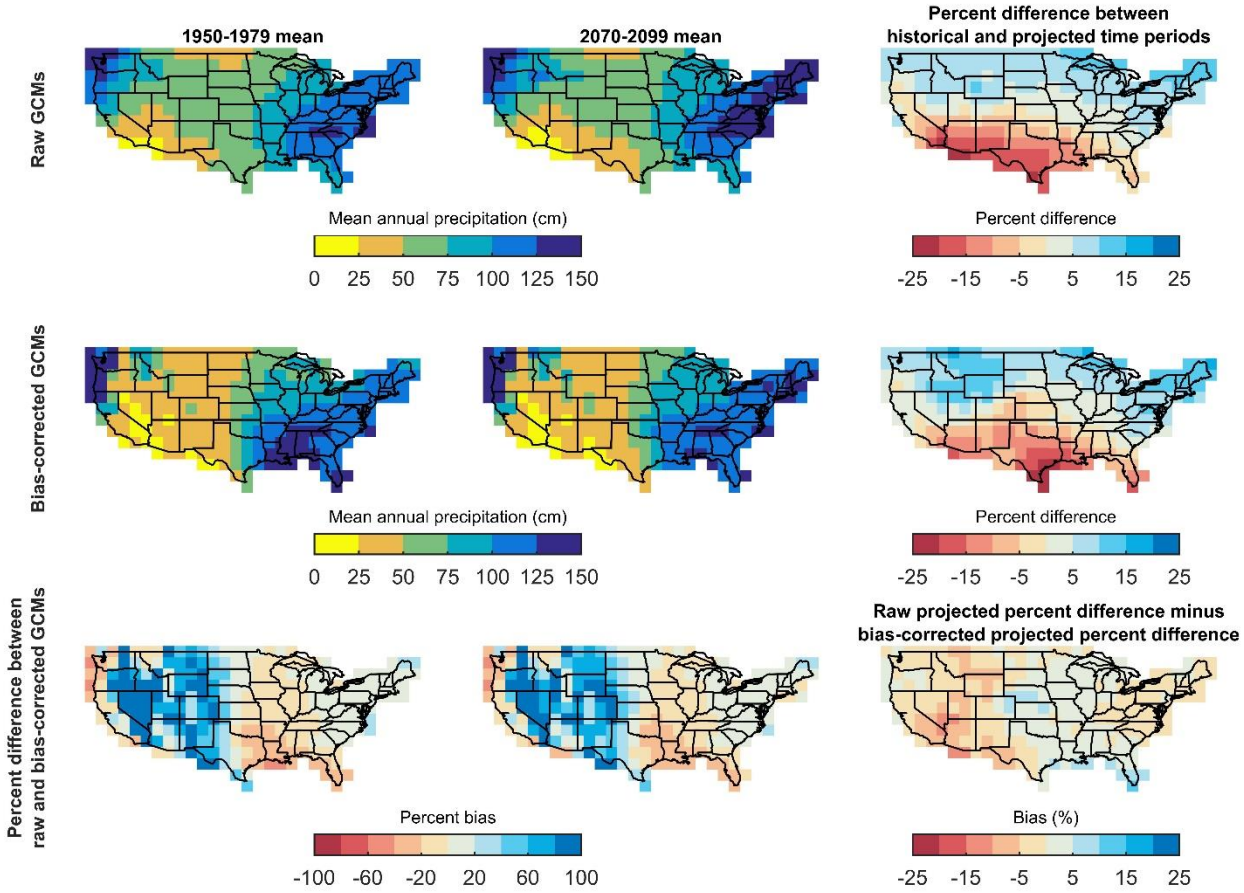
Figure 9. The interquartile range (IQR) around the 2070-2099 raw and bias-corrected GCM ensemble medians for precipitation, potential evapotranspiration, Aridity Index, and June, July, August PDSI.

760
761 **Table 1.** General Circulation Models used in this study for Representative Concentration Pathway 8.5

Model Name	Model Country	Model Agency
bcc-csm1-1	China	Beijing Climate Center, China Meteorological Administration
bcc-csm1-1	China	Beijing Climate Center, China Meteorological Administration
BNU-ESM	China	College of Global Change and Earth System Science, Beijing Normal University, China
CanESM2	Canada	Canadian Centre for Climate Modeling and Analysis
CNRM-CM5	France	National Centre of Meteorological Research, France
CSIRO-Mk3-6-0	Australia	Commonwealth Scientific and Industrial Research Organization/Queensland Climate Change Centre of Excellence, Australia
GFDL-ESM2M	USA	NOAA Geophysical Fluid Dynamics Laboratory, USA
GFDL-ESM2G	USA	NOAA Geophysical Fluid Dynamics Laboratory, USA
HadGEM2-ES	United Kingdom	Met Office Hadley Center, UK
HadGEM2-CC	United Kingdom	Met Office Hadley Center, UK
inmcm4	Russia	Institute for Numerical Mathematics, Russia
IPSL-CM5A-LR	France	Institut Pierre Simon Laplace, France
IPSL-CM5A-MR	France	Institut Pierre Simon Laplace, France
MIROC5	Japan	Atmosphere and Ocean Research Institute (The University of Tokyo), National Institute for Environmental Studies,and Japan Agency for Marine-Earth Science and Technology
MIROC-ESM	Japan	Japan Agency for Marine-Earth Science and Technology, Atmosphere and Ocean Research Institute (The University of Tokyo), and National Institute for Environmental Studies
MIROC-ESM-CHEM	Japan	Japan Agency for Marine-Earth Science and Technology, Atmosphere and Ocean Research Institute (The University of Tokyo), and National Institute for Environmental Studies
MRI-CGCM3	Japan	Meteorological Research Institute, Japan

762
763
764

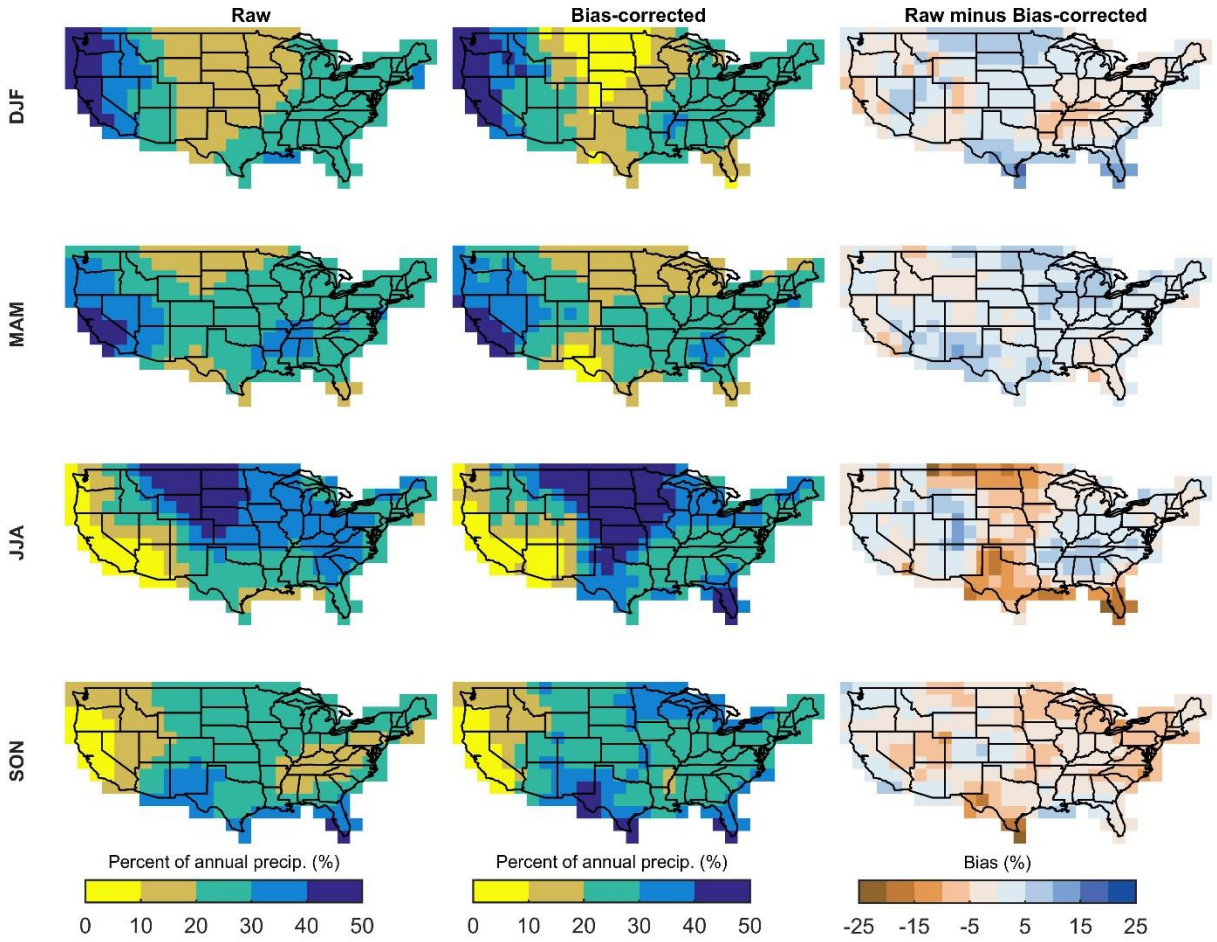
765
766 **Figure 1**. Annual mean precipitation, differences, and biases between the raw and bias-corrected
767 GCM ensemble (rows) and projections for the 1950-1979 and 2070-2099 time periods (columns).
768 Differences in projections are shown in the lower-right panel.
769



770
771
772
773
774
775
776
777
778
779
780
781

782 **Figure 2**. Percent of annual total precipitation for 1950-1979 by climatological season for the raw
783 and bias-corrected GCM ensemble as well as the difference.

784
785



786
787
788
789
790
791
792
793
794
795
796
797
798
799
800

Figure 3. Biases in annual mean potential evapotranspiration (PET) between the raw and bias-corrected GCM ensemble (rows) and PET projections for the 1950-1979 and 2070-2099 time periods (columns). Differences in projections are shown in the lower-right panel.

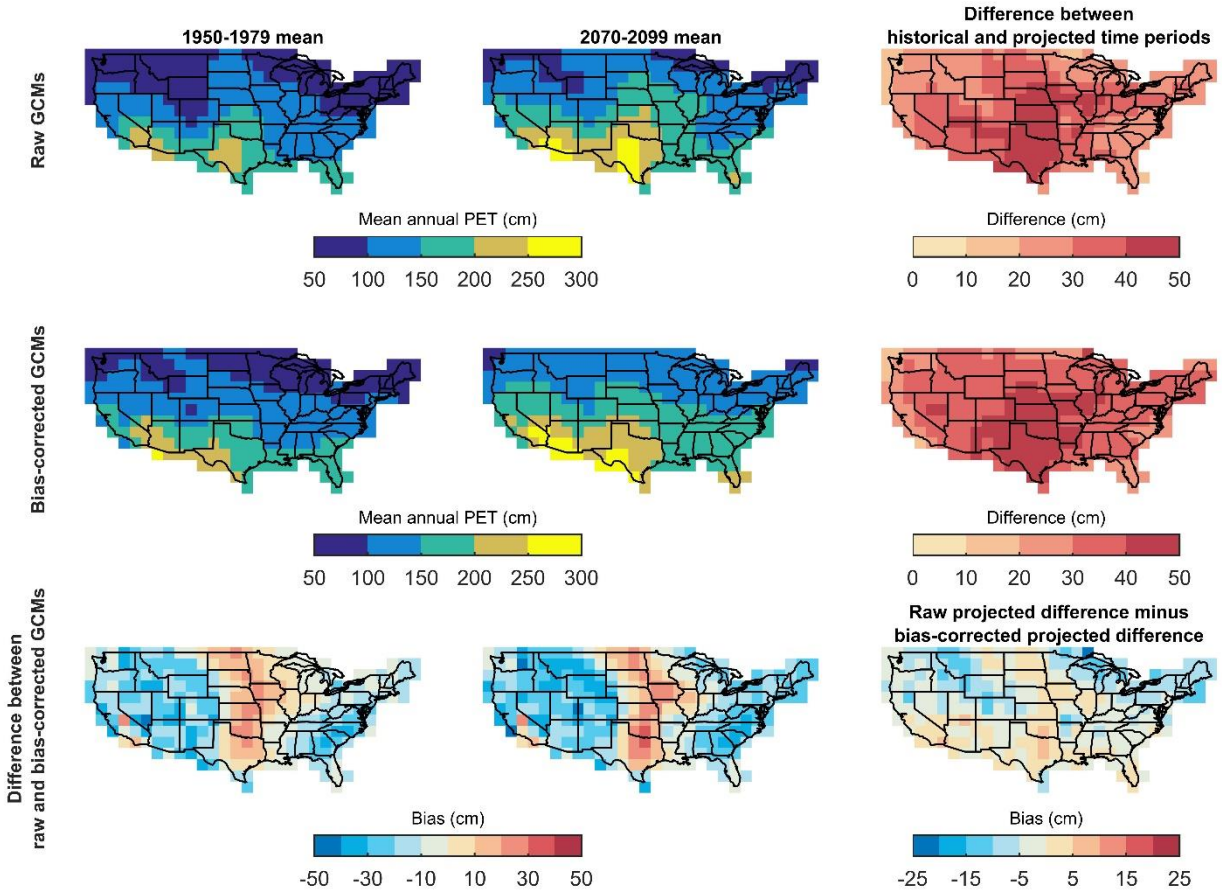


Figure 4. Biases in the climatic inputs for the Penman-Monteith annual mean potential evapotranspiration (PET) between the raw and bias-corrected GCM ensemble and PET projections for the 1950-1979 time period.

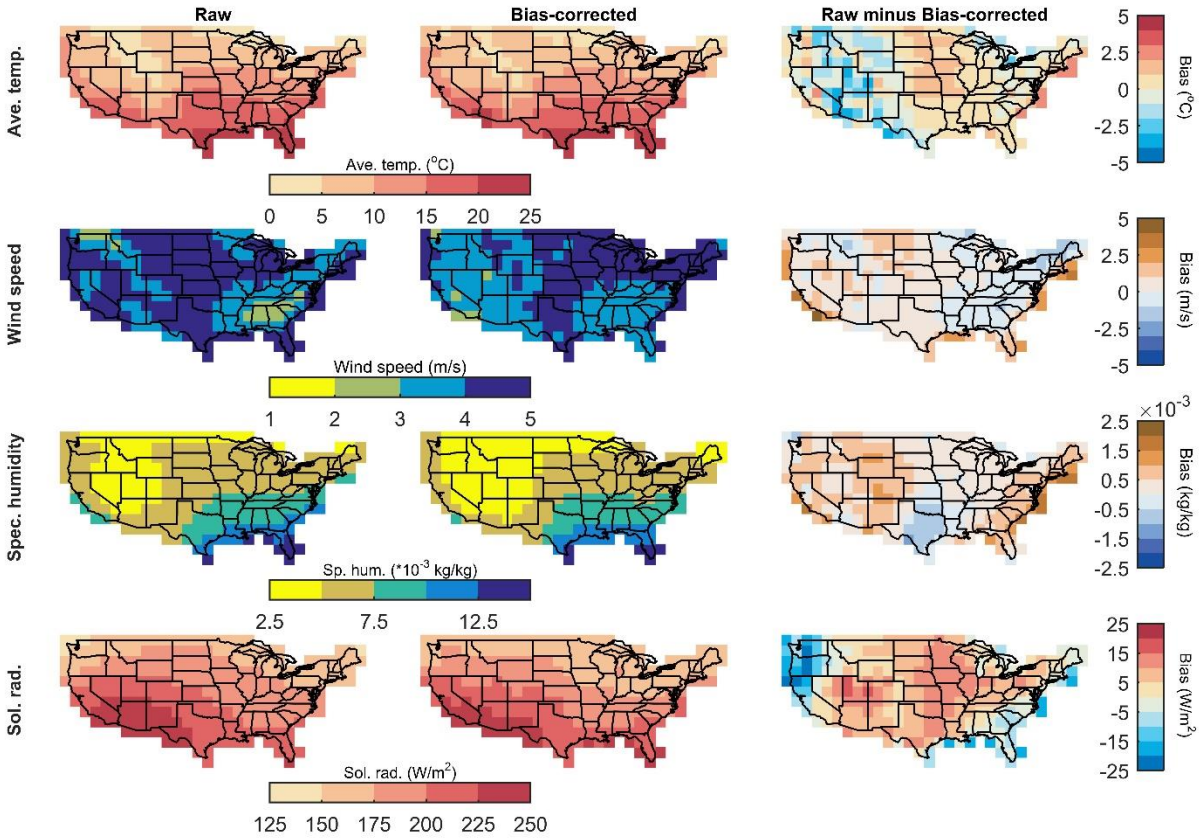


Figure 5. The number of annual deficit months (PET > P) for the raw and bias-corrected GCM ensembles (rows) for the 1950-1979 and 2070-2099 time periods (columns). Differences in projections are shown in the lower-right panel.

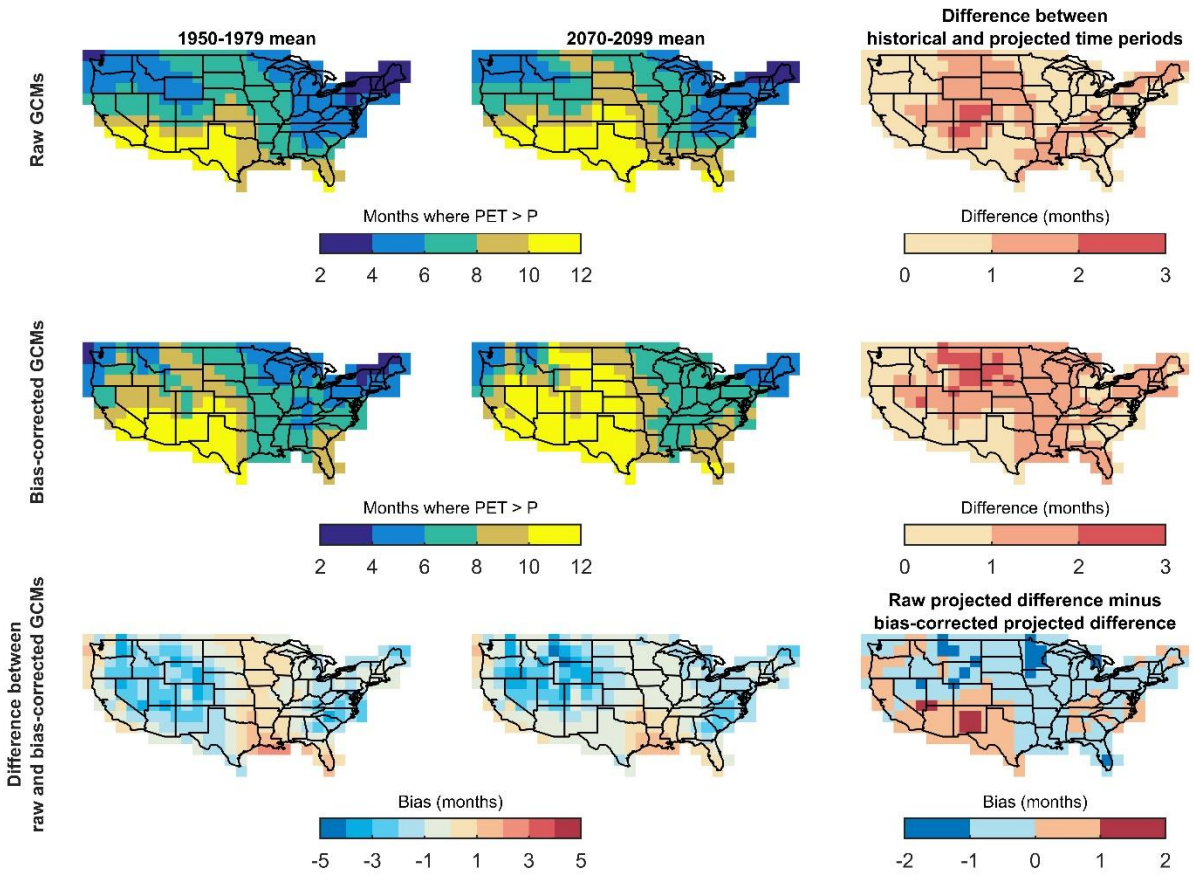
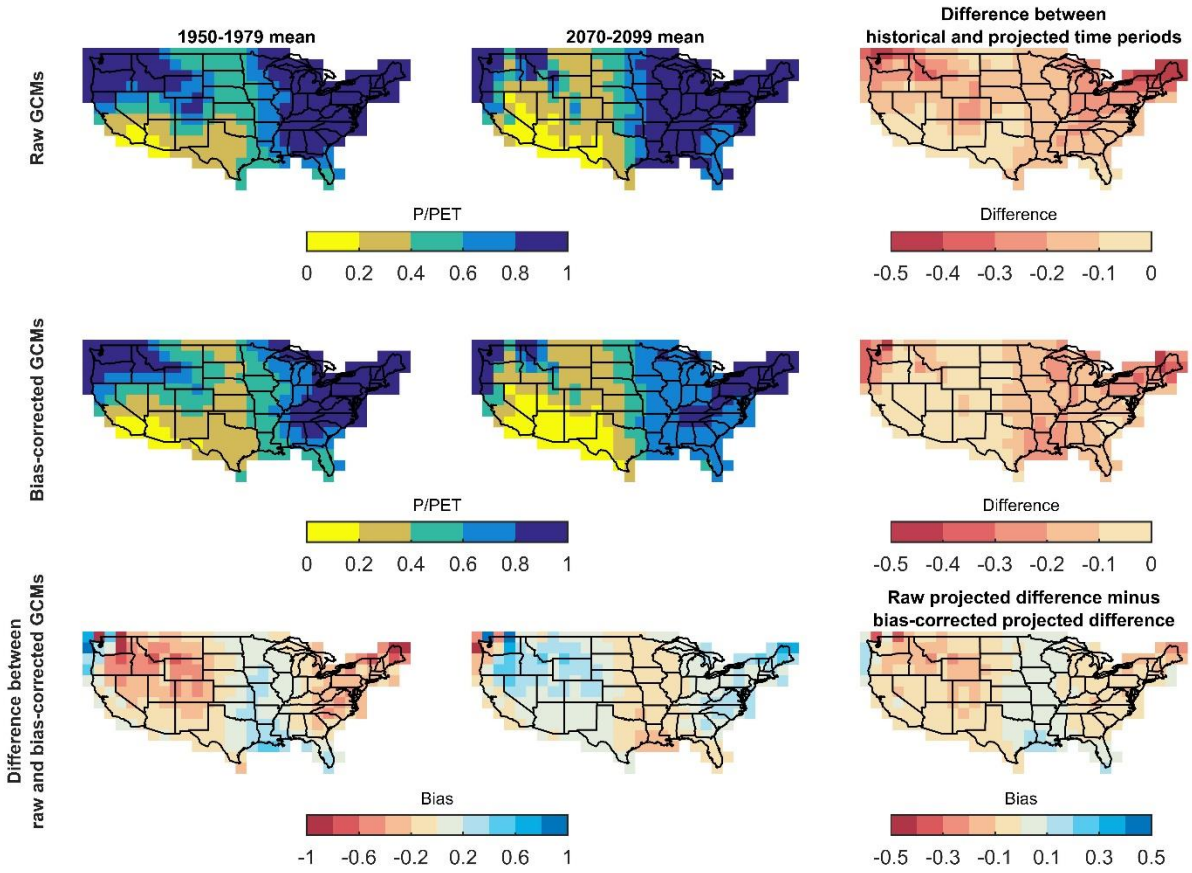
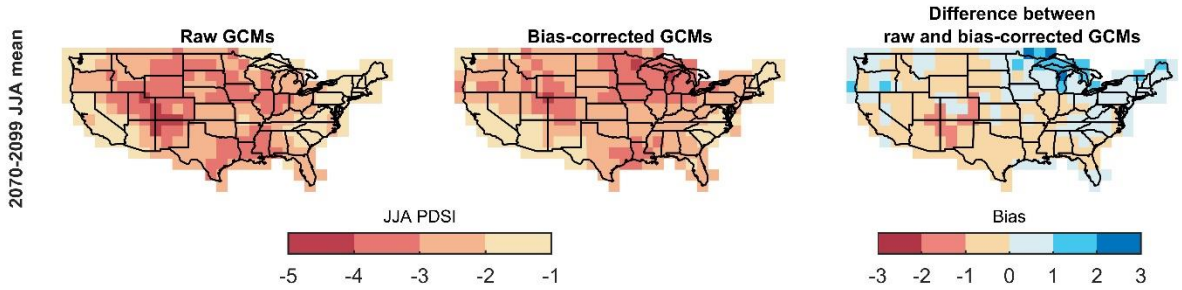


Figure 6. Aridity Index (precipitation/potential evapotranspiration) for the raw and bias-corrected GCM ensembles (rows) for the 1950-1979 and 2070-2099 time periods (columns). Differences in projections are shown in the lower-right panel.



878 **Figure 7.** Average June, July, and August PDSI for the 2070-2099 time period for the raw and
879 bias-corrected GCM ensembles as well as the difference.



880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910

Figure 8. The number of raw and bias-corrected GCMs that agree on a 2070-2099 average of a PDSI value of less than -1 (mild drought), -2 (moderate drought), and -3 (severe drought).

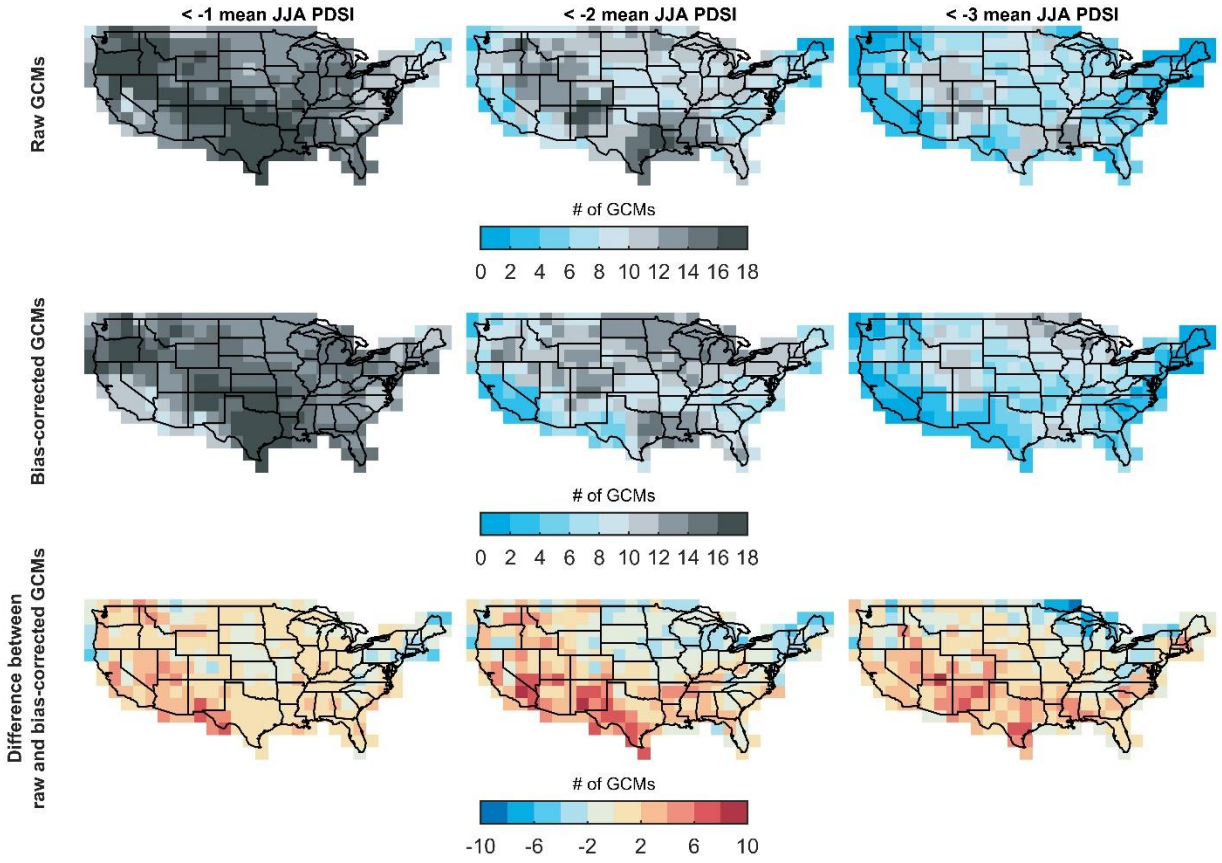


Figure 9. The interquartile range (IQR) around the 2070-2099 raw and bias-corrected GCM ensemble medians for precipitation, potential evapotranspiration, Aridity Index, and June, July, August PDSI.