Bias correction that addresses frequency dependence and preserves model-predicted mean changes

David W. Pierce[1,*], Daniel R. Cayan[1,2], Edwin P. Maurer[3], John T. Abatzoglou[4], Katherine C. Hegewisch[4]

[1]Division of Climate, Atmospheric Sciences, and Physical Oceanography, Scripps Institution of Oceanography, La Jolla, CA, 92093-0224

[2]U.S. Geological Survey, La Jolla, CA, 92037

[3]Civil Engineering Department, Santa Clara University, Santa Clara, CA, 95053-0563

[4]Department of Geography, University of Idaho, Moscow, ID, 83844

A report to the California Energy Commission

Version 1.0

12 January 2014

*Corresponding Author. Division of Climate, Atmospheric Sciences, and Physical Oceanography, Scripps Institution of Oceanography, La Jolla, CA, 92093-0224.

dpierce@ucsd.edu

Abstract

Global climate model temperature and precipitation fields need to be corrected for biases relative to observations before they can be used for climate change impact studies. Three existing bias correction methods, and a new one developed here, are applied to daily maximum temperature and precipitation from 21 climate models to investigate: 1) How bias correction alters the climate change signal of the original model; 2) How different methods affect model biases in the simulation of variance as a function of frequency. Quantile mapping (QM) and cumulative distribution function transform (CDF-t) bias correction can significantly alter the signal of change from the original climate model, with differences of up to 2°C and 30 percentage points for monthly temperature and precipitation, respectively. Equidistant quantile matching (EDCDFm) preserves model-predicted changes in daily maximum temperature, but alters model-predicted changes in precipitation by up to 30 percentage points in some locations. An extension to EDCDFm termed PresRat is introduced, which generally preserves the original model-predicted changes in precipitation by operating on ratios instead of differences, using a precipitation threshold to make the fraction of model zero-precipitation days match observations, and incorporating a final correction factor. Additionally, a frequency-dependent bias correction method is introduced that is twice as effective as standard bias correction in reducing errors in the models' simulation of variance as a function of frequency, and (unlike standard bias correction) does so while making very few locations worse.

1. Introduction

Global climate models (GCMs) are being used to explore an ever widening set of problems, some of which are sensitive to biases in the model simulated fields (IPCC, 2007). For example, daily precipitation biases can have a detrimental effect on hydrological simulations due to the non-linear nature of runoff; a moderate amount of precipitation generates little runoff if the soil is able to absorb the moisture, while doubling the precipitation might exceed the moisture storage capacity of the soil and generate much more than twice as much runoff. This non-linear relationship becomes more extreme in arid regions (Wigley and Jones, 1985), intensifying the sensitivity of runoff to GCM precipitation biases. Likewise, significant biases in surface humidity and evapotranspiration can arise from relatively small temperature biases due to the nonlinear nature of the Clausius-Clapyron equation. Unfortunately, the detrimental impacts of climate model biases on a non-linear system are not straightforward to remove.

For this reason hydrological simulations generally bias correct GCM output fields before they are used. Corrected variables include temperature and precipitation, and sometimes other relevant quantities such as downward solar radiation, humidity, or wind speed. Bias correction is often an integral part of a downscaling scheme that takes account of large scale GCM biases as well as topographical and other effects that operate at a finer scale than can be resolved by a GCM (e.g., Wood et al. 2002; Maurer et al. 2010). Here however we consider the bias correction step alone. Maraun (2013) has pointed out that bias correction is most straightforwardly applied on a spatial scale that is near the original GCM's spatial resolution, so we restrict our attention to bias correction on a grid commensurate with the original GCMs.

One common form of bias correction is quantile mapping (QM; e.g., Panofsky and Brier 1968; Wood et al. 2002; Thrasher et al. 2012), which adjusts a simulated climate variable (e.g.

temperature or precipitation) at a given location by mapping the quantiles of the simulated distribution onto the quantiles of the observations at that location. QM has been widely applied to climate model output over the U.S. (e.g., Maurer et al. 2007, 2014) and globally (Thrasher et al. 2012). QM alters both the model's mean and temporal variability, bringing them into agreement with observations over some common historical period. Gudmundsson et al. (2012) evaluate different ways of implementing QM and find that relatively straightforward non-parametric methods, such as used here, perform well compared to more complicated schemes.

Previous studies have shown that QM tends to alter the original GCM's projected trend (Hagemann et al. 2011; Pierce et al. 2013; Maurer and Pierce 2013). Whether this is a desirable feature is a research question not addressed here. However, this property certainly engenders confusion and inconsistent results, for example between bias corrected regional climate studies and GCM results assessed by the IPCC (2007, 2013). If a climate model has too much variability then QM tends to reduce variability on all timescales, suppressing the original trend. If the GCM has too little variability, QM tends to increase the trend along with variability on shorter timescales. As bias correction is a purely statistical method it fails to discriminate between the physical processes determining trends associated with anthropogenic forcing and shorter-term fluctuations associated with natural internal climate variability. From this perspective there is little justification for allowing bias correction that primarily addresses problems on synoptic, seasonal and annual timescales to change the trend as well.

Some previous schemes have addressed the problem of bias correction altering GCM-projected trends. For example, the BCSD method (Wood et al. 2002) removes temperature trends over the period to be downscaled, bias corrects and then downscales the anomalies to a fine grid, then adds back in the original GCM trend fields interpolated to the fine grid. Although straightforward and useful, this approach has the drawback that the final trend is, to first order, merely the interpolated GCM trend. This is confusing to end users who might reasonably expect that the downscaled trend will reflect more than simply the interpolated GCM trend, and means that the spatial structure of the trend is not necessarily commensurate with the spatial structure of the daily, monthly, or annual variability. In this work we address this limitation by constructing a bias correction method that retains the model-predicted change in the first place. The bias corrected fields can then be downscaled, and the final trend in the downscaled fields will be affected by the downscaling process rather than being independent of it. Pierce and Cayan (2013) addressed this issue by partitioning their future model runs into 30-yr segments (2010-2039, etc.), downscaling each segment with respect to its own climatology, then separately downscaling the global model predicted *change* in climatology in each segment and adding it back in. This preserves model changes on timescales longer than 30 years but allows shorter timescale changes to be bias corrected. The work here is more widely applicable since it is a self-contained bias correction method that preserves model-predicted changes without reference to how subsequent downscaling handles the trend. As such it is potentially applicable to downscaling methods where separately downscaling the model-predicted change is not desirable or viable.

Addressing the trend is not the only issue relevant to climate impact studies. The approaches used in BCSD and by Pierce and Cayan (2013) treat the trend differently from other timescales during the bias correction process. Yet a GCM may have too much variability on, for instance, synoptic timescales of 2-10 days but too little variability on the annual timescale. Neither a simple quantile based bias correction nor the approaches noted above address this problem. Misrepresentation of variance as a function of frequency could influence a simulation of heat waves or flooding events, while distortions in the relative importance of synoptic scale versus interannual variability could affect agriculture and ecosystems. Inaccurate partitions between short- and long-timescale precipitation variability could affect simulations of droughts and reservoir storage since the hydrological and ecological result of a given amount of annual precipitation varies greatly depending on whether the precipitation is delivered equally throughout the year or very unevenly, with a strong contrast between wet and dry seasons.

While the CMIP5 GCM simulations appear to have improved in these regards relative to prior simulations, such biases can still be substantial (Sillmann et al. 2013). Lower frequency variability in the climate system, such as El Niño-Southern Oscillation (ENSO), with an observed period of 2-7 years, are also imperfectly simulated by GCMs (e.g., Bellenger et al., 2013; Collins et al., 2013), as are the teleconnections that can drive regional precipitation and temperature variability (Sheffield et al., 2013). Since natural variability in observations and historical GCM simulations is not synchronized (e.g., Eden et al., 2012), where regional climate is influenced by low-frequency variability the biases in GCM climate output also can be expected to mimic this natural variability, which has been noted in GCM climate simulations over the U.S. (e.g., Maurer et al., 2013). A correction of GCM output to account for biases in variability on different timescales is warranted where impacts are sensitive to this variability, and has not yet been attempted.

Our first goal is to document how existing bias correction schemes alter the projected climate changes obtained from GCMs. We then propose a method that preserves the model-projected changes. Third, we document model biases as a function of frequency so that the locations and extent to which this is a problem in current state-of-the-art GCMs can be understood. Ault et al. (2012) examined model biases at the interannual and decadal timescales, however we find that bias correction is sensitive to misrepresented variability on shorter timescales as well. Lastly, we present a frequency-dependent bias correction scheme that reduces the problem of frequency-dependent model biases. Maurer et al. (2013) have already demonstrated how model biases can vary over time and at extreme percentiles; this work adds to that list by showing that there are biases in variance at different frequencies.

The rest of this report is structured as follows. In section 2 we describe the observed and model data sources. Section 3 addresses the problem of bias correction altering model-predicted changes, shows the extent to which this happens, and proposes a bias correction scheme that preserves model-predicted mean future changes. Section 4 addresses frequency-dependent model biases, documents the extent to which these are seen in the current generation of global climate models, and proposes a method for correcting these biases. The interaction of frequency-dependent bias correction with standard bias correction is also addressed. A summary and conclusions are given in section 5.

2. Data sources and time periods

2.1 Global climate models

We use daily maximum temperature and precipitation fields from 21 GCMs that participated in the Coupled Model Intercomparison Project, version 5 (CMIP5; Taylor et al., 2012), listed in Table 1. The models used are all those available from the U.S. Bureau of Reclamation (USBR) archive of regridded (1°x1° longitude-latitude) CMIP5 global climate models at the time this work was performed (ftp://gdo-dcp.ucllnl.org/pub/dcp/archive/cmip5/bcca; Maurer et al. 2014). Using the USBR regridded data has several advantages. It means we can build on work already done to obtain the raw climate model fields and regrid the disparate climate model grids to a uniform representation. Additionally, starting from the same regridded data as the USBR archive uses ensures that later work using the bias corrected fields generated here can be directly compared to the existing USBR archive results.

Historical-data are available over the period 1950-2005. Future changes over the period 2006-2099 are simulated using model output from representative concentration pathway 8.5 (RCP8.5) experiments, which correspond to a relatively high emissions scenario (van Vuuren et al. 2011).

2.2 Observations

We used observed daily maximum temperature and precipitation data from Maurer et al. (2002), as updated through 2010 (available from http://www.engr.scu.edu/~emaurer/gridded_obs/index_gridded_obs.html). The ultimate source of this gridded product is the NOAA co-operative observer weather stations, with techniques from the PRISM project (Daly et al. 1994) used to augment observed precipitation values in sparsely instrumented locations. The data are provided on a 1/8° x 1/8° latitude-longitude grid, which was aggregated to the same 1°x1° grid as the global climate model outputs.

2.3 Time periods used

The World Meteorological Organization (WMO) recommends that climatological normals be calculated over 30-year periods (a brief history of climatological normals can be found in Trewin 2007). The U.S. National Oceanic and Atmospheric Administration (NOAA) and National Climatic Data Center (NCDC) do the same (e.g., http://www.ncdc.noaa.gov/oa/climate/normals/usnormals.html). We follow this guidance by bias correcting GCM values to a 30-yr climatological record of observations, and furthermore by bias correcting contiguous 30-yr segments of climate simulations individually. For the future model projections we bias correct the periods 2010-2039, 2040-2069, and 2070-2099 separately. In the results shown below we focus on 2070-2099 as our "future" period. The climatological (historical) period is the last 30 years of the GCMs' historical runs (1976-2005), used for both the models and observations.

3. Preserving model-predicted mean changes

3.1 Overview

In this section we first evaluate the ability of three existing bias correction methods to preserve GCM-predicted future changes in daily maximum temperature and precipitation. We then propose a modification to an existing bias correction method for precipitation that preserves model-predicted mean future changes.

Both temperature and precipitation are examined because they have different spectral characteristics and we evaluate their changes in two contrasting ways: as a difference with temperature (future – historical) but as a ratio with precipitation (future / historical). This is unlike the analysis in Maurer and Pierce (2013), which evaluated precipitation changes as a difference. However, it can be useful to evaluate precipitation changes as a ratio since GCMs may have significant biases in precipitation for a variety of reasons including the inability to adequately resolve topography and its effect on precipitation, for example often extending mountain precipitation too far into a rain shadow. In certain regions and seasons model biases may be several times the local observed climatology, making it sensible to consider model changes as fractional changes relative to the model's own base climatology rather than as differences that are subsequently applied to observed climatology.

In the results shown here all the bias correction techniques are applied to daily values within each month (i.e., all days in January are bias adjusted together, then all days in February, etc.) to account for the cyclostationary nature of climate fields. More sophisticated treatments of this aspect of bias correction can be found in, for example, Piani et al. (2010), Abatzoglou and Brown (2011), and Thrasher et al. (2012). The historical period is 1976-2005 and the future period is 2070-99.

3.2 Effect of quantile mapping on model-predicted changes

In quantile mapping (QM; Wood et al. 2002) a raw GCM value *x* is converted to a bias corrected value $\hat{x}$ according to

$$\hat{x} = F_{sh}^{-1}\left(F_{gh}(x)\right) \quad (1)$$

where, using the notation from Michelangeli et al. 2009, $F(x)$ is the quantile of value *x* in the cumulative distribution function (CDF), $F^{-1}(u)$ is the value in a CDF of quantile *u*, the first subscript is *s* for the station (observed) values and *g* for the GCM (model) values, and the second subscript is *h* for the historical time period. Thus, QM bias corrects a model value by changing it to the observed value at the quantile that the model value falls in the model's historical distribution. The process is illustrated schematically, using CDFs of synthetic gamma distributions to mimic precipitation data, in Figure 1a and the caption thereof. Values off the end of the distribution are handled as described in Wood et al. (2002), i.e. by fitting a Gumbel extreme value distribution to the precipitation values and a Gaussian distribution to the temperature values.

QM has difficulty bias-correcting precipitation if the GCM has more zero-precipitation days than observed since there is no obvious prescription to determine which of the model's too numerous zero precipitation days should be assigned a non-zero value. If zero precipitation remains on these days, the QM bias-corrected time series will have a smaller mean value than observed. However, in practice current GCMs generally have too few zero precipitation values, sometimes referred to as the drizzle problem (e.g., Sun et al. 2006, Dai 2006).

QM's tendency to alter GCM-simulated trends (as noted in the introduction) was not relevant to early applications of QM in hydrological modeling, such as in Wood et al. (2002), which developed and applied QM in the context of seasonal forecasting. When lead times are on the order of a year or less there is no reason to assume a significant shift in the model distributions over the forecast period, so the behavior of QM when the mean changes appreciably was not examined. This is consistent with the fact that Eq. 1 uses only historical information, not referring to the future model-projected distributions in any way. Despite this, QM has frequently been applied to climate change simulations where the mean does change appreciably, for example in multi-decade climate simulations that include anthropogenic changes in greenhouse gasses and aerosol forcing (e.g., Harding et al. 2012).

The tendency for QM to alter model-projected changes is illustrated with the CCSM4 GCM using July daily maximum temperature in Figure 2 and December daily precipitation in Figure 3. Bias correction is applied and model-predicted changes computed using a historical (1976-2005) and future period (2070-2099). Note that some fraction of variance over the future period might arise from the anthropogenically forced trend. The monthly average change between the future and historical periods is computed as a difference for temperature (future – historical) and a ratio for precipitation (future / historical). The top row (right panel) of Figure 2 shows that QM decreases the model-predicted July daily maximum temperature change by more than 2°C in parts of Texas, Florida, and the Southeastern states, and increases it by a similar amount along the California coast. Similarly, Figure 3 shows that QM increases the model-predicted change in December precipitation by 30 percentage points over parts of the Northern Sierra Nevada and Rockies. Smaller changes can also be seen over the upper Midwest and Southeastern states.

Recently, two bias correction methods have been developed that make specific choices for how model-predicted future changes should be treated: the CDF-transform method ("CDF-t"; Michelangeli et al. 2009) and equidistant quantile matching ("EDCDFm"; Li et al. 2010). We next examine these methods to determine whether they preserve GCM-predicted mean changes in temperature and precipitation. As in Maurer and Pierce (2013) we simplify by focusing primarily on the model-predicted change in median value instead of the mean, although it should be kept in mind that changes evaluated in different ways, for example by a least-square trend, are affected by the entire distribution rather than just the change in median, and trends may differ at different quantiles.

3.3 The effect of EDCDFm bias correction on model-predicted changes

EDCDFm (Li et al. 2010) bias corrects a future value *x* that falls at quantile *u* in the future distribution by adding the historical value at *u* to the model predicted change in value at *u*:

$$\hat{x} = F_{sh}^{-1}\left(F_{gf}(x)\right) + x - F_{gh}^{-1}\left(F_{gf}(x)\right) \tag{2}$$

Where variables are defined as in Eq. 1 and the subscript *f* is for the future time period. The process is illustrated schematically in Figure 1b and the caption thereof. When bias correcting a historical run, so that $F_{gf} \equiv F_{gh}$, Eq. 2 reduces to Eq. 1. By definition Eq. 2 preserves the GCM-predicted future change in median value as long as the change is evaluated additively. EDCDFm does not necessarily preserve the model-predicted change in the mean (as opposed to median) since the quantile at which the mean falls can change if the shape of the distribution changes in the future. This does happen; for example, we find that for daily maximum temperature in the

CMIP5 models, changes in the quantile at which the mean falls can be 0.1 or more by the end of this century. However, for daily maximum temperature GCM-predicted changes are generally a weak function of quantile in the neighborhood of the mean value, so EDCDFm preserves the model-predicted change in mean value typically to within a few hundredths of a degree C (e.g., second row of Figure 2). Considering model uncertainty and natural variability this small discrepancy is irrelevant for our applications, but this should be re-evaluated if EDCDFm is applied to another climate variable that undergoes less uniform changes as a function of quantile or more exaggerated changes in the shape of the distribution, which could imply larger changes in the quantile of the mean.

As formulated by Li et al. (2010) and seen in Eq. 2, EDCDFm is an additive bias correction method that preserves model-predicted differences (as opposed to ratios), which is appropriate for temperature. As expected, Eq. 2 does not generally preserve a GCM-predicted fractional changes, i.e., (future model value − historical model value)/(historical model value). At every quantile standard EDCDFm preserves the numerator of this ratio by definition, but in the process of bias correction substitutes the observed value for the historical model value in the denominator, changing the ratio. This is illustrated in the second row of Figure 3. When evaluated multiplicatively using precipitation, EDCDFm alters the model-predicted change by more than 30 percentage points over much of the North/Central U.S. This will happen particularly when there are both large biases and large changes in the upper quantiles of a skewed precipitation distribution.

3.4 The effect of CDF-t bias correction on model-predicted changes

CDF-t bias correction (Michelangeli et al. 2009) finds a transformation that maps the GCM cumulative distribution function (CDF) of a climate variable in the historical period to the observed CDF, then applies that same mapping to the GCM's future CDF, yielding:

$$F_{sf}(x) = F_{sh}\left(F_{gh}^{-1}\left(F_{gf}(x)\right)\right). \tag{4}$$

Here, F_{sf} indicates the CDF of the bias corrected variable in the future. When bias correcting a historical run Eq. 4 reduces to QM in general, although the treatment of values off the end of the distribution (discussed further below) may come into play.

As Figure 1c makes clear, QM and EDCDFm change a model's value while preserving its quantile (a point on the F_{gf} curve is bias corrected by moving it horizontally), while CDF-t changes a model's quantile while preserving its value (a point on the F_{gf} curve is bias corrected by moving it vertically). An alternative, but equivalent, explanation is that Eq. 4 preserves model-predicted changes at quantiles, but unlike EDCDFm the model-predicted change that is preserved is taken from a different quantile than it is applied to (EDCDFm applies the model-predicted change at some quantile to that same quantile).

For example, consider the GCM-predicted change in median value. The EDCDFm bias corrected change in the median value is equal to the model predicted change at a quantile of u=0.5 (the median). However, the CDF-t bias corrected change in the median value is equal to the model predicted change at some other quantile, $u^* \neq 0.5$. It can be shown that $u^* = F_{gh}\left(F_{sh}^{-1}(0.5)\right)$, i.e. the percentile in the model historical distribution of the observed median value. So CDF-t will only preserve model-predicted changes in the median under certain special circumstances, such as when $u^* = 0.5$ (i.e., the GCM predicted the correct median over the historical period in the first place) or if the model-predicted changes are the same at different quantiles (a simple shift in the distribution).

The third rows of Figure 2 and Figure 3 show the effect of CDF-t bias correction on the original GCM-predicted change in temperature and precipitation, respectively. CDF-t alters the model predicted change in temperature by more than 1°C in the upper Midwest. The model-predicted change in precipitation is affected less by CDF-t than by quantile mapping or EDCDFm, but still can change the model prediction by more than 30 percentage points in some locations. In other months the precipitation alterations due to CDF-t bias correction can be as large as those found by quantile mapping and EDCDFm (not shown).

3.5 Bias correction that preserves model-predicted mean changes

Figure 1d, Figure 2, and Figure 3 show that the three bias correction methods considered so far, QM, EDCDFm, and CDF-t, all produce reasonable appearing yet different future CDFs and fields. Lacking theoretical guidance there is no obvious way to choose which method produces the most correct future representation. Li et al. (2010) and Maurer and Pierce (2013) use historical natural variability as a surrogate for forced changes to evaluate the quality of bias correcting future changes, but as discussed in Maurer and Pierce (2013) this approach is limited since natural variability does not necessarily arise from the same physical processes as anthropogenically forced climate change.

Each future field is determined by the assumptions of the bias correction method used to create it. QM assumes that the historical model error in value at a given *value* is preserved in the future (arrow (2) in Figure 1a), EDCDFm assumes that the historical model error in value at a given *quantile* is preserved in the future (Δ in Figure 1c), and CDF-t assumes that the historical model error in *quantile* at a given quantile is preserved in the future (arrow (2) in Figure 1b).

(The “missing” version of this quartet of bias correction methods, which would assume that the historical model error in quantile at a given *value* is preserved in the future, could also be constructed.)

Here we explore an alternative assumption: that the GCM-predicted mean change is preserved in the bias corrected future projections. The advantage of this approach is that downscaled results will then be more consistent with existing GCM analyses such as IPCC (2007, 2013).

EDCDFm already preserves model-predicted changes in temperature (evaluated additively) for all practical purposes so we adopt it for temperature here. However an amended form is required for precipitation since we evaluate its changes multiplicatively. Eq. 2 can be recast as a multiplicative scheme that preserves the model-predicted change as a ratio:

$$\hat{x} = F_{sh}^{-1}\left(F_{gf}(x)\right)\frac{x}{F_{gh}^{-1}\left(F_{gf}(x)\right)}. \qquad (3)$$

In other words, if the predicted GCM value x falls at quantile u, then the bias corrected precipitation value is the historical value at u multiplied by the model-predicted change at u evaluated as a ratio. In fact, Li et al. (2010) do this for a small number (~0.3%) of grid points that otherwise are “problematic” when bias correcting precipitation additively, although in the context of their study they did not explore the implications of Eq. 3 for preserving a model-predicted future precipitation change. Note that Eq. 3 cannot be applied at quantiles where there is no precipitation, in which case the denominator becomes zero. In this event we simply set the model-predicted change ratio to 1.

The treatment of zero-precipitation days is an important consideration for regional climate change (Polade et al. 2014). We calculate a location-specific zero-precipitation threshold for the GCM, τ, such that applying τ makes the model’s number of zero-precipitation days match observations over the historical period. Specifically, $\tau = F_{gh}(q_{\max})$ where $q_{\max}$ is the largest quantile at which $F_{sh}(q_{\max}) = 0$. We require τ >= 0.01 mm/day to avoid the possibility of very small denominators in Eq. 3. Current GCMs tend to precipitate too frequently, often at daily amounts above 0.01 mm, so this limit is rarely invoked. The GCM-predicted future fraction of zero-precipitation days, Z_{gf} , is calculated using τ with the GCM's original (non-bias corrected) future time series. The model data is then bias corrected, and the smallest Z_{gf} fraction of precipitation values are set to zero. In most cases this means that the GCM-predicted change in fraction of zero-precipitation days is preserved in the bias-corrected output. However it sometimes happens that the bias-corrected future time series has a larger fraction of zero precipitation days than Z_{gf}, which is not a correctable bias since there is no way to know which zero-precipitation days should be set to have a positive value.

Model-predicted changes in mean precipitation (evaluated multiplicatively) are generally not preserved using Eq. 3 for the same reasons noted above for the standard additive EDCDFm technique (i.e., changes in the quantile at which the mean falls). Although this results in negligible errors in temperature, precipitation distributions tend to be more skewed than temperature distributions and GCMs can show significantly varying predictions of future change as a function of quantile. Between these two effects Eq. 3 can still alter model-predicted changes in mean precipitation in some seasons and locations.

The model-predicted mean precipitation change (evaluated as a ratio) over the bias correction period can be preserved exactly if the right hand side of Eq. 3 is multiplied by a correction factor K:

$$K = \langle x \rangle / \langle \hat{x} \rangle \tag{4}$$

where brackets indicate time averaging over all days in the appropriate month (since we are implementing bias correction separately for each month). I.e., *K* is the ratio of the mean change in the original GCM to the mean change in the bias corrected GCM. We call the combination of Eqs. 3 (the ratio-preserving formulation of EDCDFm) and 4, together with the treatment of zero-precipitation days described above, the PresRat bias-correction method because it “preserves the ratio,” specifically the mean GCM-predicted future precipitation change evaluated as a ratio. In application the future / historic ratios are first computed at each quantile, then those ratios are scaled by *K*. Figure 1d includes results from PresRat (purple line) applied to the synthetic example data as well as the other methods, for comparison.

The corrections (*K*) that PresRat requires to maintain the model-predicted precipitation change are second order, arising from changes in the percentile at which the mean falls combined with differing model-predicted changes at different percentiles, and so tend to be modest. Figure 4 shows *K* for four different months averaged across all 21 GCMs. In any given month Eq. 3 tends to alter the model-predicted mean change by less than 5% in the majority of the region (white areas of Figure 4); in most other locations the mean is changed less than 20% (light and dark green and yellow areas). In some places though, especially the dry regions of California in summer, PresRat requires substantial corrections to preserve the model-predicted mean change.

By construction, PresRat preserves the model-predicted mean precipitation change exactly in CCSM4 (bottom row of Figure 3). Although it is a minor effect, it is also worth noting that PresRat allows the model to reproduce the observed historical mean in some cases where the GCM has more zero-precipitation days than observed. In this situation the other bias-correction techniques (QM, EDCDFm, and CDF-t) are unable to preserve the historical mean value since there is no way to know which of the too numerous zero-precipitation model days should be assigned a positive precipitation value (see section 3.1). This tends to be a minor effect because in this situation it is the lowest precipitation days that the GCM is missing, and since precipitation tends to have a strongly skewed distribution (especially in dry areas) the smallest precipitation days contribute little to the monthly mean. Even PresRat cannot maintain the model-predicted change if in the future period there is a month with no precipitating days to correct, which does occur in some models for particularly dry locations and months.

PresRat generally preserves mean changes in precipitation while also allowing for changes in the distribution. However a possible concern is that the multiplicative factor K is calculated to preserve the GCM-predicted mean change but alters the bias-corrected future values at all quantiles, not just the mean (i.e., PresRat does not preserve the model-predicted ratio at each quantile after K is applied). The fact that K typically alters the values by less than 5% (Figure 4) should allay this concern to some degree but it is still worth checking explicitly.

An underlying uncertainty is that there is no straightforward approach for evaluating the correctness of future distributions of climate variables. Using changes over the historical period is possible but has the drawbacks noted in Maurer and Pierce (2013). As a practical matter we compare future extreme precipitation values developed by PresRat to those from QM, CDF-t, and EDCDFm. We cannot determine whether or not the distributions produced by PresRat are correct just by comparing them with the distributions of other methods, but it is useful to know how the methods compare.

Figure 5 shows how often each of the bias correction methods (QM, CDF-t, EDCDFm, and PresRat) produces the smallest (rank 1) or largest (rank 4) 95th percentile value of future (2070-99) winter (DJF) daily precipitation in each of the 21 global models. If the 4 bias correction methods tended to produce equal values of the extrema, then on average each method would produce 25% of the values in each rank. Accordingly the values shown in Figure 5 are the difference (in percentage points) from 25%, so that positive values are seen where the bias correction method is producing more values in that rank than the other methods by the end of the century and negative values are seen where the method is producing fewer values of that rank than the other methods. This highlights the dominance of any one method in contributing values of a given rank. CDF-t and PresRat produce similar numbers of rank 1 (smallest) 95th percentile values of future winter precipitation, while EDCDFm tends to produce considerably more rank 4 (largest) 95th percentile values than the other methods. PresRat has a surfeit of rank 2 (middle of the pack) values compared to the other methods. While again emphasizing that we do not know which of these representations is the most correct, we can nonetheless infer that an end-of-century hydrological simulation using EDCDFm or QM bias corrected precipitation is likely to produce more frequent or severe winter flooding events than one using PresRat. A similar analysis for future summer (JJA) precipitation is shown in Figure 6; QM tends to show the largest (rank 4) 95th percentile precipitation values and PresRat the smallest (rank 1) by the end of this century. In other words, assuming that the model historical error at a given value is preserved (as QM does) tends to lead to precipitation values that are more extreme than is consistent with the factor that the GCM indicates the model precipitation will change by.

3.6 Multi-model ensemble results

Since bias correction affects different GCMs differently, it is useful to examine the effect of bias correction on model-predicted changes aggregated across models. Figure 7 shows projected changes (2070-99 relative to 1976-2005) in daily maximum temperature after bias correction minus projected changes in the original GCM, averaged across all 21 GCMs. It is apparent that QM and CDF-t have systematic effects, so an analysis of future temperatures would not be protected against the tendency of these bias correction techniques to alter model-predicted changes even if it used a relatively large ensemble of models. Some models show alterations considerably larger than these mean values, so using a small number of models with bias correction is potentially risky. QM, in particular, exaggerates model-predicted winter warming across much of the north-central U.S. and diminishes summer warming through much of the southeast. This provides justification for some implementations of QM (e.g., Wood et al., 2004) to remove the GCM trend prior to bias correction and replace it afterward. As outlined in Maurer and Pierce (2013), QM's modifications of the GCM trend are related to GCM misrepresentations of variability in the historical model run. CDF-t shows much smaller mean changes, but they still can exceed 0.5 °C in some locations. EDCDFm, by construction, shows very little alteration of mean model-predicted changes in future daily maximum temperature.

Of course, when considering multi-model ensemble averaging the mean result might be near zero but the individual models could have a large spread of values about zero. To determine if this is the case, Figure 8 shows the RMS difference (calculated across the 21 GCMs) between the original model-predicted future (2070-99) change in daily maximum temperature and the change after bias correction has been applied. This largely confirms the interpretation of Figure 7; QM shows the greatest tendency to alter the original model-predicted changes in daily maximum temperature, and CDF-t has both a reduced spread of results (compared to QM) and a mean closer to zero. EDCDFm shows essentially no spread between models.

A similar analysis for future (2070-99) changes in daily precipitation is shown in Figure 9. In the multi-model ensemble average, QM, CDF-t, and EDCDFm all alter the mean GCM-predicted future change in precipitation by more than 30 percentage points in some times and locations, although generally CDF-t imposes smaller changes than QM and EDCDFm. QM has a tendency to make the model-predicted changes wetter (cf. Maurer and Pierce 2013).

CDF-t tends to make forecast changes drier, for reasons that can be understood in terms of Figure 1c. To produce a point on the bias corrected future distribution it is necessary that the model historical value at the quantile being bias corrected fall within the range of observed values, as indicated by vector (2) in Figure 1c. E.g., if vector (2) were progressively moved to the right in Figure 1c, it can be seen that no historical values greater than X=16 mm/day (the maximum observed value in Figure 1c) could be bias corrected. In this event, following Michelangeli et al. 2009, the correction used is that found at the maximum valid historical value. However the GCM precipitation simulations tend to display two attributes: 1) They over-predict precipitation in dry areas, so the model CDFs are shifted to the right of the observed CDF (as depicted in Figure 1c); 2) The most extreme precipitation events increase preferentially more than others (e.g., IPCC 2007, 2013). In these situations CDF-t tends to be forced to use the maximum valid correction, which falls at a lower quantile, and so misses the preferential increase in the very highest quantiles. Figure 9 shows that this is only a modest tendency outside of dry summer California/Great basin months, but could be a consideration for regional flooding studies.

EDCDFm has mixed effects but makes the simulations strongly wetter in winter in the Rocky Mountains and Great Basin, when much of those areas receive the bulk of their annual precipitation, as well as the upper Midwest. In general EDCDFm will make predicted precipitation changes wetter in locations where the GCM simulates a wetter climatology than observed since a fixed model change (in the quantile) is being applied to a smaller historical base value. By construction PresRat has little effect on the model-predicted mean change in future precipitation, although the effect is not zero (in contrast to the results for EDCDFm with daily maximum temperature) because a few of the models do not have enough precipitating days to be corrected in certain months. This is particularly prevalent in California in July, where a number of the GCMs have no July precipitation that can be altered by the bias-correction scheme. The RMS spread of results across models (Figure 10) shows roughly comparable values for QM and EDCDFm and nearly as much for CDF-t, while PresRat has much less spread. As found in the mean results, the locations where PresRat does show model spread is due to occasional models predicting too little precipitation in some month for a correction to be applied.

3.7 Summary: preserving model-predicted mean changes

The QM and CDF-t bias correction methods generally alter model-predicted mean changes in daily maximum temperature and precipitation. EDCDFm, however, effectively preserves model-predicted changes in mean daily maximum temperature. PresRat (which is a new extension of EDCDFm to preserve ratios, add a zero-precipitation threshold, and implement a correction factor) preserves model-predicted future changes in precipitation (evaluated as a ratio) as long as there exist precipitating days in the GCM simulation that can be corrected. This is accomplished with only modest correction factors (generally less than 5%, the notable exception being the very dry California summers). The extreme values produced by PresRat are mostly consistent with the extreme values from the other bias correction methods, though it tends to produce fewer of the highest 95th percentile values than EDCDFm (in winter) or QM (in summer).

In summary, both temperature and precipitation can be bias corrected using methods that preserve global climate model-predicted future mean changes. Doing so would help minimize confusion and inconsistent results between downscaled regional climate simulations and global model analyses, such as represented by the IPCC analyses (2007, 2013). The advantage of this approach over that taken in Wood et al. (2002), where the trend is removed, bias correction performed, and the interpolated trend re-introduced, is that the model-predicted changes themselves can be downscaled rather than being only interpolated GCM fields.

4. Frequency Dependent Bias Correction

4.1 Overview

The previous section examined the effect of bias correction on GCM-predicted mean changes at long time scales (decades). In this section we address more general question of what model biases may be present across the gamut of timescales and how to address them. Quantile-based bias correction methods such as QM, EDCDFm, CDF-t, and PresRat already alter the variance spectrum of the GCM's time series if certain quantile values do not appear at random intervals in the time series, but rather preferentially at certain frequencies. For example the highest quantiles of California precipitation generally appear in winter, so the proportion of variance in the annual cycle will typically be altered by bias correction. However this effect is modest, as will be shown quantitatively below.

The frequency-dependent bias correction method developed here is designed to systematically alter the shape of the GCM's spectrum to better match observations without changing the overall variance. As such, it is intended to be applied as an additional processing step after standard bias correction has already adjusted the overall variance.

Details are given in the following sections, but in broad terms, the frequency dependent bias correction proceeds as follows. First, the variance spectra of the observations and model are calculated. The model variance error as a function of frequency is then computed as the ratio of the spectral values, (model/observed). To correct these errors the model time series is Fourier transformed to frequency space, then the amplitude of the Fourier components are adjusted so that the distribution of variance across frequencies better matches observations. The Fourier components are then inverse transformed back into a time series.

Much of the following material is devoted to examining the result to make sure the process improves the model simulation rather than degrading it. However, one caveat is that the spectral approach used here does not consider frequency-dependent biases in different seasons or months, but instead only as a collective whole over the entire time period. This potentially means that it is not feasible to expect a removal of biases across all timescales of interest by this technique.

4.2 Spectral Methodology

Since we bias correct the future model projections in 30-yr periods (section 2.3), the PresRat method outlined in section 3.4 will preserve model-predicted mean changes at periods of 30 years and longer in the future projection. Accordingly, when we consider frequency-dependent bias corrections we need only include, at most, periods from two days (the Nyquist frequency given the daily model output) to 30 years. This interval will be further refined below in light of our spectral analysis technique. Model predicted changes at these frequencies can arise from natural internal climate variability, anthropogenic causes, or both.

Numerous techniques are available to compute variance spectra (for a review, see Ghil et al. 2002). Many of the newer methods have been developed to identify narrow-band signals against a background of noise. However, in this work we are also concerned with the power in the broad parts of the spectrum that might in other applications be considered simply “noise”. This variability represents weather and climate fluctuations that affect hydrology and ecosystems across a wide range of time scales, so we seek as realistic a simulation of these fluctuations as possible. Accordingly we use relatively wide bandwidths in this work and employ the Jenkins and Watts (1969) method of computing variance spectra as the Fourier transformation of the autocovariance function. We require at least 40 degrees of freedom in the spectral estimates, which given 30 years of daily data and a Parzen lag window, means truncating the autocovariance function after 1020 lags (Jenkins and Watts 1969). Following the Jenkins and Watts recommendations the number of frequencies is set to twice the number of lags (2040), so the first non-zero frequency corresponds to a period of ~11 yrs. Longer periods are unresolved, and the frequency-dependent bias correction does not alter their relative proportion of variance.

With over 2000 frequencies spanning from 2 days to 11 years it is useful to reduce the number of frequencies at which the model error is corrected to avoid spurious over-fitting. Accordingly, the frequency-dependent model errors are calculated in a reduced set of 100 frequency bins of equal width in the logarithm of frequency. This means that higher frequency bins have multiple samples, as shown in Figure 11, with more than 5 samples per bin at periods shorter than ~80 days (purple lines). The binning therefore reduces the uncertainty in the spectral estimates for periods shorter than ~80 days. The average value of the spectrum in a bin is estimated using monotonic cubic splines (Fritch and Carlson 1980) to avoid abrupt changes in the estimate depending on whether a frequency point is barely included or excluded from a bin.

Von Storch and Zwiers (2001) note the problems in interpreting spectral plots on a logarithmic frequency axis, since the displayed area under the spectrum is no longer proportional to the variance. It is possible to maintain the property of being a spectral density if the spectral value is multiplied by frequency, or if the plotted values are integrated (as opposed to averaged) across constant widths of the logarithmic frequency axis. However these approaches change the angle of a plotted spectrum (for example, a white spectrum is then no longer flat), which can be confusing. To avoid this potentially misleading situation, values shown here are simply averaged in frequency so that the spectra appear similar to what is typically found in the literature (i.e., a white spectrum is flat).

4.3 Frequency dependent model errors

Figure 12 shows maps of the observed (1976-2005) distribution of variance in daily maximum temperature across frequencies (labeled using equivalent periods) and the multi-model ensemble errors in representing this distribution in the same period. The left column shows observations (% of total variance), the middle column shows the multi-model mean error (%) with respect to the observations, and the right hand column shows multi-model RMSE (%; i.e., at each point, the spread of values across the 21 models). The frequency-dependent bias correction is based on normalized spectra (spectral values divided by the variance of the original time series) so that it leaves the overall variance unaltered. Therefore at every location the values in the left hand column summed across frequency bands total 100%. For example, the top left panel of Figure 12 shows that in the region from western Texas north to western Kansas more than 10% of the total variance falls in the 2-10 day band, while in the region immediately to the west less than 4% does.

As expected, Figure 12 shows that the annual cycle dominates the daily maximum temperature variability over almost all of the conterminous U.S., containing on average 62% of the total variance. The exception is in locations along the California coast, where shorter period variability makes a much larger contribution to the overall variance than found elsewhere.

Reinforcing the notion that bias correction might usefully be applied as a function of frequency, the multi-model aggregate profile of model errors of daily maximum temperature (middle column) varies considerably across the spectral range. Over much of the domain there is a tendency for models, on average, to allocate less of the total variance to periods shorter than 3 months than is observed, particularly in the 10-30 day band where the mean error is −9%. RMS errors at periods shorter than the annual cycle are typically on the order of 10-15% of observed variance in those frequency bands, which implies that the mean error is relatively consistent across the models. The proportion of variance in the annual cycle is represented with virtually no mean error and a very small spread across models.

A deficiency in daily maximum temperature variability at periods shorter than the annual cycle combined with an accurate representation of the annual cycle implies that periods longer than the annual cycle must be receiving proportionally too much variance, which is confirmed by Figure 12b. Variability that occurs at periods longer than 30 months has, on average, proportionately ~40% more variance than observed, and the spread across models is large, with RMS errors of ~60%. However it should be kept in mind that the fraction of total variance contained in these long time scales is quite small (< 1% for all timescales longer than 30 months).

Figure 13 shows the same frequency-dependent analysis using daily precipitation. In contrast to daily maximum temperature, over most of the conterminous U.S. the shortest periods (2-10 days) contain the majority of the variance (on average, 62%). The exception is the wet parts of the west coast, where 10-30 day and longer period variability is nearly as important and the annual cycle contains > 7% of the total variance, more than twice the average at that frequency over the domain. The models as a group tend to simulate the short-period (2-10 day) fraction of total variance reasonably well, with a modest (5-10%) mean bias towards too much short-period variability along the west coast and upper Midwest and too little around Texas, Oklahoma, and the Gulf coast. Figure 13b shows that model-simulated precipitation variability at periods of 30 months or longer accounts for an anomalously large proportion of the total variance in the southeastern U.S., and an anomalously small proportion in the Pacific Northwest. Rupp et al., (2013) also found that models overestimate temperature variance at timescales longer than a year and underestimate precipitation variance at timescales longer than a year in the Pacific Northwest, USA. Disagreements across the models are large at these longer periods.

4.4 Correcting frequency-dependent model errors

4.4.1 Method for frequency-dependent bias correction

To correct the frequency-dependent model biases at some location, the ratio σ of the model's variance spectrum to the observed variance spectrum is computed in each of the 100 logarithmically spaced frequency bins. This step is analogous to calculating the ratio of model to observed values at each quantile in the cumulative distribution function in the PresRat method. Both spectra are computed over the historical climatology period, 1976-2005. The original model time series is then transformed to frequency space and, to bias correct the model series, the amplitude of the Fourier components are multiplied by $\sigma(f)^{-1/2}$ (the square root accounts for the fact that variance is proportional to the amplitude of the Fourier components squared). The result is then transformed back to the time domain. As typical in statistical bias correction techniques, σ is calculated over the control period and applied to both the control and future periods. This assumes that the statistics of the model error as a function of frequency do not change, but does not prevent a model from changing its future spectrum, either the overall amplitude of variance or the distribution of variance across frequencies; it just means that any model-predicted changes will be relative to the corrected model spectrum.

4.4.2 Example results for daily maximum temperature

As noted above, standard bias correction techniques such as QM, EDCDFm, and CDF-t alter the spectra of the time series they are applied to. Thus, in order to clearly demonstrate the effect of the frequency-dependent bias correction by itself, we first present results using only the frequency-dependent bias correction. We then show combined results using the frequency-dependent bias correction applied in conjunction with standard bias correction.

Typical results of the frequency-dependent bias correction using daily maximum temperature from the CCSM4 GCM are illustrated in Figure 14. The left column shows normalized spectra of observations (red), the original model (blue), and the model after frequency-dependent bias correction (green dots). For each panel values are taken from the location indicated by the purple 'x' on the inset map and shown in the panel's title (longitude, latitude). The right column shows the ratio of the model's spectral value to the observed value, both before (blue) and after (green dots) the frequency-dependent bias correction is applied.

It is useful to define a root mean squared error metric appropriate for ratios of the spectral values, which we designate as log-RMSE to differentiate it from standard RMSE measures that are appropriate to differences rather than ratios. Let $\epsilon = \ln \sigma$, then

$$\text{log-RMSE} \equiv \exp(\sqrt{\langle \epsilon^2 \rangle} - 1 \qquad (5)$$

where the angle brackets indicate the mean over the logarithmically spaced frequency values. This expression treats equal ratios of error equally (i.e., the model having twice the observed variance produces the same error as the observations having twice the model's variance), and the final −1 makes a perfect result (model variance equal observed, so σ= 1) give a log-RMSE of 0. In general, if the model values are incorrect (on average across log-spaced frequencies) by a factor of σ then the log-RMSE is $\sigma - 1$. These log-RMSE values are indicated in the right column of Figure 14. When we refer to log-RMSE below, we specifically mean the model's error in reproducing the distribution of variance across frequencies, as illustrated in Figure 14.

In some locations, such as the San Francisco region (top row of Figure 14), the ratio of the model variance to observed exhibits a notable slope (top right panel) which indicates that the model frequency errors are a systematic displacement of variance, depriving high frequencies and enriching low frequencies. At all locations the frequency-dependent bias correction improves the model's representation of how variance is distributed across frequencies. The log-RMSE typically drops by about a factor of 5 as a result of the correction. Some residual error remains due to the approximate nature of corrections calculated using discretely sampled data on a finite interval.

4.4.3 Example results for daily precipitation

Precipitation is more difficult to correct than temperature because it cannot have negative values, which limits the adjustments that the frequency-dependent bias correction can produce. There are also many days with zero precipitation, which we do not alter. In fact, to avoid potential problems with exacerbating models' drizzle problem, whereby they produce too many days of light precipitation (Sun et al. 2006; Dai 2006), we leave unmodified any model precipitation values less than 1 mm/day. Particularly in dry areas this can leave few days for the frequency-dependent bias correction to operate upon.

Precipitation results at a few example locations are shown in Figure 15 using CCSM4. It is apparent that the frequency dependent bias correction is less effective at adjusting precipitation than temperature. For example log-RMSE values only decrease by a factor of 1.3 to 2 rather than a factor of 5, as found for temperature. But although the corrections are relatively modest, they result, uniformly, in the direction of decreasing model error and so are helpful.

4.4.4 Multi-model ensemble average results

The multi-model ensemble average log-RMSE for daily maximum temperature is shown in the top row of Figure 16 both before (left column) and after (middle column) the frequency-dependent bias correction. The models systematically disagree with the observations, particularly along the west coast and in a band extending north from northern Texas. Before the frequency-dependent bias correction the mean log-RMSE of daily maximum temperature error of 0.50 indicates that the models are, on average across models, locations, and frequencies, off by a factor of 1.50 (i.e., by 50%) in their representation of the variance in any particular frequency band. After frequency-dependent bias correction the log-RMSE drops to 0.11 (nearly a factor of five decrease), indicating that the corrected models are only off by a factor of 1.11 on average.

Results for daily precipitation are shown in the bottom row of Figure 16. The models as a group tend to do worse in the Rocky Mountains and Great Basin than in most other locations. The mean log-RMSE for precipitation is approximately the same as for daily maximum temperature. However, as expected for the reasons given above, precipitation is less easily corrected than temperature; the log-RMSE for precipitation drops by less than a factor of 2 after the frequency-dependent bias correction. The pattern of log-RMSE precipitation errors after correction (Figure 16, lower center) primarily reflects the rate of occurrence of days with > 1 mm/day of precipitation (our threshold for correction). The final results are best where the most potentially correctable precipitation values exist and worst where there are few correctable days. However this does not completely explain the pattern; there are residual differences that reflect the seasonality and other aspects of the local precipitation distribution.

An important consideration is whether the frequency-dependent bias correction makes the representation of variance with frequency worse in some locations despite being better on average. This is addressed by the histograms in Figure 16 (right column), which show the difference between each location's corrected and original log-RMSE, pooled across every location and every model. On average the frequency-dependent bias correction decreases the log-RMSE for daily maximum temperature by 0.39, and this is accomplished without making any locations worse (no positive values are seen in the histogram). Even for precipitation, which shows less improvement (decrease of log-RMSE by 0.21) from the frequency-dependent bias correction than temperature, the correction virtually always decreases the log-RMSE (lower right panel of Figure 16).

4.4.5 Magnitude of the corrections

It would be potentially troubling if the modifications to the time series made by the frequency-dependent bias correction were too large. Histograms of the amplitude of the corrections pooled across all models and locations are shown in Figure 17. Any day's maximum temperature is changed less than 3°C about 95% of the time, although rarely the changes can exceed 4°C. The change in precipitation is less than 40% or 1.5 mm day^{-1} about 95% of the time, although on rare occasion can be more than 50% or 2.5 mm day^{-1}. Since the frequency-dependent bias correction operates on normalized spectra, altering the distribution of variance across frequencies without altering the overall variance, the mean changes are approximately zero for both temperature and precipitation.

Time series of daily maximum temperature before and after the frequency-dependent bias correction are shown in Figure 18, using year 2000 from the CCSM4 GCM as an example. For plotting purposes the annual mean value (shown in the upper right part of the panel) has been removed. The changes to the time series made by the frequency-dependent bias correction are small compared to the synoptic and annual timescale fluctuations in the time series. Similar time series for daily precipitation are shown in Figure 19. Again, the modifications made by the frequency-dependent bias correction are modest compared to the daily variability. The relatively constrained nature of the changes imposed by the frequency-dependent bias correction shows that the improvement in spectral properties afforded by the frequency-dependent bias correction does not come at the expense of creating an unrealistic time evolution in the final fields.

4.4.6 Combined effects of standard and frequency-dependent bias correction

In this section we explore the effect of frequency-dependent bias correction applied in conjunction with standard bias correction. Only the historical period is considered since we compare to observations. This in turn restricts this analysis to QM since the other bias correction methods differ from QM exclusively in the future period.

Figure 20 shows the multi-model mean log-RMSE across all the climate models for daily maximum temperature, both before any bias correction has been applied (panel a) and after various combinations of QM and frequency-dependent bias correction have been applied (panels b-e). QM by itself decreases the mean log-RMSE by about 0.15, compared to the frequency-dependent bias correction, which decreases the mean log-RMSE by about 0.39. So although QM helps make the models' distribution of variance across frequencies closer to observed, the improvement is considerably smaller than that achieved by the frequency-dependent bias correction. Panels d and e show the results when applying the frequency-dependent bias correction either before or after QM. On average results are slightly better when the frequency-dependent bias correction is applied after QM, although the difference is small.

Figure 21 shows the same analysis for daily precipitation. QM does a slightly poorer job of improving the models' depiction of variance across frequency than seen when operating on daily maximum temperature (a reduction in log-RMSE of 0.12 for precipitation vs. 0.15 for temperature). However, as noted above, the frequency-dependent bias correction is not as effective in correcting precipitation as temperature (log-RMSE drops by 0.21 for precipitation vs. 0.39 for temperature), although it is still provides almost twice the reduction in log-RMSE than found in QM alone (0.21 vs. 0.12). As found for daily maximum temperature, slightly better results are obtained when QM is followed by the frequency-dependent bias correction rather than the opposite order.

It was previously noted (Figure 16) that one desirable aspect of the frequency-dependent bias correction is that no location's agreement with observations becomes worse as a result of the method being applied. Figure 22 shows a similar analysis for daily maximum temperature (top row) and precipitation (bottom row) using various combinations of QM and frequency-dependent bias correction. QM degrades the agreement between the model and observations in how variance is distributed across frequencies at about 9.6% of the locations (pooled across all models) for temperature and 23% for precipitation. Of course QM was not designed to take into account the variance spectrum of the simulation so this is not a surprising result, but it is nonetheless worth pointing out this previously unidentified drawback of QM. When frequency-dependent bias correction is followed by QM (bottom right panel), 4.5% of the precipitation locations show worse agreement with observations than the original model even though the mean result is to improve the agreement. However when the order of operations is reversed, so that QM is followed by frequency-dependent bias correction, only 1.3% of the precipitation locations show a worse agreement with observations than found in the original model and no locations show a worse agreement for daily maximum temperature. These findings, along with the results from Figure 20 and Figure 21 that show a small but consistent superiority when applying QM before the frequency-dependent bias correction, are the reason we perform the operations in this order.

Although these results show that it is better to apply the frequency-dependent bias correction after QM, a point of concern is what effect this might have on the quantile matching bias correction that QM performs. Does the frequency-dependent bias correction significantly degrade the correspondence between modeled and observed quantiles that QM imposes? This is evaluated in Figure 23, which shows quantile-quantile plots comparing the quantile at which a value falls in the observed distribution to the quantile at which the same value falls in the models' distributions. Plotted values are pooled across all models and locations. If the models had a perfect representation of the observed distribution, then all the model values would fall along a straight line with slope of 1 (dashed green line in Figure 23). The box and whiskers in the figure show the distribution of model values that are found for a given observed quantile. For example, the upper left panel of Figure 23 shows that the median (0.50 quantile) observed value of daily maximum temperature is, in the median, found at the 0.55 quantile in the models, so the models as a group have a slight cold bias relative to the observations. Half the time the observed median value is found between the 0.50 and 0.60 quantile in the models; and 90% of the time the observed median value is found between the 0.45 and 0.70 quantile in the model. 5% of the time the observed median value falls either below the 0.45 quantile or above the 0.70 quantile in the model.

Figure 23 shows that, viewed across their CDFs, the models do better simulating the distribution of daily maximum temperature than precipitation; at least 25% of the models simulate the observed quantile of daily maximum temperature correctly, no matter what observed quantile is considered. For precipitation however, notably less than 25% of the models manage to simulate the observed percentile correctly at quantiles < 0.5, and at the lowest quantile plotted less than 5% of the models are able to simulate the observed percentile. The positive precipitation bias at low quantiles is consistent with the models’ drizzle problem.

For our purposes, the left two columns of Figure 23 shows that the frequency-dependent bias correction does not systematically alter the shape of the model distributions, which is by design since the method is intended to leave the overall variance unchanged. When QM is applied, either before frequency-dependent bias correction or after (right two columns of Figure 23), the agreement between observed and modeled quantiles is quite good. This is an outcome of QM by construction, and the frequency-dependent bias correction changes that result only a little.

Overall we conclude that the frequency-dependent bias correction does not inflict additional problems to the resultant adjusted model output. Furthermore, it is useful to apply since it increases the average agreement between the observed and modeled distribution of variance across frequencies without degrading the agreement at any location. It accomplishes this with relatively small and symmetric corrections (typically < 3°C or 2 mm/day) without imposing spurious behavior in time or diminishing the agreement between modeled and observed quantiles that QM imposes.

5. Summary and Conclusions

GCMs generally produce biased simulations of variables such as temperature and precipitation. It is necessary to remove these biases before using the model-simulated fields in applications that have non-linear sensitivities to biases, such as land surface or hydrological modeling. Accordingly, a bias correction step is often performed on GCM fields before use in such applications.

One problem with bias correction methods such as quantile mapping (QM; e.g., Wood et al. 2002) and the CDF-transform method (CDF-t; Michelangeli et al. 2009) is that they alter GCM-predicted mean future changes, evaluated here as 2070-99 relative to 1976-2005. Compared to the original changes produced by an ensemble of 21 GCMs with the RCP 8.5 anthropogenic greenhouse gas and aerosol scenario, QM produced warmer future daily maximum temperatures by up to 2°C across much of the upper Midwest, California coast, and Northern Rockies in January, and cooler daily maximum temperatures by up to 2°C across much of the southeastern part of the U.S. in July. CDF-t showed smaller alterations of up to 0.5°C, but they may still have consequence because they tend to persist throughout the year. When evaluated as a multiplicative change in precipitation, QM and the equidistant CDF matching method (EDCDFm; Li et al. 2010) produced wetter conditions than projected by the original global models by up to 30 percentage points across the upper Midwest and Northern Rockies in January, while CDF-t produced drier conditions by up to 20 percentage points in the Southwest U.S. in summer. These changes are large enough to make a practical difference in the results of climate impact studies, which is problematic given their widespread usage and because the magnitude of changes imposed through bias correction can be of the same order of magnitude as the model predicted changes by the end of the century. Moreover, because analyses of the projected climate changes in the original GCMs are widespread (e.g., IPCC 2007, 2013), alterations to the GCM trends may lead to inconsistencies and confusion in bias-corrected regional studies.

In the first part of this work we have demonstrated a methodology that uses existing and modified techniques to maintain model-projected climate changes even when bias correcting the global model data.

Under the assumption that bias correction should preserve the model projected future change, EDCDFm works very well for temperature projections. For precipitation projections we have introduced an extension to EDCDFm that we term PresRat, which "preserves the ratio" of future changes rather than the difference, includes a zero-precipitation threshold that makes the modeled number of zero-precipitation days match observations, and adds a correction factor that is typically < 5%. PresRat generally maintains model-predicted changes in daily precipitation. However none of the bias correction techniques, PresRat included, can preserve the model-predicted precipitation change in cases where locations that are so dry there are insufficient precipitation days to bias correct (which is rare, but does happen in some models during the dry months).

In the second part of the study we extend our examination of model biases from trends to the more general issue of the models' representation of variance across a range of timescales, and introduce a frequency-dependent bias correction method that can address inaccuracies in the GCM simulations. A comparison with observations showed that as a group, the 21 GCMs apportion too little variability of daily maximum temperature to times scales between 10 and 90 days and too much to time scales longer than 30 months. The models' simulation of daily precipitation variability was more mixed, but at long timescales (> 30 months) they show more variability than observed in the Gulf coast region and less than observed in the Pacific Northwest.

We showed that the models' simulation of variance as a function of frequency can be improved by a frequency-dependent bias correction, which is implemented as digital filter in the frequency domain. Before the frequency-dependent bias correction the model simulations tend to err in their estimate of the frequency distribution of total daily maximum temperature variance by a factor about 1.5, RMS averaged across log-spaced frequencies. After the frequency-dependent bias correction the RMS error drops to a factor of 1.11. Precipitation cannot be corrected as easily as temperature since locations typically have numerous zero-precipitation days, but the frequency dependent bias correction stills decreases the RMS error from a factor of 1.49 to 1.28. These improvements are accomplished with relatively modest alterations to the original values, typically < 3°C in daily maximum temperature and < 1.5 mm/day in daily precipitation.

The frequency-dependent bias correction improves the models' simulation of variance as a function of frequency about twice as much as standard bias correction. Additionally the frequency-dependent bias correction makes no locations worse, while standard bias correction degrades the simulated distribution of variance across frequencies at about 9.6% of the gridpoints (pooled across all 21 global models) for daily maximum temperature and 23% for precipitation. Applying the frequency-dependent bias correction subsequent to standard bias correction both increases the models' mean agreement with observations substantially (better than either technique applied alone) and reduces the fraction of degraded gridpoints to 0.0% for daily maximum temperature and 1.3% for precipitation.

Important questions about bias correction remain. This study has not addressed whether bias correction *should* be applied at any particular location given that model-observational disagreements are influenced by natural climate variability, which can be large and affect climate means over years to decades (e.g., Maraun et al. 2010; Deser et al. 2012). Likewise, it is not clear if models should be bias corrected to a particular period that tree ring or other paleoclimate evidence suggests is atypical. Although these are interesting questions, in this work we have followed the common practice of applying bias correction to the GCMs at all locations to bring them into agreement with a pre-selected recent climatological period.

Another problem with bias correction techniques that is not addressed here is that a model with a seasonal cycle of precipitation that is greatly different from observations might not preserve the GCM-predicted annual change even if all precipitation trends are preserved at the monthly time scale. This reinforces the fact that although bias correction can help make the statistics of temperature and precipitation fields from a global climate simulation more like observations, it is possible for some models in some regions to produce such a poor simulation that bias correction has little meaning. Even before bias correction care should be taken to ensure that GCMs used in a regional climate impact study capture the relevant physical processes to begin with. For example, a GCM that lacks an ENSO cycle or seasonal monsoon flow can be bias corrected and downscaled like any other model, but the result will have little meaning in areas that are influenced by ENSO or monsoonal flow.

In the end, as global climate model results continue to be applied to investigate phenomena that are sensitive to model biases, bias correction will become an ever more important step. The bias correction methods outlined here can improve these simulations, giving a clearer picture of future climate conditions for a variety of applications.

Acknowledgements

This work was sponsored by the California Energy Commission under contract CEC-500-10-041. Additional support for DWP and DC came from the USGS through the Southwest Climate Science Center, and from NOAA through the California Nevada Climate Applications Project (CNAP) Regional Integrated Science Applications (RISA) program. We also thank the U.S. Bureau of Reclamation for making available the library of regridded climate projections at http://gdo-dcp.ucllnl.org.

References

Abatzoglou, J. T. and T. J. Brown, 2012: A comparison of statistical downscaling methods suited for wildfire applications. Int. J. Climatol., 32, 772-780.

Ault, T. R., J. E. Cole, and S. St. George, 2012: The amplitude of decadal to multidecadal variability in precipitation simulated by state-of-the-art climate models. Geophys. Res. Lett., 39, L21705, doi:10.1029/2012GL053424.

Bellenger, H., Guilyardi, E., Leloup, J., Lengaigne, M., and Vialard, J., 2013: ENSO representation in climate models: from CMIP3 to CMIP5, Clim. Dyn., 1-20, 10.1007/s00382-013-1783-z.

Collins, M., AchutaRao, K., Ashok, K., Bhandari, S., Mitra, A. K., Prakash, S., Srivastava, R., and Turner, A., 2013: Observational challenges in evaluating climate models, Nature Clim. Change, 3, 940-941, 10.1038/nclimate2012.

Dai, A., 2006: Precipitation characteristics in eighteen coupled climate models. J. Climate, 19, 4605-4630.

Daly, C., R. P. Neilson, and D. L. Phillips, 1994: A Statistical-Topographic Model for Mapping Climatological Precipitation over Mountanious Terrain. J. Appl. Meteor., 33, 140-158.

Deser, C., Knutti, R., Solomon, S., and Phillips, A. S., 2012: Communication of the role of natural variability in future North American climate, Nature Clim. Change, 2, 775-779.

Eden, J. M., Widmann, M., Grawe, D., and Rast, S., 2012: Skill, Correction, and Downscaling of GCM-Simulated Precipitation, J. Climate, 25, 3970-3984, 10.1175/jcli-d-11-00254.1.

Fritsch, F. N. and R. E. Carlson, 1980: Monotone piecewise cubic interpolation. SIAM J. Numer. Anal., 17, p. 238-246.

Ghil, M., M. R. Allen, M. D. Dettinger, K. Ide, D. Kondrashov, M. E. Mann, A. W. Robertson, A. Saunders, Y. Tian, F. Varadi, and P. Yiou, 2002: Advanced spectral methods for climatic time series. Rev. Geophys., 40, doi:10.1029/2000RG000092.

Gudmundsson, L., J. B. Bremnes, J. E. Haugen, and T. Engen-Skaugen, 2012: Technical Note: Downscaling RCM precipitation to the station scale using statistical transformations – a comparison of methods. Hydrol. Earth Syst. Sci., 16, 3383-3390, doi:10.5194/hess-16-3383-2012.

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, J. Hydrometeorology, 12, 556-578, 10.1175/2011jhm1336.1.

Harding, B. L., A. W. Wood, and J. R. Prairie, 2012: The implications of climate change scenario selection for future streamflow projection in the Upper Colorado River Basin. Hydrol. Earth Syst. Sci., 16, 3989-4007.

IPCC, 2007: Climate Change 2007: Impacts, Adaptation, and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, M. L. Parry, O. F. Canziani, J. P. Palutikof, P. J. van der Linden, and C. E. Hanson, Eds., Cambridge University Press, Cambridge, UK, 976 pp.

IPCC, 2013: Climate Change 2013: The physical science basis. Working group I contribution to the IPCC fifth assessment report, T. Stocker, Q. Dahe, and G-K Plattner, coordinating lead authors. Available from http://www.ipcc.ch/report/ar5/wg1/.

Jenkins, G. M., and D. G. Watts, 1969: Spectral analysis and its applications. Holden-Day, 525 pp.

Li, H., J. Sheffield, and E. F. Wood, 2010: Bias correction of monthly precipitation and temperature fields from Intergovernmental Panel on Climate Change AR4 models using equidistant quantile matching. J. Geophys. Res. Atmos., 115 (D10101), doi:10.1029/2009JD012882.

Maraun, D., Wetterhall, F., Ireson, A. M., Chandler, R. E., Kendon, E. J., Widmann, M., Brienen, S., Rust, H. W., Sauter, T., Themeßl, M., Venema, V. K. C., Chun, K. P., Goodess, C. M., Jones, R. G., Onof, C., Vrac, M., and Thiele-Eich, I, 2010.: Precipitation downscaling under climate change: Recent developments to bridge the gap between dynamical models and the end user, Rev. Geophys., 48, RG3003, 10.1029/2009rg000314.

Maraun, D., 2013: Bias Correction, Quantile Mapping, and Downscaling: Revisiting the Inflation Issue, J. Climate, 26, 2137-2143, 10.1175/jcli-d-12-00821.1.

Maurer, E. P., A. W. Wood, J. C. Adam, and D. P. Lettenmaier, 2002: A long-term hydrologically based dataset of land surface fluxes and states for the conterminous United States. J. Climate, 15, 3237-3251.

Maurer, E. P., L. Brekke, T. Pruitt, and P. B. Duffy, 2007: Fine-resolution climate projections enhance regional climate change impact studies. EOS, Transactions of the American Geophysical Union, 88, 504, doi:10.1029/2007EO470006.

Maurer, E. P., H. G. Hidalgo, T. Das, M. D. Dettinger and D. R. Cayan, 2010: The utility of daily large-scale climate data in the assessment of climate change impacts on daily streamflow in California. Hydrol. Earth Syst. Sci., 14, 1125-1138, doi:10.5194/hess-14-1125-2010.

Maurer, E. P., T. Das, and D. R. Cayan, 2013: Errors in climate model daily precipitation and temperature output: time invariance and implications for bias correction. Hydrol. Earth Syst. Sci., 17, 2147-2159. Doi:10.519/hess-17-2147-2013.

Maurer, E. P., and D. W. Pierce, 2013: Bias correction can modify climate model-simulated precipitation changes without adverse effect on the ensemble mean. Hydrol. Earth Syst. Sci. Discuss., 10, 11585-11611, doi:10.5194/hessd-10-11585-2013.

Maurer, E. P., Brekke, L., Pruitt, T., Thrasher, B., Long, J., Duffy, P. B., Dettinger, M. D., Cayan, D., and Arnold, J, 2014: An enhanced archive facilitating climate impacts analysis, Bulletin of the American Meteorological Society, doi:10.1175/BAMS-D-1113-00126.00121 *in press*.

Michelangeli, P.-A., M. Vrac, and H. Loukos, 2009: Probabilistic downscaling approaches: Application to wind cumulative distribution functions. Geophys. Res. Lett., 36, L11708, doi:10.1029/2009GL038401.

Panofsky, H. A. and Brier, G. W.: Some Applications of Statistics to Meteorology, The Pennsylvania State University, University Park, PA, USA, 224 pp., 1968.

Piani, C., G. P. Weedon, M. Best, S. M. Gomes, P. Viterbo, S. Hagemann, and J. O. Haerter, 2010: Statistical bias correction of global simulated daily precipitation and temperature for the application of hydrological models. J. Hydrology, 395, p. 199-215.

Pierce, D.W., T. Das, D.R. Cayan, E.P. Maurer, N.L. Miller, Y. Bao, M. Kanamitsu, K. Yoshimura, M.A. Snyder, L.C. Sloan, G. Franco and M. Tyree, 2013: Probabilistic estimates of future changes in California temperature and precipitation using statistical and dynamical downscaling. Clim. Dyn., 40, 839-856. doi:10.1007/s00382-012-1337-9.

Pierce, D. W., and Cayan, D. R., 2013: The uneven response of different snow measures to human-induced climate warming. J. Clim., 26, 4148-67.

Polade, S. D., D. W. Pierce, D. R. Cayan, A. Gershunov, and M. D. Dettinger, 2014: The key role of dry days in regional climate change. *In preparation.*

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 U.S. J. Geophys. Res. Atmos, *in press.*

Sheffield, J., Langenbrunner, B., Meyerson, J. E., Neelin, J. D., Camargo, S. J., Fu, R., Hu, Q., Jiang, X., Karnauskas, K. B., Kim, S. T., Kumar, S., Kinter, J., Maloney, E. D., Mariotti, A., Pan, Z., Ruiz-Barradas, A., Nigam, S., Seager, R., Serra, Y. L., Sun, D.-Z., Wang, C., Yu, J.-Y., Johnson, N., Xie, S.-P., Zhang, T., and Zhao, M., 2013: North American Climate in CMIP5 Experiments. Part II: Evaluation of Historical Simulations of Intra-Seasonal to Decadal Variability, J. Climate, 10.1175/jcli-d-12-00593.1.

Sillmann, J., Kharin, V. V., Zhang, X., Zwiers, F. W., and Bronaugh, D., 2013: Climate extremes indices in the CMIP5 multimodel ensemble: Part 1. Model evaluation in the present climate, Journal of Geophysical Research: Atmospheres, 118, 1716-1733, 10.1002/jgrd.50203.

Sun, Y., S. Solomon, A. Dai, and R. W. Portmann, 2006: How often does it rain? J. Climate, 19, 916-934.

Taylor, K. E., R. J. Stouffer, and G. A. Meehl, 2012: An Overview of CMIP5 and the experiment design. Bull. Am. Met. Soc., 93, 485-498, doi:410.1175/BAMS-D-1111-00094.00091.

Themeßl, M., Gobiet, A., and Leuprecht, A., 2011: Empirical-statistical downscaling and error correction of daily precipitation from regional climate models, Int. J. Climatol., 31, 1530-1544, 10.1002/joc.2168.

Thrasher, B., E. P. Maurer, C. McKellar, and P. B. Duffy, 2012: Technical Note: Bias correcting climate model simulated daily temperature extremes with quantile mapping. Hydrol. Earth Syst. Sci., 16, 3309-3314. Doi:10.5194/hess-16-3309-2012.

Trewin, B. C, 2007: The role of climatological normals in a changing climate. World Meteorological Organization Technical Directive WCDMP-No. 61 and WMO-TD No. 1377. Geneva, Switzerland. Available at http://www.wmo.int/datastat/documents/WCDMPNo61_1.pdf.

van Vuuren, D. P., J. Edmonds, M. Kainuma, K. Riahi, A. Thomson, et al., 2011: The representative concentration pathways: an overview. Climatic Change, 109, 5-31. Doi: 10.1007/s10584-011-0148-z.

von Storch, H., and F. W. Zwiers, 2001: Statistical analysis in climate research. Cambridge University Press, 484 pp.

Wood, A. W., E. P. Maurer, A. Kumar, and D. P. Lettenamier, 2002: Long-range experimental hydrologic forecasting for the eastern United States. J. Geophys. Res. D., 107, 4429, doi:10.1029/2001JD000659, 2002.

Wood, A. W., Leung, L. R., Sridhar, V., and Lettenmaier, D. P., 2004: Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs, Climatic Change, 62, 189-216.

Wigley, T. M. L., and Jones, P. D.: Influences of precipitation changes and direct CO2 effects on streamflow, Nature, 314, 149-152, 1985.

Abbreviation	Model source/institution
Access1-0	Commonwealth Scientific and Industrial Research Organization (CSIRO) and Bureau of Meteorology (BOM), Australia
Bcc-csm1-1	Beijing Climate Center, China
Bnu-esm	Beijing Normal University, China
CanESM2	Canadian Centre for Climate Modelling and Analysis, Canada
CCSM4	National Center for Atmospheric Research, USA
ECSM1-BGC	National Center for Atmospheric Research, USA
CNRM-CM5	Centre National de Recherches Meteorologiques, France
CSIRO-Mk3.6.0	QCCCE & Commonwealth Scientific and Industrial Research Organization, Australia
GFDL-CM3	Geophysical Fluid Dynamics Laboratory, Princeton, USA
GFDL-ESM2G	Geophysical Fluid Dynamics Laboratory, Princeton, USA
GFDL-ESM2M	Geophysical Fluid Dynamics Laboratory, Princeton, USA
INMCM4	Institute of Numerical Mathematics Russian Academy of Sciences, Russia
IPSL-CM5a-LR	Institut Pierre-Simon Laplace, France
IPSL-CM5a-MR	Institut Pierre-Simon Laplace, France
MIROC-ESM	Japan Agency for Marine-Earth Science and Technology, and National Inst. For Environ. Studies, Japan
MIROC-ESM-CHEM	Japan Agency for Marine-Earth Science and Technology, and National Inst. For Environ. Studies, Japan
MIROC5	Atmosphere and Ocean Research Institute and Nat. Inst. For Environ. Studies, Japan
MPI-ESM-LR	Max Planck Institute for Meteorology, Germany
MPI-ESM-MR	Max Planck Institute for Meteorology, Germany
MRI-CGCM3	Meteorological Rsearch Institute, Japan
NorESM1-m	Norwegian Climate Centre

Table 1. The GCMs used in this work and their originating institutions.

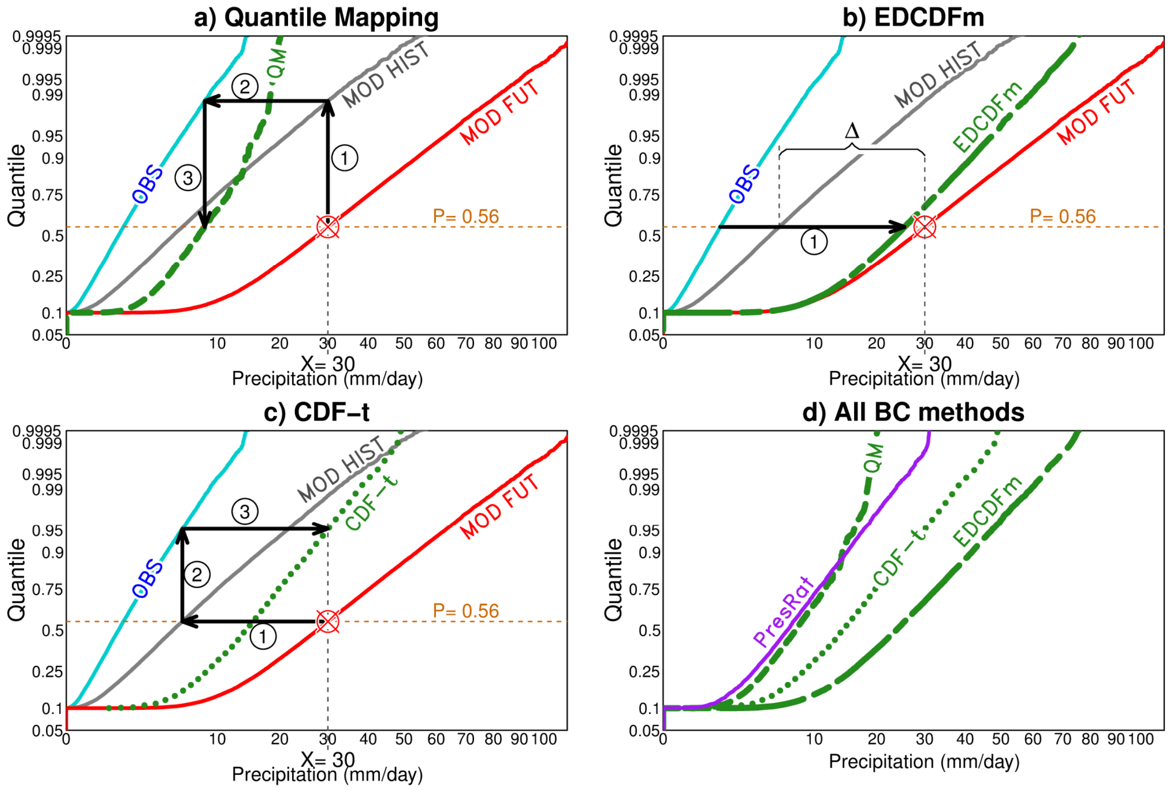

Figure 1. Cumulative distribution functions (CDFs) of synthetic daily precipitation data schematically illustrating how each bias correction method constructs the model's bias corrected future CDF (green dotted/dashed lines). The solid blue, grey, and red lines are the same in all panels and show the observed (1976-2005), model historical (1976-2005), and model future (2070-2099) CDFs, respectively. The example point being corrected is X=30 mm/day, which falls at the 0.56 quantile in the model future distribution (dotted orange line). a) Quantile mapping (QM): starting at the point to be corrected, go vertically to the grey line (1), horizontally to the blue line (2), and vertically to the original percentile (3). b) Equidistant CDF-matching (EDCDFm): at the quantile of the point being corrected, compute the offset from the model historical value to the model future value (Δ), then add Δ to the observed value at the percentile being corrected (1). c) The CDF-transform (CDF-t) method; starting at the point to be corrected, go horizontally to the grey line (1), vertically to the blue line (2), and horizontally to the original value (3). d) Final results from all 3 bias correction methods (dotted/dashed green lines), along with the PresRat method (solid purple line) for comparison. Note that the X axis uses a square root transformation and the Y axis uses an inverse error function ("probability plot") transformation.

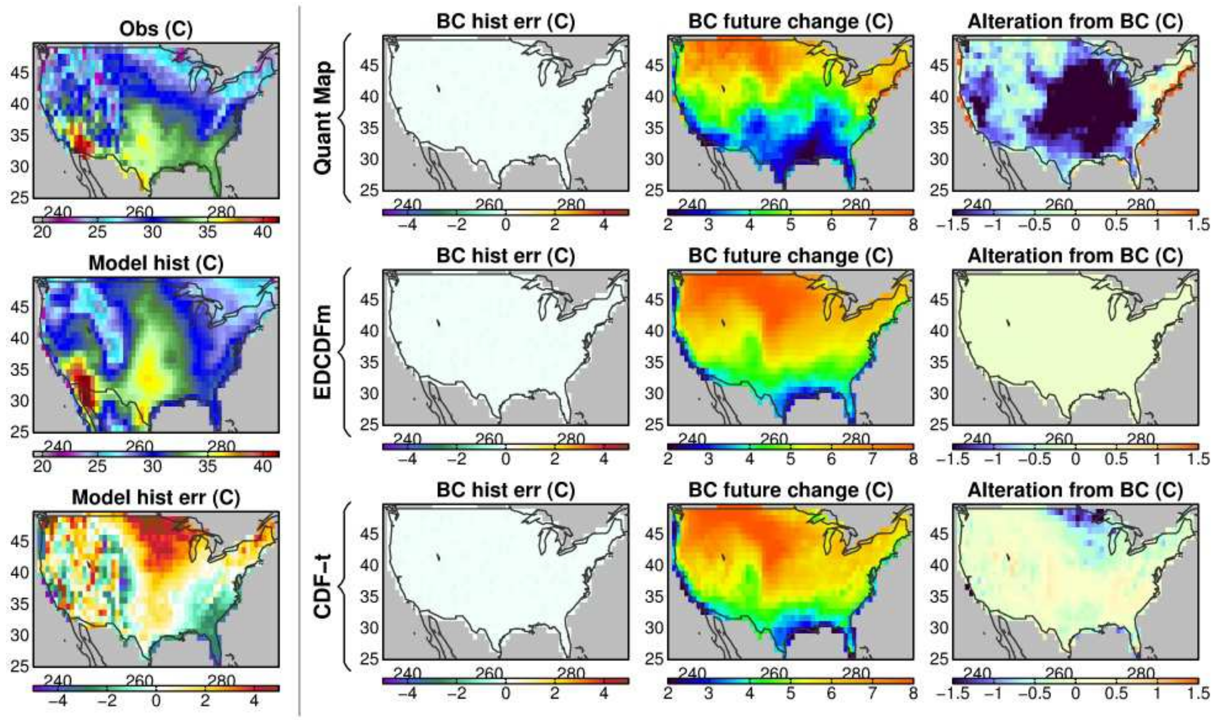

Figure 2. Illustration of how bias correction can alter the model-predicted future change in monthly-averaged maximum daily temperature, shown for July using the CCSM4 GCM. Left column: the observations, model simulation over the historical period (1976-2005; °C), and model error with respect to observations without any bias correction (°C). Right part of figure: For each of the bias correction methods indicated (quantile mapping (upper row), EDCDFm (middle row), and CDF-t (lower row)) shown are the model error with respect to observations over the historical period after bias correction has been applied (°C), the model-projected future change (2070-2099) after bias correction using the indicated method (°C), and the amount that the bias correction method alters the original model-predicted change (°C).

/cir1/compare_BC_methods_v3_tasmax_ccsm4.R.ps

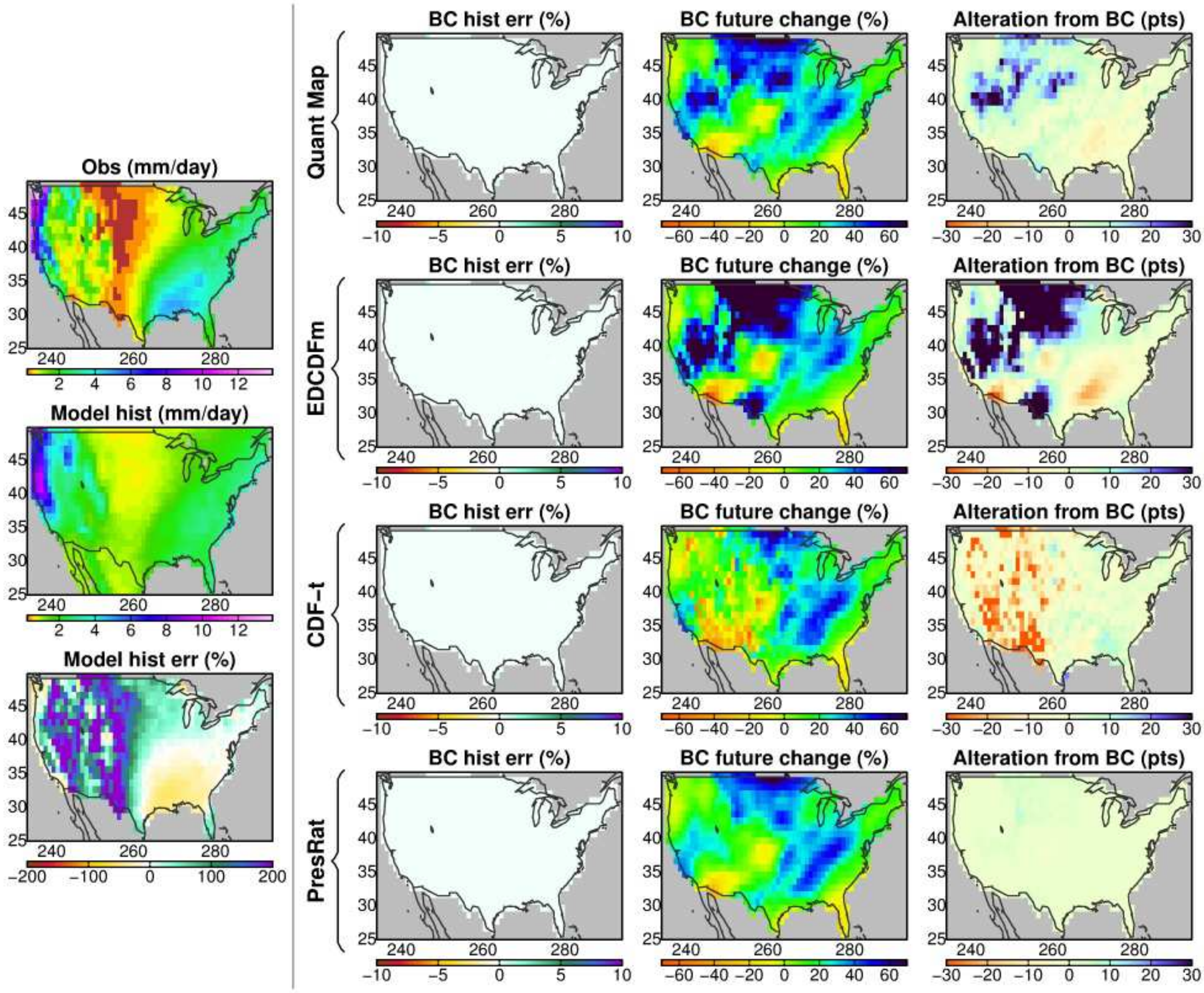

Figure 3. Illustration of how bias correction can alter the model-predicted future change in precipitation, shown for December using CCSM4. Left column: the observations, model simulation over the historical period (1976-2005; mm/day), and model error over the historical period with respect to observations without any bias correction (%). Right part of figure: For each of the four bias correction methods indicated, shown are the model error with respect to observations over the historical period after bias correction has been applied (%), the model-predicted change in future (2070-2099) precipitation field after bias correction with the indicated method (%), and the amount that the bias correction method alters the original model-predicted change in precipitation between the future and historical period (percentage points).

/cir1/cmip5_regrid/compare_BC_methods_v3_ccsm4.R.ps

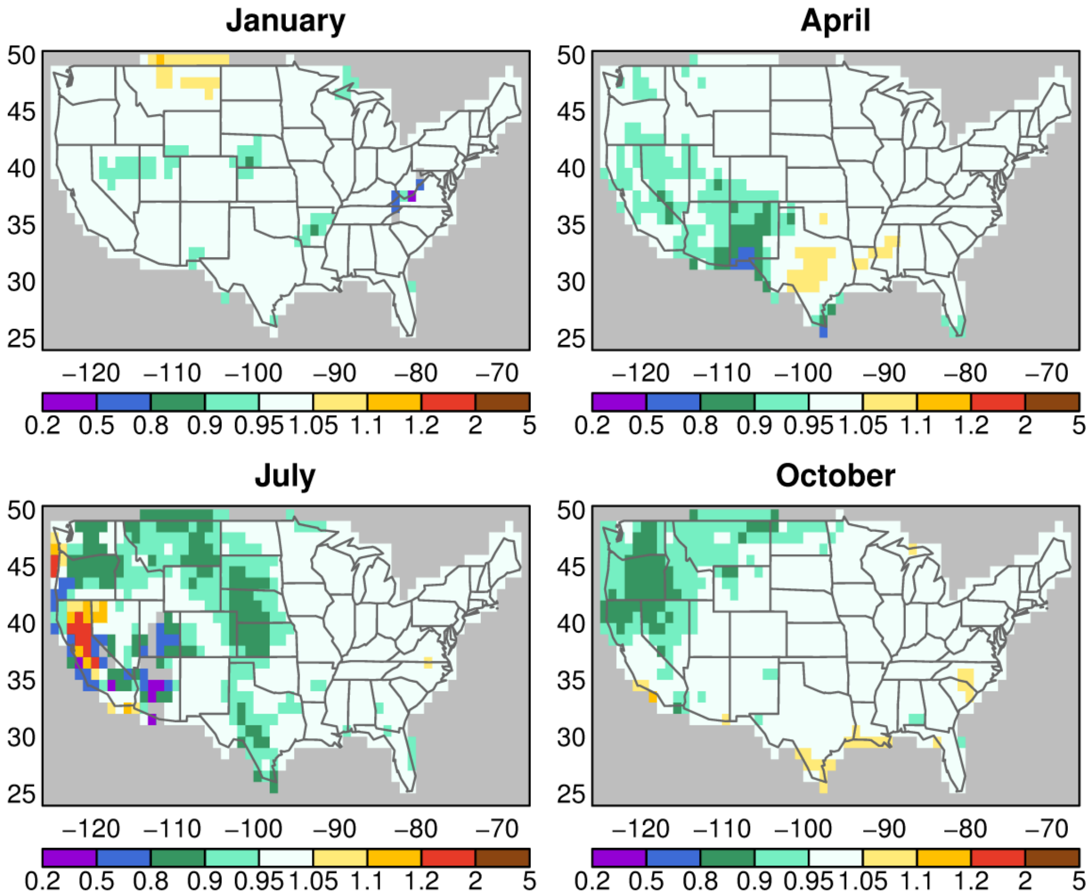

Figure 4. Correction factors, *K*, for the PresRat scheme that are necessary to preserve model-predicted changes (2070-2099 vs. 1976-2005) in mean precipitation, illustrated for four months. Values are averaged across 21 GCMs. White areas are within 5% of unity.

plot_presrat_factors_allmods.R.gif

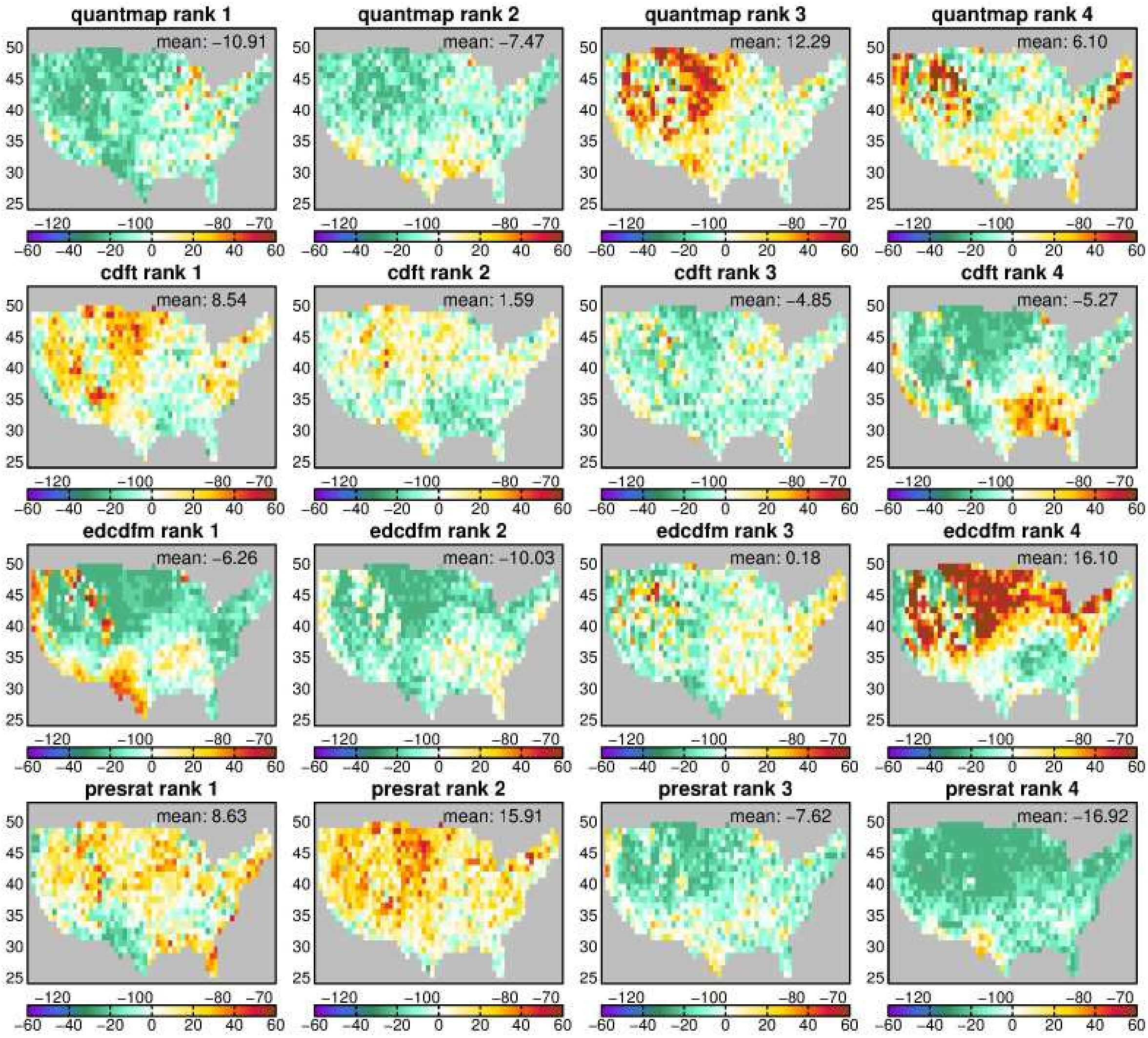

Figure 5. Percent of the 21 GCMs in which the indicated bias correction method (rows) produces a winter (DJF) 95th percentile daily precipitation value of the indicated rank (columns; 1=smallest value across the bias correction methods; 4=largest). Plotted values are relative to 25%, which is the expected value assuming all 4 bias correction methods produce extrema of equal magnitude. Yellows and reds show where a particular bias correction method produces more values of the indicated rank than expected; greens and blues show where it produces less values of the indicated rank than expected.

compare_BC_method_ptiles_pctextr_v3.R

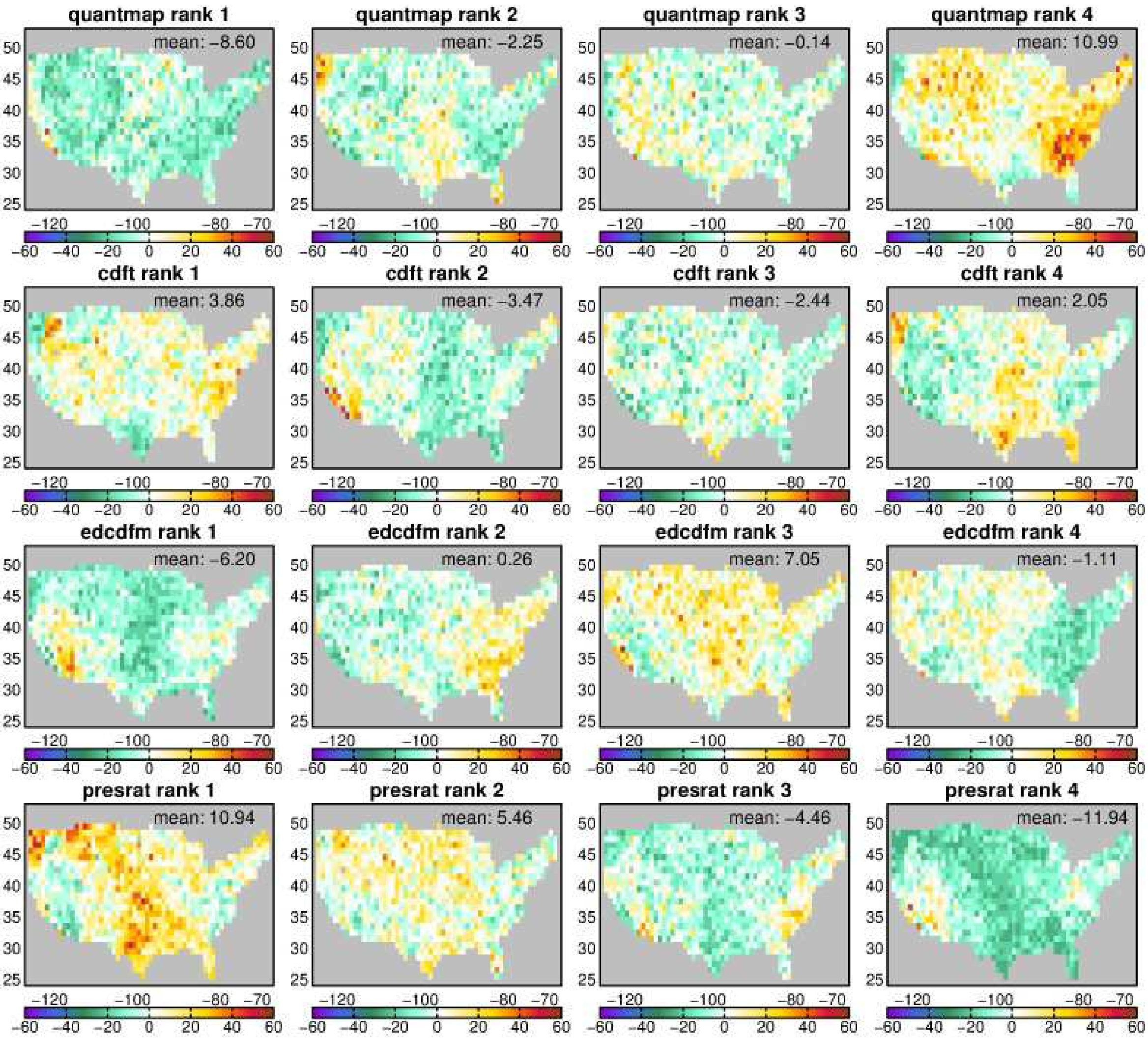

Figure 6. Same as Figure 5, but for summer (JJA).

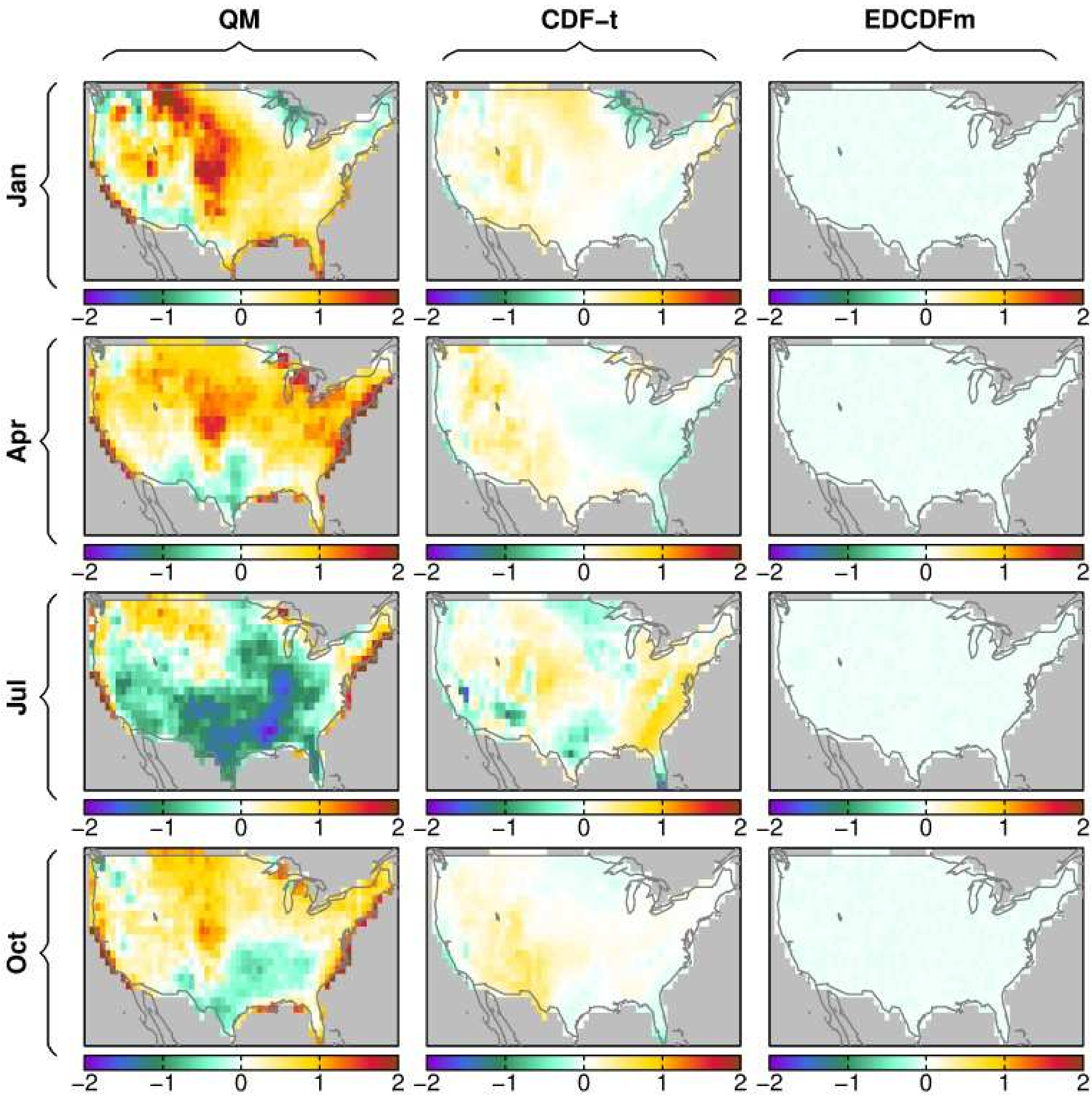

Figure 7. Bias corrected minus original GCM change in daily maximum temperature (°C) over the period 2070-2099 relative to 1976-2005, shown for 4 months (rows) and 3 bias correction methods (columns). Values are ensemble averaged across all 21 GCMs.

compare_BC_methods_allmods_tasmax_summary_v2.R.ps

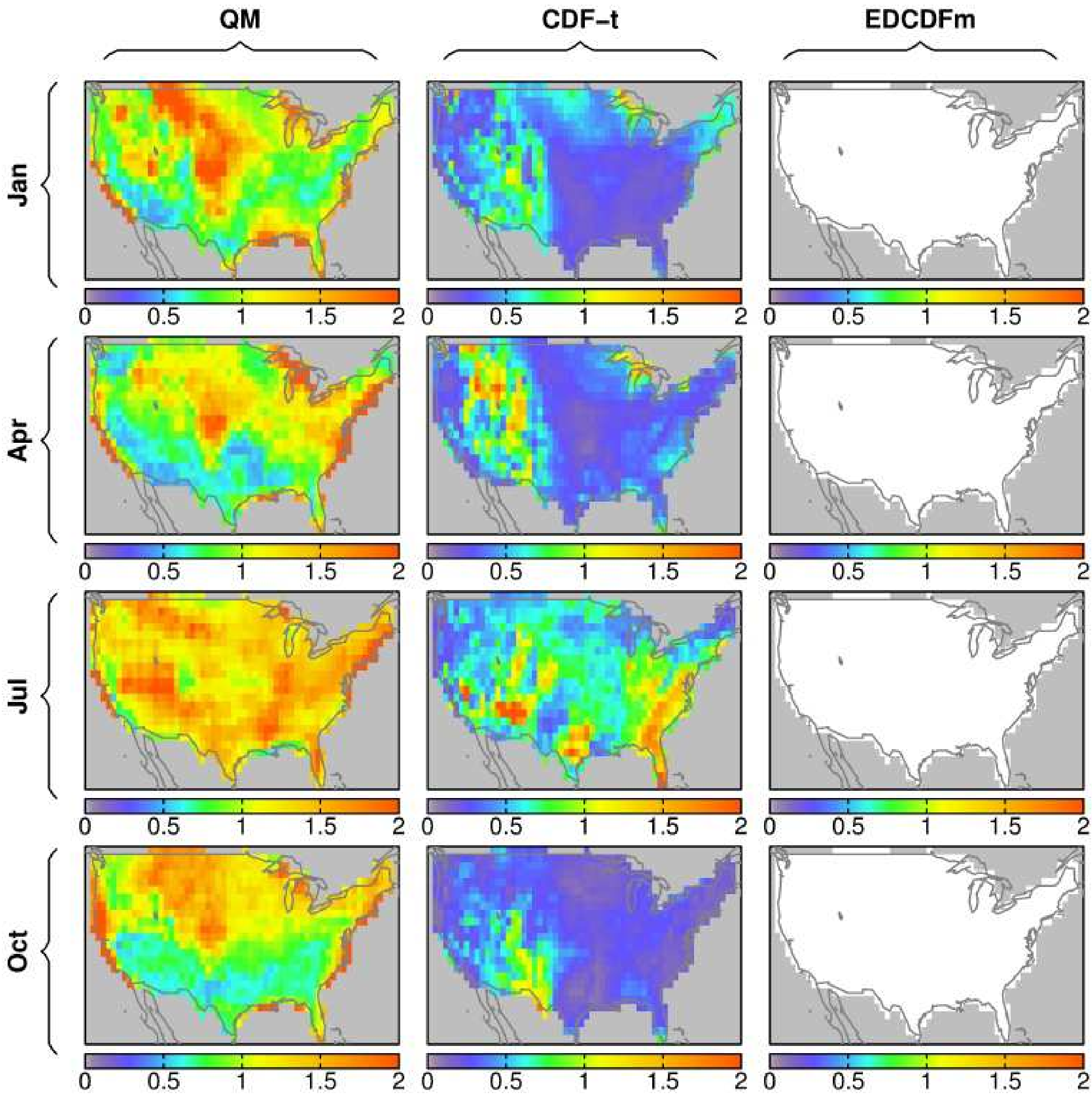

Figure 8. As in Figure 7, but for the RMS difference (°C), obtained from 21 models, between the original model-predicted future change in daily maximum temperature and the model-predicted change after bias correction has been applied.

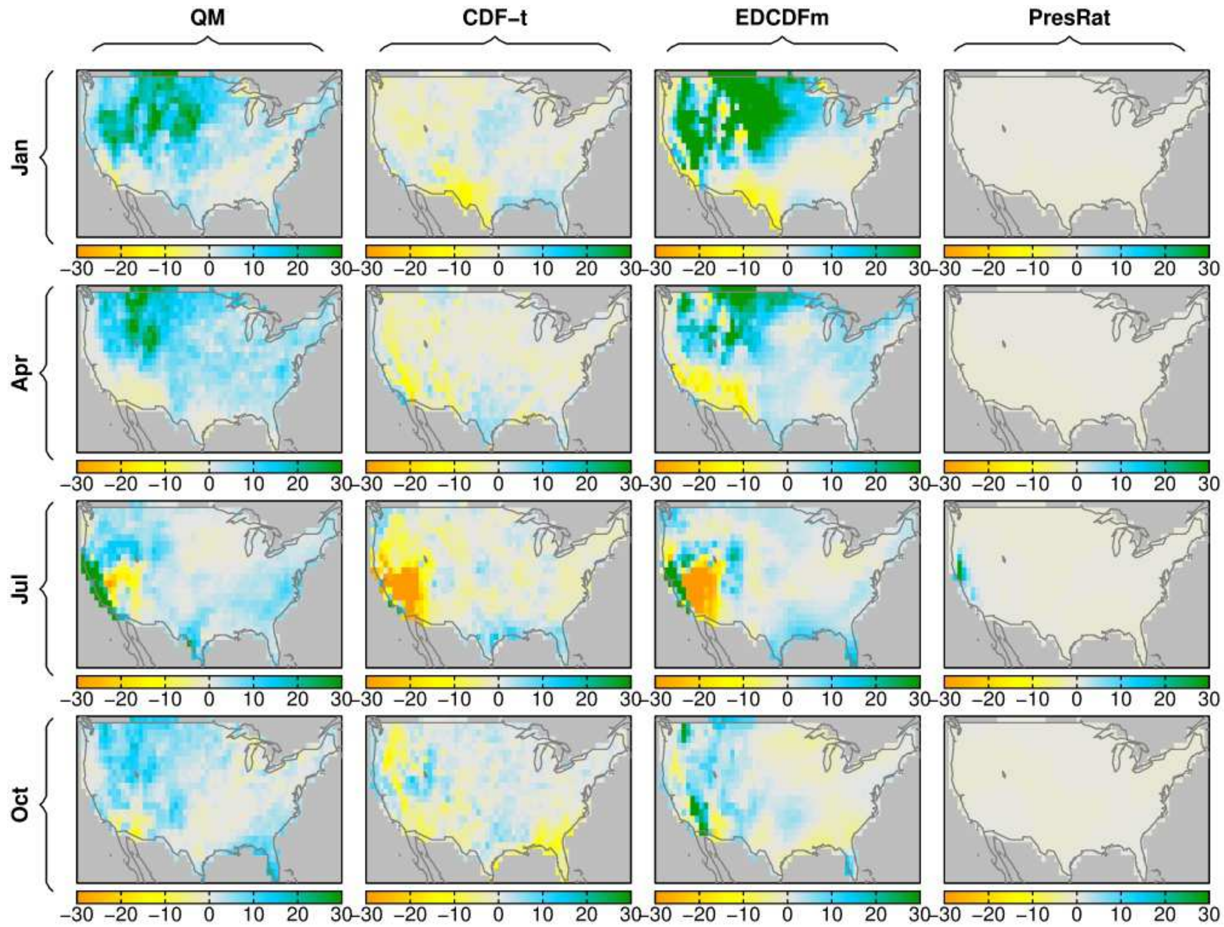

Figure 9. How bias correction alters model-predicted change in future daily precipitation (percentage points), shown for 4 months (rows) and 3 bias correction methods (columns). Values are ensemble averaged across all 21 GCMs.

compare_BC_methods_allmods_pr_summary_v2.R

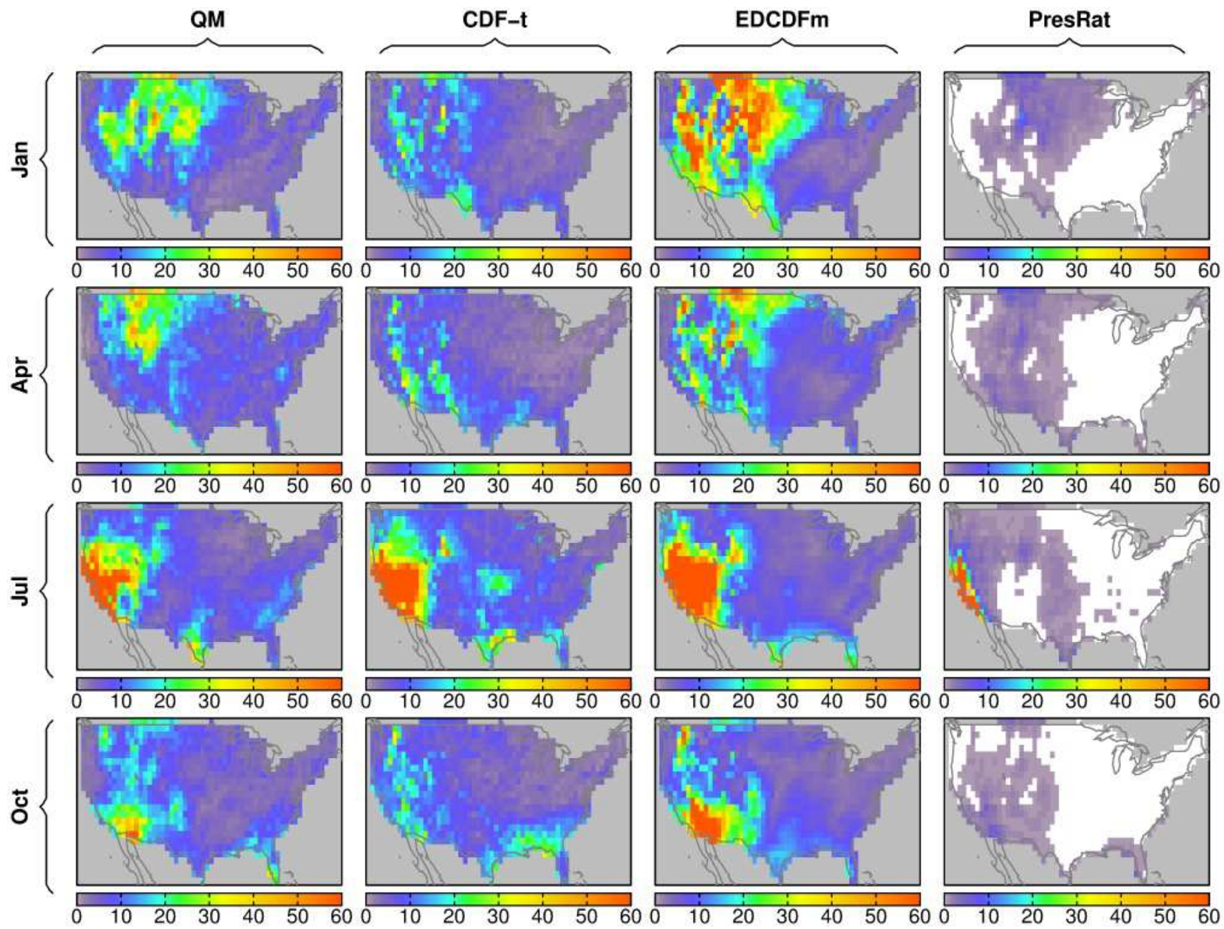

Figure 10. As in Figure 9, but for the RMS difference (percentage points).between the original model-predicted future change in daily precipitation and the model-predicted change after bias correction has been applied.

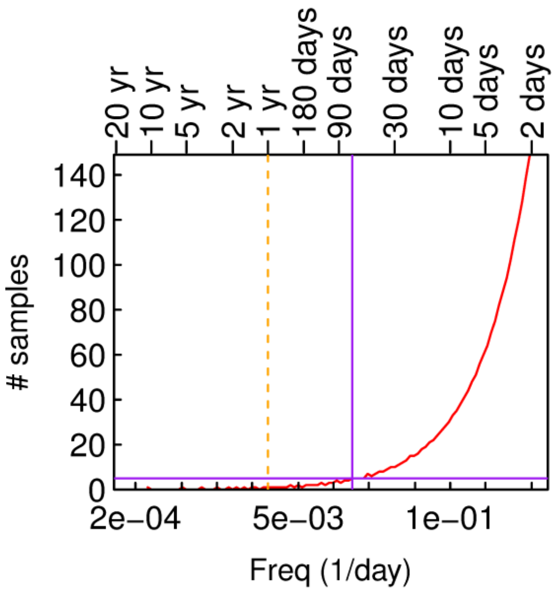

Figure 11. Number of samples per constant-width bin in the logarithm of frequency. The period corresponding to the frequency is shown along the top axis. The vertical dashed orange line shows the annual cycle. The purple lines show that the number of samples per logarithmic frequency bin exceeds 5 at periods less than about 80 days.

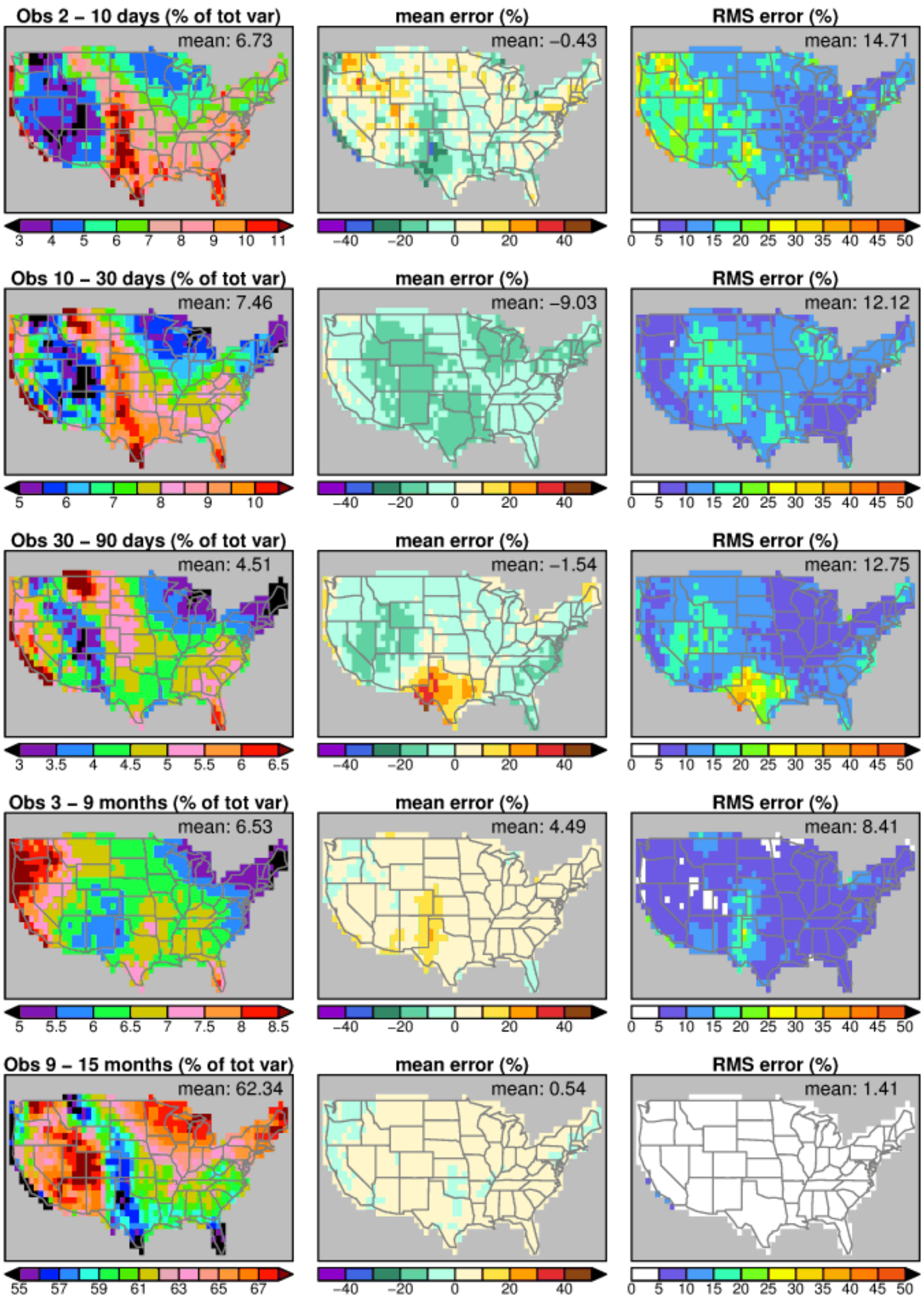

Figure 12a. Left column: proportion (%) of total variance of daily maximum temperature that falls in the frequency band whose period is indicated in left hand column, from observations over the period 1976-2005. Note that the color range varies substantially by frequency band. Middle column: the multi-model mean error (%) for the same quantity in the GCMs, relative to the observations. Right column: the multi-model RMSE (%). Figure continues.

compare_spec_mod_vs_obs_norm_tasmax_summary_v3.R

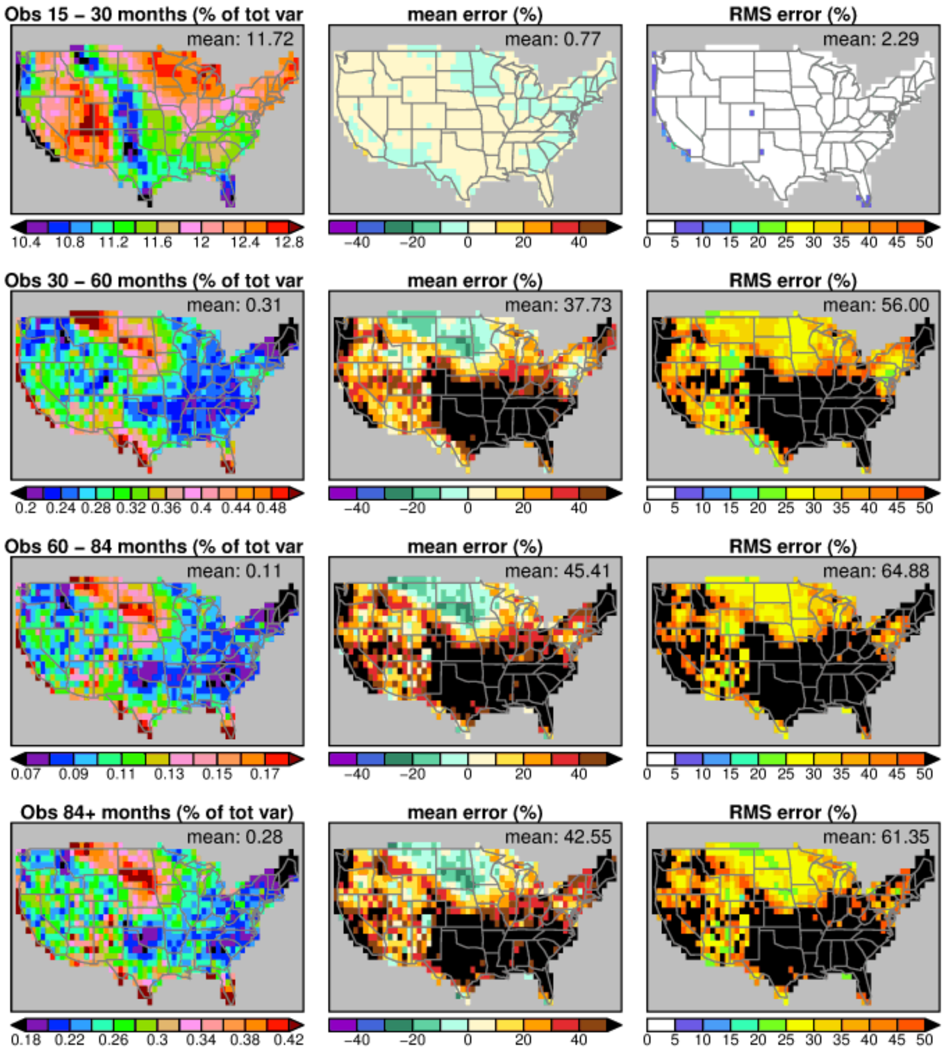

Figure 12b. As in Figure 12, but for frequency bands whose periods are indicated in left hand column.

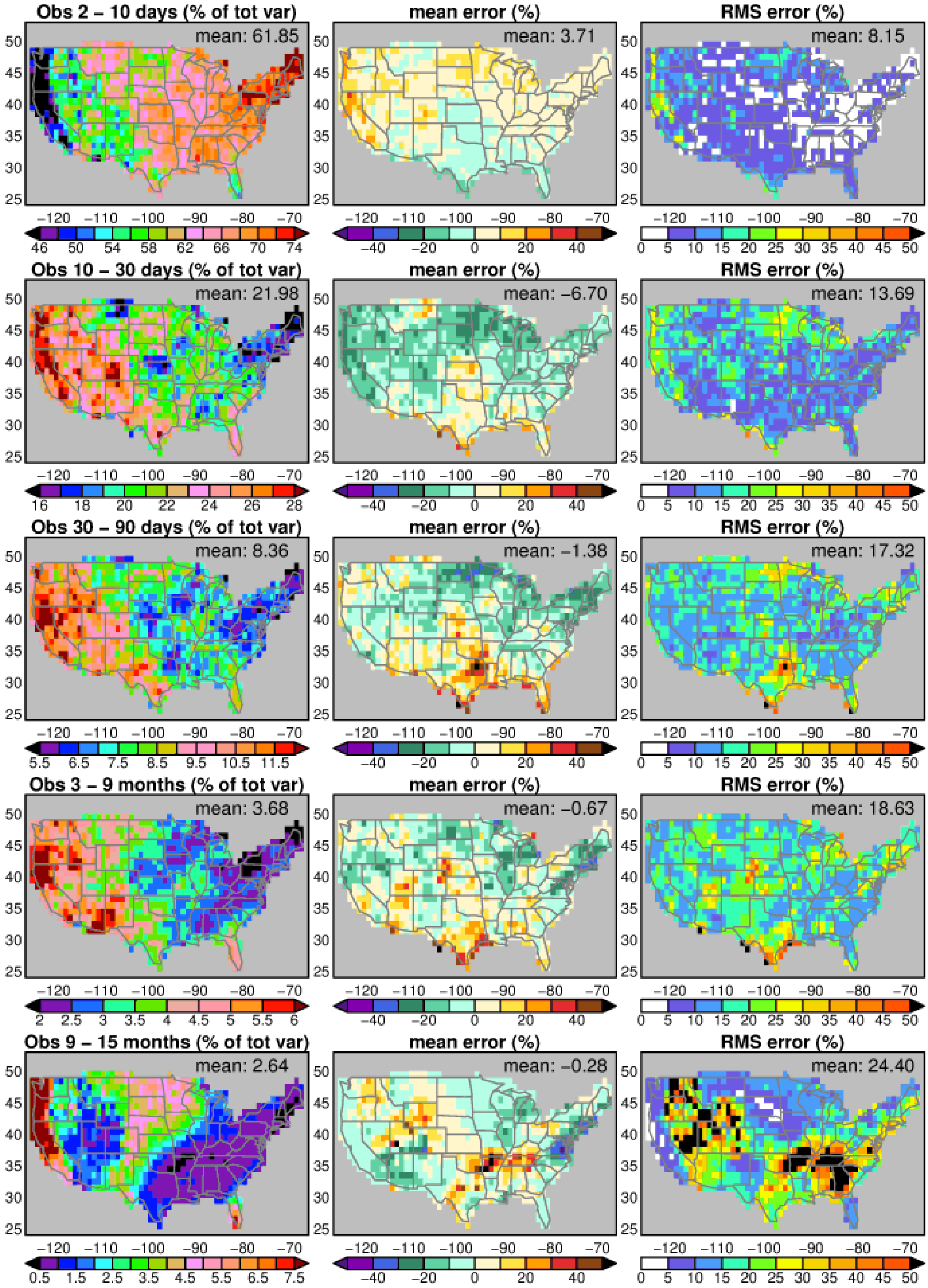

Figure 13a. As in Figure 12a, but for daily precipitation. Figure continues.

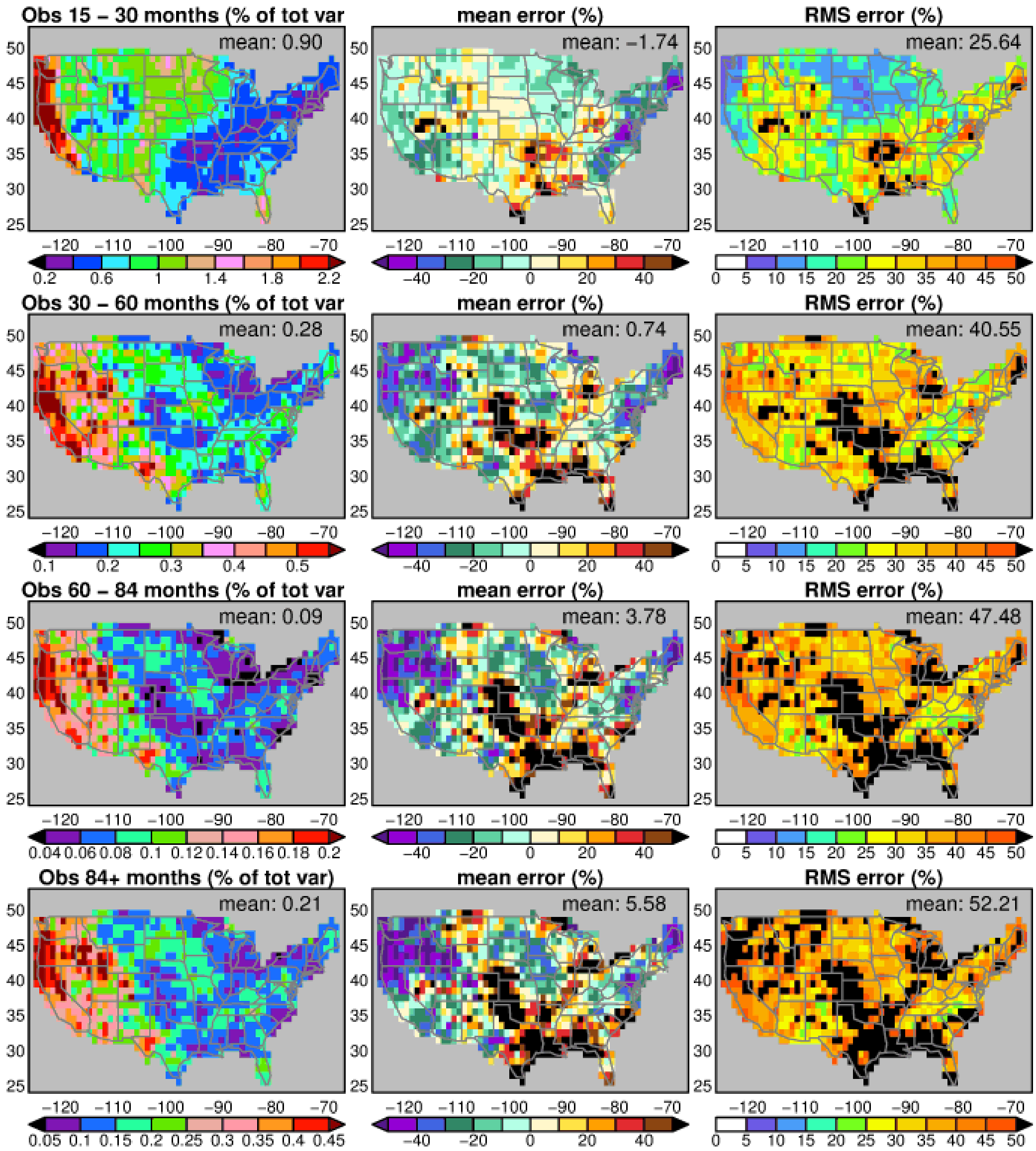

Figure 13b, continued.

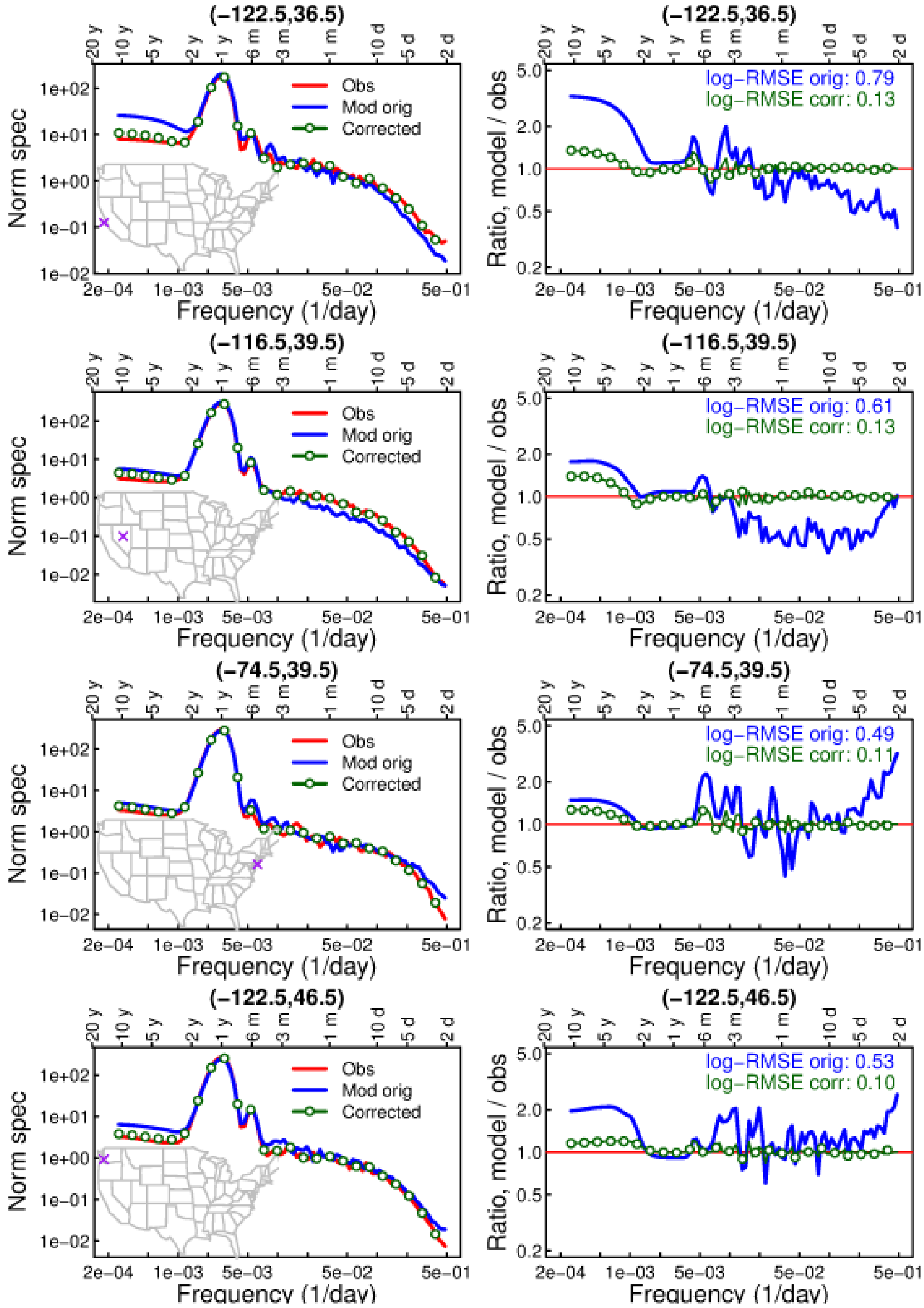

Figure 14. Left column: Normalized spectra of daily maximum temperature from observations (red line), CCSM4 (blue line), and CCSM4 after frequency-dependent bias correction (green dots and line), taken at the location indicated by the purple cross on inset map, coordinates given in the panel title. Right column: Ratio of CCSM4 spectral power to observations (blue line) and ratio of CCSM4 to observations after frequency-dependent bias correction to observations (green dots and line).

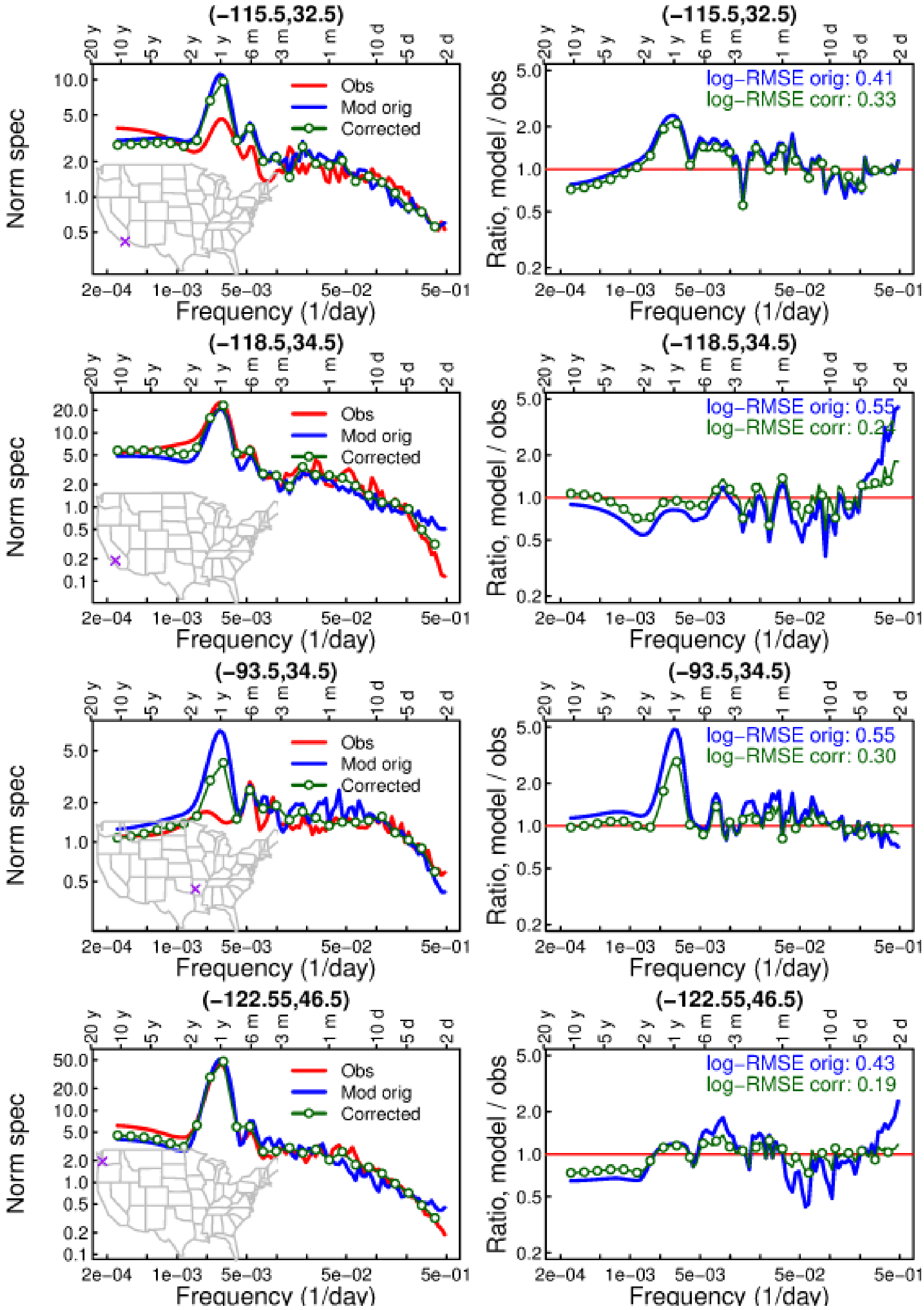

Figure 15. Same as Figure 14, but for daily precipitation.

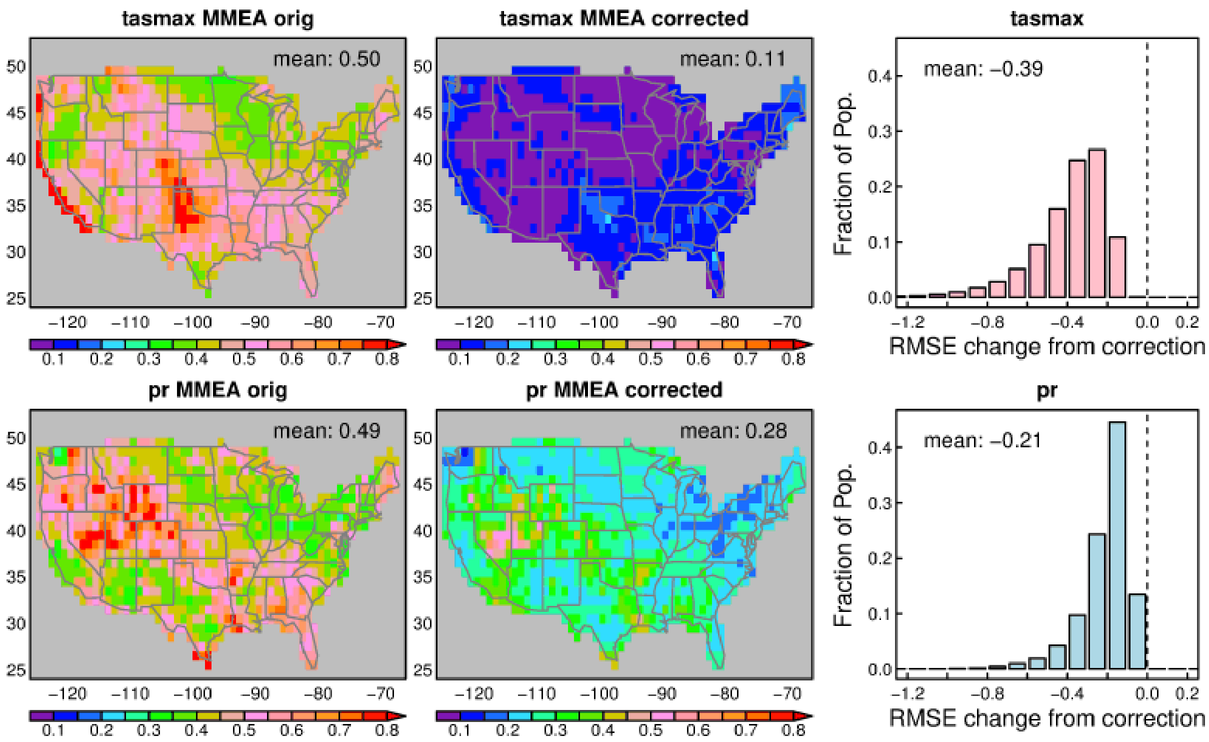

Figure 16. For daily maximum temperature (top row) and precipitation (bottom row), the multi-model ensemble average log-RMSE in simulating the observed distribution of variance across frequency, both before the frequency-dependent bias correction (left column) and after (middle column). Right: histograms of how the frequency-dependent bias correction changes the log-RMSE, taken over all models and all locations.

/cir1/cmip5_regrid/get_rmse_all_models.R

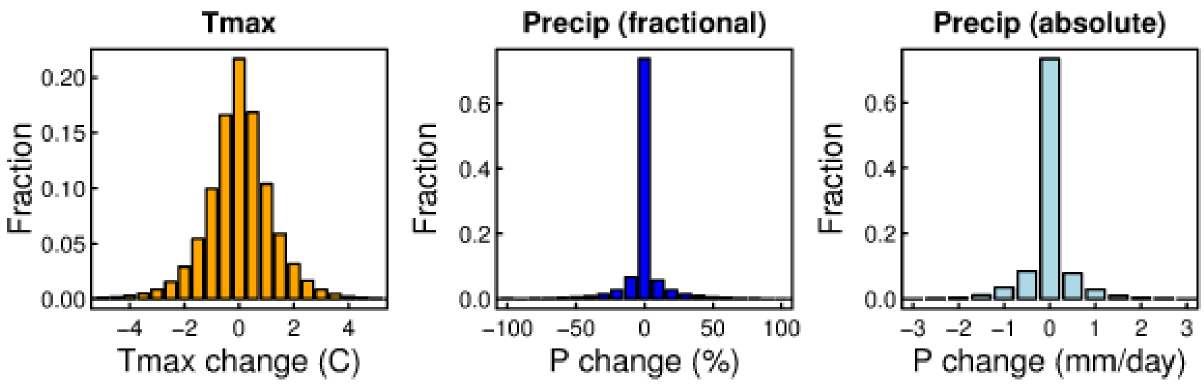

Figure 17. Histograms of how much the frequency-dependent bias correction alters the daily temperature (left, °C) and precipitation (right two panels). The precipitation results are given both as the fraction change (%) and absolute change (mm/day). Results are shown for all the models across all points in the conterminous U.S.

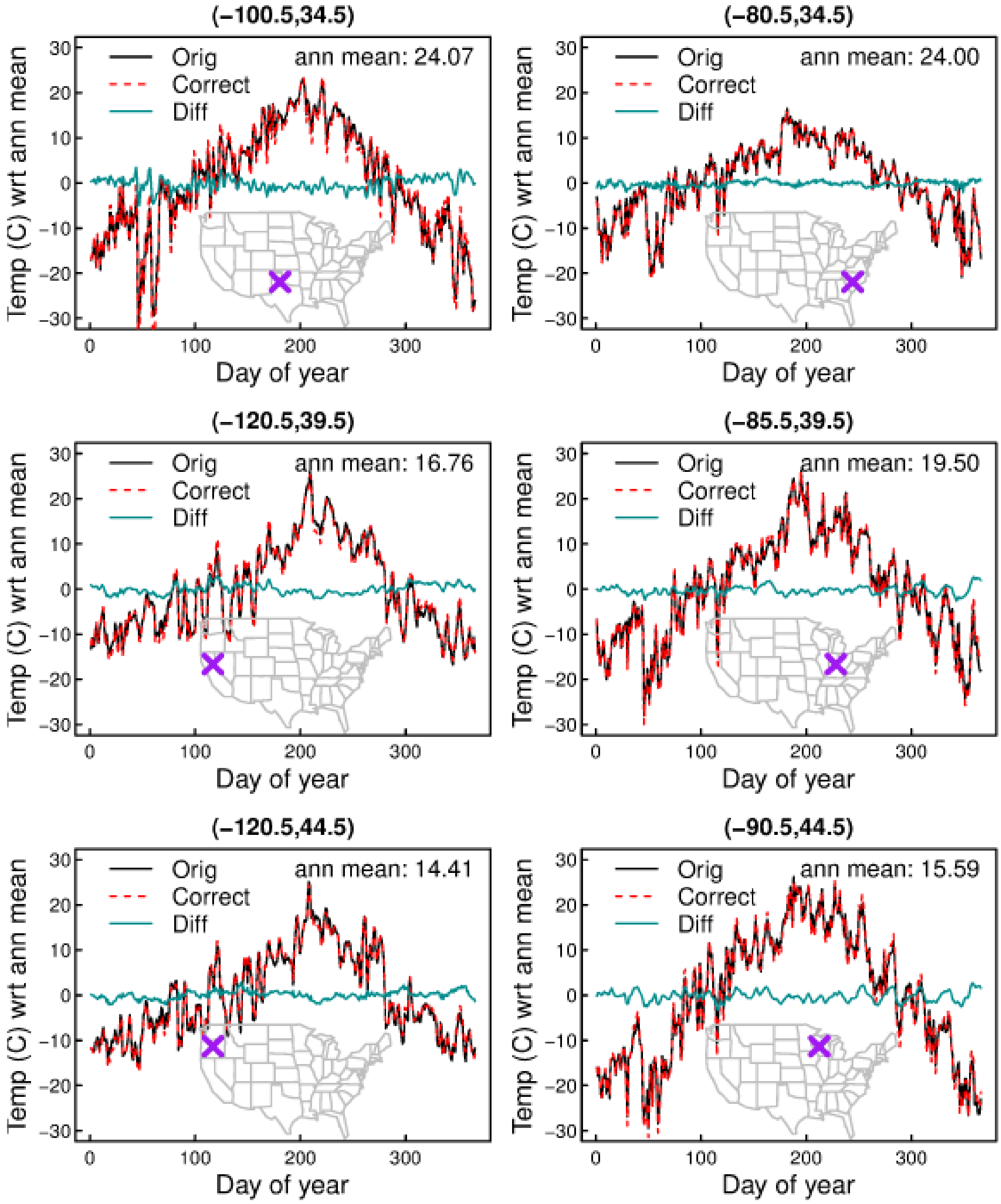

Figure 18. Example one-year time series of daily maximum temperature at the location marked by the purple 'X' both before (black line) and after (dotted red line) the frequency-dependent bias correction. Values have had the annual mean removed; the value of the annual mean is shown in the upper right part of the panel. The blue line is the time series of the correction, i.e., the corrected time series minus original. Values are from CCSM4 in year 2000.

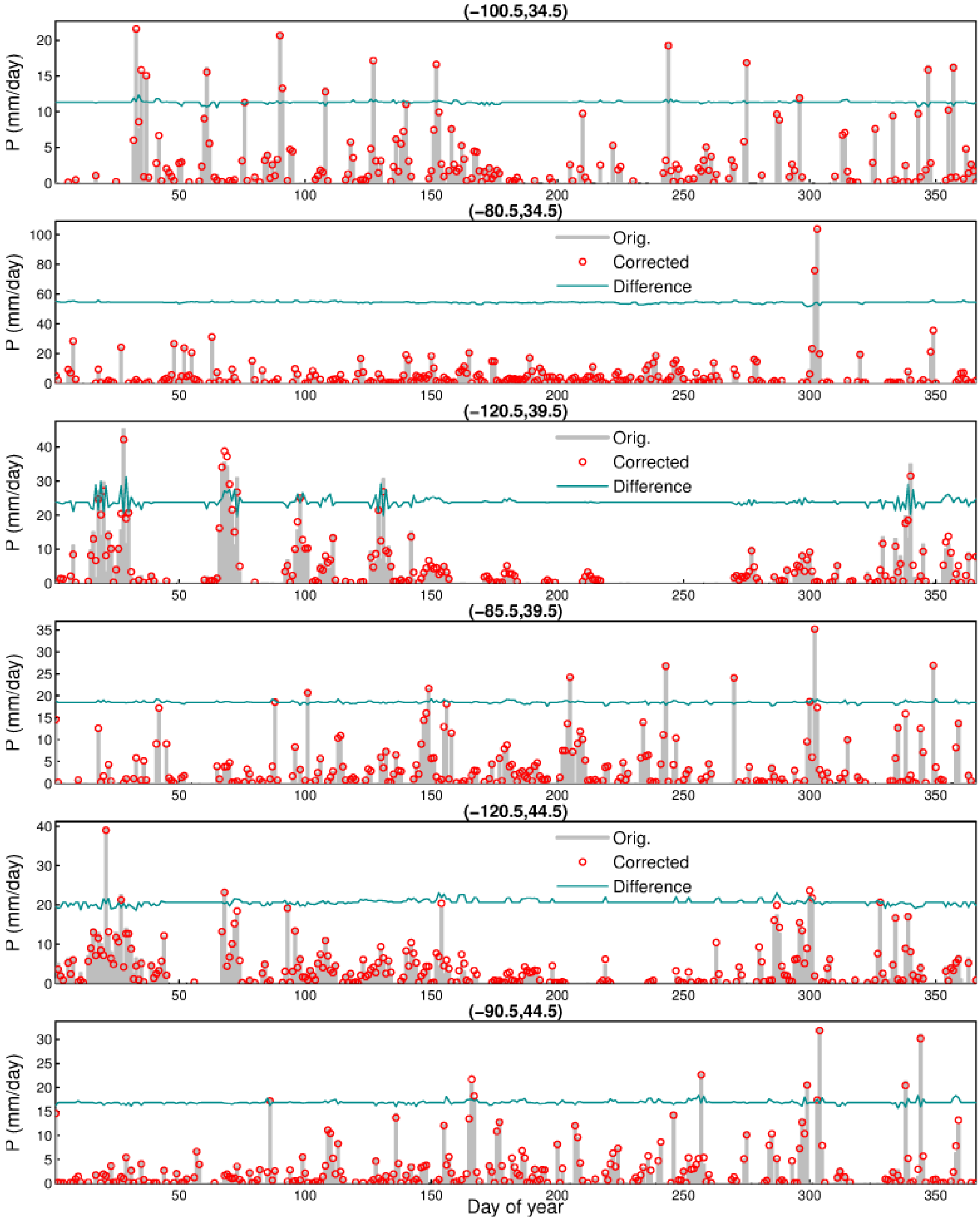

Figure 19. Example one-year time series of daily precipitation at the location indicated in the panel title (longitude, latitude), both before (grey lines) and after (red circle) the frequency-dependent bias correction. The blue line is the time series of the correction, i.e., the corrected time series minus original, with no scaling but offset to be vertically centered in the middle of the panel. Values are from CCSM4 in year 2000.

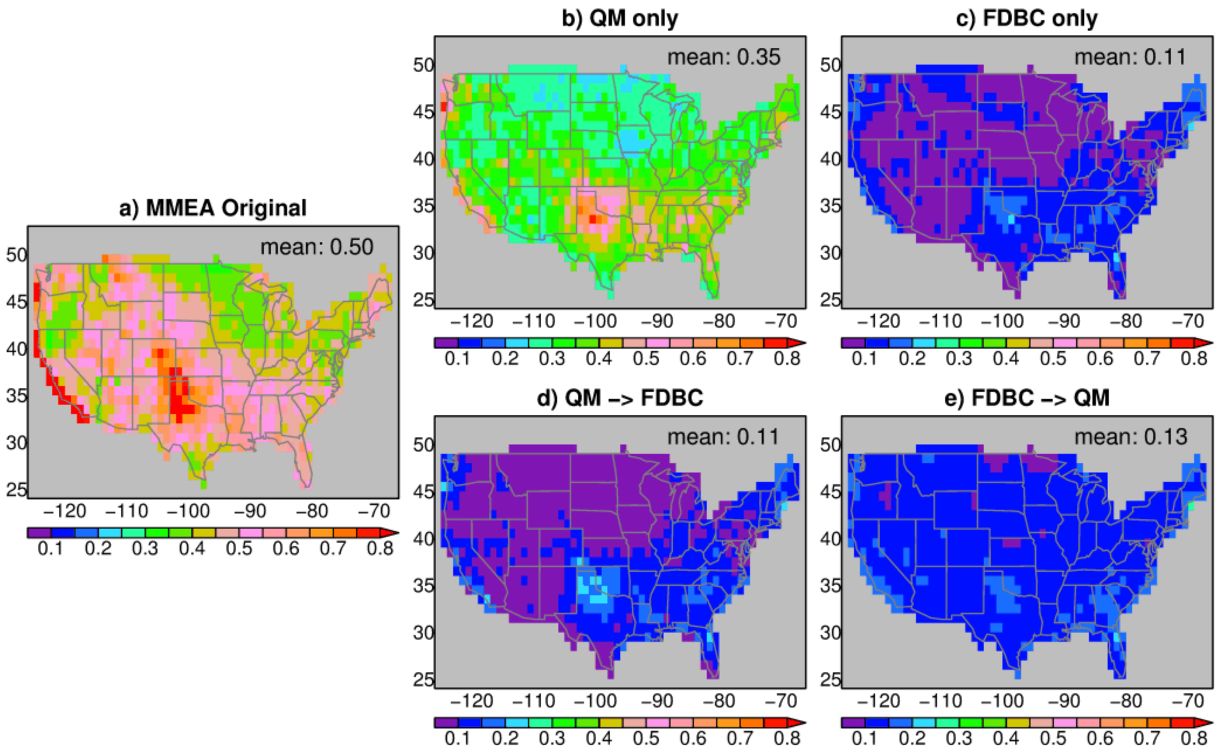

Figure 20. Multi-model ensemble mean log-RMSE for daily maximum temperature. a) the multi-model ensemble average value before any bias correction has been applied. b) With only quantile mapping (QM) applied. c) With only frequency-dependent bias correction (FDBC) applied. d) With QM applied first, then FDBC. e) With FDBC applied first, then QM.

get_rmse_all_models_bc_srs_both_v2_presentation.R

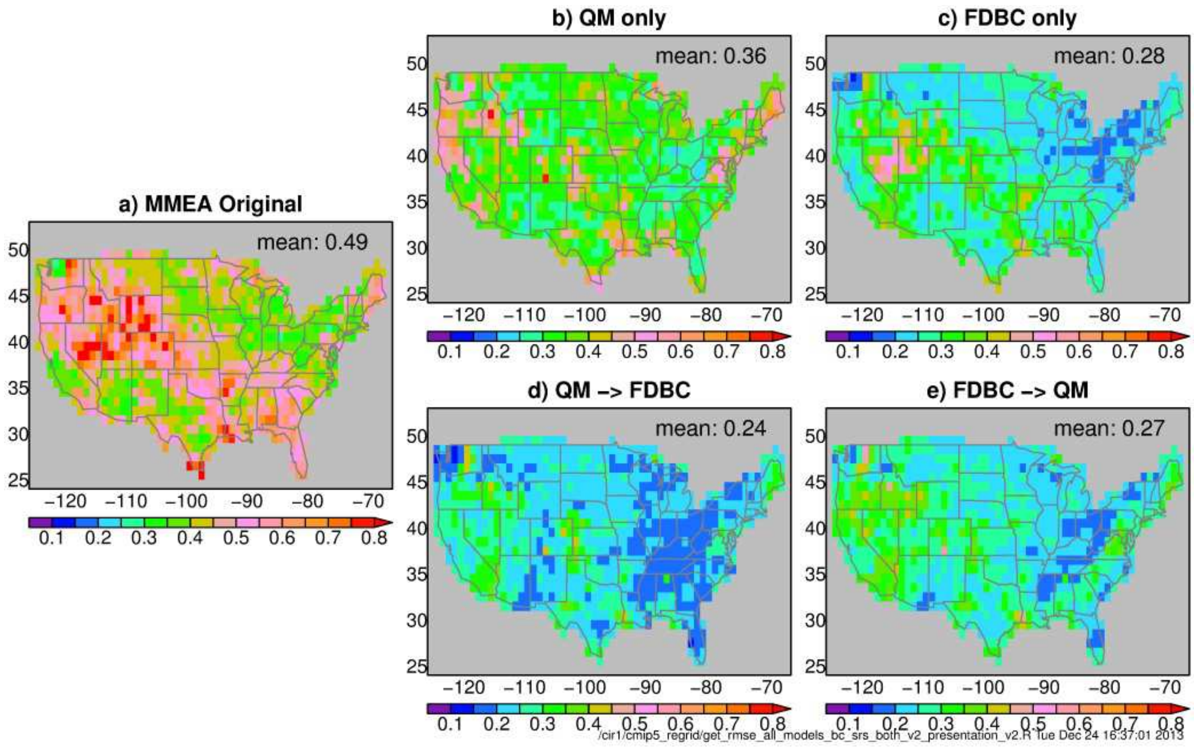

Figure 21. As in Figure 20, but for daily precipitation.

get_rmse_all_models_bc_srs_both_v2_presentation.R.ps

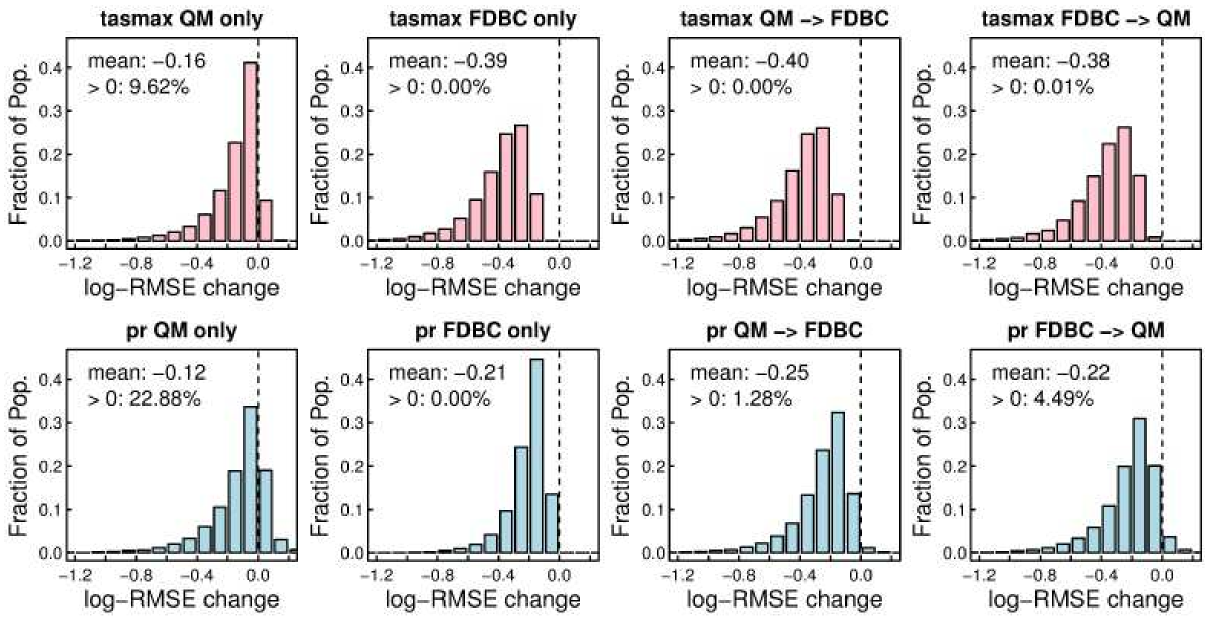

Figure 22. Histograms of the change in log-RMSE for the models' simulation of the variance spectra of daily maximum temperature (top row) and precipitation (bottom row) when the indicated bias correction method is applied. QM: quantile mapping. FDBC: frequency-dependent bias correction. Also indicated in each panel are the mean value and percent of values greater than zero. Values are pooled over all models and locations.

get_rmse_all_models_bc_srs_both_v2_presentation_hists.R.ps

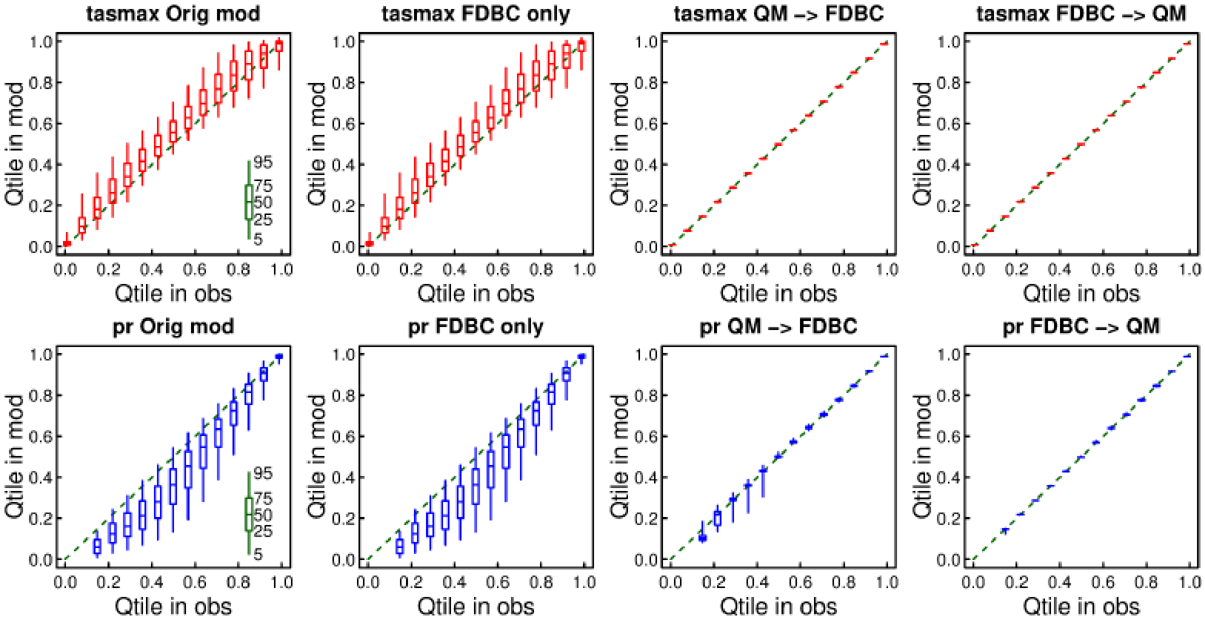

Figure 23. Quantile-quantile plots showing how well the GCMs simulate the distribution of daily maximum temperature (top) and daily precipitation (bottom), both before (left column) and after various combinations (described in Figure 19) of quantile mapping (QM) and frequency-dependent bias correction (FDBC). In each panel the dotted green line shows a 1-to-1 relationship, which would be perfect agreement between the model and observations. The box and whiskers show the distribution of model quantile values as indicated in the legend, pooled across all models and all locations. Values are obtained from the control period, 1976-2005.

get_rmse_all_models_bc_srs_both_v2_presentation_QQ.R.gif