

Bias Correction of GCM Precipitation by Quantile Mapping: How Well Do Methods Preserve Changes in Quantiles and Extremes?

ALEX J. CANNON,* STEPHEN R. SOBIE, AND TREVOR Q. MURDOCK

Pacific Climate Impacts Consortium, University of Victoria, Victoria, British Columbia, Canada

(Manuscript received 30 October 2014, in final form 18 June 2015)

ABSTRACT

Quantile mapping bias correction algorithms are commonly used to correct systematic distributional biases in precipitation outputs from climate models. Although they are effective at removing historical biases relative to observations, it has been found that quantile mapping can artificially corrupt future model-projected trends. Previous studies on the modification of precipitation trends by quantile mapping have focused on mean quantities, with less attention paid to extremes. This article investigates the extent to which quantile mapping algorithms modify global climate model (GCM) trends in mean precipitation and precipitation extremes indices. First, a bias correction algorithm, quantile delta mapping (QDM), that explicitly preserves relative changes in precipitation quantiles is presented. QDM is compared on synthetic data with detrended quantile mapping (DQM), which is designed to preserve trends in the mean, and with standard quantile mapping (QM). Next, methods are applied to phase 5 of the Coupled Model Intercomparison Project (CMIP5) daily precipitation projections over Canada. Performance is assessed based on precipitation extremes indices and results from a generalized extreme value analysis applied to annual precipitation maxima. QM can inflate the magnitude of relative trends in precipitation extremes with respect to the raw GCM, often substantially, as compared to DQM and especially QDM. The degree of corruption in the GCM trends by QM is particularly large for changes in long period return values. By the 2080s, relative changes in excess of +500% with respect to historical conditions are noted at some locations for 20-yr return values, with maximum changes by DQM and QDM nearing +240% and +140%, respectively, whereas raw GCM changes are never projected to exceed +120%.

1. Introduction

Simulated precipitation outputs from global climate models (GCMs) and regional climate models (RCMs) can exhibit large systematic biases relative to observational datasets (Mearns et al. 2012; Sillmann et al. 2013). As GCM and RCM precipitation series are used as inputs to process models (e.g., Hagemann et al. 2011; Muerth et al. 2013) and gridded statistical downscaling models (e.g., Wood et al. 2004; Maurer and Hidalgo 2008; Maurer et al. 2010), algorithms have been developed to correct and minimize these biases as sources of error in subsequent modeling chains.

Systematic errors in climate model outputs can be ascribed to different sources. For example, Eden et al. (2012) classify errors in GCM precipitation fields as being due to 1) unrealistic large-scale variability or response to climate forcings, 2) unpredictable internal variability that differs from observations (e.g., as might happen if the sampled historical period happens to coincide with the positive phase of the Pacific decadal oscillation in observations and the negative phase in the climate model), and 3) errors in convective parameterizations and unresolved subgrid-scale orography. Univariate statistical bias correction algorithms that are applied on a grid cell basis can, in principle, only correct the third source of error. By nudging the large-scale atmospheric state of a GCM toward an atmospheric reanalysis, Eden et al. (2012) were able to remove the influence of the first two sources of error and show that, over much of the globe, simulated precipitation is a good predictor of observed precipitation and that bias correction can, in theory, be used to reduce the third source of error. In reality, the effectiveness of bias correction

* Current affiliation: Climate Data and Analysis Section, Climate Research Division, Environment Canada.

Corresponding author address: Alex J. Cannon, Canadian Centre for Climate Modelling and Analysis, University of Victoria, P.O. Box 1700 STN CSC, Victoria, BC V8W 2Y2, Canada.
E-mail: alex.cannon@ec.gc.ca

methods will be limited by the first and second sources of error, although the influence of the second—called by Maraun (2012) “variability related apparent bias changes”—can be mitigated to some extent, such as by calibrating postprocessing corrections on sufficiently long historical records (Teng et al. 2015).

While not without controversy (Ehret et al. 2012; Maraun 2013), bias correction is, in practice, a common component of climate change impacts studies. Corrections can be to the modeled mean, variance, and also higher moments of a distribution, with many methods now applying bias corrections to all quantiles (Gudmundsson et al. 2012). So-called quantile mapping bias correction algorithms have been reviewed in the context of hydrological impacts studies and have been found to outperform simpler bias correction methods that correct only the mean or mean and variance of precipitation series (Teutschbein and Seibert 2012; Gudmundsson et al. 2012; Chen et al. 2013). Of quantile mapping algorithms, those that rely on nonparametric estimates of quantiles have tended to outperform those that fit a parametric distribution to data (Gudmundsson et al. 2012).

Recently, studies have highlighted a potentially serious problem with this form of bias correction, in particular when used to assess projected climate change impacts, namely that long-term changes in simulated series can artificially deteriorate following quantile mapping (Hagemann et al. 2011; Themeßl et al. 2012; Maraun 2013; Maurer and Pierce 2014). Maurer and Pierce (2014) evaluated the effect of quantile mapping on GCM-simulated precipitation changes between two historical periods. For individual models, modifications of projected trends in seasonal mean precipitation by quantile mapping were, in some locations, as large as the original GCM-projected changes. While not examined outside of a synthetic example, it was noted that quantile mapping can affect trends in extreme quantiles differently than trends in the mean, and that further work is needed to assess the impact of quantile mapping on the tails of the distribution. Using a nonstationary generalized extreme value (GEV) analysis, Maraun (2013) investigated the role of bias correction in modifying relative trends in annual precipitation maxima from a RCM. In this case, the RCM underestimated observed variability, which led to substantial amplification by quantile mapping of modeled trends in extremes that were of large magnitude relative to interannual variability.

Should trends in bias-corrected outputs match those of the model itself? For historical simulations, results from Maurer and Pierce (2014) were inconclusive. For future projections, which are subject to much larger external forcing signals relative to natural variability than the historical experiments analyzed by Maurer and

Pierce (2014), it has been argued that trends should be preserved so that the climate sensitivity of the model is not affected by bias correction (Hempel et al. 2013). In the case of variables related to atmospheric moisture, such as precipitation, preserving the relative change signal is also important for maintaining physical scaling relationships with model-projected temperature changes, for example as indicated by the Clausius–Clapeyron equation [$\sim 7\%$ increase in column water vapor per degree (K) increase in temperature; O’Gorman and Muller 2010].

More fundamentally, the question of whether or not trends in climate model outputs should be preserved depends, in part, on the physical realism of the projected trends. For instance, on a global scale, simulated surface temperature variability on annual to centennial time scales has been found to be consistent with observations in the latest generation of GCMs (see Fig. 9.33 of Flato et al. 2013). On the other hand, it has been demonstrated that RCMs can modify trends at the coarse scale due to improved representation of land–atmosphere interactions and local winds (Teichmann et al. 2013), and, at very high resolution, their ability to resolve local convective processes (Kendon et al. 2014). If one believes that such trends are more physically realistic than those from the driving GCM, then a successful bias correction algorithm applied at the coarse scale should indeed have the ability to modify these trends. This, of course, would require additional information to which a traditional univariate quantile mapping algorithm would not have access. If, however, trends in the climate model are believed to be realistic, a more modest goal then would be to avoid artificial deterioration of trends that arise simply as a statistical artifact of quantile mapping or related methods. For example, the use of trend-preserving quantile mapping methods with subdaily GCM precipitation extremes may be questionable, whereas their application may be appropriate at longer accumulation scales where both observational and model studies show changes that are consistent with Clausius–Clapeyron scaling (Westra et al. 2014). With this view, the objective would then be for the bias correction to correct systematic distributional biases relative to historical observations and, to the extent possible subject to this correction, respect the available signal from the climate model by maintaining its model-projected future trends. For precipitation, as mentioned previously, this means preserving relative changes.

In the literature, the trend-preserving bias correction applied by Hempel et al. (2013) to phase 5 of the Coupled Model Intercomparison Project (CMIP5) representative concentration pathway (RCP) GCM experiments (Taylor et al. 2012) was explicitly designed to preserve relative trends in monthly mean precipitation, while correcting daily variability about the trend via quantile mapping. This

preserves monthly trends, but quantile mapping of the anomalies could still modify trends in the daily extremes. Similarly, the detrended quantile mapping algorithm examined by [Bürger et al. \(2013\)](#) uses quantile mapping to bias correct model projections that have had their long-term mean trends removed. Following bias correction, the removed climate change signals are then reintroduced. Again, this will tend to maintain the modeled long-term trend in the mean, but does not guarantee that trends in precipitation extremes, as governed by the tails of the distribution, are preserved. Finally, the quantile delta change or quantile perturbation methods considered for precipitation projections by [Olsson et al. \(2009\)](#), [Willems and Vrac \(2011\)](#), and [Sunyer et al. \(2014\)](#) modify historical observations by superimposing relative trends in quantiles from a climate model overtop the observed series. In this case, modeled trends in all quantiles, including in the tails, are preserved, but the algorithm does not explicitly bias-correct the daily time series from the climate model, instead falling back to adjustment of the observed series. In addition, preserving relative changes in quantiles does not mean that relative changes in the mean will be preserved. Tradeoffs that go along with the choice to preserve relative trends in the mean or in the quantiles of a bias-corrected series have yet to be explored in a systematic manner.

This study aims to assess these tradeoffs, and, more generally, to investigate the extent to which quantile mapping algorithms modify GCM-projected trends in mean precipitation and indices of precipitation extremes. While many of the analyses and arguments made in this paper could also be applied to interval variables, like surface air temperature (measured in °C), our focus is on bias correction of precipitation (and, more generally, ratio variables; i.e., quantitative variables with an absolute zero; [Finkelstein and Leaning 1984](#)) and hence exclusively on preservation of relative changes in precipitation quantiles by quantile mapping. As we show in [appendix A](#), it is straightforward to preserve absolute changes in quantiles of interval variables, such as temperature, by quantile mapping. However, given that natural variability relative to projected anthropogenically forced trends is much larger for precipitation than for temperature ([Deser et al. 2012](#)), we have chosen to consider precipitation only.

Also, we note that quantile mapping is often applied for two very different reasons: 1) as a bias correction applied to climate model and observed fields at similar scales and 2) for downscaling from coarse climate model scales to finer observed scales. In this study, we are only concerned with quantile mapping as a bias correction algorithm, that is, when the observed and modeled data

have comparable spatial resolutions or have been appropriately regridded to the same resolution, for instance as is common when quantile mapping is applied as the bias correction step of a larger downscaling framework ([Wood et al. 2004](#)). Because precipitation varies on scales that are much shorter than the grid spacing of a typical climate model, finescale spatial variability can be distorted substantially when quantile mapping is used directly for downscaling purposes ([Maraun 2013](#); [Gutmann et al. 2014](#)). Furthermore, in regions of complex terrain, climate model simulations may systematically displace precipitation features (e.g., rain shadows) such that univariate bias corrections applied to locally coincident pairs of observed and model grid cells may be less appropriate than nonlocal corrections ([Maraun and Widmann 2015](#)). More generally, univariate quantile mapping, which is applied on a grid cell basis, can distort spatial variability of the underlying climate model, potentially leading to inconsistency in modes of spatial variability and underlying model physics ([Bürger et al. 2011](#); [Rocheta et al. 2014](#)). This points to the need for a multivariate bias correction algorithm rather than the univariate methods under consideration here. However, we note that univariate algorithms can be used in a nonlocal context through careful pairing of observed and model grid cells, and further that multivariate bias corrections often rely on univariate quantile mapping algorithms as a key part of their application ([Li et al. 2014](#); [Vrac and Friederichs 2015](#)).

Following [Olsson et al. \(2009\)](#), [Willems and Vrac \(2011\)](#), and [Bürger et al. \(2013\)](#), we first describe a bias correction algorithm, quantile delta mapping (QDM), that explicitly preserves relative changes in simulated precipitation quantiles from a climate model. Systematic distributional biases relative to observations in a historical baseline period are corrected by first detrending all projected future quantiles from a model and then applying quantile mapping to the detrended series. After detrending and quantile mapping, projected trends in the modeled quantiles are reintroduced on top of the bias-corrected outputs, thus ensuring that the climate sensitivity of the underlying climate model, at least so far as quantiles are concerned, is unaffected by the bias correction. QDM is based on the quantile delta change/perturbation and detrended quantile mapping methods cited above, and, as we show in an appendix, is related to the equidistant CDF matching algorithm of [Li et al. \(2010\)](#). In essence, these methods are quantile mapping forms of the simple “delta change” approach to climate model projections ([Olsson et al. 2009](#)).

Methods are first demonstrated using synthetic data and are then applied to daily GCM-simulated precipitation outputs over Canada. In the latter case, the

ability of the algorithms to correct historical biases and preserve relative changes in mean quantities and annual extremes, is assessed using 1) the suite of precipitation indices recommended by the World Meteorological Organization's Expert Team on Climate Change Detection and Indices (ETCCDI; Zhang et al. 2011) and 2) results from a GEV analysis applied to annual precipitation maxima. QDM is compared against detrended quantile mapping, which is designed to preserve trends in the mean, and standard quantile mapping. There are three important reasons for analyzing algorithm performance in terms of extremes. First, previous studies that have looked at the influence of bias correction on simulated precipitation changes over large spatial domains have focused on the mean rather than extremes (Maurer and Pierce 2014). Second, while bias correction methods are calibrated on daily precipitation series, they are, as argued by Bürger et al. (2012), not explicitly tuned to replicate distributions of annual extremes, so this provides a stringent and relatively independent test of their performance. Finally, in terms of providing relevant information for engineering planning, return period calculations are of key interest (Koutsoyiannis 2004).

2. Quantile mapping and detrended quantile mapping

Quantile mapping (QM) equates cumulative distribution functions (CDFs) $F_{o,h}$ and $F_{m,h}$ of, respectively, observed data $x_{o,h}$, denoted by the subscript o , and modeled data $x_{m,h}$, denoted by the subscript m , in a historical period, denoted by the subscript h . This leads to the following transfer function,

$$\hat{x}_{m,p}(t) = F_{o,h}^{-1}\{F_{m,h}[x_{m,p}(t)]\}, \quad (1)$$

for bias correction of $x_{m,p}(t)$, a modeled value at time t within some projected period, denoted by the subscript p . If CDFs and inverse CDFs (i.e., quantile functions) are estimated empirically from the data, the algorithm can be illustrated with the aid of a quantile–quantile plot, which is the scatterplot between empirical quantiles of observed and modeled data (i.e., the sorted values in each sample when the number of observed and modeled samples are the same). In this case, QM amounts to a lookup table whose entries are found by interpolating between points in the quantile–quantile plot of the historical data. The transfer function is constructed using information from the historical period exclusively; information provided by the future model projections is ignored.

QM, like all statistical postprocessing algorithms, relies strongly on an assumption that the climate model

biases to be corrected are stationary (i.e., that characteristics in the historical period will persist into the future). As it is beyond the scope of this paper to address this assumption, we instead point to studies by Maraun et al. (2010) and Maraun (2012) for more insight.

For empirical CDFs, Eq. (1) is only defined over the historical range of the modeled dataset. If a projected value falls outside the historical range, then some form of extrapolation is required, for example using parametric distributions following Wood et al. (2004) or the constant correction approach of Boé et al. (2007). Regardless, one way in which frequent extrapolation can be avoided is to explicitly account for changes in the projected values, for example by first removing the modeled trend in the long-term mean prior to quantile mapping, which will shift the future distribution so that it tends to lie within the support of the historical distribution, and then reimposing it afterward. For a ratio variable like precipitation, trends are removed and then reimposed by scaling and rescaling:

$$\hat{x}_{m,p}(t) = F_{o,h}^{-1}\left\{F_{m,h}\left[\frac{\bar{x}_{m,h}x_{m,p}(t)}{\bar{x}_{m,p}(t)}\right]\right\}\frac{\bar{x}_{m,p}(t)}{\bar{x}_{m,h}}, \quad (2)$$

where $\bar{x}_{m,h}$ and $\bar{x}_{m,p}(t)$ are, respectively, estimates of the long-term modeled mean over the historical period and at time t in the projected period p . (For an interval variable, trend removal/reimposition would be performed additively rather than multiplicatively.) Relative to QM, detrended quantile mapping (DQM), as defined in Eq. (2), incorporates additional, albeit limited, information about the climate model outputs in the projected period, in this case in the form of the projected mean. Depending on the degree of extrapolation still required after detrending (and the means by which extrapolation is handled), the climate change signal from DQM will tend to match that of the underlying climate model. This applies to the mean, but does not necessarily apply to all quantiles, such as those in the tails of the distribution that define climate extremes.

3. Quantile delta mapping

The QDM for precipitation preserves model-projected relative changes in quantiles, while at the same time correcting systematic biases in quantiles of a modeled series with respect to observed values. Preservation of relative changes follows directly from the quantile delta change (Olsson et al. 2009) and quantile perturbation (Willems and Vrac 2011) methods, both of which apply simulated relative changes in quantiles overtop observed historical series. The idea is similar to DQM [Eq. (2)], except that relative changes in all

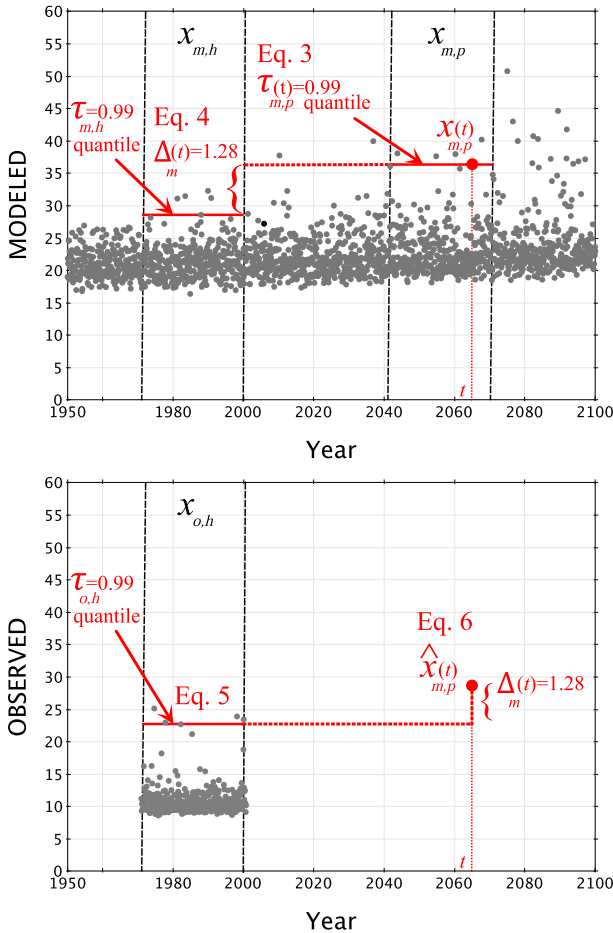


FIG. 1. A graphical illustration of the quantile delta mapping algorithm applied to the modeled data point at time $t = 2065$. (top) Modeled data; two time windows, historical (1971–2000) and projected (2041–70), are demarcated by vertical dashed lines. (bottom) Observed data for the historical period that are used as the basis for bias correction of the modeled data. Labeled equation numbers correspond to those presented in section 3.

modeled quantiles are accounted for rather than only relative changes in the modeled mean. As shown in Fig. 1, the algorithm combines two steps in sequence: first, future model outputs are detrended by quantile and bias corrected to observations by quantile mapping; second, model-projected relative changes in quantiles are superimposed on the bias-corrected model outputs.

QDM starts with the time-dependent CDF of the model projected series $x_{m,p}$, for example as estimated from the empirical CDF over a time window around t :

$$\tau_{m,p}(t) = F_{m,p}^{(t)}[x_{m,p}(t)], \quad \tau_{m,p}(t) \in \{0, 1\}, \quad (3)$$

where $\tau_{m,p}$ is the nonexceedance probability associated with the value at time t .

In the top panel of Fig. 1, a model projected value at time $t = 2065$, $x_{m,p}(2065) = 36.5$, corresponds to a value of $\tau_{m,p}(2065) = 0.99$ (i.e., the 99th percentile of model projected values over the 2050s). The corresponding modeled $\tau_{m,p}$ quantile in the historical period can be found by the entering this value into the historical inverse CDF, $F_{m,h}^{-1}$. In Fig. 1, this yields a value of $F_{m,h}^{-1}(0.99) = 28.5$. The relative change in quantiles between the historical period and time t is then given by

$$\Delta_m(t) = \frac{F_{m,p}^{(t)-1}[\tau_{m,p}(t)]}{F_{m,h}^{-1}[\tau_{m,p}(t)]} = \frac{x_{m,p}(t)}{F_{m,h}^{-1}[\tau_{m,p}(t)]} \quad (4)$$

or, in this example, $\Delta_m(2065) = 36.5/28.5 = 1.28$. The modeled $\tau_{m,p}$ quantile at time t can be bias corrected by applying the inverse CDF estimated from observed values $x_{o,h}$ over the historical period

$$\hat{x}_{o,m,h;p}(t) = F_{o,h}^{-1}[\tau_{m,p}(t)], \quad (5)$$

which, as shown in the bottom panel of Fig. 1, yields a value of $\hat{x}_{o,m,h;p}(2065) = 22.6$. Finally, the bias-corrected future projection at time t is given by applying the relative change $\Delta_m(t)$ multiplicatively to this historical bias-corrected value,

$$\hat{x}_{m,p}(t) = \hat{x}_{o,m,h;p}(t)\Delta_m(t), \quad (6)$$

which, in Fig. 1, gives a value of $\hat{x}_{m,p}(2065) = 22.6 \times 1.28 = 28.9$. To preserve absolute rather than relative changes in quantiles, Eqs. (4) and (6) can simply be applied additively rather than multiplicatively. In appendix A, we show that a variety of bias correction methods, despite differences in motivation, description, and presentation, share this property; depending on formulation, they will either preserve relative or absolute changes in modeled quantiles.

Detrending by quantile and then quantile mapping means that $x_{o,m,h;p}$ takes on the statistical characteristics of the historical observations; this is where biases between the modeled and observed quantiles are corrected. Multiplying the bias-corrected series by Δ_m reintroduces the projected climate change signal; this is where relative changes in modeled quantiles are preserved. The bias correction reduces to standard QM if the modeled distribution does not change between the historical and projected periods. If distributions are estimated empirically based on series of equal lengths, for example 30-yr time slices, then there is a one-to-one correspondence between the sorted historical and projected values. As a result, model-projected values that lie outside the range of the historical period are dealt with as part of the algorithm when the climate change signal is reintroduced in Eq. (6).

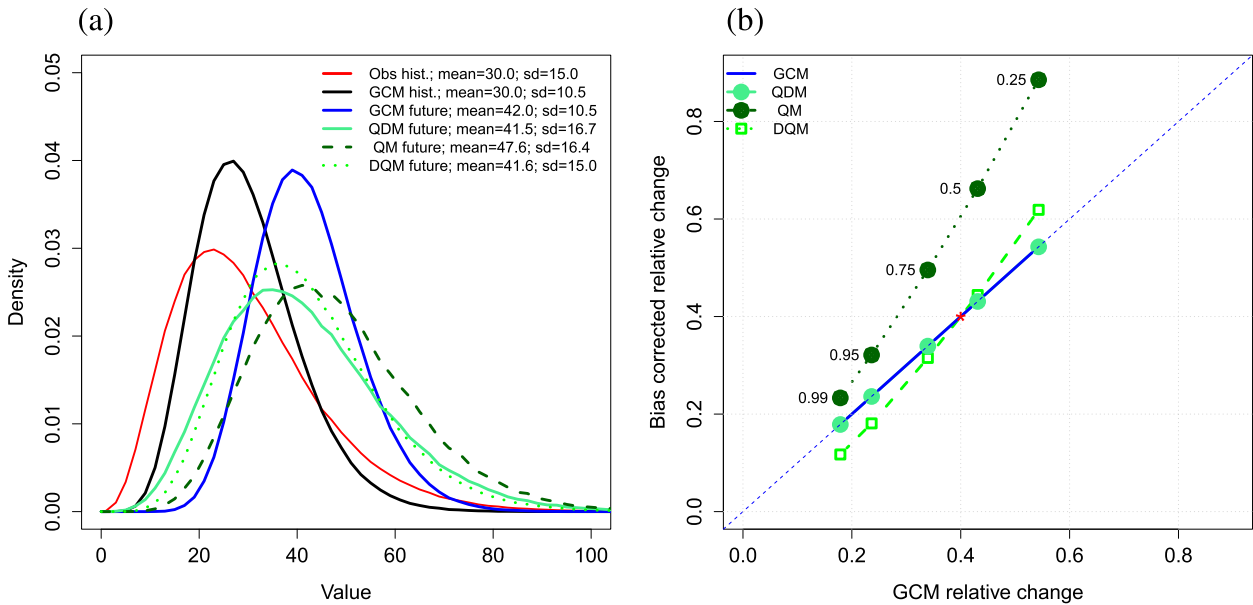


FIG. 2. (a) Probability density functions for the synthetic observed, GCM, QM, DQM, and QDM outputs in section 4 and (b) corresponding relative changes in the $\tau = 0.25, 0.5, 0.75, 0.95,$ and 0.99 quantiles between the historical and future periods for the synthetic GCM and bias-corrected QM, DQM, and QDM data. If the bias correction preserves relative changes, points should lie on the 1:1 line. The red asterisk shows the prescribed 40% relative change in the GCM mean.

4. Synthetic example

Following the example presented by Maurer and Pierce (2014), consider synthetic observed and GCM outputs generated by a gamma distribution with shape parameter k and scale parameter θ , $X \sim \text{gamma}(k, \theta)$, which has the probability density function

$$f(x; k, \theta) = \frac{x^{k-1} \exp(-x/\theta)}{\theta^k \Gamma(k)}, \quad x > 0, \quad k > 0, \quad \theta > 0, \quad (7)$$

where $\Gamma(k)$ is the gamma function evaluated at k . The mean of the distribution is given by $\mu = k\theta$ and the standard deviation by $\sigma = \sqrt{k\theta}$. Synthetic observed historical data $x_{o,h}$ are drawn from a gamma(4, 7.5) distribution ($\mu = 30, \sigma = 15$), historical GCM data $x_{m,h}$ from a gamma(8.15, 3.68) distribution ($\mu = 30, \sigma = 10.5$), and projected GCM data $x_{m,p}$ from a gamma(16, 2.63) distribution ($\mu = 42, \sigma = 10.5$). In this synthetic example, the observed and modeled data for the historical period have the same mean, but the GCM underestimates the observed standard deviation by 30%; the model projects a 40% increase in the mean for the future period, with no accompanying change in the standard deviation. Probability density functions for the distributions are shown in Fig. 2a.

Future distributions obtained by applying the QM, DQM, and QDM bias correction algorithms to the synthetic outputs are shown in Fig. 2a; as all three algorithms

perfectly reproduce the observed distribution in red, bias-corrected historical distributions are not shown separately. As demonstrated by Maurer and Pierce (2014), underestimation of variance for a bounded, right-skewed variable will result in an amplification of projected trends in the mean by QM (and vice versa for overestimation of variance), in this case leading to a projected increase of 58.6% versus 40% for the raw GCM. For these distributions, neither DQM nor QDM leads to similar inflation of the trend magnitude for the mean. Figure 2b shows bias-corrected relative changes in $\tau = 0.25, 0.5, 0.75, 0.95,$ and 0.99 quantiles between the historical and future periods, with deviations from the 1:1 line indicating modification of the underlying GCM trends by the bias correction algorithm. By design, the relative change in each quantile is preserved perfectly by QDM, whereas QM and DQM can both distort the modeled change signals, in some instances, substantially.

In this particular example, both DQM and QDM closely reproduce the prescribed relative change in the mean, with QDM also reproducing changes in the quantiles. In the case of DQM, note that correcting for the mean does not correct for the entire distribution. Similarly, for QDM, relative changes in the mean may not, in the more general case, be preserved, as this property only strictly applies to relative changes in the quantiles (including the median). In both algorithms, all distributional properties (mean, quantiles, etc.) are forced to match those of observations in the historical period, but only the future change signals in

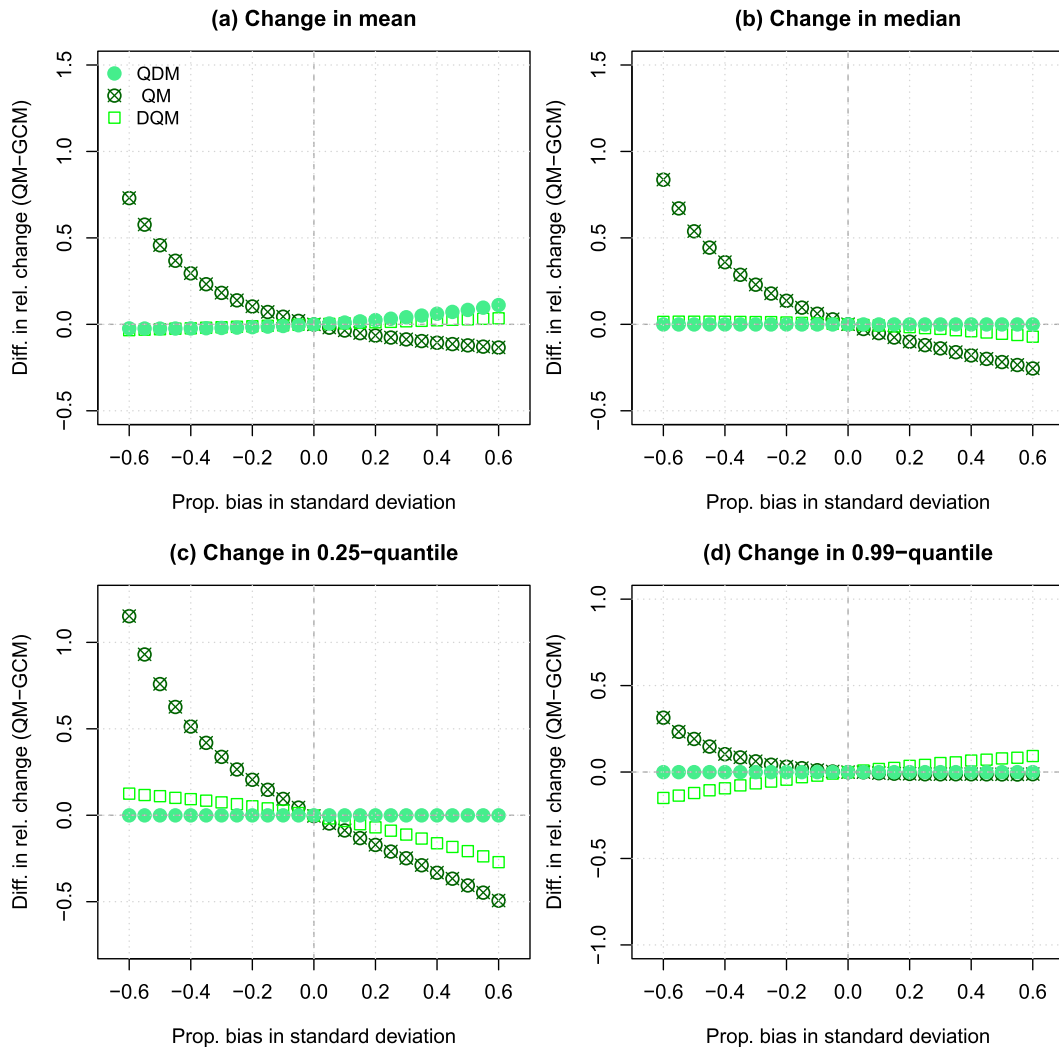


FIG. 3. Plots showing differences between QM, DQM, and QDM relative changes in the (a) mean, (b) median, (c) $\tau = 0.25$ quantile, and (d) $\tau = 0.99$ quantile with respect to GCM-projected relative changes. If a bias correction algorithm preserves GCM-projected relative changes, points should lie along a horizontal line along zero. Synthetic data are based on the distributions from section 4, but with the historical and future GCM standard deviation set between -60% and $+60\%$ of the historical observed standard deviation.

particular properties (mean for DQM and quantiles for QDM) are guaranteed to match those from the GCM following bias correction.

To illustrate, consider the following example based on the synthetic outputs above. In this case, outputs again follow gamma distributions, except that the historical and future standard deviation for the GCM in each simulation is adjusted to lie between $\pm 60\%$ of the historical observed standard deviation, rather than being fixed at -30% . Again, the GCM projects a constant 40% increase in the mean for the future period.

QM, DQM, and QDM bias correction algorithms are applied to the synthetic outputs and results are presented in Fig. 3 for relative changes in the mean, median,

and $\tau = 0.25$ and $\tau = 0.99$ quantiles. Results for QM are consistent with those reported by Maurer and Pierce (2014). Model overestimates of variance in the historical period lead to inflated trends in the mean and quantiles after bias correction, whereas underestimates typically lead to suppressed trend magnitudes. Deviations from the GCM-projected trend in the median are small for DQM, whereas differences in changes for the $\tau = 0.25$ and $\tau = 0.99$ quantiles are of larger magnitude. For QDM, trends in the three quantiles are reproduced perfectly, while the mean is also reproduced well. The largest differences from the GCM-projected mean are of approximately equal magnitude to the largest differences in the median noted for DQM. Thus, despite the fact that trends

in the mean are not guaranteed to be preserved by QDM, accounting for changes in all the quantiles in this particular example provides a reasonably tight constraint on the mean. In the next section, we explore how well this holds when bias correcting real GCM outputs.

5. Bias correction of GCM precipitation

a. Data and implementation of methods

In this section, QM, DQM, and QDM bias correction algorithms are applied to daily precipitation outputs from transient GCM runs. Analyses are conducted on simulated daily precipitation amounts from three GCMs, MIROC5 (run r3i1p1) (Watanabe et al. 2010), CanESM2 (run r1i1p1) (Arora et al. 2011), and CCSM4 (run r1i1p1) (Gent et al. 2011), obtained from the CMIP5 archive (Taylor et al. 2012). For each model run, outputs have been extracted over the Canadian landmass for historical (1950–2005) and RCP8.5 (2006–2100) forcing scenarios (van Vuuren et al. 2011) and regridded to a common 1.4° grid. We focus on RCP8.5 as it provides the largest increase in greenhouse gas emissions over time, leading to the highest greenhouse gas concentration levels by the end of the twenty-first century (radiative forcing of 8.5 W m^{-2} in 2100). Canada-wide gridded daily precipitation observations on a $1/12^\circ$ grid were obtained from Natural Resources Canada (McKenney et al. 2011; Hopkinson et al. 2011) and spatially aggregated to the GCM grid spacing. We treat these observations as truth, but note that significant uncertainty has been found for historical precipitation datasets (Sillmann et al. 2013).

QM is applied using the fitQmapQUANT code by Gudmundsson (2014), DQM is applied using the same code but following Eq. (2), and QDM is implemented by the authors. Extrapolation for QM and DQM is handled using the constant correction of Boé et al. (2007). When applying bias correction algorithms to daily precipitation amounts, the mixed discrete-continuous nature of the data needs to be taken into consideration. In this study, dry days are treated as censored values falling below a trace amount of 0.05 mm day^{-1} . Zeros in the observed and modeled data are replaced with nonzero uniform random values below the trace threshold prior to bias correction; following bias correction, values below the threshold are set to zero. This measure corrects biases in wet day frequency. To correct biases in the seasonal cycle, the quantile mapping algorithms are applied to pooled daily data falling within sliding 3-month windows centered on the month of interest. For example, correction of data from December would include days from November to January, correction of data from January would include days from December to February, and so forth. Time-dependent means and

empirical quantiles of projected data are calculated over 30-yr sliding windows centered on the year of interest.

Methods are first tested in terms of their ability to correct historical GCM simulations and replicate the distribution of observed annual ETCCDI precipitation extremes indices (Zhang et al. 2011). Second, the ability of methods to preserve relative changes in suitable precipitation indices is tested on model projections for the twenty-first century under the RCP8.5 forcing scenario. In this context, relative changes in a subset of the ETCCDI indices calculated following bias correction are compared against model-projected changes. Finally, bias correction algorithms are evaluated in terms of their ability to reproduce observed and simulated changes in 20-yr return period magnitudes of annual precipitation maxima estimated using extreme value theory (see appendix B). Because of the large number of ETCCDI indices, results for these variables are summarized in terms of aggregate statistics over the Canada-wide domain for the 2071–2100 (2080s) period only. Spatial patterns for each GCM and bias correction method, as well as analyses for the full time range of the future scenarios, will be presented in more detail when evaluating performance on the 20-yr return period magnitudes.

b. Historical precipitation extremes indices

Following Bürger et al. (2012), performance is first evaluated by applying the bias correction algorithms to daily historical GCM outputs and then checking to see if distributions of the 11 ETCCDI annual precipitation extremes indices shown in Table 1 are consistent with observations. The ETCCDI indices are representative of so-called moderate extremes, rather than the more infrequent extremes considered later by the 20-yr return period analysis. R1mm, R10mm, and R20mm count instances of precipitation amounts above fixed magnitude thresholds, whereas R95pTOT and R99pTOT measure annual accumulated precipitation for events above relative percentile thresholds; CDD and CWD count consecutive days below and above a 1-mm wet day precipitation threshold; Rx1day and Rx5day measure annual maximum 1- and 5-day precipitation amounts; finally, PRCPTOT and SDII are representative of annual mean accumulations and intensities.

The QM, DQM, and QDM algorithms are calibrated using observations from the 1971–2000 period and then ETCCDI indices are calculated over the 1971–2000 calibration period and a split-sample validation period consisting of the combined 1950–70 and 2001–05 out-of-calibration years. For each GCM grid cell, distributions of the ETCCDI indices calculated before and after bias

TABLE 1. The 11 ETCCDI precipitation extremes indices used to validate the historical performance of the bias correction algorithms. The final column indicates whether a given index is used to assess the reproduction of relative changes between historical and future periods. These indices do not depend strongly on fixed magnitude thresholds that would be affected by bias correction. Routines for calculating ETCCDI indices are provided by Bronaugh (2014).

| Index | Description | Relative change |
|---------|--|-----------------|
| Rx1day | Annual maximum 1-day precipitation | x |
| Rx5day | Annual maximum consecutive 5-day precipitation | x |
| SDII | Simple daily precipitation intensity index in wet days | |
| R1mm | Annual count of days when precipitation \geq 1 mm (wet days) | |
| R10mm | Annual count of days when precipitation \geq 10 mm | |
| R20mm | Annual count of days when precipitation \geq 20 mm | |
| CDD | Maximum length of dry spell (consecutive dry days $<$ 1 mm) | |
| CWD | Maximum length of wet spell (consecutive wet days \geq 1 mm) | |
| R95pTOT | Annual total precipitation when daily wet day amount $>$ 95th percentile | x |
| R99pTOT | Annual total precipitation when daily wet day amount $>$ 99th percentile | x |
| PRCPTOT | Annual total precipitation in wet days | x |

correction are compared against distributions calculated from observations using the Kolmogorov–Smirnov (K-S) test at a 1% significance level. The null hypothesis is that the two samples are drawn from the same distribution. A model is considered to pass the diagnostic test at a grid cell if the null hypothesis is not rejected; that is, none of the quantiles of the modeled distribution falls outside the 99% confidence interval of the observed quantiles (Bürger et al. 2012).

K-S test results for the historical period are summarized in Figs. 4 and 5 for both the raw GCM and bias-corrected outputs. Over the calibration period, the median percentage of grid cells passing the K-S test for the 11 ETCCDI indices lies between 40% and 60% for the raw GCM outputs. As DQM and QDM reduce to QM within the calibration sample, the three bias correction methods perform equivalently over the calibration period; the median number of tests passed exceeds 99.5% of grid cells for all GCMs. [Note that the rejection rate is smaller than anticipated based on a nominal 1% significance level. Given that the K-S test is being applied in a diagnostic rather than formal significance testing context, we have chosen to continue with its use for consistency with Bürger et al. (2012).] In the calibration period, median values of the K-S test statistic D , the maximum difference between cumulative distribution

functions, lie between 0.4 and 0.5 for the raw GCM, with the three quantile mapping algorithms reducing this to \sim 0.17. Despite algorithms being applied to daily GCM data, distributions of the annual ETCCDI indices are successfully corrected by quantile mapping. As expected, the raw GCM outputs pass the K-S diagnostic test at a similar rate over the validation period as the calibration period. For reference, Fig. 6 shows the proportion of grid cells passed for each ETCCDI index in the validation period. Validation performance for the bias correction algorithms is relatively stable and consistent across variables and between methods for the three GCMs, in all cases showing a modest drop relative to the calibration period (median of 91% tests passed and a K-S D statistic of 0.29 for QM, DQM, and QDM). Because GCM-projected changes due to external forcings over the historical period are small relative to natural variability, use of additional information about trends in the mean or quantiles by DQM and QDM does not lead to improvements in historical skill relative to QM. In general, all three algorithms are robust and capable of correcting large systematic biases that are present in the GCM representations of the ETCCDI indices over the historical simulation period.

c. Relative changes in precipitation extremes indices

Following validation of historical performance, the ability of the bias correction algorithms to preserve relative changes in annual precipitation indices for future GCM simulations is evaluated using the PRCPTOT, R95pTOT, R99pTOT, Rx1day, and Rx5day indices. Of the 11 ETCCDI indices, these 5 are selected as they do not depend on a large, fixed-magnitude threshold (e.g., R10mm or R20mm) or they are only weakly dependent on the 1-mm wet day threshold. Avoidance of fixed thresholds means that it can reasonably be argued that relative changes in the raw GCM indices should be preserved by bias correction.

Thirty-year climatological mean values of each index are computed at each grid cell for the 1971–2000 (1980s) historical period and the 2080s future time slice from the RCP8.5 scenario. Projected relative changes for the 2080s period are computed with respect to the 1980s for the raw GCM outputs and the QM, DQM, and QDM bias-corrected outputs. Root-mean-square error (RMSE) and bias statistics taken with respect to the GCM-projected relative changes are then calculated over the Canada-wide domain for the different methods. RMSE results are summarized in Fig. 7.

Based on values of RMSE aggregated over the five ETCCDI indices and three GCMs, QDM and DQM outperform QM by a large margin. QDM, which performs additional corrections on the quantiles, reproduces GCM-projected changes marginally better than DQM overall.

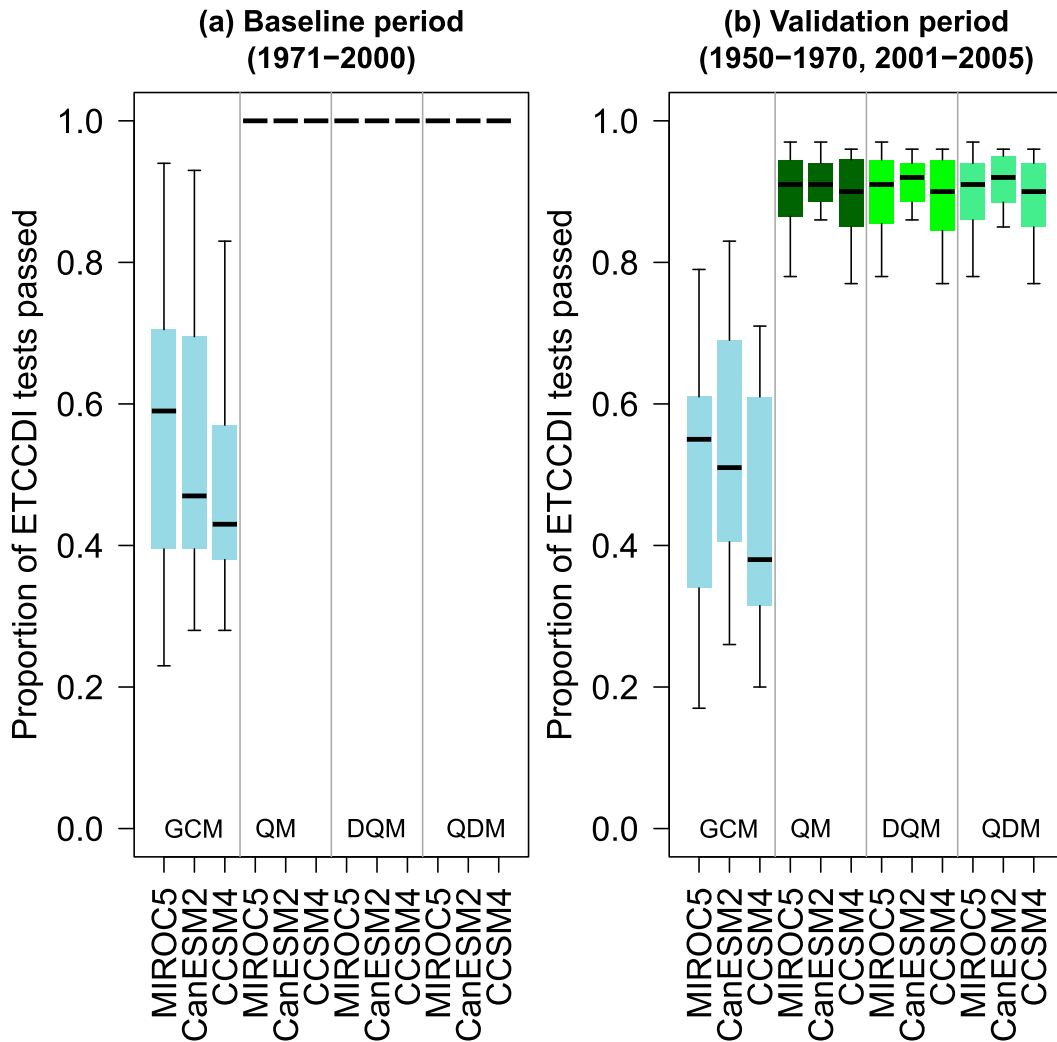


FIG. 4. Boxplots showing distributions, over the 11 ETCCDI indices, of the proportion of grid cells passing local K-S diagnostic tests in the (a) 1971–2000 calibration period and (b) 1950–70 and 2001–05 out-of-sample validation periods. Distributions are summarized for each GCM and bias correction method.

These relative ranks hold for all of the indices except PRCPTOT, the ETCCDI index that is most representative of the mean of the distribution. In this case, QDM is the worst performer for MIROC5 and CanESM2 but is the best for CCSM4. As noted earlier, QDM is not guaranteed to preserve changes in the mean, so poorer PRCPTOT performance for some GCMs is not unexpected.

To investigate performance in more detail, distributions of GCM-projected relative changes in the ETCCDI indices are shown in Fig. 8 for the 2080s; these are accompanied by distributions of QM, DQM, and QDM biases relative to the raw GCM changes. Consistent with RMSE results shown in Fig. 7, biases for DQM and QDM are of similar magnitude (although QDM has a large impact in a small number of cases, particularly for R99pTOT), and are generally much

smaller than those exhibited by QM. QM shows a marked tendency toward heavily right-skewed distributions, characterized by large positive outliers that are of comparable magnitude to the underlying GCM change signals themselves.

d. Generalized extreme value analysis

Next, bias correction methods are evaluated based on their ability to preserve relative changes in estimated magnitudes of 20-yr return period daily maximum precipitation. Twenty-year return magnitudes, rather than those for longer return periods, and hence more extreme events, have been reported for CMIP3 and CMIP5 GCMs by Kharin et al. (2007, 2013) as a means of striking a balance between the rarity of the extreme event and uncertainty in the estimated return magnitudes. This

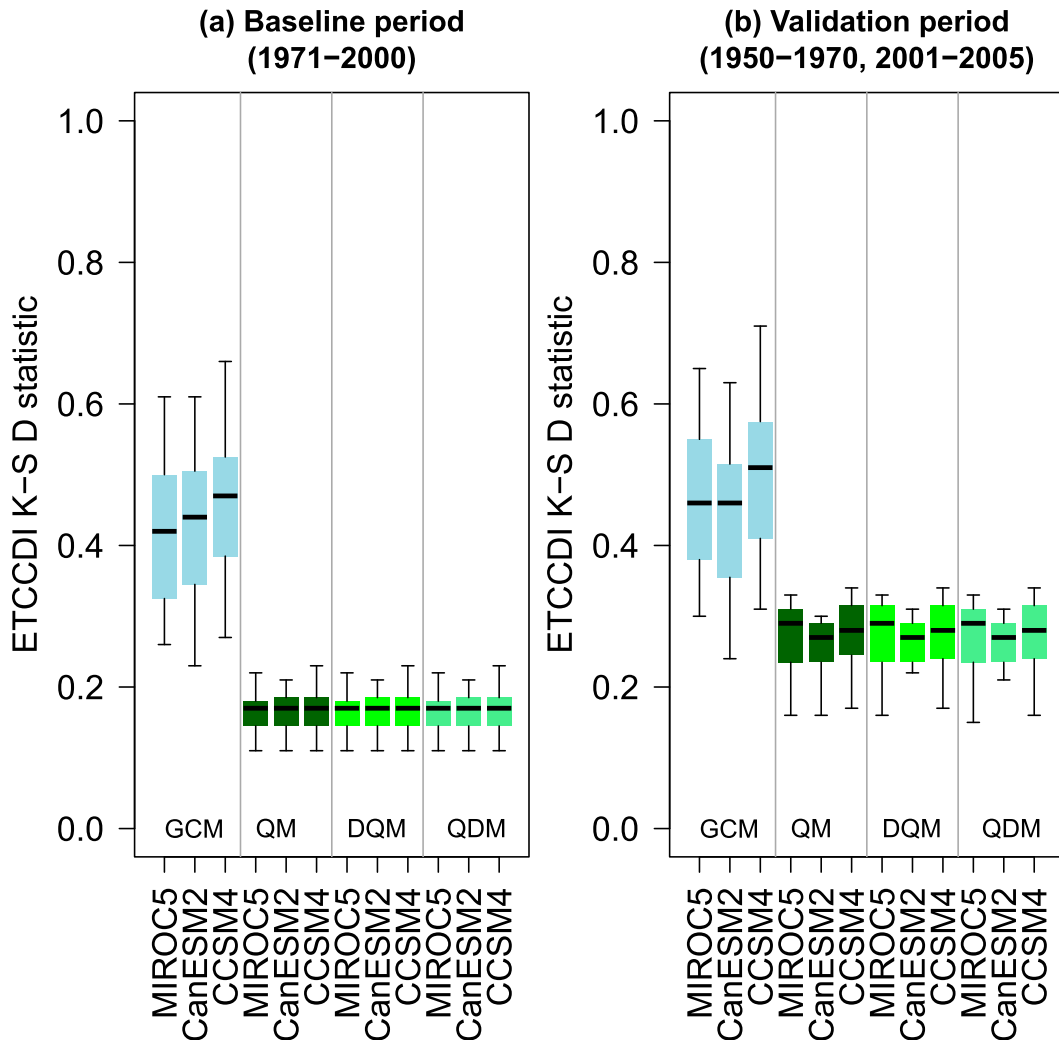


FIG. 5. Boxplots showing distributions, over the 11 ETCCDI indices, of the local K-S D test statistics (the maximum difference between cumulative distribution functions) in the (a) 1971–2000 calibration period and (b) 1950–70 and 2001–05 out-of-sample validation periods. Distributions are summarized for each GCM and bias correction method.

complements the analysis of annual ETCCDI indices presented above, which provides information for more modest extremes.

Return magnitudes are estimated by fitting a stationary GEV distribution to annual Rx1day series via the method of generalized maximum likelihood (Martins and Stedinger 2000), following Cannon (2010), using the shape parameter prior of Papalexiou and Koutsoyiannis (2013) (see appendix B). GEV parameters and return magnitudes are estimated separately for the 1980s historical period, as well as the 2020s (2011–40), 2050s (2041–70), and 2080s periods for the RCP8.5 scenario.

Estimated 20-yr return magnitudes for the 1980s are shown in Fig. 9 for the observations and raw GCM outputs. In terms of spatial patterns, historical 20-yr return magnitudes for the three GCMs are broadly

consistent with observations, exhibiting anomaly correlations between 0.77 and 0.87. However, as indicated by slope and intercept values of best-fit lines between the GCM-simulated and observed magnitudes in Fig. 9, the GCMs, most notably CanESM2 and CCSM4, are biased with respect to observations. A slope value that differs from one indicates the presence of conditional bias (i.e., the bias depends on the magnitude of the modeled value); an intercept value that differs from zero indicates unconditional bias, the degree to which the modeled mean differs from the observed mean after correcting for conditional bias (Stewart and Reagan-Cirincione 1991). Following adjustment of the daily values by QM, DQM, and QDM, biases in the estimated 20-yr return magnitudes of annual maxima in the 1980s calibration period are removed. Maps (not shown) are

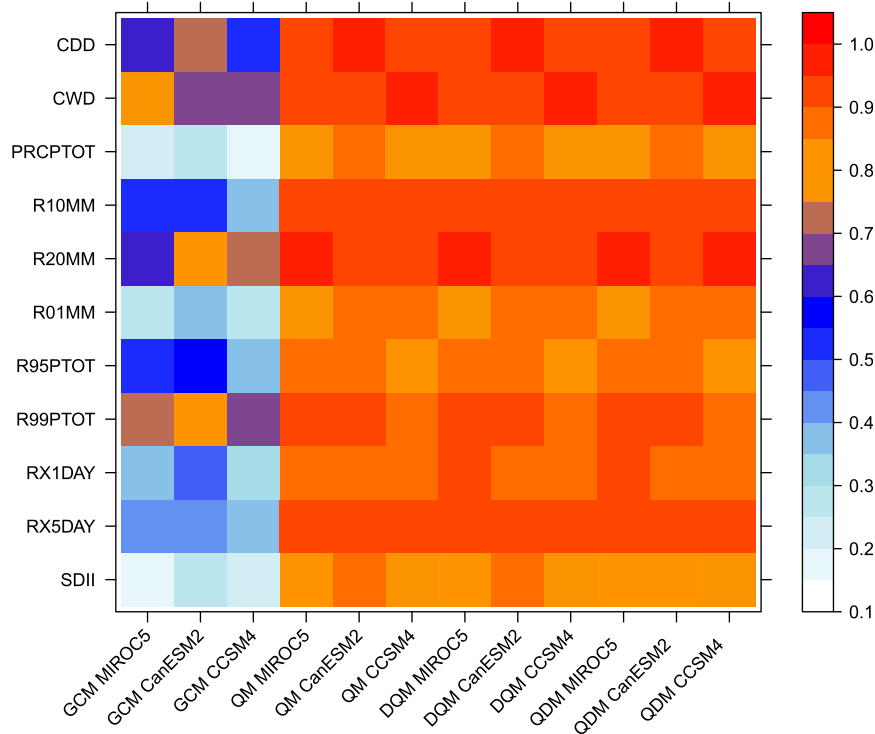


FIG. 6. Plot showing the proportion of grid cells passing local K-S diagnostic tests for each ETCCDI index over the 1950–70 and 2001–05 validation periods.

visually indistinguishable from observations, with anomaly correlations >0.99 , slopes between 0.98 and 1.01, and intercepts between 0.06 and 0.44 mm day⁻¹ for the three bias-corrected GCMs. The corresponding correlation, slope, and intercept values for the 1950–70 and 2001–05 out-of-calibration years have ranges of 0.91–0.93, 1.03–1.07, and 0.92–1.73 mm day⁻¹ respectively, which are substantial improvements relative to the raw GCM values.

After verifying that the bias correction algorithms are capable of correcting historical 20-yr return magnitudes, we turn to performance of the projected change signal. Figure 10 shows distributions of projected relative changes, taken with respect to the 1980s, for raw outputs from the three GCMs, as well as bias-corrected outputs from QM, DQM, and QDM. Results are presented for the 2020s, 2050s, and 2080s time slices. Raw GCM-projected changes in the 2020s are similar across the three GCMs, with increases in the 20-yr return magnitude projected at the majority of grid cells; the median increase over the domain ranges from +5% for CCSM4 to +8% for MIROC5. By the 2050s, this increases to from +10% to +18%, and by the 2080s to from +18% to +37%. The distribution of changes is slightly right skewed, with a small number of positive grid cell outliers seen in most of the GCM/time slice combinations. Relative to the raw GCM outputs, QM modifies the projected relative

change signal by a substantial margin. Median relative changes are amplified by factors of 1.45–2.44, with values for the 2020s, 2050s, and 2080s ranging from +11% to +15%, +19% to +28%, and +29%

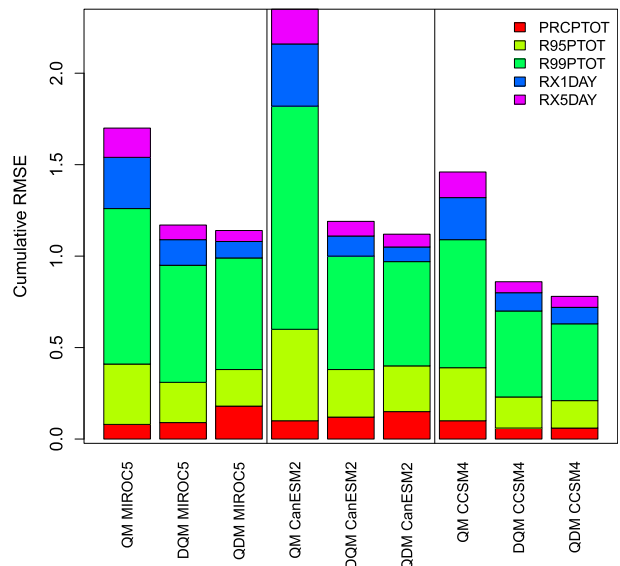


FIG. 7. Stacked bar plots showing cumulative values of RMSE for QM, DQM, and QDM bias-corrected relative changes in ETCCDI indices taken with respect to GCM-projected changes for the 2080s.

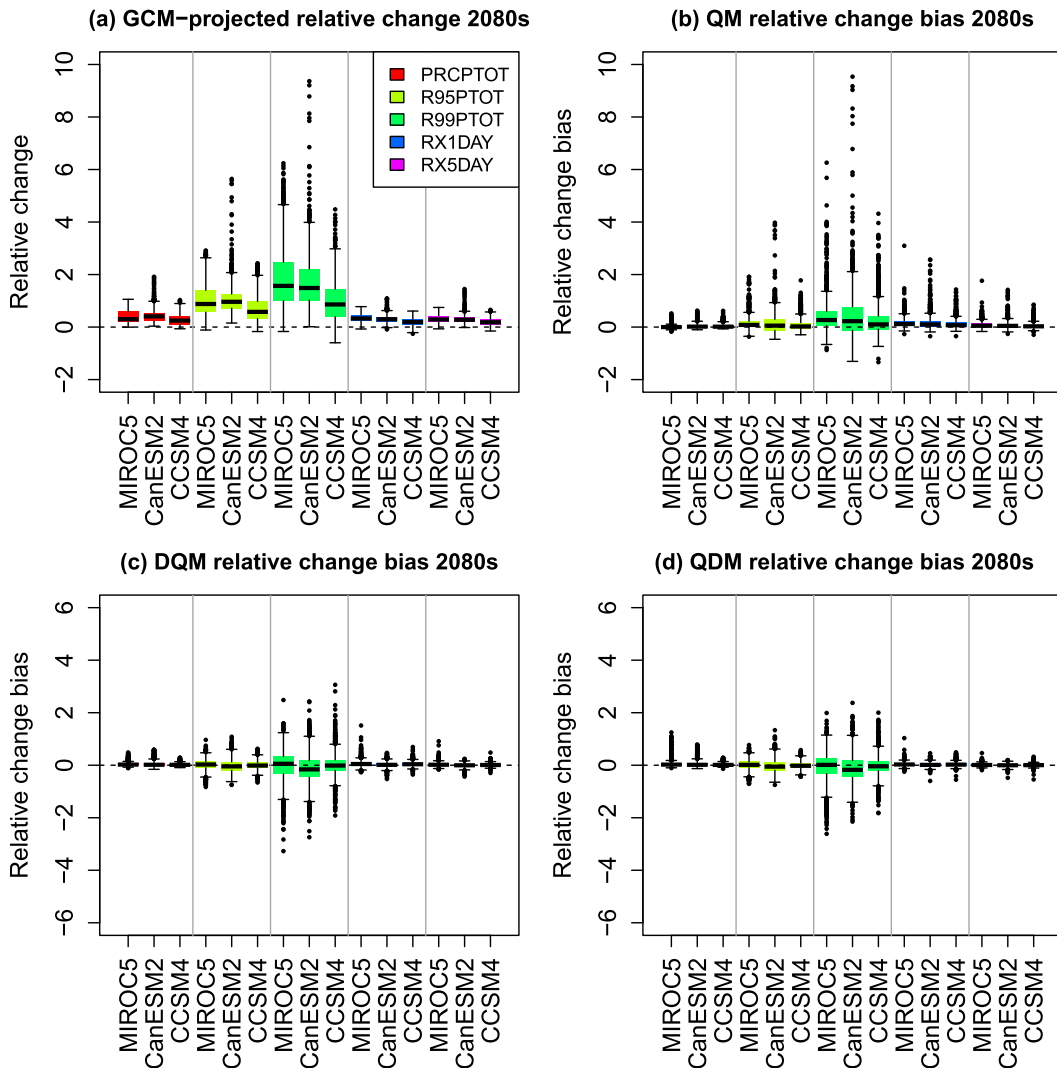


FIG. 8. Boxplots showing the distribution of (a) GCM-projected relative changes in ETCCDI indices at each grid cell, and errors in bias-corrected relative changes taken with respect to GCM-projected changes for (b) QM, (c) DQM, and (d) QDM methods for the 2080s. The median bias over the Canada-wide domain is given by the central line within each box in (b)–(d).

to +53% respectively. The distribution is heavily right skewed; relative changes in excess of +300% are projected at some grid cells in each GCM/time slice combination, with a maximum of +573% noted for MIROC5 in the 2080s. For comparison, the maximum projected change by MIROC5 in the 2080s is just +117%. Note that we do not claim that the GCM-projected trends are physically realistic, only that the QM can alter these trends considerably as an artifact of its application. With that said, taking into account GCM-projected changes in surface air temperature (e.g., median of +6 K for summer and +10 K by 2100 for winter over the northern portion of the study domain; IPCC 2013), the changes in extreme precipitation greater than +300% that are projected by QM are well in excess

of levels anticipated by Clausius–Clapeyron or observed super Clausius–Clapeyron scaling relationships in the extratropics (O’Gorman 2012; Berg et al. 2013).

Consistent with results from the synthetic example in section 4 and Maraun (2013) and Maurer and Pierce (2014), we find that the largest QM outliers tend to occur at grid cells where the raw GCM underestimates, relative to observations, the variance of the fitted GEV distribution in the historical period (appendix B). In this case, underestimation of historical observed variance is noted at 68% of grid cells where the projected difference in relative change between QM and the raw GCM exceeds 200% points (i.e., the arithmetic difference between the two percentages); similarly, all

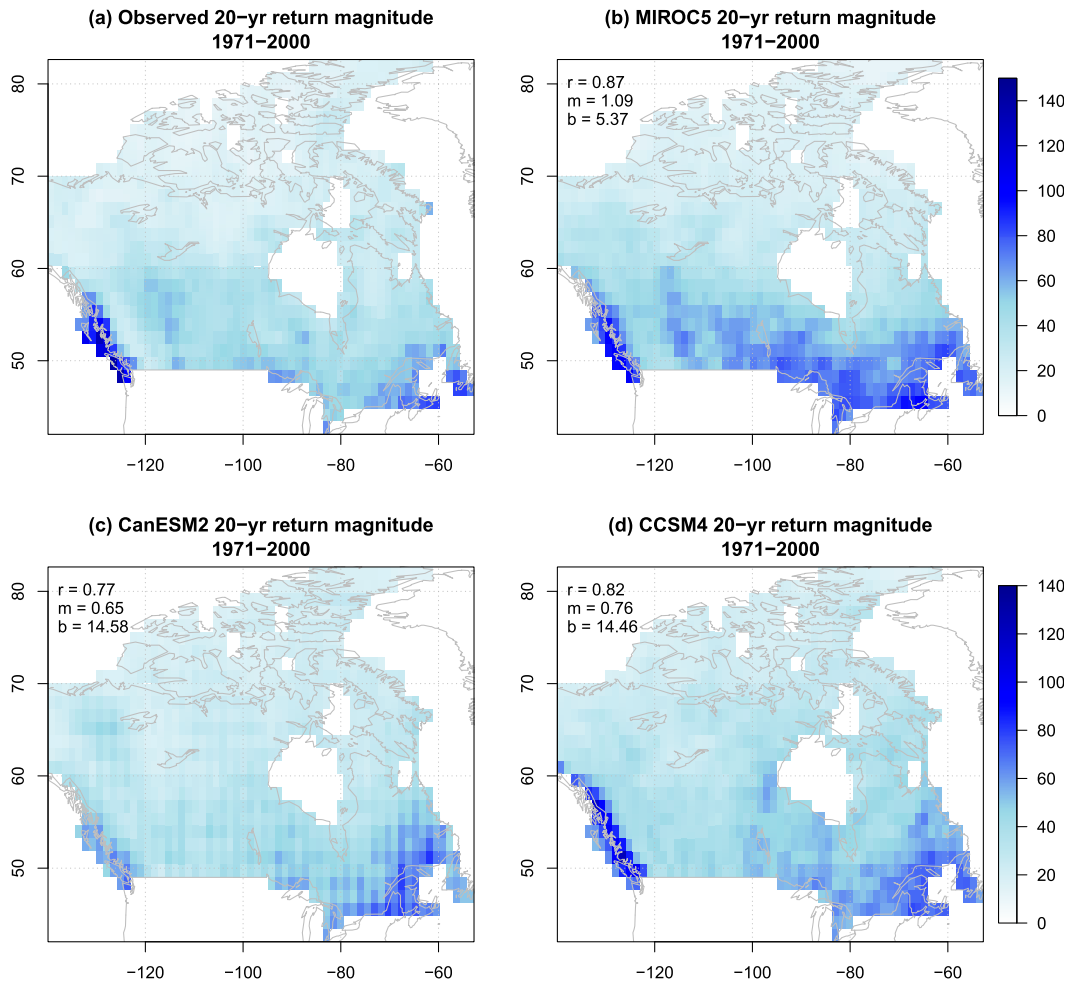


FIG. 9. Estimated 20-yr return magnitudes (mm day^{-1}) for the 1980s historical period based on (a) observations and raw (b) MIROC5, (c) CanESM2, and (d) CCSM4 outputs. Values of the spatial anomaly correlation (r), slope (m), and intercept (b , mm day^{-1}) for the best-fit line between the GCM-simulated and observed magnitudes are shown in the top-left corner of (b)–(d).

projected differences in excess of 300% points occur at grid cells where the GCM underestimates the historical variance.

Distortion of the average change signal is much more modest for DQM, with median relative changes amplified by factors ranging between 1.15 in the 2080s for MIROC5, with the time slice and GCM exhibiting the largest amplitude changes, and 2.07 in the 2020s for CCSM4, with the time slice and GCM with the smallest amplitude changes. While large outliers are still present in the DQM outputs, the largest projected relative change is +241% for MIROC5 in the 2080s, which is less than half that of QM. Finally, distortion of the average change signal is the smallest for QDM, with median relative changes amplified by factors ranging between 1.1 in the 2080s for MIROC5 and 1.7 in the

2020s for CCSM4. Similarly, the number of outliers and their magnitudes are more consistent with the underlying GCMs than for either QM or DQM, with a maximum projected change in 20-yr return magnitude of +138% for MIROC5 in the 2080s.

For reference, spatial patterns of relative changes in 20-yr return magnitude for the 2080s are shown in Figs. 11–13. Anomaly correlations taken with respect to the GCM-projected relative change field are lowest for QM (0.52–0.62), followed by DQM, (0.73–0.82), with QDM most closely reproducing results from the raw GCM (0.82–0.86). Similarly, slopes and intercepts of best-fit lines between the bias-corrected and GCM outputs show that QM suffers from the largest conditional and unconditional bias, with both DQM and QDM replicating the GCM projections more closely.

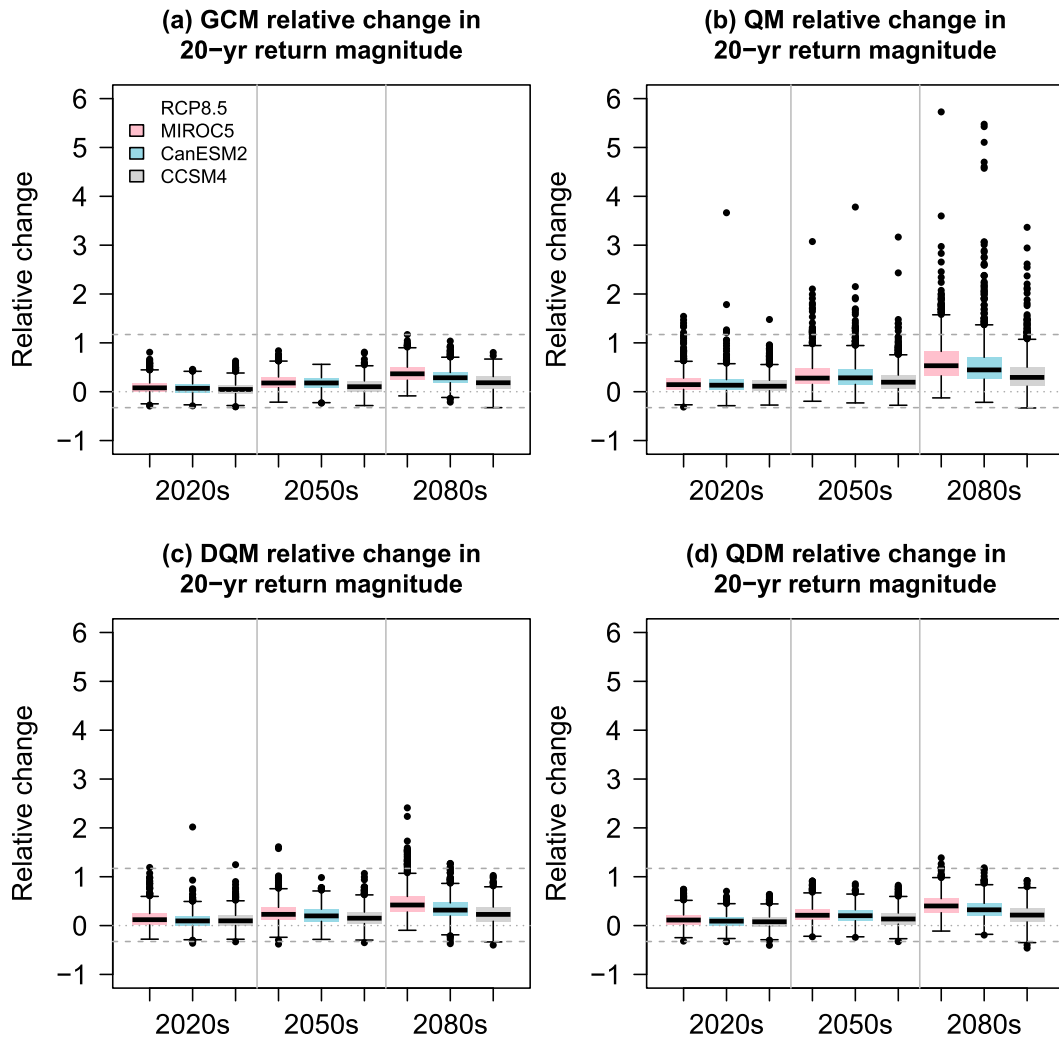


FIG. 10. Boxplots showing distributions of grid cell relative changes in 20-yr annual maximum precipitation return magnitudes for the 2020s, 2050s, and 2080s periods taken with respect to the 1980s historical period for (a) raw GCM, (b) QM, (c) DQM, and (d) QDM outputs. For reference, horizontal dashed lines show the minimum and maximum relative change taken over the GCM grid cells.

6. Conclusions

Quantile mapping algorithms are commonly used to bias correct daily precipitation series from climate models so that distributional properties more closely match those of historical observations. Furthermore, it has been argued that projected trends should be preserved following bias correction so that the underlying model's climate sensitivity is respected. For variables related to atmospheric moisture, like precipitation, preserving the relative change signal is necessary to maintain physical scaling relationships with model-projected temperature changes (e.g., the Clausius–Clapeyron equation). Thus, for precipitation, the goal is to correct systematic distributional biases relative to historical observations, and then, to the extent possible

subject to this correction, maintain model-projected relative changes.

Quantile mapping algorithms, however, have been shown to modify the magnitude of projected trends in mean precipitation quantities (Hagemann et al. 2011; Themeßl et al. 2012; Maraun 2013; Maurer and Pierce 2014). In this study, we investigated the extent to which quantile mapping algorithms corrupt relative trends in precipitation extremes, as characterized by annual ETCCDI indices, 20-yr return magnitudes of annual precipitation maxima, and quantiles lying in the tails of the distribution. Following the quantile delta/perturbation methods (Olsson et al. 2009; Willems and Vrac 2011), we first presented a form of quantile mapping, the QDM bias correction algorithm, that explicitly preserves relative

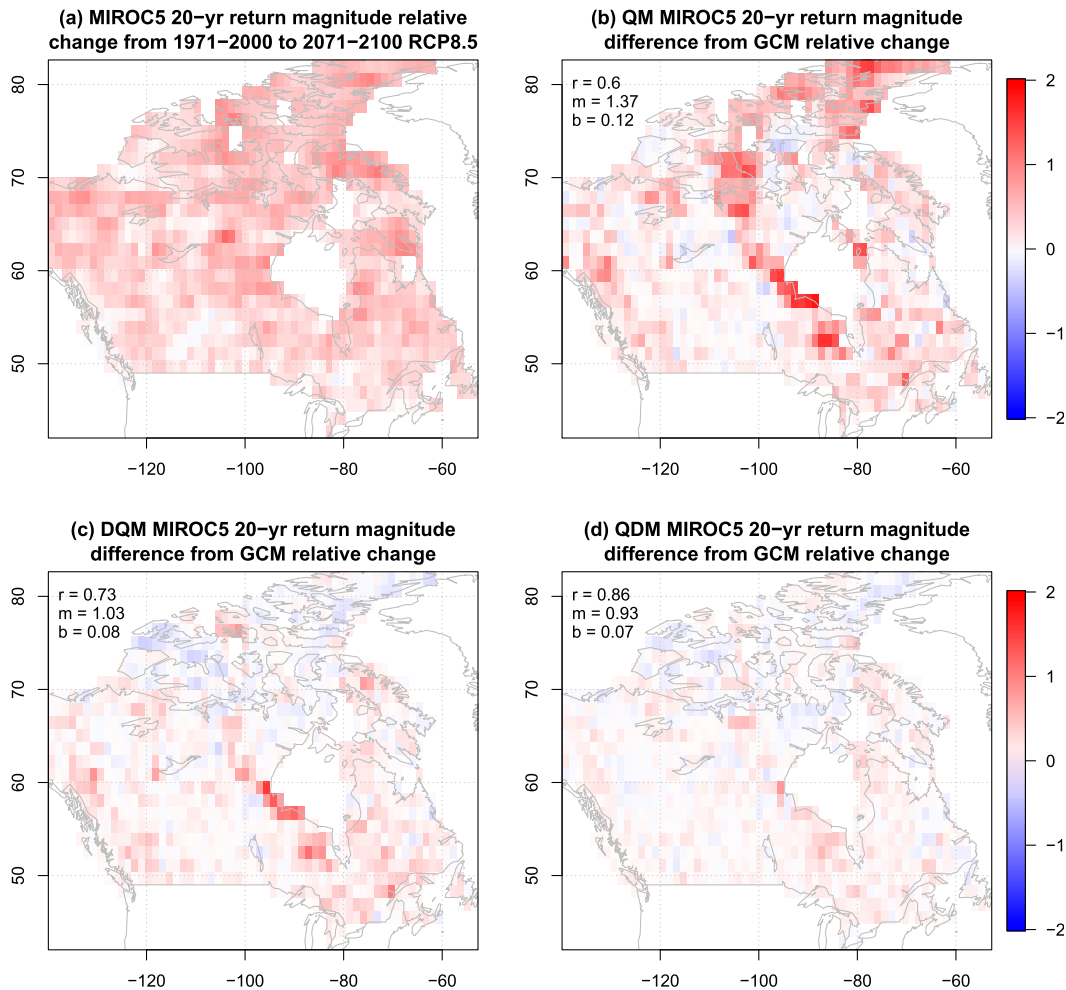


FIG. 11. Spatial distribution of relative changes in 20-yr return magnitude annual maximum precipitation for the 2080s for (a) raw MIROC5 outputs. Differences in relative changes with respect to (a) for 20-yr return magnitude precipitation amounts following bias correction are shown for (b) QM, (c) DQM, and (d) QDM outputs. The color range has been adjusted so that values falling outside the range $\pm 200\%$ are plotted as dark red or dark blue. Values of the spatial anomaly correlation (r), slope (m), and intercept (b) for the best-fit line between the GCM-simulated and bias-corrected changes are shown in the top-left corner of (b)–(d).

changes in all quantiles of a distribution. In [appendix A](#), we showed that the quantile delta/perturbation, QDM, and equidistant/equiratio CDF matching ([Li et al. 2010](#); [Wang and Chen 2014](#)) methods are all fundamentally similar. Depending on formulation, they will either preserve relative or absolute changes in quantiles simulated by a climate model. We then compared QDM against DQM, which preserves changes in the mean, and against the standard QM algorithm on synthetic data and daily CMIP5 projections from three GCMs.

Based on these analyses, QM is found to inflate relative trends in precipitation extremes indices projected by GCMs, often by a substantial amount, whereas DQM and QDM are less prone to this issue; both of these algorithms perform roughly equally well in terms of

reproducing relative changes in the annual ETCCDI indices. For the more extreme 20-yr return values of annual precipitation maxima, QM modifies the projected GCM trends by a very large margin. By the 2080s, relative changes in excess of $+500\%$ are noted at some locations for 20-yr return values, whereas raw GCM changes do not exceed $+120\%$. QM thus inflates the maximum change signal by over a factor of 4. For comparison, maximum changes by DQM and QDM near $+240\%$ and $+140\%$, respectively, an inflation of 2.0 and 1.2 times the raw GCM signal, respectively.

As noted earlier, we do not claim that the GCM-projected or bias-corrected trends are physically realistic, only that the QM can alter the underlying GCM trends considerably as an artifact of its application. If the

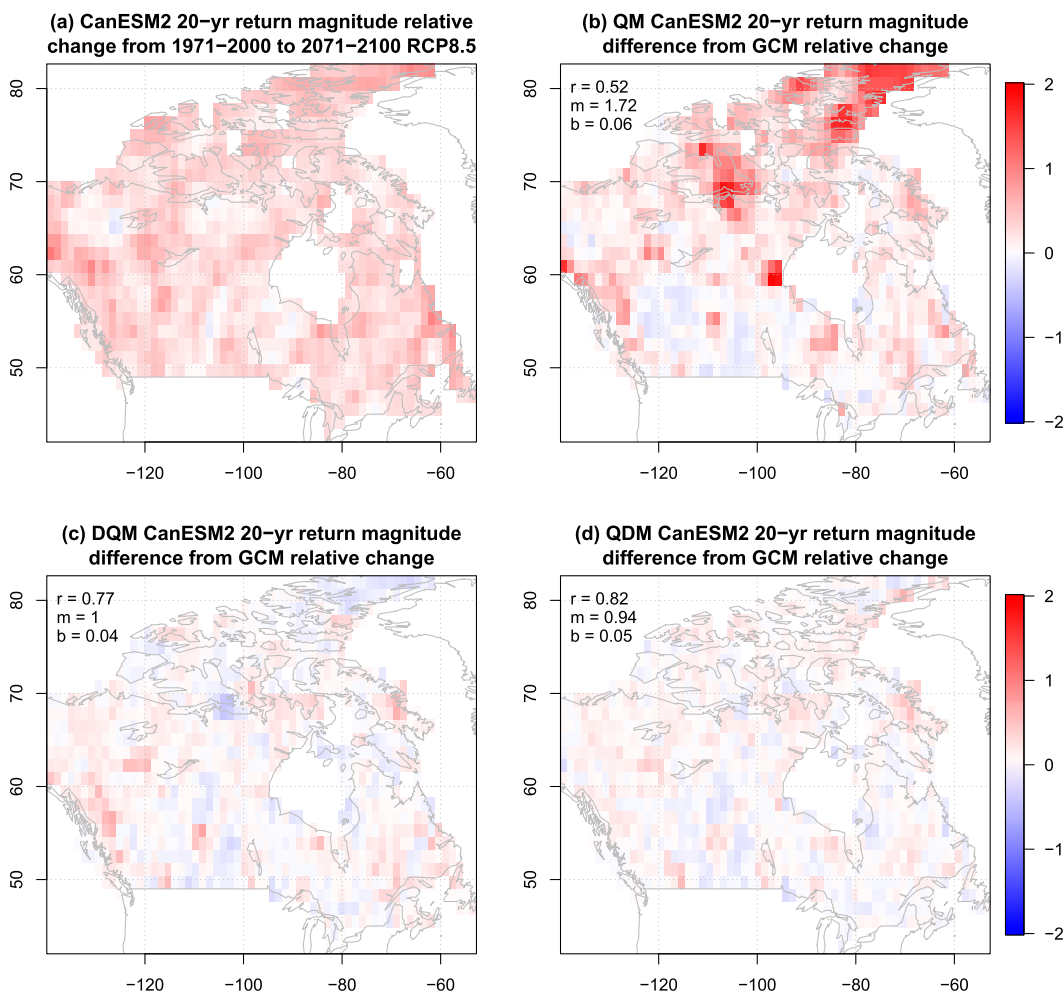


FIG. 12. As in Fig. 11, but for raw and bias-corrected CanESM2 outputs.

goal is to reproduce relative trends in precipitation extremes as originally simulated by a climate model, then the use of standard QM for bias correction cannot be recommended. Instead, some form of quantile mapping that takes into account future trends, minimally in the mean, or, more comprehensively, in all quantiles, is preferred. While it is outside the scope of this paper, one could invoke physical arguments as a means of assessing the realism of projected trends in precipitation extremes by a bias correction algorithm. For example, following Westra et al. (2013), a nonstationary GEV analysis of observed precipitation extremes with temperature anomalies as a covariate could be used to diagnose the observed scaling behavior of precipitation with temperature and assess its adherence to Clausius–Clapeyron and other physical scaling relationships. Climate model output could be analyzed similarly, and an assessment of the extent to which models reproduce observed sensitivities, both historically and in the future, could then be made. If models and observations

show similar sensitivities, but a bias correction algorithm does not, then there would be a strong physical basis for rejecting that algorithm.

In terms of the quantile mapping algorithms, we only considered the simplest of methods, namely calculating empirical CDFs over 30-yr sliding windows, for estimating future distribution and quantile functions. It is possible that more parsimonious methods, such as asynchronous regression (Stoner et al. 2013), may be more robust. Similarly, more flexible time-dependent quantile regression (Koenker and Schorfheide 1994; Cannon 2011) or parametric conditional density estimators (Cannon 2008, 2012), both of which can accommodate mixed discrete-continuous variables such as precipitation, may also perform better in this application. In addition, we did not consider applying bias corrections separately on different time scales, as recommended by Hempel et al. (2013), but note here that quantile mapping algorithms can be applied in this manner. Finally, we recommend that close attention

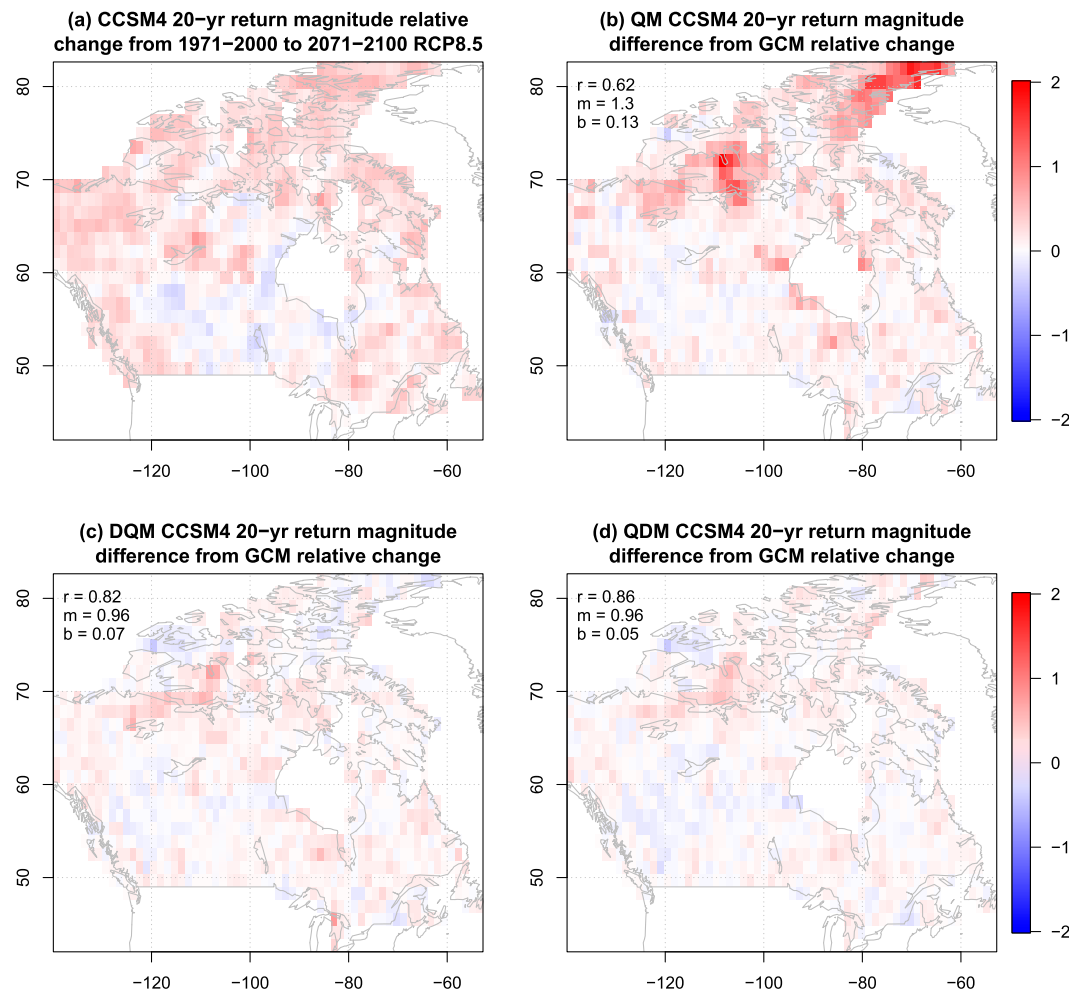


FIG. 13. As in Fig. 11, but for raw and bias-corrected CCSM4 outputs.

be paid to other issues raised in recent critiques of quantile mapping algorithms (Ehret et al. 2012; Maraun 2013). Care must be taken when applying quantile mapping algorithms, as is the case with any postprocessing technique for climate model outputs, to ensure their fitness for the purpose at hand.

Acknowledgments. We thank the Landscape Analysis and Applications section of the Canadian Forest Service, Natural Resources Canada, for developing and making available the historical daily gridded climate data for Canada. We acknowledge the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP5, and we thank the climate modeling groups for producing and making available their GCM output. For CMIP, the U.S. Department of Energy's Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led development of software infrastructure

in partnership with the Global Organization for Earth System Science Portals. Dr. F.W. Zwiers (PCIC) gave comments on the draft manuscript and provided valuable insight on the potential for inflation of extremes by statistical bias correction algorithms. Finally, we gratefully acknowledge the constructive comments and suggestions of the anonymous reviewers.

APPENDIX A

QDM and Equidistant CDF Matching

As shown in the body of the paper, the QDM method described in section 3 preserves the relative change signal in modeled quantiles of a ratio variable, with the underlying algorithm following directly from the quantile perturbation/quantile delta change and DQM methods (Olsson et al. 2009; Willems and Vrac 2011; Bürger et al. 2013). We show here that an additive form

of the QDM algorithm, despite differences in motivation, description, and presentation, is equivalent to the equidistant CDF matching algorithm of Li et al. (2010). Equidistant CDF matching is described by Li et al. (2010, p. 6) as follows: “For a given percentile, we assume that the difference between the model and observed value during the training period also applies to the future period, which means the adjustment function remains the same. However, the difference between the CDFs for the future and historic periods is also taken into account.” Similarly, Wang and Chen (2014, p. 1) state: “Though it incorporates the change in distribution, the fundamental assumption of this method is that the difference between modeled and observed values over the reference period will be preserved in a future period.” The emphasis is foremost on the historical bias in the climate model. Following Olsson et al. (2009) and Willems and Vrac (2011), QDM focuses on preserving the future change signal in quantiles simulated by a climate model. Given that equidistant CDF matching starts with its eye on the historical bias and QDM on the future change signal, it might be expected that the two approaches are fundamentally different. As we show here, however, this is not the case.

To demonstrate for additive changes, we follow the presentation in section 3. The nonexceedance probability of a model projected variable $x_{m,p}$ at time t is

$$\tau_{m,p}(t) = F_{m,p}^{(t)}[x_{m,p}(t)], \quad \tau_{m,p}(t) \in \{0, 1\}, \quad (\text{A1})$$

with the corresponding modeled $\tau_{m,p}$ quantile in the historical period given by entering this value into the historical inverse CDF, $F_{m,h}^{-1}$. The absolute change in quantiles between the historical and future periods is then given by

$$\begin{aligned} \Delta_m(t) &= F_{m,p}^{(t)-1}[\tau_{m,p}(t)] - F_{m,h}^{-1}[\tau_{m,p}(t)] \\ &= x_{m,p}(t) - F_{m,h}^{-1}[\tau_{m,p}(t)]. \end{aligned} \quad (\text{A2})$$

The modeled $\tau_{m,p}$ quantile at time t is bias corrected by applying the inverse CDF estimated from observed values $x_{o,h}$ over the historical period

$$\hat{x}_{o,m,h,p}(t) = F_{o,h}^{-1}[\tau_{m,p}(t)]. \quad (\text{A3})$$

Finally, the bias-corrected future projection at time t is given by adding the change signal in quantiles $\Delta_m(t)$ to this historical bias-corrected value

$$\hat{x}_{m,p}(t) = \hat{x}_{o,m,h,p}(t) + \Delta_m(t) \quad (\text{A4})$$

which means that model-projected absolute changes are preserved. Combining Eqs. (A1)–(A4) gives

$$\begin{aligned} \hat{x}_{m,p}(t) &= F_{o,h}^{-1}\{F_{m,p}^{(t)}[x_{m,p}(t)]\} + x_{m,p}(t) \\ &\quad - F_{m,h}^{-1}\{F_{m,p}^{(t)}[x_{m,p}(t)]\}. \end{aligned}$$

After rearranging, this results in

$$\begin{aligned} \hat{x}_{m,p}(t) &= x_{m,p}(t) + F_{o,h}^{-1}\{F_{m,p}^{(t)}[x_{m,p}(t)]\} \\ &\quad - F_{m,h}^{-1}\{F_{m,p}^{(t)}[x_{m,p}(t)]\}, \end{aligned}$$

which is the same as the equidistant CDF matching method of Li et al. [2010, their Eq. (2)]. Thus, the additive form of QDM and equidistant CDF matching both preserve absolute changes in quantiles, a property which may not have been clear in the original exposition of equidistant CDF matching (and which is more suited to variables like temperature rather than precipitation). Finally, we note that a multiplicative (i.e., ratio) rather than additive form of equidistant CDF matching was raised in passing by Li et al. (2010, p. 10)¹: “For precipitation, an issue arises in some dry regions where the model has essentially no skill at all [...] For these grid points, we use the ratio of the future to the current climate in the model instead of the difference as the transfer function to adjust the observed CDFs.” Thus, the equidistant/equiratio CDF matching, quantile delta/perturbation, and QDM methods are all fundamentally similar. Depending on formulation, they will either preserve relative or absolute changes—the “deltas”—in quantiles simulated by a climate model.

APPENDIX B

Generalized Extreme Value Distribution

Methods used to fit the GEV distribution parallel those described by Cannon (2015a,b) and the text here follows from these references. The extreme value theorem states that the block maximum of a sequence of properly normalized independent and identically distributed random variables can only converge to the GEV distribution as the number of variables tends to infinity. Because of its role as a limiting distribution for block maxima, the GEV distribution is often recommended for analyzing annual precipitation extremes (Coles 2001). The GEV distribution, $\text{GEV}(\xi, \alpha, \kappa)$, is specified by three parameters: the location ξ , the scale α ($\alpha > 0$), and the shape κ . The probability density function of a series x drawn from a GEV distribution is given by

¹ Note that this equiratio form of CDF matching was rediscovered by Wang and Chen (2014).

$$f(x; \xi, \alpha, \kappa) = \frac{1}{\alpha} \left\{ 1 - \kappa \frac{(x - \xi)}{\alpha} \right\}^{(1-\kappa)/\kappa} \exp \left[- \left\{ 1 - \kappa \frac{(x - \xi)}{\alpha} \right\}^{1/\kappa} \right], \quad \kappa \neq 0, \quad 1 - \kappa \frac{(x - \xi)}{\alpha} > 0 \quad \text{and}$$

$$f(x; \xi, \alpha, \mu) = \frac{1}{\alpha} \exp \left[- \frac{(x - \xi)}{\alpha} \right] \exp \left[- \exp \left\{ - \frac{(x - \xi)}{\alpha} \right\} \right], \quad \kappa = 0, \quad (B1)$$

with the corresponding cumulative distribution function (CDF) given by

$$F(x; \xi, \alpha, \kappa) = \exp \left[- \left\{ 1 - \kappa \frac{(x - \xi)}{\alpha} \right\}^{1/\kappa} \right], \quad \kappa \neq 0 \quad \text{and}$$

$$F(x; \xi, \alpha, \kappa) = \exp \left[- \exp \left\{ - \frac{(x - \xi)}{\alpha} \right\} \right], \quad \kappa = 0. \quad (B2)$$

The inverse CDF or quantile function, which determines x_τ quantiles of the GEV distribution with probability $\tau \in \{0, 1\}$ —that is, for a return period of $T = 1/(1 - \tau)$ years—is given by

$$x_\tau = F^{-1}(\tau; \xi, \alpha, \kappa) = \xi + \frac{\alpha}{\kappa} [1 - (-\log \tau)^\kappa], \quad \kappa \neq 0 \quad \text{and}$$

$$x_\tau = F^{-1}(\tau; \xi, \alpha, \kappa) = \xi - \alpha \log(-\log \tau), \quad \kappa = 0. \quad (B3)$$

For later reference, the variance of the GEV distribution for $\kappa \neq 0$, $\kappa > -0.5$ is given by $\alpha^2(g_2 - g_1^2)/\kappa^2$ where $g_j = \Gamma(1 + j\kappa)$ and Γ is the gamma function.

The sign of the shape parameter κ determines which of three forms the GEV distribution takes. Following the convention in hydroclimatology, $\kappa = 0$ corresponds to the two parameter Gumbel distribution with unbounded tails; $\kappa < 0$ to the Fréchet distribution, which is bounded below and has a heavy upper tail; and $\kappa > 0$ to the reversed Weibull distribution, which is bounded above. Recent studies on Canadian rainfall series have demonstrated that rainfall extremes are well represented by the Fréchet form of the GEV distribution ($\kappa < 0$) (Adamowski et al. 1996; Alila 1999), which is consistent with analyses of global precipitation records that recommend excluding the reversed Weibull form of the GEV distribution and constraining $\kappa \in \{-0.5, 0\}$ during parameter estimation (Papalexiou and Koutsoyiannis 2013). To accomplish this, we adopt the method of generalized maximum likelihood (Martins and Stedinger 2000), as implemented by Cannon (2010), using a shifted beta(44.7, 71.1) prior for κ following Papalexiou and Koutsoyiannis (2013).

REFERENCES

- Adamowski, K., Y. Alila, and P. J. Pilon, 1996: Regional rainfall distribution for Canada. *Atmos. Res.*, **42**, 75–88, doi:10.1016/0169-8095(95)00054-2.
- Alila, Y., 1999: A hierarchical approach for the regionalization of precipitation annual maxima in Canada. *J. Geophys. Res.*, **104** (D24), 31 645–31 655, doi:10.1029/1999JD900764.
- Arora, V., and Coauthors, 2011: Carbon emission limits required to satisfy future representative concentration pathways of greenhouse gases. *Geophys. Res. Lett.*, **38**, L05805, doi:10.1029/2010GL046270.
- Berg, P., C. Moseley, and J. O. Haerter, 2013: Strong increase in convective precipitation in response to higher temperatures. *Nat. Geosci.*, **6**, 181–185, doi:10.1038/ngeo1731.
- Boé, J., L. Terray, F. Habets, and E. Martin, 2007: Statistical and dynamical downscaling of the Seine basin climate for hydro-meteorological studies. *Int. J. Climatol.*, **27**, 1643–1655, doi:10.1002/joc.1602.
- Bronaugh, D., 2014: climdex.pcic: PCIC implementation of Climdex routines. Pacific Climate Impacts Consortium, R package version 1.1-1 [Available online at <http://CRAN.R-project.org/package=climdex.pcic>].
- Bürger, G., J. Schulla, and A. Werner, 2011: Estimates of future flow, including extremes, of the Columbia River headwaters. *Water Resour. Res.*, **47**, W10520, doi:10.1029/2010WR009716.
- , T. Q. Murdock, A. T. Werner, S. R. Sobie, and A. J. Cannon, 2012: Downscaling extremes—An intercomparison of multiple statistical methods for present climate. *J. Climate*, **25**, 4366–4388, doi:10.1175/JCLI-D-11-00408.1.
- , S. R. Sobie, A. J. Cannon, A. T. Werner, and T. Q. Murdock, 2013: Downscaling extremes: An intercomparison of multiple methods for future climate. *J. Climate*, **26**, 3429–3449, doi:10.1175/JCLI-D-12-00249.1.
- Cannon, A. J., 2008: Probabilistic multisite precipitation downscaling by an expanded Bernoulli-gamma density network. *J. Hydrometeorol.*, **9**, 1284–1300, doi:10.1175/2008JHM960.1.
- , 2010: A flexible nonlinear modelling framework for non-stationary generalized extreme value analysis in hydroclimatology. *Hydrol. Processes*, **24**, 673–685, doi:10.1002/hyp.7506.
- , 2011: Quantile regression neural networks: Implementation in R and application to precipitation downscaling. *Comput. Geosci.*, **37**, 1277–1284, doi:10.1016/j.cageo.2010.07.005.
- , 2012: Neural networks for probabilistic environmental prediction: Conditional Density Estimation Network Creation and Evaluation (CaDENCE) in R. *Comput. Geosci.*, **41**, 126–135, doi:10.1016/j.cageo.2011.08.023.
- , 2015a: An intercomparison of regional and at-site rainfall extreme value analyses in southern British Columbia, Canada. *Can. J. Civ. Eng.*, **42**, 107–119, doi:10.1139/cjce-2014-0361.
- , 2015b: Revisiting the nonlinear relationship between ENSO and winter extreme station precipitation in North America. *Int. J. Climatol.*, doi:10.1002/joc.4263, in press.
- Chen, J., F. P. Brissette, D. Chaumont, and M. Braun, 2013: Finding appropriate bias correction methods in downscaling precipitation for hydrologic impact studies over North America. *Water Resour. Res.*, **49**, 4187–4205, doi:10.1002/wrcr.20331.
- Coles, S., 2001: *An Introduction to Statistical Modeling of Extreme Values*. Springer, 208 pp.

- Deser, C., A. Phillips, V. Bourdette, and H. Teng, 2012: Uncertainty in climate change projections: The role of internal variability. *Climate Dyn.*, **38**, 527–546, doi:10.1007/s00382-010-0977-x.
- Eden, J. M., M. Widmann, D. Grawe, and S. Rast, 2012: Skill, correction, and downscaling of GCM-simulated precipitation. *J. Climate*, **25**, 3970–3984, doi:10.1175/JCLI-D-11-00254.1.
- Ehret, U., E. Zehe, V. Wulfmeyer, K. Warrach-Sagi, and J. Liebert, 2012: HESS opinions: “Should we apply bias correction to global and regional climate model data?” *Hydrol. Earth Syst. Sci.*, **16**, 3391–3404, doi:10.5194/hess-16-3391-2012.
- Finkelstein, L., and M. Leaning, 1984: A review of the fundamental concepts of measurement. *Measurement*, **2**, 25–34, doi:10.1016/0263-2241(84)90020-4.
- Flato, G., and Coauthors, 2013: Evaluation of climate models. *Climate Change 2013: The Physical Science Basis*. T. F. Stocker et al., Eds., Cambridge University Press, 741–866.
- Gent, P. R., and Coauthors, 2011: The Community Climate System Model version 4. *J. Climate*, **24**, 4973–4991, doi:10.1175/2011JCLI4083.1.
- Gudmundsson, L., 2014: qmap: Statistical transformations for post-processing climate model output. R package version 1.0-3 [Available online at <http://cran.r-project.org/web/packages/qmap/>.]
- , J. Bremnes, J. 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.
- Gutmann, E., T. Pruiitt, M. P. Clark, L. Brekke, J. R. Arnold, D. A. Raff, and R. M. Rasmussen, 2014: An intercomparison of statistical downscaling methods used for water resource assessments in the United States. *Water Resour. Res.*, **50**, 7167–7186, doi:10.1002/2014WR015559.
- 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. Hydrometeorol.*, **12**, 556–578, doi:10.1175/2011JHM1336.1.
- Hempel, S., K. Frieler, L. Warszawski, J. Schewe, and F. Piontek, 2013: A trend-preserving bias correction—The ISI-MIP approach. *Earth System Dyn.*, **4**, 219–236, doi:10.5194/esd-4-219-2013.
- Hopkinson, R. F., D. W. McKenney, E. J. Milewska, M. F. Hutchinson, P. Papadopol, and L. A. Vincent, 2011: Impact of aligning climatological day on gridding daily maximum-minimum temperature and precipitation over Canada. *J. Appl. Meteor. Climatol.*, **50**, 1654–1665, doi:10.1175/2011JAMC2684.1.
- IPCC, 2013: Annex I: Atlas of global and regional climate projections (and RCP8.5 supplementary material). *Climate Change 2013: The Physical Science Basis*, T. F. Stocker et al., Eds., Cambridge University Press, 1311–1393. [Available online at <http://www.climatechange2013.org/report/full-report/>.]
- Kendon, E. J., N. M. Roberts, H. J. Fowler, M. J. Roberts, S. C. Chan, and C. A. Senior, 2014: Heavier summer downpours with climate change revealed by weather forecast resolution model. *Nat. Climate Change*, **4**, 570–576, doi:10.1038/nclimate2258.
- Kharin, V. V., F. W. Zwiers, X. Zhang, and G. C. Hegerl, 2007: Changes in temperature and precipitation extremes in the IPCC ensemble of global coupled model simulations. *J. Climate*, **20**, 1419–1444, doi:10.1175/JCLI4066.1.
- , —, —, and M. Wehner, 2013: Changes in temperature and precipitation extremes in the CMIP5 ensemble. *Climatic Change*, **119**, 345–357, doi:10.1007/s10584-013-0705-8.
- Koenker, R., and F. Schorfheide, 1994: Quantile spline models for global temperature change. *Climatic Change*, **28**, 395–404, doi:10.1007/BF01104081.
- Koutsoyiannis, D., 2004: Statistics of extremes and estimation of extreme rainfall: I. Theoretical investigation. *Hydrol. Sci. J.*, **49**, 575–590, doi:10.1623/hysj.49.4.575.54430.
- Li, C., E. Sinha, D. E. Horton, N. S. Diffenbaugh, and A. M. Michalak, 2014: Joint bias correction of temperature and precipitation in climate model simulations. *J. Geophys. Res. Atmos.*, **119**, 13 153–13 162, doi:10.1002/2014JD022514.
- 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.*, **115**, D10101, doi:10.1029/2009JD012882.
- Maraun, D., 2012: Nonstationarities of regional climate model biases in European seasonal mean temperature and precipitation sums. *Geophys. Res. Lett.*, **39**, L06706, doi:10.1029/2012GL051210.
- , 2013: Bias correction, quantile mapping, and downscaling: Revisiting the inflation issue. *J. Climate*, **26**, 2137–2143, doi:10.1175/JCLI-D-12-00821.1.
- , and M. Widmann, 2015: The representation of location by regional climate models in complex terrain. *Hydrol. Earth Syst. Sci. Discuss.*, **12**, 3011–3028, doi:10.5194/hessd-12-3011-2015.
- , and Coauthors, 2010: Precipitation downscaling under climate change: Recent developments to bridge the gap between dynamical models and the end user. *Rev. Geophys.*, **48**, RG3003, doi:10.1029/2009RG000314.
- Martins, E. S., and J. R. Stedinger, 2000: Generalized maximum-likelihood generalized extreme-value quantile estimators for hydrologic data. *Water Resour. Res.*, **36**, 737–744, doi:10.1029/1999WR900330.
- Maurer, E. P., and H. Hidalgo, 2008: Utility of daily vs. monthly large-scale climate data: An intercomparison of two statistical downscaling methods. *Hydrol. Earth Syst. Sci.*, **12**, 551–563, doi:10.5194/hess-12-551-2008.
- , and D. W. Pierce, 2014: Bias correction can modify climate model simulated precipitation changes without adverse effect on the ensemble mean. *Hydrol. Earth Syst. Sci.*, **18**, 915–925, doi:10.5194/hess-18-915-2014.
- , H. Hidalgo, T. Das, M. Dettinger, and D. 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.
- McKenney, D. W., and Coauthors, 2011: Customized spatial climate models for North America. *Bull. Amer. Meteor. Soc.*, **92**, 1611–1622, doi:10.1175/2011BAMS3132.1.
- Mearns, L. O., and Coauthors, 2012: The North American Regional Climate Change Assessment Program: Overview of phase I results. *Bull. Amer. Meteor. Soc.*, **93**, 1337–1362, doi:10.1175/BAMS-D-11-00223.1.
- Muerth, M., and Coauthors, 2013: On the need for bias correction in regional climate scenarios to assess climate change impacts on river runoff. *Hydrol. Earth Syst. Sci.*, **17**, 1189–1204, doi:10.5194/hess-17-1189-2013.
- O’Gorman, P. A., 2012: Sensitivity of tropical precipitation extremes to climate change. *Nat. Geosci.*, **5**, 697–700, doi:10.1038/ngeo1568.
- , and C. Muller, 2010: How closely do changes in surface and column water vapor follow Clausius–Clapeyron scaling in climate change simulations? *Environ. Res. Lett.*, **5**, 025207, doi:10.1088/1748-9326/5/2/025207.

- Olsson, J., K. Berggren, M. Olofsson, and M. Viklander, 2009: Applying climate model precipitation scenarios for urban hydrological assessment: A case study in Kalmar City, Sweden. *Atmos. Res.*, **92**, 364–375, doi:10.1016/j.atmosres.2009.01.015.
- Papalexiou, S. M., and D. Koutsoyiannis, 2013: Battle of extreme value distributions: A global survey on extreme daily rainfall. *Water Resour. Res.*, **49**, 187–201, doi:10.1029/2012WR012557.
- Rocheta, E., J. P. Evans, and A. Sharma, 2014: Assessing atmospheric bias correction for dynamical consistency using potential vorticity. *Environ. Res. Lett.*, **9**, 124010, doi:10.1088/1748-9326/9/12/124010.
- Sillmann, J., V. Kharin, X. Zhang, F. Zwiers, and D. Bronaugh, 2013: Climate extremes indices in the CMIP5 multimodel ensemble: Part 1. Model evaluation in the present climate. *J. Geophys. Res. Atmos.*, **118**, 1716–1733, doi:10.1002/jgrd.50203.
- Stewart, T. R., and P. Reagan-Cirincione, 1991: Coefficients for debiasing forecasts. *Mon. Wea. Rev.*, **119**, 2047–2051, doi:10.1175/1520-0493(1991)119<2047:CFDF>2.0.CO;2.
- Stoner, A. M., K. Hayhoe, X. Yang, and D. J. Wuebbles, 2013: An asynchronous regional regression model for statistical downscaling of daily climate variables. *Int. J. Climatol.*, **33**, 2473–2494, doi:10.1002/joc.3603.
- Sunyer, M. A., and Coauthors, 2014: Inter-comparison of statistical downscaling methods for projection of extreme precipitation in Europe. *Hydrol. Earth Syst. Sci. Discuss.*, **11**, 6167–6214, doi:10.5194/hessd-11-6167-2014.
- Taylor, K. E., R. J. Stouffer, and G. A. Meehl, 2012: An overview of CMIP5 and the experiment design. *Bull. Amer. Meteor. Soc.*, **93**, 485–498, doi:10.1175/BAMS-D-11-00094.1.
- Teichmann, C., and Coauthors, 2013: How does a regional climate model modify the projected climate change signal of the driving GCM: A study over different CORDEX regions using REMO. *Atmosphere*, **4**, 214–236, doi:10.3390/atmos4020214.
- Teng, J., N. Potter, F. Chiew, L. Zhang, B. Wang, J. Vaze, and J. Evans, 2015: How does bias correction of regional climate model precipitation affect modelled runoff? *Hydrol. Earth Syst. Sci.*, **19**, 711–728, doi:10.5194/hess-19-711-2015.
- Teutschbein, C., and J. Seibert, 2012: Bias correction of regional climate model simulations for hydrological climate-change impact studies: Review and evaluation of different methods. *J. Hydrol.*, **456–457**, 12–29, doi:10.1016/j.jhydrol.2012.05.052.
- Thiemeßl, M. J., A. Gobiet, and G. Heinrich, 2012: Empirical-statistical downscaling and error correction of regional climate models and its impact on the climate change signal. *Climatic Change*, **112**, 449–468, doi:10.1007/s10584-011-0224-4.
- van Vuuren, D. P., and Coauthors, 2011: The representative concentration pathways: An overview. *Climatic Change*, **109**, 5–31, doi:10.1007/s10584-011-0148-z.
- Vrac, M., and P. Friederichs, 2015: Multivariate—intervariable, spatial, and temporal—bias correction. *J. Climate*, **28**, 218–237, doi:10.1175/JCLI-D-14-00059.1.
- Wang, L., and W. Chen, 2014: Equiratio cumulative distribution function matching as an improvement to the equidistant approach in bias correction of precipitation. *Atmos. Sci. Lett.*, **15**, 1–6, doi:10.1002/asl2.454.
- Watanabe, M., and Coauthors, 2010: Improved climate simulation by MIROC5: Mean states, variability, and climate sensitivity. *J. Climate*, **23**, 6312–6335, doi:10.1175/2010JCLI3679.1.
- Westra, S., L. V. Alexander, and F. W. Zwiers, 2013: Global increasing trends in annual maximum daily precipitation. *J. Climate*, **26**, 3904–3918, doi:10.1175/JCLI-D-12-00502.1.
- , and Coauthors, 2014: Future changes to the intensity and frequency of short-duration extreme rainfall. *Rev. Geophys.*, **52**, 522–555, doi:10.1002/2014RG000464.
- Willems, P., and M. Vrac, 2011: Statistical precipitation downscaling for small-scale hydrological impact investigations of climate change. *J. Hydrol.*, **402**, 193–205, doi:10.1016/j.jhydrol.2011.02.030.
- Wood, A. W., L. R. Leung, V. Sridhar, and D. Lettenmaier, 2004: Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs. *Climatic Change*, **62**, 189–216, doi:10.1023/B:CLIM.0000013685.99609.9e.
- Zhang, X., L. Alexander, G. C. Hegerl, P. Jones, A. Klein Tank, T. C. Peterson, B. Trewin, and F. W. Zwiers, 2011: Indices for monitoring changes in extremes based on daily temperature and precipitation data. *Wiley Interdiscip. Rev.: Climate Change*, **2**, 851–870, doi:10.1002/wcc.147.