Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2019-139
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 17 April 2019

Hydrology and Earth System Sciences
Open Access
Discussions

Bias in downscaled rainfall characteristics

Nicholas J. Potter[1], Francis H.S. Chiew[1], Stephen P. Charles[2], Guobin Fu[2], Hongxing Zheng[1], Lu Zhang[1]

[1]CSIRO Land and Water, Canberra ACT, 2601, Australia
[2]CSIRO Land and Water, Floreat WA, 6148, Australia

Correspondence to: Nicholas J. Potter (Nick.Potter@csiro.au)

Abstract. Dynamical downscaling of future projections of global climate model outputs can potentially provide useful information about plausible and possible changes to water resource availability, for which there is increasing demand for regional water resource planning processes. By explicitly modelling climate processes within and across global climate model gridcells for a region, dynamical downscaling can provide higher resolution hydroclimate projections, as well as independent (from historical timeseries) and physically plausible future rainfall timeseries for hydrological modelling applications. However, since rainfall is not typically constrained to observations by these methods, there is often a need for bias correction before use in hydrological modelling. Many bias correction methods (such as scaling, empirical and distributional mapping) have been proposed in the literature, but methods that treat daily amounts only (and not sequencing) can result in residual biases in certain rainfall characteristics, which flow through to biases and problems with subsequently modelled runoff. We apply quantile-quantile mapping to rainfall dynamically downscaled by NARCliM in the State of Victoria, Australia and examine the effect of this on: (i) biases both before and after bias correction in different rainfall metrics; (ii) change signals in metrics in comparison to the bias; and (iii) the effect of bias correction on wet-wet and dry-dry transition probabilities. After bias correction, persistence of wet states is under-correlated (i.e. more random than observations), and this results in a significant bias (underestimation) of runoff using hydrological models calibrated on historical data. A novel representation of quantile-quantile mapping is developed based on lag-one transition probabilities of dry and wet states, and we use this to explain residual biases in transition probabilities. This demonstrates that any quantile mapping bias correction methods are unable to correct the underestimation of autocorrelation of rainfall sequencing, which suggests that new methods are needed to properly bias correct dynamical downscaling rainfall outputs.

1 Introduction

There is a growing and on-going need for information about plausible and possible changes to water resource availability in the future due to climate change. Of most interest is possible changes to rainfall characteristics, particularly those that could affect runoff and streamflow. Information on future changes to rainfall are typically derived from ensembles of global climate models (GCMs), however the spatial resolution of these models are too coarse to provide information at the scale needed for hydrological impact modelling (i.e. catchments or gauges). Downscaling is the process by which finer scale spatial detail is extracted from the larger scale GCM change information (Maraun et al., 2010). Many water resource studies use 'empirical

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2019-139
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 17 April 2019

Hydrology and Earth System Sciences
Open Access
Discussions

scaling', where historical rainfall observations are scaled (perhaps annually or seasonally) for direct use. These methods are relatively simple to use, and results from empirical scaling typically lie in the middle of the range of results from other downscaling methods (Chiew et al., 2010; Potter et al., 2018). However, as empirical scaling methods rely on the historical record of rainfall, future changes in rainfall sequencing (e.g. changes to multi-day accumulations and wet/dry transitions) and consequent effects on runoff cannot be modelled. Dynamical downscaling, in which a regional climate model (RCM) of finer spatial resolution than the host GCM generates rainfall sequences by simulating physical climatic processes, does generate rainfall sequences independent from historical observations. However, challenges remain with using dynamical downscaling output since rainfall (and other climate variables) are not explicitly constrained by observations (see, e.g., Piani et al., 2010; Chen et al., 2011; Teutschbein and Siebert, 2012). As such, dynamical downscaling outputs typically need to be bias corrected for direct use in hydrological models. In particular, a common feature of dynamical downscaling is the tendency to underpredict the occurrence of zero- and low-rainfall days, which is sometimes known as the drizzle effect (e.g. Maraun, 2013). RCM output has been bias corrected for applications in Tasmania (Bennett et al., 2014) and central coast of New South Wales (Lockart et al., 2014) but both these studies found residual biases in multi-day rainfall events, dry spell durations, and autocorrelation of rainfall occurrences. Themeßl et al. (2012) bias corrected RCM output over Europe and found residual biases in rainfall extremes and a modification of the climate change signal.

Water resources in Victoria are shared by urban users, irrigators, industry and the environment. Long-term water strategies and shorter term sustainable water strategies are required for Victoria's water regions by the Victorian Water Act. A key aspect of these water planning processes is accounting for scenarios of climate change as determined by the available science. Cool-season rainfall in Victoria since 2000 has averaged 15% less than the long-term average during the 20th century (Hope et al., 2017). This has been linked to the observed expansion of Hadley Cell circulation (Post et al., 2014). The median scenario of climate change for Victoria typically has reduced rainfall and runoff later in the century, with slightly larger percentage declines in the western parts of the State (Post et al., 2012; Potter et al., 2016). Providing better information to improve water planning processes includes developing finer spatial resolution projections as well as different metrics of daily rainfall amounts and occurrences. The dynamically downscaled NARCliM climate projections (NSW/ACT Regional Climate Modelling, Evans et al., 2014), described in Sect. 2.1 allows the opportunity to provide these improvements, but require bias correction in order to produce sufficiently correct daily rainfall distributions.

The underlying assumption of bias correction is that the RCM output faithfully represents the climate processes responsible for rainfall, although the amounts themselves may not be accurate. Water resource projection modelling is concerned with future changes, and so an argument could be made that although the rainfall amounts are biased for hindcast (historical) simulations, they will presumably be equally biased for future simulations, so that changes can be inferred from comparing biased historical and future rainfall and runoff. However, the sensitivity of runoff to rainfall means that biased rainfall can have large effects on the change signal of runoff (Teng et al., 2015). Furthermore, hydrological models are calibrated to historical rainfall and runoff sequences, and since the distribution of runoff is usually highly skewed, using biased rainfall sequences can distort the distribution of runoff, thus creating large biases in high and low runoff amounts. This makes inferences on the changes to runoff characteristics highly uncertain when biased rainfall inputs are used.

Bias correction identifies a relationship or mapping between observed historical rainfall and hindcast RCM rainfall. This mapping when applied to hindcast RCM rainfall results in a distribution of rainfall identical (or very similar, depending on the methods) to the historical observations. This mapping can then be applied to future RCM rainfall, resulting in unbiased future rainfall sequences. Of course, applying the relationship into the future assumes the bias in RCM rainfall does not change into the future or for different (wetter or drier) climate periods. Bias-correction methods (see Schmidli et al., 2006; Boe et al., 2007; Lenderink et al., 2007; Christensen et al., 2008; Piani et al., 2010; Themeßl et al., 2011; Teng et al., 2015;) fall into three main categories:

- Scaling or change-factor methods;
- Empirical quantile-quantile mapping (QQM); and
- Distributional QQM.

Scaling methods simply consider the change in mean, and apply a constant factor to correct bias in RCM rainfall. Quantile-quantile mapping matches each quantile (or a selection of quantiles) of the two distributions. This can be done using the empirical cumulative density or fitting a distribution to both observed and hindcast RCM daily rainfall amounts.

Teng et al. (2015) demonstrated that representing daily rainfall distributions with double-gamma distributions was largely identical to empirical QQM, implying that distributional and empirical approaches give similar results so long as the distribution is sufficiently flexible. Arguably, the choice between empirical or distributional mapping is a representation of the bias-variance tradeoff problem. Empirical mapping will reduce bias to zero, but at the cost of increasing the variance of predictions, since the mapping will be very sensitive to individual amounts. Distributional mapping fits the data across the entire rainfall distribution, but can result in the hindcast RCM rainfall not being mapped exactly to the historical distribution. For this study we apply empirical quantile-quantile mapping for each season across integral percentiles as described below. Overall there is a small but relatively unimportant difference between different methods for QQM.

Whereas quantile-quantile mapping can effectively reduce historical error in daily rainfall amounts to zero, albeit with some of the caveats already mentioned, the bias corrected rainfall timeseries could still harbour biases and unrealistic characteristics that will result in runoff biases after being routed through a rainfall-runoff model. Specifically, QQM bias correction cannot remove biases in rainfall sequencing and multi-day accumulations that might not be readily apparent when considering only the daily distribution of rainfall amounts (Addor and Seibert, 2014). Unfortunately it is not easy to tell exactly which characteristics of rainfall drive runoff generation, and in general the sensitivity will depend on catchment physical characteristics, storm type and intensity, as well as antecedent moisture and groundwater stores (Goodrich and Woolhiser, 1991; Bell and Moore, 2000; Beven, 2001). Spectral and multifractal approaches (e.g. Milly and Wetherald, 2002; Matsoukas et al., 2000; Tessier et al., 1996) show that rainfall variability at shorter timescales is by and large incorporated into soil moisture buffers thus dampening runoff variability at these timescales. However, over timescales of several days and greater, variability in runoff matches variability in rainfall more and more closely. As such, it is evident that large, intense rainfall

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2019-139
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 17 April 2019

events (measured perhaps by the upper tail of the rainfall distribution), more seasonal rainfall regimes (Wolock and McCabe, 1999), relatively larger variability of rainfall (Potter and Chiew, 2011), and large multi-day accumulations of rainfall are most important for runoff generation (Addor and Seibert, 2014), particularly for high flow events (Jaun et al., 2008), and we focus on these kinds of rainfall metrics in this study.

The main aim of the paper is to investigate the effect of bias correction on rainfall characteristics relevant to runoff generation. Specifically, we investigate whether key rainfall metrics contain biases in dynamically downscaled GCM hindcasts relative to observations. We examine if bias correction acts to either enhance or moderate any such biases, and whether bias correction affects change signals (i.e. GCM future relative to GCM historical). Section 2 describes the data and methods used in this study, Sect. 3 presents results, and Sect. 4 and 5 contain discussion and conclusions.

2 Data and Methods

2.1 Daily Data

Observed rainfall data is obtained from the Australian Water Availability Project (AWAP; Jones et al., 2009). This is a 0.05°×0.05° gridded dataset interpolating observations from point rainfall records from the Australian Bureau of Meteorology. We use rainfall projections output from the Weather Research and Forecasting (WRF) model, which is a mesoscale atmospheric model with many applications both in numerical weather prediction and climate projections (Skamarock and Klemp, 2008). WRF reads in output from a GCM along its lateral and lower boundaries and simulates the climate on a finer resolution within those boundaries (Evans et al., 2014). The NARCliM projections use three different configurations of WRF in combination with four CMIP3 (Meehl et al., 2007) GCMs (MIROC, ECHAM, CCCMA and CSIRO Mk3.0) that were selected to represent GCM uncertainty based on their skill and independence (Evans et al., 2013). (Note that a newer runs of NARCliM, currently being processed, use the CMIP5 model ensemble, and preparations are in place to take advantage of the upcoming CMIP6 ensemble.) The historical (baseline) period for WRF is 1990–2009, and we relate WRF historical rainfall features to observations and NCEP/NCAR reanalysis over the same period. Realisations of future rainfall follow the A2 Scenario of the Special Report on Emissions Scenarios (SRES) (Nakićenović et al., 2000) at 2060–2079. Change signals presented later are thus averages over 2060–2079 compared to 1990–2009. The AWAP rainfall observations are projected over a regular latitude-longitude grid, and to be commensurate with this, the WRF outputs are bilinearly interpolated to a 0.1°×0.1° grid aligned with the AWAP grid. The AWAP rainfall data is then regridded by using weighted averages of AWAP gridcells overlapping the 0.1°×0.1° AWAP grid.

2.2 Quantile-quantile bias correction

Quantile-quantile bias correction works by estimating the cumulative density function for observed and modelled historical daily rainfall amounts: F_o and F_m. These are then combined to produce a mapping function:

$$P_o = F_o^{-1} \circ F_m\,(P_m)$$

The map $F_o^{-1} \circ F_m$ thus returns the observed daily distribution (or approximately so, depending on the method) when applied to the modelled historical timeseries, and a bias corrected future timeseries when applied to the modelled future timeseries. We use the R package 'qmap' (Gudmundsson et al., 2012), to estimate the cumulative density functions for the mapping function. This estimates quantiles for both observed and modelled non-zero rainfall at integral percentiles including 0 (minimum) and 100 (maximum). The quantile for a particular daily rainfall amount is then estimated using linear interpolation between percentiles, and linear extrapolation in case of future modelled rainfall lying outside the historical distribution. Compared to using the empirical distributions directly, differences between the bias-corrected modelled rainfall distribution and the distribution of observations can occur (of the order of 2–3%) because the interpolation between large rainfall percentiles (particularly 99 to 100) will not match observed percentiles exactly. As noted by Teng et al. (2012), sufficiently flexible approaches to bias correction give very similar results. QQM bias correction in this way was applied separately to each three-month season (i.e. DJF, MAM, JJA and SON) in each grid cell independently.

2.3 Defining Transition Probabilites

Whereas QQM bias correction can correct the daily distribution exactly, daily bias correction is not set up to correct sequences and accumulations Addor and Seibert (2014). To this end, we consider not only how bias correction affects daily metrics of rainfall, but also the sequencing of wet days that produce runoff. One way of measuring this is through transition probabilities of wet and dry sequences. To this end, we consider a simple two-state Markov Chain rainfall occurrence model. Here, the probability of a wet or dry day depends on whether the previous day was wet or dry. A "dry" day can be defined as either zero rainfall or rainfall below a given threshold (such as 1 mm). Define the wet-to-wet and dry-to-dry transition probabilities as $w = \Pr(W|W)$ and $d = \Pr(D|D)$. These determine the Markov Chain since $\Pr(D|W) = 1 - w$ and $\Pr(W|D) = 1 - d$. The probability of dry day occurrences, p is fully determined by w and d parameters, as given by Cox and Miller (1965):

$$p = \frac{1 - w}{2 - d - w}$$

Equivalently,

$$w = 1 - p\frac{1 - d}{1 - p}$$

If a series of occurrences of dry and wet days has zero autocorrelation (i.e. the state probability is independent of the rainfall state in the previous day), then it follows that $\Pr(D|D) = \Pr(D|W) = p$. As such, the diagonal line where $p = d$ (dashed line in Figure) corresponds to an independently (i.e. zero autocorrelation) series of occurrences. The area above and to the right of the diagonal line corresponds to a series of occurrences with positive autocorrelation (i.e. $\Pr(D|D) > p$ and so dry sequences are more likely to persist), whereas the area below and to the left of the line corresponds to series with negative autocorrelation. Figure is used in Sect. 3.3 as a novel way to represent the relationships between state-transition probabilities and rainfall quantiles to investigate the effect of bias correction on transition probabilities.

3 Results

3.1 Change Signal versus Bias

Figure 3 shows mean annual rainfall across Victoria from observations, downscaled (NCEP/NCAR) reanalysis (2nd column) and GCM-downscaled (3rd column, historical; 4th column, future). ECHAM5-R1 was chosen as a representative GCM ensemble member as it had the median historical regional rainfall across Victoria from the WRF model ensemble. The spatial rainfall fields from both the reanalysis and GCM both show reasonable agreement with the spatial pattern of the observed mean annual rainfall, with relatively larger rainfall across the mountain ranges to the east of the State and across the southern coast. There is less rainfall to the more arid north-west region as well. However, both rainfall fields are evidently positively biased with around 100–200 mm of excess rainfall consistently across most of the State, relatively more for the far Eastern part of Victoria for the reanalysis data and across the mountain ranges for the GCM data, interspersed with small patches of negative bias (observations larger than downscaled data). Averaged across Victoria, the mean absolute bias is approximately 26% for the reanalysis and 43% for the GCM data. In comparison, the absolute change signal averages 3%, rising to around 10% in the Eastern part of Victoria.

3.2 Bias Correction of WRF

Consistent with Fig. 3, the raw WRF rainfall is mostly wetter (positive bias) across Victoria (Fig. 4), except for a tendency towards underprediction in the south-east coast for some models. The quantile-quantile bias correction method is formulated to correct quantiles of rainfall exactly so that bias-corrected mean annual rainfall, as well as any quantiles, are approximately equal. However, the method used here does not correct high rainfall quantiles (e.g. P99 and above) exactly, due to the interpolation between quantiles as described in Sect. 2.2 (Fig. 5). This residual bias appears to be randomly distributed spatially and results in bias-corrected mean annual rainfall not being exactly corrected. However, this effect is generally less than 5% of the mean annual rainfall. This effect can be removed entirely by using empirical density functions rather than interpolated values, but with the effect of increasing prediction uncertainty. This residual bias in QQM bias corrected rainfall is more pronounced at P99 (Fig. 5) where interpolation between P99 and P100 (maximum) can have a larger effect.

Figures 6 to 8 show the distribution of bias for different rainfall metrics before bias correction and the residual bias after QQM bias correction. The bias in raw WRF ranges from 5% to 50% for all percentiles (Fig. 6), increasing as the percentile increases (i.e. more relative bias for higher rainfall amounts). After bias correction, as with Fig. 4, bias at all percentiles is effectively reduced to zero after bias correction, although the residual bias is relatively larger at larger percentiles (higher rainfall), similarly to Fig. 5. WRF rainfall is overestimated at all seasons and months before bias correction (Fig. 7), although winter rainfall is relatively less biased than summer rainfall. Bias correction reduces bias to zero annually and seasonally, since QQM is applied to each season separately. Since the intra-seasonal relative monthly rainfall amounts are not exactly equal to the observed amounts, seasonal bias correction occasionally overcorrects bias, particularly in February, April, May and June, with the bias corrected rainfall in these months being less than observed whereas they were overpredicted before bias correction.

Overall though, the absolute relative bias is reduced and closer to zero in all months compared to the raw RCM monthly amounts.

Figure 8 shows relative bias in rainfall sequencing related metrics. Autocorrelation of rainfall amounts is underpredicted before bias correction, and the magnitude of this bias actually increases after bias correction. Whereas QQM reduces bias in dry-dry transition probabilities (i.e. the probability of dry sequences persisting), bias in wet-wet transition probabilities increases after bias correction so that, similarly to autocorrelation, the probability of wet spells persisting is significantly underpredicted after bias correction. We examine this in more detail in Sect. 3.3.1 using the transition probability framework developed in Sect. 2.3. This results in the bias in mean and maximum dry spells being well corrected, whereas maximum 3-day rainfall accumulation and wet-spell occurrences all have negative bias after QQM bias correction. Different percentiles of 3-day rainfall accumulation (calculated as percentiles of a 3-day moving sum of rainfall timeseries) have different residual bias (Fig. 9). 3-day accumulation percentiles below the 80th are all slightly overestimated after bias correction, but above the 80th percentile, a large residual underestimation is present. This reduces at around the 99th percentile but is moderated somewhat for the 3-day maximum (i.e. 100th percentile). As noted by Olson et al. (2016), WRF models were selected according to their skill in reproducing selected 2-week periods of heavy rainfall. This provides a potential explanation for the smaller bias in 3-day maxima relative to 3-day 99% rainfall.

It is likely that underpredicting wet spell occurrences and persistence (Fig. 8 and Fig. 9) will result in runoff from the bias corrected rainfall being underpredicted too. To explore this, runoff was modelled from bias corrected rainfall and observed PET using GR4J calibrated on observed rainfall and PET data using 90 catchments in and around Victoria. Ungauged areas use parameters donated from the nearest neighbour calibration catchment. Further details are provided by Charles et al. (2019, submitted). Figure 10 plots the percentage difference in ensemble-median mean annual runoff for each $0.1° \times 0.1°$ cell compared to mean annual runoff modelled using AWAP observed rainfall. The ensemble median of runoff across Victoria is underpredicted by between 10–20% across almost all of Victoria suggesting that the residual bias in wet spell occurrences and persistence is problematic for runoff modelling. Whereas the smallest percentage biases appear to be over the high-runoff producing region, this region has the highest absolute biases with bias in runoff of more than −20 mm. Characteristics and biases of runoff from bias corrected WRF rainfall is explored in more detail by Charles et al. (2019, submitted).

3.3 Residual Bias in Rainfall State Transition Probabilities

A number of studies have highlighted the residual bias in rainfall sequencing after QQM bias correction (e.g. Addor and Seibert, 2014). The results above demonstrate that dry-dry transition probabilities have low residual bias (possibly due to the emphasis in QQM on preserving zero-rain occurrences), but that wet-wet transition probabilities have more bias after QQM bias correction. This results in the persistence of wet spells being underestimated even though the volumetric amount of rainfall is, by design of QQM bias correction, equal to observed rainfall at any gridpoint.

After bias correction, dry-dry transition probabilities for a 1 mm threshold are reduced, but still have a small negative bias (Fig. 8). Figure 11 shows the dry-dry transition probabilities calculated from observed rainfall (top left), bias corrected reanalysis (middle column), bias corrected historical GCM (right column). Both the bias-corrected reanalysis and bias-corrected GCM results show a spatial pattern very similar to observations with higher dry-dry transition probabilities to the northwest of the State, and at similar places along the southern coastline. However, the reanalysis and more so the GCM result has a lower dry-dry transition probability across almost all of the region. As such, dry spells from the bias-corrected model output are likely to be shorter in duration and less common than that from the observed rainfall (although bias correction does reduce the bias in dry spells somewhat compared to the bias in the raw data as seen in Fig. 8).

Whereas the dry-dry transition probabilities were largest in the north-west, drier, part of Victoria, the wet-wet transition probabilities are largest over the high-runoff producing region, which corresponds to the high-relief, high-altitude part of the State. As with the dry-dry transition probabilities, both bias-corrected reanalysis and bias-corrected GCMs reproduce the spatial pattern of wet-wet transition probabilities, but there is considerable residual bias in these probabilities across the entire region. The residual bias in GCM transition probabilities is over 10% over almost all of Victoria. This results in underestimation of wet spell occurrences and durations and multiday accumulations of rainfall (Fig. 8). The bias in wet-wet transition probabilities is more problematic for modelling runoff than the bias in dry-dry transition probabilities, not only because it is of larger magnitude, but because:

- runoff is sensitive to multiday wet spells
- the larger wet-wet probabilities occur in high runoff producing areas, which we would like to model correctly for regional water availability projections
- QQM bias correction reduces the bias in dry-dry transition probabilities, but increases the magnitude of the bias in wet-wet transition probabilities (Fig. 8).

Figure 13 shows the observed (green), raw (blue) and bias-corrected (red) historical GCM rainfall amounts for a sample grid cell overlaid on the transition-probability space developed in Sect. 2.3 (i.e. Fig. 2). Other grid cells and GCMs show very similar responses (as can be seen in the low spread of results for *d* and *w* in Fig. 8. Quantile-quantile mapping bias correction equates the quantiles q for each rainfall amount such that equal rainfall amounts for observations and bias-corrected rainfall occur on the same probability contours (orange lines in Fig. 13), with raw values translated along the rainfall amount curve. That is, values on the blue line in Fig. 13 map to corresponding values on the red line; the slight variation between the lines is due to different bias corrections in each season. As such, wet-wet and dry-dry transition probabilities for a given rainfall threshold (e.g. 1 mm) for bias-corrected rainfall are equal to the transition probabilities for the corresponding amount in the raw data. For example, in Fig. 13, the exceedance probability for 1 mm in the observed data (black line) is 0.774. The corresponding quantile in the raw data is 2.675 mm. This amount is mapped to 1 mm in the bias-corrected data, and the corresponding wet-wet and dry-dry transition probabilities for 2.675 mm are identical to the transition probabilities for 1 mm in the bias-corrected data. Recall from Sect. 2.3 that the slope −1 line in Fig. 13 where p=d corresponds to an independent sequence of events and northeast of this line in the transition-probability space represents more (positive) serial correlation. This implies that the observed rainfall timeseries contains more correlation structure in the sequence of wet and dry spell occurrences than the modelled rainfall sequence, and that QQM bias correction cannot rectify this since daily QQM retains the

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2019-139
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 17 April 2019

autocorrelation structure of the raw time series since daily amounts are simply rescaled. We surmise that a bias correction method that adjusts occurrences is needed to properly correct biases for hydrological modelling.

3.4 Change signals

Here we examine change signals in rainfall metrics (i.e. percentage difference in RCM future relative to RCM historical averages) specifically looking at whether bias correction alters the change signals. Figure 14 shows the change signals in different rainfall percentiles. For the raw data, there is a small decrease in low to moderate rainfall amounts less than the non-zero 40th percentile and a future increase in non-zero percentiles above 50%. The nature of QQM bias correction means that raw and bias corrected equal percentiles cannot be compared directly. Nevertheless a similar pattern is found with the bias corrected data, namely that larger rainfall amounts have larger relative changes than smaller rainfall amounts.

Figure 15 shows change signals in mean annual, seasonal and monthly average rainfall. The magnitude of the median change signal in mean annual rainfall is around –5%, and seasonal changes are comparable to the annual change except for SON rain which has a decrease projected by the WRF ensemble of around 20%. Compared to the raw bias in mean and seasonal rainfall (Fig. 7) of between 25%–50%, these change signals are between one-half and one-tenth of the bias. After bias correction, there is little difference in the magnitude and direction of change in seasonal and monthly averages. However, the mean annual rainfall change is moderated somewhat, and this is somewhat problematic since mean annual changes are most often considered in regional projection applications. Charles et al. (2019, submitted) discuss this effect in more detail.

Although the residual bias in rainfall sequencing metrics is not eliminated, and in some cases (e.g. wet-wet transition probabilities) is actually increased after bias correction, Fig. 16 shows that the change signals in rainfall sequencing metrics is largely unaffected by bias correction. With the exception of maximum 3-day rainfall accumulation, the distributions of change signals are largely identical before and after bias correction. The difference in change signal for maximum 3-day rainfall accumulation is related to the relatively larger increase in large rainfall events (e.g. P99 in Fig. 14).

4 Discussion

We demonstrate in Sect. 3.4 that change signals (future mean relative to historical) in rainfall metrics can be considerably smaller than the bias (modelled historical relative to observed historical). On the one hand, this seems problematic since biases in processes can be considered so large that the changes are insignificant. On the other hand, there is no particular legitimacy for this viewpoint, certainly not from a statistical sense. The magnitude of bias does not provide any sort of confidence level in changes to rainfall metrics. However, given such relatively large biases, it is reasonable to assume that there are some errors in the way particular climate processes are modelled, either through the host GCM or the RCM. It would be desirable to understand the reasons and climatic process responsible for biases and assess whether these processes are unrealistic, as well as whether these biases render the changes physically implausible. Such an assessment is beyond the scope of this paper however.

In general, bias correction does not tend to alter the change signals in rainfall metrics (with the exception of 3-day accumulation and low rainfall percentiles). Nevertheless, small differences in rainfall metrics can result in large differences in runoff metrics and other water availability measures (e.g. low flows and high flows). High runoff and even average runoff amounts can be very sensitive to 3-day rainfall accumulation, which we saw can be altered through daily bias correction. Charles et al. (in prep) discuss the effect on runoff of residual biases in rainfall metrics after daily bias correction. Bias correction can and does affect results of regional water availability assessments, and it is recommended that bias correction is included in any uncertainty analysis undertaken.

Another important consideration is the relevant metrics to be considered by end users. Bias correction by season, for example, can alter change signals annually, and care must be taken as to which metrics are of interest, and which are the most appropriate bias correction methods to apply in order to properly account for the metrics of interest. Certainly, caution must be applied when considering rainfall and runoff metrics that were not considered when applying bias correction to projections. Low flow metrics are particularly problematic (Potter et al., 2018), where different downscaling and bias correction methods can give very different answers.

Although daily bias correction methods as outlined in this paper tend to result in residual bias in multi-day metrics, generally change signals in transition probabilities are very similar before and after bias correction. This information could thus potentially be extracted from RCMs to drive local weather generation or stochastic methods to provide future rainfall projections that can be suitable for local hydrological projections. Maintaining interannual and muti-decadal correlations, as well as spatial correlations between rainfall gauges, remains a challenge for stochastic methods, however.

5 Conclusions

Projections of future changes to rainfall and runoff from dynamically downscaled climate models often necessitates a form of bias correction to rainfall fields to obtain sufficiently realistic rainfall inputs for hydrological models. Dynamical downscaling offers potential benefits to regional hydroclimate projections, such as the ability to better model daily rainfall metrics, low flow metrics (after modelling runoff with a hydrological model), and finer spatial scale information, but comes with challenges related to bias. Whereas bias in rainfall amounts can be corrected using quantile-quantile mapping (QQM) methods, biases in rainfall occurrences (such as rainfall autocorrelation, dry-dry and wet-wet transition probabilities) are not properly corrected with QQM.

The relative magnitude of change signals (future RCM to historical RCM) of the different rainfall metrics examined here is typically less than the magnitude of the bias. Mean annual rainfall change is an order of magnitude smaller than the bias in mean annual rainfall but seasonal changes are more like half of the bias in seasonal averages. Although this might call into question the validity of the change signal, one approach is to assume that the magnitudes of the changes are responsive to changing greenhouse gas emissions, insofar as the changing atmospheric processes are realistically modelled by the dynamical downscaling process. Indeed this is the basic premise behind empirical scaling, i.e. that the change is the authentic signature of the climate modelling especially since the RCMs are not explicitly tuned to observed rainfall.

Individual percentiles and seasonal totals are, by design, effectively reduced to zero using QQM. Some interpolation and extrapolation occurs in the approach used here, so there is some random residual bias in higher percentiles (i.e. high rainfall amounts). This can be eliminated altogether by using the exact empirical density functions, but at the cost of increased predictive uncertainty. Using empirical densities also raises problems with extrapolation past historical amounts. Monthly totals retain some residual bias because of compensating biases within each season due to small errors in rainfall seasonality by the RCMs. Metrics associated with rainfall sequencing (e.g. serial correlation, wet-wet and dry-dry state transition probabilities and quantiles of 3-day accumulation) all have significant residual bias, particularly so for wet-wet state transition probabilities in which the magnitude of bias in raw RCM historical runs is amplified after bias correction. This leads to a considerable underestimation of mean annual runoff after rainfall is routed through a hydrological model because runoff is very sensitive to multiday accumulations of rainfall and sequencing of wet spells in particular.

An analysis of the lag-one transition probabilities (i.e. wet state to wet state and dry state to dry state) showed that WRF rainfall had transitions to different states that are more random (i.e. more independent) compared to observed rainfall. QQM bias correction is unable to correct these transition probabilities as QQM retains the transition probabilities for any particular quantile. Since persistence of wet spells is critical for runoff generation, a different approach to bias correction is needed to successfully use WRF for runoff projections that can correct rainfall sequencing to better represent the observed correlation structure in wet and dry occurrences.

Change signals in annual, seasonal and monthly average rainfall as well as rainfall sequencing metrics are largely preserved after bias correction, with the exception of maximum 3-day rainfall accumulation. However there is a slight tendency for DJF and MAM rainfall change signals to increase after bias correction and this leads to a tangible reduction in the magnitude of the projected decrease in mean annual rainfall. This is problematic for applications since mean annual change is the most commonly used metric for hydroclimate projections. One possible solution is to rescale the bias corrected rainfall according to raw changes signals but this depends on whether we believe the raw or the bias corrected change signal is correct. However, the fact that rainfall sequencing metrics (such as state transition probabilities and daily rainfall autocorrelation) are largely unchanged by bias correction suggests the possibility of using this information to drive either weather-generation models or stochastic/resampling-based bias correction methods to produce hydrologically realistic rainfall sequences for hydroclimate projection applications.

Acknowledgments

This study was funded by the Victorian Department of Environment, Land, Water and Planning (DELWP) through the Victorian Water and Climate Initiative.

References

Addor, N., and Seibert, J.: Bias correction for hydrological impact studies – beyond the daily perspective, Hydrol. Process., 28, 4823–4828, https://doi.org/10.1002/hyp.10238, 2014.

Bell, V. A., and Moore, R. J.: The sensitivity of catchment runoff models to rainfall data at different spatial scales, Hydrol. Earth Syst. Sc., 4, 653–667, https://doi.org/10.5194/hess-4-653-2000, 2000.

Bennett, J. C., Grose, M. R., Corney, S. P., White, C. J., Holz, G. K., Katzfey, J. J., Post, D. A., and Bindoff, N. L.: Performance of an empirical bias-correction of a high-resolution climate dataset, Int. J. Climatol., 34, 2189–2204, https://doi.org/10.1002/joc.3830, 2014.

Beven, J. K.: Rainfall-runoff modelling: The primer, Wiley, Chichester, England, 2001.

Boe, J., Terray, L., Habets, F., and Martin, E.: Statistical and dynamical downscaling of the Seine basin climate for hydro-meteorological studies, Int. J. Climatol., 27, 1643–1655, https://doi.org/10.1002/joc.1602, 2007.

Charles, S. P., Chiew, F. H. S., Potter, N. J., Zheng, H., Fu, G., and Zhang, L.: Impact of dynamically downscaled rainfall biases on projected runoff changes, submitted to Hydrol. Earth Syst. Sc., 2019 (submitted).

Chen, J., Brissette, F. P., and Leconte, R.: Uncertainty of downscaling method in quantifying the impact of climate change on hydrology, J. Hydrol., 401, 190–202, https://doi.org/10.1016/j.jhydrol.2011.02.020, 2011.

Chiew, F. H. S., Kirono, D. G. C., Kent, D. M., Frost, A. J., Charles, S. P., Timbal, B., Nguyen, K. C., and Fu, G.: Comparison of runoff modelled using rainfall from different downscaling methods for historical and future climates, J. Hydrol. 387, 10–23, https://doi.org/10.1016/j.jhydrol.2010.03.025, 2010.

Christensen, J. H., Boberg, F., Christensen, O. B., and Lucas-Picher, P.: On the need for bias correction of regional climate change projections of temperature and precipitation, Geophys. Res. Lett., 35, L20709, https://doi.org/10.1029/2008GL035694, 2008.

Cox, D. R., and Miller, H. D.: The theory of stochastic processes, Methuen, London, 1965.

Evans, J. P., Ji, F., Abramowitz, G., and Ekström, M.: Optimally choosing small ensemble members to produce robust climate simulations, Environ. Res. Lett., 8, https://doi.org/10.1088/1748-9326/8/4/044050, 2013.

Evans, J. P., Ji, F., Lee, C., Smith, P., Argüeso, D., Fita, L.: Design of a regional climate modelling projection ensemble experiment – NARCliM, Geosci. Model Dev., 7, 621–629. https://doi.org/10.5194/gmd-7-621-2014, 2014.

Goodrich, D. C., and Woolisher, D. A.: Catchment hydrology, Rev. Geophys. (Suppl.), 29, 202–209, 1991.

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

Hope, P., Timbal, B., Hendon, H., Ekström, M., and Potter, N.: A synthesis of findings from the Victorian Climate Initiative (VicCI), Bureau of Meteorology, Australia, 56pp., 2017.

Jaun, S., Ahrens, B., Walser, A., Ewen, T., and Schär, C.: A probabilistic view on the August 2005 floods in the upper Rhine catchment, Nat. Hazard. Earth Sys., 8, 281–291, https://doi.org/10.5194/nhess-8-281-2008, 2008.

Jones, D. A., Wang, W., Fawcett, R.: High-quality spatial climate data-sets for Australia. Aust. Meteorol. Ocean., 58, 233–248, 2009.

Lenderink, G., Buishand, A., and van Deursen, W.: Estimates of future discharges of the river Rhine using two scenario methodologies: direct versus delta approach, Hydrol. Earth Syst. Sc., 11, 1145–1159, https://doi.org/10.5194/hess-11-1145-2007, 2007.

Lockart, N., Willgoose, G., Kuczera, G., Kiem, A. S., Chowdhury, A., Manage, N. P., Zhang, L. Y., and Twomey, C.: Case study on the use of dynamically downscaled climate model data for assessing water security in the Lower Hunter region of the eastern seaboard of Australia, Journal of Southern Hemisphere Earth Systems Science, 66, 177–202, 2016.

Maraun, D.: Bias Correction, Quantile Mapping, and Downscaling: Revisiting the Inflation Issue, J. Climate, 26, 2137–2143, https://doi.org/10.1175/JCLI-D-12-00821.1, 2013.

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.: Precipitation downscaling under climate change: Recent developments to bridge the gap between dynamical models and the end user, Rev. Geophys., 48, RG3003, https://doi.org/10.1029/2009RG000314, 2010.

Matsoukas, C., Islam, S., and Rodriguez-Iturbe, I.: Detrended fluctuation analysis of rainfall and streamflow time series, J. Geophys. Res.-Atmos., 105, D23, 29165–29172, https://doi.org/10.1029/2000JD900419, 2000.

Meehl, G. A., Covey, C., Delworth, T., Latif, M., Mcavaney, B., Mitchell, J. F. B., Stouffer, R. J., and Taylor, K.E.: The WCRP CMIP3 multimodel dataset - A new era in climate change research, B. Am. Meteorol. Soc., 88, 1383–1394, https://doi.org/10.1175/BAMS-88-9-1383, 2007.

Milly, P. C. D., and Wetherald, R. D.: Macroscale water fluxes 3. Effects of land processes on variability of monthly river discharge, Water Resour. Res., 38, 1235, https://doi.org 10.1029/2001WR000761, 2002.

Olson, R., Evans, J. P., Di Luca, A., and Argüeso, D.: The NARCliM project: model agreement and significance of climate projections, Clim. Res., 69, 209–227, https://doi.org/10.3354/cr01403, 2016.

Piani, C., Weedon, G. P., Best, M., Gomes, S. M., Viterbo, P., Hagemann, S., and Haerter, J. O.: Statistical bias correction of global simulated daily precipitation and temperature for the application of hydrological models, J. Hydrol., 395, 199–215, https://doi.org/10.1016/j.jhydrol.2010.10.024, 2010.

Post, D. A., Chiew, F. H. S., Teng, J., Wang, B., and Marvanek, S.: Projected changes in climate and runoff for south-eastern Australia under 1°C and 2°C of global warming. A SEACI Phase 2 special report, Commonwealth Scientific and Industrial Research Organisation, Australia, 2012.

Post, D. A., Timbal, B., Chiew, F. H. S., Hendon, H. H., Nguyen, H., and Moran, R.: Decrease in southeastern Australian water availability linked to ongoing Hadley cell expansion. Earth's Future, 2, 231–238, https://doi.org/10.1002/2013EF000194, 2014.

Potter, N. J., and Chiew, F.H.S.: An investigation into changes in climate characteristics causing the recent very low runoff in the southern Murray-Darling Basin using rainfall-runoff models, Water Resour. Res., 47, W00G10, https://doi.org/10.1029/2010WR010333, 2011.

Potter, N. J., Chiew, F. H. S., Zheng, H., Ekström, M., and Zhang, L.: Hydroclimate projections for Victoria at 2040 and 2065. Commonwealth Scientific and Industrial Research Organisation, Australia, 2016.

Potter, N. J., Ekström, M., Chiew, F. H. S., Zhang, L., and Fu, G. Change-signal impacts in downscaled data and its influence on hydroclimate projections, J. Hydrol., 564, 12–28, https://doi.org/10.1016/j.jhydrol.2018.06.018, 2018.

Schmidli, J., Frei, C., and Vidale, P. L.: Downscaling from GCM precipitation: a benchmark for dynamical and statistical downscaling methods, Int. J. Climatol., 26, 679–689, https://doi.org/10.1002/joc.1287, 2006.

Skamarock, W. C., Klemp, J. B.: A time-split nonhydrostatic atmospheric model for weather research and forecasting applications, J. Comput. Phys., 227, 3465–3485. https://doi.org/10.1016/j.jcp.2007.01.037, 2008.

Teng, J., Potter, N. J., Chiew, F. H. S., Zhang, L., Wang, B., Vaze, J., and Evans, J.P.: How does bias correction of regional climate model precipitation affect modelled runoff? Hydrol. Earth Syst. Sc., 19, 711–728. https://doi.org/10.5194/hess-19-711-2015, 2015.

Tessier, Y., Lovejoy, S., Hubert, P., Schertzer, D., and Pecknold, S.: Multifractal analysis and modeling of rainfall and river flows and scaling, causal transfer functions, J. Geophys. Res.-Atmos., 101, 26427–26440, https://doi.org/10.1029/96JD01799, 1996.

Teutschbein, C., and Seibert, J.: Is bias correction of regional climate model (RCM) simulations possible for non-stationary conditions? Hydrol. Earth Syst. Sc., 17, 5061–5077, https://doi.org/10.5194/hess-17-5061-2013, 2012.

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

Themeßl, M. J., Gobiet, A., and Heinrich, G.: Empirical-statistical downscaling and error correction of regional climate models and its impact on the climate change signal, Climatic Change, 112, 449–468, https://doi.org/10.1007/s10584-011-0224-4, 2012.

Wolock, D. M., and McCabe, G. J.: Explaining spatial variability in mean annual runoff in the conterminous United States, Clim. Res., 11, 149–159, 1999.

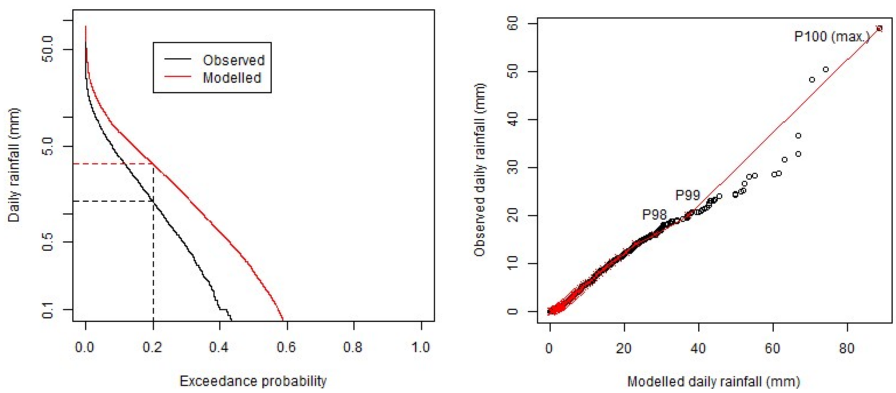


Figure 1: Schematic of QQM bias correction. The left panel shows the empirical cumulative density functions for both observed and modelled rainfall in a given gridcell. Percentiles are estimated from both distributions, which are equated in bias correction to generate a mapping function (right panel). Values lying between percentiles or outside the modelled maximum value are interpolated or extrapolated linearly.

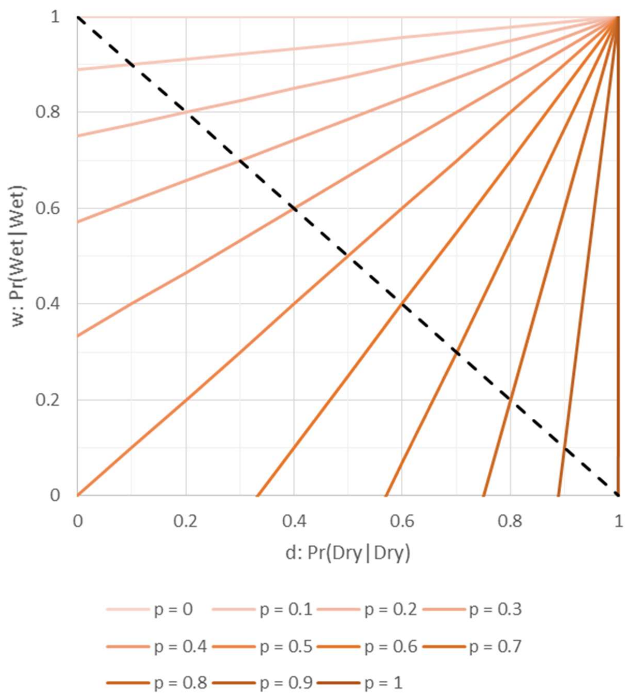


Figure 2: ***p*** **(quantile probability) as a function of wet-wet transition probabilities and dry-dry transition probabilities.**

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2019-139
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 17 April 2019

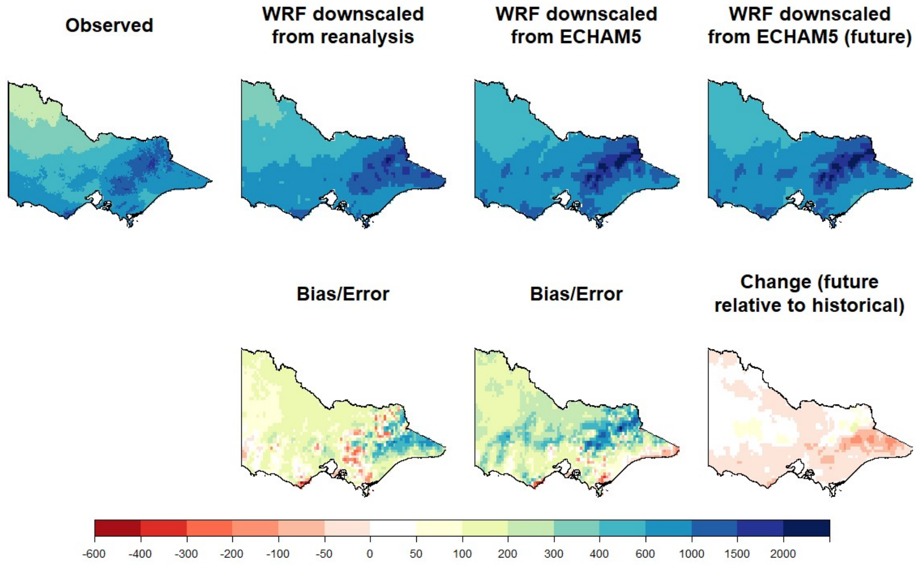


Figure 3: Magnitude of change signal of mean annual rainfall compared to bias.

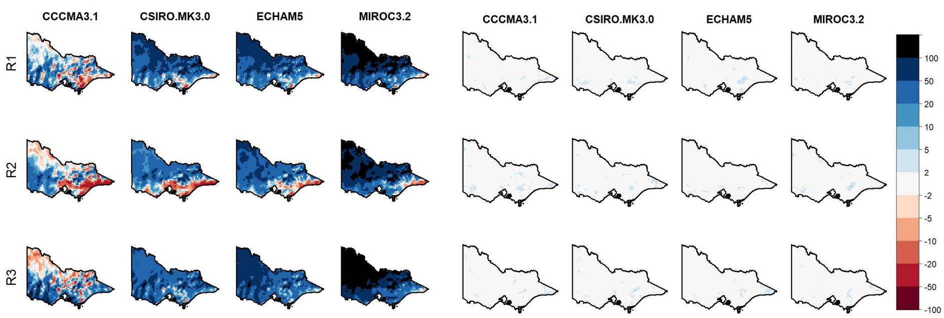


Figure 4: Percentage bias in mean annual rainfall (raw data, left panels; residual bias, right panels).

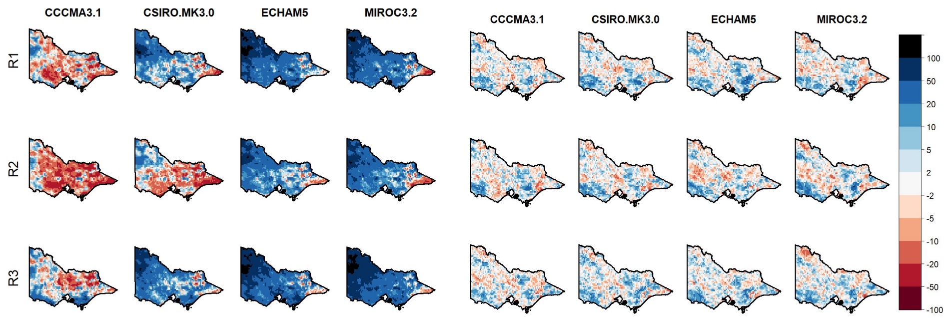


Figure 5: Percentage bias in P99: 99th percentile of rainfall (raw data, left panels; residual bias, right panels).

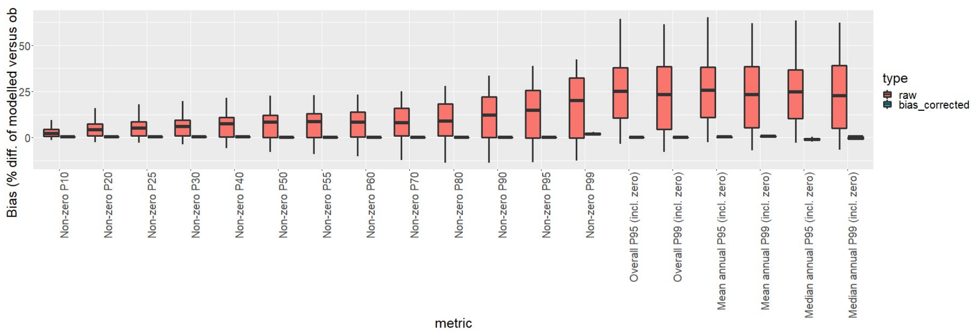


Figure 6: Relative bias (modelled compared to AWAP) of rainfall percentiles before and after bias correction. The range of results represents the spread of GCM hindcast spatial averages from the WRF 12-model ensemble

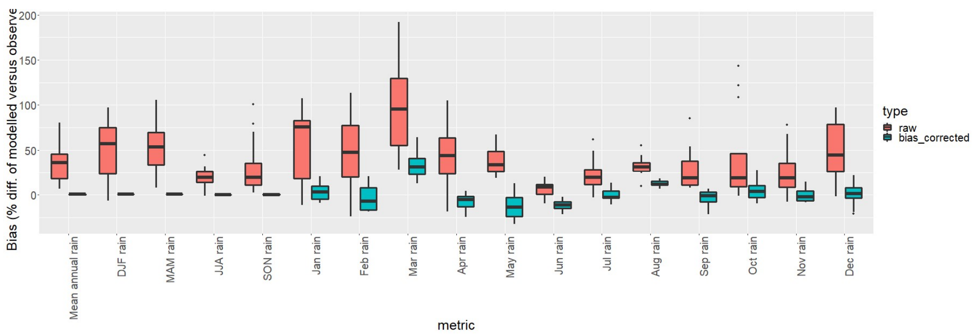


Figure 7: Relative bias in annual, seasonal and monthly averages

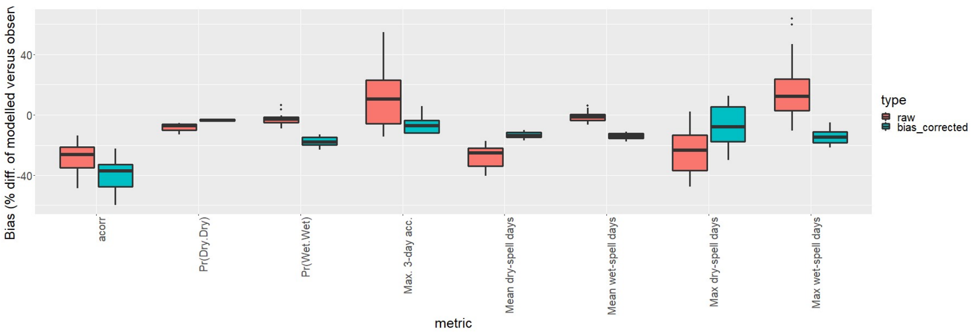


Figure 8: Relative bias in rainfall sequencing related metrics

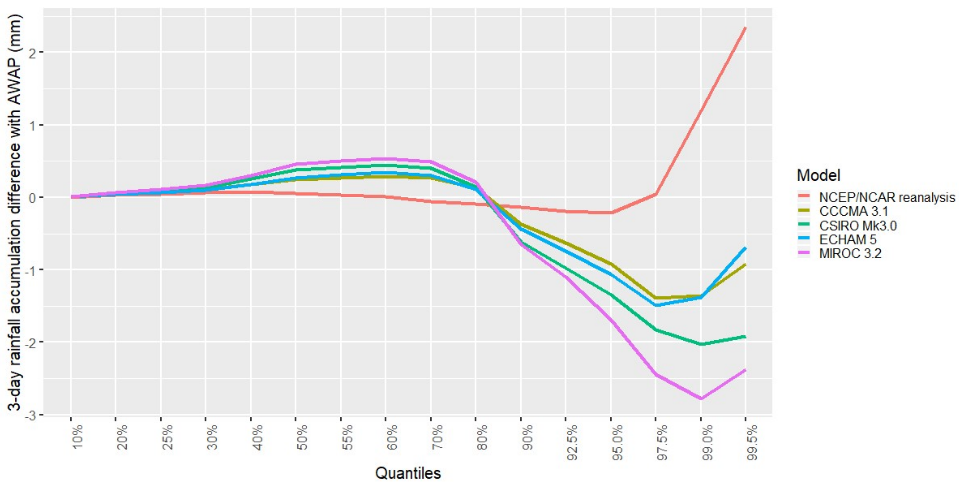


Figure 9: Biases in spatial average 3-day rainfall accumulation percentiles. Here a 3-day moving average filter was applied to each hindcast timeseries and quantiles taken at increasing values.

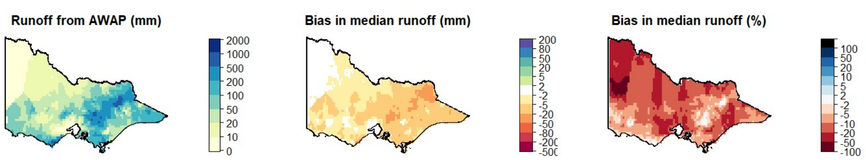


Figure 10: Absolute (mm) and percentage bias in median runoff compared to modelled runoff from AWAP rainfall.

Hydrology and Earth System Sciences
Open Access
Discussions

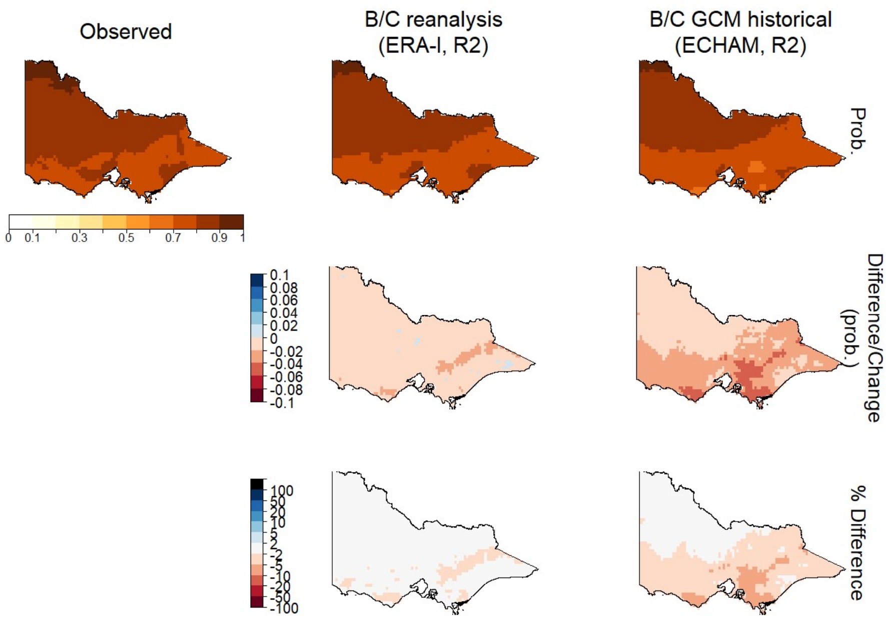


Figure 11: Dry-dry transition probabilities (1mm threshold)

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2019-139
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 17 April 2019

Hydrology and Earth System Sciences
Open Access
Discussions

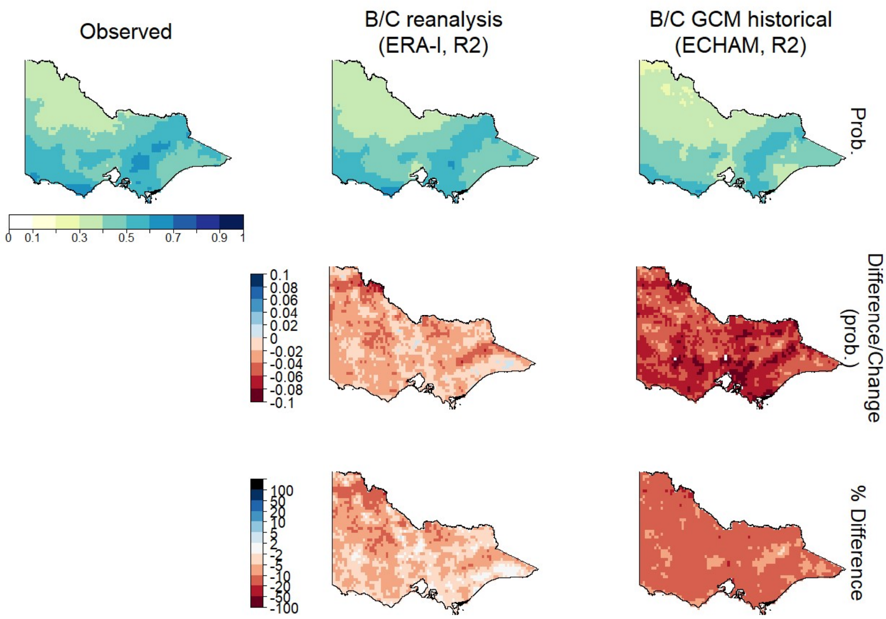


Figure 12: Wet-wet transition probabilities (1mm threshold)

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2019-139
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 17 April 2019

Hydrology and Earth System Sciences
Open Access
Discussions

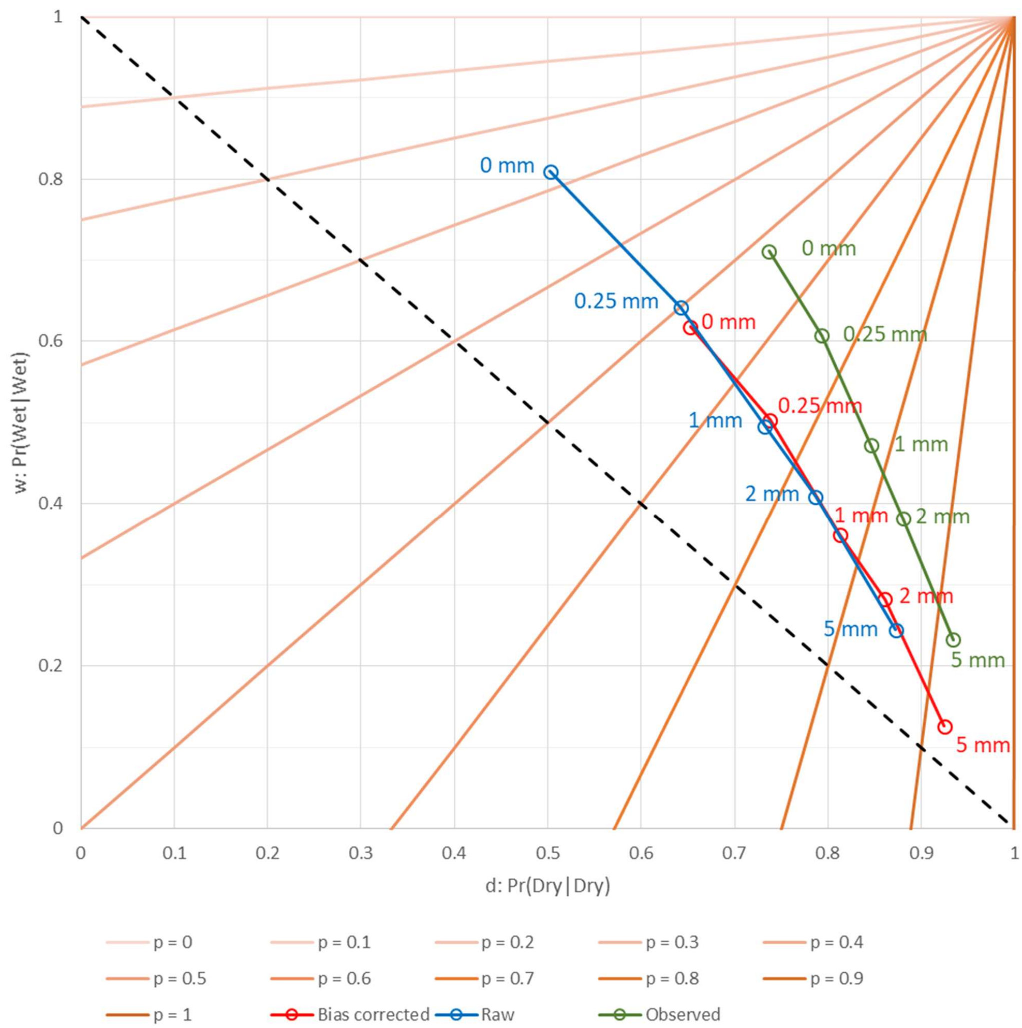


Figure 13: An alternative perspective on quantile-quantile mapping: daily rainfall amounts and associated probabilities plotted in *d*-*w* space (cf. Figure 2). Quantile-quantile mapping works by mapping amounts from the raw data (blue curve) to the probability contours (dashed lines) corresponding to the observed amount.

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2019-139
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 17 April 2019

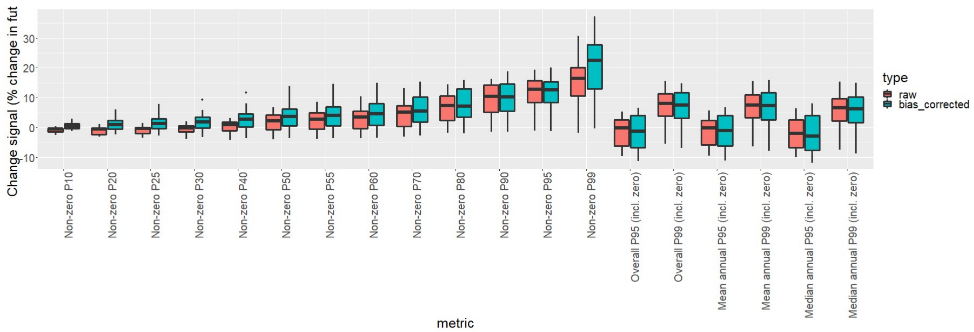
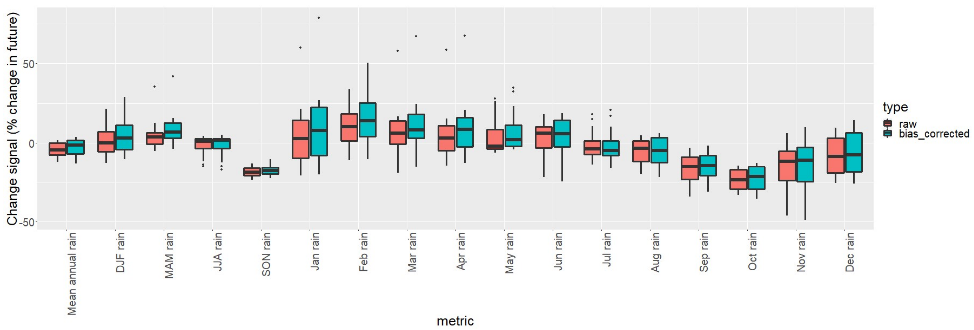


Figure 14: Change signal (percentage difference of RCM future relative to RCM historical) in rainfall percentiles before and after bias correction.



5 **Figure 15: Change signal in mean annual, seasonal and monthly averages before and after bias correction**

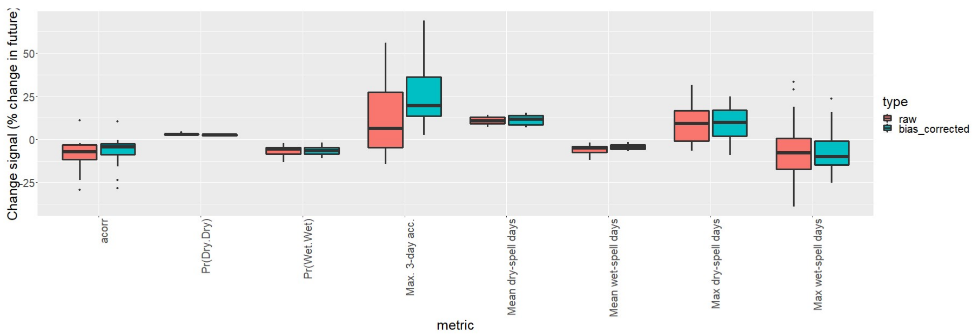


Figure 16: Change signal in rainfall sequencing metrics before and after bias correction