

Research papers

A rank-based approach for correcting systematic biases in spatial disaggregation of coarse-scale climate simulations



Jannatun Nahar, Fiona Johnson^{*}, Ashish Sharma

School of Civil and Environmental Engineering, University of New South Wales, Sydney, Australia

ARTICLE INFO

Article history:

Received 4 December 2016
Received in revised form 19 May 2017
Accepted 23 May 2017
Available online 25 May 2017
This manuscript was handled by A. Bardossy, Editor-in-Chief, with the assistance of Bruno Merz, Associate Editor

Keywords:

Statistical downscaling
BCSD
Bias correction
Rank-BCSD

ABSTRACT

Use of General Circulation Model (GCM) precipitation and evapotranspiration sequences for hydrologic modelling can result in unrealistic simulations due to the coarse scales at which GCMs operate and the systematic biases they contain. The Bias Correction Spatial Disaggregation (BCSD) method is a popular statistical downscaling and bias correction method developed to address this issue. The advantage of BCSD is its ability to reduce biases in the distribution of precipitation totals at the GCM scale and then introduce more realistic variability at finer scales than simpler spatial interpolation schemes. Although BCSD corrects biases at the GCM scale before disaggregation; at finer spatial scales biases are re-introduced by the assumptions made in the spatial disaggregation process. Our study focuses on this limitation of BCSD and proposes a rank-based approach that aims to reduce the spatial disaggregation bias especially for both low and high precipitation extremes.

BCSD requires the specification of a multiplicative bias correction anomaly field that represents the ratio of the fine scale precipitation to the disaggregated precipitation. It is shown that there is significant temporal variation in the anomalies, which is masked when a mean anomaly field is used. This can be improved by modelling the anomalies in rank-space. Results from the application of the rank-BCSD procedure improve the match between the distributions of observed and downscaled precipitation at the fine scale compared to the original BCSD approach. Further improvements in the distribution are identified when a scaling correction to preserve mass in the disaggregation process is implemented. An assessment of the approach using a single GCM over Australia shows clear advantages especially in the simulation of particularly low and high downscaled precipitation amounts.

© 2017 Elsevier B.V. All rights reserved.

1. Introduction

Climate change is likely to have major impacts on human and natural ecosystems and in particular water resources systems. The current generation of General Circulation Models (GCMs) have reasonably coarse spatial resolutions that precludes their use in many impact assessments (Gutmann et al., 2014). For better planning and management of water resources, climate projections are required at spatial scales that are finer than the GCM resolution. To reduce the gap between policy relevant information and climate model data, downscaling is used to transfer large-scale GCMs output to regionally relevant scales (Christensen et al., 2008; Mejia et al., 2012). Downscaling aims to retain all the large-scale

information provided by the climate model, and to add finer-scale information that climate model could not resolve (Kanamitsu and Kanamitsu, 2007).

Because computational requirements prohibit the widespread use of dynamic downscaling, for long term simulations statistical downscaling approaches are very common (Fowler et al., 2007; Wilby et al., 1998; Wood et al., 2004). This is particularly the case when multiple GCMs and/or multiple greenhouse gas emission scenarios need to be considered (Ahmed et al., 2013; Maurer et al., 2013). The Bias Correction Spatial Disaggregation (BCSD) method (Wood et al., 2004) is a popular approach that has been used in a number of studies to assess the hydrological impacts of climate change (Christensen et al., 2008; Maurer and Hidalgo, 2008; Payne et al., 2004; Shrestha et al., 2014; VanRheenen et al., 2004; Vicuna et al., 2007).

BCSD uses a three step approach for the downscaling. Monthly GCM simulations are bias corrected at the GCM grid scale using Quantile Mapping (QM) (Ines and Hansen, 2006; Wood et al., 2002). The simulations are then spatially disaggregated to match

^{*} Corresponding author at: Water Research Centre, School of Civil and Environmental Engineering, University of New South Wales, Kensington, New South Wales 2052, Australia.

E-mail address: f.johnson@unsw.edu.au (F. Johnson).

the fine-scale resolution of the observation data. Finally the monthly data is temporally disaggregated to produce a daily time series. Due to the popularity of the BCSD approach, a number of modifications have been proposed suggesting the bias correction step followed by spatial disaggregation of raw GCM output (Abatzoglou and Brown, 2012; Ahmed et al., 2013). These have generally focused on improving the final daily time series. It was found that this provided better performance for both daily maximum and minimum temperatures as well as precipitation (Ahmed et al., 2013). Abatzoglou and Brown (2012) used a 15-day moving window to define the probability distribution for bias correction, leading to improvements in the resulting daily time series.

An aspect of BCSD that has received limited attention is whether the spatial disaggregation step interferes with the bias correction at the coarse scale. Ahmed et al. (2013) comment on the sensitivity of the downscaled results to the initial coarse spatial resolution at which the bias correction is applied. This sensitivity is demonstrated in Fig. 1 where the GCM and BCSD simulations are compared to observations at coarse and fine spatial scales. For a single grid cell, Fig. 1a demonstrates that before applying the quantile mapping the probability distributions of the GCM and observations are very different. This bias is completely removed (as expected) following the bias correction. Once the GCM is bias corrected, then the simulations are spatially disaggregated. This step ensures that the spatial variability at the fine scale is appropriate (Wood et al., 2004). However it is clear from Fig. 1b that the spatial disaggregation leads to biases in the simulations at the finer scale, even though at the GCM scale the precipitation simulations are unbiased. Simple solutions to address this issue are developed in this paper.

The BCSD method was originally developed for downscaling monthly precipitation and temperature with the temporal disaggregation step as described above to generate daily time series. In principle, daily GCM output could also be downscaled directly using the BCSD approach. The modified BCSD method by Abatzoglou and Brown (2012) has already been used for daily data to improve the downscaling skill in reproducing fine scale temporal statistics. However spatial variability of daily precipitation is much higher than for monthly data and the interpolation scheme is unlikely to be able to represent this variability (Hwang and Graham, 2013). Hence, this paper will focus on the BCSD applied

to month data. Although it is not often the case that hydrologic simulations are made with monthly climate data, there are important implications of improving the monthly BCSD simulations for applications such as drought assessments where retaining the spatial and temporal characteristics of precipitation events is paramount.

This research thus examines the assumptions behind the BCSD method and highlights possible improvements to the approach. The approach adopted here is based on the BCSD algorithm detailed in Wood et al. (2004). Different versions of this algorithm have been reported in Wood et al. (2002) and (Maurer, 2007), which result in subtle differences in the downscaled field obtained. It should be noted that, although the original BCSD is applied for both temperature and precipitation, here the focus is on precipitation because the spatial variability of monthly precipitation data is much larger than for temperature data. Nevertheless, any improvement in the BCSD method could be used for temperature as well. The outline of the paper is as follows. Section 2 provides information about the observed and model data. In Section 3 details of the BCSD method are presented along with the proposed modifications to address the spatial disaggregation problems highlighted above. Results are provided in Section 4, with a discussion and conclusions in Section 5.

2. Data

Fine scale gridded monthly observed precipitation data were used to correct the precipitation simulations from a single GCM from the Coupled Model Inter-comparison Project 5 (CMIP5), namely CSIRO-MK3.6 model. Twenty years (1980–1999) precipitation from current climate was used for the analyses as per Wood et al. (2004). A single GCM is appropriate for use here because the bias correction step will ensure a perfect match in the distributions between the GCM and observed data at the GCM scale, allowing the focus to be on the spatial disaggregation step, rather than addressing uncertainties of climate model simulations.

The observed precipitation data is provided by the Australian Bureau of Meteorology and the monthly product has a spatial resolution of 0.25°. This gridded dataset is produced using an optimised Barnes successive correction technique that applies a weighted averaging process to the station data (Jones et al., 2009).

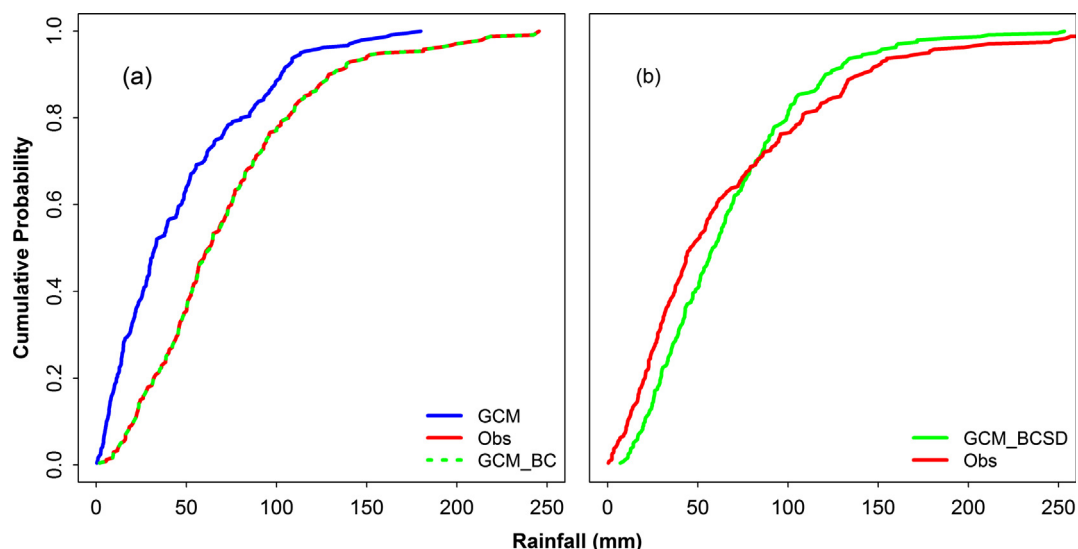


Fig. 1. CDFs of monthly precipitation from a single grid comparing the observations and GCM simulations at a) coarse and b) fine scale. At the coarse scale, QM matches the GCM simulations with the observations. At the fine scale the BCSD simulations do not match the observations.

3. Methodology

The process of BCSD is well described by Wood et al. (2004). The important details of the method are shown in Fig. 2 and summarised here to motivate the proposed modifications. The monthly precipitation from the GCM simulation and observations are denoted as P^g and P^o respectively. The unit spatial domain is the coarse GCM grid cell which will be disaggregated to a single grid cell at the fine-scale of the observations, denoted by the subscript i,j . The original observations at fine scale are thereby denoted as $P_{i,j,t}^o$. Superscripts are used to indicate the different steps in the BCSD process, g' refers to the bias corrected GCM simulations, b refers to data that has been bilinearly interpolated and d indicates the spatially disaggregated data. All calculations are performed separately for each month of the year to allow for seasonal variations in the biases and the spatial patterns of precipitation. Thus the subscript t is used to represent the precipitation for year t for the month of interest.

The first step of BCSD applies QM to the GCM precipitation by matching the empirical CDFs of P^o and P^g for each month. GCM biases are corrected at native scale to exactly match the modelled distribution with the observed. Therefore, any feature that is not correctly picked up by the GCM can be corrected without allowing it to propagate in downscaled product. As shown in Fig. 2a, the QM requires that the fine-scale observations have been aggregated to the same grid resolution as the GCM (Maraun, 2013), in this case $1.875^\circ \times 1.85^\circ$. Quantile mapping can remove large systematic biases in the GCM simulations, even for daily rainfall extremes (Cannon et al., 2015). For future simulations, it is assumed that the biases for each quantile remain constant over time (Li et al., 2010; Shrestha et al., 2014). Although, the assumption of stationary biases is questionable (Christensen et al., 2008; Ehret et al., 2012; Maraun, 2012; Maraun et al., 2010; Nahar et al., 2017), for consistency with previous BCSD applications, this assumption is

retained. Also for these relatively short periods, it has been shown that no consistent trends in bias are found suggesting that the assumption is acceptable (Nahar et al., 2017). Furthermore, while QM is used here for illustrating the proposed method, it can be replaced with other suitable correction alternatives that modify systematic low-frequency variability biases, of considerable relevance in hydrological applications (Johnson and Sharma, 2012; Mehrotra and Sharma, 2012; Nguyen et al., 2016). The bias corrected precipitation ($P_t^{g'}$) is then bi-linearly interpolated to the fine-scale grid. However, this leads to precipitation ($P_{i,j,t}^{g',b}$) that lacks the correct spatial variability. Therefore a correction step is required to produce the final spatially disaggregated simulations. The focus of this paper is on this correction step and whether it can be improved compared to previous implementations of BCSD.

The loss of spatial variability that results from bilinear interpolation from the coarse scale to the fine scale is demonstrated in panels b)–d) of Fig. 2. Climatological average precipitation over Australia is shown in Fig. 2b for observed precipitation at the original resolution of 0.25° . The precipitation is aggregated to the scale of the GCM using area-averaging (Fig. 2c). Using bilinear interpolation to return the data back to the original fine resolution (Fig. 2d) leads to a substantially smoother pattern of precipitation across Australia. BCSD proposes that the smooth precipitation fields can be corrected by applying a multiplicative correction factor that represents the difference between Fig. 2b and Fig. 2d. This correction factor is termed the anomaly field by Wood et al. (2004) and it is calculated as the ratio of the original monthly mean observations ($P_{i,j,t}^o$) and the aggregated and then bilinearly interpolated monthly mean observations ($P_{i,j,t}^{o,b}$). In Fig. 2a the anomaly is denoted as $\bar{\Delta}_{i,j,t}$ where the overbar represents the anomaly field calculated from the climatological mean precipitation for each month and is applied to the bilinearly interpolated GCM precipitation ($P_{i,j,t}^{g',b}$) to create the disaggregated product ($P_{i,j,t}^{g',d}$).

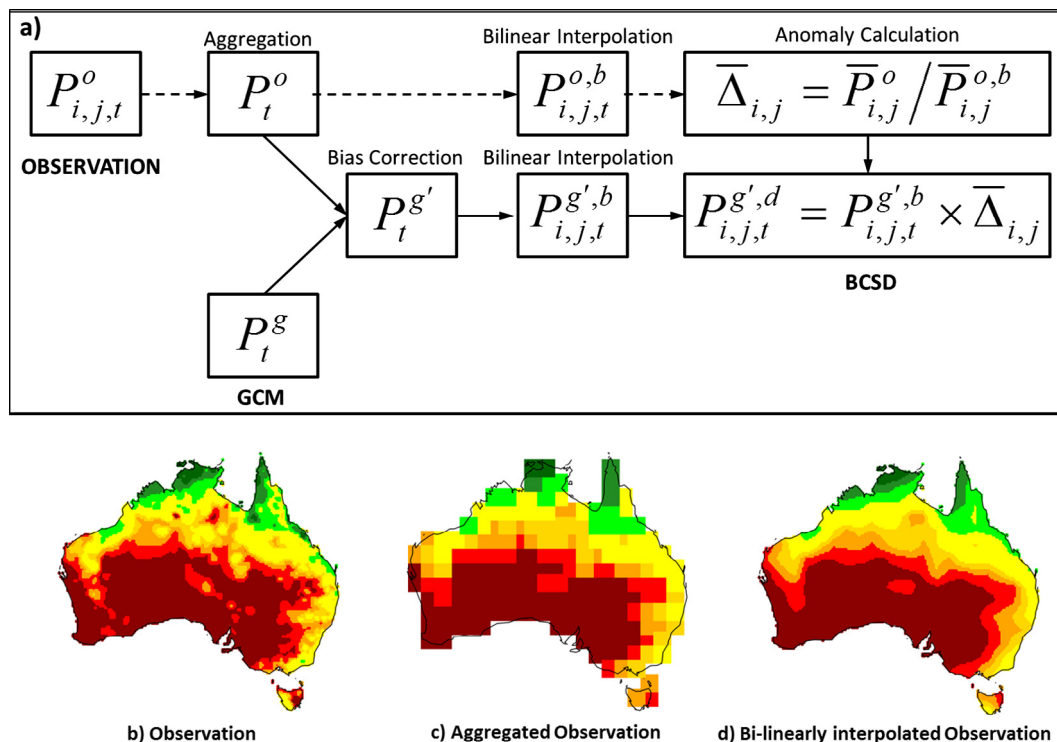


Fig. 2. a) Flowchart of BCSD process. b) Original fine scale observation, c) aggregated observation at GCM scale, d) bi-linearly interpolated aggregated observation at fine scale. The anomaly field is the ratio of b and d which gives a measure of spatial information loss due to the aggregation and disaggregation process.

4. Results

4.1. Exploring the nature of the spatial anomalies

As demonstrated by the results presented in Fig. 1, the focus of this work is on how to reduce the bias in the downscaled GCM simulations. The bias is clear for low and high precipitation amounts whilst the median precipitation estimates are reasonably unbiased. This suggests that the anomaly field does not represent the full range of variation of disaggregated precipitation. Because the anomaly field is calculated as the ratio of the original and interpolated mean monthly observations, it does not reflect the true spatial variability when precipitation in any particular month is high or low. The use of climatological mean precipitation for the observations and interpolated data results in some information being lost. If the anomaly field is considered for each month and year of data separately, it is found that the range of anomalies is quite substantial and may vary from 0.2 to around 2.5.

To highlight the difference in projections between using mean anomaly field and the full range of anomaly values, a resampling scheme has been applied such that all 20 possible anomaly values for each month of the year are used to calculate the disaggregated precipitation amounts. The results from this are shown in Fig. 3 where an ensemble of disaggregated precipitation has been created by repeating the resampling the monthly anomalies 20 times. This ensemble shows the potential uncertainty in the BCSD disaggregated precipitation and suggests that the biases in the high and low precipitation may be able to be addressed as some of the ensemble members are much closer to the observations than the original BCSD estimate.

Fig. 3 suggests that the anomaly field could be used in a way that produces more realistic values when compared to the observations. However, this requires an understanding of how the anomaly field varies temporally. Since the anomaly field is created only from observations, it is possible to explore the relationship between the anomalies and the observed precipitation for each month. In Fig. 4, the anomalies are compared to the monthly precipitation for all the fine grid locations for the months of January and July. The anomalies are grouped according to latitude bands to allow for varying relationships with the diverse climate across Australia. There are clear relationships between the size of the anomaly and the precipitation amounts within the different latitude bands, particularly evident for the largest precipitation events. Similar results are found for other months.

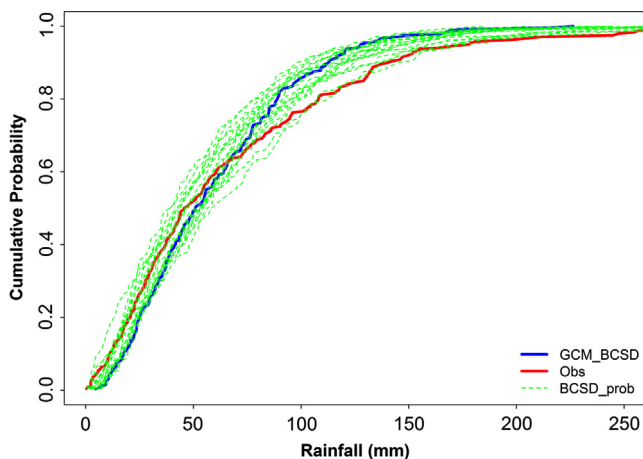


Fig. 3. Range of CDFs when possible values of the anomalies are used. Thin green lines represent CDFs when the probabilities of the anomaly for each month were used. Red line shows CDF of observations and the blue line shows CDF of BCSD using the mean anomaly field. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Spearman's rank correlation, which evaluates the monotonic relationship between two continuous variables (Yue et al., 2002), has been used to test the strength of the relationship between the monthly observed precipitation and monthly anomaly at individual grid scale as shown in Fig. 5. It is found that the relationship between the precipitation amounts and anomalies is significant (at the 0.05 level) at most locations as shown in Fig. 5. In general, the anomalies increase with rainfall amount, leading to higher correlations in high rainfall areas. However, in some areas such as Tasmania, the strongest correlations occur in the areas with lower rainfall totals whilst the highest rainfalls were found to not have significant rainfall-anomaly correlations. It is therefore important to explore the spatial variation of locations where the additional information from the rainfall-anomaly relationship could be useful for improving the BCSD method. As evident in Fig. 5, correlations between precipitation amounts and anomalies in July tend to be greater than in January for most of the country.

What does a significant positive correlation between the anomalies and precipitation amounts mean in practice? When the monthly precipitation at a fine grid cell is high, the interpolation process is less likely to have recreated this high amount. This indicates that high precipitation tend to be associated with more spatially discrete precipitation events. There is more variation within the larger grid than for small precipitation amounts.

It is worth noting that there are some areas where the correlations between the precipitation amounts and anomalies are not very strong and are even sometimes negative. Most of the points are in areas of relatively high relief. One hypothesis is at these locations the interpolation process may be introducing too much error into the precipitation amounts because the neighboring locations have different precipitation profiles due to the variations in topography. Jones et al. (2009) suggested that the smoothness of the local topography or lack of the gauge station influences the estimation errors in a gridded rainfall product and the accuracy of the analyses will also be limited in regions where the station network is insufficient. The negative correlations in the center of Australia are therefore, more likely related to possible inconsistencies in the AWAP data in this area due to low station densities.

4.2. Suggested modification to BCSD

The results of the rank correlation test indicate that larger anomalies are more likely to occur for higher precipitation. This finding can be used to address the shortcomings of BCSD by using all the anomalies for each grid cell rather than a single value representing a mean correction. To achieve this, the anomalies are applied to the bias corrected GCM simulations according to the rank of the precipitation, for example the largest of the twenty anomalies is used for the largest precipitation total for that month. This is shown in Eq. (1) where $P_{i,j,t,m}^{g,b}$ is the m -th ranked bilinearly interpolated precipitation amount and $\Delta_{i,j,m}$ is the corresponding m -th ranked anomaly. The new bias corrected and spatially disaggregated product $P_{i,j,t,m}^{g,dr}$ is denoted as the ranked BCSD (rBCSD) and is calculated using Eq. (1).

$$P_{i,j,t}^{g,dr} = P_{i,j,t,m}^{g,b} \times \Delta_{i,j,m} \quad (1)$$

Does the use of rank anomalies improve the match between the observations and GCM at the fine scale? Fig. 6a shows the CDFs of observations, BCSD and rBCSD at the same location as Fig. 1. The improvements produced by rBCSD are clear with the overestimated high precipitation amounts (>75 mm) substantially reduced when the rank version is used instead of original BCSD. Improvements are also seen across the full distribution of precipitation.

Are similar improvements seen in all locations? To summarise the performance of the BCSD and rBCSD, the two-sample

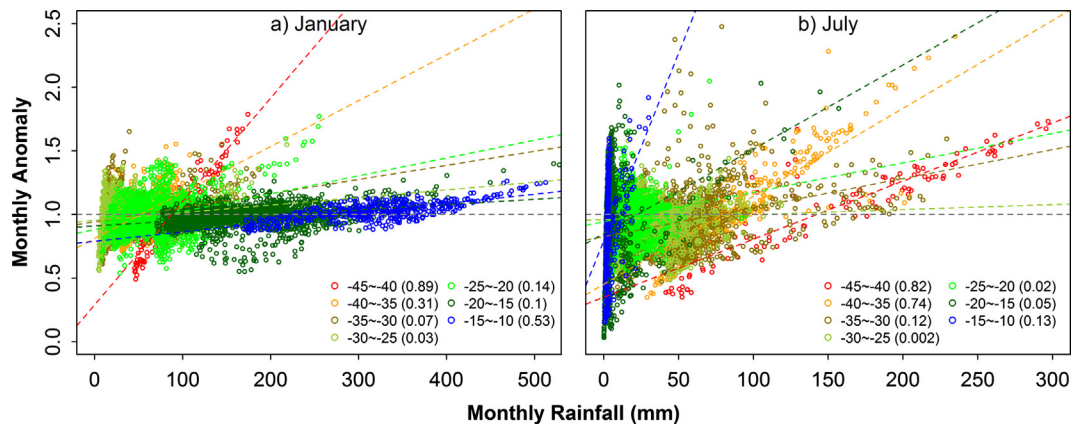


Fig. 4. Monthly anomaly vs monthly precipitation plot for a) January and b) July. Anomalies are shown in different colours depending on the latitude band. Anomalies increase with increasing precipitation. R^2 values for these regression lines are shown in brackets. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

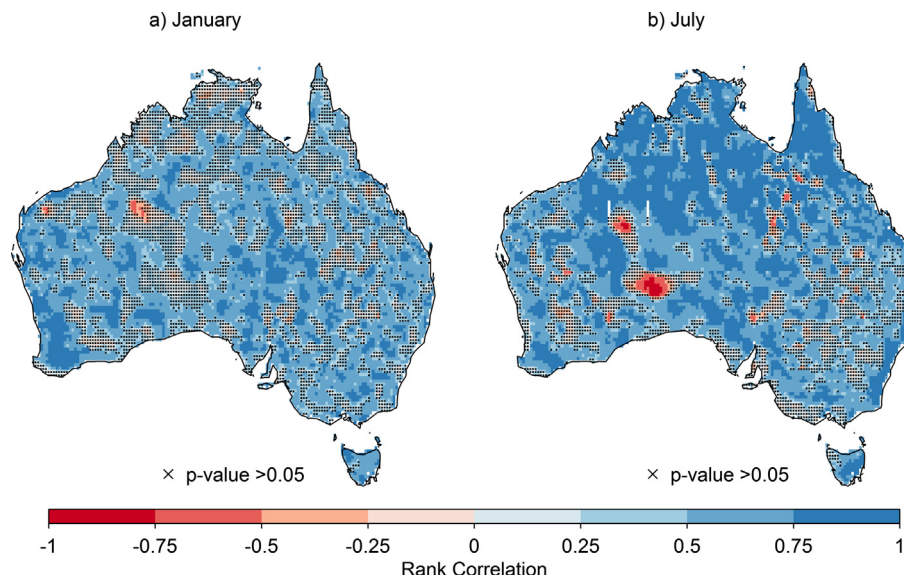


Fig. 5. Rank correlation between anomaly and observed rainfall for a) January and b) July over Australia. Dark blue means high correlation and red means negative correlation. Black cross indicates the location where the p-value is not significant; at other locations p-values are significant. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Kolmogorov–Smirnov test (K-S test) was used (Sheskin, 2003). The K-S test is applied separately to compare BCS and rBCS to the observations. A better match in the CDFs will lead to a smaller value for the K-S test statistic (the maximum distance between the two distribution functions). The average value of the K-S test statistic is 0.137 for the BCS and is 0.064 for the rBCS, showing that the rBCS provides a better match to the observations. The K-S test statistics are plotted in Fig. 6(b) along with contours of the number of grid points due to the large number of overlapping points. This clearly reveals that the rBCS in the vast majority of locations (85% of grid locations) provides a better match to the observations than the BCS. Additional testing using the Anderson–Darling test (Anderson and Darling, 1952) resulted in similar conclusions.

4.3. Mass conservation

Mass conservation is an important aspect that must be considered for any disaggregation scheme, particularly when considering hydrological downscaling applications where the water balance is

a key driver of the system. Because the anomaly field is based on the climatological mean, BCS ensures that the mass balance is properly preserved in the disaggregation process. However the ranked anomalies proposed above can lead to problems with the mass balance because there is no spatial information included in the disaggregation as the calculations for each fine grid proceed separately. Thus a number of grids may be assigned particularly high or low anomalies at the same time step causing the overall precipitation aggregated at the coarse scale to become biased. Fig. 7a shows the mass balance is preserved in BCS, while Fig. 7b shows that rBCS has affected the mass conservation. The bias tends to be high, such that the high precipitation amounts at adjacent cells contribute too much precipitation at the coarse GCM scale. To overcome this problem, rank BCS simulations were scaled to ensure mass conservation. A scaling factor (ε) is adopted, which is the ratio of quantile corrected GCM ($P_t^{g^*}$) and aggregated rBCS ($P_t^{g^*,dr}$). The scaling factor is applied at the fine scale and is constant for all fine grids within a coarse GCM cell. The post-processing bias correction approach of Abatzoglou and Brown (2012) could also be used to address this problem. After scaling,

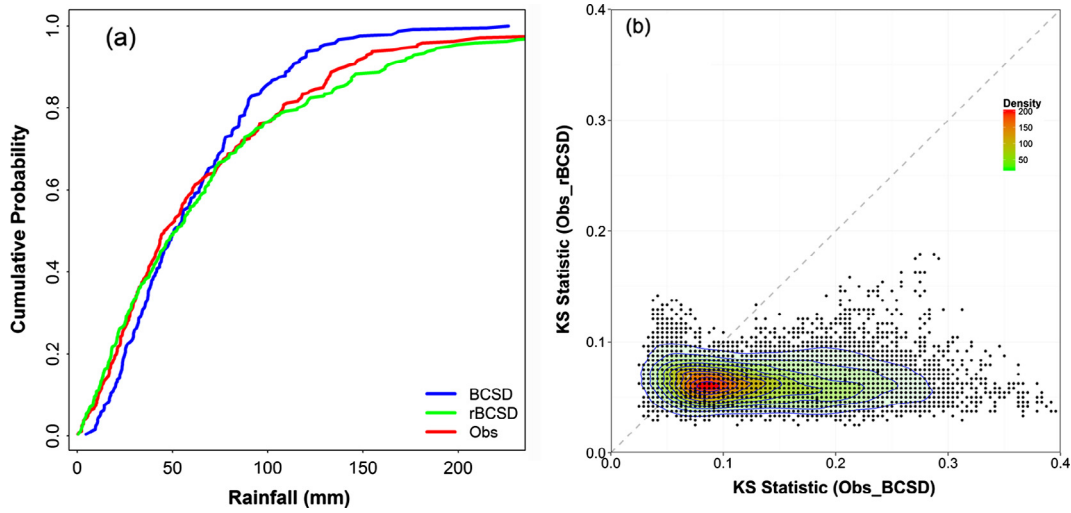


Fig. 6. (a) Improvement in CDF when ranked anomaly field is used. The red line represents the CDF of the observation; the blue line is the CDF of current BCSD, and light green line is CDF for Ranked BCSD. (b) K-S statistics between observed and Rank BCSD compared to K-S statistics between observed and normal BCSD at grid level. The point density is shown with colored contours in descending order from red to green. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

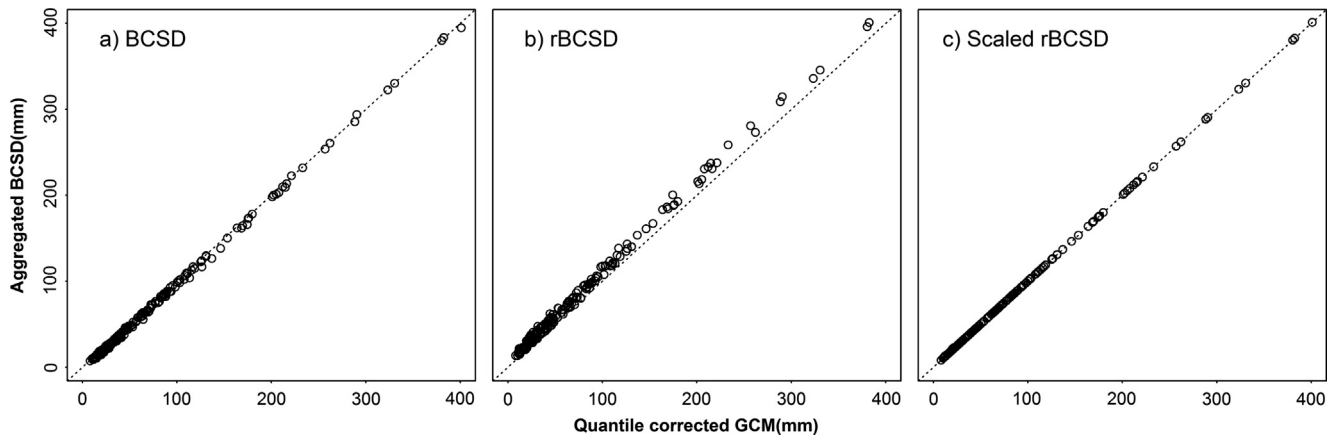


Fig. 7. Monthly precipitation amounts calculated at the GCM scale for (a) BCSD, (b) rBCSD and (c) rBCSD after scaling.

the final BCSD product ($p_{ij,t}^{g',ds}$) is calculated as the product of rBCSD and the scaling factor following the Eq. (2).

$$p_{ij,t}^{g',ds} = p_{ij,t}^{g',dr} \times \varepsilon \quad (2)$$

By design, this scaled rBCSD when aggregated to GCM scale perfectly preserves the total mass at this scale as shown in Fig. 7c.

Clearly the scaling factor could result in other attributes of the downscaled precipitation being affected. To assess this impact the K-S statistics have been recalculated for the scaled rBCSD (Fig. 8a). The mean value of the K-S statistic was 0.137 for BCSD and 0.064 for the rBCSD case; the scaled rBCSD also has a mean K-S test statistic of 0.064 which shows that the scaling on average does not lead to deterioration in the fine scale precipitation distributions. Although the average K-S test statistics remains unchanged, the improvements are now seen at 91% of grid locations.

Fig. 8b shows the locations where each method (BCSD and scaled rBCSD) leads to the best representation of the observed precipitation distribution. As suggested by Fig. 8a, at the vast majority of locations the scaled rBCSD leads to the better results. There are some locations where the ranked anomalies do not led to improvements. To identify the reason behind this, the difference in the K-S statistics of BCSD and scaled rBCSD is plotted against the correla-

tion between the observations and anomaly (Fig. 9). The plot clearly shows that a poor rank correlation between precipitation and anomaly is associated with less improvement in K-S statistics in rBCSD.

As discussed earlier, one possible reason for the low correlations between precipitation amounts and the anomaly values could be the influence of topography. High elevation can change the precipitation patterns and influence spatial correlations and therefore the rank correlation between precipitation and anomalies may change. Simple statistical downscaling methods often fail to adequately reproduce spatial and temporal variability, especially in regions of complex terrain characterized by steep spatial gradients (Abatzoglou and Brown, 2012). Accurate representation of spatiotemporal variability in climate variables such precipitation is important for accurately predicting regional hydrologic behavior (Hwang and Graham, 2014) and hence further work is recommended to address these remaining issues with the BCSD disaggregation process.

4.4. Practical applications of monthly rBCSD estimates

Correctly representing monthly precipitation statistics is important for large scale water availability assessments and in particular

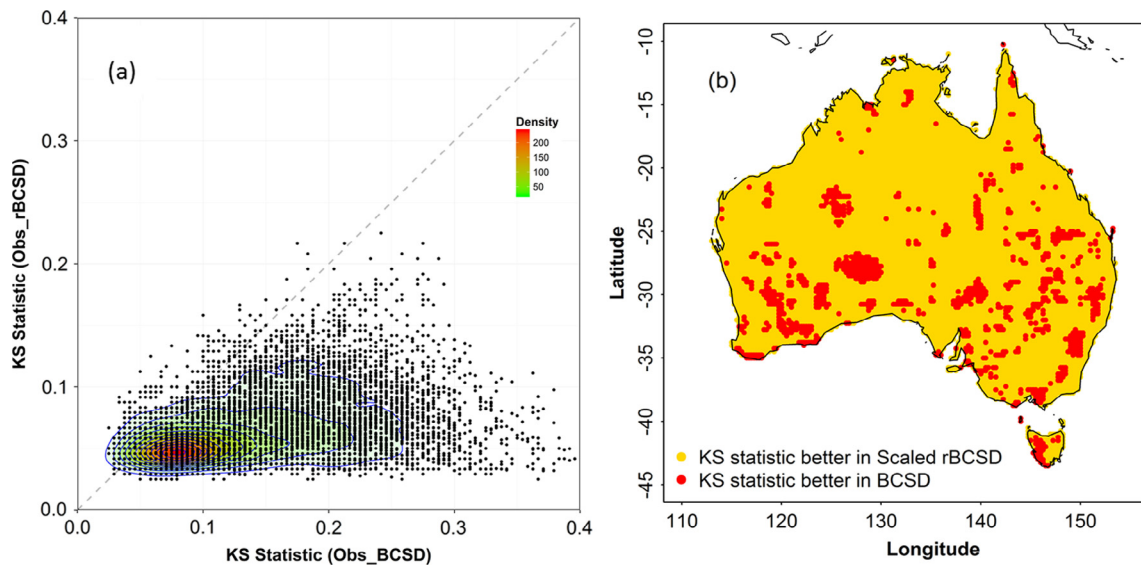


Fig. 8. (a) K-S statistics between observed and scaled rBCSD Vs K-S statistics between observed and normal BCSD. (b) Improvement in precipitation distribution in rank BCSD. Yellow shows K-S statistic is better in rank BCSD and red means the K-S statistic is better in original BCSD.

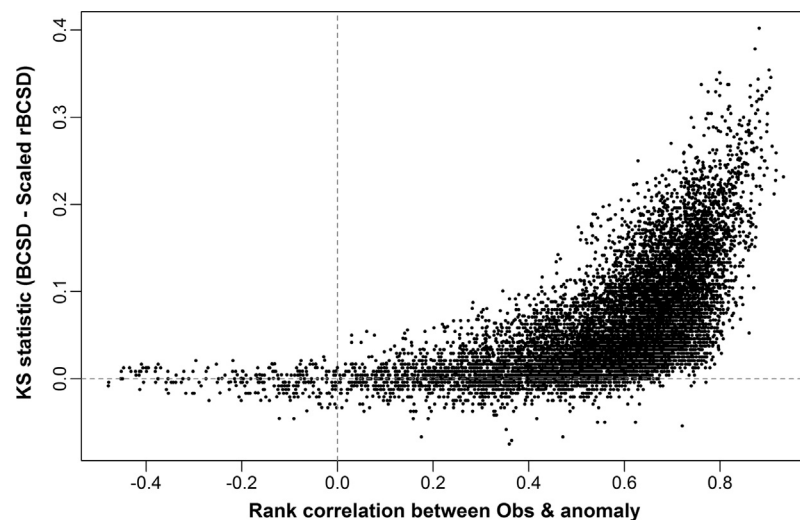


Fig. 9. Difference in K-S statistic of BCSD and scaled rBCSD compared to the rank correlation of precipitation amounts and the anomalies.

drought evolution, particularly in Australia where the duration of significant droughts is several months to many years (Kiem and Franks, 2004). The previous section presented results showing the improvement in the scaled rBCSD over the full distribution of precipitation. However to focus on the precipitation amounts that are likely to cause flood and moisture and streamflow deficits, the 5th and 95th percentile of precipitation anomalies have been considered from both BCSD and scaled rBCSD simulations and the results are shown in Fig. 10. For all three data sets (observations, BCSD and scaled rBCSD), the precipitation anomaly series has been calculated by removing the monthly climatological mean precipitation. Completely dry months have been removed from this analysis. The set of all precipitation anomalies at each grid cell (12 months $\times$ 20 years) have then been ranked and the 5th and 95th percentile values extracted. These are then compared for the observations to the two downscaled products. There is a clear bias in the representation of relatively dry months in the BCSD simulations which on average are larger than the observations. Higher rainfalls are under estimated by the BCSD method. The

scaled rBCSD improves the simulations of the dry and wet months with much smaller biases. The temporal pattern of dry periods is also an important driver of droughts and bias correction approaches have been developed that can address the modelling of persistence in GCMs (Johnson and Sharma, 2012). Future work could consider implementing such an approach within the BCSD logic.

4.5. Validation

The ultimate aim is to use the new rank based BCSD method for downscaling future GCM simulation and thus improving local climate predictions. As in all statistical downscaling and bias correction approaches, the anomalies estimated over the current observed period are assumed to apply in the future. To assess the implications of this assumption, a validation process has been designed such that the anomalies calculated from calibration period (80 years) have been applied to validation period (20 years). For example, if the calibration period used for creating the anomaly

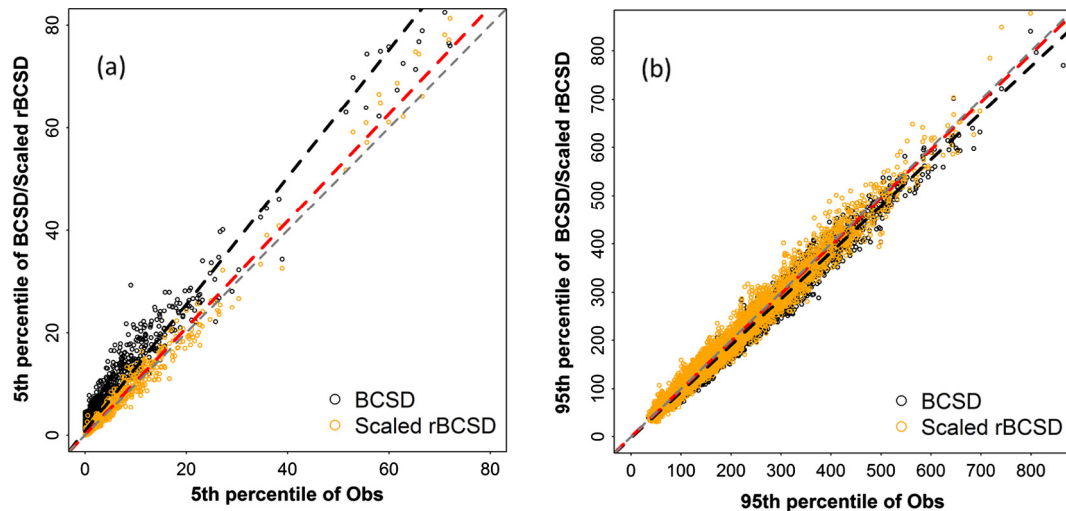


Fig. 10. Comparison of the 5th and 95th percentile precipitation in BCSD and scaled rBCSD with that of observation at all grid locations. The black dotted line represents the regression line for BCSD. The regression line for scaled rBCSD (red dotted line) almost matches with 1:1 line (grey dotted line). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

fields is 1900–1979, the validation period will be 1980–1999. With the 100 years of data used in this study, five different validation periods have been tested.

The bias corrected GCM simulations (using quantile matching) for the validation period are disaggregated using the anomaly information from the observations in the calibration period. The bias correction of the GCM also uses the data from the calibration period. The BCSD and scaled rBCSD simulations are compared to the observations using a number of different statistics. The root mean square error (RMSE) is calculated at each grid cell to assess the representation of the climatological average precipitation in each month. The full distribution of all precipitation amounts at each grid cell compared to the observations is again tested with the K-S test. Finally the 5th and 95th percentiles from BCSD and scaled rBCSD simulations have also been compared.

The RMSE results are presented in Table 1. The BCSD method leads to the smallest errors for all five of the validation periods. This is because the use of the mean anomaly in the BCSD method leads to better improvements in the mean rainfall. However, the simulations at higher and lower end of the distribution rather than

the mean are more important in ensuring that extreme events like floods and droughts are properly captured and these are considered next.

To represent the modelling of the full distribution of precipitation amounts, the K-S statistics averaged over all grids are presented in Table 2. The rBCSD method outperforms the BCSD in all five validation tests, with improvements seen at 80–90% of grid cells.

The average K-S statistics across the five validation periods are shown in Fig. 11 which shows the range of values of the K-S statistic compared to the means presented in Table 2. In Fig. 11a a density plot of the two sets of K-S statistics is provided. Most of the points are below 1 to 1 line indicating improvements in the K-S statistic for the scaled rBCSD method. As evident from Table 2, improvements from the scaled rBCSD method are apparent at most locations across the continent. Locations where the BCSD method performs better in validation are broadly the same as found previously (e.g. Fig. 8). This suggests that in general the relationships between rainfall amounts and anomalies hold when applied to a different period. Thus similar improvements can be expected when

Table 1
Average RMSE for different validation periods.

Calibration period	Validation period	RMSE		
		BCSD	rBCSD	Scaled rBCSD
1900–1979	1980–1999	10.11	18.55	10.64
1900–1959, 1980–1999	1960–1979	9.42	16.94	9.92
1900–1939, 1960–1999	1940–1959	10.66	19.33	11.13
1900–1919, 1940–1999	1920–1939	12.00	22.32	12.47
1920–1999	1900–1919	10.50	20.33	11.02

Table 2
Spatially averaged KS-statistics for different validation periods.

Calibration period	Validation period	KS-statistics		rBCSD leads to best results (% of grid cells)
		BCSD	Scaled rBCSD	
1900–1979	1980–1999	0.14	0.097	81
1900–1959, 1980–1999	1960–1979	0.16	0.100	89
1900–1939, 1960–1999	1940–1959	0.18	0.12	92
1900–1919, 1940–1999	1920–1939	0.18	0.12	91
1920–1999	1900–1919	0.16	0.107	86

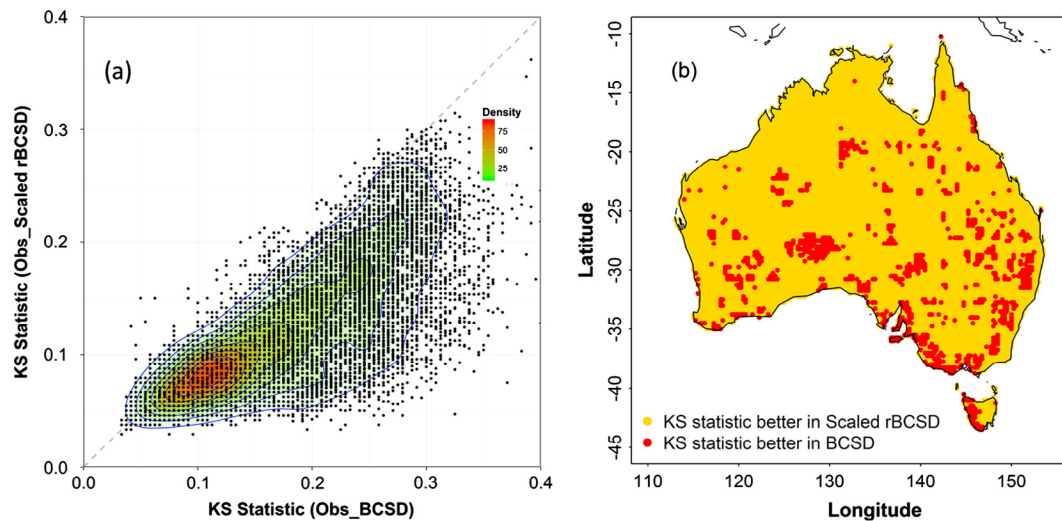


Fig. 11. (a) Density plot of K-S statistics between observed and scaled rBCSD Vs K-S statistics between observed and BCSD in validation. (b) Improvement in precipitation distribution in scaled rank BCSD.

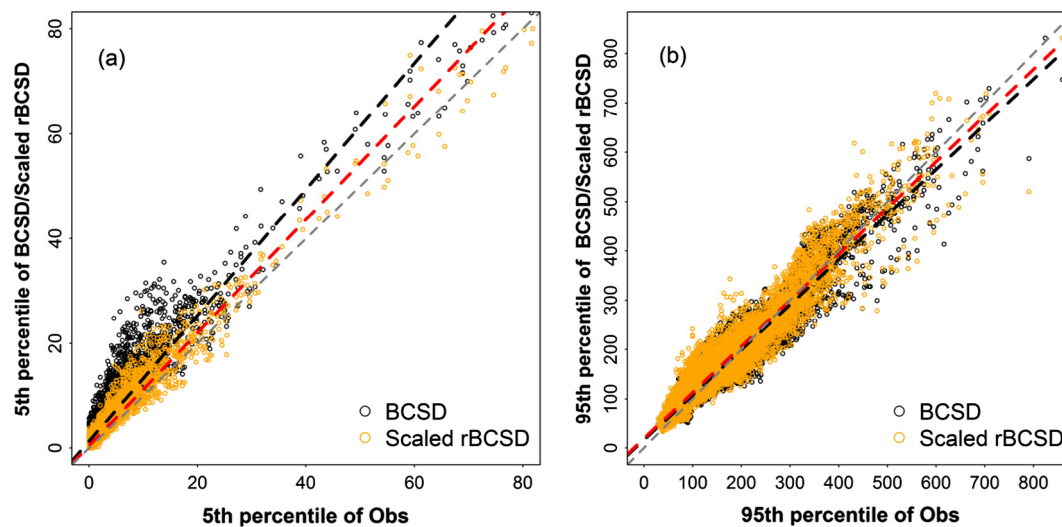


Fig. 12. Comparison of 5th and 95th percentile precipitation in BCSD and scaled rBCSD with that of observation at all grid locations for validation. The black and red dotted line represents the regression line for BCSD and scaled rBCSD.

the method is used with future climate simulations, with the caveat that the assumptions of bias stationarity are probably stronger when moving into the different future climate.

Finally the 5th and 95th percentiles of BCSD and scaled rBCSD have been calculated and compared to those of observation. Fig. 12 shows the results, where scaled rBCSD performs well in simulating 5th and 95th percentile of rainfall events compared to BCSD. As found in the testing reported in Section 4.4 the biases in the BCSD are particularly evident for low precipitation amounts and high precipitation months tend to be underestimated. Both problems are addressed even in the validation tests with the scaled rBCSD method.

5. Conclusions

GCM simulations are vital for climate prediction but cannot be directly used for hydrologic modelling as they do not properly represent climate variability at the local scale. From a policy perspective, better planning and management of water resources

information requires information on the future climate at time scales ranging from hours to years depending on the application and a range of spatial scales. Downscaling can help to address these requirements, particularly when catchments are the spatial scale of interest. This study proposes an improvement to an existing method of spatial downscaling that is capable of producing more accurate precipitation at relevant scales.

We have shown that the popular downscaling method BCSD leads to biases in the precipitation projections at the fine scale, but that these can be easily addressed by relating the spatial variability correction in BCSD to the corresponding precipitation amount. When the spatial variability correction is applied according to the ranked precipitation, improvements are found in the representation of the fine scale precipitation in the new rBCSD approach. Ranking the anomalies could possibly alter the spatial covariance structure, which is a common issue in all existing bias correction procedure unless a spatial bias correction method is adopted. A simple correction is applied that can also ensure that precipitation totals at the GCM scale are also conserved even when the ranking is applied. It is recommended that this new approach is

adopted in studies where the BCSD is being considered. Future work could focus on a combined approach that incorporates the scaled rBCSD method and preserves the observed spatial covariance structure of the precipitation field.

Acknowledgements

The comments of three anonymous reviewers have greatly improved the presentation of this work. The authors gratefully acknowledge the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP, and thank the climate modelling group "Commonwealth Scientific and Industrial Research Organization in collaboration with Queensland Climate Change Centre of Excellence" (CSIRO-QCCCE) for producing and making available their model output that served as the basis of the results reported here. 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. Precipitation observation data used in this research is available from <http://www.bom.gov.au/climate/how/newproducts/IDCrainpercgrids.shtml>. The authors also acknowledge the Australian Research Council for partial funding for this work.

References

- Abatzoglou, J.T., Brown, T.J., 2012. A comparison of statistical downscaling methods suited for wildfire applications. *Int. J. Climatol.* 32 (5), 772–780.
- Ahmed, K.F., Wang, G., Silander, J., Wilson, A.M., Allen, J.M., Horton, R., Anyah, R., 2013. Statistical downscaling and bias correction of climate model outputs for climate change impact assessment in the U.S. northeast. *Global Planet. Change* 100, 320–332. <http://dx.doi.org/10.1016/j.gloplacha.2012.11.003>.
- Anderson, T.W., Darling, D.A., 1952. Asymptotic theory of certain "goodness of fit" criteria based on stochastic processes. *Ann. Math. Stat.*, 193–212.
- Cannon, A.J., Sobie, S.R., Murdock, T.Q., 2015. Bias correction of GCM precipitation by quantile mapping: how well do methods preserve changes in quantiles and extremes? *J. Clim.* 28 (17), 6938–6959.
- Christensen, J.H., Boberg, F., Christensen, O.B., Lucas-Picher, P., 2008. On the need for bias correction of regional climate change projections of temperature and precipitation. *Geophys. Res. Lett.* 35 (20). <http://dx.doi.org/10.1029/2008GL035694>.
- Ehret, U., Zehe, E., Wulfmeyer, V., Warrach-Sagi, K., Liebert, J., 2012. HESS Opinions "Should we apply bias correction to global and regional climate model data?". *Hydrol. Earth Syst. Sci. Discuss.* 9 (4), 5355–5387.
- Fowler, H.J., Blenkinsop, S., Tebaldi, C., 2007. Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling. *Int. J. Climatol.* 27 (12), 1547–1578.
- Gutmann, E., Pruitt, T., Clark, M.P., Brekke, L., Arnold, J.R., Raff, D.A., Rasmussen, R. M., 2014. An intercomparison of statistical downscaling methods used for water resource assessments in the United States. *Water Resour. Res.* 50 (9), 7167–7186.
- Hwang, S., Graham, W.D., 2013. Development and comparative evaluation of a stochastic analog method to downscale daily GCM precipitation. *Hydrol. Earth Syst. Sci.* 17 (11), 4481–4502.
- Hwang, S., Graham, W.D., 2014. Assessment of alternative methods for statistically downscaling daily GCM precipitation outputs to simulate regional streamflow. *J. Am. Water Resour. Assoc.* 50 (4), 1010–1032.
- Ines, A.V., Hansen, J.W., 2006. Bias correction of daily GCM rainfall for crop simulation studies. *Agric. For. Meteorol.* 138 (1), 44–53.
- Johnson, F., Sharma, A., 2012. A nesting model for bias correction of variability at multiple time scales in general circulation model precipitation simulations. *Water Resour. Res.* 48 (1). <http://dx.doi.org/10.1029/2011WR010464>.
- Jones, D.A., Wang, W., Fawcett, R., 2009. High-quality spatial climate data-sets for Australia. *Austr. Meteorol. Oceanogr. J.* 58 (4), 233.
- Kanamaru, H., Kanamitsu, M., 2007. Scale-selective bias correction in a downscaling of global analysis using a regional model. *Mon. Weather Rev.* 135 (2), 334–350.
- Kiem, A.S., Franks, S.W., 2004. Multi-decadal variability of drought risk, eastern Australia. *Hydrol. Process.* 18 (11), 2039–2050.
- Li, H., Sheffield, J., Wood, E.F., 2010. Bias correction of monthly precipitation and temperature fields from Intergovernmental Panel on Climate Change AR4 models using equidistant quantile matching. *J. Geophys. Res.* (1984–2012) 115 (D10).
- Maraun, D., 2012. Nonstationarities of regional climate model biases in European seasonal mean temperature and precipitation sums. *Geophys. Res. Lett.* 39 (6).
- Maraun, D., 2013. Bias correction, quantile mapping, and downscaling: revisiting the inflation issue. *J. Clim.* 26 (6), 2137–2143.
- Maraun, D., Wetterhall, F., Ireson, A., Chandler, R., Kendon, E., Widmann, M., Brienen, S., Rust, H., Sauter, T., Themeßl, M., 2010. Precipitation downscaling under climate change: recent developments to bridge the gap between dynamical models and the end user. *Rev. Geophys.* 48 (3).
- Maurer, E.P., 2007. Uncertainty in hydrologic impacts of climate change in the Sierra Nevada, California, under two emissions scenarios. *Clim. Change* 82 (3–4), 309–325.
- Maurer, E.P., Hidalgo, H.G., 2008. Utility of daily vs. monthly large-scale climate data: an intercomparison of two statistical downscaling methods. *Hydrol. Earth Syst. Sci.* 12 (2), 551–563. <http://dx.doi.org/10.5194/hess-12-551-2008>.
- Maurer, E., Das, T., Cayan, D., 2013. Errors in climate model daily precipitation and temperature output: time invariance and implications for bias correction. *Hydrol. Earth Syst. Sci.* 17 (6), 2147–2159.
- Mehrotra, R., Sharma, A., 2012. An improved standardization procedure to remove systematic low frequency variability biases in GCM simulations. *Water Resour. Res.* 48 (12). <http://dx.doi.org/10.1029/2012WR012446>.
- Mejia, J.F., Huntington, J., Hatchett, B., Koracin, D., Niswonger, R.G., 2012. Linking global climate models to an integrated hydrologic model: using an individual station downscaling approach. *J. Contemp. Water Res. Educ.* 147 (1), 17–27.
- Nahar, J., Johnson, F., Sharma, A., 2017. Assessing the extent of non-stationary biases in GCMs. *J. Hydrol.* Accepted (press).
- Nguyen, H., Mehrotra, R., Sharma, A., 2016. Correcting for systematic biases in GCM simulations in the frequency domain. *J. Hydrol.* 538, 117–126.
- Payne, J.T., Wood, A.W., Hamlet, A.F., Palmer, R.N., Lettenmaier, D.P., 2004. Mitigating the effects of climate change on the water resources of the Columbia River basin. *Clim. Change* 62 (1–3), 233–256.
- Sheskin, D.J., 2003. *Handbook of Parametric and Nonparametric Statistical Procedures*. CRC Press.
- Shrestha, R.R., Schnorbus, M.A., Werner, A.T., Zwiers, F.W., 2014. Evaluating hydroclimatic change signals from statistically and dynamically downscaled GCMs and hydrologic models. *J. Hydrometeorol.* 15 (2), 844–860.
- VanRheenen, N.T., Wood, A.W., Palmer, R.N., Lettenmaier, D.P., 2004. Potential implications of PCM climate change scenarios for Sacramento-San Joaquin River Basin hydrology and water resources. *Clim. Change* 62 (1–3), 257–281.
- Vicuna, S., Maurer, E.P., Joyce, B., Dracup, J.A., Purkey, D., 2007. *The sensitivity of California water resources to climate change scenarios*. Wiley Online Library.
- Wilby, R.L., Wigley, T., Conway, D., Jones, P., Hewitson, B., Main, J., Wilks, D., 1998. Statistical downscaling of general circulation model output: a comparison of methods. *Water Resour. Res.* 34 (11), 2995–3008.
- Wood, A.W., Maurer, E.P., Kumar, A., Lettenmaier, D.P., 2002. Long-range experimental hydrologic forecasting for the eastern United States. *J. Geophys. Res.* (1984–2012) 107 (D20) ACL 6-1-ACL 6-15.
- Wood, A.W., Leung, L.R., Sridhar, V., Lettenmaier, D., 2004. Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs. *Clim. Change* 62 (1–3), 189–216.
- Yue, S., Pilon, P., Cavadias, G., 2002. Power of the Mann-Kendall and Spearman's rho tests for detecting monotonic trends in hydrological series. *J. Hydrol.* 259 (1), 254–271.