Performance of quantile-quantile bias-correction for use in hydroclimatological projections

J.C. Bennett[a,b], **M.R. Grose**[c], **D.A. Post**[d], **F.L.N. Ling**[a], **S.P. Corney**[c] **and N.L. Bindoff**[b,c,e]

[a]*Entura, Cambridge Park, Tasmania.*
[b]*Institute of Marine and Antarctic Studies, University of Tasmania, Sandy Bay, Tasmania.*
[c]*Antarctic Climate and Ecosystems CRC, University of Tasmania, Sandy Bay, Tasmania.*
[d]*CSIRO Land and Water, Black Mountain, ACT.*
[e]*Centre for Australian Climate and Weather Research, CSIRO Marine and Atmospheric Research, Hobart, Tasmania.*
Email: *James.Bennett@entura.com.au*

Abstract: Quantile-quantile bias-correction has been used in several northern hemisphere studies to improve the utility of regional climate model (RCM) outputs, however the technique is rarely used in Australia. The technique has the advantage of preserving complex changes in the RCM projections – e.g. to weather systems, dry spells, rainfall intensities, mean rainfalls, rainfall extremes – in hydrological modelling. The technique also has the advantage of retaining the physical correlation between rainfall and evapotranspiration.

We apply a quantile-quantile bias-correction to an ensemble of six fine-scale (10 km) regional climate simulations for 1961-2100 over Tasmania generated by the Climate Futures for Tasmania project (http://www.acecrc.org.au/Research/Climate%20Futures). The regional climate simulations show a high degree of skill in replicating spatial patterns of mean annual rainfall (spatial correlation R^2=0.75) and require relatively modest bias-corrections. Multiplicative bias-corrections are calculated for daily values of rainfall and evapotranspiration for the calibration period 1961-2007. Bias-corrections are calculated independently for each grid cell, for each season and for each percentile. The bias-correction substantially improves spatial correlation between modelled and observed seasonal and annual rainfall (spatial correlation R^2>0.99). We use split sample cross-validation to find that the projected changes to rainfall are insensitive to the period chosen to train the bias-correction.

Spatial relationships of daily rainfalls are not explicitly accounted for by the bias-correction. To test the simulation of spatial relationships of rainfall we aggregate rainfalls to the seven Tasmanian Bureau of Meteorology forecast zones. The behaviour of biases of aggregated rainfall for each zone is compared to the behaviour of biases in individual grid cells within that zone. These comparisons indicate that the bias-corrected RCM outputs tend to overestimate large regional rainfall events and underestimate small regional rainfall events.

We test the performance of the bias-correction by using bias-corrected rainfall and evapotranspiration as inputs to five hydrological models (AWBM, IHACRES, Sacramento, SIMHYD, SMAR-G). Performance of the hydrological models is assessed at 86 flow gauges across Tasmania. Performance of hydrological models with bias-corrected RCM inputs varies between models. AWBM, Sacramento, SIMHYD and SMAR-G simulate runoff realistically with bias-corrected RCM inputs, while IHACRES does not. The SIMHYD model gives the most realistic results (median bias -3%) while the IHACRES model gives the poorest results (median bias -21%). Our study supports the findings of northern-hemisphere studies that quantile-quantile bias-correction can effectively couple RCM outputs to hydrological models. Historical Tasmanian stream flows are realistically simulated with bias-corrected RCM outputs, while the climate change signal is successfully retained.

Keywords: *Bias-correction, Regional Climate Modelling,, Hydrological Modelling, Climate Futures for Tasmania*

1. INTRODUCTION

It is well established that regional climate models (RCMs) do not produce outputs that are sufficiently accurate for direct input into hydrological modelling (Wood *et al.* 2004). This problem has been solved by applying various bias-correction techniques to RCM outputs to make them suitable for hydroclimatological studies. Simple perturbation of observed rainfall and evaporation has been the most popular of these techniques in Australia (e.g. Chiew et al. (2009)), However, as Fowler and Kilsby (2007) point out, simple perturbation does not account for changes to rainfall variability or for changes in the sequences of wet and dry days, even though these are likely to have significant impacts on water availability. Quantile-quantile bias-correction removes biases across the entire frequency distribution of a given variable, and has been shown to be highly effective at removing biases from climate model outputs while retaining changes to rainfall frequency and variability (Ines and Hansen 2006). Quantile-quantile bias-correction has been shown to be effective for coupling regional climate model (RCM) output directly to hydrological models (Fowler and Kilsby 2007). This method has the considerable advantage of transferring the complex suite of rainfall changes projected by RCMs to hydrological models, including changes to rainfall distributions, rainfall characteristics at short temporal scales (e.g. maximum daily precipitation, autocorrelation of rain days) and longer time scales (annual and seasonal). This allows more meaningful assessments of future runoff variability.

Quantile-quantile bias-correction has been successfully employed to couple RCMs to hydrological models in several northern hemisphere studies (e.g. Wood et al. (2004), Fowler and Kilsby (2007)). This paper tests the suitability of this method in the Australian context by assessing its performance over Tasmania.

2. METHODS

2.1. Regional climate modelling

We use the six regional climate simulations produced by the Climate Futures for Tasmania project (http://www.acecrc.org.au/Research/Climate%20Futures) covering the period 1961-2100. Climate simulations are produced by downscaling global climate models (GCMs) with the CSIRO conformal cubic atmospheric model (CCAM) (McGregor and Dix 2008). CCAM is a global atmospheric model that uses a stretched grid to increase the resolution over the region of interest. Unlike limited area models, CCAM has no lateral boundaries. For this study CCAM is configured to be forced only by GCM sea surface temperatures (SSTs) and sea ice concentration. Six GCMs from the coupled model intercomparison project (CMIP-3) are downscaled under the SRES A2 emissions scenario: CSIRO-Mk3.5, ECHAM5/MPI-OM, GFDL-CM2.0, GFDL-CM2.1, MIROC3.2(medres) and UKMO-HadCM3. Biases inherent in the GCM SSTs are removed using a simple additive bias-correction before downscaling (Katzfey *et al.* 2009). The downscaling is carried out in two stages. The first stage is forced only using the bias-corrected GCM SSTs and sea-ice concentration and achieves a horizontal resolution of 0.5° (~50 km) over Australia. The second stage is forced using the same bias-corrected GCM SSTs and sea-ice concentration along with spectral nudging of the atmosphere from the corresponding 0.5° simulations, achieving a horizontal resolution of 0.1° (~10 km) over Tasmania.

2.2. Quantile-quantile bias-correction

We use a multiplicative quantile-quantile bias-correction to align simulated daily rainfalls and areal potential evapotranspiration (APET) to 0.1° (~10 km) gridded interpolated observations. In this paper, we focus on the results for rainfall. Gridded observations are aggregated from the 0.05° (~5 km) gridded SILO dataset (Jeffrey *et al.* 2001).

Unlike many other studies using quantile-quantile bias-correction (e.g. Ines and Hansen et al. (2006)) we do not assume a shape for the frequency distribution of rainfall or APET. Rather, we calculate correction factors at each half percentile from 0.5 to 99.5 for the entire distribution (including days of zero rain). Percentiles are calculated using linear interpolation between closest ranks. Correction factors are calculated independently for each season (DJF, MAM, JJA, SON) and at each grid cell for the calibration period 1961-2007. Our method has the advantage of being entirely predicated on data, rather than on assumed frequency distributions. Our method does not explicitly align the number of rain days in observed and bias-corrected time series in the manner of Ines and Hansen et al. (2006). We force any rain day with rainfall of less than 0.2 mm to zero in both observed and modelled rain time series. The threshold of 0.2 mm is chosen because it is the lower resolution limit of the Bureau of Meteorology (BoM) rain gauges used as the basis of the

SILO dataset. Rainfalls of <0.2 mm per day in the observed rain timeseries (and, by extension, in the bias-corrected rain timeseries) are more precise than the accuracy of gauged rainfalls allows. The entire simulation (1961-2100) is detrended with a 30-year moving average to remove any long-term changes in rainfall regimes. Each day from this detrended series is consigned to a percentile 'bin' between integer percentiles (i.e. percentile bins of 0-1, 1-2, …, 98-99, 99-100). The bias-correction is then applied to the entire uncorrected simulation based on these bins: the factor calculated for the 0.5[th] percentile is applied to the 0-1 percentile bin, the factor for the 1.5[th] percentile is matched to the 1-2 percentile bin, and so on. Finally, bias-corrected CCAM outputs are regridded from the 0.1° (~10 km) CCAM grid to a 0.05° (~5 km) grid to be compatible with the hydrological models.

2.3. Testing performance of the bias-corrected RCM outputs

The bias correction is evaluated using split-sample cross-validation. We compare two cases of split-sample cross-validation, in addition to the bias-correction calculated for the entire calibration period, which we call the *47-year* case (Table 1).

The bias-correction corrects biases only at individual grid cells. It takes no account of the spatial correlation of rain storms across catchments. We assess the ability of the bias-corrected RCM outputs to replicate the spatial characteristics of rainfall by aggregating modelled and SILO rainfalls to the Bureau of Meteorology (BoM) forecast zones (Fig 1). Biases of aggregated rainfall in a given zone are compared to the biases of all individual grid cells within that zone.

Table 1 Periods used to train and test bias-correction

Name	Calibration Period	Validation Period
47-year case	1961-2007	-
Case 1	1962-1984	1985-2007
Case 2	Odd years 1961-2005 (1961,1963,… ,2003,2005)	Even years 1962-2006 (1962,1964,… ,2002,2004)

2.4. Hydrological modelling

We use the five rainfall-runoff models calibrated by Viney et al. (2009) to 90 stream flow records for 1975-2007 for catchments around Tasmania: AWBM (Boughton 2004), IHACRES (Post and Jakeman 1999), Sacramento (Burnash *et al.* 1973), SIMHYD (Chiew *et al.* 2002) with Muskingum routing (Tan *et al.* 2005), and SMAR-G (Goswami *et al.* 2002). These models produce runoff distributed on a 0.05° grid covering all of Tasmania. Stream records were from catchments that had negligible human influence on flows. To achieve Tasmania-wide coverage with the five runoff models, Viney et al. (2009)

Districts
1. North East & Flinders Island
2. East Coast
3. Central North & Midlands
4. South East Huon & Channel & Lower Derwent Valley
5. Central Plateau & Upper Derwent Valley West & South Coast & Highlands
6. Northwest Coast & King Island

Figure 1 Bureau of Meteorology (BoM) forecast zones for Tasmania. Source: **www.bom.gov.au**

assigned model parameters to ungauged catchments from their nearest gauged neighbour. We use the bias-corrected CCAM rainfall and APET outputs as direct inputs to the rainfall-runoff models.
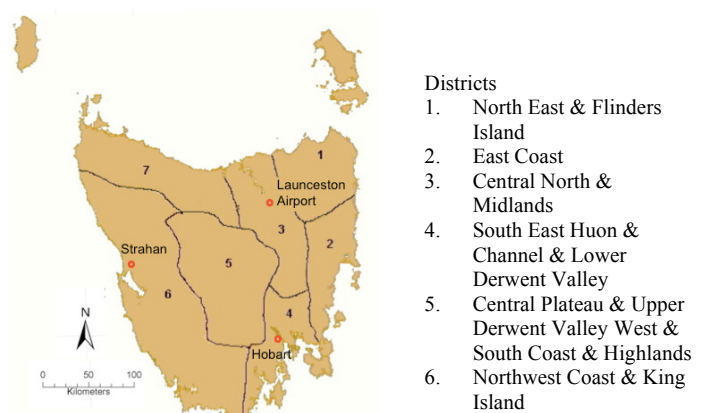
3. RESULTS

3.1. Ability of quantile-quantile bias-correction to remove a range of biases

All results presented here are calculated from averages of the six climate simulations. The bias-correction proved very successful at correcting a variety of rainfall characteristics (Table 2). As expected, spatial correlations of simulated rainfall with observations for mean annual and seasonal rainfall become near

perfectly correlated with observations, as does the variance of daily rainfall (represented as the coefficient of variation, CV) (Table 2). The spatial correlation of the number of raindays is substantially improved by

the bias-correction. The CCAM simulations tend to have more rain days than observations. The additional rain days in the modelled timeseries are forced to no-rain days by the bias-correction, resulting in the close match between the number of observed and bias-corrected CCAM raindays. Interestingly, the variance in annual rainfall (CV) is also improved in the drier east of Tasmania (Table 2, Fig 2).

Table 2 Spatial correlations of simulated rainfall with SILO gridded observations. CCAM statistics are the mean of six simulations.

Spatial Correlations (R^2)		CCAM and SILO	Bias-corrected CCAM and SILO
Mean Rainfall (mm)	DJF	0.64	0.9994
	JJA	0.76	0.9998
	Annual	0.75	>0.9999
Coefficient of Variation (CV)	Annual CV	0.60	0.76
	Daily CV	0.79	0.99
Raindays (> 1 mm)	DJF	0.26	0.9998
	JJA	0.26	0.9998
	Annual	0.27	>0.9999

Rainfall in this region is more sporadic and less strongly seasonal than over wetter western Tasmania. As a consequence, correcting daily variance has the effect of partly correcting annual variance in the east (Fig 2).
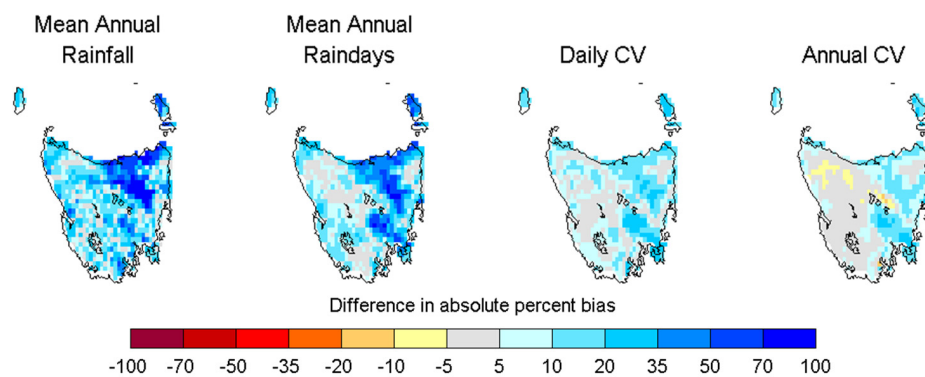

Figure 2 Absolute biases of uncorrected CCAM outputs minus absolute biases of bias-corrected CCAM-outputs. Maps (l-r) show improvement in biases in mean annual rainfall, mean annual raindays > 1 mm, daily coefficient of variation (CV) and annual CV. CCAM statistics are the mean of the six simulations.

Biases in mean annual rainfall and raindays and daily and annual CV are effectively reduced by the bias-correction. Fig 2 shows the difference in absolute biases between uncorrected RCM rainfalls and bias-corrected RCM rainfalls. In almost all regions, biases in mean annual rainfall and raindays are reduced. Representation of daily CV is improved over the east coast, resulting in improved annual CV in this area. There are slight increases in biases in annual CV in the north-west of Tasmania.
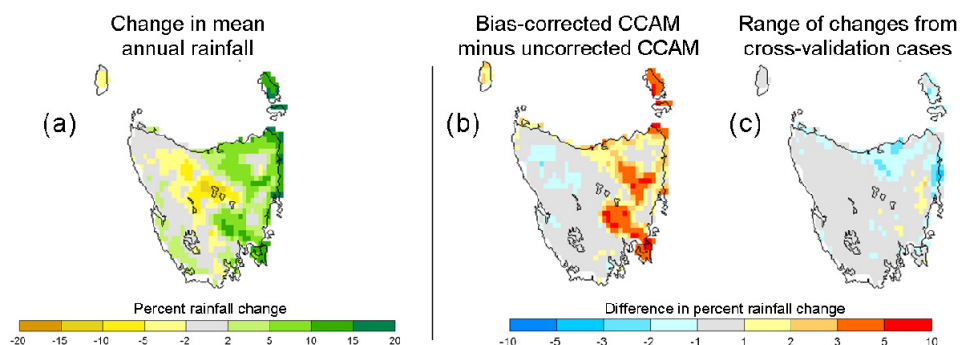

Fig 3 Effects of bias-correction on projected future rainfall changes. (a) shows changes to bias-corrected rainfall from 1961-1990 to 2070-2099; (b) shows percent change of rainfall from bias-corrected CCAM outputs minus percent rainfall change from uncorrected CCAM outputs; (c) shows range of percent changes from the three cross-validation cases (Table 1). CCAM changes are the mean of the six CCAM simulations.

3.2.   Comparing bias-corrected and uncorrected climate simulations

Rainfall changes projected with uncorrected CCAM outputs differ in magnitude from changes projected with bias-corrected CCAM outputs (Fig 3a) in some parts of Tasmania. Uncorrected projections for mean annual rainfall from the mean of the six simulations are as much as 8% drier in the south-east of Tasmania than bias-corrected projections (Fig 3b). Uncorrected and bias-corrected projections always agree on the sign of changes (not shown).
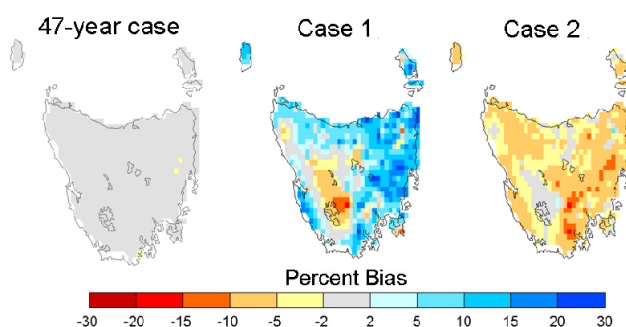

Fig 4 Performance of bias-correction under cross-validation. Left map shows biases for mean annual rainfall of 47-year case for the calibration period (1961-2007) compared to observations, centre and right maps show biases of Case 1 and Case 2 compared to validation periods (Table 1). CCAM statistics are the mean of the six simulations.

3.3.   Split sample cross-validation

Biases in mean annual rainfall of the 47-year case are virtually eliminated for the calibration period (1961-2007) (Fig 4). Cross-validation Cases 1 and 2 show that biases increase in the order of ±10% when the bias-correction is validated against a period independent of the calibration period (Fig 4), indicating that the bias-correction is somewhat sensitive to the period chosen for calibration. Despite this, the choice of calibration period has almost no effect on the magnitude of projected changes. The range of changes in mean annual rainfall from 1961-1990 to 2070-2099 calculated from Case 1, Case 2 and the 47-year case are shown in Fig 3c. No matter which case is used, the changes do not vary by more than ±1% for almost all of Tasmania, and all differences are smaller than 2%.

3.4.   Characteristics of rainfall bias over regions

Rainfall biases for total rainfall in the BoM forecast zones (Fig 1) and biases for the individual grid cells contained in each BoM forecast zone are given in Fig 5 for larger rainfalls (70th–100th percentile) using outputs from CCAM forced by the CSIRO-Mk3.5 GCM. The behaviour of spatial biases is very similar for all climate simulations (not shown). We concentrate on larger rainfalls because a higher proportion of rainfall is converted to runoff in large rainfall events. Biases in larger rainfall events could magnify biases in runoff. The bias-corrected outputs tend to underestimate regional rainfall totals for lower rainfall events and over-estimate regional rainfall totals for higher rainfall events (Fig 5). This contrasts the behaviour of individual cells within regions, which have biases that are approximately normally distributed around zero irrespective of event size (Fig 5), excepting when either modelled or observed rainfalls are zero (not shown). Biases at individual grid cells are not perfectly eliminated by the bias-correction because the bias correction is applied to the length of the simulation (1961-2100) rather than only to the calibration period (1961-2007).

3.5.   Hydrological modelling

Fig 6a shows frequency distributions of biases at 86 gauge sites in Tasmania for the five runoff models when forced by CCAM and SILO. AWBM, SIMHYD, Sacramento and SMAR-G showed similar median biases when forced by CCAM as when forced by SILO. Overall, biases are larger for runoff forced by CCAM. Interquartile ranges of biases in mean flows increase for all hydrological models forced by CCAM. SIMHYD and AWBM showed the least change in median bias and the smallest increase in interquantile range of biases. The IHACRES model performs poorly when forced with CCAM in relation to the other hydrological models, with a significant decline in median bias and a much greater interquartile range.

All hydrological models display a tendency to predict lower flow volumes when forced by CCAM (Fig 6a). This is highlighted by looking at the biases in the high flows (Q_{95}) (Fig 6b) of hydrological models forced by CCAM when compared to flows modelled by forcing hydrological models with SILO at 86 gauge sites. The CCAM-forced hydrological models produced lower Q_{95} flows than SILO-forced hydrological models at more than 70% of catchments.
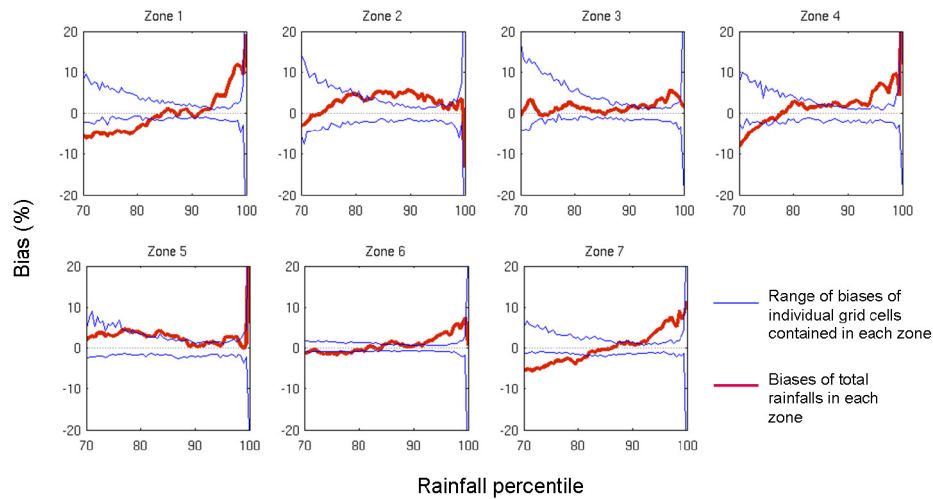
Figure 5 Spatial biases of bias-corrected rainfalls produced by CCAM forced by CSIRO-Mk3.5 for the period 1961-2007. Red lines show biases of total rainfall over BoM forecast zones (see Fig 1) of bias-corrected CCAM rainfalls compared to SILO. Blue lines describe the range of biases, here defined as the 5th and 95th percentile of biases, at all individual grid cells contained within each forecast zone.
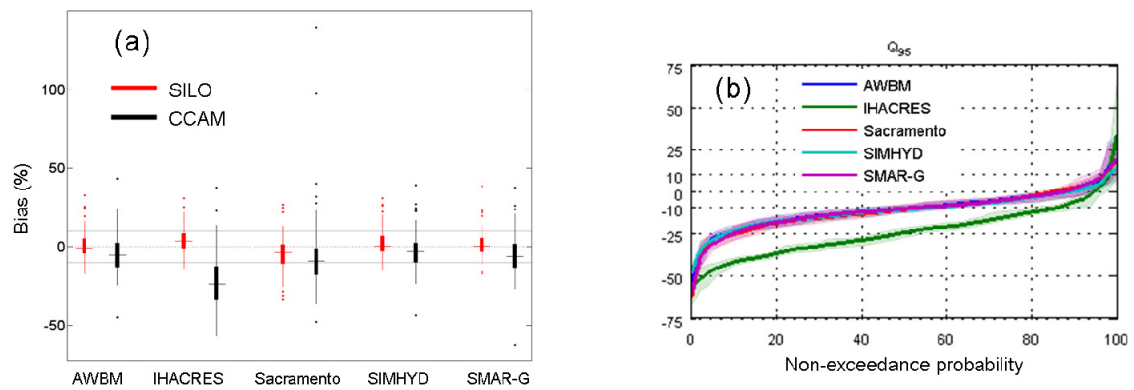

Fig 6 (a) Box plots of biases for hydrological models forced by SILO inputs (red) and by CCAM inputs (black) at 86 catchments calculated from the average of the six climate simulations; boxes indicate 25th, 50th and 75th percentiles, whiskers are two standard deviations from the mean, points are outliers, grey lines show biases of ±10%. (b) Cumulative probability plot of biases in high flows (Q_{95}) forced by CCAM calculated against Q_{95} flows forced by SILO; shaded areas show range of the six climate simulations.

4. DISCUSSION AND CONCLUSIONS

The quantile-quantile bias-correction method used in this study proved effective at reducing a range of biases from regional climate simulations of rainfall over Tasmania, including mean annual and seasonal rainfall and the number of raindays. It also reduced the bias in interannual variability in low rainfall regions. Cross-validation shows that the performance of quantile-quantile bias-correction declines in independent validation periods, however the validation biases of ± 10% are still small compared to the biases of > 50% for the uncorrected model output (Fig 2). Further, projected future changes to mean rainfall are insensitive to the period chosen to calibrate the bias-correction, indicating that the choice of training period is not a crucial consideration when assessing future change. There was some discrepancy in changes projected by uncorrected model outputs and bias-corrected regional outputs. These tended to be in the drier, eastern parts of Tasmania, particularly in areas that showed marked future increases in rainfall. The causes of the discrepancy are unclear.

Biases in total rainfall over regions behaved differently from biases at individual grid cells. The tendency of the bias-corrected CCAM rainfalls to over-predict large regional rainfall events in all regions of Tasmania could be expected to lead to an even larger positive bias in runoff. However, runoff models forced by bias-corrected CCAM inputs tended to underpredict flows modelled with the same hydrological models forced

by SILO. We speculate that this may be due to poor representation of the temporal characteristics of rainfall events, either by CCAM or by the bias-correction. More investigation is required to determine the causes of the spatial biases and the tendency of hydrological models to under-predict flows when forced by CCAM.

We have shown that quantile-quantile bias-correction can effectively couple regional climate simulations to hydrological models in the Australian context. The method produces realistic river flows for a number of hydrological models and catchments, particularly when using the SIMHYD and AWBM models. IHACRES proved highly sensitive to changes in inputs, a finding echoed by Viney et al. (2009), who found that IHACRES performed poorly under spatial cross-validation in Tasmania. Care should be taken to test the suitability of hydrological models for direct coupling to RCMs if using the methods described here.

A considerable amount of effort has been invested in recent years in Australia to build complicated sets of hydrological models for basin-scale hydroclimatological studies based on GCM inputs (e.g. Chiew et al (2009)). We have shown that the potential exists to use quantile-quantile bias-correction to update these studies with higher resolution climate projections when these projections become available.

ACKNOWLEDGMENTS

REFERENCES

Boughton WC (2004) The Australian water balance model. *Environmental Modelling & Software* **19**, 943-956.

Burnash RJC, Ferral RL, McGuire RA (1973) A generalized streamflow simulation system-conceptual modeling for digital computers. Joint Federal and State River Forecast Center, Sacramento.

Chiew FHS, Peel MC, Western AW (2002) Application and testing of the simple rainfall-runoff model SIMHYD. In 'Mathematical models of small watershed hydrology and applications'. (Eds VP Singh, DK Frevert) pp. 335–367. (Water Resources Publications: Littleton, USA).

Chiew FHS, Teng J, Vaze J, Post DA, Perraud JM, Kirono DGC, Viney NR (2009) Estimating climate change impact on runoff across southeast Australia: Method, results and implications of the modeling method. *Water Resources Research* **45**, W10414.

Fowler HJ, Kilsby CG (2007) Using regional climate model data to simulate historical and future river flows in northwest England. *Climatic Change* **80**, 337-367.

Goswami M, O'Connor KM, Shamseldin AY (2002) Structures and performances of five rainfall-runoff models for continuous river-flow simulation. In '1st Biennial Meeting of International Environmental Modeling and Software Society'. Lugano, Switzerland.

Ines AVM, Hansen JW (2006) Bias correction of daily GCM rainfall for crop simulation studies. *Agricultural and Forest Meteorology* **138**, 44-53.

Jeffrey SJ, Carter JO, Moodie KM, Beswick AR (2001) Using spatial interpolation to construct a comprehensive archive of Australian climate data. *Environmental Modelling and Software* **16**, 22.

Katzfey JJ, McGregor J, Nguyen KC, Thatcher M (2009) Dynamical downscaling techniques: impacts on regional climate change signals. In '18th World IMACS Congress and MODSIM09 International Congress on Modelling and Simulation'. Cairns. http://www.mssanz.org.au/modsim09

McGregor JL, Dix MR (2008) An updated description of the conformal-cubic atmospheric model. In 'High Resolution Numerical Modelling of the Atmosphere and Ocean'. (Eds K Hamilton, W Ohfuchi) pp. 51-76. (Springer: New York).

Post DA, Jakeman AJ (1999) Predicting the daily streamflow of ungauged catchments in SE Australia by regionalising the parameters of a lumped conceptual rainfall–runoff model. *Ecological Modeling* **123**, 91–104.

Tan KS, Chiew FHS, Grayson, R.B., Scanlon PJ, Siriwardena L (2005) Calibration of a daily rainfall-runoff model to estimate high daily flows. In 'Congress on Modelling and Simulation (MODSIM 2005)'. Melbourne, Australia. http://www.mssanz.org.au/modsim05/index.htm

Viney NR, Post DA, Yang A, Willis M, Robinson KA, Bennett JC, Ling FLN, Marvanek S (2009) Rainfall-runoff modelling for Tasmania. CSIRO Water for a Healthy Country Flagship, Australia. http://www.csiro.au/science/TASSY.html

Wood A, Leung LR, Sridhar V, Lettenmaier DP (2004) Hydrologic implications of dynamical and statistical approaches to downscaling climate outputs. *Climate Change* **62**, 189-216.