

An improved statistical analogue downscaling procedure for seasonal precipitation forecast

Quanxi Shao · Ming Li

Published online: 28 June 2012

© Crown Copyright as represented by: Simon Barry 2012

Abstract Seasonal forecasting can be highly valuable for water resources management. Hydrological models (either lumped conceptual rainfall-runoff models or physically based distributed models) can be used to simulate streamflows and update catchment conditions (e.g. soil moisture status) using rainfall records and other catchment data. However, in order to use any hydrological model for skillful seasonal forecasting, rainfall forecast at relevant spatial and/or temporal scales is required. Together with downscaling, general circulation models are probably the only tools for making such seasonal predictions. The Predictive Ocean Atmosphere Model for Australia (POAMA) is a state-of-the-art seasonal climate forecast system developed by the Australian Bureau of Meteorology. Based on the preliminary assessment on the performance of existing statistical downscaling methods used in Australia, this paper is devoted to develop an analogue downscaling method by modifying the Euclidian distance in the selection of similar weather pattern. Such a modification consists of multivariate Box–Cox transformation and then standardization to make the resulted POAMA and observed climate pattern more similar. For the predictors used in Timbal and Fernandez (CAWCR Technical Report No. 004, 2008), we also considered whether the POAMA precipitation provides useful information in the analogue method. Using the high quality station data in the Murray Darling Basin of Australia, we found that the modified analogue method has potential to improve the seasonal precipitation forecast using POAMA outputs. Finally, we found that in the analogue method, the precipitation from POAMA

should not be used in the calculation of similarity. The findings would then help to improve the seasonal forecast of streamflows in Australia.

Keywords Bias correction · POAMA · Quantile–quantile (Q–Q) transformation · Standardization · Statistical downscaling

1 Introduction

Skillful seasonal forecasting (with a lead time of one to several months) of streamflow can be highly valuable for water resources management, especially in agriculture and water allocation. Hydrological models (either lumped conceptual rainfall-runoff models or physically based distributed models) can be used to simulate streamflows and update catchment conditions (e.g. soil moisture status) using rainfall records and other catchment data. However, in order to predict future streamflow using any hydrological model, rainfall forecast at relevant spatial and/or temporal scales is required. Together with downscaling, general circulation models (GCMs) are probably the only tools for making such seasonal predictions. POAMA (The Predictive Ocean Atmosphere Model for Australia) is a state-of-the-art seasonal climate forecast system developed by the Australian Bureau of Meteorology (the bureau) based on a coupled ocean/atmosphere model and ocean/atmosphere/land observation assimilation systems (Alves et al. 2003). It produces a set of ensembles for climatic variables based on different initial atmospheric conditions. However, like other GCMs, POAMA produces only large spatial scale predictions over Australia, which cannot be used directly as inputs for catchment scale hydrological models. Such predictions do not capture the large degree of

Q. Shao (✉) · M. Li
CSIRO Mathematics, Informatics and Statistics,
Private Bag 5, Wembley, WA 6913, Australia
e-mail: Quanxi.shao@csiro.au

spatial variability in rainfall. Robust downscaling of the large scale rainfall predictions from POAMA to rainfall at a catchment or point scale is needed to drive the hydrological models for streamflow forecasting.

The term “downscaling” has been adopted in literature to describe a set of techniques that relate the local and regional scale climate variables to the large scale atmospheric and oceanic forcing variables. It bridges the gap between what the current GCMs are able to provide and the information users require. There are two general categories of downscaling: dynamic downscaling and statistical downscaling. Dynamic downscaling uses nested regional climate models with intensive computing. In contrast, statistical downscaling establishes transfer functions between different scales and can easily be implemented in practice. A general review of different statistical downscaling methods is given in Wilby and Wigley (1997). Incorporated with POAMA outputs, Shao et al. (2010) recommended the analogue method proposed by Timbal et al. (2008) (hereafter, called Timbal’s analogue method for short) as the preferred statistical downscaling method for seasonal forecasting of precipitation in Australia from a comparison with the nonhomogenous hidden Markov model (NHMM) (Hughes et al. 1999; Charles et al. 1999) and the Generalised Linear Modelling of Daily Climate Sequences (GLIMCLIM) (Chandler and Wheatler 2002).

One of main error sources in the analogue method is the mismatch of POAMA outputs and real observations, which makes the weather patterns identified by two datasets different especially in the case of seasonal forecasting. As a result, the downscaled local variable based on the POAMA outputs may not be as accurate as we expect. In this study, we explore such a mismatch in terms of bias and dispersion and further develop an improvement of Timbal’s analogue method to address the bias correction of POAMA outputs. An application to the south Murray Darling (SMD) region is assessed using precipitations at 35 high quality stations.

The paper is organized as following. Section 2 describe the data used, followed by methodology development. The results are presented in Sect. 4. The paper is concluded by conclusions and discussion.

2 Data

POAMA (The Predictive Ocean Atmosphere Model for Australia; see <http://poama.bom.gov.au/>), is developed by the Australian Bureau of Meteorology (the bureau). POAMA 1.5 is the version used during our development, and is available to public. POAMA hindcast (1980–2006) is where POAMA was run for a past period of time with actual (or as near as known) forcing and realistic starting conditions. This allows researchers to see how well that model performs. We

used the POAMA grids latitude = $[-25.07, -40.96] \times$ longitude = $[140, 155]$, which covers the Southern part of the Murray Darling basin (SMD) and contains 7×7 POAMA grids (Fig. 1). The grids were optimized for the SMD region when the bureau’s analogue downscaling method was developed (Timbal et al. 2008) and was also used for downscaling POAMA (Charles et al. 2010). The grids are named (1, 1) to (7, 7) with (i, j) representing the grid located at the i th node from west and j th node from north. We used the NCEP/NCAR reanalysis as observations. The NCEP/NACR reanalysis is regridded by interpolation to match the POAMA grids. The available NCEP/NACR reanalysis data used are from 1958 to 2006. Three variables are examined: mean sea level pressure (MSLP); v-wind at 850 hPa (V850) and precipitation (PRCP), which are used in the analogue downscaling method for this region (Timbal et al. 2008). In this study, the local weather is referred to the 24-h rainfall total between 9 am local time, the calendar window is chosen to be 30 days, and the reference period is from 1958 to 2006.

To assess the performance of our method, we applied it to 35 high quality stations in the Murray Darling Basin of Australia (see Fig. 1) which are of best interest for Australian water resources management (CSIRO 2008) and carefully selected by our projects team for testing the impact on streamflow via dynamic hydrological models (Wang et al. 2011). The daily rainfall data are from 1980 to 2000. Three lead times are considered, including 1, 2 and 3 months.

3 Methodology

3.1 A review of analogue downscaling method

An analogue in a given day is the single and best day selected from the historical pool which has the closest weather pattern. The similarity between two weather

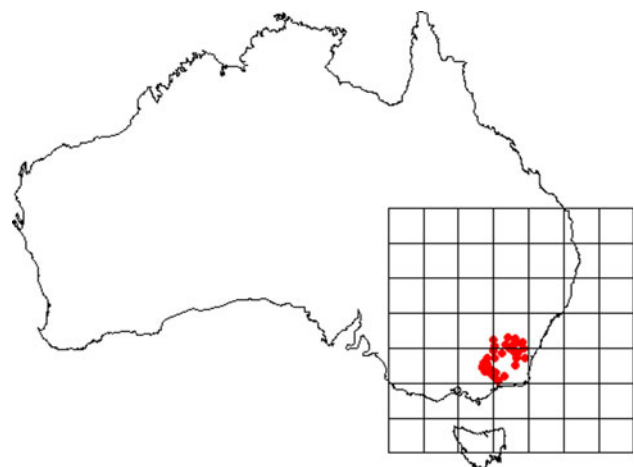


Fig. 1 POAMA grids used for downscaling and the high quality stations used for assessment

patterns is defined by the Euclidean metric of a set of atmospheric predictors. Such an idea is based on the assumptions that the simultaneously observed local weather is associated to the large-scale weather pattern and that such weather patterns dominate at least the occurrence and to certain extent the amount of precipitation. The analogous weather pattern is based on the search of the analogous vector formed by large scale climate variables. The feasibility of this idea is investigated first by Lorenz (1969) and explored further by many researchers (Kruizinga and Murphy 1983; Zorita et al. 1995; Zorita and von Storch 1999; Timbal et al. 2008).

Let the weather pattern be characterized by a vector of large-scale predictors $X = (x_1, \dots, x_m)'$, which has a series of historical observed values $\{X_1^O, X_2^O, \dots, X_n^O\}$ with $X_i^O = (x_{i,1}^O, \dots, x_{i,m}^O)'$. The large-scale predictors are usually a collection of climate variables at regular GCM grids in a large region. The large region does not have to be a regular rectangular shape (as we used in here) but needs to be sufficiently large so that the weather pattern can be identified.

For a target time t , a vector of model-derived predictors (e.g. POAMA outputs) $X_t^M = (x_{t,1}^M, \dots, x_{t,m}^M)'$ is used for forecasting. The dissimilarity between the model-derived predictor vector X_t^M at the target time t and the observed predictor vector X_i at the historical time i is measured by the Euclidean distance as

$$d(X_t^M, X_i^O) = \sum_{k=1}^m (x_{t,k}^M - x_{i,k}^O)^2 \tag{1}$$

The best analogue time $a(t)$ at time t is defined as the day in the historical pool with minimum dissimilarity,

$$a(t) = \arg \min_{i=1, \dots, n} d(X_t^M, X_i^O) \tag{2}$$

The predictor vector can be the grid values of selected variables in a region which covers the area of interest. The size of region needs to be large enough so that the weather pattern is representative. Timbal et al. (2008) used the standardization for each grid by applying the long term monthly mean and standardization (removing the mean and then dividing by the standardization). With the consideration of seasonality, the search of analogue weather pattern is excised within a calendar window spanning a calibrated period of months. As suggested by Timbal et al. (2008), searching analogues is only carried out in the same season over years.

3.2 A review of existing bias correction method

The present statistical downscaling excises calibrate the models to historical data (e.g. NCEP/NCAR reanalysis data from the National Oceanic and Atmospheric Administration

(NOAA)) and then the calibrated models are applied to GCMs (e.g. POAMA) for long term climate prediction. That is, the calibration is independent of the GCM output. While the GCM's climatology can be used to simulate long term climate changes, it is not directly applicable to seasonal forecasting. It is well known that GCM precipitation output cannot be used to force hydrological or other impact models without some form of downscaling that includes bias correction if realistic output is sought (Sharma et al. 2007; Hansen et al. 2006; Feddersen and Andersen 2005). In general, each GCM (including POAMA) has its own climatology which can be different from the observed climatology that drives historical data (e.g. NCEP reanalysis). Typically, (a) GCM output exhibits less variability than the NCEP reanalysis; (b) The average of GCM values can be significantly different from that of observed values. Such a mismatch between GCM and historical data should be addressed before statistical analogue method is used.

Standardization (Timbal et al. 2008; Charles et al. 2010) tends to match the first two orders of distributional statistics by shifting (i.e. addition) the GCM outputs by the difference of the average NCEP/NCAR and GCM values and scaling (i.e. multiplication) the GCM outputs by the ratio of the standard deviations of NCEP and GCM values. Let μ_O (or μ_M) and σ_O (or σ_M) be the mean and standard deviation of historical observed data (or GCM outputs), respectively. (Here the index of individual predictor variable is ignored for simplicity.) Standardization transforms a particular GCM data x^M as follows

$$S(x^M) = \frac{\sigma_O}{\sigma_M} (x^M - \mu_M) + \mu_O \tag{3}$$

It is easily shown that the standardized GCM data have the same mean and variance as the historical data. Standardization is essentially a linear transformation and the correlation between NCEP/NCAR and GCM remains same after transformation. Standardization normalises GCM outputs to match mean and standard deviation with observations. If the distribution of a predictor does not follow a normal distribution, the correction of mean and standard deviation may not ensure the distribution of transformed GCM outputs close to the observed one. Standardization, therefore, is more appropriate for normal distributed data than heavily skewed data.

Alternatively, in order to obtain a better match between the distributions of GCM output and historical data, quantile–quantile (Q–Q) transformation provides a perfect match in term of distribution by assigning the GCM percentiles associated output values to the corresponding percentiles associated NCEP outputs. Let F_O (or F_M) be the distribution of historical observed data (or GCM output), and quantile–quantile transformation can be simply represented by

$$Q(x^M) = F_o^{-1}F_M(x^M) \quad (4)$$

for a particular GCM data. In practice, theoretical distributions are replaced by the empirical distribution estimated from historical observations. The Q–Q transformed POAMA has the exactly same distribution to NCEP in the period of interest, and therefore all distributional characteristics, such as mean and variance, are preserved. Unlike standardization, the Q–Q transformation is generally non-linear and therefore may lose the correlation between GCM and NCEP.

The Q–Q transformation has been used widely as a post bias correction for the GCM output to translate long-range climate predictions to the realm of hydrology in hydrological modeling (Wood et al. 2002, 2005; Luo and Wood 2008; Ines and Hansen 2006; Sharma et al. 2007; Piani et al. 2010). The Q–Q transformation has some variations including empirical Q–Q transformation, distribution-based Q–Q transformation and a mix of empirical and distribution-based transformation.

The Q–Q transformation ensures a perfect match (or nearly perfect in the parametric Q–Q transformations) between the NCEP and GCM data for the period used for calibration. The disadvantage of the Q–Q transformation is two-fold: (1) The Q–Q transformation is dependence on time scale of available data. A Q–Q transformation based on daily data has no relationship with the transformation established on the same set of data but aggregated to monthly totals. In other words, the Q–Q transformation does not own invariant property at different times scales. (2) The Q–Q transformation is often a non-linear transformation and loses the correlation of original large-scale predictors and a local-scale variable.

Although the Q–Q transformation is a monotonic transformation, it could not control the change of correlation after the transformation because of the nonlinearity in the transformation. However, if the Q–Q transformation is linear, it will reserve the correlation. Therefore, it will encourage us to find a distribution which gives a linear Q–Q transformation. An obvious one is normal distribution.

Two different time scales are used in bias correction: daily and monthly scales. However, the bias corrected results can be different for different time scales, making it difficult to be interpreted. From a statistical point of view, the standardization only considers mean and standard deviation and therefore normality on data is implicitly assumed in order to have an invariant property at different time scales.

3.3 Two-step bias correction

Based on the review of the existing bias correction approaches, we develop a two-step bias correction and incorporate the analogue method to improve seasonal forecasting of rainfall.

3.3.1 Step 1: Box–Cox transformation

The Box–Cox transformation (Box and Cox 1964) is widely used in statistical analysis to transform data to a normal distribution. The transformation is applicable for both univariate and multivariate data by using appropriate objective function in estimation. In this study, we apply a bivariate Box–Cox transformation to historical variable x_k^O and corresponding GCM variable x_k^M

$$x_k^{O,B} = \tau(x_k^O; \lambda_O, \alpha_O) = \begin{cases} \frac{(x_k^O + \alpha_O)^{\lambda_O} - 1}{\lambda_O}, & \text{if } \lambda_O \neq 0, \\ \ln(x_k^O + \alpha_O), & \text{if } \lambda_O = 0; \end{cases} \quad (5)$$

$$x_k^{M,B} = \tau(x_k^M; \lambda_M, \alpha_M) = \begin{cases} \frac{(x_k^M + \alpha_M)^{\lambda_M} - 1}{\lambda_M}, & \text{if } \lambda_M \neq 0, \\ \ln(x_k^M + \alpha_M), & \text{if } \lambda_M = 0, \end{cases} \quad (6)$$

to make the joint distribution of $(x_k^{O,B}, x_k^{M,B})$ as close to the bivariate normal distribution as possible. The transformation parameters (λ_O, α_O) and (λ_M, α_M) are estimated from a likelihood approach. The main reason of employing simultaneous transformation rather than individual transformations separately is that we aim to match the NCEP/NCAR and POAMA data to achieve a better downscaled precipitation.

3.3.2 Step 2: standardization

After mitigating the possible skewness of the original data from the Box–Cox transformation, we standardize the transformed NCEP/NCAR and POAMA data $(x_k^{O,B}, x_k^{M,B})$ as follows

$$x_k^{O,N} = N(x_k^{O,B}) = \frac{x_k^{O,B} - \mu_{O,B}}{\sigma_{O,B}} \quad (7)$$

$$x_k^{M,N} = N(x_k^{M,B}) = \frac{x_k^{M,B} - \mu_{M,B}}{\sigma_{M,B}} \quad (8)$$

The standardization used in this study is slightly different from the one defined by Eq. (3) and is applied to both NCEP/NCAR and POAMA data in order to make the normalized data with zero mean and unit variance.

We apply this two-step bias correction to all large-scale predictors, calculate the dissimilarity measure defined by Eq. (1) based on $\{x_1^{O,N}, x_2^{O,N}, \dots, x_m^{O,N}, x_1^{M,N}, x_2^{M,N}, \dots, x_m^{M,N}\}$ and search the analogue day by Eq. (2). A flowchart for the procedure of improved analogue method is given in Fig. 2.

3.4 Assessment of variables used for forecast

Traditionally, a statistical downscaling method is assessed by NCEP/NCAR reanalysis data in both calibration and validation. The final model is then applied to a GCM to

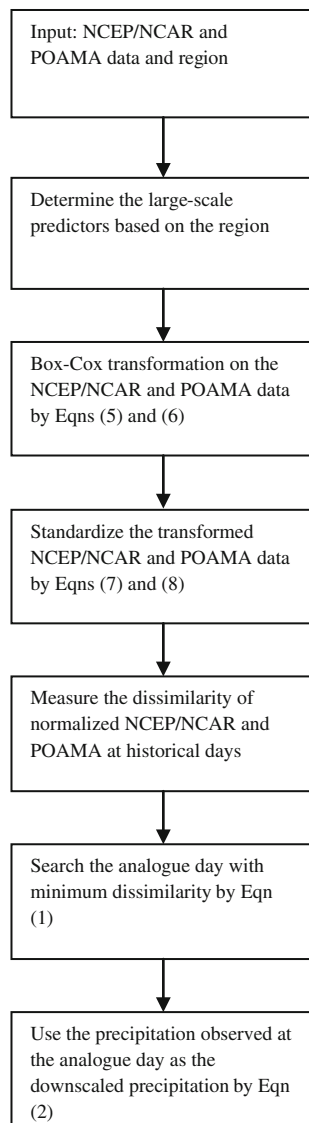


Fig. 2 Flow chart of improved analogue downscaling method

evaluate the future climate. However, for seasonal forecast, we have to assess further whether or not a variable used in calibration and validation based on NCEP/NCAR reanalysis data is still suitable when the variable is replaced by the GCM output, because the GCM may have very low skill on this variable even after bias correction. In this case, the use of low skill variable will not improve but may reduce the forecasting skill on the precipitation by adding more noise in the model. Therefore, we need re-assess if any variable should be excluded for forecasting purpose based on POAMA output.

Note that most statistical downscaling methods do not use the precipitation as a predictor to downscale precipitation due to the lack of skill but Timbal's analogue method sometimes uses it as a predictor. In this paper we

will also assess the usefulness of POAMA precipitation in the sense of seasonal forecasting.

4 Results

In this paper, all the computations are implemented in an R package (<http://cran.r-project.org/package=car>) and the codes are relatively simple. Readers can request the codes from the authors.

Unlike other statistical downscaling methods, the analogue statistical downscaling method does not need to parameter calibration as the similarity is defined by the Euclidian distance (see Eq. (1)). In order to obtain the bivariate Box–Cox transformation and standardization, we used all possible matched data between POAMA hindcast (for each lead time) and NCEP/NCAR from 1980–2006. To assess the performance of our improved analogue statistical downscaling method, we need to compare the downscaled and the observed rainfalls for the high quality stations. To do it, for each date in 1980–2000 (which is the span of observations for the high-quality stations), the analogue date is found as the date which has the minimum Euclidian distance between its POAMA hindcast and the full historical NCEP/NCAR dataset from 1958 to 2006 excluding the one-year window centred at the date. That is we choose the analogue date having the highest similarity to the given date and use the rainfall in the analogue date as the forecasted rainfall of the given date.

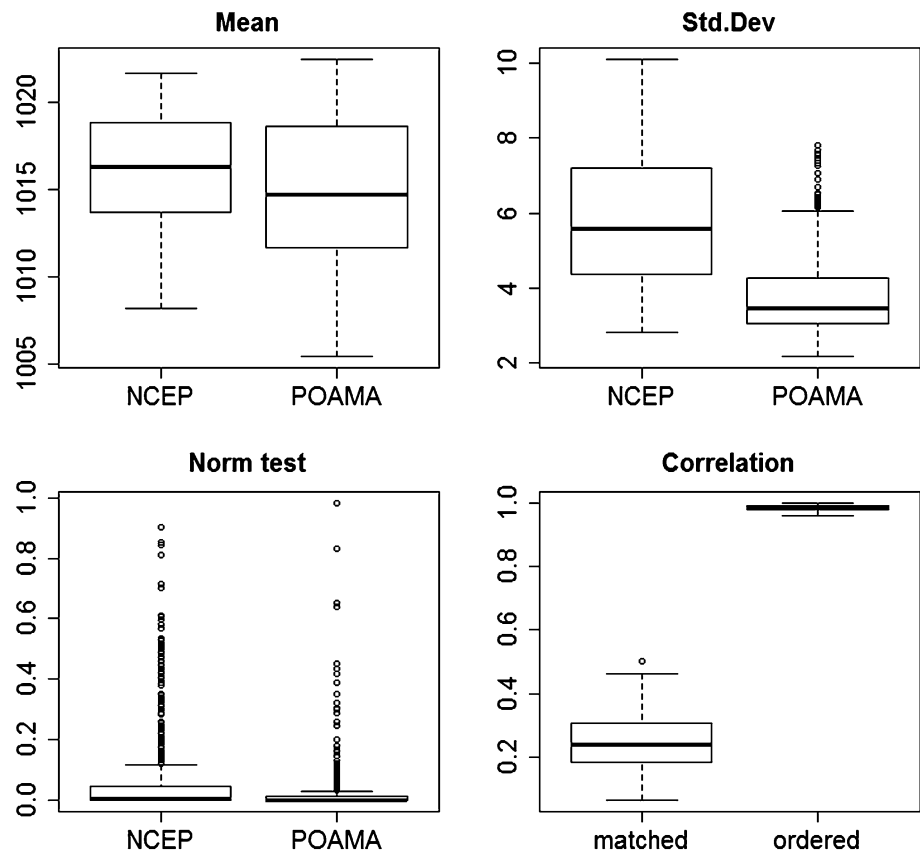
4.1 Assessment of different bias corrections for the POAMA outputs

Several comparisons are used to assess the POAMA outputs against NCEP/NCAR data. The following statistics are used for comparison: (i) mean and standard deviation; (ii) correlation with NCEP; (iii) correlation for ordered observations (to check the quantile–quantile bias correction which will be defined in the next section); and (iv) normality testing.

As an example, Fig. 3 gives the summary of daily NCEP/NCAR and POAMA MSLP outputs across all grids with lead time of 2 months. It can be seen that in general NCEP has larger mean and standard deviation than POAMA. The Lilliefors test (Dallal and Wilkinson 1986) was performed for normality test. Both NCEP and POAMA outputs are frequently not normally distributed (large P-values mean that the data are normally distributed). The POAMA outputs have some skills with positive correlations between NCEP and POAMA paired (matched) data and with quite high correlations between ordered NCEP and POAMA data.

Note that the bias correction by Q–Q transformation can be performed at either daily or monthly scale. We

Fig. 3 Summary of NCEP and POAMA daily MSLP with lead time = 2 month



conducted the bias correction at the daily time scale. To compare our bias correction with other methods, we summarize the results for all grids and months in Table 1 for different bias correction methods, together with the results for the original data. As a demonstration, we show the xy plot of NCEP and POAMA daily (Fig. 4) data and summarized data to monthly scale (Fig. 5) for MSLP at grid (3, 3) with lead time of 2 months. It can be seen that there are some correlations between the original POAMA and NCEP data at daily scale with lead time of 1 month but the correlation decreases fast for the lead times of 2 and 3 months, meaning that there are some skills in POAMA up to 1 month but the skill decreases fast for the lead times of 2 and 3 months. At monthly scale (Fig. 5), the correlation is higher than that that at daily scale. However, the POAMA outputs are relatively flat (with a smaller standard deviation) compared with NCEP data.

At the original daily scale, the standardization approach (referring to row 3 of Fig. 4) retains the correlation (row 3 and column 1) and corrects the distributional mismatch except for upper and lower tails (row 3 and column 2) but does not change the normality of the data (row 3 and column 3). The Q–Q transformation (row 4) results in perfect distribution match as expected (row 4 and column 2) but destroys the correlation (row 4 and column 1) and does not change the normality of the data either (row 4 and

column 3). The proposed Box–Cox transformation approach (referring to Row 2 of Figs. 4, 5) retains the correlation (row 2 and column 1), makes two distributions close to each other even at two tails (row 2 and column 2) and approximately changes the data to normal distribution (row 2 and column 3). The findings are also evident from the qq plot between NCEP and POAMA values (row 1 and column 2 of Fig. 4), from which we can see the distributional mismatch between POAMA and NCEP outputs.

To further understand the performance of bias correction, the bias corrected daily data are also summed to monthly scale and the same statistics are given in the second part of Table 1 and demonstrated in Fig. 5. The normality is important for the use of standardization. However, it can also be seen that both NCEP and POAMA data are not normally distributed (row 1 and column 3 of Figs. 4, 5) at monthly scale although there are some correction.

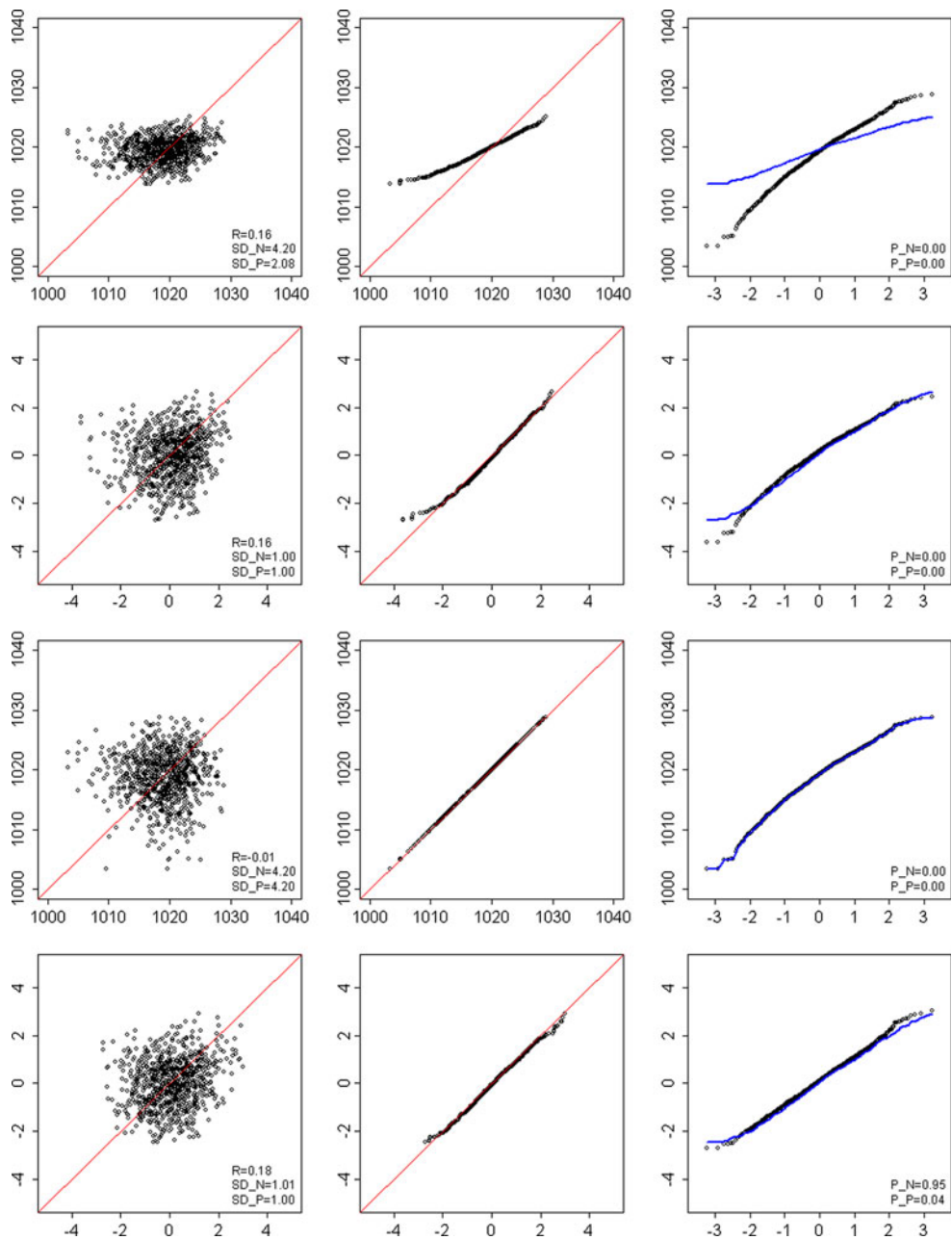
4.2 Performance of different bias corrections for the analogue downscaling

As the purpose of the bias correction is to improve the performance of statistical downscaling methods, it is important to assess the performance of different bias corrections in term of the analogue downscaling results. To do this, different bias correction methods were applied to the

Table 1 Summary of statistics for the original and transformed data cross all grids in three lead times

	Lead time = 1 month						Lead time = 2 month						Lead time = 3 month					
	Mean	SD	Cor	Normal test (%)	Mean	SD	Cor	Normal test (%)	Mean	SD	Cor	Normal test (%)	Mean	SD	Cor	Normal test (%)		
Daily data																		
MSLP	Org	(1016.2, 1014.8)	(5.87, 3.78)	0.25	(24.32, 12.07)	(1016.2, 1014.8)	(5.87, 2.21)	0.06	(23.98, 70.41)	(1016.21, 1015.81)	(5.88, 2.16)	0.05	(23.81, 77.55)					
	Std	(0.05, 0)	(1, 1)	0.25	(24.32, 12.07)	(0.05, 0)	(1, 1)	0.06	(23.98, 70.41)	(0.05, 0)	(1, 1)	0.05	(23.81, 77.55)					
	QQ	(1016.2, 1016.2)	(5.87, 5.87)	0	(24.32, 24.32)	(1016.2, 1016.2)	(5.87, 5.87)	0	(23.98, 23.98)	(1016.21, 1016.21)	(5.88, 5.88)	0	(23.81, 23.81)					
V850	Bet	(0.05, 0)	(1.01, 1)	0.25	(79.08, 44.05)	(0.05, 0)	(1.01, 1)	0.06	(79.08, 92.52)	(0.05, 0)	(1.01, 1)	0.06	(78.74, 93.03)					
	Org	(0.35, -0.37)	(5.08, 2.79)	0.14	(67.35, 23.81)	(0.34, -0.1)	(5.08, 1.77)	0.01	(66.5, 92.52)	(0.34, -0.27)	(5.08, 1.7)	0.02	(65.82, 88.95)					
	Std	(-0.01, 0)	(1, 1)	0.14	(67.35, 23.81)	(-0.01, 0)	(1, 1)	0.01	(66.5, 92.52)	(-0.01, 0)	(1, 1)	0.02	(65.82, 88.95)					
Monthly data	QQ	(0.35, 0.35)	(5.08, 5.08)	0	(67.35, 67.35)	(0.34, 0.34)	(5.08, 5.08)	0	(66.5, 66.5)	(0.34, 0.34)	(5.08, 5.08)	0	(65.82, 65.82)					
	Bet	(-0.01, 0)	(1, 1)	0.14	(85.71, 53.57)	(-0.01, 0)	(1, 1)	0.01	(86.56, 98.13)	(-0.01, 0)	(1, 1)	0.02	(86.39, 97.79)					
	Org	(1016.2, 1014.8)	(2.35, 1.65)	0.44	(96.09, 96.94)	(1016.2, 1014.8)	(2.34, 1.06)	0.16	(96.09, 94.05)	(1016.21, 1015.81)	(2.35, 1.07)	0.04	(96.43, 90.14)					
V850	Std	(0.05, 0)	(0.4, 0.43)	0.44	(96.09, 96.94)	(0.05, 0)	(0.4, 0.48)	0.16	(96.09, 94.05)	(0.05, 0)	(0.4, 0.49)	0.04	(96.43, 90.14)					
	QQ	(1016.2, 1016.2)	(2.35, 1.26)	-0.02	(96.09, 93.37)	(1016.2, 1016.2)	(2.34, 1.17)	0	(96.09, 94.9)	(1016.21, 1016.21)	(2.35, 1.19)	-0.03	(96.43, 93.2)					
	Bet	(0.05, 0)	(0.41, 0.44)	0.44	(96.43, 96.94)	(0.05, 0)	(0.4, 0.48)	0.16	(96.6, 95.24)	(0.05, 0)	(0.41, 0.49)	0.04	(96.77, 91.16)					
Monthly data	Org	(0.35, -0.37)	(1.27, 0.84)	0.28	(94.22, 94.22)	(0.34, -0.1)	(1.27, 0.52)	0.09	(94.22, 95.58)	(0.34, -0.27)	(1.27, 0.48)	0.08	(94.05, 95.07)					
	Std	(-0.01, 0)	(0.25, 0.31)	0.28	(94.22, 94.22)	(-0.01, 0)	(0.25, 0.3)	0.09	(94.22, 95.58)	(-0.01, 0)	(0.25, 0.29)	0.08	(94.05, 95.07)					
	QQ	(0.35, 0.35)	(1.27, 0.95)	0.02	(94.22, 95.24)	(0.34, 0.34)	(1.27, 0.93)	0	(94.22, 95.58)	(0.34, 0.34)	(1.27, 0.95)	-0.03	(94.05, 95.07)					
Bet	(-0.01, 0)	(0.25, 0.31)	0.28	(94.22, 95.41)	(-0.01, 0)	(0.25, 0.3)	0.09	(93.71, 95.24)	(-0.01, 0)	(0.25, 0.29)	0.08	(93.37, 96.26)						

Fig. 4 Plot of NCEP (x axis) and POAMA (y axis) daily MSLP at grid (3, 3) with lead time = 2 months. The plots are organized in row by column by following. Row 1 is the result for observations, row 2 for standardization, row 3 for Q–Q transformation and row 4 for our Box–Cox transformation plus standardization. Column 1 gives xy plots, column 2 qq plots and column 3 Q–normal plots in which dots represent the NCEP data and line represents the POAMA data

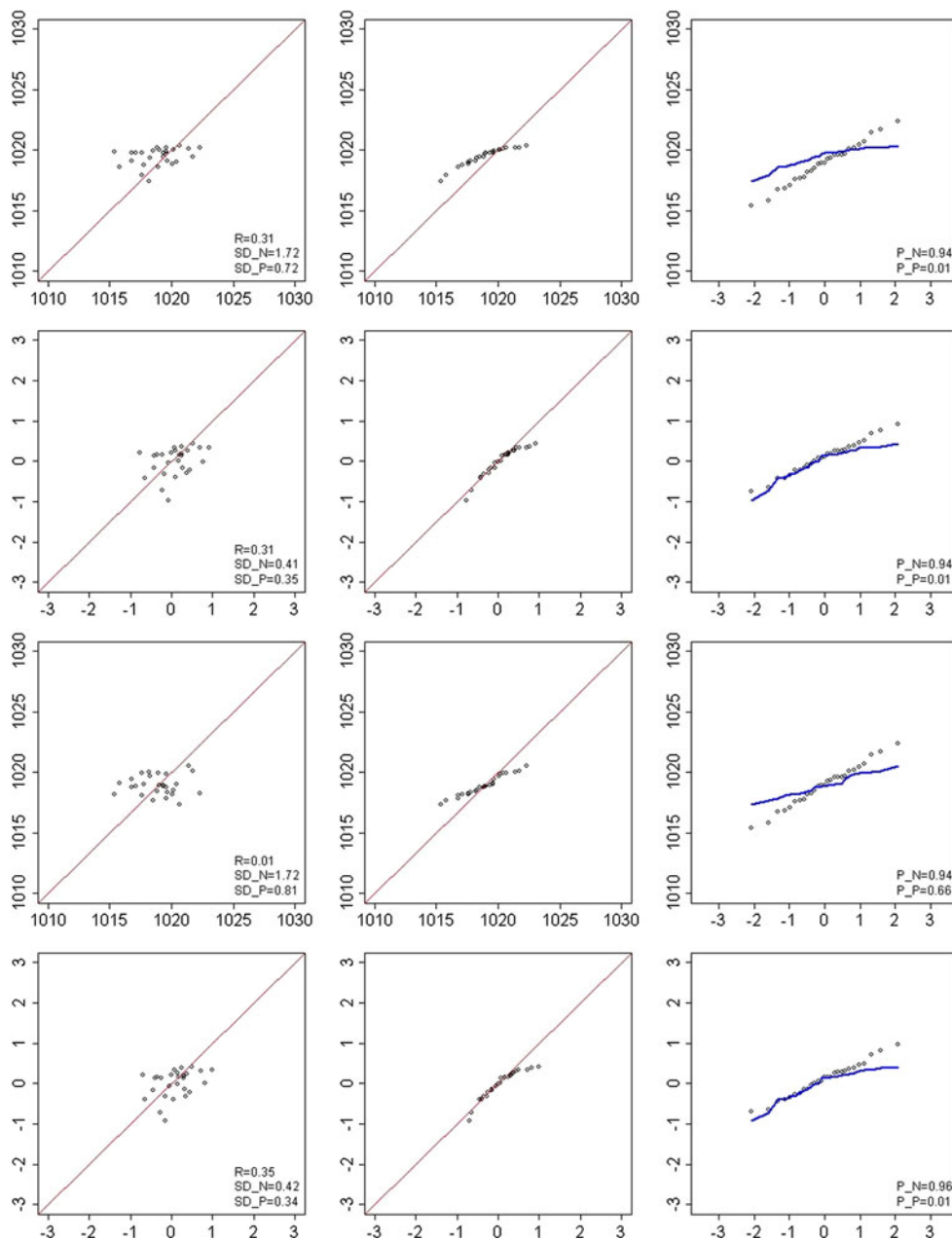


analogue downscaling method, including (1) the standardized NCEP and POAMA data; (2) the original NCEP and Q–Q transformed POAMA data and (3) the Box–Cox transformed NCEP and POAMA data.

In this paper we also assess the usefulness of POAMA precipitation in the sense of forecasting. To do this, we provide the results with and without precipitation as predictor. Precipitation is observed with many zero values, and this results that the distribution of precipitation is highly skewed and hardly possible to be transformed to normal. Therefore, we do not carry out the Box–Cox transformation on precipitation and use the standardized precipitation instead.

Before applying the analogue downscaling method for the purpose of forecasting, we examine its performance when the NCEP/NCAR reanalysis data are used (that is, its performance in calibration and validation). Figure 6 gives that correlation of observed (NCEP/NCAR reanalysis) and downscaled precipitations using bias corrections by standardization and Box–Cox transformation for each month. It can be seen that for each bias correction, the use of NCEP/NCAR reanalysis precipitation improves the downscaling result. For a given set of predictors (i.e., with and without precipitation), the bias correction by standardization has more chance (8 of 12 months in both cases) of out-performing that by the Box–Cox transformation. It is not

Fig. 5 Plot of NCEP (x axis) and POAMA (y axis) monthly MSLP at grid (3, 3) with lead time = 2 months. The plots are organized in row by column by following. Row 1 is the result for observations, row 2 for standardization, row 3 for Q–Q transformation and row 4 for our Box–Cox transformation plus standardization. Column 1 gives xy plots, column 2 qq plots and column 3 Q–normal plots in which dots represent the NCEP data and line represents the POAMA data



surprising as the NCEP/NCAR reanalysis precipitation has high skill and contains more information on the station precipitation, and it is not necessary to conduct data transformation as the original data contain more precise information than transformed data.

As we are interested in the seasonal forecasting by using POAMA, we selected three lead times (1, 2 and 3 months corresponding to 30, 60 and 90 days, respectively) in preparing POAMA data. Note that most statistical downscaling methods do not use the precipitation as a predictor to downscale precipitation due to the lack of skill but Timbal’s analogue method sometimes uses it as a predictor. In the following study, we provide the results with and without precipitation as predictor. Precipitation is observed

with many zero values, and causes the distribution of precipitation to be highly skewed and hardly possible to be transformed to normal. Therefore, we do not carry out the Box–Cox transformation on precipitation and use the standardized precipitation instead.

The correlations of Timbal’s analogue method using different bias correction methods with and without precipitation for June with three lead times are given in Table 2. The boxplots of Timbal’s analogue method for the winter rainfall forecasting using different bias correction methods with and without precipitation are given in Fig. 7. The following findings can be concluded. (1) For a given lead time and predictor set, our new bias correction always performed best, followed by the standardization. (2) For a

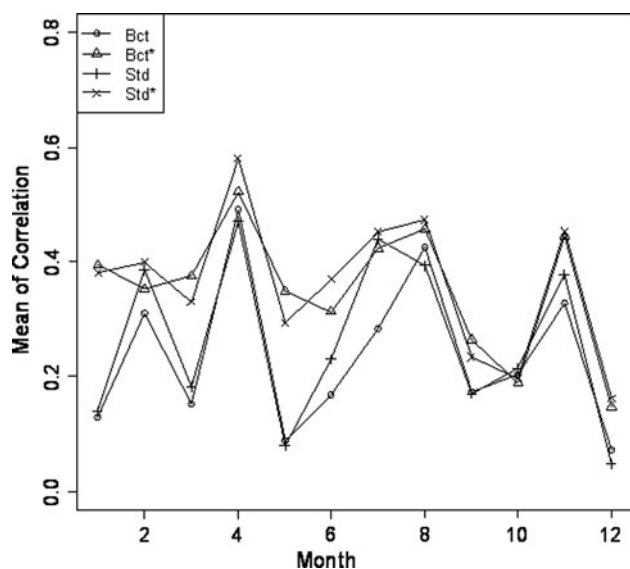


Fig. 6 The mean of correlation between downscaled and observed monthly rainfall across 35 high quality stations. The model with (or without) *asterisk* denote the use (or no use) of NCEP/NCAR reanalysis daily precipitation

Table 2 The median of the correlation between monthly observed and downscaled rainfall over 35 high quality stations

Method	Lead time (month)		
	1	2	3
Std ^a	0.316	0.211	-0.018
Std	0.335	0.193	0.175
QQ ^a	0.143	0.090	-0.049
QQ	0.084	0.137	-0.157
Bct ^a	0.360	0.216	0.080
Bct	0.381	0.287	0.192

^a (or not) denotes the use (or not) of daily precipitation hr24_precp

given bias correction and predictor set, the skill of downscaling method decreases as the lead time increases, except the Q–Q transformation without precipitation at lead time of 2 month. (3) For our Box–Cox bias correction and standardization, the use of precipitation generally decreases the skill of downscaling result.

5 Discussions and conclusion

Our improved analogue method has the same dissimilarity measurement as Eq. (1) but based on a different bias correction to solve the mismatch between the historical data and GCM output. Our bias correction is a two-stage procedure: (i) bivariate Box–Cox transformation to make both NCEP and GCM outputs as normally distributed as

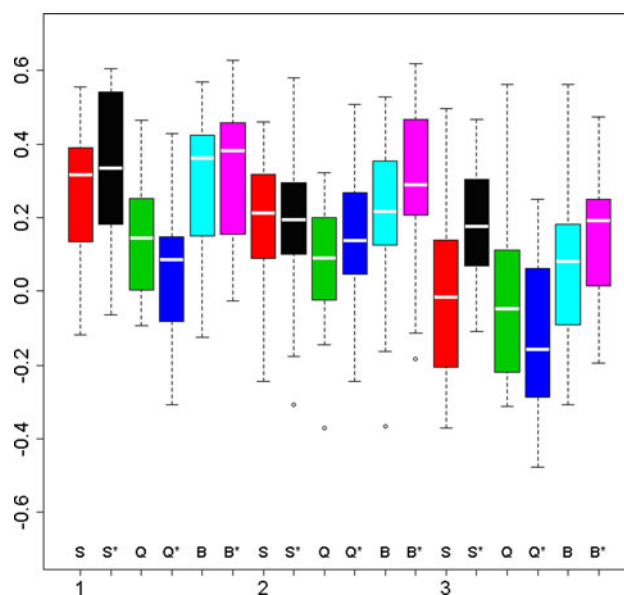


Fig. 7 Boxplots of the correlation between monthly observed and downscaled rainfall over 35 high quality stations from six methods including: standardization (denoted by S), Q–Q transformation (denoted by Q) and Box–Cox transformation (denoted by B), with (no *superscript*) and without (denoted by *asterisk*) daily precipitation, for three lead times (1, 2 and 3 months). The POAMA data used are from 1980 to 2000 and NCEP reanalysis used are from 1960 to 2000

possible and (ii) normalization to match mean and standard deviation. The proposed bias correction is based on the following considerations in statistics. (1) Both the standardization and Q–Q transformation try to rectify some distortion of GCM outputs and are complementary to each other. The standardization reserves the correlation between the NCEP and GCM but does not change the shape of their distributions. (2) In contrast, the Q–Q transformation ensures a perfect match (or nearly perfect in the parametric Q–Q transformations) between the NECP and GCM data in the period of interest (e.g. all the POAMA 1 month forecast daily data in January from 1980 to 2006). (3) Although the Q–Q transformation is a monotonic transformation, it could not control the change of correlation after the transformation because of the nonlinearity in the transformation. However, if the Q–Q transformation is linear, it will reserve the correlation. Therefore, it will encourage us to find a distribution which gives a linear Q–Q transformation. An obvious one is normal distribution. (4) Note also that standardization is a linear transformation from the original data to standardized data and that for normal data, standardization will produce the same transformation at different time scale under mild assumptions (including independent data and autocorrelated data). However, such the scale invariant property does not hold in a general distribution. (5) The normal distribution is determined by two parameters: mean and standard deviation, and thus the Q–Q transformation makes the distribution for normally

distributed data by just matching these two parameters. Therefore, if both GCM and reanalysis data are normally distributed, the Q–Q transformation and standardization are almost identical (subject to sample error). (6) As GCM's climatology (modeled system) may be different from the observed climatology in a nonlinear way, we cannot simply expect the NCEP and GCM output have linear relationship. We can see from our exploration in the next section that the observed and modeled outputs are not highly (linearly) correlated. In fact, the Q–Q transformation does not assume that GCM and reanalysis data are linearly related. (7) Many statistical tools require normality assumption on variables. Otherwise the analysis, estimation and inference may be inefficient, or more seriously, incorrect. (8) Data transformations are widely used in statistics and have been proved to be powerful. Furthermore, there are powerful mathematical transformations to make variables as close to normal distribution as possible.

We conclude that the proposed two-stage bias correction is not only statistically sound but also was proven useful when being applied in conjunction with Timbal's analogue method. The proposed two-bias correction was shown to transform the non-normal POAMA data to normal distribution and to largely retain the correlation between NCEP/NCAR and POAMA data. The correlation between observed and forecasted rainfall aggregated in monthly totals has been improved by using the proposed two-stage bias correction. Nevertheless, the forecasting skill reduces with a long lead time. Furthermore, we found that the precipitation should not be used in the calculation of similarity when POAMA is used for forecasting. This may be caused by the large uncertainty of precipitation output from POAMA. Further research is being conducted to assess the performances of our new bias correction in different regions using different statistical downscaling methods and to integrate our new bias correction to a new statistical downscaling method being developed.

It is an important and interesting research to evaluate the performance of our improved analogue statistical downscaling in other regions. We will extend our improved analogue statistical downscaling methods to other Australian regions and assess its performance once high quality data are available. Furthermore, it is also important to assess the impact of this improved analogue statistical downscaling on seasonal streamflow forecasting. The results will be reported in separate papers.

The two-step bias-correction presented in this paper is in fact a data transformation technique which tends to make the transformed data as closely to be normally distributed as possible, because many statistical methods require normality in data. It would be an interesting research topic to see the effect of newly transformed data on different statistical downscaling methods.

Acknowledgments This research is supported by the Water Information Research and Development Alliance (WIRADA) between the Australian Bureau of Meteorology (BoM) Water Division and the CSIRO Water for a Healthy Country Flagship Program. The authors wish to acknowledge useful discussions from CSIRO colleagues QJ Wang, Enli Wang and Hongxing Zheng and BoM collaborators Narendra Kumar Tuteja, Bertrand Timbal, Daehyok Shin and Sri Srikanthan. We would like also express our thanks to the editor, an Associate Editor and two anonymous referees for their valuable comments which helped to improve the quality of this paper.

References

- Alves O, Wang G, Zhong A, Smith N, Tseitki F, Warren G, Schiller A, Godfrey S, Meyers G (2003) POAMA: bureau of meteorology operational coupled model seasonal forecast system. In: Proceedings of the national drought forum, Brisbane, April 2003, pp 49–56. Available from DPI Publications, Department of Primary Industries, GPO Box 46, Brisbane, QLD, 4001, Australia
- Box George EP, Cox DR (1964) An analysis of transformations. *J R Stat Soc B* 26(2):211–252
- Chandler RE, Wheeler HS (2002) Analysis of rainfall variability using generalized linear models: a case study from the west of Ireland. *Water Resour Res* 38(10):1192. doi:10.1029/2001WR000906
- Charles SP, Bates C, Hughes JP (1999) A spatiotemporal model for downscaling precipitation occurrence and amounts. *J Geophys Res* 104((D24)):31657–31669
- Charles A, Fernandez E, Timbal B, Hendon H (2010) Analogue downscaling of POAMA seasonal rainfall forecasts. Technical report: WIRADA project 4.2 activity B: improved climate predictions at hydrologically-relevant time and space scales from the POAMA seasonal climate forecasts
- CSIRO (2008) Water availability in the Murray-Darling Basin. A report to the Australian government from the CSIRO Murray-Darling basin sustainable yields project. CSIRO, Australia
- Dallal GE, Wilkinson L (1986) An analytic approximation to the distribution of Lilliefors' test for normality. *Am Stat* 40:294–296
- Fedderson H, Andersen U (2005) A method for statistical downscaling of seasonal ensemble predictions. *Tellus* 57A:398–408
- Hansen JW, Challinor A, Ines A, Wheeler T, Moron V (2006) Translating climate forecasts into agricultural terms: advances and challenges. *Clim Res* 33:27–41
- Hughes JP, Guttorp P, Charles SP (1999) A non-homogeneous hidden Markov model for precipitation occurrence. *Appl Stat* 48:15–30
- Ines AVM, Hansen JW (2006) Bias correction of daily GCM rainfall for crop simulation studies. *Agric For Meteorol* 138:44–53
- Kruizinga S, Murphy A (1983) Use of an analogue procedure to formulate objective probabilistic temperature forecasts in the Netherlands. *Mon Weather Rev* 111:2244–2254
- Lorenz EN (1969) Atmospheric predictability of a 28-variable atmospheric model. *Tellus* 27:321–333
- Luo L, Wood ER (2008) Use of Bayesian merging techniques in a multimodel seasonal hydrologic ensemble prediction system for the Eastern United States. *J Hydrometeorology* 9:866–884. doi:10.1175/2008JHM980.1
- Piani C, Haerter JO, Coppola E (2010) Statistical bias correction for daily precipitation in regional climate models over Europe. *Theor Appl Climatol* 99:187–192
- Shao Q, Li M, Chan C (2010). Assessing the performance of statistical downscaling techniques on POAMA to Murrumbidgee sub-catchments. Water for a Healthy Country National Research Flagship, CSIRO
- Sharma D, Das Gupta A, Babel MS et al (2007) Spatial disaggregation of bias-corrected GCM precipitation for improved

- hydrologic simulation: ping river basin Thailand. *Hydrol Earth Syst Sci* 11(4):1373–1390
- Timbal B, Li Z, Fernandez E (2008) The Bureau of meteorology statistical downscaling model graphical user interface: user manual and software documentation. CAWCR Technical Report No. 004. Bureau of Meteorology, Australia
- Wang E, Zheng H, Shao Q (2011) Ensemble streamflow forecasting using rainfall-runoff modelling with conditional parameterisation and GCM predictions in East Australia (submitted)
- Wilby RL, Wigley TML (1997) Downscaling general circulation model output: a review of methods and limitations. *Prog Phys Geogr* 21(4):530–548
- Wood AW, Maurer EP, Kumar A, Lettenmaier DP ((2002) Longrange experimental hydrologic forecasting for the eastern United States. *J Geophys Res* 107(D20):4429. doi:[10.1029/2001JD000659](https://doi.org/10.1029/2001JD000659)
- Wood AW, Kumar A, Lettenmaier DP (2005) A retrospective assessment of National Centers for Environmental Prediction climate model-based ensemble hydrologic forecasting in the western United States. *J Geophys Res* 110:D04105. doi:[10.1029/2004JD004508](https://doi.org/10.1029/2004JD004508)
- Zorita E, von Storch H (1999) The analog method as a simple statistical downscaling technique: comparison with more complicated methods. *J Clim* 12:2474–2489
- Zorita E, Hughes J, Lettenmaier D, von Storch H (1995) Stochastic downscaling of regional circulation patterns for climate model diagnosis and estimation of local precipitation. *J Clim* 8: 1023–1042