18th World IMACS / MODSIM Congress, Cairns, Australia 13-17 July 2009
http://mssanz.org.au/modsim09

A multisite daily rainfall data generation model for climate change conditions

Srikanthan, R.

eWater CRC, Water Division, Bureau of Meteorology, Melbourne, Victoria
Email: r.srikanthan@bom.gov.au

Abstract: Concerns over climate change caused by increasing concentration of CO_2 and other trace gases in the atmosphere have increased in recent years. A major effect of climate change may be alterations in regional hydrologic cycles and changes in regional water availability. Predictions produced from General Circulation Models (GCMs) are the main source at present to get information about the future climate. However, the resolution of the GCMs is too coarse to use in hydrologic modelling to assess the impact of climate change. As a result, a number of statistical downscaling techniques (Hughes et al., 1999; Chandler, 2002; Charles et al., 2004; Mehrotra and Sharma, 2005) have been developed to obtain rainfall data at catchment scale. While the techniques produce reasonable results (see Frost et al., 2009) for a comparison of models applied within Australia), they can be complicated and time consuming to apply. As a result, simple techniques, such as constant or daily scaling, are still used to obtain daily rainfall data under future climate for either a single site or catchment average rainfall. There are difficulties in applying these methods to multiple sites and maintaining spatially consistent rainfall values. At present, there is no simple technique available to obtain rainfall data for future climate for multiple stations apart from the constant scaling. The eWater CRC has already developed a multi-site two-part model for the generation of spatially consistent daily rainfall data for a number of sites for the current climate. In this paper, this model is modified by adjusting the parameters based on the GCM output to take into account of the effect of climate change. The parameters modified are the means and standard deviations at the monthly and annual time scales and the transition probabilities and gamma distribution parameters at the daily time scale. The cross correlations were not modified.

The nested multisite two-part model was applied to 30 stations in the Murrumbidgee River catchment for the current climate (1981 – 2000). One hundred replicates, each of length 20 years, were generated to evaluate the performance of the model. A number statistics at the daily, monthly and annual time scale were calculated and the results indicated a satisfactory performance of the model. For the future climate (2046 – 2065) with A2 emission case, daily rainfall data were obtained from the Max Planck Institute GCM simulation. Ten grid points lying in and around the catchment were used to adjust the parameters of the model. The ratios of the parameters for the future and current climate GCM daily rainfall were obtained first at the ten grid points and then interpolated to 30 sites using inverse distance weighting. The ratios calculated monthly appeared to vary considerably between successive months due to small sample size (20 years). To overcome this, the ratios were averaged over the four seasons and the same value was used for the three months within a season. Again, one hundred replicates each of length 20 years were generated and the above statistics were calculated. The statistics calculated from the generated data were compared with the adjusted values of the parameters used in the model. The comparison showed that the model satisfactorily preserved the input model parameters for the future climate.

Keywords: *Climate Change, Stochastic Model, Daily Rainfall*

1. INTRODUCTION

Concerns over climate change caused by increasing concentration of CO_2 and other trace gases in the atmosphere have increased in recent years. A major effect of climate change may be alterations in regional hydrologic cycles and changes in regional water availability. The main source of climate change projections are the General Circulation Models (GCMs). While current GCMs perform reasonably well in simulating the present climate with respect to annual and seasonal averages over large areas, they are considerably less reliable in regional scale information that are necessary for hydrologic studies. As a result, the climate change impact studies need to use a spectrum of climate change scenarios and to reflect the different possible CO_2 scenarios. These are generally constructed using observed records of temperature and rainfall adjusted to reflect climate changes obtained from monthly average GCM results.

Most of the early work on the impacts of climate change used historical data adjusted for the climate change (Lettenmaier and Gan, 1990; Panagoulia, 1992). The rainfall records were multiplied by the monthly precipitation ratios for the CO_2-doubling and control runs. The main limitation of this method is that it does not take into account the dynamic changes in the temporal distribution of rainfall nor the changes in spatial rainfall at the sub-grid scale. In addition, it also reduces the magnitude of the extreme rainfall in areas where GCM runs indicates a decrease in mean annual rainfall. However, most GCM runs predict an increase in extreme daily rainfall in areas where a decrease in mean annual rainfall is predicted.

The most popular technique used to obtain rainfall data at catchment scale for future climate conditions is the use of statistical downscaling. A number of statistical downscaling techniques (Hughes et al., 1999; Chandler, 2002; Charles et al., 2004; Mehrotra and Sharma, 2005) have been developed over the years. These techniques are quite complex, cumbersome and are not easy to learn and apply. As a result, simpler methods have been developed so that these can be easily applied.

Wilby et al. (2002) developed a software for statistical downscaling model for obtaining daily rainfall and temperature at a single station. Harrold and Jones (2003) proposed a daily scaling method to overcome some of the limitations of the constant scaling method. Two approaches are used to apply the daily scaling method. In the first approach, the ratio of the ranked GCM daily rainfall for the future and current climate is used to scale the ranked historical daily rainfall data (Chiew et al., 2003; Harrold et al., 2005). Like the constant scaling method, this approach also ignores the changes in the temporal distribution of rainfall. However, this method changes different rainfall amounts differently. The second approach uses the ratio of the historical daily rainfall for a given percentile to GCM rainfall for the same percentile for the present climate to modify the GCM daily rainfall for the future climate. This approach incorporates changes in the temporal distribution as simulated by the GCM for the future conditions. A major drawback of modifying the historical rainfall is that there is only one sequence which does not represent the sampling variability of the future rainfall.

In this paper, the nested multisite two-part model was modified to generate daily rainfall under future climate conditions. The mean and standard deviation of annual and monthly rainfall along with the rainfall occurrence from the GCM projections were used to adjust the parameters of the model. The paper is structured as follows. The rainfall data and the study region are briefly described in Section 2. Section 3 describes the multisite two-part model, while Section 4 describes the adjustment procedures for the parameters of the model for future climate conditions. The results from the application of the modified model are presented in Section 5, and the conclusions drawn from the results in Section 6.

2. GAUGE RAINFALL AND GCM SIMULATION DATA

The Murrumbidgee River catchment is located in southern New South Wales and the Australian Capital Territory lies within the catchment. The area of the catchment is 81563 km^2. Thirty stations were selected, whose locations are shown in Figure 1. The available daily rainfall data is 110 years long covering the period 1890 to 1999. The mean annual rainfall varies from about 340 to 970 mm, while the average number of wet days per year varies from 57 to 106. Even though a long length of data is available for these stations, only 20 years (1981 – 2000) of data were used to calibrate the model. This period is the same as the period for which the GCM was used to estimate the parameters for current climate conditions. The Max Planck Institute's GCM results were used to obtain the daily rainfall for the current climate (1981 – 2000) and future climate (2046 – 2065) for the ten grid points shown in Figure 1. The A2 emission scenario was used in the future climate GCM runs.

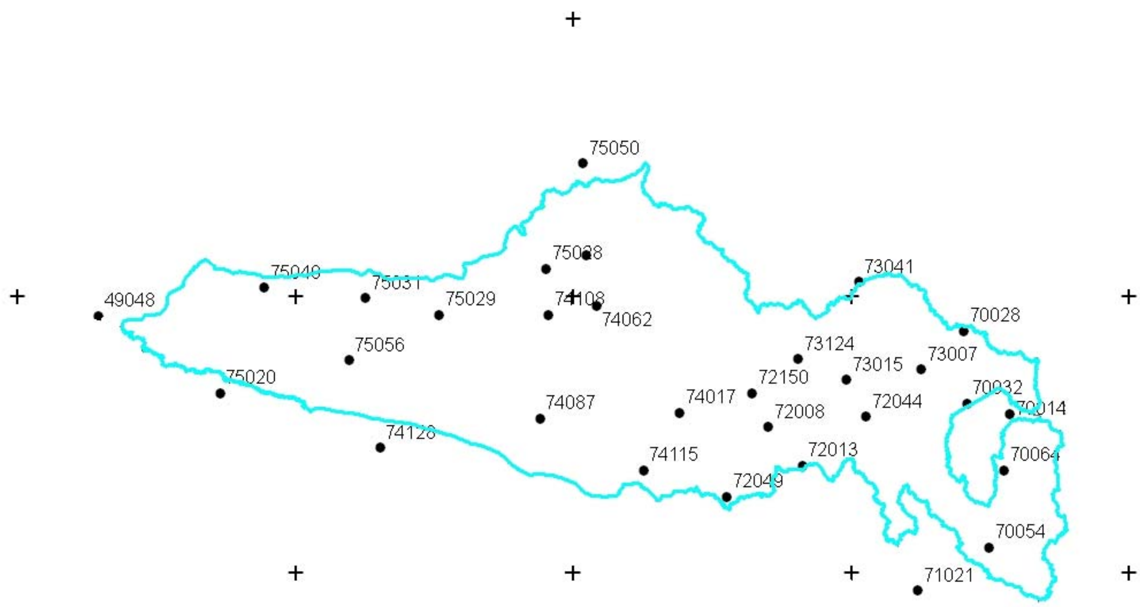

Figure 1. Locations of the rainfall stations (•) and the GCM grid points (+) used.

3. NESTED MULTISITE TWO-PART MODEL

The nested multi-site daily rainfall model consists of three parts: occurrence, amounts and nesting. These three parts are briefly described below. For a more detailed description and derivation, readers are referred to Srikanthan and Pegram (2009).

3.1. Multi-site rainfall occurrence model

A first-order two-state Markov chain is used to determine the occurrence of rainfall at each site. For each site, the Markov chain has two transition probabilities. The individual models are driven by serially independent but cross-correlated random numbers to preserve the spatial correlation in the rainfall occurrence process. The method used to derive the required matrix to transform independent random numbers to cross correlated random numbers is given in Srikanthan and Pegram (2009).

3.2. Rainfall amounts model

The rainfall amounts on wet days are generated by using a Gamma distribution, which has been found to fit better than the routinely used distributions, exponential and Weibull. As was detailed above for the occurrence model, the spatial correlation in the daily rainfall amounts is preserved by using a vector of suitably cross-correlated uniform variates.

Once the daily rainfalls at all sites are generated for a given month, the monthly rainfall totals, $\tilde{x}_i^k$, at each site are obtained by summing the daily rainfall values. Their cross-correlations are calculated and the monthly totals modified by using a multi-site monthly model (Srikanthan and Pegram, 2009) to preserve the monthly spatial and serial correlations.

$$\boldsymbol{X_i} = \boldsymbol{A_i}\,\boldsymbol{X_{i-1}} + \boldsymbol{B_i}\,\boldsymbol{a_i} \qquad (1)$$

where $\boldsymbol{A_i}$ and $\boldsymbol{B_i}$ are coefficient matrices and $\boldsymbol{X_i}$ is the adjusted standardised monthly rainfall (zero mean and unit variance) vector for month *i*. The details of the estimation procedure for the coefficient matrices are presented in Srikanthan and Pegram (2009).

After the adjustment, the monthly rainfall, x_i^k, at each site is obtained by multiplying the standardised value X_i^k by the standard deviation and adding the mean. Once the values for the twelve months of a year (j) have been adjusted, the generated monthly values are aggregated to obtain the annual values ($\tilde{z}_j^k$). The aggregated annual values are standardised to have zero mean and unit variance and then modified as above, by using a multi-site model to preserve the annual characteristics.

$$\boldsymbol{Z_j} = \boldsymbol{P}\, \boldsymbol{Z_{j-1}} + \boldsymbol{Q}\, \boldsymbol{b_j} \tag{2}$$

where $\boldsymbol{P}$ and $\boldsymbol{Q}$ are coefficient matrices to preserve the lag zero and lag one cross correlations, $\boldsymbol{b}_j$ is the already generated standardised annual value before adjustment and $\boldsymbol{Z}_j$ is the adjusted standardised annual rainfall (zero mean and unit variance) vector. After the adjustment, the annual rainfall at each site is again obtained by appropriate scaling and shifting.

Each generated monthly rainfall value is multiplied by the ratio $z_j^k / \tilde{z}_j^k$. This will preserve the annual characteristics. The modified monthly rainfall values are used to adjust the daily rainfall values. Rather than adjusting the daily rainfall values twice, the adjustment to the daily rainfall values can be carried out in one step by multiplying the generated rainfall values for each month (i) by the ratio $x_i^k z_j^k / \tilde{x}_i^k \tilde{z}_j^k$. If the lag one cross correlations (monthly or annual) are all small, a contemporaneous multi-site model can be used for nesting. In this case, the matrix ***A*** *or* ***P*** becomes a diagonal matrix with diagonal elements being the lag one autocorrelations. In this case, one only needs to estimate the other matrix ***B*** *or* ***Q***.

4. ADJUSTMENT OF MODEL PARAMETERS FOR CLIMATE CHANGE

The nested multisite two-part model has three sets of parameters: annual, monthly and daily. At the annual and monthly levels, only the mean and standard deviations are adjusted for future climate conditions. The ratio of the means and standard deviations for the future and current climate conditions are calculated for the ten grid points. For each site, the required ratio is computed from the nearby grid values by using inverse distance as weights, and then used for the adjustment. The cross correlations are not adjusted. The daily model parameters are adjusted as follows.

Wilks (1992) presented a method to adapt stochastic daily weather generation models for generation of synthetic daily time series consistent with assumed future climates. The assumed climates were specified by the monthly means and variances of rainfall and temperature. For a two-part rainfall model with gamma distribution for rainfall amounts, there are four parameters (p_{11}, p_{01}, α, β). The shape and scale parameters of the gamma distribution are denoted by α and β. The transition probabilities (p_{11} and p_{01}) are convenient for Monte Carlo simulation. However, these are replaced by the unconditional probability of a wet day (π) and a dependence parameter (d).

$$\pi = p_{01}/(1 + p_{01} - p_{11}) \tag{3}$$

$$d = p_{11} - p_{01} \tag{4}$$

Denoting the parameters for the changed conditions with primes, the ratios of the monthly means and variances result in:

$$\frac{\mu'}{\mu} = \frac{\pi'\alpha'\beta'}{\pi\alpha\beta} \tag{5}$$

$$\frac{\sigma'^2}{\sigma^2} = \frac{\pi'\alpha'\beta'^2\left[1+\alpha'(1-\pi')\dfrac{1+d'}{1-d'}\right]}{\pi\alpha\beta^2\left[1+\alpha(1-\pi)\dfrac{1+d}{1-d}\right]} \tag{6}$$

The left hand side of the above two equations has known monthly means and variances under the present and changed conditions. All the variables without prime on the right hand side are known as well. Hence, there are four unknowns (π', d', α', β') and two equations. Two additional constraints are required to solve the equations. The nature of these constraints will depend on other available information. One of the simplest form is to assume no change in the precipitation occurrence process so that the only changes are in the gamma distribution parameters. Bates et al. (1994) set $\pi' = \pi\pi_2/\pi_1$ and $d' = dd_2/d_1$ where the subscripts 1 and 2 denote the GCM values for a nearby GCM grid cell for control and doubled CO_2 runs respectively. In this study, this approach was used to estimate the gamma parameters α' and β' for the future climate conditions. From Equations (3) and (4), α' and β' are obtained from

$$\alpha' = \left(\frac{\mu'\sigma}{\mu\sigma'}\right)^2 \frac{\pi}{\pi'} \frac{\alpha}{(1+\alpha(1-\pi)D - \left(\frac{\mu'\sigma}{\mu\sigma'}\right)^2 \frac{\pi}{\pi'}\alpha(1-\pi')D'} \tag{7}$$

$$\beta' = \frac{\mu'\pi\alpha\beta}{\mu\pi'\alpha'} \qquad \text{where } D = \frac{1+d}{1-d} \quad \text{and} \quad D' = \frac{1+d'}{1-d'} \tag{8}$$

As for the annual and monthly cases, the daily cross correlations are not adjusted.

5. APPLICATION OF THE MODEL TO FUTURE CLIMATE CONDITIONS

The nested multisite two-part model was first calibrated using the daily rainfall data for the period 1981 – 2000. The calibrated model was used to generate 100 replicates, each of length 20 years, and a number of statistics were calculated at daily, monthly and annual levels. It can be seen from Figure 2 that the annual mean and standard deviation are preserved well, while the monthly mean and standard deviation are satisfactorily preserved (Figure 3), with some evidence of positive bias for the mean and underestimation of variability. The mean number of wet days, maximum, mean and standard deviation of daily rainfall are also satisfactorily preserved (Figure 4). However, there is some evidence of overestimation of wet daily mean and standard deviation.

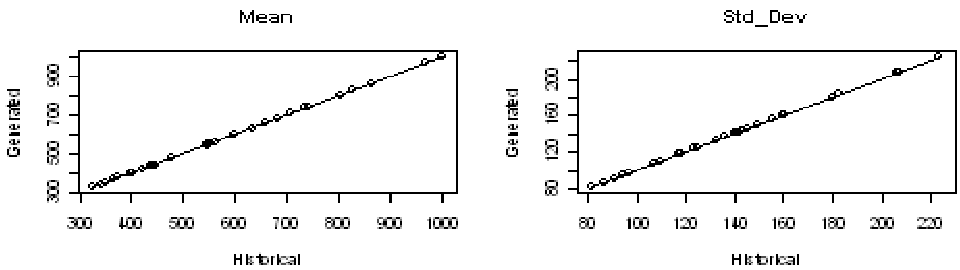

Figure 2. Comparison of annual mean and standard deviation for the current climate.

The ratios of the magnitude of the model parameters for the future to the current climate were estimated using the GCM rainfall data. The ratios appear to vary considerably between adjacent months. This might be due to short sample length (20 years) from which the ratios were estimated. It was decided to average the ratios within seasons to minimise the large variability among months. As before, 100 replicates, each of length 20 years, were generated and the above statistics were calculated. For comparison, these were plotted against the adjusted values of the parameters used in the data generation for the future climate conditions in Figures 5 – 7. These figures indicate a satisfactory performance of the model for the future climate.

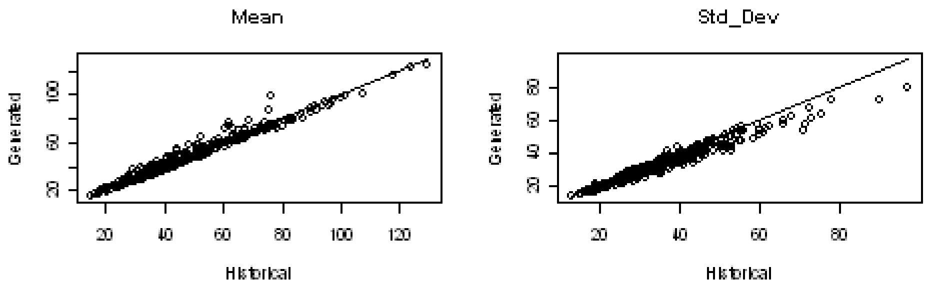

Figure 3. Comparison of monthly mean and standard deviation for the current climate.

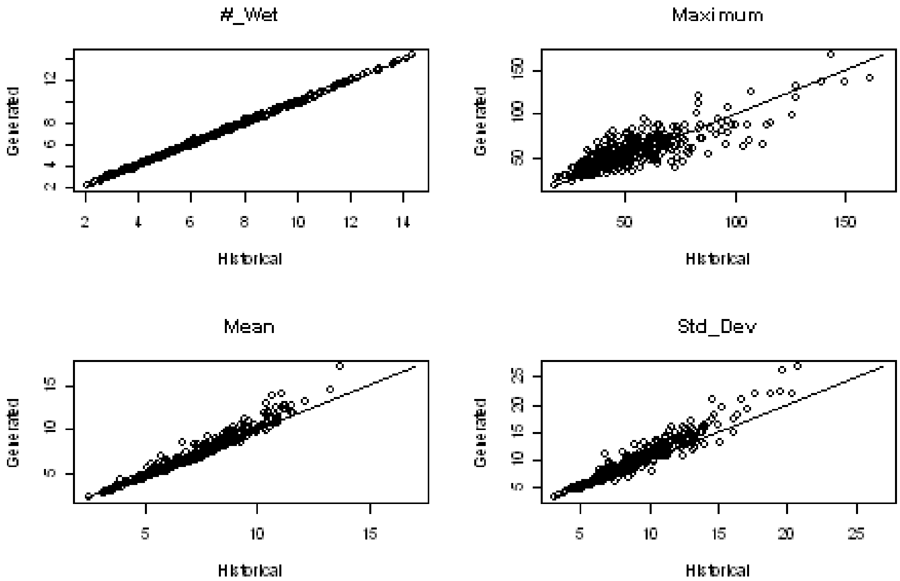

Figure 4. Comparison of mean monthly number of wet days, maximum, mean and standard deviation of daily rainfall for the current climate.

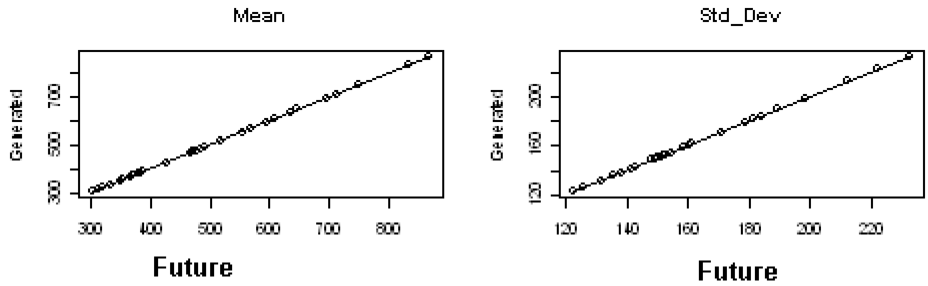

Figure 5. Comparison of annual mean and standard deviation for the future climate. The generated (y axis) refers to the generated data for the future climate conditions while the "Future" (x axis) refers to the modified parameters input to the data generation model

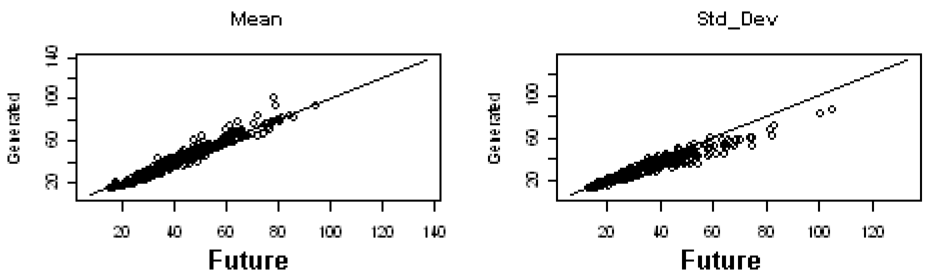

Figure 6. Comparison of monthly mean and standard deviation for the future climate.

6. CONCLUSIONS

The parameters of the nested multisite two-part model were modified for future climatic conditions using the rainfall data from GCM simulations. The parameters modified are the means and standard deviations at the monthly and annual time scale and the transition probabilities and gamma distribution parameters at the daily time scale. The cross correlations were not modified. The model was applied to 30 rainfall stations in the Murrumbidgee River catchment and found to perform satisfactorily. Work is in progress to improve the model in terms of preserving the cross correlations as well as to use upper air data from the GCM.

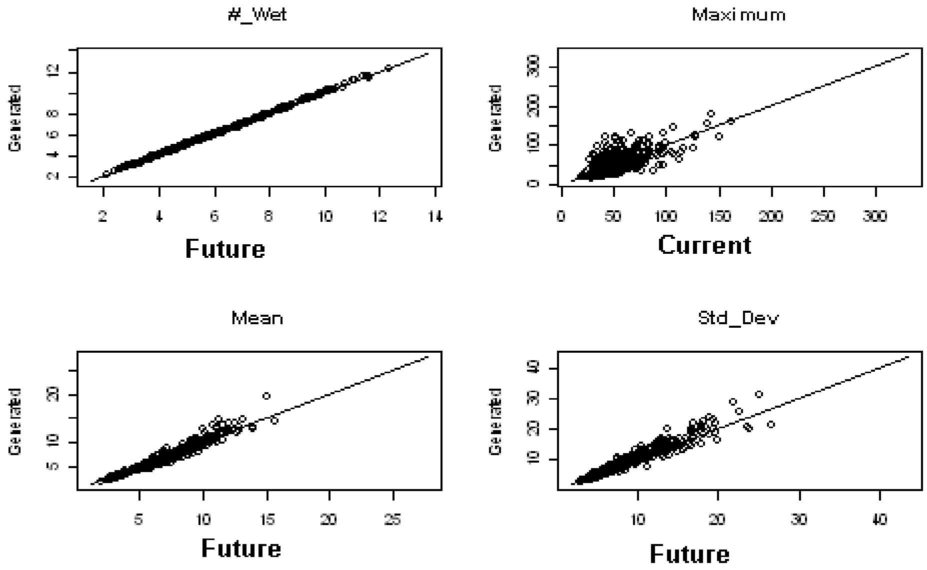

Figure 7. Comparison of mean monthly number of wet days, maximum, mean and standard deviation of daily rainfall for the future climate.

ACKNOWLEDGMENTS

The GCM rainfall data used in this paper was supplied by Dr Francis Chiew and Dr Ang Yang.

REFERENCES

Bates, B. C., Jakeman, A. J., Charles, S. P., Sumner, N. R. and Fleming, P. M., (1994) Impact of climate change on Australia’s surface water resources. In: *Greenhouse, Coping with climate change*, W. J.Bouma, G. I. Pearman and M. R. Manning, (Eds.), 248-262.

Chandler, R. E. (2002). GLIMCLIM: Generalized linear modelling for daily climate time series (software and user guide). Department of Statistical Science, University College London.

Charles, S. P., B. C. Bates, I. N. Smith, J. P. Hughes (2004). Statistical downscaling of daily precipitation from observed and modelled atmospheric fields. *Hydrological Processes*, 18(8): 1373-1394.

Chiew, F. H. S., Harrold, T. I., Siriwardena, L., Jones, R. N. and Srikanthan, R. (2003) Simulation of climate change impacts on runoff using rainfall scenarios that consider daily patterns of change in GCMs. MODSIM 2003, Townsville, 1, 154-159.

Frost, A. F, Charles, S. P., Mehrotra, R., Timbal, B., Nguyen, K. C., Chiew, F. H. S., Fu, G., Chandler, R. E., McGregor, J., Kirono, D., Fernandez, E., and Kent, D. (2009 - In preparation) A comparison of multi-site daily rainfall downscaling techniques under Australian conditions.

Harrold, T. I. and Jones, R. N. (2003) Generation of local scale scenarios using changes in GCM rainfall: a refinement of the perturbation method. MODSIM 2003, Townsville, 1, 11-16.

Harrold, T. I., Chiew, F. H. S. and Siriwardena, L. (2005) A method for estimating climate change impacts on mean and extreme rainfall and runoff.MODSIM 2005, Melbourne, 497-504.

Hughes, J. P., Guttorp, P. and Charles, S. P. (1999) A non-homogeneous hidden Markov model for precipitation occurrence. *Applied Statistics*, 48(1), 15-30.

Lettenmaier, D. P. and Gan, T. Y. (1990), Hydrologic sensitivities of the Sacramento-San Joaquin River basin, Califronia for global warming. *Water Resources Research*, 26(1): 69-86.

Panagoulia, D. (1992) Impacts of GISS-modelled climate changes on catchment hydrology. *Hydrological Sciences Journal*, **37**, 141-163.

Mehrotra, R. and A. Sharma (2005). A nonparametric nonhomogeneous hidden Markov model for downscaling of multisite daily rainfall occurrences. *Journal of Geophysical Research* 110(D16108).

Srikanthan, R. and Pegram, G. G. S. (2009) A nested multisite daily rainfall stochastic generation model. Accepted for publication in Journal of Hydrology.

Wilks, D. S. (1992) Adopting stochastic weather generation algorithms for climate change studies. *Climate Change*, **22**, 67-84.

Wilby, R. L., Dawson, C. W. and Barrow, E. M. (2002), SDSM - a decision support tool for the assessment of regional climate change impacts, *Environmental Modelling and Software*, 17(2), 145- 157.