Hydrologic Response of SWAT to Single Site and Multi-Site Daily Rainfall Generation Models

[1]Watson, B.M., [2]R. Srikanthan, [1]S. Selvalingam, and [1]M. Ghafouri

[1]School of Engineering and Technology, Deakin University, [2]Bureau of Meteorology and CRC for Catchment Hydrology, E-Mail: bmwatson@deakin.edu.au

Keywords: *Rainfall generation; stochastic model; spatial correlation; multi-site; SWAT.*

EXTENDED ABSTRACT

Stochastic models must be relied upon to generate the rainfall sequences needed to drive hydrologic models that are used to predict the impacts of land use change. It is well established that the quality of generated rainfall sequences directly affects the output of hydrologic models. Therefore, there is a need to evaluate the response of hydrologic models to different rainfall generation models. SWAT is a hydrologic model that is used extensively around the world by government agencies and catchment management authorities to predict the long-term impacts of land use change on the water yield and water quality of large, heterogeneous catchments. Given that SWAT is widely used for routine planning and decision making, it is critical that research be conducted to evaluate the impact that rainfall sequences generated by different stochastic models have on SWAT runoff predictions.

Single site stochastic models have been used exclusively to date for supplying rainfall sequences to hydrologic models. These models, however, do not preserve spatial correlations among multiple stations. Therefore, the resulting series at multiple sites are independent of each other (Qian *et al.*, 2002). However, in recent times there has been a push to use spatially correlated rainfall sequences as input to hydrologic models that are applied to large scale catchments. The spatially correlated rainfall sequences are generated using multi-site stochastic models. The aim of this study was to evaluate the response of SWAT to a single site model (DMAn) and a multi-site model (MS2P) to determine if using spatially correlated rainfall sequences as input to SWAT would result in the preservation of more statistical characteristics of the historical runoff than from using rainfall series generated independently of one another. The study area was the Woady Yaloak River catchment (306 km^2) located in Victoria.

The first stage of the evaluation procedure was to assess the ability of each model for preserving the statistical characteristics of the historical rainfall record at annual, monthly and daily time scales. It was found that there was little distinction between the performances of the models with respect to preserving the statistical characteristics of the historical rainfall record at annual, monthly and daily time scales. Both models preserved the mean and standard deviation of annual, monthly and daily rainfall satisfactorily. The models also preserved the lag-one autocorrelation coefficient of annual rainfall and the serial correlation of monthly rainfall. However, the models failed to reproduce the skew coefficient at all time scales.

The second stage of the evaluation procedure involved transforming the rainfall sequences generated by the models to daily runoff sequences using a calibrated version of SWAT and determining how well the statistical characteristics of the synthetic runoff series compared to the statistical characteristics of the runoff derived from the historical rainfall record. The mean annual runoff was slightly underestimated by the models. Both models underestimated the standard deviation of annual runoff, although the MS2P model did outperform the DMAn model for this statistic. The mean monthly and daily runoff was preserved reasonably well. In contrast, the standard deviation of the monthly and daily runoff was reproduced poorly by both models.

Overall, the results of this study indicate that there was little distinction between using rainfall sequences from a single site model and a multi-site model as input into SWAT. This was despite the rainfall sequences from the multi-site model being spatially correlated. It is strongly recommended that further studies be conducted to determine whether using spatially correlated synthetic rainfall series as input into distributed hydrologic models are significantly superior to using rainfall series that are generated independently of each other.

1. INTRODUCTION

The Soil and Water Assessment Tool (SWAT) is a hydrologic model that was developed to predict the impacts of land use change on the water yield and water quality of large, heterogeneous catchments (Neitsch *et al*., 2001). To assess the impacts that might occur as a result of changes in land use, stochastic models must be relied upon to provide sequences of rainfall that are statistically consistent with the historical record as input into SWAT.

To date, single site stochastic models have been used almost exclusively to provide the rainfall sequences needed to drive hydrologic models. A considerable amount of research has been carried out on the generation of daily rainfall at individual sites in the past few decades (Srikanthan and McMahon, 2001). This is because many single site models, such as those based on Markov chains, are easy to formulate and are based on a relatively simple stochastic process (Mehrota *et al*., 2005). It is mainly for this reason that single site models continue to be used extensively today.

However, single site models are not capable of reproducing the spatial correlation of rainfall events at multiple sites. The rainfall series generated by single site models at different sites are independent of each other, whereas in reality very strong spatial correlation exists in rainfall data (Qian *et al*., 2002). It is now widely accepted that multi-station rainfall simulation can only be achieved with stochastic models that preserve the spatial correlation between stations (Mehrota *et al*., 2005). Srikanthan and McMahon (2001) provide a comprehensive review of a number of multi-site models developed for daily rainfall generation. Only a limited amount of research has been conducted on multi-site models compared to single site models, although multi-site models have received much more attention of late.

It has been claimed that the need to simulate spatially correlated rainfall sequences over a region is very important when used as input into catchment scale hydrologic models (Wilks and Wilby, 1999; Srikanthan and McMahon, 2001; Harmel *et al*., 2002; Qian *et al*., 2002; Mehrota *et al*., 2005). For example, Qian *et al*. (2002) reported that:

> “It is necessary to preserve the spatial correlation in simulations of the weather series corresponding to certain climate scenarios as input to impact models, especially for hydrologic models, in which the spatial distribution of precipitation may have essential effects on the discharge of a river and the formation of floods.”

Srikanthan and McMahon (2001) also stated that:

> “If hydrological and land management changes are required simultaneously across larger regions, then the spatial dependence between the weather inputs at different sites have to be accommodated. This is particularly important to the simulation of rainfall, which displays the largest variability in time and space.”

Despite these claims, it has yet to be established whether using spatially correlated rainfall series generated by multi-site models are superior to using rainfall series generated independently at multiple sites by single site models for input into distributed hydrologic models. That is, it is still to be determined whether there is any clear distinction between the ability of single site and multi-site models to preserve characteristics of annual, monthly and daily runoff when used as input into distributed hydrologic models.

A number of studies have been reported in the literature comparing the response of hydrologic models to different single site stochastic models (Harmel *et al*., 2000; Siriwardena *et al*., 2002). However, the authors are not aware of any studies that have considered the response of a distributed hydrologic model to single site and multi-site models. Therefore, there is a need to quantify the impact of using spatially correlated rainfall sequences on the output of hydrologic models.

The aim of this study was to compare the hydrologic response of SWAT to a single-site rainfall generation model and a multi-site rainfall generation model. The models were first assessed on their ability to preserve the statistical characteristics of the historical rainfall. The generated rainfall sequences were then transformed to runoff sequences using a calibrated version of SWAT. The statistical characteristics of the runoff sequences derived using the synthetic rainfall series were compared to the statistical characteristics of the runoff sequence derived from the historical rainfall using SWAT.

2. RAINFALL GENERATION MODELS

The DMAn and MS2P models are both two part models. These types of models are comprised of two components: (1) a model of rainfall occurrence which provides a sequence of dry and wet days; and (2) a model of rainfall amounts which determines the amount of rainfall on a wet day. A brief description of the models is provided below. Due to space limitations, equations are not included and readers are referred to the literature cited for complete details of each model.

2.1. Single site model

The single site model (DMAn) was developed by Srikanthan (2004). The occurrence of rainfall is simulated using a two-state first-order Markov chain. A day can be either dry or wet (two states) with the probability of rainfall occurrence depending on whether the previous day was wet or dry (first order). The amount of rainfall on a wet day is simulated using a gamma distribution. The daily rainfall amount model is nested in the monthly Thomas-Fiering model, which is in turn nested in a first order autoregressive annual model. Most two-part models are not capable of preserving the monthly and annual characteristics of rainfall. However, Srikanthan (2004) showed that nesting a daily model in monthly and annual models enables the monthly and annual characteristics to be preserved.

2.2. Multi-site model

The multi-site model (MS2P) was developed by Srikanthan (2005). It is based on the model originally developed by Wilks (1998), who extended the familiar two-part model to generate rainfall simultaneously at multiple sites by driving a collection of individual models with serially-independent but spatially-correlated random numbers. As with the DMAn model, the MS2P model utilises a two-state Markov chain for rainfall occurrences and a gamma distribution for rainfall amounts. The daily rainfall amount model has also been nested in monthly and annual models.

Individual models are fitted to each of the sites first. Given a network of N stations, there are $N(N-1)/2$ pairwise correlations that should be maintained in the rainfall occurrence process. This is achieved by using correlated uniform variates derived from standard Gaussian variates. The spatial correlation in the daily rainfall amounts is preserved by using a vector of correlated uniform variates (Srikanthan, 2005).

3. HYDROLOGIC MODEL

SWAT is a physically-based, semi-distributed hydrologic model that operates continuously on a daily time step. It was developed specifically to predict the water yield and water quality of large-scale, heterogeneous catchments. SWAT is a long-term yield model that is capable of simulating a number of different physical processes that occur in a catchment including hydrology, weather, sedimentation, soil temperature, crop growth, nutrients, pesticides and agricultural management.

SWAT divides a catchment into any number of subcatchments, which are further divided into Hydrologic Response Units (HRUs). HRUs are lumped land areas that are composed of unique land use and soil combinations. The water balance is initially calculated for each HRU. The output from a subcatchment is the sum of outputs from all HRUs defined within that subcatchment. Finally, the outputs from the subcatchments are routed through the channel network.

A complete description of SWAT can be found in Neitsch *et al*. (2001).

4. STUDY AREA AND DATA

The study area was the Woady Yaloak River catchment (306 km^2) in southwest Victoria, Australia (Figure 1). The main land use categories are grazing livestock, cereal crops, eucalyptus forests and pine plantations. Soils throughout the catchment are predominantly duplex. The climate is highly seasonal with cold, wet winters and hot, dry summers.

The Woady Yaloak River catchment was subdivided into six subcatchments and 19 HRUs. The three rainfall stations shown in Figure 1 were used to calibrate and validate the model for the periods 1978-1989 and 1990-2001 respectively. SWAT assigns to a given subcatchment the rainfall from the station closest to the centroid of that subcatchment. Rainfall from Ballarat was assigned to subcatchments 1 and 6, rainfall from Rokewood was assigned to subcatchments 2, 4 and 5, while rainfall from Skipton was assigned to subcatchment 3. The performance of SWAT for predicting runoff is given in Watson *et al*. (2004).

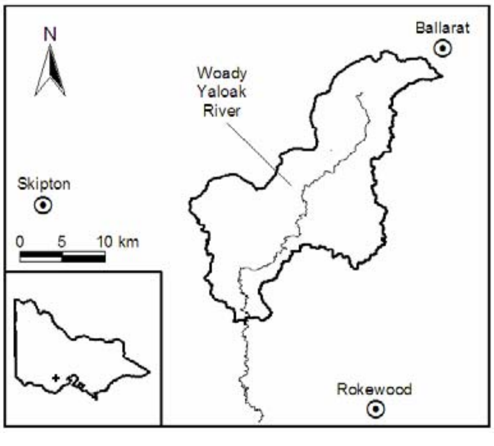

Figure 1. Location of study area.

5. RESULTS AND DISCUSSION

5.1. Rainfall

Continuous rainfall data from 1919 to 2001 (83 years) was obtained from the three stations shown in Figure 1. One hundred replicates of daily rainfall, each being 83 years in length, were generated at each station using the DMAn and MS2P models. Statistical characteristics were estimated for each of the 100 replicates and averaged. The models were evaluated by comparing the statistics of the historical and generated rainfall at annual, monthly and daily time scales including the first three moments (mean, standard deviation and skew coefficient). Due to space constraints only results from Ballarat are presented. The results for Ballarat are representative of the results for the other stations. Statistics of the historical and generated annual rainfall at Ballarat are presented in Table 1.

Table 1. Statistics of annual rainfall.

Statistic	Hist	DMAn	MS2P
Mean (mm)	702	703	705
Std. dev. (mm)	132	131	133
Skew coefficient	-0.02	0.25	0.32
R1	-0.12	-0.11	-0.12
No. wet days	169	169	169

The mean and standard deviation of the annual rainfall was preserved by the models. The lag one autocorrelation coefficient (R1) was also reproduced satisfactorily. However, both models substantially overestimated the skew coefficient. The number of wet days per year was reproduced by the models.

Figures 2 to 5 show the mean, standard deviation, skew coefficient and serial correlation coefficient of the monthly rainfall at Ballarat. The serial correlation coefficient represents the correlation between the rainfall amounts for two consecutive months. Both models preserved the mean, standard deviation and serial correlation coefficient. However, neither model managed to reproduce the skew coefficient.

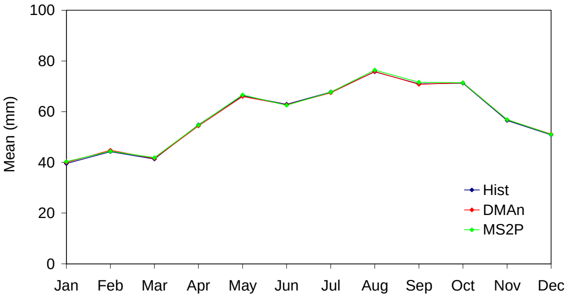

Figure 2. Mean monthly rainfall.

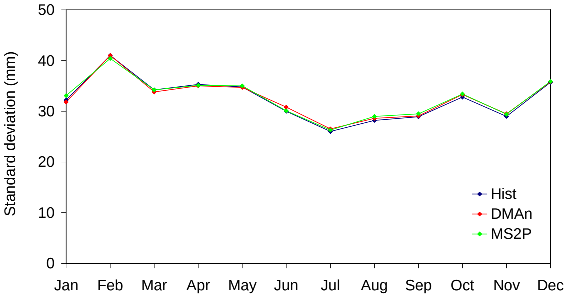

Figure 3. Standard deviation of monthly rainfall.

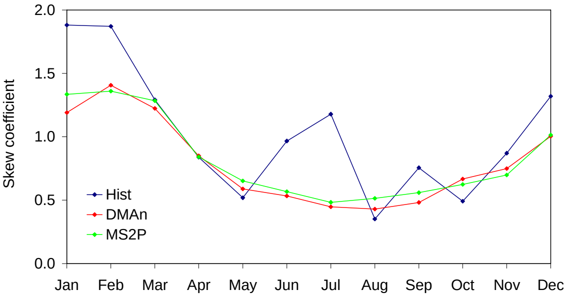

Figure 4. Skew coefficient of monthly rainfall.

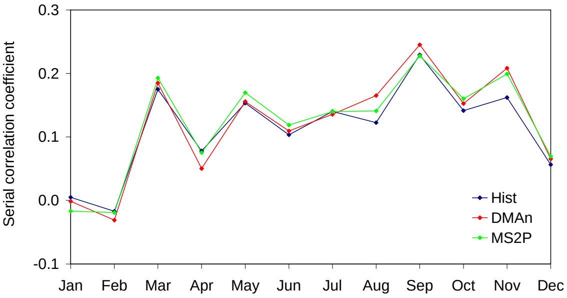

Figure 5. Serial correlation coefficient of monthly rainfall.

The mean, standard deviation and skew coefficient of the daily rainfall at Ballarat are presented in Figures 6 to 8. The models reproduced the mean and standard deviation of the daily rainfall, but failed to preserve the skew coefficient adequately.

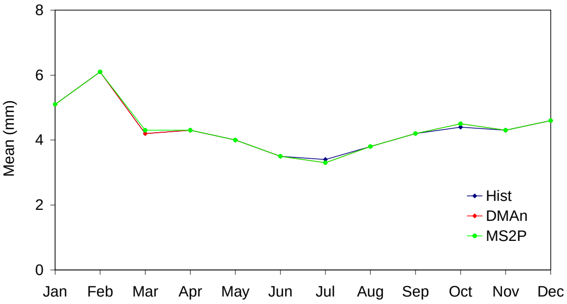

Figure 6. Mean daily rainfall.

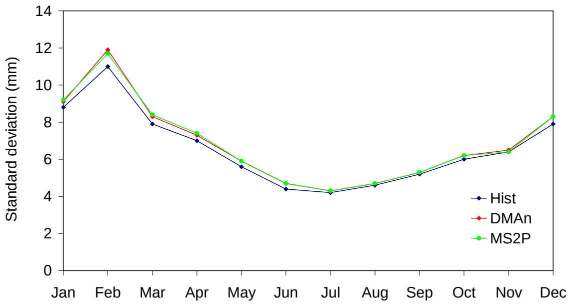

Figure 7. Standard deviation of daily rainfall.

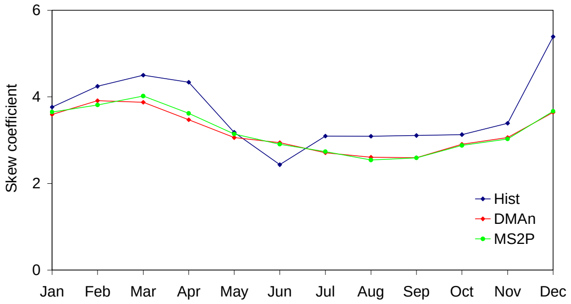

Figure 8. Skew coefficient of daily rainfall.

Since the MS2P model is a multi-site model, it is important to determine if the cross correlations between the Ballarat, Rokewood and Skipton stations are preserved. The cross correlations between the rainfall occurrences and rainfall amounts at sites *i* and *j* are obtained from:

$$r^{ij} = \frac{1}{(n-1)s^i s^j}\sum\left(x_t^i - \bar{x}^i\right)\left(x_t^j - \bar{x}^j\right) \qquad (1)$$

where x_t is the daily rainfall occurrence or daily rainfall amount, n is the number of data values, $\bar{x}$ is the mean and s is the standard deviation.

The cross correlations between daily rainfall occurrences are presented in Figure 9 while the cross correlations between daily rainfall amounts are presented in Figure 10. It can be observed that the cross correlations between daily rainfall occurrences were preserved exceptionally well. The cross correlations between daily rainfall amounts were preserved reasonably well, although the model did have a tendency to slightly underestimate the historical values. Although not shown here, the cross correlations between monthly and annual rainfall amounts were underestimated by the model. These results are consistent with the findings of Srikanthan (2005). It is important to note that the cross correlations were also calculated for the DMAn model, but they were found to be close to zero at each time scale.

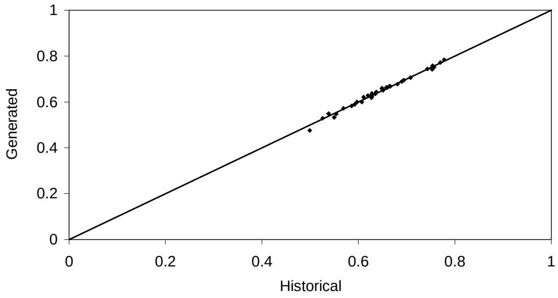

Figure 9. Cross correlations between daily rainfall occurrences for the MS2P model.

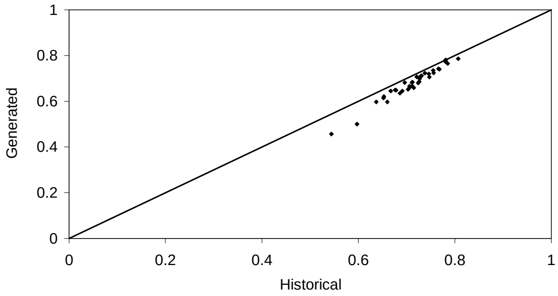

Figure 10. Cross correlations between daily rainfall amounts for the MS2P model.

5.2. Runoff

The calibrated version of SWAT was used to derive 100 replicates of runoff from the rainfall sequences generated by the DMAn and MS2P models. A sequence of daily runoff was also derived with SWAT using the historical rainfall data for the period of record. This sequence is referred to as the historical runoff. It is important to point out that the sequence referred to as the historical runoff is not the observed runoff. Using the runoff derived from historical rainfall provided a reference against which the performances of the rainfall generation models could be assessed (Siriwardena *et al.*, 2002). The historical and generated runoff sequences were 83 years in length. The rainfall generation models were evaluated by comparing the mean and standard deviation of the historical and generated runoff at annual, monthly and daily time scales. Values for the mean and standard deviation of the generated runoff were the average values of the 100 replicates.

The mean and standard deviation of the annual and monthly runoff are presented in Tables 2 and 3 respectively.

Table 2. Mean annual and monthly runoff.

	Mean (mm)		
Month	**Hist**	**DMAn**	**MS2P**
Jan	1.8	1.5	1.6
Feb	1.8	1.3	1.2
Mar	0.9	0.8	0.7
Apr	0.7	0.8	0.9
May	1.6	1.4	1.5
Jun	2.7	2.8	2.8
Jul	5.6	5.5	5.5
Aug	9.4	9.2	9.3
Sep	10.5	10.5	10.6
Oct	8.8	9.1	9.2
Nov	5.2	4.8	4.7
Dec	3.3	2.8	2.8
Annual	52.1	50.6	50.8

Table 3. Standard deviation of annual and monthly runoff.

	Standard deviation (mm)		
Month	**Hist**	**DMAn**	**MS2P**
Jan	2.1	1.2	1.6
Feb	8.3	2.8	2.9
Mar	2.6	1.4	1.3
Apr	0.9	1.6	2.1
May	3.8	3.0	3.3
Jun	4.6	3.6	4.1
Jul	5.9	4.4	5.2
Aug	7.2	5.3	6.7
Sep	8.4	6.2	7.7
Oct	7.3	7.1	8.4
Nov	5.9	4.1	4.2
Dec	4.9	2.4	2.4
Annual	31.7	23.2	29.2

The mean annual runoff was marginally underestimated by both models. The MS2P model slightly underestimated the standard deviation of annual runoff, whereas the DMAn model underestimated this statistic substantially. The mean monthly runoff was preserved reasonably well by both models, with only small differences between the historical and generated runoff values being observed for most months. There was little distinction between the models with respect to their performance for preserving the standard deviation of monthly runoff, with this statistic generally being underestimated by the models for most months. Harmel *et al*. (2000) also found that there were large errors between the standard deviation of the monthly historical runoff and the standard deviation of the monthly runoff derived with SWAT using rainfall sequences produced by three different daily rainfall generation models.

Figures 11 and 12 show the mean and standard deviation of the daily runoff for each month. Both models managed to preserve the mean daily runoff reasonably well. The standard deviation of daily runoff was not reproduced satisfactorily by either model over the entire year.

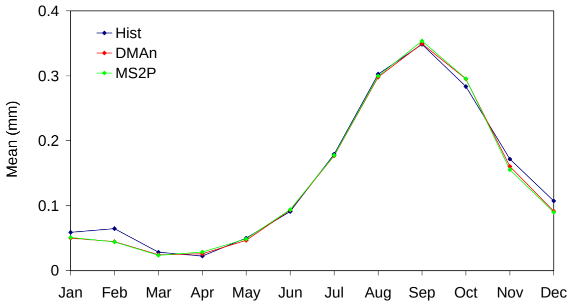

Figure 11. Mean daily runoff.

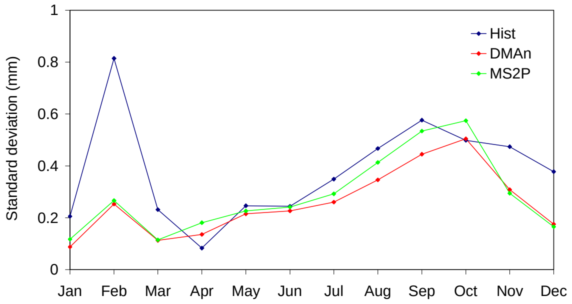

Figure 12. Standard deviation of daily runoff.

It is acknowledged that the density of rainfall stations used in this study was low. However, as these were the only stations from which continuous long-term rainfall data was available little else could be done. It is important to point out that this situation is typical for most regions across Australia. Although this study provided important insights into the response of a distributed hydrologic model to a single site model and a multi-site model, a considerable amount of research still needs to be conducted to compare differences between using single site and multi-site models for supplying rainfall sequences to hydrologic models. Therefore, plans are currently underway to apply the single site and multi-site models used in this study to several USDA-ARS research catchments in the United States, where SWAT has already been applied as part of other hydrologic studies (J. Arnold, USDA, pers. comm., 2005). The spatial variability of rainfall over these catchments is represented extremely well because there are 20-30 rainfall stations located in and around the catchments. This will provide a stringent test of model performance in terms of their ability to preserve the statistical characteristics of the historical rainfall and runoff.

6. CONCLUSIONS

There was little distinction between the performances of the DMAn and MS2P models with respect to preserving the statistical characteristics of the historical rainfall at all three rainfall stations. The mean and standard deviation of annual, monthly and daily rainfall was preserved by both models. The models managed to preserve the lag one autocorrelation coefficient of annual rainfall and the serial correlation coefficient of monthly rainfall satisfactorily. However, the skew coefficient was not preserved by either model at any time scale. Although a variety of statistics can be employed to evaluate stochastic models, it is impossible to satisfactorily fulfill all criteria at each time scale. Despite the skew coefficient not being preserved adequately, this is not regarded as a major drawback because it is considered more important to preserve the first two moments (mean and standard deviation) of the data (Siriwardena *et al*., 2002). The MS2P model was also shown to preserve the daily spatial cross correlations. However, the spatial cross correlations at monthly and annual time scales were underestimated.

The mean annual runoff was slightly underestimated by both models. The models managed to reproduce the mean monthly and daily runoff reasonably well. The MS2P model marginally underestimated the standard deviation of annual runoff, whereas the DMAn model underestimated this statistic considerably. The standard deviation of the monthly and daily runoff was preserved poorly by both models.

Overall, the results of this study indicate that there was little distinction between using rainfall sequences from a single site model and a multi-site model as input into SWAT. Despite the rainfall sequences from the MS2P model being spatially correlated, it had little impact when they were transformed to runoff sequences using the calibrated version of SWAT. It is strongly recommended that further testing be carried out to determine whether using spatially correlated synthetic rainfall series as input into distributed hydrologic models offer any significant advantages over using rainfall series that are generated independently of each other.

7. ACKNOWLEDGMENTS

The first author acknowledges the support of an Australian Postgraduate Award Scholarship.

8. REFERENCES

Harmel, R.D., C.W. Richardson, and K.W. King. (2000), Hydrologic response of a small watershed model to generated precipitation, *Transactions of the ASAE*, 43, 1483–1488.

Harmel, R.D., G. Johnson, and C.W. Richardson. (2002), The GEM experience: Weather generator technology development in the USDA, *Bulletin of the American Meteorological Society*, 83, 954-957.

Mehrotra, R., R. Srikanthan, and A. Sharma. (2005), Comparison of three approaches for stochastic simulation of multi-site precipitation occurrence, *20th Hydrology and Water Resources Symposium*, Canberra.

Neitsch, S.L., J.G. Arnold, J.R. Kiniry, and J.R. Williams. (2001), *Soil and Water Assessment Tool Theoretical Documentation*, Grassland Soil and Water Research Laboratory and Blackland Research Center, Texas.

Qian, B., J. Corte-Real, and H. Xu. (2002), Multisite stochastic weather models for impact studies, *International Journal of Climatology*, 22, 1377-1397.

Siriwardena, L., R. Srikanthan, and T.A. McMahon. (2002), *Evaluation of two daily rainfall data generation models*, Technical Report 02/14, CRC for Catchment Hydrology.

Srikanthan, R. (2004), Stochastic generation of daily rainfall data using a nested model, *57th Canadian Water Resources Association Annual Congress*, Montreal.

Srikanthan, R. (2005), *Stochastic generation of daily rainfall at a number of sites*, Technical Report 05/7, CRC for Catchment Hydrology.

Srikanthan, R., and T.A. McMahon. (2001), Stochastic generation of annual, monthly and daily climate data: A review, *Hydrology and Earth System Sciences*, 5, 653-670.

Watson, B.M., S. Selvalingam, and M. Ghafouri. (2004), Modelling saturation excess runoff with SWAT, *8th National Conference on Hydraulics in Water Engineering*, Queensland.

Wilks, D.S. (1998), Multisite generalization of a daily stochastic precipitation generation model, *Journal of Hydrology*, 210, 178-191.

Wilks, D.S., and R.L. Wilby. (1999), The weather generation game: A review of stochastic weather models, *Progress in Physical Geography*, 23, 329-357.