USE OF SPLIT SAMPLE APPROACH IN EVALUATION OF STOCHASTIC DAILY RAINFALL GENERATION MODELS

(Date received: 6.11.2006)

Aminuddin bin Mohd. Baki

Faculty of Civil Engineering, Universiti Teknologi MARA, 40450 UiTM, Shah Alam, Selangor
Email: aminbaki2@gmail.com

ABSTRACT

This paper reviews the use of split sample approach to test the ability of stochastic daily rainfall generation model to generate rainfall data for the future. The catchment adopted is Kangaroo Valley in New South Wales, Australia. Total data of 101 years long are divided into two sets: Earlier Period (80 years) and Later Period (the subsequent 21 years). The model adopted is the 8x8 Transition Probability Matrices Model, using two variations the Shifted Exponential Distribution and Box-Cox Power Transformation for the eighth class. Model parameters including transition probability matrices, exponential distribution parameters and Box-Cox Power Distribution parameters were computed using the data from the Earlier Period. The comparisons of statistical measures were made against the Later Period. Comparisons were made using daily statistical measures, daily extremes, monthly statistical measures, monthly extremes, annual statistical measures, annual extremes and serial correlation coefficients. The results shown that in general satisfactory statistical comparisons were made between the generated data based on Earlier Period against the statistics of the Later Period. In conclusion, the stochastic daily rainfall generation model can be used to generate synthetic data for planning and forecasting.

Keywords: *Split Sample, Stochastic Daily Rainfall Generation, Transition Probability Matrices*

1. INTRODUCTION

Hydrological forecasting generally required long-term data [1]. Unfortunately, in some cases recorded data of sufficiently long duration is not available. One of the methods to overcome this problem of short duration is to use stochastic data generation. To use the method of stochastic data generation, analysis of the hydrologic time series needs to be conducted. The hydrologic time series consists of two contributing factors: random factors and persistence (stochastically deterministic factor). Stochastic modelling used the stochastic properties of observed time series to generate synthetic long-term time series. The statistical and stochastic properties of the observed time series are assumed to represent the population properties, and the synthetic long-term time series are assumed to come from the same population [2].

Rainfall is regarded as the most basic weather variable, independent of temperature and evaporation [3]. Thus, generation of long-term synthetic rainfall data can provide basic set of weather variable for long term forecasting [4]. Rainfall is the key input variable which activates flow and mass transport in hydrological system for simulating and forecasting rainfall in space and time can play an important role in enhancing understanding of hydrological system response, and in the design and operation of water resource. Long term historical records of hydrological information such as rainfall and runoff data form the basis of planning and design of major water resources projects [5] Detailed knowledge of rainfall characteristics in ones catchments is essential for improvement in planning and design of drainage networks.

There are many studies conducted in generating stochastic daily rainfall data [6]. However, most of the statistical comparisons were made against the data used to compute the model parameters. Thus, those comparisons evaluated the ability of the model to re-produce the stochastic properties of the daily rainfall time series by assuming that both the recorded and generated rainfall are samples from the rainfall data population time series.

In this study however, a split sample approach will be utilised to test the ability of stochastic daily rainfall data generation model to generate data for the future. The significance of the findings of this study would be to support the use of stochastic modelling, which in turn can be very useful in cases of scarcity of data. Stochastic modelling can generate synthetic data for planning and forecasting purposes.

2. CATCHMENT DESCRIPTION

The catchment selected for this study is Kangaroo Valley, which is located about 150km south of Sydney, and about 50km west of the east coast of New South Wales, Australia. The map is shown

Table 1: Catchment characteristics [7]

CHARACTERISTICS	MEASURES
National Index	215220
Area	330 km^2
Stream Length	34.5 km
Average Slope	1.35%
Annual Rainfall	1637.0 mm
Annual Runoff	934.2 mm
Annual Pan Evaporation	1773.4 mm
Climate	Temperate
Vegetation	Rainforest, Hedgeland, Sedgeland, Grassland

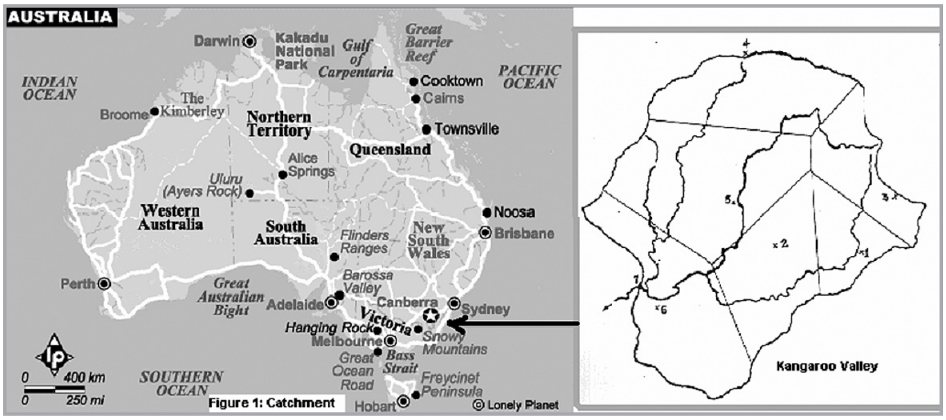

Figure 1: Catchment

in Figure 1 and catchment characteristics are shown in Table 1. Data from 1890 to 1990 will be used in this study, from 1890 to 1969 (80 years) for the Earlier Period and from 1970 to 1990 (the subsequent 21 years) for the Later Period.

3. TRANSITION PROBABILITY MATRICES MODEL

Haan *et al.* [8] used a multi-state Markov chain approach to model the distribution of rainfall. They used seven states to describe rainfall behaviour based on rainfall depths. The first state is dry (no rain) and six others are wet (with rainfall). Uniform distributions were assumed for states 2 to 6, and a shifted exponential distribution for the seventh state (unbounded).

Haan *et al.* [8] mentioned that persistence and periodicities could be observed in daily weather patterns. The persistence is modelled by a Markov chain.

Consider the following [8]:

$$P(E_{nj}|E_{n-1jn-1},\ldots, E_{1j1}) = P(E_{nj}|E_{n-1jn-1}) \tag{1}$$

where for x_1, x_2, … as the observations of daily rainfall, then $E_{i,j}$ ($i = 1, 2,\ldots c$, and $j = 0, 1,\ldots, c$), where c is the number of classes or states, and if $P(E_{nj}|E_{n-1j})$ does not depend on n, then these transition probabilities (denoted P_{ij}), and the Markov chain is stationary. The Transition Probability Matrices (TPM) is the collection of P_{ij} between classes in (c + 1) x (c + 1) matrices.

Periodicities mean that the weather pattern undergoes a cyclical behaviour within a year. Within a season, the weather pattern can be assumed to be stationary. Therefore, the TPM can be assumed to be stationary within each season:

$$(P_{ij}(k))(i, j = 0, 1,\ldots, c \text{ and } k = 1,\ldots, s) \tag{2}$$

where k denotes the k^{th} season and s is the total number of seasons.

The probability distributions had to be fitted to each class. It was assumed that the same set of distributions would model each season. Therefore, (c + 1) cumulative distribution functions are used:

$$F_m(x)\ (m = 0,\ldots, c) \tag{3}$$

where $F_m(x) = P$ (rainfall < x | rainfall belongs to class m).

A uniform distribution was assumed for all wet states, except for the last one. For the highest class, a shifted exponential distribution was found to be the most suitable [8]:

$$F_{last}(x) = e^{(x - ncl)/\eta} \tag{4}$$

where ncl is the lower boundary of the last class and η is a constant found by maximum likelihood:

$$\eta = \bar{x} - ncl \tag{5}$$

where $\bar{x}$ is the mean daily rainfall greater than ncl.

Haan *et al.* [8] adopted the months to be the seasons. Seasons follow an annual cycle, and by using months to represent seasons, the cyclical pattern can be satisfactorily represented. Hence, the TPM can be assumed to be stationary within a month. They also adopted 7 classes of daily rainfall states after testing up to 12 classes. These values were found to be satisfactory for the Kentucky basin. Therefore, twelve sets of (7x7) matrices needed to be found from the recorded data.

Baki [13] tested six variations of the TPM model: 6x6 TPM (called SE6), 7x7 TPM (called SE7) and 8x8 TPM (called SE8), all three with shifted exponential distribution for the last class and linear distribution for the other classes, and 6x6 TPM (called BC6), 7x7 TPM (called BC7) and 8x8 TPM (called BC8), all three with Box-Cox Power transformation for the last class and linear distribution for the other classes. The last (highest) class has closed lower bound and open higher bound. The class boundaries are shown in Table 2.

Srikanthan and McMahon [9] developed a modified TPM model based on the TPM Model of Haan *et al.* [8]. The exception was that the daily rainfall data is transformed using the Box-Cox Power Transformation [10] instead of a shifted exponential distribution for the last class. Srikanthan and McMahon [11] used TPM Model in their development of automatic evaluation of stochastically generated rainfall data. Srikhathan *et al.* [12] also used TPM Model in their comparison of daily rainfall data generation models. Baki [13] found that in general, all six variations used (three sets of matrices using Shifted Exponential and three sets of matrices using Box-Cox power transformation) were equally satisfactory as the differences between the six variations are minimal. This was consistent with past research as Haan *et al.* [8] found that the number of classes did not affect the accuracy of the TPM Model to a great extent. Therefore, the selection between the six variations is not very critical. In overall considerations, the transition probability matrices model (TPM) is proven to be satisfactory. Thus, this study adopted two variations of the TPM Model, namely the 8x8 TPM with Shifted Exponential Distribution and Box-Cox Power Transformation (referred to as SE8 and BC8, respectively).

Table 2: Class boundaries for TPM model [8]

CLASS	LOWER LIMIT (mm)	UPPER LIMIT (mm)		
		6x6	7x7	8x8
1	0.0	0.0	0.0	0.0
2	0.1	0.9	0.9	0.9
3	1.0	2.9	2.9	2.9
4	3.0	6.9	6.9	6.9
5	7.0	14.9	14.9	14.9
6	15.0		30.9	30.9
7	31.0 (for 7x7 & 8x8)	N/A		62.9
8	63.0 (for 8x8)	N/A	N/A	

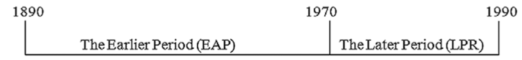

Figure 2: The Split of Data into the Earlier Period and the Later Period

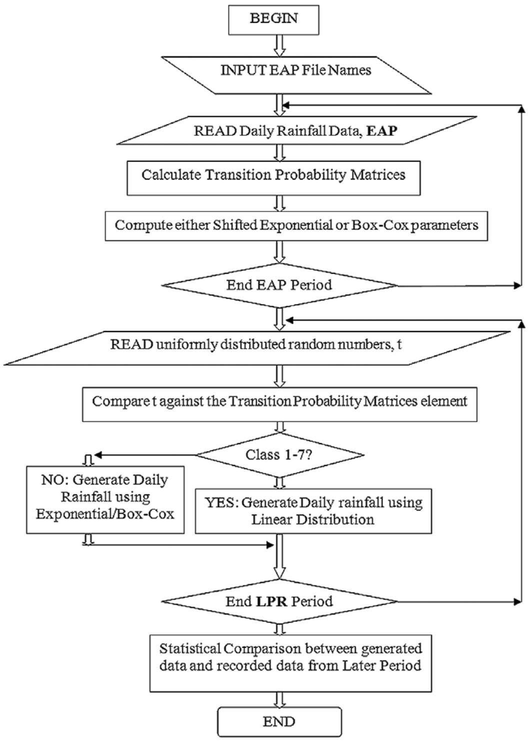

Figure 3: Flowchart of the split sample approach

4. METHODOLOGY OF THE SPLIT SAMPLE APPROACH

The split sample approach used in this study is to utilise data from the Earlier Period to compute the model parameters and then used these parameters to generate data for the Later Period. Figure 2 shows the data split into two periods: the Earlier Period (EAP) and the Later Period (LPR). Initially, data from the Earlier Period (80 years) was used to compute the 8x8 transition probability matrices using maximum likelihood method. The parameters of the shifted exponential distribution and Box-Cox Power Transformation were also computed using class 8 data from the Earlier Period. Daily rainfall data generation were carried for the Later Period using those parameters based on the Earlier Period. Ten replicates of synthetic daily rainfall data, each with the length equals to the Later Period (21 years). Statistical comparisons will then be made between the generated data to those of the Later Period.

The Split Sample approach tested the ability of rainfall data generation models to generate data for the future (since the Later Period succeeds the Earlier Period). The comparison was made between the statistics of the two independent samples (generated data using parameters from EAP and the recorded data from LPR). Figure 3 shows the flowchart of the split sample approach.

5. RESULTS

Stochastic rainfall data generation model 8x8 TPM with shifted exponential distribution (SE8) and Box-Cox Power Transformation (BC8) used parameters developed in the Earlier Period to generated data for the Later Period. The recorded data from both Earlier Period (EAP) and Later Period (LPR) were also presented together with the results of the two variations of TPM model. Slight variations were observed between EAP and LPR due to sampling errors.

Ten replicates of data were generated for each TPM model variation, each with the length equal to LPR (21 years), even though the parameters were developed from EAP (80 years). The initial state of rainfall used was 1. Tables 2 to 6 show the comparison between the average of ten replicates of generated data of both variations of TPM model (SE8 and BC8) and statistics of the recorded data (EAP and LPR). Comparisons made include daily statistics and extremes (Table 3), monthly statistics (Table 4), monthly extremes (Table 5), annual statistics and extremes, and lag-one serial correlation coefficients (Table 6).

Table 3 illustrates that both SE8 and BC8 generated data, had daily means, and standard deviations that were reasonably close to the statistics of data from the Later Period. In terms of overall daily means, SE8 and BC8 are 4.6mm and 4.4mm, respectively compared to EAP and LPR of 4.4mm and 4.6mm, respectively. For overall daily standard deviations, SE8 and BC8 are 15.7mm and 14.9mm, respectively compared to EAP and LPR of 15.6mm and 15.4mm, respectively. For overall daily skews, SE8 and BC8

Table 3: Daily rainfall statistical comparisons

	Means (mm)				**Standard Deviations** (mm)				**Skews** (mm)				**Daily Maxima** (mm)			
MON	**EAP**	**LPR**	**SE8**	**BC8**	**EAP**	**LPR**	**SE8**	**BC8**	**EAP**	**LPR**	**SE8**	**BC8**	**EAP**	**LPR**	**SE8**	**BC8**
Jan	4.8	4.9	5.3	4.9	16.4	10.3	16.2	15.3	11.3	3.1	6.6	6.8	423.5	66.4	506.5	249.1
Feb	5.5	6.3	5.2	5.0	17.5	16.5	16.4	14.7	7.4	5.7	6.9	6.0	308.9	157.4	314.6	212.3
Mar	5.8	7.2	5.7	6.1	18.2	24.0	17.2	17.7	6.3	6.9	6.1	5.6	242.9	291.6	301.0	212.9
Apr	4.9	5.3	5.0	5.2	16.5	14.8	16.2	17.2	7.0	5.7	6.2	6.1	258.1	191.1	329.7	206.8
May	4.8	4.7	5.0	4.9	17.7	14.1	17.3	16.7	7.9	6.0	6.8	7.0	268.0	177.0	375.0	244.2
Jun	6.1	5.4	6.3	6.1	19.3	17.7	19.6	18.9	5.6	6.9	6.1	5.3	236.2	211.5	510.7	195.4
Jul	4.6	2.2	4.3	4.5	18.0	10.8	15.5	16.2	8.3	12.7	7.1	7.1	285.2	194.2	386.4	275.6
Aug	3.1	3.8	3.2	3.0	11.4	21.8	12.4	11.1	8.2	9.9	8.9	7.7	196.9	262.0	345.9	203.3
Sep	3.1	3.0	3.2	2.9	9.9	9.3	10.0	9.2	6.5	6.3	5.7	5.7	141.4	111.0	130.9	137.9
Oct	3.7	4.7	4.1	3.8	15.0	14.9	16.7	14.3	9.2	7.6	9.8	8.3	243.8	208.6	466.0	244.3
Nov	2.9	4.3	3.0	2.8	8.9	12.3	8.6	8.1	6.8	6.6	5.5	6.2	137.6	157.7	131.4	146.5
Dec	4.1	3.1	4.3	3.9	132	9.1	13.8	11.8	7.3	5.7	6.8	6.3	190.1	97.4	289.4	193.9
All	**4.4**	**4.6**	**4.6**	**4.4**	**15.6**	**15.4**	**15.7**	**14.9**	**8.1**	**8.5**	**8.4**	**7.1**	**423.5**	**291.6**	**510.7**	**275.6**

Table 4: Monthly statistical comparisons

	Means (mm)				**Standard Deviations** (mm)				**Skews** (mm)			
MON	**EAP**	**LPR**	**SE8**	**BC8**	**EAP**	**LPR**	**SE8**	**BC8**	**EAP**	**LPR**	**SE8**	**BC8**
Jan	148.9	152.3	162.9	152.3	140.9	114.7	123.3	119.9	2.3	1.1	1.0	1.3
Feb	161.1	178.7	147.5	140.3	162.4	135.2	129.5	129.5	2.1	0.8	1.5	1.6
Mar	181.0	221.7	175.7	188.7	181.6	167.1	146.0	161.9	1.8	1.0	1.5	1.1
Apr	146.6	159.7	151.5	155.8	125.6	154.2	132.2	147.5	1.7	1.5	1.3	1.5
May	147.5	144.7	155.8	150.6	177.6	125.5	157.7	137.1	2.9	1.1	1.5	1.5
Jun	183.8	162.4	188.9	194.6	194.6	164.5	184.6	167.5	1.6	1.5	1.3	1.0
Jul	142.7	67.8	133.0	140.0	146.0	65.0	139.3	132.3	1.9	1.7	1.6	1.1
Aug	96.6	119.1	98.6	92.1	107.6	174.9	114.3	96.7	2.0	1.7	2.0	1.7
Sep	92.0	88.6	63.4	87.6	75.1	77.8	79.9	79.4	1.4	0.9	0.9	1.3
Oct	114.6	146.6	127.9	117.5	137.9	146.7	124.0	127.0	3.1	1.2	1.8	1.9
Nov	88.3	129.9	89.7	82.6	81.3	94.0	74.1	65.4	2.4	1.2	1.6	1.6
Dec	126.2	95.3	134.5	122.0	118.0	83.4	130.6	106.3	2.0	1.4	2.0	1.6
All	**135.8**	**138.9**	**138.5**	**134.5**	**144.7**	**133.8**	**138.7**	**134.1**	**2.4**	**1.5**	**2.3**	**2.0**

Table 5: Monthly extreme comparisons

	Means (mm)				**Standard Deviations** (mm)			
MON	**EAP**	**LPR**	**SE8**	**BC8**	**EAP**	**LPR**	**SE8**	**BC8**
Jan	5.1	20.4	10.7	8.4	815.8	435.4	786.8	588.1
Feb	2.4	39.3	6.5	16.0	851.5	470.9	759.4	847.9
Mar	4.6	14.4	9.5	3.7	792.7	662.3	894.8	886.9
Apr	5.0	23.6	5.0	1.5	717.4	630.5	894.8	756.1
May	0.7	5.1	0.3	3.3	1079.9	489.3	967.5	957.8
Jun	1.8	16.2	1.2	0.3	982.2	576.4	1039.4	991.9
Jul	5.1	0.0	0.9	0.5	686.8	261.2	1287.1	748.6
Aug	1.7	5.3	0.5	1.0	551.4	543.3	742.0	541.2
Sep	4.6	11.9	3.8	2.6	387.0	248.9	399.4	420.9
Oct	3.4	2.6	9.5	6.6	828.4	471.6	637.7	1038.9
Nov	1.5	9.6	1.3	4.7	519.4	391.2	753.6	519.6
Dec	5.9	9.6	8.2	3.3	700.8	328.3	1041.9	724.8
All	**0.7**	**0.0**	**0.3**	**0.3**	**1079.9**	**662.3**	**1287.1**	**1038.9**

Table 6: Annual statistical comparisons and lag-one serial correlation coefficients

Measures	**EAP**	**LPR**	**SE8**	**BC8**
Average (mm)	1629.2	1666.8	1662.5	1613.5
Standard Deviations (mm)	515.0	537.2	456.8	487.4
Skew (mm)	0.5	0.7	0.5	0.5
Maximum (mm)	3103.1	2950.0	3197.8	3228.5
Minimum (mm)	684.9	902.5	621.6	275.6
Lag-one Serial Correlation Coefficient, r_1	0.436	0.465	0.393	0.405

are 8.4mm and 7.1mm, respectively compared to EAP and LPR of 8.1mm and 8.5mm, respectively. There are slight deviations in the monthly distribution of daily means, standard deviations and skews. Most of the deviations observed were in fact due to the difference between statistical measures of the Earlier and Later data (two separate samples from the same population). Nevertheless, for daily means, differences between generated and Later data were mostly less than 2 mm. For daily standard deviations, both models were generally satisfactory except for extreme values in March and August. For daily skews, most values were satisfactory, except for January and July where there are differences between the two (2) recorded data sets (EAP and LPR) which result in differences between generated and Later data. For daily maxima, most generated maxima are greater than the recorded values (both EAP and LPR) indicating ability of the model to simulate extreme events beyond the available record.

For the monthly statistical comparisons, both variations of TPM model are generally satisfactory compared to the Later data (shown in Table 4). For monthly means, SE8 and BC8 are 138.5mm and 134.5mm, respectively compared to EAP and LPR of 135.8mm and 138.9mm, respectively. For monthly standard deviations, SE8 and BC8 are 138.7mm and 134.1mm, respectively compared to EAP and LPR of 144.7mm and 133.8mm, respectively. For monthly skews, SE8 and BC8 are 2.3mm and 2.0mm, respectively compared to EAP and LPR of 2.4mm and 1.5mm, respectively. Satisfactory monthly distribution of monthly means was observed. However, comparisons of monthly distribution of monthly standard deviations indicate significant differences were observed on four occasions. Nevertheless for monthly skews, satisfactory values were generated.

For both monthly maxima and minima, both model variations generated greater maxima and lower minima (except for overall minimum), indicating ability to simulate extreme events beyond the available record (as illustrated in Table 5). Table 5 shows that for this approach to be satisfactory, the data needs to be stationary in terms of its statistical variations, as indicated by some variations in monthly statistics due to differences between the two (2) sets of recorded data.

In Table 6, annual statistics including average, standard deviations and skews for both TPM model are satisfactory compared to the Later Period. The average annual rainfall for EAP and LPR are 1629.2mm and 1666.8mm, respectively, compared to SE8 and BC8 of 1662.5mm and 1613.5mm, respectively. For the annual extremes, higher annual maximum and lower annual minimum were generated, indicating ability of model to generate extreme events beyond the available data. The serial correlation coefficients were also reasonably satisfactory generated by both TPM models (SE8 and BC8 with 0.393 and 0.405, respectively) compared to the recorded data (EAP and LPR with 0.436 and 0.469, respectively).

Generally, in comparing the two variations of the TPM models, the shifted exponential distribution seems to be slightly more satisfactory compared to the Box-Cox Power transformation (as shown in Tables 3 to 6). Nevertheless, the purpose of this study is to illustrate the ability of stochastic daily rainfall generation models to generate data for the future, not so much comparing between the two TPM model variations.

6. DISCUSSIONS

In general, the use of Earlier Period data to generate data for the Later Period shows satisfactory statistical comparisons for both variations of TPM model (SE8 and BC8). In conclusion, apart from a few slight differences, the daily rainfall data generation generally yielded satisfactory statistical comparisons between the generated data (using historical data from the Earlier Period) and the future data (Later Period). These findings supported the use of historical data statistics for stochastic long-term data generation. Stochastic data generation seem to be more accepted in the developed countries compared to developing countries. However, the advancement in computer technology in recent times has provided opportunities for much faster data generation to be made. This should encourage revisit of the stochastic data generation approach as potential tools in forecasting and planning. These stochastically generated data can then be utilised for planning or forecasting purposes in cases of limited data availability.

7. CONCLUSIONS

The study intended to illustrate the applicability of using stochastic rainfall data generation to generate rainfall data for the future by using the split sample approach. The statistical properties such as average, standard deviations and skews of the source daily rainfall data from (EAP) were reproduced well by the stochastic model used, namely the 8x8 TPM with Box-Cox Power Transformation (BC8) and 8x8 TPM with Shifted Exponential Distribution (SE8). These statistical measures were found to be comparable to the statistical measures of daily rainfall data in the subsequent period (LPR). Thus, the study has met its objectives of using the split sample approach to illustrate that stochastic daily rainfall data generation can be used to generate daily rainfall data. ■

REFERENCES

[1] D.P. Loucks, Water Resource Systems Planning and Analysis, Prentice-Hall, 1981.

[2] A.S. Goodman, Principles of Water Resources Planning, Prentice-Hall, 1984.

[3] J.W. Jones, R.D. Colwick, and E.D. Threadgill, "A Simulated Environmental Model of Temperature, Evaporation, Rainfall and Soil Moisture", Transaction of American Society of Agricultural Engineers, Vol.15(2), PP.366-372, 1972.

[4] P. Purcell, Design of Water Resources Systems, Thomas Telford, 2003.

[5] Achela, D, Fernando, K and Jayawardena, A.W., Generation and Forcasting of Monsoon Rainfall Data, Proc. 20th WEDC Conference of Affordable Water Supply and Sanitation, Colombo, Sri Lanka, 1994, pg 310-313, 1994.

[6] R. Srikanthan, and T.A. McMahon, "Stochastic of Annual, Monthly and Daily Climate Data: A Review", Hydrology and Earth System Sciences, Australia, Vol.5(4), 653-670, 2001.

[7] A. Baki, "Objective Functions in the Optimisation of Daily Rainfall-Runoff Modelling", *JURUTERA*: The Monthly Bulletin of The Institution of Engineers, Malaysia, No.9, pp.11-15, Sept. 1996.

[8] C.T. Haan, D.M. Allen and J.O. Street, "A Markov Chain Model of Daily Rainfall", Water Resources Research, Vol.12(3), pp.443-449, 1976.

[9] R. Srikanthan, and T.A. McMahon, "Stochastic Simulation of Daily Rainfall for Australian Stations", Transactions of the American Society of Agricultural Engineers, Vol.26, pp.754-759, 1983.

[10] G.E.P. Box and O.R. Cox, "The Analysis of Transformations", Journal of Royal Statistical Society, Series B, Vol.26(2), pp.211-252, 1964.

[11] R. Srikanthan, and T.A. McMahon, "Automatic Evaluation of Stochastically Generated Rainfall Data", Australian Journal of Water Resources, Vol.8(2), pp.195-201, 2005.

[12] R. Srikanthan, L. Siriwardena, and T.A. McMahon, "Comparison of Two Daily Rainfall Data Generation Models", Australian Journal of Water Resources, Vol.8(2), pp.203-212, 2005.

[13] A.B.M. Baki, "Stochastic Rainfall Data Generation Using Transition Probability Model", Proceedings of the Seventh Annual IEM Water Resources Colloquium 2005, The Institution of Engineers Malaysia, Petaling Jaya, Malaysia, pp.9-1 to 9-9, 18 June 2005.

PROFILE

ENGR. ASSOC. PROF. DR AMINUDDIN BIN MOHD BAKI, FIEM, P.ENG.

Engr. Aminuddin is currently an Associate Professor in Water Resources and Environmental Systems, Faculty of Civil Engineering, Universiti Teknologi MARA Malaysia in Shah Alam. He has his Executive Diploma in QM (with Distinction) from Universiti Teknologi Malaysia (UTM). He obtained his Bachelor of Engineering (Hons)/Bachelor of Commerce (Civil Engineering and Management Studies) from the University of Wollongong, Australia. He completed his MBA from Irish International University. He graduated with his PhD (Civil Engineering) from the University of Wollongong, Australia.