A DAILY RAINFALL GENERATING MODEL FOR WATER YIELD AND FLOOD STUDIES

W. C. Boughton

Report 99/9

June 1999

COOPERATIVE RESEARCH CENTRE FOR
CATCHMENT HYDROLOGY

Boughton, W.C. (Walter C.)

A daily rainfall generating model for water yield and flood studies.

Bibliography.
ISBN 1 876006 49 8.

1. Rain and rainfall – Australia – Measurement – Mathematical models. 2. Rainfall frequencies – Australia – Methodology. 3. Rainfall frequencies – Australia – Computer Programs. I. Cooperative Research Centre for Catchment Hydrology. II. Title. (Series: Report (Cooperative Research Centre for Catchment Hydrology); 99/9).

551.5770287

Keywords

Rainfall/Runoff Relationship
Modelling (Hydrological)
Precipitation (Atmospheric)
Floods and Flooding
Frequency Analysis
Design Data
Forecasting
Water Yields
Catchment Areas

PREFACE

A core project in the CRC's Flood Program is Project FL1: "Holistic approach to rainfall-based design flood estimation". Its aim is to reduce the uncertainty in design flood estimates by accounting for the interaction and joint probability of the different flood producing components, i.e. a holistic approach. This project has followed two different ways to account for the interdependence of the factors which affect the amount of runoff from rainfall (e.g. rainfall intensity, duration, pattern; antecedent wetness; baseflow). The first has used an approach which considered the joint probabilities of these factors directly.

The second approach has considered continuous rainfall-runoff modelling; part of which involved the generation of rainfall for input to the models. This report gives details of the methodology developed for this, and a case study showing its application.

The author of the report, Dr Walter Boughton, has contributed much to the CRC's Flood Program as an Honorary Research Fellow. It is a pleasure to acknowledge his important role in the program of research for this project.

Russell Mein
Program Leader – Flood Program

ABSTRACT

Daily rainfall records form a major hydrological data base in Australia, but the common 50-100 years of available record at a station do not give adequate information about long term risks of droughts or floods. Transition Probability Matrix models have been used in prior studies to generate long sequences of daily rainfalls, but the model most commonly used in Australia seriously under-estimates the variance of annual totals of rainfall. An extension of the basic TPM model is made to incorporate a variance adjustment parameter which enables easy calibration of the model to match the frequency distribution of annual totals in the available record. Annual maxima daily rainfalls are calibrated against estimates of 24-hour Probable Maximum Precipitation or similar daily rainfalls with a long average recurrence interval. The model is programmed for simple calibration of the 2 parameters and use on a personal computer.

TABLE OF CONTENTS

PREFACE **iii**

ABSTRACT **v**

1. INTRODUCTION **1**

2. DESCRIPTION OF THE MODEL **2**

- 2.1 The Transition Probability Matrix (TPM) Model 2
- 2.2 Generation within each state 4
- 2.3 Adjustment of the variance of annual totals 5
- 2.4 Calibration of the model 7

3. TEST DATA **8**

- 3.1 Boggy Creek data 8
- 3.2 Brisbane 8

4. RESULTS **8**

- 4.1 Run to run variation in DAYRNGEN 8
- 4.2 Variation in results of DAYRNGEN due to parameter A 9
- 4.3 Calibration of model to Boggy Creek data 10
  - (i) Calibration to annual rainfalls with DAYRNGEN 10
  - (ii) Calibration to annual maxima daily rainfalls with PMPRAIN 11
  - (iii) Statistics of actual and generated rainfalls 14
- 4.4 Calibration of the model to Brisbane data 15

5. DISCUSSION **19**

6. CONCLUSION **20**

7. REFERENCES **21**

1. INTRODUCTION

In most areas of Australia where hydrological design studies are undertaken, there is a substantial data base of daily rainfalls, commonly with 50 to 100 years of record. Daily rainfall records are usually much longer than associated streamflow records, and the use of daily rainfall-runoff models to extend streamflow records is well established. Rainfall-runoff modelling with daily data has been applied mainly to water yield and water supply studies; however, there is a lot of current interest in the use of these models for design flood estimation and flood forecasting work (e.g. see Boughton and Carroll, 1993, Srikanthan et al, 1994, Boughton and Hill, 1997).

Although 50 to 100 years of daily rainfall records are commonly available at many locations in Australia, this length of record gives very limited information about the risk of floods or droughts with average recurrence intervals greater than a few decades. There has been a steadily increasing interest in the stochastic generation of long sequences of daily rainfalls for periods ranging from 100 to 1,000,000 years in order to give a better definition of extreme droughts and floods. Srikanthan and McMahon (1985) reviewed daily rainfall generating models, and they describe the various approaches which have been reported in the literature. Their review is the best source of references about the topic. They tested three models with daily rainfall data from five stations (Melbourne, Mackay, Perth, Alice Springs and Darwin) and recommended the transition probability matrix model for generating daily rainfalls in Australia. Chapman (1994) tested 5 daily rainfall generating models with several methods of evaluating the model parameters, and reported that the Srikanthan-McMahon (TPM) model performed particularly well when calibrated with long rainfall records.

Boughton and Hill (1997) used a transition probability matrix model to generate 1,000,000 years of daily rainfalls as part of a design flood estimation procedure. They reported a problem in that the frequency distribution of annual totals of generated rainfall had a variance that was substantially less than the variance of annual totals of recorded rainfall, i.e. the severity of drought periods was substantially under-estimated. Subsequent testing confirmed that the common form of transition probability matrix model will reproduce the averages of daily, monthly and annual rainfalls, but in all tests the variance of the generated annual totals was substantially less than the actual variance.

This paper describes a daily rainfall generating model which overcomes the problem mentioned above and accurately reproduces the daily, monthly and annual statistics of rainfall.

2. DESCRIPTION OF THE MODEL

2.1 The Transition Probability Matrix (TPM) Model

The model used to generate synthetic sequences of daily rainfalls is a modification of a method described by Haan et al (1976) and modified for use in Australia by Srikanthan and McMahon (1985). The model is referred to as a transition probability matrix (TPM) or a multi-state Markov chain model.

Daily rainfalls are divided into a number of states as shown in Table 1.

Table 1. States of Daily Rainfall

State	Rain			
1	zero			
2	zero	< rain	<=	0.9 mm
3	0.9	< rain	<=	2.9 mm
4	2.9	< rain	<=	6.9 mm
5	6.9	< rain	<=	14.9 mm
6	14.9	< rain	(no upper limit)	

Srikanthan and McMahon (1985) used different numbers of states in different months of the year and at the 15 locations in Australia at which they tested their TPM model. The number of states that they used are shown in Table 2. In the present study, 6 states were used at both locations tested.

The probabilities for rain in one state to be followed by rain on the next day in the same or another state are collated into a matrix - the transition probability matrix (TPM) - using as many whole years of rainfall record as are available. Seasonality of rain is modelled by using 12 TPMs, one for each calendar month. The TPM for January on the Boggy Creek catchment is shown in Table 3. The values in Table 3 are probabilities of going to the next state on the following day from the current state of today's rain.

Table 2. Number of states for various stations (from Srikanthan & McMahon, 1985)

Station	J	F	M	A	M	J	J	A	S	O	N	D
Melbourne	6	6	6	6	6	6	6	6	6	6	6	6
Sydney	7	7	7	7	7	7	7	7	7	7	7	7
Monto	6	6	6	6	6	6	6	6	6	6	6	6
Cowra	6	6	6	6	6	6	6	6	6	6	6	6
Mackay	7	7	7	7	7	7	7	7	7	7	7	7
Brisbane	7	7	7	7	7	7	7	7	7	7	7	7
Darwin	7	7	7	7	3	2	2	2	3	7	7	7
Broome	7	7	7	3	3	3	3	3	3	3	3	4
Perth	6	6	6	6	6	6	6	6	6	6	6	6
Adelaide	6	6	6	6	6	6	6	6	6	6	6	6
Alice Springs	4	4	4	4	4	4	4	4	4	4	4	4
Kalgoorlie	5	5	5	5	5	5	5	5	5	5	5	5
Onslow	4	4	4	3	4	3	4	3	3	3	3	3
Bamboo Springs	6	6	6	5	5	5	2	2	2	2	2	5
Lerderderg catchment	6	6	6	6	6	6	6	6	6	6	6	6

Table 3. Transition Probability Matrix for January

Current State	Next state					
	1	2	3	4	5	6
1	0.843	0.056	0.037	0.027	0.023	0.014
2	0.598	0.189	0.061	0.076	0.038	0.038
3	0.425	0.152	0.141	0.141	0.076	0.065
4	0.546	0.047	0.070	0.128	0.128	0.081
5	0.413	0.143	0.127	0.079	0.048	0.190
6	0.435	0.081	0.064	0.081	0.129	0.210

2.2 Generation within each state

The selection of the next state from the current state is made using a random number. Using Table 3 as an example, assume that the current state is 1 (i.e. zero rain). If the random number is less than or equal to 0.843, then the next state is 1. If the random number is between 0.843 and 0.843+0.056 = 0.899, then the next state is 2, etc.

The frequency of occurrence of daily rainfalls declines between zero and 15 mm, being greatest at 0.1 mm/day and least at 14.9 mm/day. The relative frequency of daily rainfalls in this range is shown in Figure 1. Note that the vertical scale is logarithmic because of the big range of values. When the random number generator determines a rainfall within one of States 2, 3, 4 or 5, the distribution across each range is generated as an exponentially decreasing distribution. This approximates the pattern shown in Figure 1.

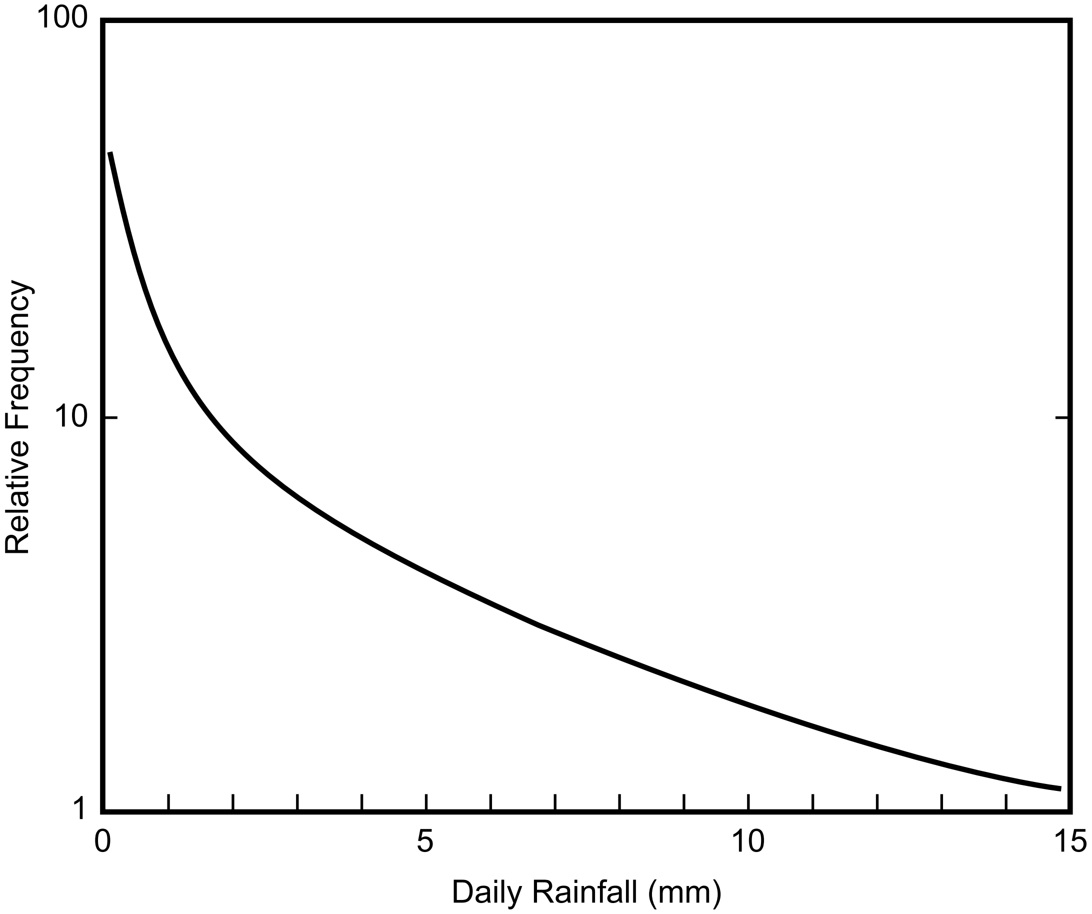

Figure 1. Relative frequency of daily rainfalls in the range 0.1 to 14.9 mm – Boggy Creek data

In State 6, the generated values have no upper limit, and a frequency distribution is used with a second random number to determine the value of the rainfall. The distribution used is the log-Boughton distribution (Boughton, 1980, Boughton and Shirley, 1983). The relationship between average recurrence interval T years and frequency factor K is given by the equation :

$$K = A + \frac{C}{\ln\ln(T/(T-1)) - A}$$

where A is a shape parameter similar to a skew parameter, and C is a function of A. Setting C = A*(A + 0.3665) forces the distribution to make K = 0.0 when T = 2 years. The mean and standard deviation are determined for each month of the year from the daily rainfalls equal to or greater than 15.0 mm in the month. The shape parameter A is determined by trial and error adjustment to match generated rainfalls with estimates of Probable Maximum Precipitation or Forge rainfalls or other estimates of daily rainfalls with long average recurrence intervals.

2.3 Adjustment of variance of annual totals

When a sequence of daily rainfalls is generated with a TPM model of the type described above, and the daily values are summed to annual totals, the frequency distribution of annual totals has a variance that is significantly less than the variance of the annual totals of the actual rainfall. This is illustrated in Figure 2(a) which shows the frequency distribution of annual totals from 2000 years of generated rainfalls compared with the frequency distribution of annual totals of recorded rainfalls, using 58 years of data from the Boggy Creek catchment. Although the averages of daily, monthly and annual rainfalls are reproduced adequately, the simple form of TPM model substantially under-estimates the variance of the annual totals, and hence substantially under-estimates the severity of severe droughts. This negates much of the purpose in using data generation for design of water supply systems.

This problem is overcome in the model by deliberately increasing the variance of the annual totals. In operation, the model generates and stores one year of daily rainfalls, and sums the values to an annual total. The difference between the generated annual total (GAT) and the mean of recorded annual totals (mean) is increased by an adjustment factor (F), and this gives an adjusted annual total (AAT) :

AAT = mean + (GAT - mean)*F

Ratio = AAT / GAT

The ratio AAT / GAT is then used to multiply each daily rainfall in the year of generated rainfalls which have been stored. When the generated annual total is less than the mean, the ratio is less than 1.0 and each daily rainfall is decreased by the ratio. When the generated annual total is more than the mean, the ratio is greater than 1.0 and each daily rainfall is increased. This procedure retains the pattern of daily rainfalls in very dry or very wet years, and merely adjusts each daily value to increase the variance of the annual totals.

Setting F = 1.0 makes AAT = GAT and Ratio = 1.0, i.e. there is no adjustment of the generated daily rainfalls. Values of F, when calibrated to match the frequency distributions of actual and calculated annual totals of rainfall, are usually in the range 1.0 to 2.0.

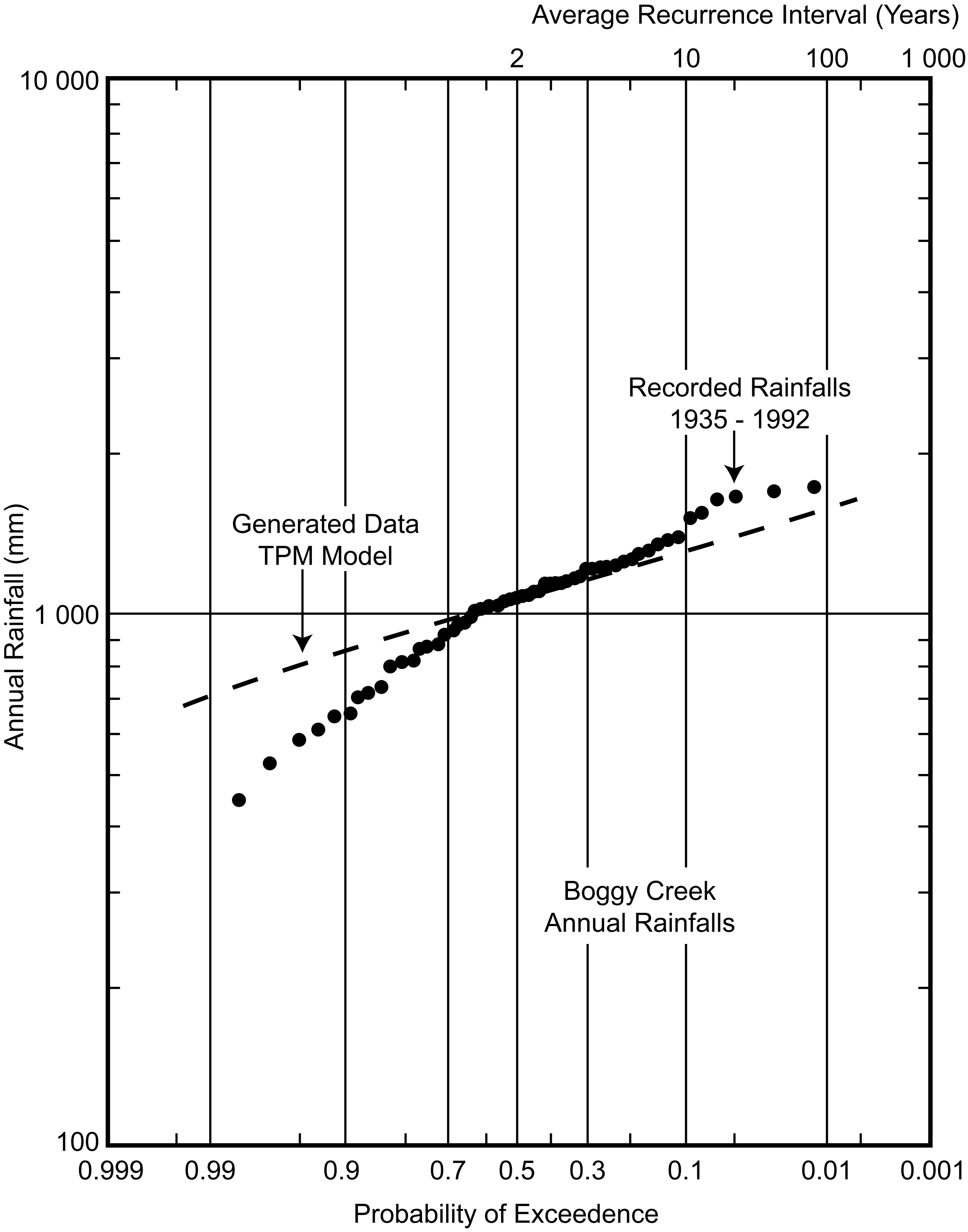

Figure 2(a) Boggy Creek data – comparison of actual and generated distributions of annual rainfalls before adjustment of variance (F = 1.0)

2.4 Calibration of the model

If the model was to be used in the absence of a data set for calibration, it would be necessary to assume values for the 6 x 6 matrices of probabilities for the 12 months of the year. At present, there is no established way to regionalise such a number of parameters. Where rainfall data are available for calibration, the matrix probabilities are automatically calculated by the computer program in which the TPM model is coded, and there are only 2 parameters for a user to evaluate.

One parameter, F, adjusts the variance of the annual totals so that extreme dry periods can be properly simulated. The second parameter, A, is used to adjust the frequency distribution for values >= 15.0 mm/day which affects the annual maxima daily values which are used in flood studies.

The generating model has been coded in Turbo Pascal 6.0 into two programs :

> DAYRNGEN generates 2000 years of daily rainfalls per run and is used to calibrate the annual total variance adjustment factor F. It is also used to generate and store daily, monthly or annual rainfalls for any chosen period up to 2000 years.
>
> PMPRAIN generates 1,000,000 years of rainfall per run, and is used to calibrate the parameter A in the log-Boughton distribution of daily rainfalls >= 15.0 mm/day. It is also used to select annual maxima daily rainfalls which are the usual data of interest when the model is used in flood studies.

To calibrate the model, a value of A is assumed (default setting is A = 10.0). DAYRNGEN is then used to find the value of the annual total variance adjustment factor F by trial and error comparing of generated and actual annual totals of runoff. The value of F is used with PMPRAIN and the value of A is found by trial and error adjustment such that the annual maxima daily rainfalls match with estimates of 24-hour PMP rainfall or CRC-FORGE estimates of daily rainfalls to average recurrence intervals of 2000 years, or similar calibration data.

If the value of parameter A found in the calibration use of PMPRAIN is significantly different from the initially assumed value, then DAYRNGEN is used again to re-calibrate parameter F. Similarly, if the value of F changes significantly in the second calibration, then PMPRAIN is run again to re-calibrate parameter A. It would be a very unusual set of data that required more than two uses of each program to calibrate the model. With the test data of Boggy Creek and Brisbane rainfalls, the second calibration runs of each program made very little difference to the results. The variation in results from change in the parameter values is documented in Section 4.

3. TEST DATA

Two data sets are used to demonstrate the daily rainfall generating model. Both data sets are 9 am to 9 am daily rainfalls. The average annual rainfalls are similar, 1072 mm and 1083 mm; however, the Boggy Creek rainfalls are Thiessen weighted catchment average rainfalls in a winter maximum climate while Brisbane is a single station in a summer maximum climate.

3.1 Boggy Creek data

These daily rainfall data are Thiessen weighted catchment average values using records from 4 rainfall stations in the 108 km^2 Boggy Creek catchment at Angleside, station number 403226 (Hill, 1994). A 58-year period, 1935-1992, of daily rainfalls were available for calibrating and testing the generating model.

These data were used by Boughton and Hill (1997) for testing a design flood estimation procedure involving daily rainfall generation and a rainfall-runoff model.

3.2 Brisbane

The Brisbane data set is composed of 80 years, 1900-1979, of daily rainfalls from the Brisbane meteorological office. These are single station data.

4. RESULTS

4.1 Run to run variation in DAYRNGEN

With fixed values of the two parameters, A and F, there will be variation in the results of DAYRNGEN from run to run because of the use of random numbers in the generation process. This variation can be of the same order of magnitude as the variation due to change in parameter values.

Table 4 shows the variation in results from 5 consecutive runs of DAYRNGEN using the same values of A = 12.0 and F = 1.65. Each run was 2000 years of generated rainfalls to produce the values for the probabilities of exceedance (Pe) in the table.

Table 4. Annual totals of generated rainfall (mm) for Boggy Creek catchment
[A = 12.0, F = 1.65]

Pe	Run number					Ave	Range
	1	2	3	4	5		
0.01	1922	1871	1870	1868	1871	1880	1868-1922
0.1	1490	1483	1504	1491	1496	1493	1483-1504
0.3	1234	1228	1231	1249	1234	1235	1228-1249
0.5	1055	1049	1061	1070	1055	1058	1049-1070
0.7	89	889	896	906	882	893	882-906
0.9	683	673	677	697	661	678	661-697
0.99	410	411	411	451	398	416	398-451

4.2 Variation in results of DAYRNGEN due to parameter A

The assumption of a value of parameter A when making the first calibration of parameter F has little effect on the calibration of F for a wide range of values of A. Table 5 shows the change in generated annual totals using DAYRNGEN due to change in the value of parameter A with F = 1.65. Each column of results for a value of A are averages of 5 runs, as in Table 4.

Table 5. Annual total of generated rainfalls (mm) for Boggy Creek catchment

Pe	Generated rainfall (mm)		
	A=10	A=12	A=14
0.01	1868	1880	1908
0.1	1472	1493	1516
0.3	210	1235	1244
0.5	1042	1058	1069
0.7	877	893	902
0.9	670	678	684
0.99	394	416	421

Table 5 shows that, for a given probability of exceedance, the difference in generated annual total rainfall between A = 10 and A = 14 is less than the difference between individual runs with A = 12, as shown in Table 4. The significance of this is that the effect of the initial assumption of a value of A is quickly eliminated by the second iteration of programs DAYRNGEN and PMPRAIN.

4.3 Calibration of model to Boggy Creek data

(i) Calibration to annual rainfalls with DAYRNGEN

Table 6 shows the calibration of parameter F to the Boggy Creek data using DAYRNGEN. Each of the values shown in the range from F = 1.0 to F = 2.0 are the averages from 5 runs for each setting of F.

Table 6. Annual Rainfall (mm/yr) for Boggy Creek data

Pe	LB Fit	F					
		1.0	1.3	1.5	1.65	1.8	2.0
0.01	1700	1565	1715	1820	1880	1945	2075
0.1	1440	1325	1405	1445	1490	1530	1585
0.3	1225	1165	1190	1215	1235	1240	1255
0.5	1070	1065	1060	1060	1060	1060	1045
0.7	915	965	930	910	895	880	850
0.9	65	830	760	705	675	635	585
0.99	410	665	535	450	415	345	268

The values in the column headed "LB Fit" come from fitting the log-Boughton frequency distribution to the 58 annual totals of recorded rainfall, 1935-1992, and calculating values for the probabilities of exceedance, Pe, from 0.01 to 0.99. Because of the large number (58) of data points, the choice of distribution has very little effect on the calculated values. When DAYRNGEN is used in practice, the daily rainfall generating model is fitted directly to the data points - the fitted distribution is used in Table 6 only for purposes of comparison.

The problem of using a Transition Probability Matrix model without adjustment of the annual totals is illustrated in Table 6 by the column headed "F = 1.0". The variance of the annual totals is substantially under-estimated as shown by the result for Pe = 0.99, i.e. estimated 665 mm versus actual 409 mm, and the under-estimation at Pe = 0.01, i.e. estimated 1565 mm versus actual 1700 mm. The severity of severe droughts would be badly under-estimated if there was no adjustment of the annual totals.

As the value of parameter F is increased from 1.0 to 2.0, the variance of the annual totals is also increased. Figure 2(b) shows the results from F = 1.65 plotted against the actual 58 annual totals. This confirms that the model with variance adjustment can satisfactorily reproduce the probability of severe drought years, but calibration of parameter F is essential. Change in the value of F from 1.5 to 1.8 around the calibrated value of 1.65 produces changes of +9% to -17% in the estimation of the Pe = 0.99 annual rainfall, i.e. the 1 in 100 years drought. While this is not large sensitivity, the effect on estimated streamflow will be much greater than the effects on rainfall in very dry years, so accurate calibration of F is needed when the rainfall model is to be used with a rainfall-runoff model for drought studies.

(ii) Calibration to annual maxima daily rainfalls with PMPRAIN

When the daily rainfall generating model is to be used for flood studies, it is important that the largest daily rainfalls are modelled correctly, while the annual totals are of much less significance. The values of the largest daily rainfalls are determined mainly by the parameter A in the log-Boughton frequency distribution which is used to generate daily rainfalls in state 6, i.e. >= 15.0 mm/day.

The effect of change in the value of A is illustrated in Table 7, which shows annual maxima daily rainfalls generated by PMPRAIN with three values of A. With each value, 1,000,000 years of daily rainfalls were generated and the annual maxima series selected from the results. There were two estimates of 24-hour PMP rainfall for catchments close to the Boggy Creek catchment, the values being 470 mm/day and 665 mm/day. The assumed ARI for the PMP rainfalls is 1,000,000 years. The generated value of 549 mm/day for this ARI using A = 12 is closest of those shown in Table 7, and the value of 12 is selected as the calibrated value of parameter A. Figure 3 shows a comparison of the actual annual maxima daily rainfalls with generated values using A = 12.0 and F = 1.65.

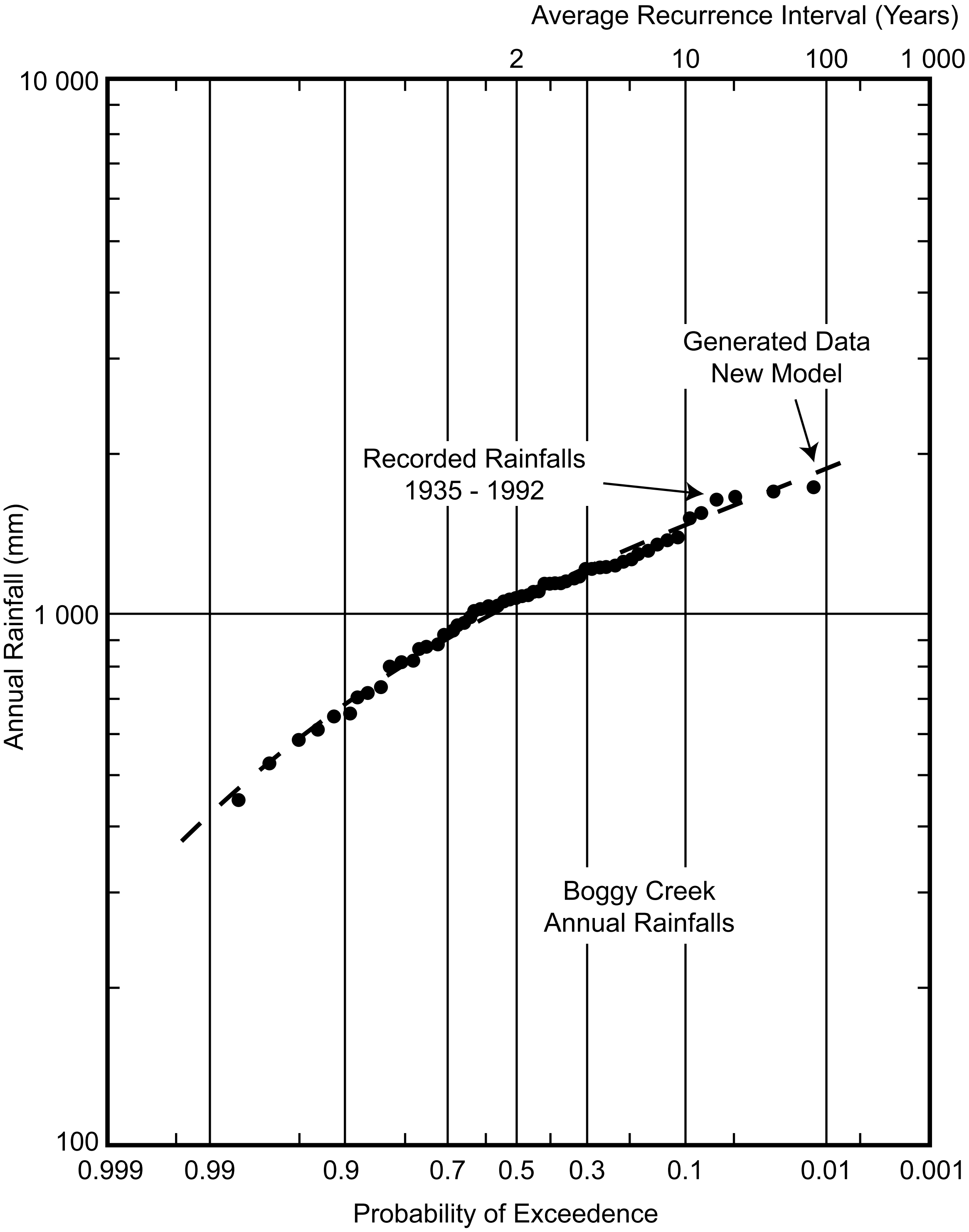

Figure 2(b) Boggy Creek data – comparison of actual and generated distributions of annual rainfalls after adjustments of variance (F = 1.65)

Table 7. Rainfalls generated by PMRAIN

ARI	Annual maxima daily rain (mm)		
Years	A = 10	A = 12	A = 14
1,000,000	387	549	619
100,000	333	420	515
10,000	264	323	383
1,000	196	233	269
100	139	155	173
10	91	98	103

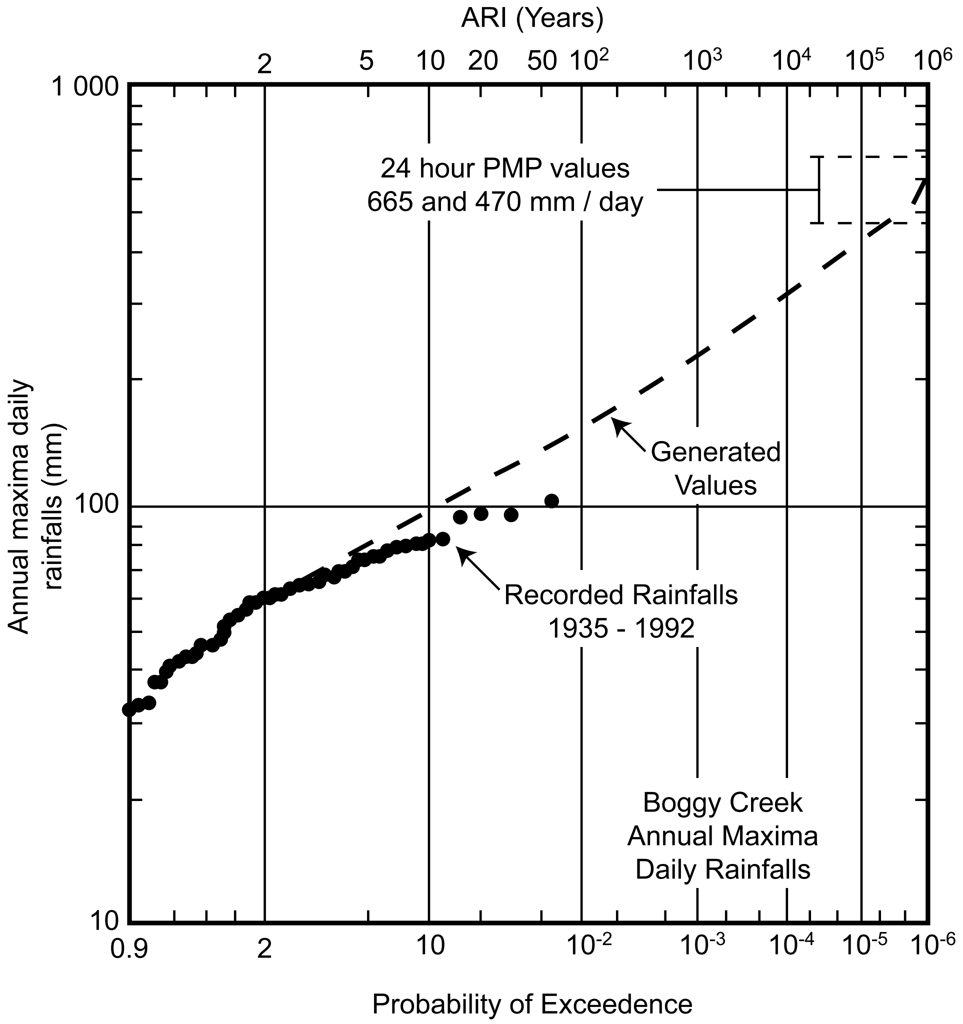

Figure 3 Annual maxima daily rainfalls for Boggy Creek.
Comparison of generated and actual data.

Because the mean and standard deviation are fixed by the 58 years of actual data, there is only a small spread of values at ARI = 10 years, but the range increases substantially as the ARI increases. Each of the values in Table 7 was obtained from a single run of program PMPRAIN. Each run of 1,000,000 years takes about 4.5 hours on a 100 megahertz Pentium PC which incorporates a math co-processor and a 16 KB cache memory. There is variation in the results from run to run, and it is advisable in practice to use the average of several runs to offset this variation. The use of replicate runs provides confidence limits on the generated value of the PMP as well as giving a better estimate of the value.

If the value of A is determined from the actual data, similar to fitting a frequency distribution to annual maxima floods, then very minor changes in the data (e.g. one very dry year or adding a few more years of data) can have a substantial effect on the generated values for very large ARIs (see Boughton and Hill, 1997). For this reason, it is recommended that the parameter A be calibrated against estimates of 24-hour probable maximum precipitation, which are set at an assumed ARI (see Australian Rainfall and Runoff, chapter 13).

(iii) Statistics of actual and generated rainfalls

Table 8 shows monthly and annual averages of actual and generated rainfalls using the Boggy Creek data. The generated values are the averages from 2000 years of generated daily rainfalls. The parameter values used to obtain these values are A = 12 (see Table 7 above) and F = 1.65 (see Table 6 above).

Table 8. Monthly and annual rainfall averages (mm) for Boggy Creek [A=12, F = 1.65]

Month	Actual	Generated
1	56.3	58.0
2	47.4	48.3
3	58.8	60.3
4	78.8	78.3
5	108.5	111.1
6	116.0	116.1
7	136.9	131.9
8	136.3	132.8
9	101.1	101.3
10	98.3	95.7
11	71.5	70.6
12	62.1	66.6
Year	1071.9	1070.6

The matching of monthly and annual averages in Table 8, together with the matching of the frequency distribution of annual totals as in Figure 2(b) and the matching of annual maxima daily rainfalls by the calibration of parameter A, illustrates that the generating model satisfactorily reproduces the characteristics of daily rainfalls.

4.4 Calibration of the model to Brisbane data

The calibration of the model to the Brisbane data followed the procedure described in detail for the Boggy Creek data in Section 4.3 above. Using the default value of A = 10, parameter F was calibrated using DAYRNGEN. Then A was varied using PMPRAIN. A second round with the two programs resulted in the calibrated parameter values of A = 8.2 and F = 1.38.

A comparison of the generated frequency distribution of annual totals of rainfall against the actual 80 years of data is given in Figure 4. Figure 4(a) shows the results from the TPM model without adjustment of the annual totals, i.e. F = 1.0. Figure 4(b) shows the calibrated results with F = 1.38.

An estimate of the 24-hour Probable Maximum Precipitation for the region around the Brisbane rainfall station is 950 mm (Ruffini, personal communication). This was used to calibrate parameter A with PMPRAIN. The frequency distribution of generated annual maxima daily rainfalls is shown in Figure 5.

Table 9 shows a comparison of the average monthly and annual generated rainfalls with the actual averages from the 80 years of input data. The generated averages were from 2000 years of generated rainfalls.

Table 9. Monthly and annual rainfall averages (mm) for Brisbane [A = 8.2, F = 1.38]

Month	Actual	Generated
1	154.3	157.8
2	157.7	153.7
3	140.0	139.2
4	76.8	79.3
5	63.7	66.0
6	67.7	66.1
7	50.7	50.1
8	35.9	35.8
9	44.3	43.7
10	77.7	77.5
11	96.4	96.3
12	124.1	123.3
Year	1083.3	1088.8

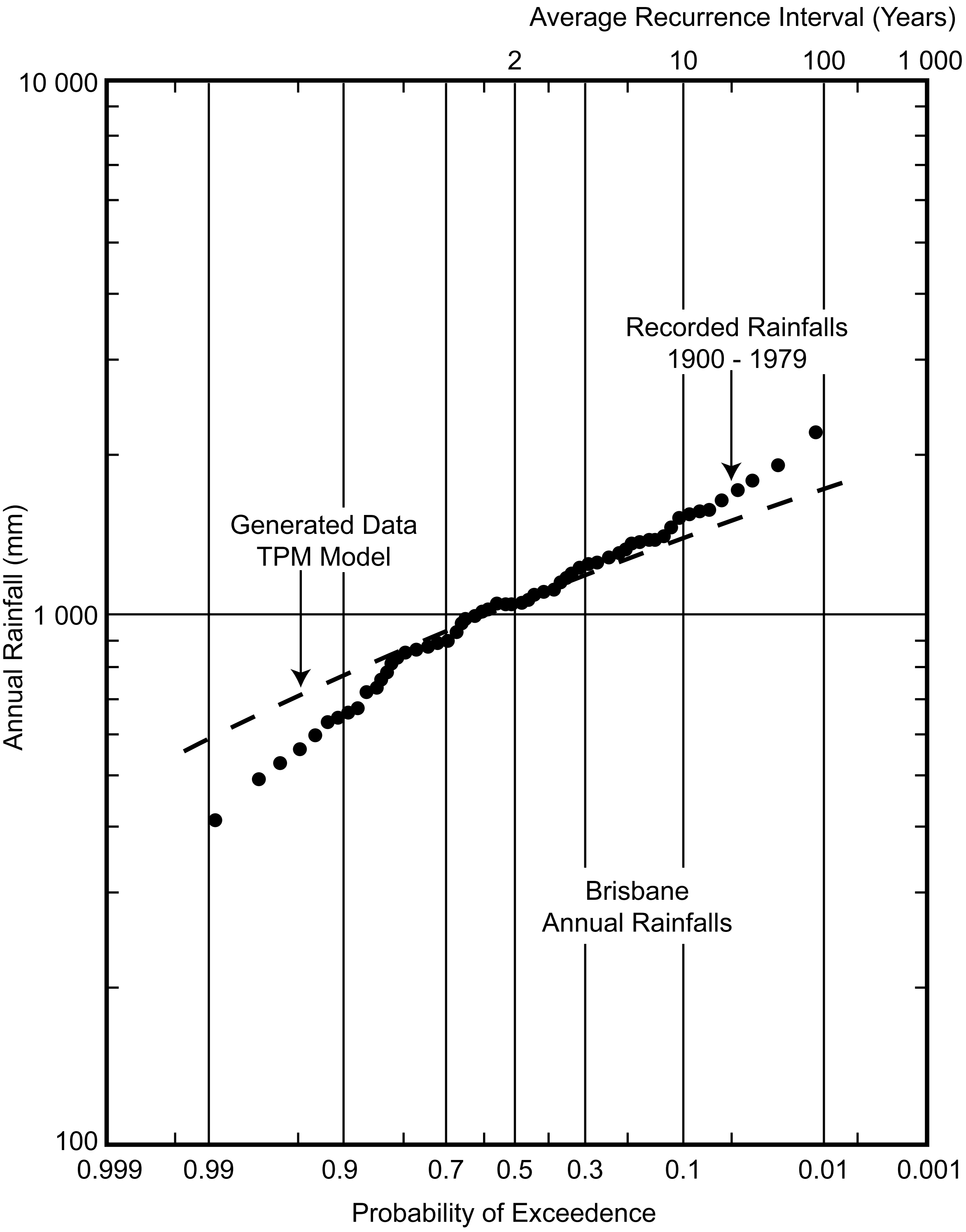

Figure 4(a) Brisbane data – comparison of actual and generated distributions of annual rainfalls before adjustment of variance [F = 1.0]

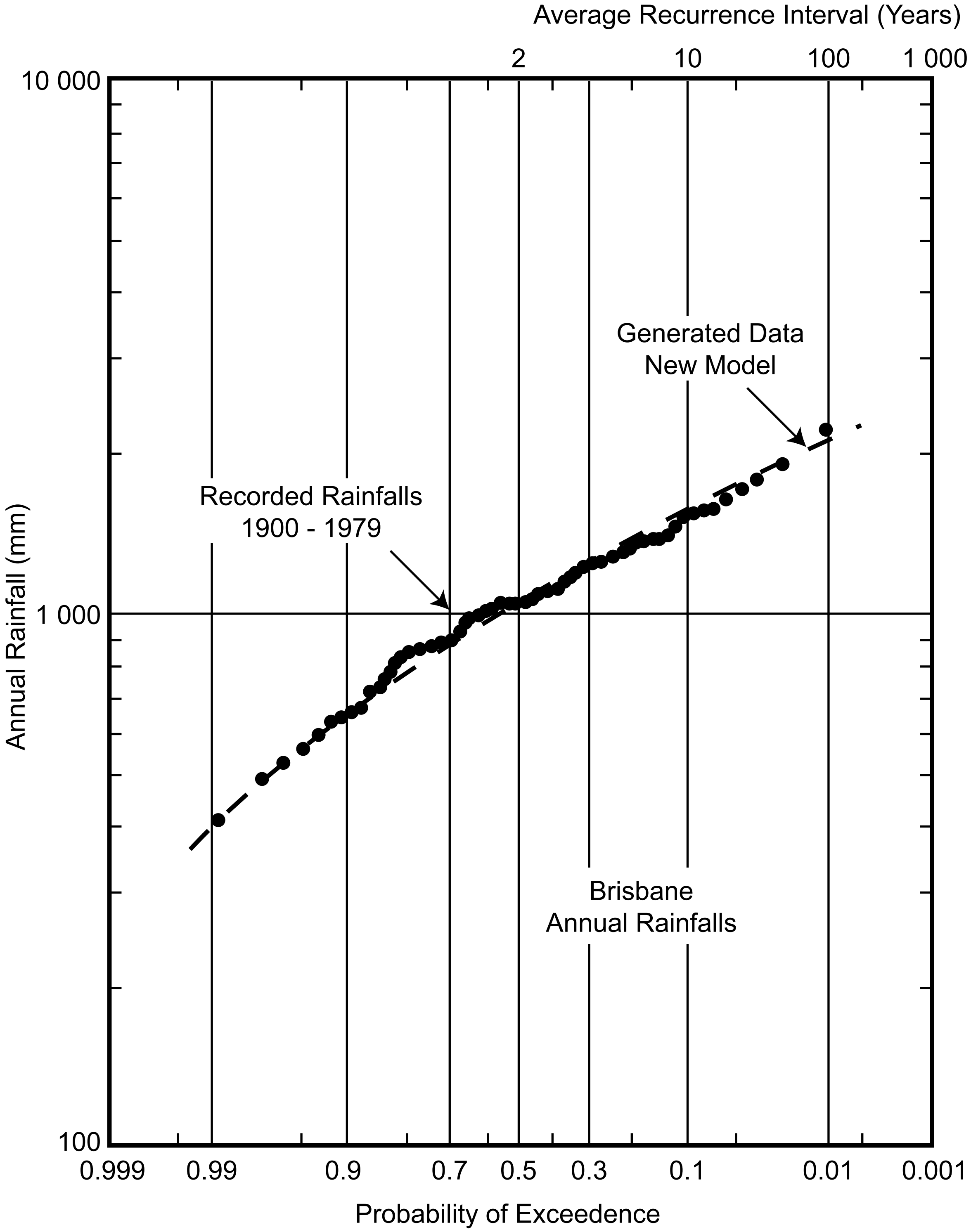

Figure 4(b) Brisbane data – comparison of actual and generated distributions of annual rainfalls after adjustment of variance [F = 1.38]

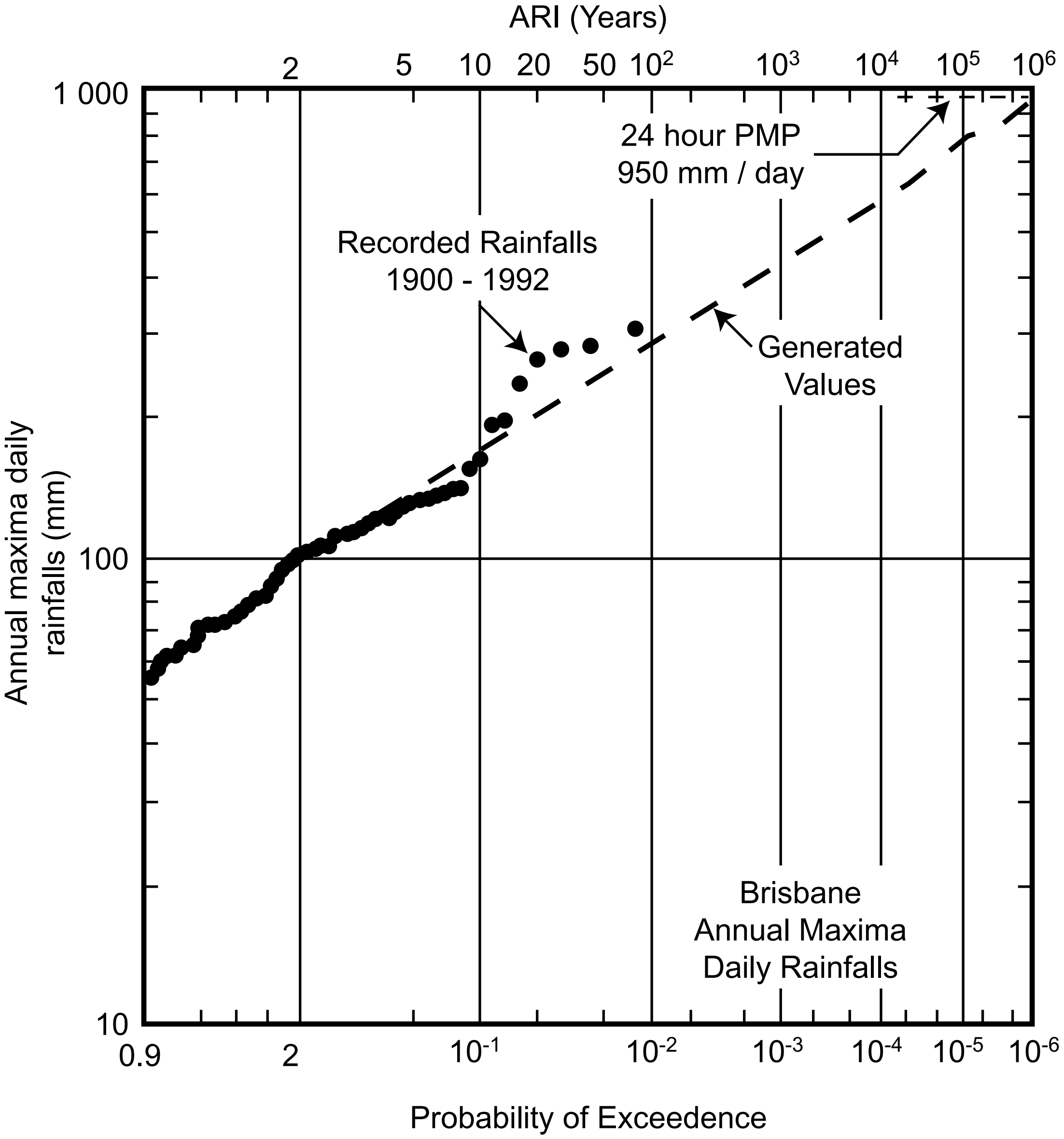

Figure 5. Annual maxima daily rainfalls for Brisbane.
Comparison of generated and actual data.

5. DISCUSSION

A brief review was made of short duration (hourly and daily) data generation models for both rainfall and runoff. In the studies seen, comparison of generated with actual data relied heavily on comparisons of the means and standard deviations of the short duration data. This can be very misleading. Table 10 shows how well the average monthly and average annual rainfalls for Boggy Creek and Brisbane are reproduced without any adjustment of the annual totals, i.e. F = 1.0.

Table 10. Comparison of actual and generated average monthly and annual rainfalls (mm) [F = 1.0]

Month	Boggy Creek		Brisbane	
	Act	Gen	Act	Gen
Jan	56.3	56.8	154.3	152.6
Feb	47.4	48.0	151.7	151.8
Mar	58.8	61.0	140.0	141.5
Apr	78.8	78.2	76.8	77.6
May	108.5	111.2	63.7	65.4
Jun	116.0	116.1	67.7	69.5
Jul	136.9	135.7	50.7	49.6
Aug	136.3	132.7	35.9	36.9
Sep	101.1	101.8	44.3	44.1
Oct	98.3	98.3	77.7	76.4
Nov	71.5	74.0	96.4	96.1
Dec	62.1	62.7	124.1	123.9
Year	1071.9	1076.3	1083.3	1085.5

The generated values are the averages from 2000 years of generated daily rainfalls in each case.

These results, if taken alone, suggest that the model is reproducing the general statistics of the rainfall; however, Figures 2(a) and 4(a) show that the variances of the annual totals are very much in error.

Srikanthan and McMahon (1985) checked the standard deviations of annual totals of rainfall that were generated by their daily rainfall generating model. Table 11 shows their results for the 15 rainfall stations used in their study

Table 11. Standard deviation of actual and generated annual totals of rainfall (mm) (Actual values from Srikanthan and McMahon, 1985)

Location	Act	Gen
Melbourne	128	111
Sydney	330	292
Monto	221	169
Cowra	212	143
Mackay	544	471
Brisbane	364	232
Darwin	253	270
Broome	260	206
Perth	166	140
Adelaide	108	99
Alice Springs	114	108
Kalgoorlie	106	80
Onslow	234	171
Bamboo Springs	139	120
Lerderderg	210	143

By selecting and highlighting these results from the Srikanthan and McMahon study, the under-estimation of the standard deviations of the annual totals is obvious; however, it is only with the hindsight of the results shown in Figures 2(a) and 4(a) that it is possible to select and highlight in this way. Srikanthan and McMahon made extensive checks of their daily rainfall generating model including the average monthly and annual number of wet days, the mean, standard deviation and coefficient of skewness of dry and wet spells for each month, the maximum daily rainfall in each month, the mean, standard deviation and coefficient of skewness of rainfall depths on wet days in each month, the correlation between rainfall depth and the length of wet spell in each month, and the longest wet and dry spell in the record or replicate. It was only when the problem with annual totals became apparent in the Boughton and Hill (1997) study that it was possible to identify the problem and to develop the variance adjusted model described in this paper.

6. CONCLUSION

The Transition Probability Matrix model is both simple and robust for daily rainfall generation. The structure of the model ensures that the generated data matches the monthly statistics used in the monthly probability matrices - means and standard deviations - and the frequencies of wet and dry spells of various durations. It is significant that the error in the variance of annual totals, which occurred in the earlier versions of TPM models, occurred in a statistic that was not constrained by the structure of the model.

The under-estimation of the variance of generated annual totals of rainfall is the main problem addressed in this paper and is corrected by the variance adjusted generating model. Because the monthly probability matrices are the same as in the earlier versions, the model accurately matches the monthly statistics and wet and dry spells as before. The introduction of the adjustment to the variance of the annual totals now ensures that the probability of exceedance of very dry years are satisfactorily reproduced. The calibration of the variance adjustment is very simple and users of the computer program do not require special training or lengthy experience.

The new model had been designed to be useable for both drought (i.e. water supply) and flood studies. The latter require the model to reproduce the frequency distribution of annual maxima daily flows. The model has been developed so that the annual maxima daily flows can be calibrated against estimates of 24-hour Probable Maximum Precipitation or similar estimates of large daily rainfalls. The use of annual maxima daily rainfalls in flood estimation is described elsewhere (Boughton and Hill, 1997).

7. REFERENCES

Allen, D. M. and Haan, C. T., 1975 : "Stochastic simulation of daily rainfall", Res. Rept. 82, Univ. of Kentucky, Water Resources Inst., Lexington.

Boughton, W. C. and Hill, P., 1997 : "A design flood estimation procedure using data generation and a daily water balance model" Report 97/8, CRC for Catchment Hydrology, Monash Univ., Melbourne.

Chapman, T., 1994 : "Stochastic models for daily rainfall", The Institution of Engineers, Australia, National Conference Publication 94/15, pp.7-12.

Haan, C. T., Allen, D. M. and Street, J. D., 1976 : "A Markov chain model of daily rainfall", Water Resources Res., 12(3), pp.443-449.

Nandakumar, N., Weinmann, E., Mein, R. and Nathan, R., 1997 : "Estimation of extreme rainfalls for Victoria using the CRC-FORGE method", Report 97/4, CRC for Catchment Hydrology, Monash Univ., Melbourne.

Srikanthan, R. and McMahon, T. A., 1985 : "Stochastic generation of rainfall and evaporation data", Australian Water Resources Council, Tech. Pap. no.84, AGPS, Canberra, 301 pp.

Srikanthan, R., Elliott, J. and Adams, G., 1994 : "A review of real-time flood forecasting methods", Report 94/2, CRC for Catchment Hydrology, Monash Univ., Melbourne.

The Cooperative Research Centre for Catchment Hydrology
is a cooperative venture formed under the Commonwealth CRC Program between:

Bureau of Meteorology
CSIRO Land and Water
Department of Land and Water Conservation, NSW
Department of Natural Resources and Environment, Vic
Goulburn-Murray Water
Melbourne Water
Monash University
Murray-Darling Basin Commission
Southern Rural Water
The University of Melbourne
Wimmera-Mallee Water

Associates
Department of Natural Resources, Qld
Hydro-Electric Corporation, Tas
State Forests of NSW

COOPERATIVE RESEARCH CENTRE FOR
CATCHMENT HYDROLOGY

Centre Office
Department of Civil Engineering, PO Box 60, Monash University, Victoria, 3800 Australia.
Telephone: (03) 9905 2704 Facsimile: (03) 9905 5033 International +61 3 9905 2704
http://www.catchment.crc.org.au