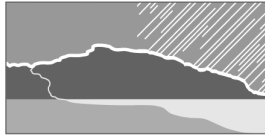


COOPERATIVE RESEARCH CENTRE FOR



CATCHMENT HYDROLOGY

## ***WORKING DOCUMENT***

# **STOCHASTIC MODELLING OF DAILY RAINFALL**

**S. L. Zhou  
R. Srikanthan  
T.A. McMahon**

**Working Document 02/5  
August 2002**

**In view of the preliminary nature of CRC Working Documents, permission for quoting or reproduction from this Working Document is required from a Deputy Director, CRC for Catchment Hydrology by persons other than CRC funded staff or listed in-kind staff from CRC Parties**

**For further information contact the Centre Office on  
tel: 03 9905 2704**

**S. L. Zhou  
R. Srikanthan  
T.A. McMahon**

**Stochastic modelling of daily rainfall**

**Working Document 02/5**

**© Cooperative Research Centre for Catchment Hydrology, 2002  
[www.catchment.crc.org.au/publications](http://www.catchment.crc.org.au/publications)**

## **SUMMARY**

Stochastic generation of rainfall data offers an alternative to the use of observed records. This paper presents an evaluation of daily rainfall generation models at 21 stations across Australia. The models are the Transition Probability Matrix (TPM) method, the Daily and Monthly Mixed (DMM) algorithm, and a variation of each model. A goal of stochastic modelling is to generate synthetic data that are representative of the statistical characteristics of the historical data. One hundred replicates of length equal to the historical data were generated using the above models. Preliminary assessment of the models suggests that overall both the TPM and DMM models preserved key statistical characteristics of the historical rainfall at the annual, monthly and daily levels. The DMM model was unable to preserve the amounts of rainfall on solitary wet days and the TPM model needed to be modified by the empirical adjustment factor to preserve the annual variability.

## CONTENTS

<b>1. INTRODUCTION.....</b>	<b>1</b>
<b>2. RAINFALL DATA.....</b>	<b>2</b>
<b>3. DAILY RAINFALL MODELS.....</b>	<b>3</b>
3.1 TRANSITION PROBABILITY MATRIX (TPM) MODEL.....	3
3.2 MODIFIED TRANSITION PROBABILITY MATRIX (TPMB) MODEL.....	5
3.3 DAILY AND MONTHLY MIXED (DMM) MODEL .....	6
3.4 MODIFIED DAILY AND MONTHLY MIXED (DMMM) MODEL.....	7
<b>4. MODEL EVALUATION.....</b>	<b>8</b>
4.1 PARAMETERS FOR MODEL EVALUATION.....	8
4.2 RESULTS.....	10
4.2.1 <i>Annual rainfall</i> .....	10
4.2.2 <i>Monthly rainfall</i> .....	11
4.2.3 <i>Daily rainfall</i> .....	12
<b>5. CONCLUSIONS.....</b>	<b>13</b>
<b>6. REFERENCES .....</b>	<b>14</b>

## LIST OF TABLES

Table 1:	Details of the rainfall stations selected.....	2
Table 2:	State boundaries for the TPM models.....	3
Table 3:	Number of states used for the 21 stations in the daily TPM model. ....	5
Table 4:	Selection of annual, monthly and daily statistics. ....	15
Table 5:	Assessment of daily models. ....	17

## LIST OF FIGURES

Figure 1:	Locations of the rainfall stations used and CRCCH focus catchments. ....	3
Figure A1:	Mean annual rainfall. ....	19
Figure A2:	Standard deviation of annual rainfall. ....	19
Figure A3:	Lag one auto-correlation. ....	19
Figure A4:	Minimum 5-year rainfall sums. ....	19
Figure A5:	Coefficient of skewness of monthly rainfall. ....	20
Figure A6:	Number of months with no rain. ....	21
Figure A7:	Average number of wet days. ....	22
Figure A8:	Mean daily rainfall. ....	23
Figure A9:	Mean daily rainfall for solitary wet days. ....	24
Figure A10:	Mean wet spell length. ....	25
Figure A11:	Longest dry spell. ....	26
Figure A12:	Maximum daily rainfall. ....	27
Figure A13:	Comparison of observed and TPMb model generated mean annual rainfall. ....	28
Figure A14:	Comparison of observed and TPMb model generated standard deviation of annual rainfall. ....	28
Figure A15:	Comparison of observed and TPMb model generated lag one auto-correlation of annual rainfall. ....	28
Figure A16:	Comparison of observed and TPMb model generated minimum 5-year rainfall sums. ....	28
Figure A17:	Comparison of observed and TPMb model generated coefficient of skewness of monthly rainfall. ....	29
Figure A18:	Comparison of observed and TPMb model generated number of months with no rain. ....	30
Figure A19:	Comparison of observed and TPMb model generated average number of wet days. ....	31
Figure A20:	Comparison of observed and TPMb model generated mean daily rainfall. ....	32
Figure A21:	Comparison of observed and TPMb model generated mean daily rainfall for solitary wet days. ....	33
Figure A22:	Comparison of observed and TPMb model generated mean wet spell length. ....	34
Figure A23:	Comparison of observed and TPMb model generated longest dry spell. ....	35
Figure A24:	Comparison of observed and TPMb model generated maximum daily rainfall. ....	36

## 1. Introduction

Stochastic generation of daily rainfall series is an increasingly used procedure for hydrological and water resources studies. There has been considerable development of daily single-site models such as two-parts models, transition probability matrix models, resampling models and time series models. A comprehensive review of approaches to modelling daily rainfall is given by Srikanthan and McMahon (1985, 2000).

Most stochastic models of daily rainfall consist of two parts, a model for the occurrence of dry and wet days and a model for the generation of the rainfall amount on wet days. Models of rainfall occurrence are commonly based on Markov chains which specify the state of each day as wet or dry and develop a relation between the state of the current day and the states of the preceding days. Models used for rainfall amounts on wet days include the exponential, the Gamma, and skewed normal distribution.

Srikanthan and McMahon (1985), in a study covering the main climatic regions of Australia, combined the modelling of rainfall occurrences and rainfall amounts by extending the Markov chain concept to a multi-state model or transition probability matrix (TPM). The daily rainfalls are grouped into states (maximum of seven) of specified sizes, and the probabilities are calculated for transition from each state to any other. Transition to the lowest state specifies the occurrences of dry days, to the other states gives the occurrences of wet days. The state of the largest magnitude is modelled by a power transformation and intermediate states are modelled by a linear distribution.

Boughton (1999) observed that the TPM model underestimates the standard deviation of annual rainfall and proposed an empirical adjustment to match the observed standard deviation.

Wang and Nathan (2002) developed a daily and monthly mixed (DMM) algorithm for the generation of daily rainfall. Daily rainfall data are generated month by month using two sets of parameters for the Gamma distribution: one estimated from the daily rainfall data and the other from monthly rainfall data. A pro-rata adjustment is made to the generated daily rainfalls from the daily Gamma parameters to match the serially correlated monthly rainfalls.

This report presents the application and evaluation of stochastic data generation models of rainfall using data from 21 stations located in various parts of Australia. In this study, two basic daily models were examined namely the Transition Probability Matrix (TPM) method (Srikanthan and McMahon, 1985) and the Daily and Monthly Mixed (DMM) algorithm (Wang and Nathan, 2002), and a variation of these models. The models were assessed by a number of statistics. The daily rainfall data sets and the models are briefly described below, followed by an assessment of the performance of the models.

## 2. Rainfall Data

Twenty one stations located across Australia with long daily rainfall records were selected. The details of the stations are given in Table 1 and their locations are shown in Figure 1 along with the CRCCH focus catchments. The data length varies from 42 to 125 years. The mean annual rainfall varies from 180 to 1490 mm and the average number of wet days per year is in the range 24 to 153.

Table 1: Details of the rainfall stations selected.

Number	Name	Latitude	Longitude	Record length (years)	Mean annual rainfall (mm)	CV of annual rainfall	No of wet days per year
004035	Roebourne PO	-20.78	117.12	65	315	0.59	24
006036	Meedo	-25.66	114.62	57	209	0.50	27
010037	Cuttening	-31.73	117.76	70	303	0.26	61
012065	Norseman PO	-32.20	121.78	49	303	0.32	70
014016	Darwin PO	-12.40	130.80	42	1490	0.19	96
015531	Hermannsburg	-23.90	132.80	44	222	0.44	31
017031	Marree	-29.65	138.06	55	180	0.55	34
023721	Happy Valley Res	-35.07	138.57	85	639	0.22	133
028004	Palmerville	-16.00	144.08	87	1040	0.30	83
035070	Taroom PO	-25.65	149.80	102	668	0.28	62
040082	Gatton-Lawes	-27.55	152.33	58	815	0.24	97
044026	Cunnamulla PO	-28.07	145.68	97	374	0.43	44
055045	Pine Cliff	-31.50	150.45	66	550	0.27	62
068034	Jervis bay	-35.10	150.80	74	1254	0.37	141
075035	Hay	-34.48	144.98	45	328	0.31	51
077030	Narraport	-36.01	143.03	83	353	0.30	68
086071	Melbourne RO	-37.82	144.97	125	660	0.19	153
086117	Toorourrong Res	-37.48	145.15	99	803	0.20	145
088060	Wallaby Creek	-37.45	145.22	86	1219	0.20	165
091013	Carrick	-41.55	147.02	50	713	0.205	124
092037	Swan Island	-40.70	148.10	47	604	0.24	138



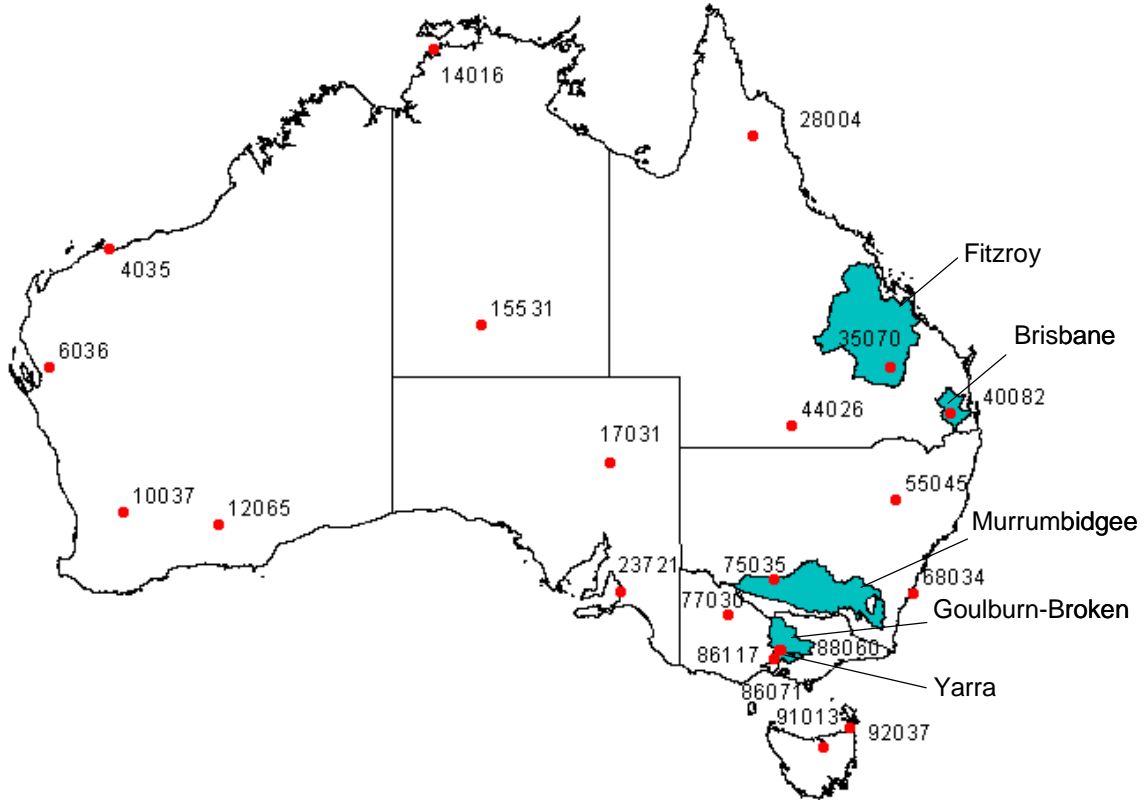


Figure 1: Locations of the rainfall stations used and CRCCH focus catchments.

### 3. Daily Rainfall Models

#### 3.1 Transition Probability Matrix (TPM) model

In the Transition Probability Matrix (TPM) model, the seasonality in the occurrence and magnitude of daily rainfall are taken into account by considering each month separately. The daily rainfall is divided into a number of states up to a maximum of seven. State 1 is dry (no rainfall) and the other states are wet. The state boundaries are given in Table 2.

Table 2: State boundaries for the TPM models.

State Number	State Boundary in mm	
	Lower	Upper
1	0.0	0.0
2	0.1	0.9
3	1.0	2.9
4	3.0	6.9
5	7.0	14.9
6	15.0	30.9
7	31.0	$\infty$

The transition probabilities are estimated from

$$p_{ij}(k) = f_{ij}(k) / \sum_{j=1}^C f_{ij}(k) \quad i = 1, 2, \dots, C; \quad k = 1, 2, \dots, 12 \quad (1)$$

Where  $p_{ij}(k)$  = probability of transition from state  $i$  to state  $j$  within month  $k$ ,  
 $f_{ij}(k)$  = historical frequency of transition from state  $i$  to state  $j$  within month  $k$ ,  
 $C$  = the maximum number of states.

The rainfall values in the intermediate states are modelled by a linear distribution. In an intermediate state, the derivation given by Srikanthan and McMahon (in Appendix G, 1985) shows that

$$x = x_l + (x_u - x_l) \frac{B - \sqrt{B^2 - 4AU}}{2A} \quad (2)$$

where  $x$  is a value in the intermediate state, and  $x_l$  and  $x_u$  are the lower and upper state boundaries of the state and  $U$  is a uniformly distributed random number between 0 and 1.  $A = 0.3$  and  $B = 1.3$  are the empirical values used in the study.

The rainfall values in the largest state are modelled by a power transformation. The exponent,  $\lambda$ , in the power transformation

$$y = (x - x_l)^\lambda \quad (3)$$

is obtained by trial and error so that the skewness of the transformed value,  $y$ , is close to zero. In Equation (3)  $x$  is a value in the largest state and  $x_l$  is the lower state boundary of the largest state.

The daily rainfall data are generated by following the steps set out below (Srikanthan and McMahon, 1985). Assume that the initial state is dry (i.e., state one).

- Step 1: Generate a uniformly distributed random number  $U$  between 0 and 1. Using the appropriate TPM for the month, determine the state of the next day.
- Step 2: If the state is wet, go to Step 3. Otherwise, set the rainfall depth to zero and go to Step 1.
- Step 3: Calculate the rainfall depth by using a linear distribution for the intermediate states and a power transformation for the largest state.
- Step 4: Repeat Steps 1 to 3 until the required length of daily rainfall data is achieved.

The number of states for each month used for the TPM model are given in Table 3. The number of states chosen was based on the recommendation of Srikanthan and McMahon (1985).

Table 3: Number of states used for the 21 stations in the daily TPM model.

Station	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Roebourne Post Office	6	6	6	5	5	5	2	2	2	2	2	5
Meedo	6	6	6	5	5	5	2	2	2	2	2	5
Cuttening	6	6	6	6	6	6	6	6	6	6	6	6
Norseman Post Office	5	5	5	5	5	5	5	5	5	5	5	5
Darwin Post Office	7	7	7	7	3	2	2	2	3	7	7	7
Hermannsburg Mission	4	4	4	4	4	4	4	4	4	4	4	4
Marree	4	4	4	4	4	4	4	4	4	4	4	4
Happy Valley Reservoir	6	6	6	6	6	6	6	6	6	6	6	6
Palmerville	7	7	7	7	3	2	2	2	3	7	7	7
Taroom Post Office	7	7	7	7	7	7	7	7	7	7	7	7
Gatton-Lawes	7	7	7	7	7	7	7	7	7	7	7	7
Cunnamulla Post Office	4	4	4	4	4	4	4	4	4	4	4	4
Pine Cliff	6	6	6	6	6	6	6	6	6	6	6	6
Jervis bay (Pt Perpendicular)	7	7	7	7	7	7	7	7	7	7	7	7
Hay	6	6	6	6	6	6	6	6	6	6	6	6
Narraport	6	6	6	6	6	6	6	6	6	6	6	6
Melbourne Regional Office	6	6	6	6	6	6	6	6	6	6	6	6
Toorourrong Reservoir	6	6	6	6	6	6	6	6	6	6	6	6
Wallaby Creek (Weir)	6	6	6	6	6	6	6	6	6	6	6	6
Carrick	6	6	6	6	6	6	6	6	6	6	6	6
Swan Island	6	6	6	6	6	6	6	6	6	6	6	6

### 3.2 Modified Transition Probability Matrix (TPMb) model

To improve the annual variability of rainfall, a modified version of the TPM model (TPMb) was formulated using the Boughton adjustment factor.

Boughton (1999) proposed an empirical adjustment factor ( $F$ ) in the TPM procedure to match the generated standard deviation of the annual rainfall with the observed value. The adjustment factor is obtained by trial and error until the frequency distribution of the observed and generated annual rainfalls agrees. The generated daily rainfall in each year is multiplied by the following ratio:

$$Ratio_i = \{M + (T_i - M)F\}/T_i \quad (4)$$

where  $M$  = the observed mean annual rainfall, and  
 $T_i$  = the generated annual rainfall for year  $i$ .

Since the slope of the frequency curve is proportional to the standard deviation, the adjustment factor can be directly obtained as a ratio of the standard deviation of the observed and generated annual rainfall.

$$F = StDev_o / StDev_g \quad (5)$$

The standard deviation of the generated annual rainfall is estimated from a number of replicates and averaged. In this study, we used generated mean annual rainfall averaged from a number of replicates instead of observed mean annual rainfall in Equation (4).

### 3.3 Daily and Monthly Mixed (DMM) model

In the DMM model, the occurrence of rainfall is determined by a first order Markov chain using the two transitional probabilities:  $p_{W|D}$ , the conditional probability of a wet day given that the previous day was dry;  $p_{W|W}$ , the conditional probability of a wet day given that the previous day was wet. The unconditional probability of a wet day can be derived as

$$\pi = \frac{P_{W|D}}{1 + p_{W|D} - p_{W|W}} \quad (6)$$

The rainfall depth is obtained from a Gamma distribution whose probability density function is given by

$$f(x) = \frac{(x/\beta)^{\alpha-1} \exp(-x/\beta)}{\beta \Gamma(\alpha)} \quad (7)$$

where  $\alpha$  is the shape parameter and  $\beta$  the scale parameter. It is assumed that the rainfall amounts in different wet days are unrelated. The mean and variance of the Gamma distribution are given by

$$\mu(x) = \alpha\beta \quad (8)$$

$$\sigma^2(x) = \alpha\beta^2 \quad (9)$$

Seasonal variation is modelled by allowing the model parameters to vary with each calendar month.

Given a two-state first-order Markov chain model for rainfall occurrence and the Gamma distribution model for daily rainfall amount, the expectation of the rainfall total,  $X$ , over a month of  $N$  days is

$$\mu(X) = N\pi\alpha\beta \quad (10)$$

and the variance of  $X$  is (Katz 1985)

$$\sigma^2(X) \approx N\pi\alpha\beta^2 \left[ 1 + \alpha(1-\pi) \frac{1 + p_{W|W} - p_{W|D}}{1 - p_{W|W} + p_{W|D}} \right] \quad (11)$$

The daily and monthly mixed model of Wang and Nathan involves the following steps:

- Step 1: For month  $k$  ( $k = 1, \dots, 12$ ), generate a sequence of wet and dry days for the whole month using a two-state first order Markov chain.
- Step 2: For any wet day in that month, generate a daily rainfall amount  $x^d$  from a Gamma distribution with parameters  $\alpha = \alpha^d$  and  $\beta = \beta^d$  which are estimated from the mean and variance of daily rainfall amounts by using Equations (8) and (9).
- Step 3: For the same wet day, generate a twin rainfall amount  $x^m$  with exactly the same probability of non-exceedence from another Gamma distribution with parameters  $\alpha = \alpha^m$  and  $\beta = \beta^m$  which are estimated from the mean and variance of monthly rainfall using Equations (10) and (11).
- Step 4: Manipulate the monthly total of the daily rainfall generated in Step 3,  $\tilde{X}_i = \sum x^m$ , to produce a new monthly total  $X_i$  by using the Thomas-Fiering monthly model.

$$\frac{X_i - \mu(X_i)}{\sigma(X_i)} = \rho \frac{X_{i-1} - \mu(X_{i-1})}{\sigma(X_{i-1})} + (1 - \rho^2)^{1/2} \frac{\tilde{X}_i - \mu(X_i)}{\sigma(X_i)} \quad (12)$$

The subscripts  $i-1$  and  $i$  in Equation (12) denote the previous and current months respectively. The lag one autocorrelation is estimated from the non-seasonal data.

- Step 5: Produce a new daily rainfall series  $x$  for that month by multiplying all the  $x^d$  by a factor ( $X_i / \sum x^d$ ).

The core of the model is to generate two daily rainfall series which resemble each other, the first reproducing daily statistics and the second reproducing monthly statistics, and subsequently use the second series (after incorporation of autocorrelation in monthly rainfall) to adjust the first series.

### 3.4 Modified Daily and Monthly Mixed (DMMm) model

In this study, a variation of the DMM model (DMMm) was formulated to investigate the necessity of generating two daily rainfall series. The procedure of the DMMm model is very similar to the DMM model except that Step 3 is bypassed. The steps are:

Step 1 & 2: As the DMM model

- Step 3: Manipulate the monthly total of the daily rainfall generated in Step 2,  $\tilde{X}_i = \sum x^d$ , to produce a new monthly total  $X_i$  by using the Thomas-Fiering monthly model.

$$\frac{X_i - \mu(X_i)}{\sigma(X_i)} = \rho \frac{X_{i-1} - \mu(X_{i-1})}{\sigma(X_{i-1})} + (1 - \rho^2)^{1/2} \frac{\tilde{X}_i - \mu'(X_i)}{\sigma'(X_i)} \quad (13)$$

The mean  $\mu'(X_i)$  and standard deviation  $\sigma'(X_i)$  used in Equation (13) are obtained from Equations (10) and (11) using the daily Gamma parameters obtained in Step 2.

Step 4: Produce a new daily rainfall series  $x$  for that month by multiplying all  $x^d$  by a factor  $(X_i / \sum x^d)$ .

In this variation, only one sequence rather than two of daily rainfalls is generated but at the same time the daily rainfalls are adjusted to match the monthly characteristics. The lag-1 autocorrelation in monthly rainfall is considered by Equation (13).

## 4. Model Evaluation

A key requirement in stochastic data generation is that the synthetic sequences be statistically consistent with the observed characteristics of the historical record. In this section, we present the parameters for model evaluation and discuss the performance of the TPM, TPMb, DMM and DMMm models described in Section 3 in respect of annual, monthly and daily statistics of rainfall.

### 4.1 Parameters for model evaluation

The performance of the daily models is evaluated using a number of statistics. These include:

- Annual
  - (1) Mean annual rainfall (mm)
  - (2) Standard deviation of annual rainfall (SD)
  - (3) Coefficient of skewness (CS)
  - (4) Lag one auto-correlation ( $r$ )
  - (5) Maximum annual rainfall
  - (6) Minimum annual rainfall
  - (7) Minimum 2-, 3-, 5-, 7-, and 10-year rainfall sums
  - (8) Adjusted range
- Monthly
  - (9) Mean monthly rainfall (mm)
  - (10) Standard deviation of monthly rainfall (SD)
  - (11) Coefficient of skewness of monthly rainfall (CS)
  - (12) Serial correlation of monthly rainfall ( $r$ )
  - (13) Maximum monthly rainfall
  - (14) Minimum monthly rainfall
  - (15) Frequency of months with no rain
  - (16) Average number of wet days for each month
- Daily
  - (17) Mean daily rainfall for each month (mm)
  - (18) Standard deviation of daily rainfall for each month (SD)

- (19) Coefficient of skewness of daily rainfall for each month (CS)
- (20) Mean daily rainfall on a solitary wet day for each month (mm)
- (21) Mean daily rainfall on wet days bounded on one side by a wet day for each month (mm)
- (22) Mean daily rainfall on wet days bounded on each side by a wet day for each month (mm)
- (23) Correlation between rainfall depth and duration of wet spells over one or more days
- (24) Mean dry spell length for each month (days)
- (25) Standard deviation of dry spell length for each month
- (26) Coefficient of skewness of dry spell length for each month
- (27) Mean wet spell length for each month (days)
- (28) Standard deviation of wet spell length for each month
- (29) Coefficient of skewness of wet spell length for each month
- (30) Longest dry spell for each month (days)
- (31) Longest wet spell for each month (days)
- (32) Longest wet and dry spells using different threshold values (days)
- (33) Mean of monthly maximum daily rainfall
- (34) Maximum daily rainfall for each month
- (35) Coefficient of skewness of monthly maximum daily rainfall
- (36) Frequency distribution of annual maximum daily rainfall
- (37) Mean of annual maximum daily rainfall
- (38) CV of annual maximum daily rainfall
- (39) Coefficient of skewness of annual maximum daily rainfall

These parameters represent annual statistics from (1) to (8), monthly statistics from (9) to (16), and daily statistics from (17) to (32). The parameters from (33) to (39) represent the annual, monthly and daily statistics of maximum daily rainfall.

The mean, standard deviation, skewness and lag-1 autocorrelation are estimated from the following expressions.

$$\bar{x} = \frac{1}{n} \sum_{t=1}^n x_t \quad (14)$$

$$SD = \sqrt{\frac{1}{(n-1)} \sum_{t=1}^n (x_t - \bar{x})^2} \quad (15)$$

$$CS = \frac{n}{(n-1)(n-2)SD^3} \sum_{t=1}^n (x_t - \bar{x})^3 \quad (16)$$

$$r = \frac{1}{(n-1)SD^2} \sum_{t=1}^{n-1} (x_{t+1} - \bar{x})(x_t - \bar{x}) \quad (17)$$

In the above equations,  $x_t$  represents the annual, monthly or daily rainfall and  $n$  the number of data. The mean estimated from the historical data was used in computing the standard deviation, skewness and correlation. For convenience, the maximum and minimum rainfall

were standardised by dividing by the historical mean rainfall. The adjusted range ( $R$ ) was obtained from

$$R = \max\{D_k\} - \min\{D_k\} \quad k = 1, 2, \dots, n \quad (18)$$

where  $D_k = \sum_{t=1}^k (x_t - \bar{x})$

## 4.2 Results

One hundred replicates each of length equal to the historical data were generated using the models described in Section 3. Table 4 contains a selection of annual, monthly and daily statistics, which are the average of the parameters from the 100 replicates along with the historical values. The number in bracket represents the reference number to the statistics in the above section.

For rest of this section, we discuss the model performance in respect of annual, monthly and daily statistics based on Table 4 and Figures A1 – A12. Table 5 presents the results of model assessment for selected statistics. The first column in the table provides the reference number to the statistics in the above section.

As noted in the next section, TPMb model outperformed the other three models and is recommended as the most satisfactory daily stochastic model. In view of this, the average statistics for 100 replicates were compared with the historical values for the 21 test stations in Figures A13 – A24. The range, 25%, 75% values are shown in the figures. To reduce the number of plots in Figures A5 – A12 and A17 – A24, only four months January, April, July, and October are shown.

### 4.2.1 Annual rainfall

#### Mean

The DMM and DMMm models reproduced adequately the annual mean as shown in Figure A1. The TPM and TPMb models slightly overestimated the observed means for nearly all stations, the maximum difference between historical and stochastic data being less than 4% (Table 4). Figure A13 shows the comparison of observed and TPMb model generated mean annual rainfall.

#### Standard deviation

The TPM, DMM and DMMm models underestimated the observed standard deviation of annual rainfalls in most cases as shown in Figure A2. A large discrepancy of the statistics between historical and generated data was observed at Jervis bay (#068034). The TPM model performed particularly poorly to reproduce this statistics. This shortcoming of the TPM model has been recognised (Boughton, 1999). The TPMb model, on the other hand, preserved



satisfactorily the historical standard deviation in all cases. Figure A2 shows that the DMM and DMMm models produced very similar results. Figure A14 shows the comparison of the standard deviation of annual rainfalls for historical and TPMb generated data.

### **Lag-1 auto-correlation**

All four models did not reproduce the historical lag-1 auto-correlation statistics of annual rainfalls satisfactorily, as illustrated in Figure A3. However, the DMM and DMMm models performed slightly better for the stations with positive observed serial correlation coefficients. As indicated in Figure A15, the observed serial correlation coefficients are all within the ranges between the minimum and maximum values generated by the TPMb model.

### **Minimum 2-, 3-, 5-, 7-, and 10-year sums**

The TPM, TPMb, DMM and DMMm models reproduced the observed minimum 2-, 3-, 5-, 7-, and 10-year sums of annual rainfalls satisfactorily. Figure A4 depicts the results for the ratios of minimum 5-year sums to the observed means for the 21 stations. Figure A16 shows the comparison of observed and TPMb model generated results.

## **4.2.2 Monthly rainfall**

### **Mean**

All four models preserved the observed mean of monthly rainfall very well and Table 4 presents the results for the months January and July for the 21 stations.

### **Standard deviation**

All four models reproduced the observed standard deviation of monthly rainfall satisfactorily although the TPM model slightly underestimated the statistics for most of the 21 stations. The detailed results are not presented here.

### **Skewness**

The TPM, TPMb, DMM and DMMm models reproduced the observed coefficients of skewness of monthly data satisfactorily (although underestimated) in most cases as shown in Figure A5 for the months January, April, July and October. Figure A17 presents the comparison of observed and TPMb model generated results for the same months as Figure A5.

### **Number of months with no rain**

The three models – TPM, TPMb and DMMm – reproduced adequately the observed number of months with no rain as shown in Figure A6, but the DMM model could not preserve the statistics for some stations. In Figure A18 the comparison of observed and TPMb model generated results shows that the number of months of no rain are preserved in the TPMb model.

### **Number of wet days**

As shown in Figure A7, all the models satisfactorily preserved the number of wet days for the 21 stations. Figure A19 presents the comparison of observed and generated results for the TPMb model.

### **4.2.3 Daily rainfall**

#### **Mean daily rainfall**

The TPM, TPMb, DMM and DMMm models adequately reproduced the observed mean in most cases as shown in Figure A8. Figure A20 presents the comparison of historical and TPMb model generated results.

#### **Mean daily rainfall for different types of wet days**

Chapman (1994) has mentioned the importance of modelling solitary, one-sided and two-sided wet days. The TPM and TPMb models adequately reproduced the observed mean daily rainfall for solitary wet days in all cases as shown in Figure A9 and a comparison of observed and TPMb model generated results is shown in Figure A21. However, the DMM and DMMm models overestimated the mean in most cases.

To improve the performance of the DMM model on this statistics, another modified version was investigated. In this version of the DMM model (DMM3), the rainfall depth is obtained from three separate Gamma distributions for solitary, one-side and two-side rain days, using the three following guidelines:

- For each month, if the numbers of wet days for all three cases - solitary, one-side and two-side rain days, are not less than 18, use three separate Gamma distributions.
- If the number of solitary days or total number of one-side and two-side rain days is less than 18, use single Gamma distribution only for all three cases.
- If the total number of one-side and two-side rain days is not less than 18, but the number of either one-side rainfall or double-side rainfall is less than 18, use two separate Gamma distributions, one for solitary days and one for the other two cases.

In this study, based on the data we arbitrarily chose 18 rain days as a threshold for simplicity. However, this number should be used as a guideline.

Figure A9 shows that the DMM3 model significantly improved the performance of the DMM model, but still overestimated the statistics.

#### **Mean wet spell length**

All four models satisfactorily preserved the historical mean wet spells in all stations as shown in Figure A10. Figure A22 presents the comparison of observed and TPMb model generated results. Overall this is satisfactory, although the generation for a July initiated spell is not adequately preserved for Roebourne.

### **Longest dry spell**

The TPM, TPMb and DMMm models reproduced the observed longest dry spells satisfactorily in most cases as shown in Figure A11, but the DMM model overestimated the parameter in some cases. Figure A23 shows the TPMb model generated data preserves the historical longest dry spell.

### **Maximum daily rainfall**

Finally all the models reproduced the observed maximum daily rainfall satisfactorily in most cases as shown in Figure A12. Figure A24 presents the comparison of observed and TPMb model generated results.

## **5. Conclusions**

The daily stochastic models were assessed by a number of statistics at annual, monthly and daily levels using rainfall data from 21 stations across Australia. The TPM model was found to preserve key statistical characteristics of the historical rainfall at the annual, monthly and daily levels so long as the Boughton adjustment is applied. The DMM model performed well in reproducing these statistics. The DMM and DMMm models produced very similar results in terms of the above analysis.

The TPM model preserved the rainfall depth for solitary wet days, which the DMM model did not handle satisfactorily. This factor is considered important for estimating runoff using generated daily rainfall in a rainfall-runoff model. The DMM model can be improved using a Gamma distribution with three sets of parameters for the rainfall amounts on solitary, one-side and two-side rain days. Another important difference between the two types of models was that the DMM/DMMm model performed slightly better in reproducing the serial correlation, although all the four models were unable to preserve this statistic satisfactorily.

## 6. References

Boughton, W. C. 1999. A daily rainfall generating model for water yield and flood studies. Report 99/9, CRC for Catchment Hydrology, Monash University, Melbourne, 21pp.

Chapman, T. 1994. Stochastic models for daily rainfall. In: Water Down Under 1994. Institution of Engineers Australia, Adelaide, 7-12.

Katz, R. W. 1985. Probabilistic models. In Probability, statistics and decision making in the atmospheric sciences, edited by A. H. Murphy and R W Katz, Westview, 251-288.

Srikanthan, R. and McMahon T.A., 1985. Stochastic generation of rainfall and evaporation data. AWRC Technical Paper No. 84, 301pp.

Srikanthan, R. and McMahon T.A., 2000. Stochastic generation of climate data: a review. Report 00/16, CRC for Catchment Hydrology, Monash University, Melbourne, 41pp.

Wang Q.J. and Nathan R.J., 2002. A daily and monthly mixed algorithm for stochastic generation of rainfall time series. Proceedings of the 27<sup>th</sup> Hydrology and Water Resources Symposium, Melbourne, 20 - 23 May 2002.





Table 5: Assessment of daily models.

<i>No</i>	<i>Parameters</i>	<i>TPM</i>	<i>TPMb</i>	<i>DMM</i>	<i>DMMm</i>
<b>Annual</b>					
(1)	Mean	✓	✓	✓	✓
(2)	SD	✗	✓	✓	✓
(4)	Lag-1 Serial Corr	✗	✗	✗	✗
(7)	Min <i>n</i> -years	✓	✓	✓	✓
<b>Monthly</b>					
(9)	Mean	✓	✓	✓	✓
(10)	SD	✓	✓	✓	✓
(11)	CS	✓	✓	✓	✓
(13)	Max	✓	✓	✓	✓
(14)	Min	✓	✓	✓	✓
(15)	No rain months	✓	✓	✗	✓
(16)	No of wet days	✓	✓	✓	✓
<b>Daily</b>					
(17)	Mean	✓	✓	✓	✓
(18)	SD	✓	✓	✓	✓
(19)	CS	✓	✓	✓	✓
(20)	Solitary rainfall	✓	✓	✗	✗
(21)	One-sided rainfall	✓	✓	✗	✗
(22)	Two-sided rainfall	✓	✓	✓	✓
(23)	Corr depth & dur	✓	✓	✓	✓
(24)	Mean dry spell	✓	✓	✓	✓
(25)	SD of dry spell	✓	✓	✓	✓
(27)	Mean wet spell	✓	✓	✓	✓
(28)	SD of wet spell	✓	✓	✓	✓
(30)	Longest dry spell	✓	✓	✗	✓
(31)	Longest wet spell	✓	✓	✓	✓
(34)	Max daily rainfall	✓	✓	✓	✓

TPM – Transition probability matrix method (Srikanthan and McMahon, 1985)

TPMb – Boughton's adjustment to the above model

DMM – Daily and monthly mixed algorithm (Wang and Nathan, manuscript)

DMMm – Modified DMM model

✓ – adequately modelled

✗ – poorly modelled

## **Appendix A: Figures – Annual, Monthly and Daily Rainfall Data and Model Results**

(All statistics calculated from generated data are based on one hundred replicates)



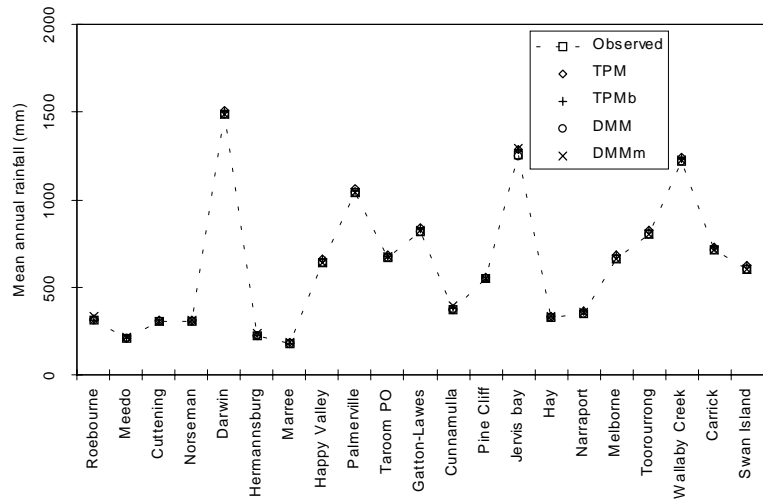


Figure A1: Mean annual rainfall.

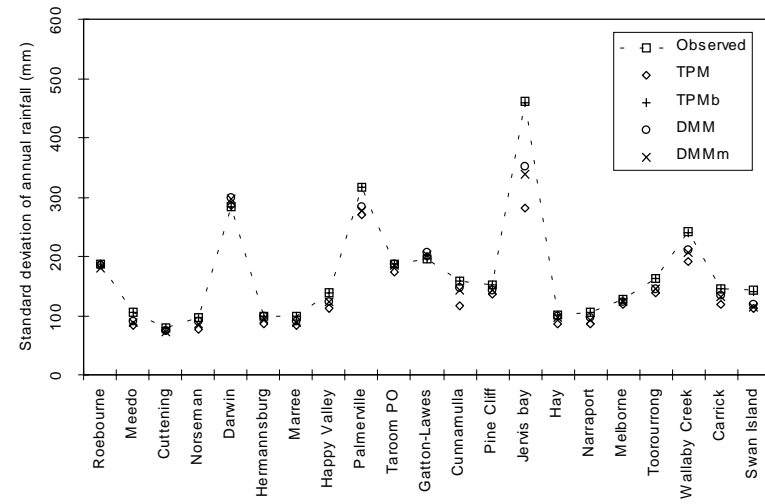


Figure A2: Standard deviation of annual rainfall.

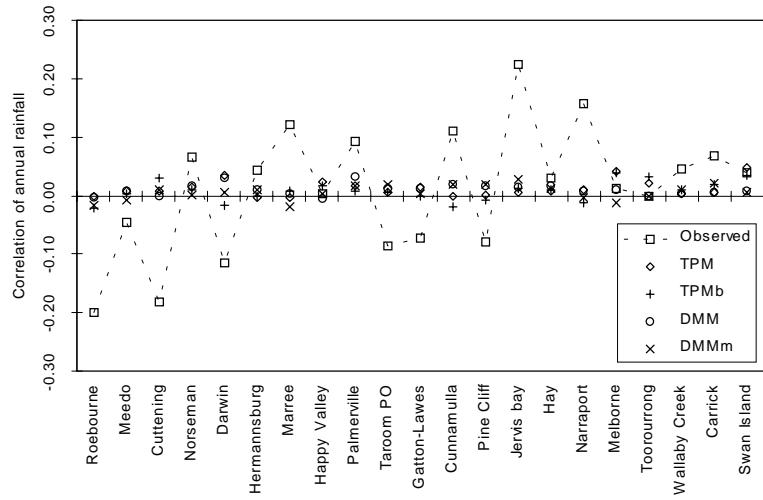


Figure A3: Lag one auto-correlation.

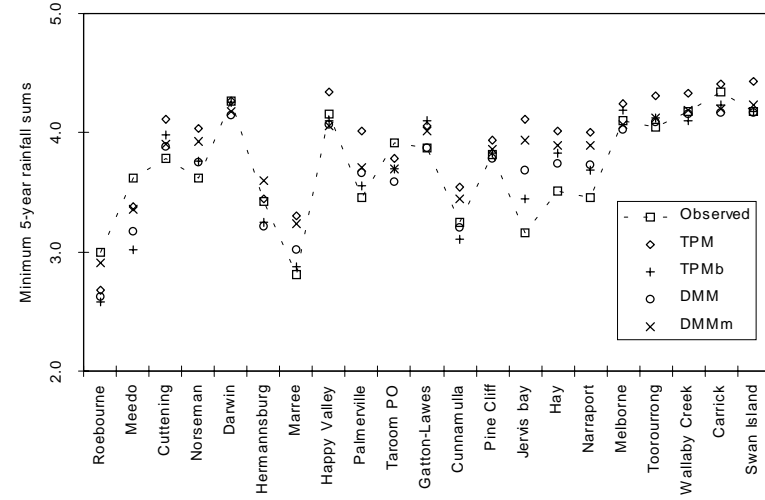


Figure A4: Minimum 5-year rainfall sums.

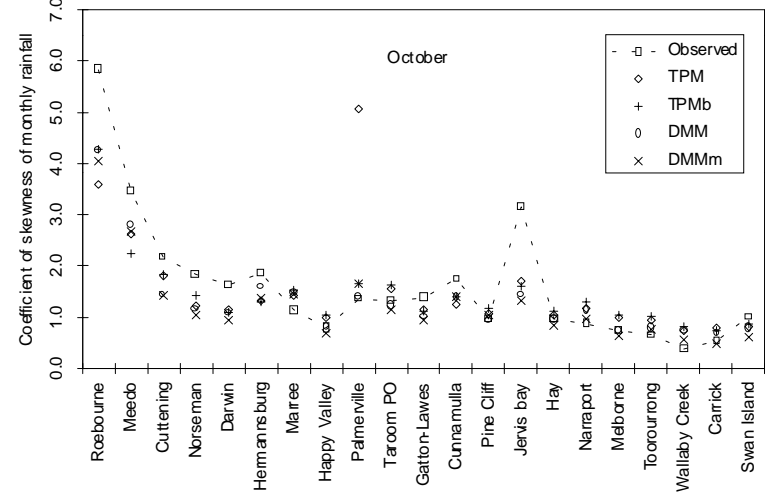
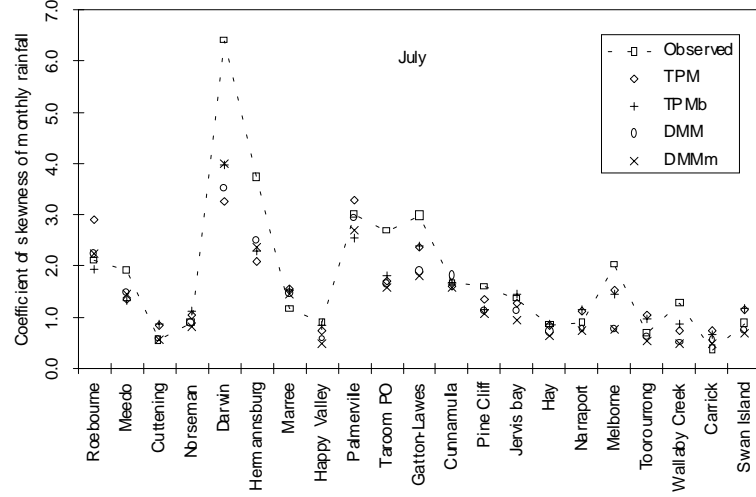
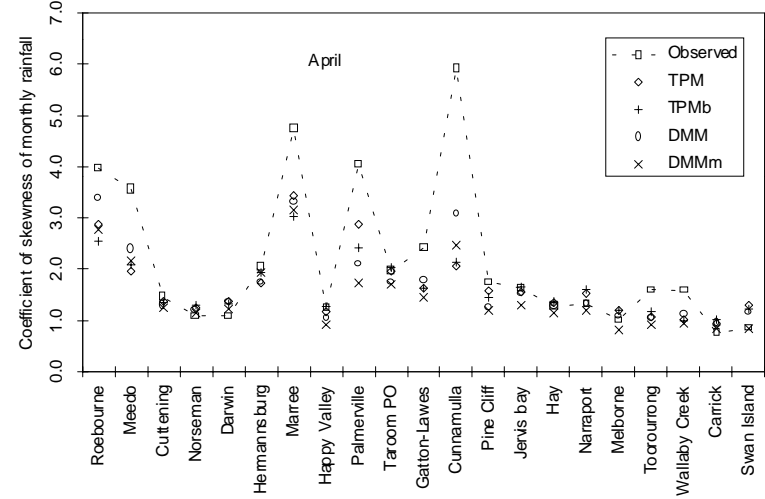
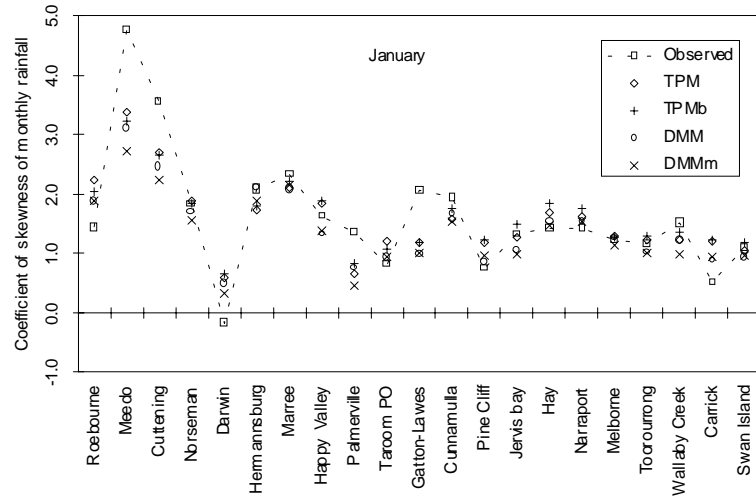


Figure A5: Coefficient of skewness of monthly rainfall.

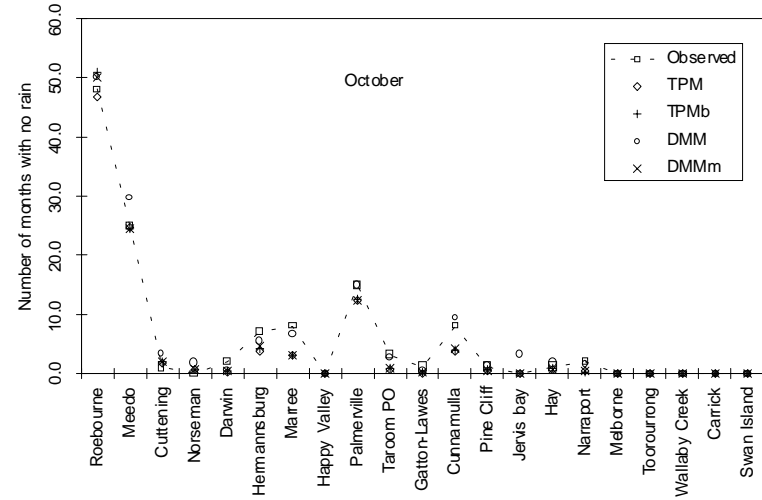
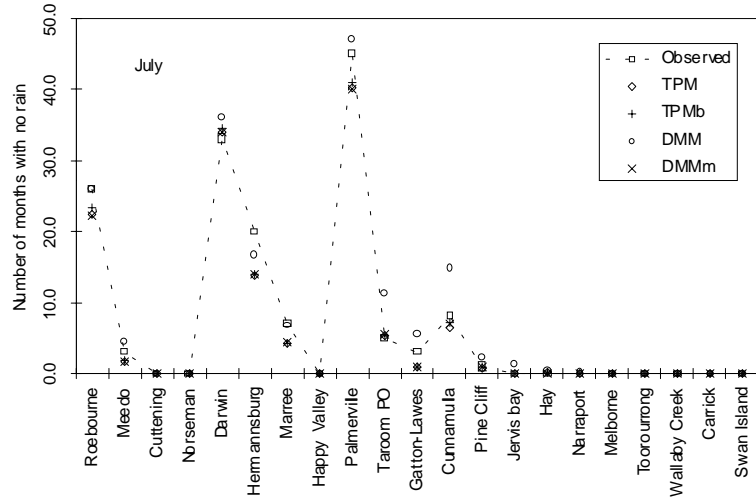
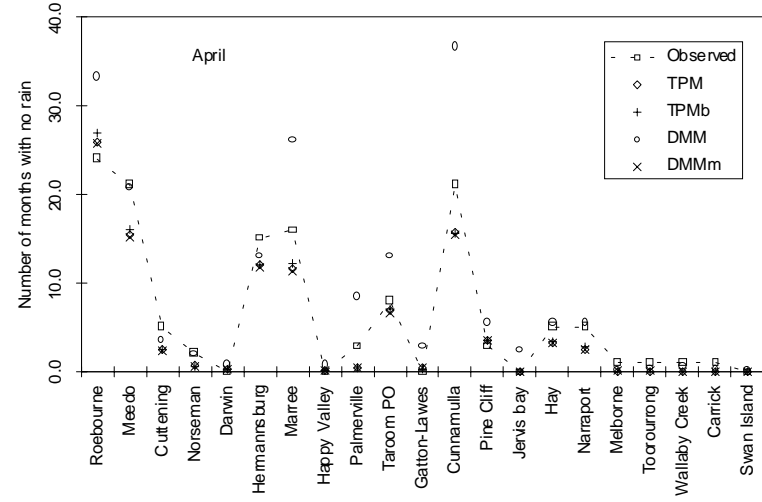
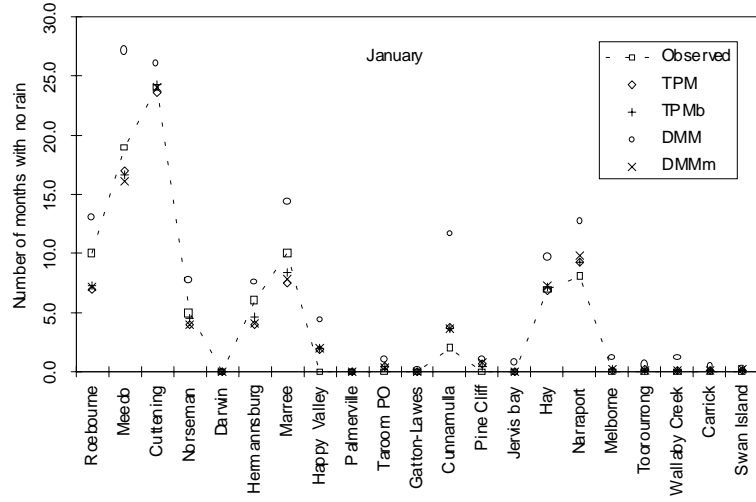


Figure A6: Number of months with no rain.

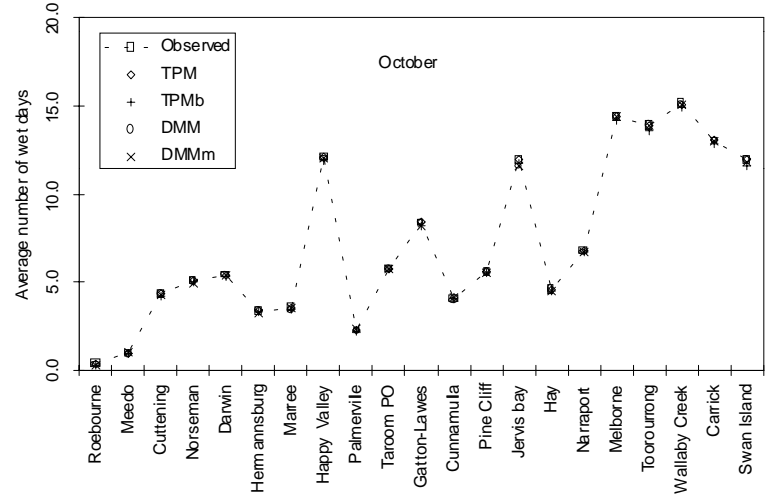
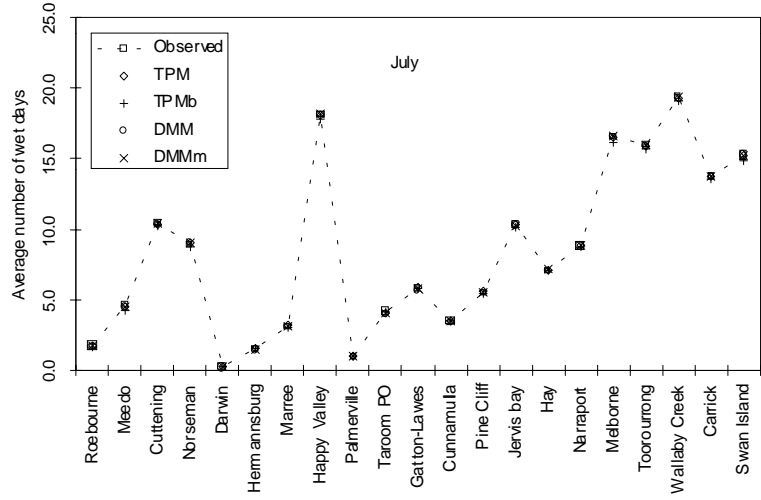
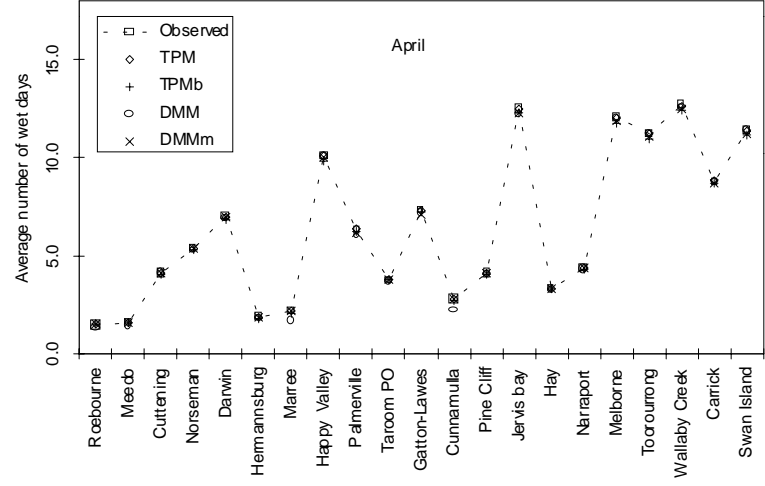
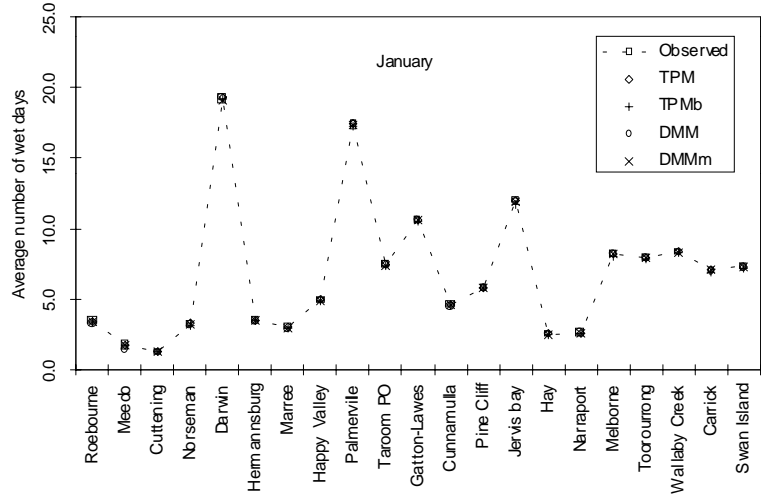


Figure A7: Average number of wet days.

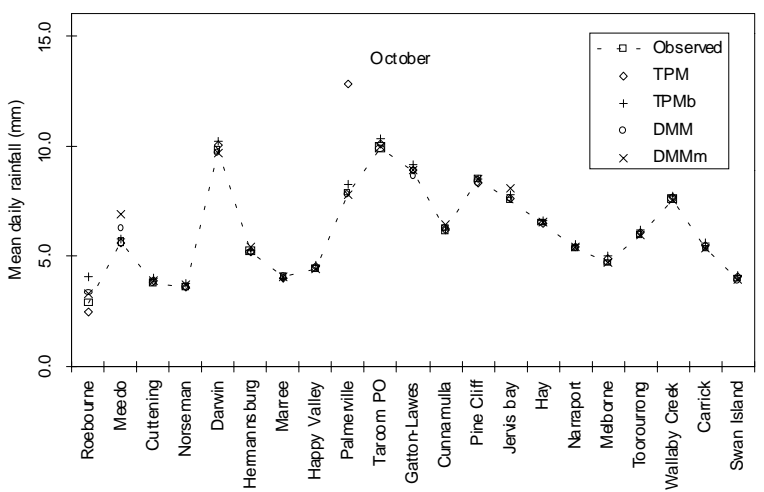
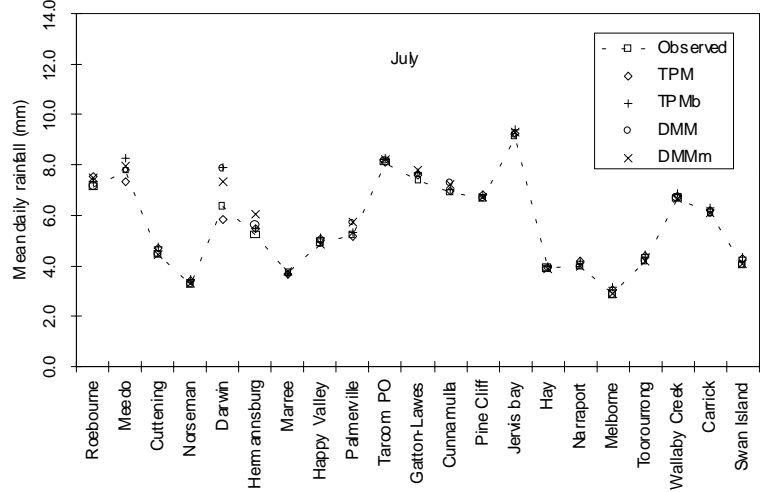
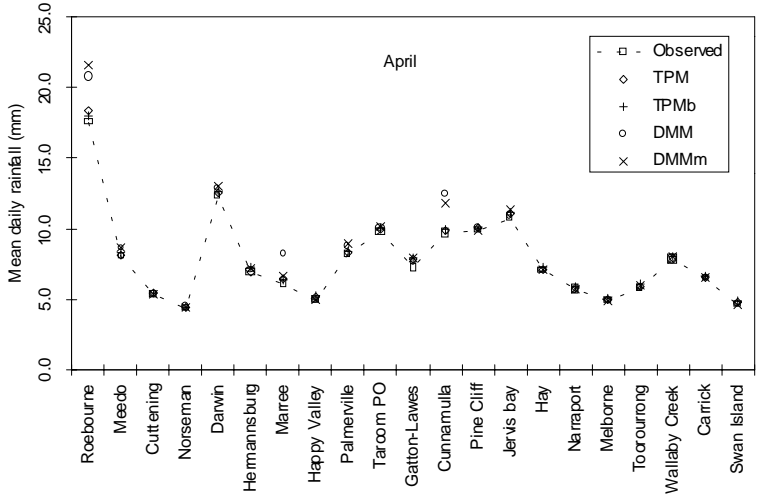
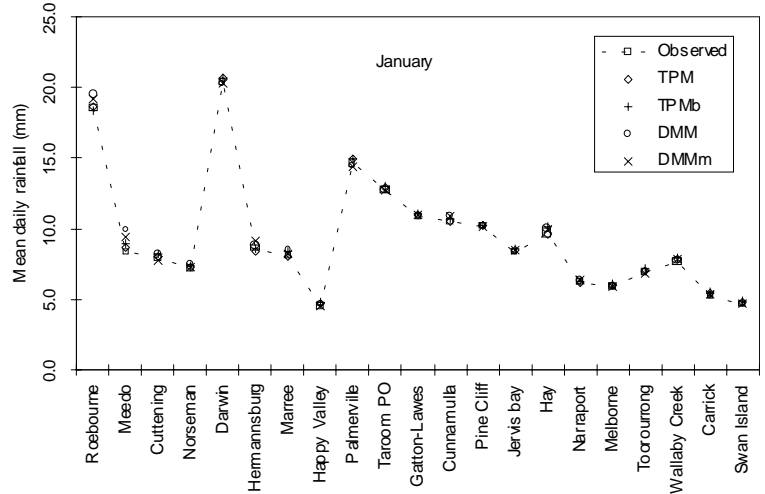


Figure A8: Mean daily rainfall.

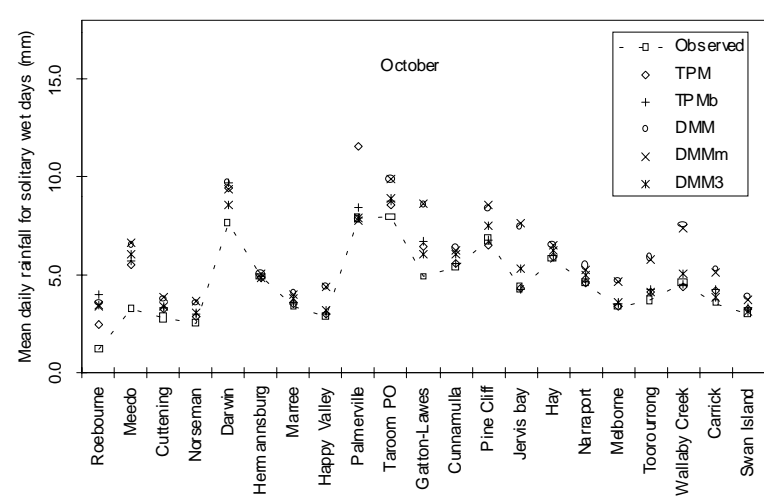
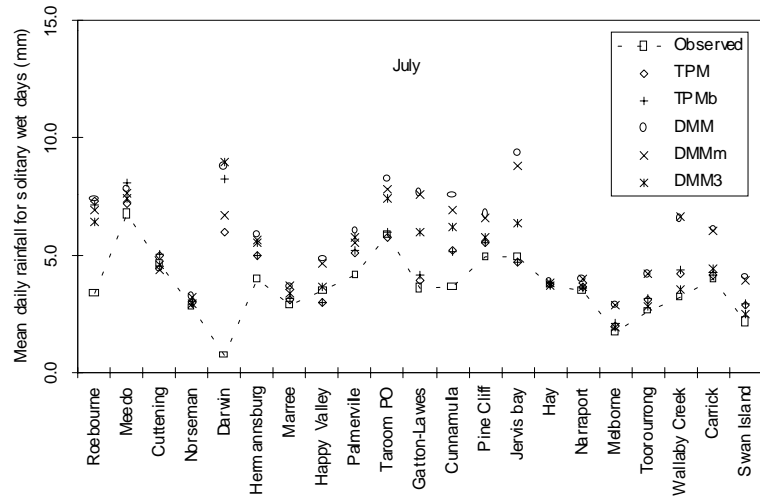
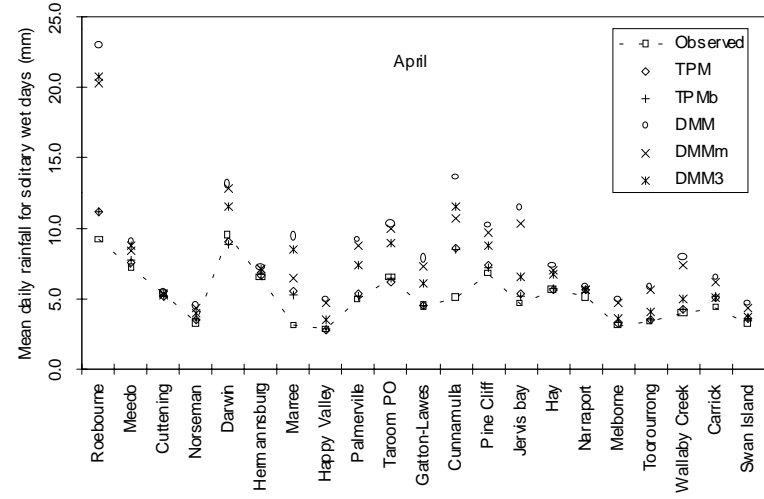
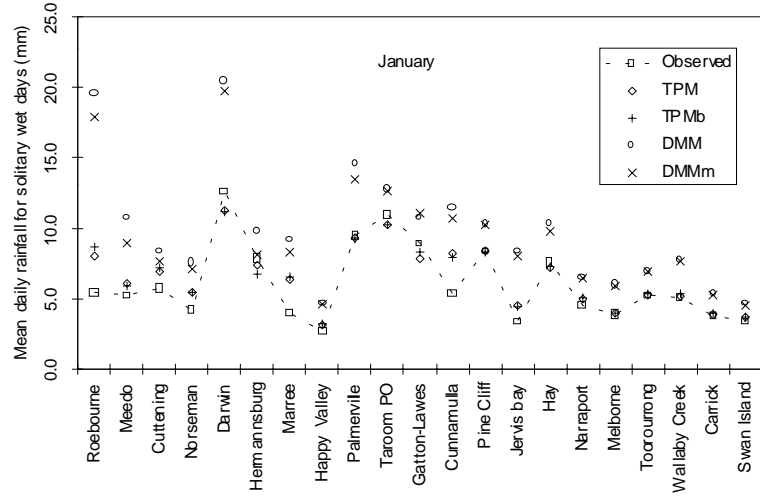


Figure A9: Mean daily rainfall for solitary wet days.

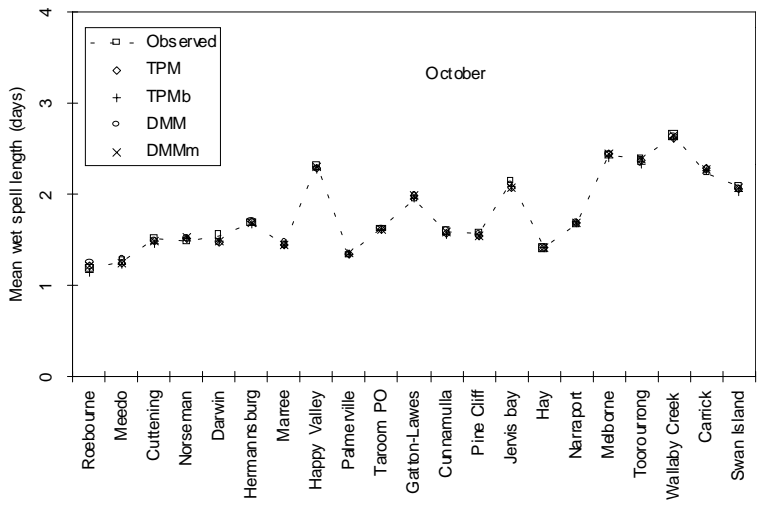
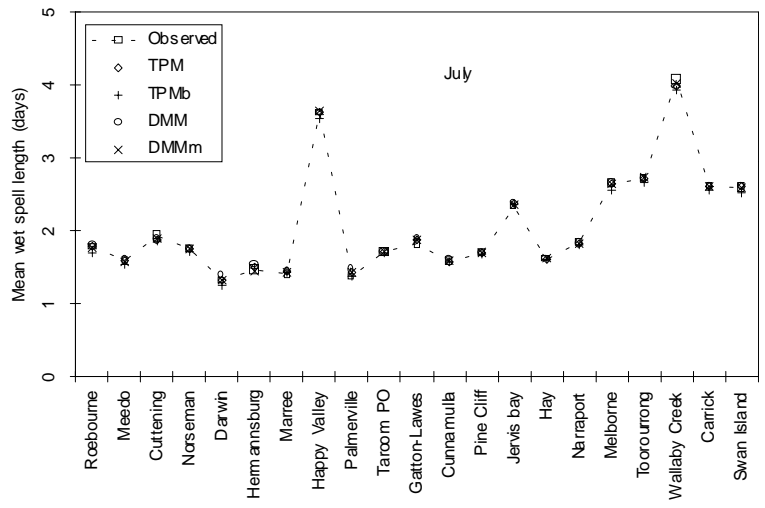
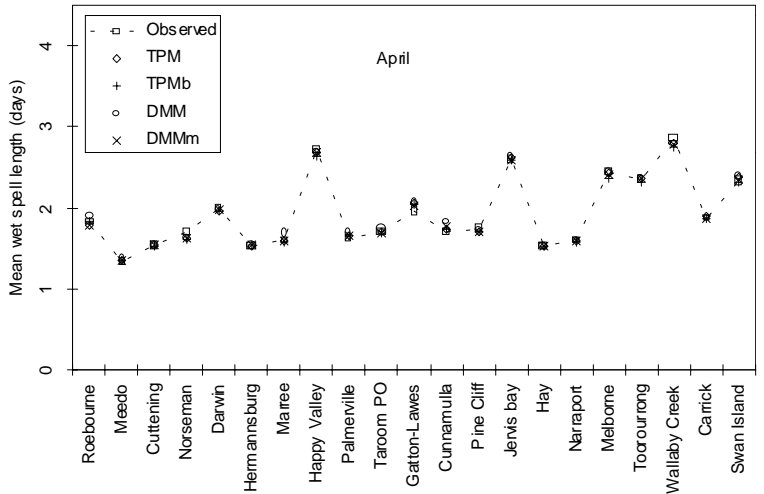
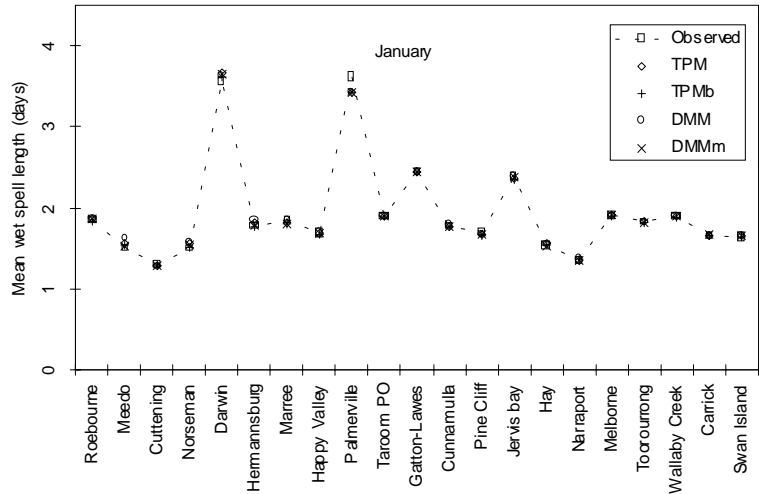


Figure A10: Mean wet spell length.

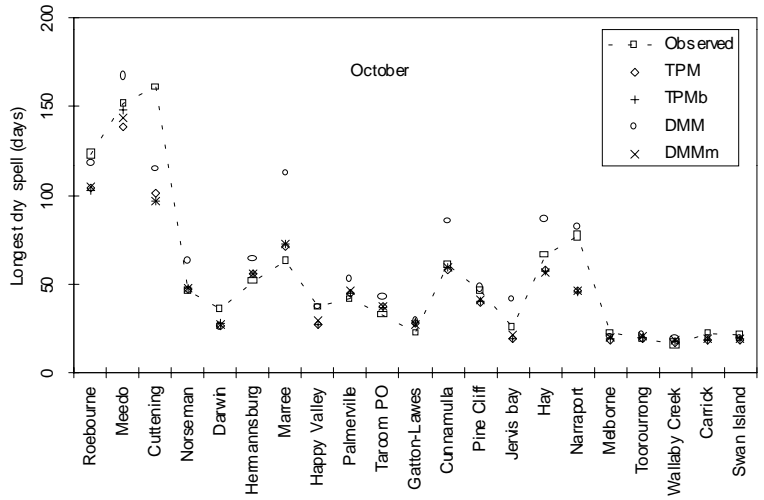
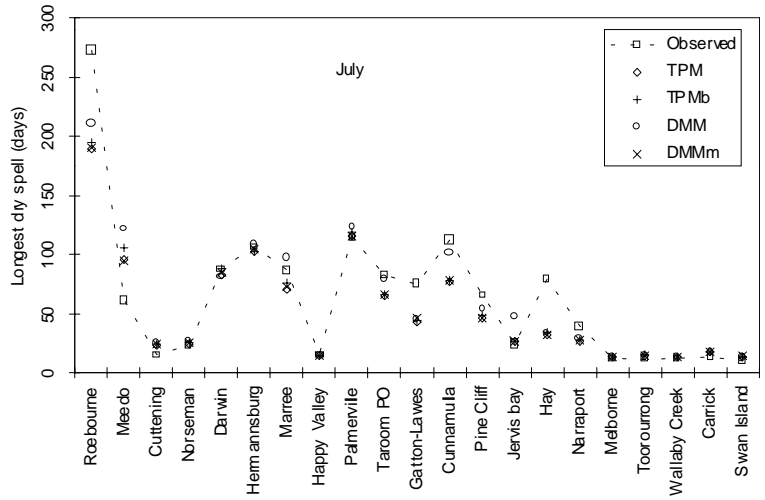
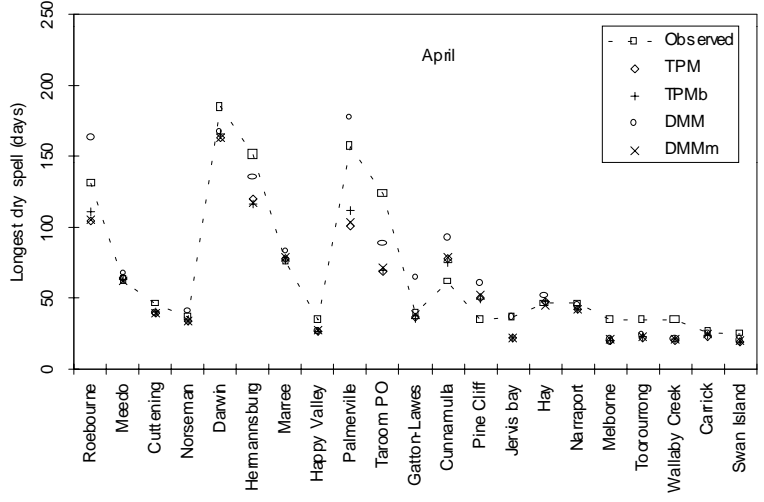
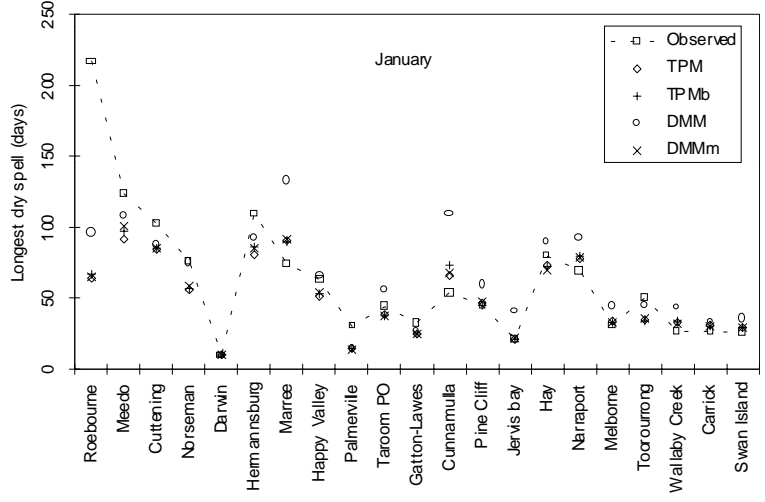


Figure A11: Longest dry spell.



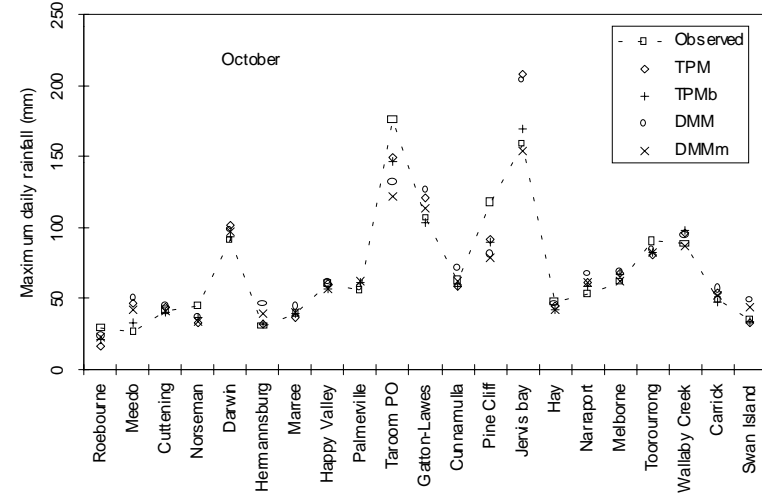
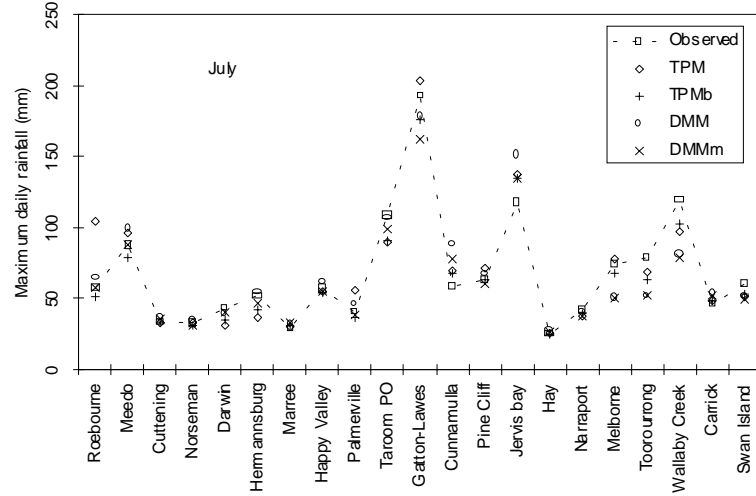
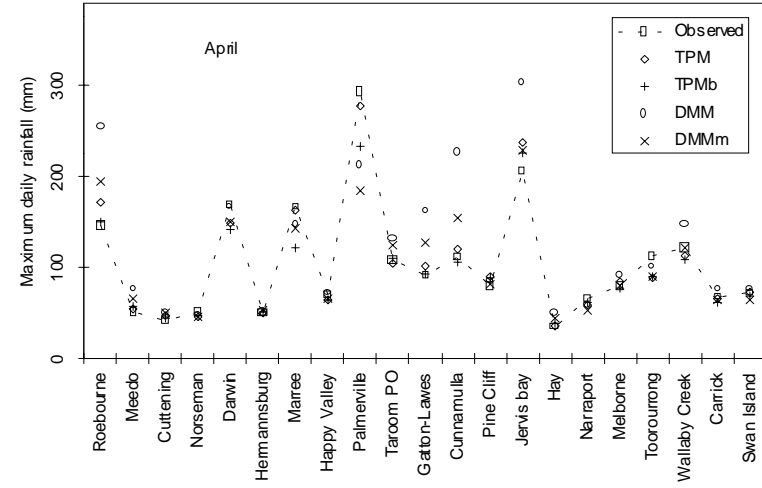
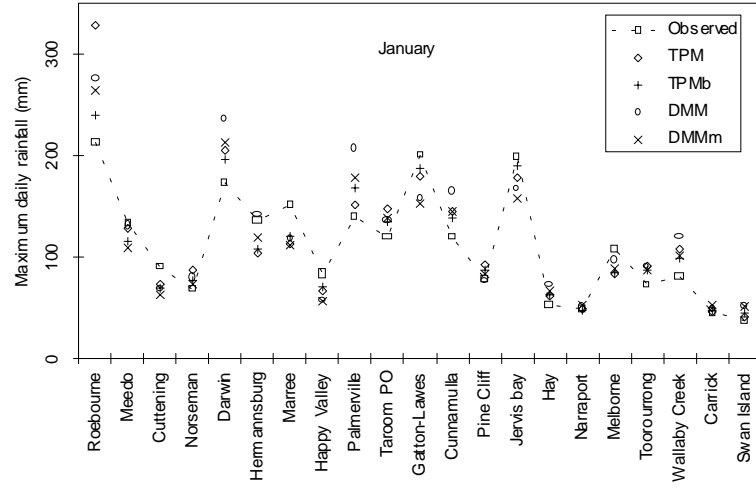


Figure A12: Maximum daily rainfall.

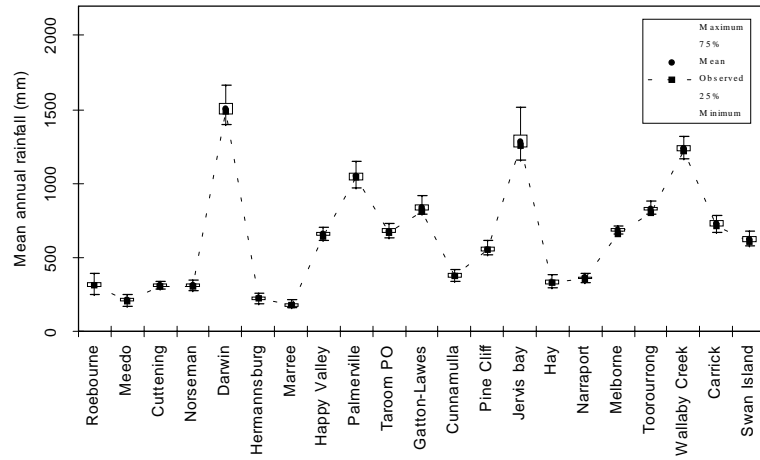


Figure A13: Comparison of observed and TPMb model generated mean annual rainfall.

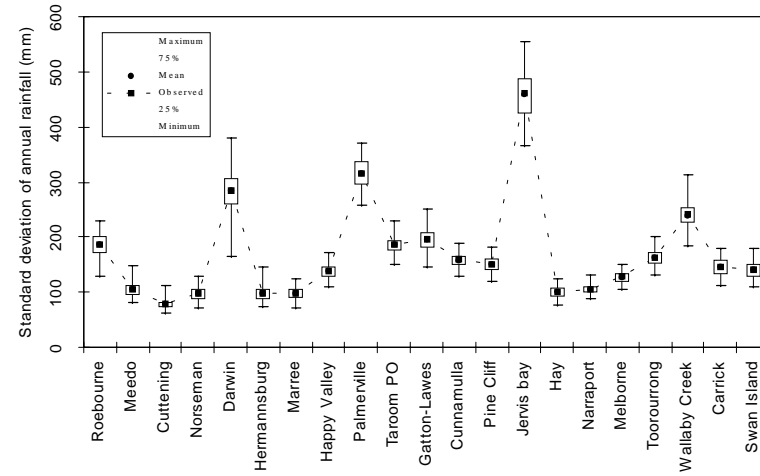


Figure A14: Comparison of observed and TPMb model generated standard deviation of annual rainfall.

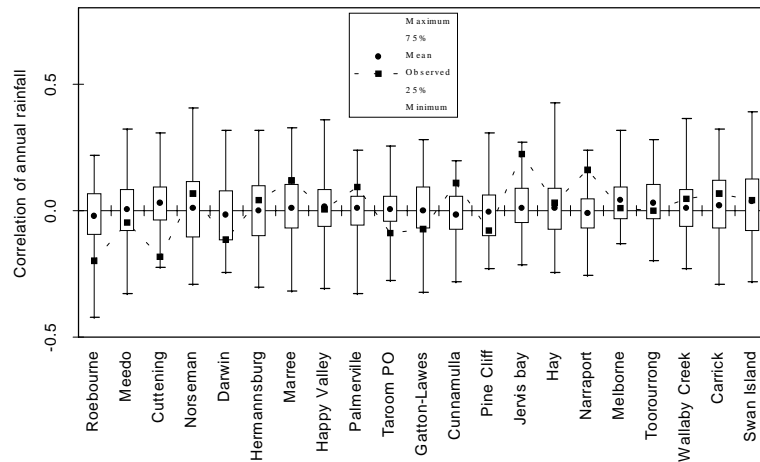


Figure A15: Comparison of observed and TPMb model generated lag one auto-correlation of annual rainfall.

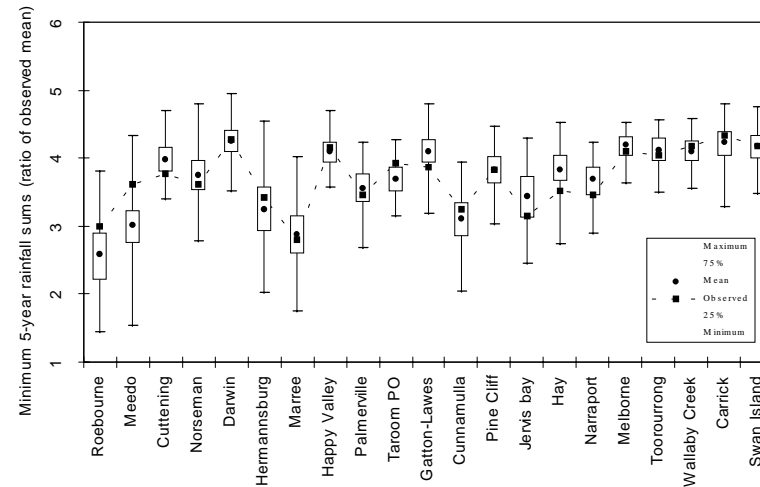


Figure A16: Comparison of observed and TPMb model generated minimum 5-year rainfall sums.

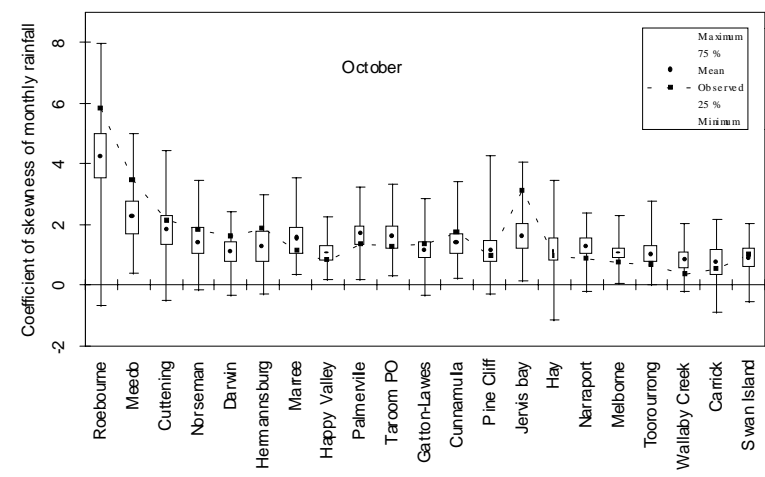
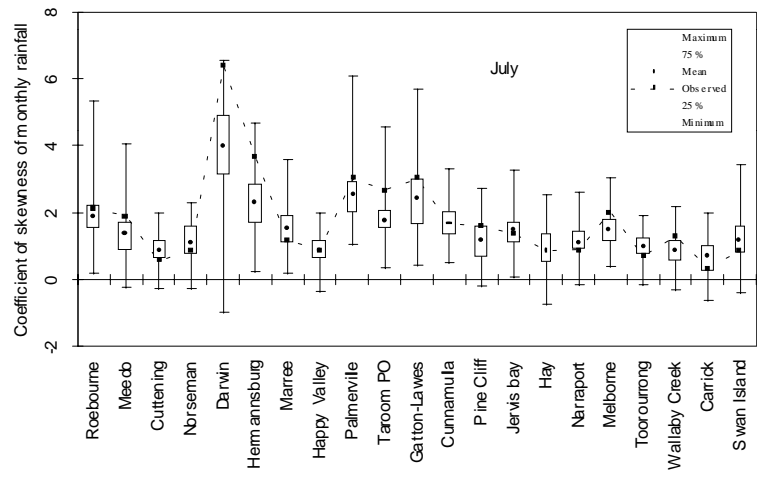
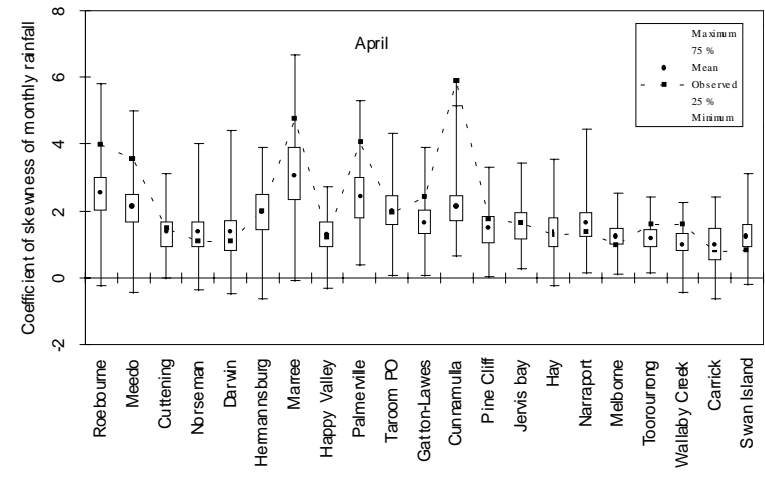
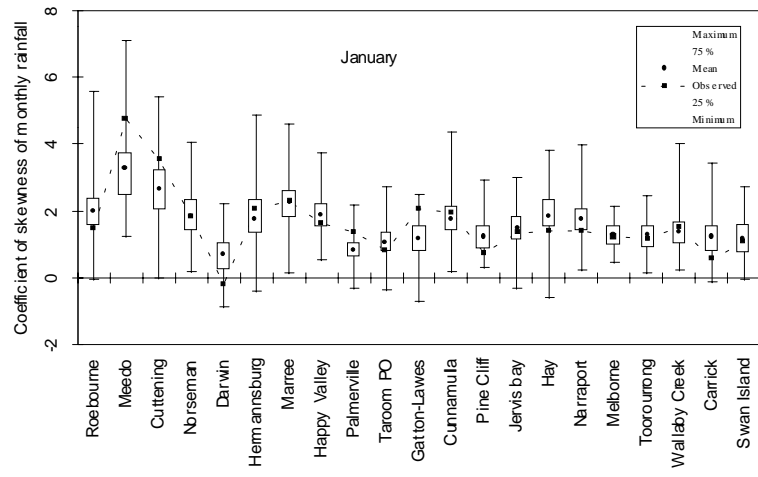


Figure A17: Comparison of observed and TPMb model generated coefficient of skewness of monthly rainfall.

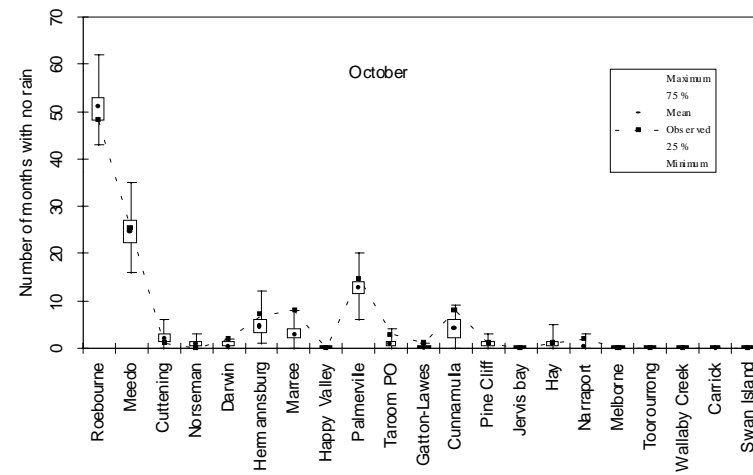
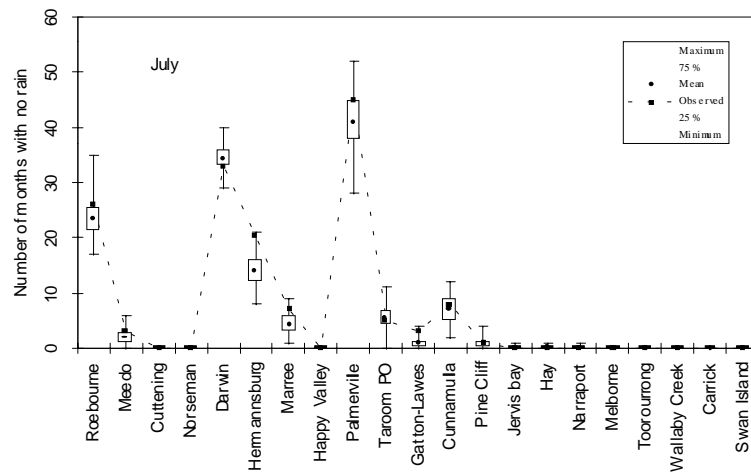
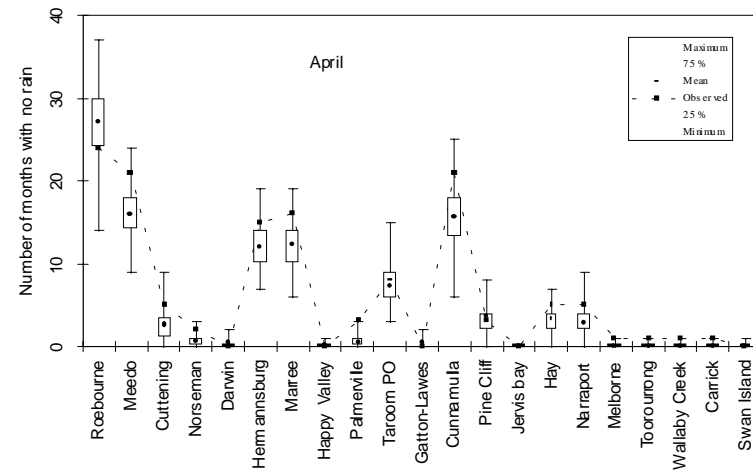
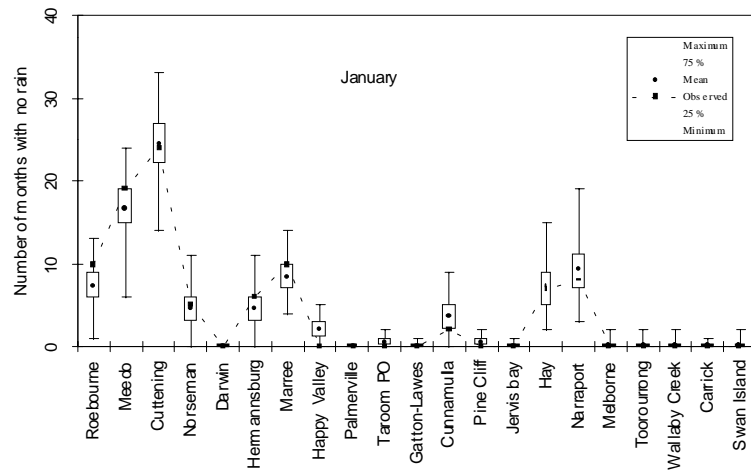


Figure A18: Comparison of observed and TPMb model generated number of months with no rain.

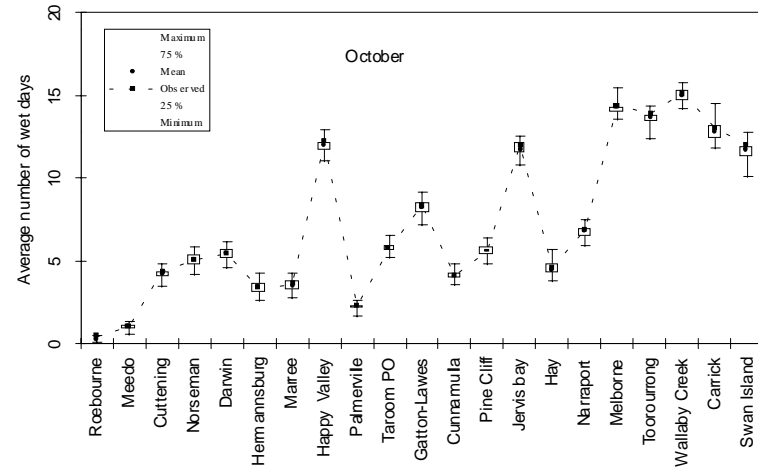
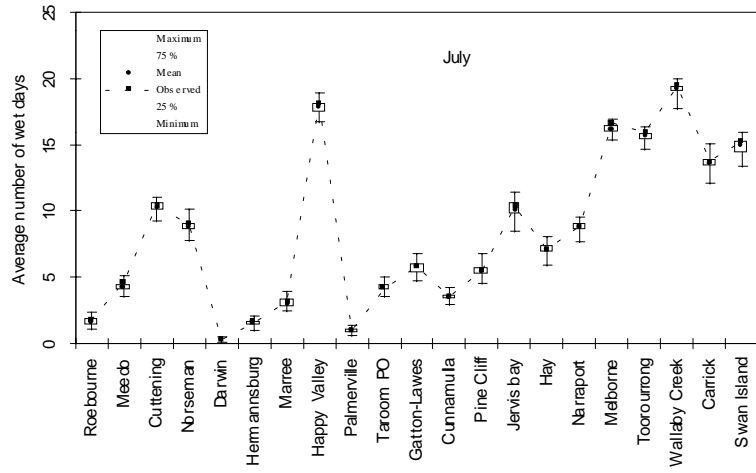
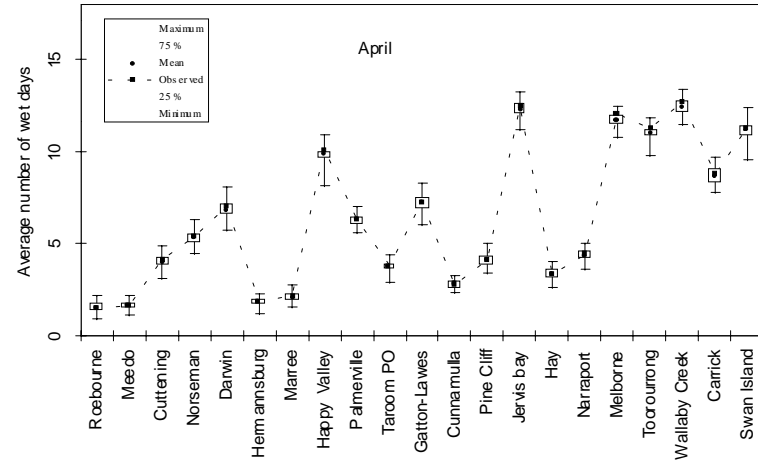
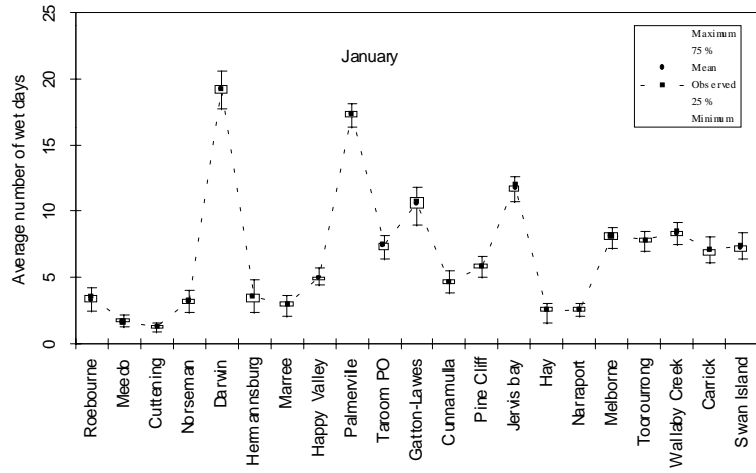


Figure A19: Comparison of observed and TPMb model generated average number of wet days.

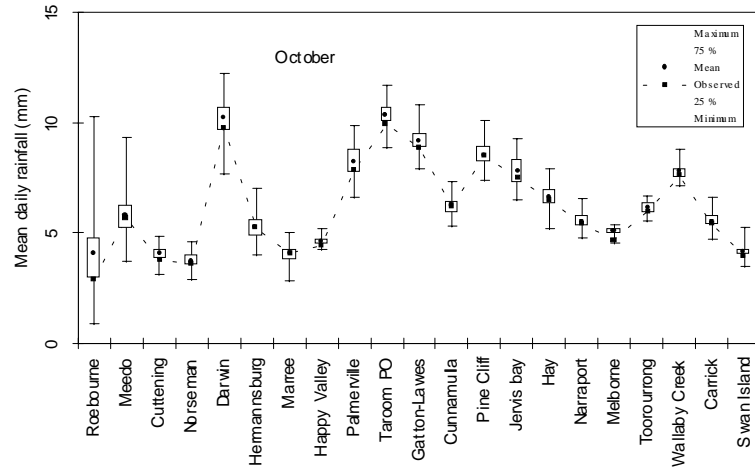
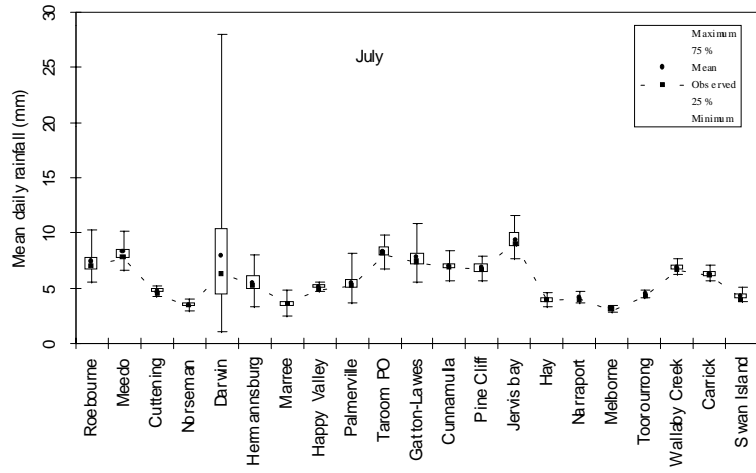
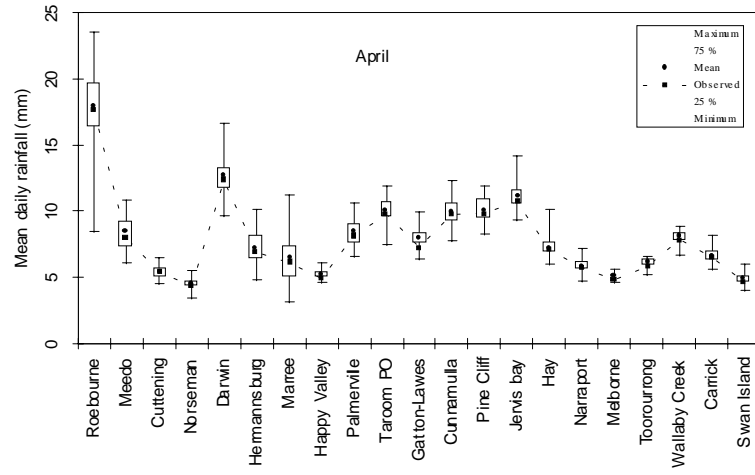
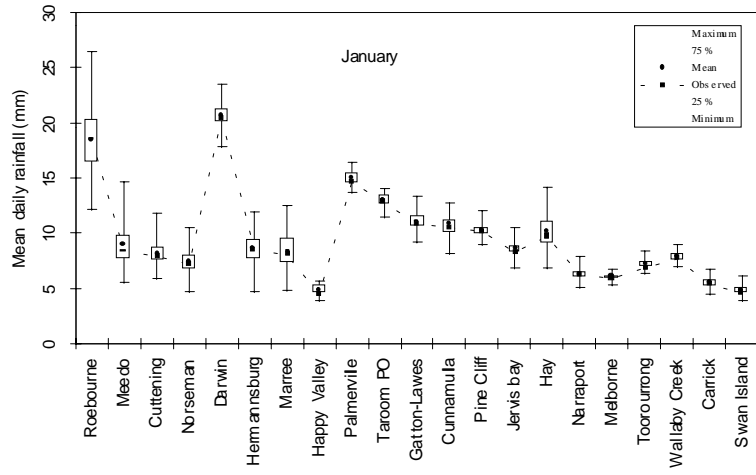


Figure A20: Comparison of observed and TPMb model generated mean daily rainfall.

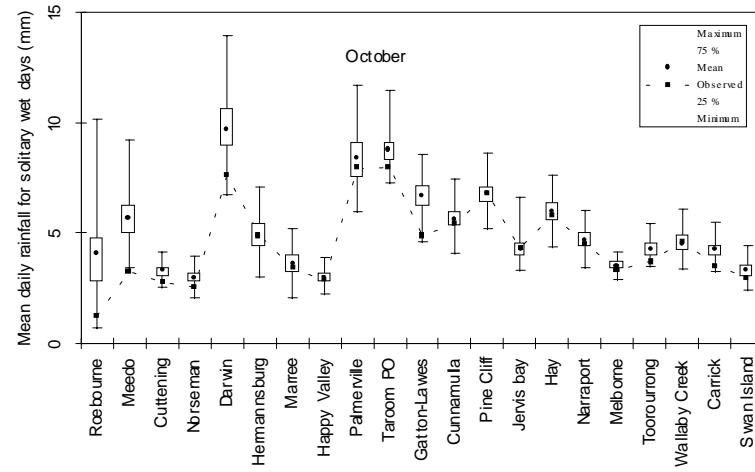
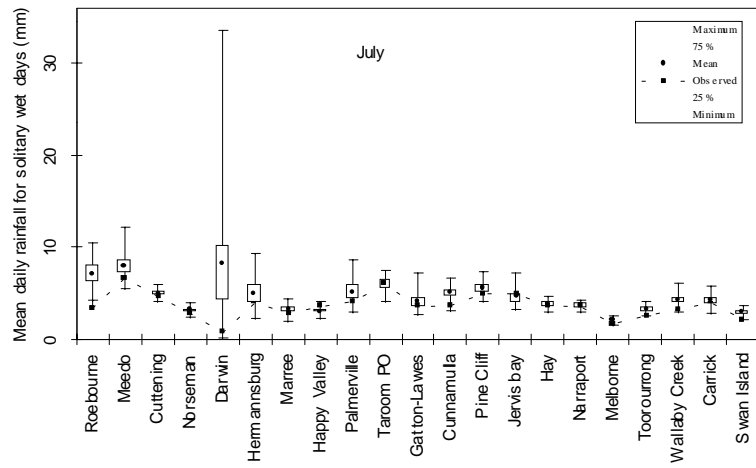
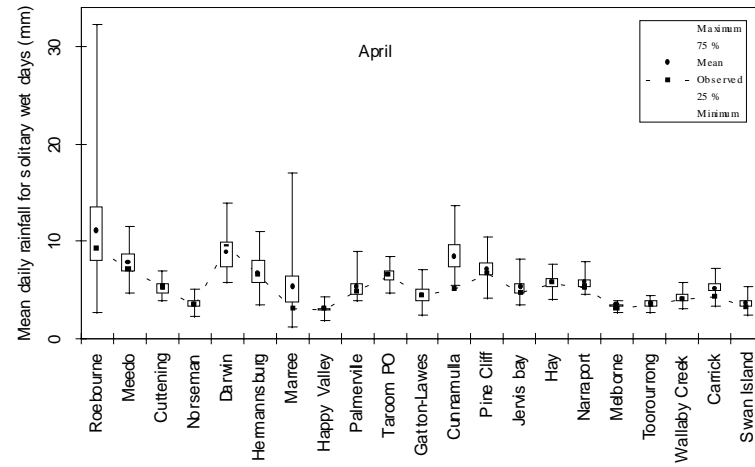
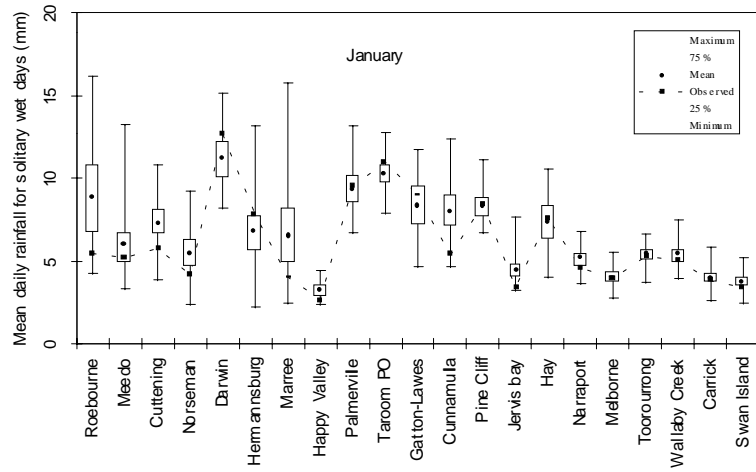


Figure A21: Comparison of observed and TPMb model generated mean daily rainfall for solitary wet days.

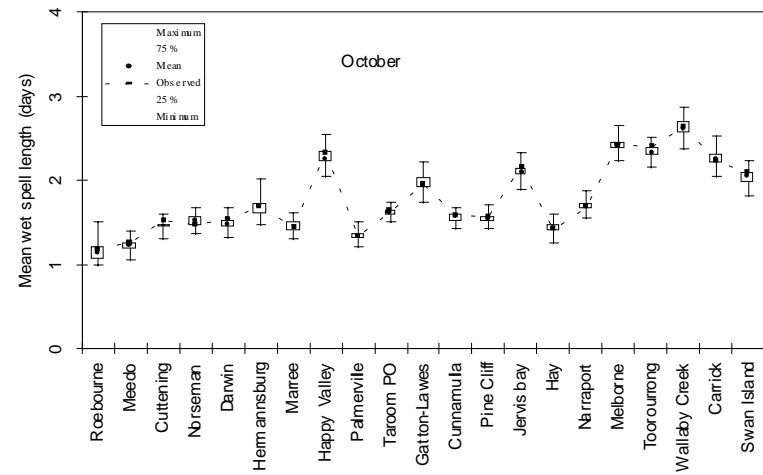
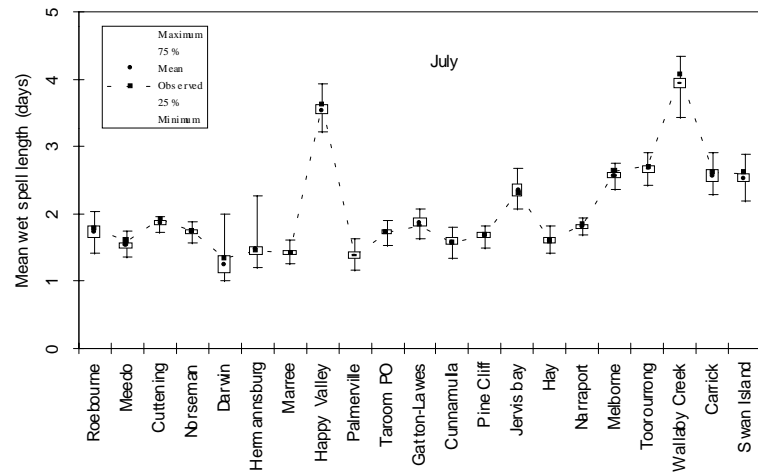
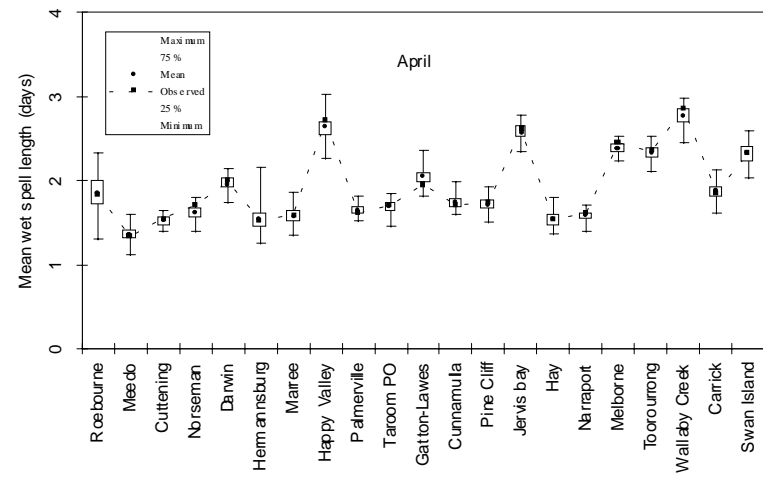
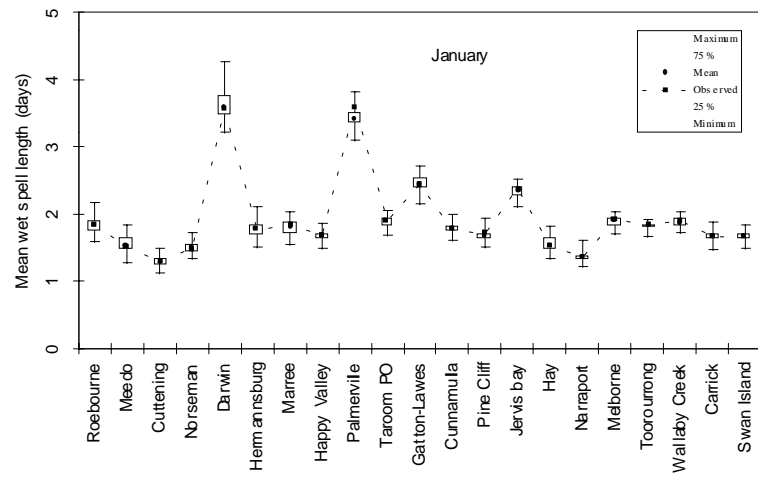


Figure A22: Comparison of observed and TPMb model generated mean wet spell length.



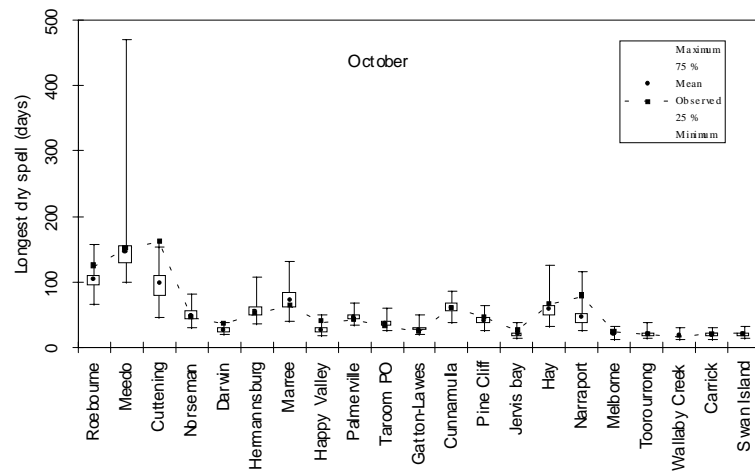
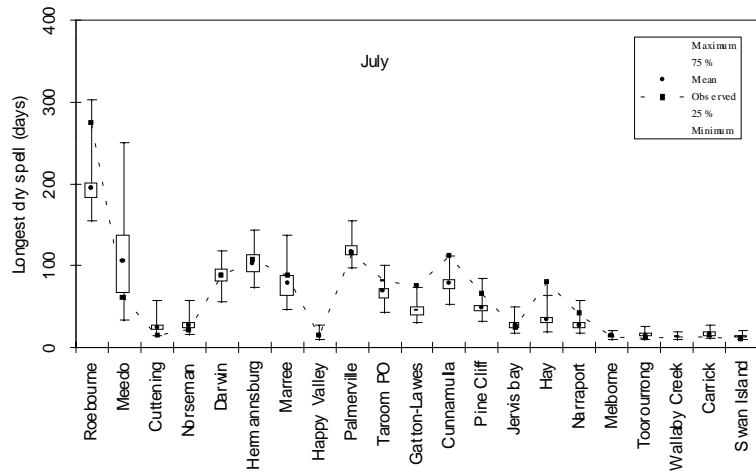
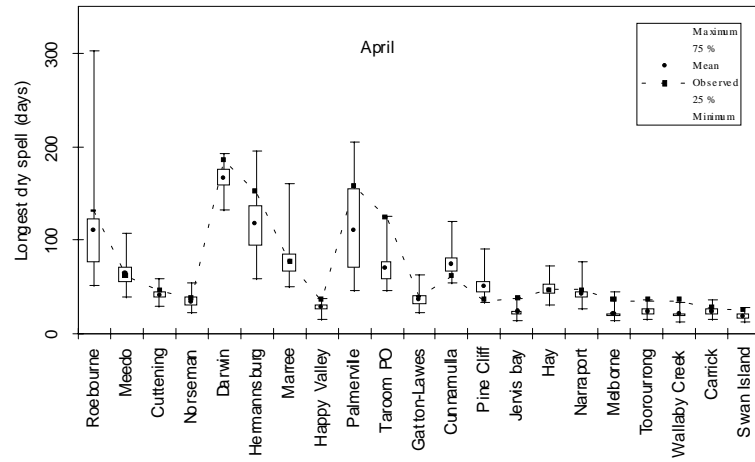
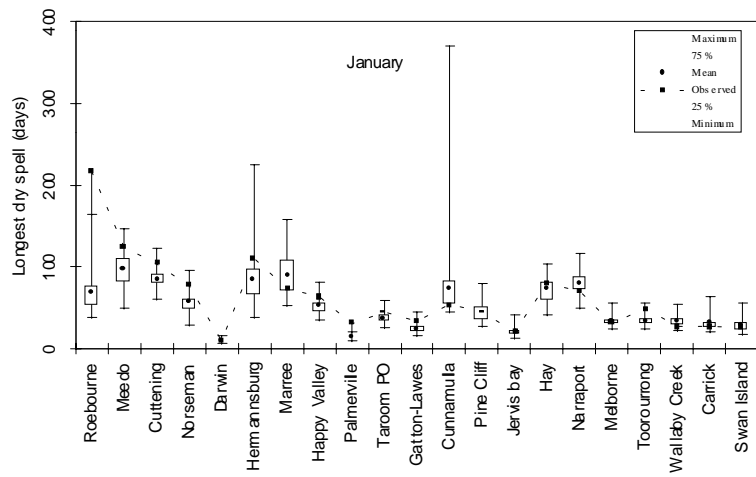


Figure A23: Comparison of observed and TPMb model generated longest dry spell.

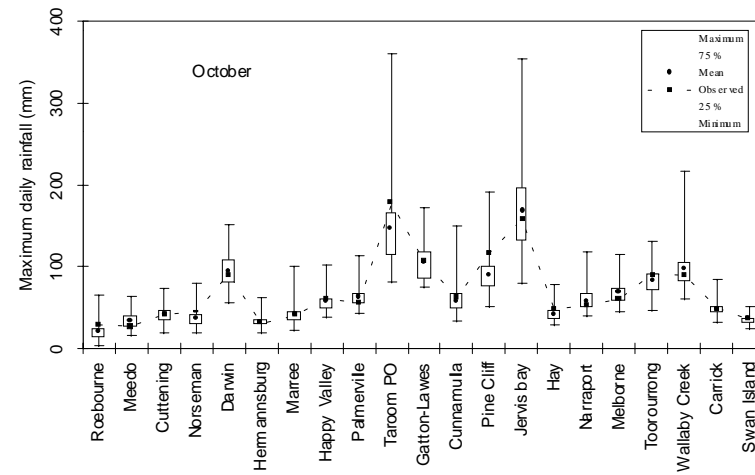
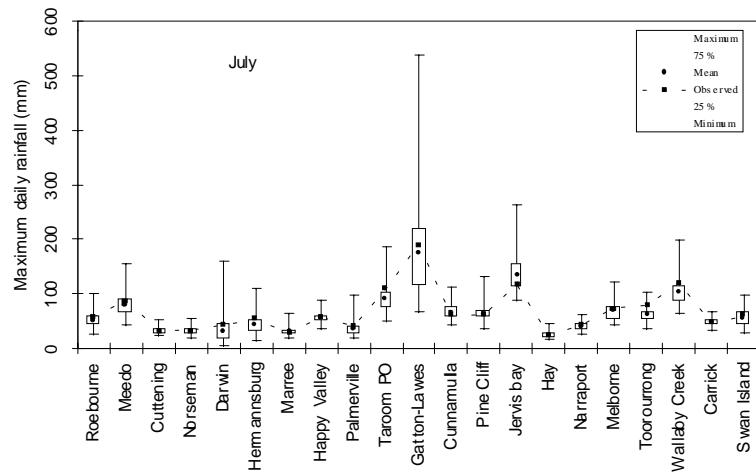
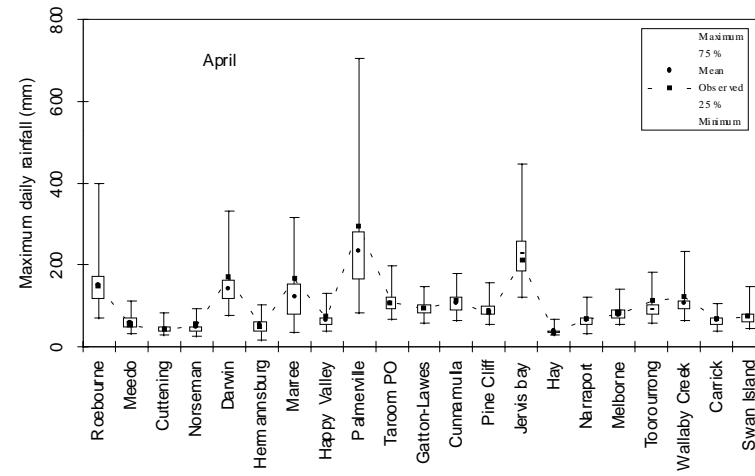
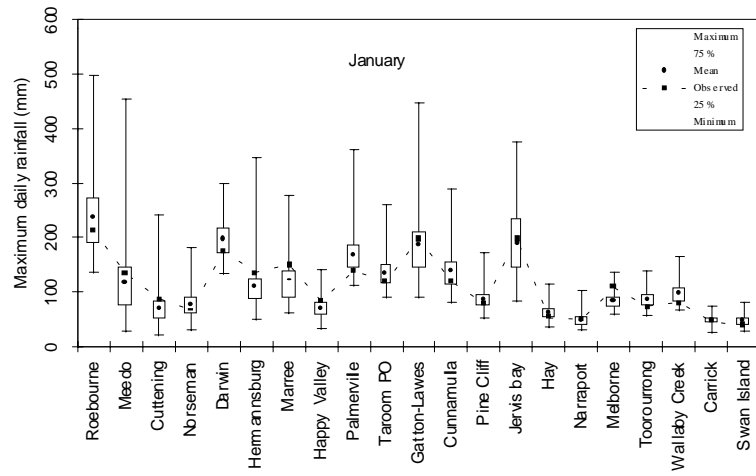


Figure A24: Comparison of observed and TPMb model generated maximum daily rainfall.