

# **Hydrological analysis relevant to surface water storage at Jabiluka**

FHS Chiew & QJ Wang

Centre for Environmental Applied Hydrology  
Department of Civil and Environmental Engineering  
University of Melbourne

**March 1999**

**Melbourne Enterprises International Limited**

**A Company of the University of Melbourne**

# Executive Summary

This consultancy is part of an investigation into hydrological issues relating to the water management system proposed for the Jabiluka mine site. The report has been prepared for the Supervising Scientist at Jabiru.

The objective of this consultancy is to estimate the storage capacity required to store surface runoff and other water within the total containment zone (TCZ) of the Jabiluka project, for various probabilities of exceedance. The consideration here is for a 241 400 m<sup>2</sup> TCZ area, of which 90 000 m<sup>2</sup> is the water storage area. This relates to the Kinhill-ERA storage capacity calculations reported in the Jabiluka Public Environment Report Appendix B1. For the calculations, we assume that the bunds and other drainage diversion structures can effectively prevent all water outside the TCZ from entering the TCZ and vice-versa.

The main inflows into the storage are surface runoff and mine dewatering inflows. The main losses from the storage are evaporation, water disposal through the underground ventilation system and water used for milling and other mining operations.

To estimate the storage capacity, 50 000 sets of 30 years of daily rainfall and monthly pan evaporation data are stochastically generated to simulate the storage water balance on a daily time step. This approach thus mimics 50 000 possibilities in the climate for a 30-year mining operation (a total of 1.5 million years), assuming stationarity in the climate.

The approach used here to estimate the storage capacity is significantly different from the approach used by Kinhill-ERA. The main limitation in the Kinhill-ERA approach is that, with their constructed extreme rainfall sequence, a probability of exceedance cannot be directly attributed to the derived storage capacity. There are also several differences in the water balance and modelling considerations in the two approaches. Unlike the Kinhill-ERA method, the approach used here simulates the storage water balance on a daily time step, considers the inter-annual variability in evaporation and the inverse relationship between evaporation and rainfall, uses more conservative pan coefficients, takes into account the lower ventilation loss in the Wet season, and allows the monthly distribution of annual rainfall to change from year to year.

Based on the 50 000 sets of daily storage water balance simulations, a storage capacity of 939 000 m<sup>3</sup> (an equivalent depth of 10.4 m in the 90 000 m<sup>2</sup> storage area) is required so that there is only a 0.01% probability of the storage being exceeded over a 30 year mine life. The estimate derived in the Kinhill-ERA simulations is 25% smaller (706 000 m<sup>3</sup>), and based on the simulations here, has a 0.6% probability of exceedance. The final Kinhill-ERA recommended storage capacity is 810 000 m<sup>3</sup> and this has a 0.08% probability of being exceeded in the 30-year mine life.

The analyses also indicate that any exceedance of the design volume will most likely occur in the first few years of the mining operation. This is because the typical evaporation and water usage from the storage in the later years are significantly greater than the typical inflows into the storage. It should be noted that the simulation results are based on the assumption that the water usage (particularly the volumes associated with the mill requirement and ventilation loss) specified in the Jabiluka PER is achievable.

# Contents

Executive Summary	ii
Contents	iii
1 Introduction	1
2 Storage water balance components	1
2.1 Runoff	2
2.2 Mine dewatering and contaminated laundry	3
2.3 Evaporation from the storage	3
2.4 Mill requirement	5
2.5 Ore wetdown and plant washdown	5
2.6 Mine ventilation and dust suppression	6
3 Rainfall, pan evaporation and data generation	6
3.1 Rainfall	6
3.2 Relationship between pan evaporation and precipitation	9
3.3 Stochastic generation of daily rainfall data	11
3.4 Stochastic generation of monthly pan evaporation data	13
4 Estimation of storage capacity	15
4.1 Storage water balance	15
5 Discussion of storage water balance simulations	16
5.1 A typical simulation run	16
5.2 The approach adopted here versus Kinhill-ERA approach	16
5.3 Sensitivity analyses	18
References	20
Appendix A The DMM algorithm for stochastic generation of rainfall series	23
Appendix B An Algorithm for stochastic generation of monthly evaporation	26

# 1 Introduction

The report is prepared for the Supervising Scientist at Jabiru. It describes part of an investigation into hydrological issues relating to the water management system proposed for the Jabiluka project.

Specifically, the objective is to estimate the water storage capacity required to store surface runoff and other water within the total containment zone (TCZ) of the Jabiluka project. The water storage volume is calculated for a range of probabilities up to 0.002% that the pond design volume would be exceeded over a 30-year mine life. In this study, 50 000 sets of 30 years of daily rainfall and monthly pan evaporation data are stochastically generated to simulate the storage water balance.

The approach used by Kinhill and Energy Resource Australia (ERA) is reviewed and the pond design compared with the estimates derived here. The Kinhill-ERA approach is described in The Jabiluka Mill Alternative Public Environment Report and The Jabiluka Mill Alternative Public Environment Report Technical Appendices (hereon referred to as Jabiluka PER Appendices) (1998). The two reports also provide background to many other issues.

The structural design of the storage and other features of the mine site are not considered here. This study also assumes that the bunds and other drainage diversion structures will prevent all water outside the TCZ from entering the TCZ and vice versa.

The storage water balance components are discussed in section 2. Some of the water inflows into the storage and losses from the storage are discussed in detail, while elsewhere, the values used by Kinhill-ERA are adopted. Section 3 describes the selection of the climate stations used here, the rainfall and pan evaporation characteristics in the area and the stochastic generation of 1.5 million years of daily rainfall and monthly pan evaporation data. Section 4 describes the approach used to estimate the storage capacity, and presents the storage capacity estimates for various probabilities of exceedance of the design volume in the 30-year mine life. Section 5 reviews the storage water balance simulations carried out here and the Kinhill-ERA simulations. The differences between the two approaches, and the effects of the various parameters on the storage water balance are discussed.

## 2 Storage water balance components

The consideration here is for a 241 400 m<sup>2</sup> TCZ area, of which 90 000 m<sup>2</sup> is the water storage area. This relates to Appendix B1 of the Jabiluka PER Appendices report.

The annual volumes of water inflows into the storage and losses from the storage are summarised in tables 2.1 and 2.2 respectively (see also tables B1.3 and B1.4 in the Jabiluka PER Appendices). The results are dependent mainly on the water balance considerations, and the values adopted in the tables are discussed below.

**Table 2.1** Inflows into the storage

Source	Area (m <sup>2</sup> )	Annual volume (m <sup>3</sup> )
Runoff		
Ore and waste stockpile	11 000	60% of rainfall
Hardstand area	33 500	95% of rainfall
Mill/mine	106 900	80% of rainfall
Pond	90 000	100% of rainfall
Mine dewatering and contaminated laundry		73 365

All the calculations throughout this report assume a water year that starts in September, to ensure that consecutive wet summer months are taken into account.

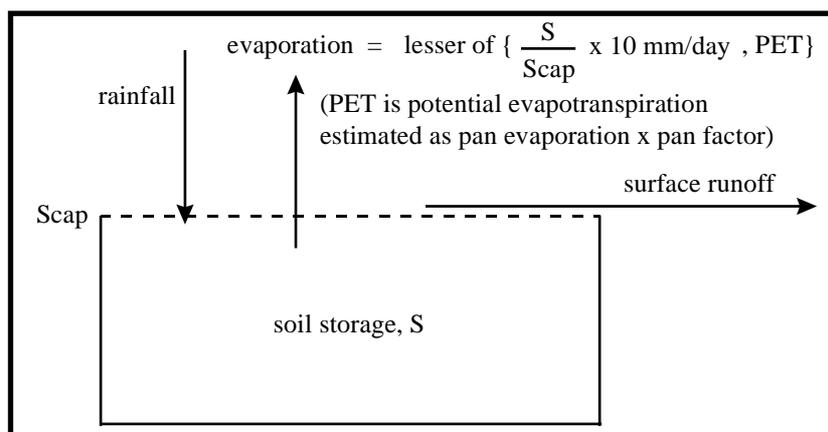
**Table 2.2** Losses from the storage (annual volumes in m<sup>3</sup>)

	Year 1	Year 2	Year 3	Year 4	Year 5	Year 6	Year 7	Year 8	Year 9	Year 10	Year 11-30
Evaporation	see Section 2.3										
Mill requirement	0	180000									
Ore wetdown and plant washdown	800	1200	3500				7000			10000	
Mine ventilation and dust suppression	0	15000	30000	45000	60000	75000	90000				

## 2.1 Runoff

There are four types of TCZ catchment land use — an area to store ore and hold the waste stockpile, a hardstand area, a mill/mine area and the pond area. The runoff coefficients used in the Jabiluka PER Appendices are very conservative and are adopted here (see table 2.1). Adopting these values, almost 90% of the rainfall falling onto the TCZ becomes runoff.

A simple conceptual water balance model (see fig 2.1) is used here to simulate the rainfall-runoff process. The soil water storage capacity (Scap in fig 2.1) is the only parameter in the model, and this is optimised such that the total runoff is the same as that estimated using the above runoff coefficients. Although the total estimated runoff is the same, the model allows for higher runoff coefficients during wet periods because the soil is closer to saturation. Rainfall data from Oenpelli and evaporation data from Jabiru (see section 3) between September 1972 and August 1998 are used to optimise the soil water capacity in the model. The resulting soil capacities for the oil/waste stockpile, hardstand and mill/mine areas are 8.7, 0.75 and 3.6 mm respectively. These small soil water capacities reflect the conservative considerations in estimating runoff.



**Figure 2.1** Daily conceptual soil water balance model used to simulate runoff

## **2.2 Mine dewatering and contaminated laundry**

The mine dewatering represents water inflows into the mine that will be collected in a sump and pumped into the storage. The value in table B1.3 of the Jabiluka PER Appendices is used (73 000 m<sup>3</sup>/year). The bottom of page B1–7 and the top of page B1–8 in the PER Appendices discuss some of the problems that can arise from using this value. The volume of contaminated laundry (365 m<sup>3</sup>/year) is negligible compared to all the other variables. The mine dewatering and contaminated laundry are simulated by adding a constant daily volume of 201 m<sup>3</sup> into the storage (73 365 m<sup>3</sup>/year).

## **2.3 Evaporation from the storage**

The approach used by Kinhill-ERA for estimating pond evaporation at Jabiluka is described in the PER Appendix B1 (page B1–8). Kinhill-ERA used the long-term (1972–1997) averages of monthly pan evaporation, together with monthly pan evaporation coefficients recommended by Hatton (1997).

### **Inter-annual variability of evaporation and inverse relationship between evaporation and rainfall**

The inter-annual variability of evaporation is not accounted for by Kinhill-ERA. The standard deviation of annual pan evaporation (based on September 1971 to August 1998 Jabiru data) is 122 mm. Using the average pan coefficient of 0.77 recommended by Hatton (1997), the standard deviation of the storage evaporation is 94 mm. This is smaller than the standard deviation of annual rainfall (293 mm, based on 1911–1998 data at Oenpelli).

In terms of the water balance, rainfall applies to the whole TCZ, while evaporation from the storage covers an area which is only 37% of the TCZ area. The inter-annual variability of evaporation will therefore have a smaller impact on the water balance than inter-annual variability of rainfall.

The correlation between evaporation and rainfall is not accounted for by Kinhill-ERA. The correlation coefficient of annual pan evaporation and annual rainfall at Jabiluka is -0.43 (based on 27 years of data — see section 3.2), which is statistically different from zero at a significance level of  $p=0.025$ . This inverse relationship between evaporation and rainfall means that the pond storage needs to be larger than if the inverse relationship is not considered.

Both the inter-annual variability and the inverse relationship between evaporation and rainfall are taken into account in the simulations carried out here.

### **Pan coefficients**

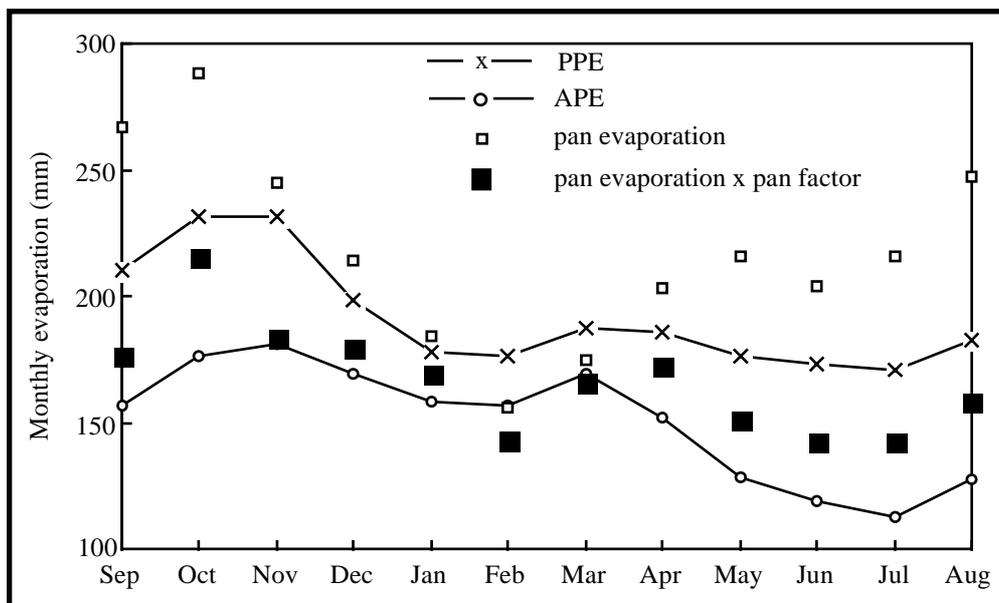
The pan coefficients recommended by Hatton (1997) for the Ranger mine were adopted by Kinhill-ERA for the storage water balance simulations reported in the Jabiluka PER Appendices. Hatton derived these coefficients after examining mainly the work by McQuade (1993). The final coefficients were set equal to those given by Vardavas (1987) except for April and October where Hatton used higher pan factors.

To examine whether the pan coefficients are appropriate, the average monthly evaporation rates (using Jabiru data from September 1971 to August 1998) derived using these coefficients (average monthly pan evaporation times the pan factors) are compared with point potential evapotranspiration (PPE) and areal potential evapotranspiration (APE) estimates. The PPE and APE estimates are extracted from the digital maps of evapotranspiration of Australia prepared by the Cooperative Research Centre for Catchment Hydrology (Wang et al

1999). In simple terms, the PPE is the rate of evapotranspiration from a small wet area in an existing environment, while APE is the rate of evapotranspiration if a large area is well watered. The Jabiluka storage evaporation rate is expected to be somewhere between the PPE and APE, and closer to APE during the wet season than during the dry season due to less advective energy in the wet season.

The comparison in figure 2.2 indicates that the estimated monthly storage evaporation rates (using Hatton's pan factors) are generally reasonable. For February and March, the estimated storage evaporation rates are lower than the APE. This could be due to errors in the recorded pan evaporation (because of the difficulty in accurately estimating pan evaporation when rainfall is high) or errors in the APE estimates.

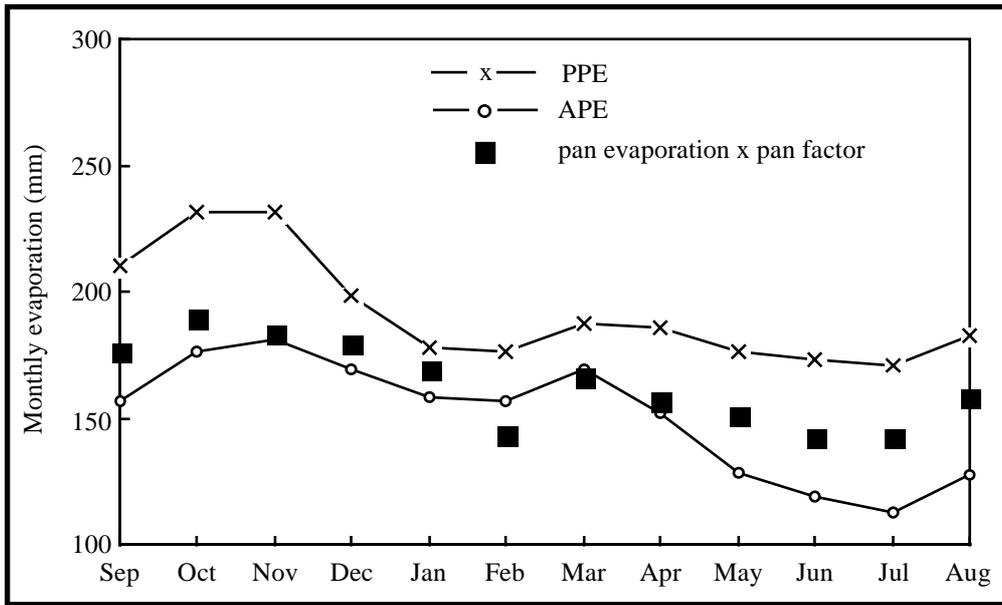
The plots in figure 2.2 also suggest that the pan coefficients for April and October may be high. These were the two months where Hatton used higher pan coefficients compared to the coefficients given by Vardavas. Figure 2.3 shows the evaporation rates calculated using Vardavas' original pan coefficients, and they appear to be in more realistic agreement with the seasonal pattern of the APE and PPE, compared to the rates calculated using Hatton's factors for October and April.



**Figure 2.2** Comparison of average monthly storage evaporation estimates, pan evaporation rates and point potential and areal potential evapotranspiration rates (storage evaporation estimated using Hatton's pan coefficients)

The pan coefficients of Vardavas are used here (see fig 2.3 and table 2.3). This results in on average, about 16 mm (1450 m<sup>3</sup>) and 26 mm (2350 m<sup>3</sup>) less evaporation from the storage in October and April respectively.

The use of climate data to compute evapotranspiration using Penman's combination equation is not investigated here. In any case, both the APE and PPE maps are derived from evapotranspiration rates estimated for a large number of climate stations throughout Australia (Wang et al 1999). The APE and PPE estimates are derived from climate data (solar radiation, temperature and vapour pressure deficit) using Morton's (1983) energy method.



**Figure 2.3** Comparison of average monthly storage evaporation estimates and point potential and areal potential evapotranspiration rates (storage evaporation estimated using Vardavas' pan coefficients)

**Table 2.3** Pan coefficients used here to estimate evaporation from the storage

Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug
0.66	0.66	0.75	0.84	0.92	0.92	0.95	0.77	0.70	0.70	0.66	0.64

**Effect of salinity buildup on storage evaporation**

The water salinity in the storage will have a negligible effect on the evaporation rates. There will not be any substantial buildup of salinity in the storage because of the continual inflow of surface runoff replacing the water drawn from the storage for milling and other mining operations. In any case, the salinity would have to approach 20% by weight before significant reduction in evaporation can be expected (Hatton 1997).

**2.4 Mill requirement**

From the second year of the mining operation onwards, an annual volume of 180 000 m<sup>3</sup> (493 m<sup>3</sup>/day) will be consumed by the mill. This value is given in table B1.3 of the PER Appendices and is used here.

**2.5 Ore wetdown and plant washdown**

The value adopted by Kinhill-ERA is also used here (see table 2.2). This varies from a volume of 800 m<sup>3</sup> in the first year of the mining operation to 10 000 m<sup>3</sup> from year 12 onwards. This volume is relatively small compared to the other variables.

## 2.6 Mine ventilation and dust suppression

As the mine develops, a significant loss of water will occur in the underground ventilation system. Air passing through the underground workings and stopes will evaporate free water which will be subsequently discharged via the ventilation exhaust system. The annual ventilation loss volumes used by Kinhill-ERA are adopted here (as agreed in the consultancy brief). The volumes vary from 15 000 m<sup>3</sup> in the second year of mining operation to 90 000 m<sup>3</sup> from year 7 onwards (see table 2.2). The ventilation loss is a significant component of the water balance, and the values used by Kinhill-ERA have been questioned by Wasson et al (1998). This aspect is not investigated here.

In the Kinhill-ERA storage water balance simulations, a constant ventilation loss is assumed through the year. However, the evaporation potential through the ventilation system is greater in the dry season than the wet season because of the greater moisture deficit in the dry season. This is taken into account here by attributing 16% of the total water disposal from the ventilation system to the four wettest months (December to March) and the remaining 84% to the other months (see table 2.4), as suggested in the Jabiluka PER Appendices (middle of page B1–9).

**Table 2.4** Distribution of annual water disposal through the ventilation system through the year

Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug
11%	11%	9%	4%	4%	4%	4%	9%	11%	11%	11%	11%

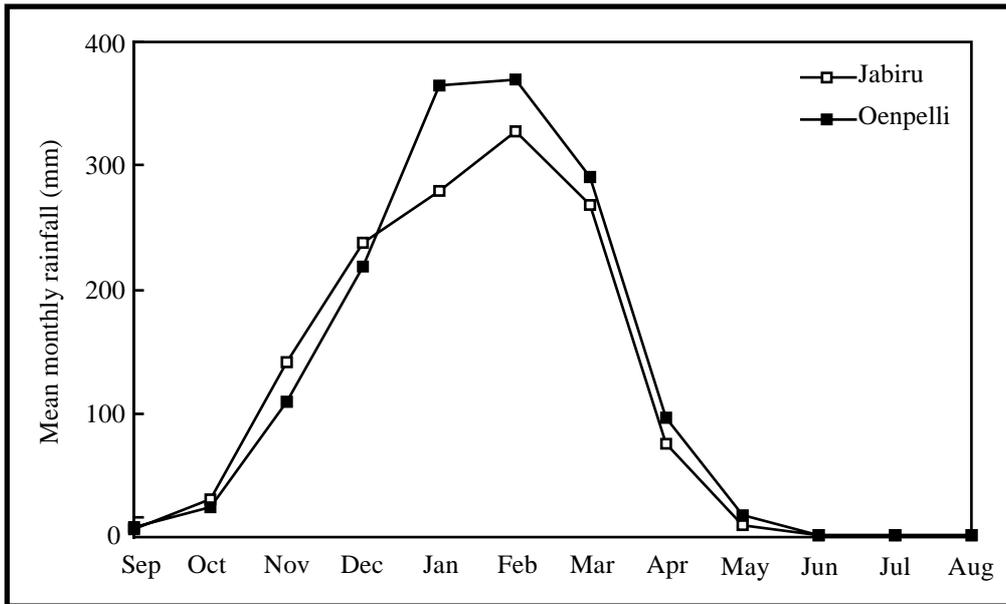
## 3 Rainfall, pan evaporation and data generation

### 3.1 Rainfall

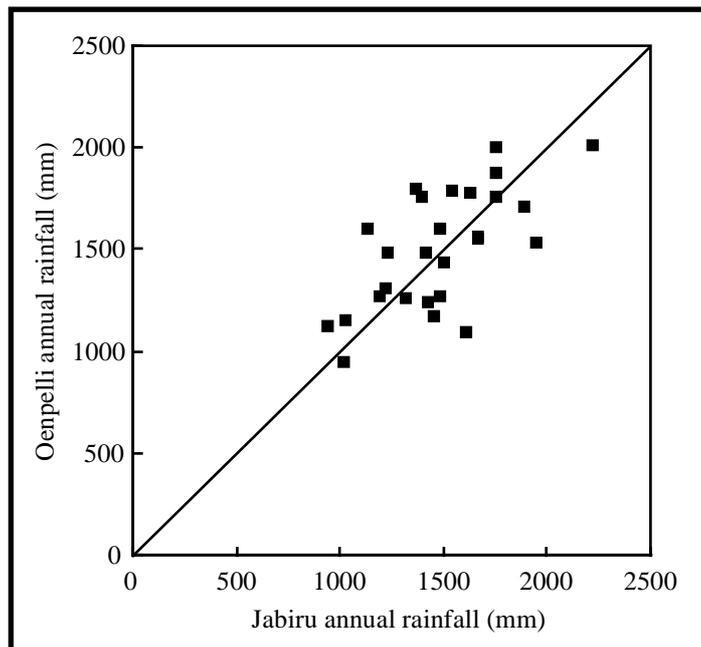
There are two rainfall stations close to the Jabiluka site with good records, at Jabiru (15 km south of the Jabiluka site) and at Oenpelli (25 km north-east). The data at Jabiru extends from 1971 to 1998 and the data at Oenpelli covers the period from 1911 to 1998. The complete set of daily records at Jabiru were obtained from ERA. The Oenpelli daily rainfall data were obtained from the National Climate Centre (Bureau of Meteorology). There were very few missing records, and these were infilled by the Hydrology Section of the Bureau of Meteorology (Bruce Stewart, pers comm).

The Jabiru and Oenpelli rainfall are compared below using concurrent data between September 1971 and August 1998. Figure 3.1 shows that the rainfall distribution through the year is similar at the two locations, although the rainfall is slightly higher in Oenpelli in the wetter months. The plots in figures 3.2 and 3.3 also indicate that the annual and monthly rainfall at the two locations are relatively similar. The correlation of annual rainfall (Sep–Aug) at the two locations is 0.67, and the correlations of monthly rainfall are 0.92, 0.79, 0.92 and 0.65 in spring (Sep–Nov), summer (Dec–Feb), autumn (Mar–May) and winter (Jun–Aug) respectively. The duration plot in figure 3.4 also indicates that the daily rainfall characteristics at the two locations are similar.

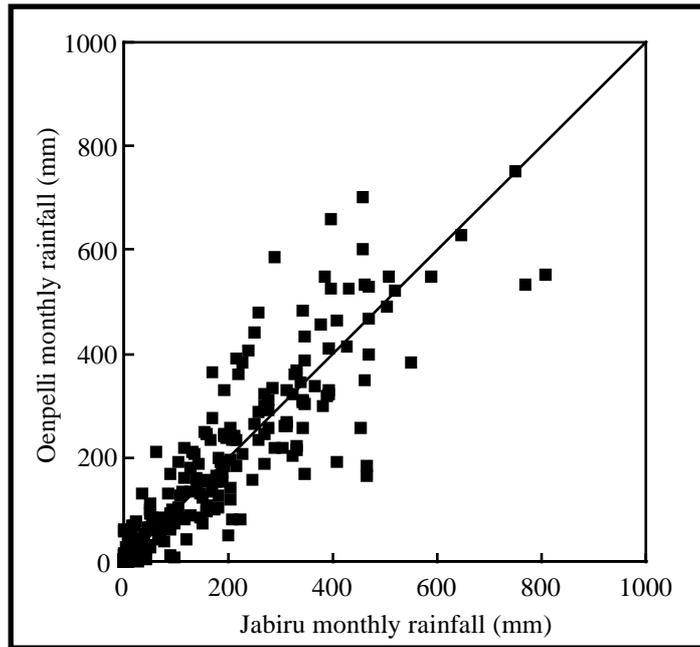
The mean annual rainfall over this period (1971–1998) is 1500 mm at Oenpelli and 1480 mm at Jabiru. This average rainfall over this period is higher than the 1911 to 1998 average (1400 mm). As the rainfall data at the two locations are relatively the same, the rainfall data at Oenpelli between 1911–1998 will be used in the analyses, because it has a much longer record.



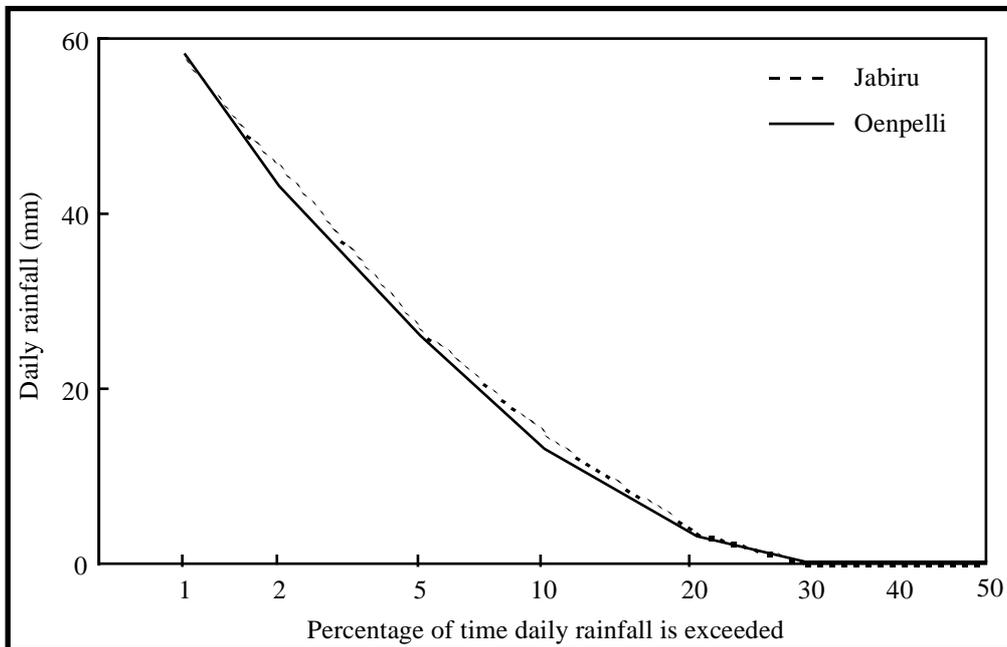
**Figure 3.1** Mean monthly precipitation at Jabiru and Oenpelli (data from September 1971 to August 1998)



**Figure 3.2** Comparison of annual rainfall at Jabiru and Oenpelli (data from September 1971 to August 1998)



**Figure 3.3** Comparison of monthly rainfall at Jabiru and Oenpelli (data from September 1971 to August 1998)



**Figure 3.4** Comparison of daily rainfall characteristics at Jabiru and Oenpelli (data from September 1971 to August 1998)

### 3.2 Relationship between pan evaporation and precipitation

There is a fairly strong inverse relationship between the pan evaporation and rainfall data. The plots in figure 3.5 show the relationship between annual pan evaporation and rainfall and the relationships between monthly pan evaporation and rainfall over the four seasons. Pan evaporation data at Jabiru and rainfall data at Oenpelli between September 1971 and August 1998 are used here. The evaporation data at Jabiru were obtained from ERA. There were very few missing records (2%) and these were infilled by relating the values at Jabiru to those at Darwin.

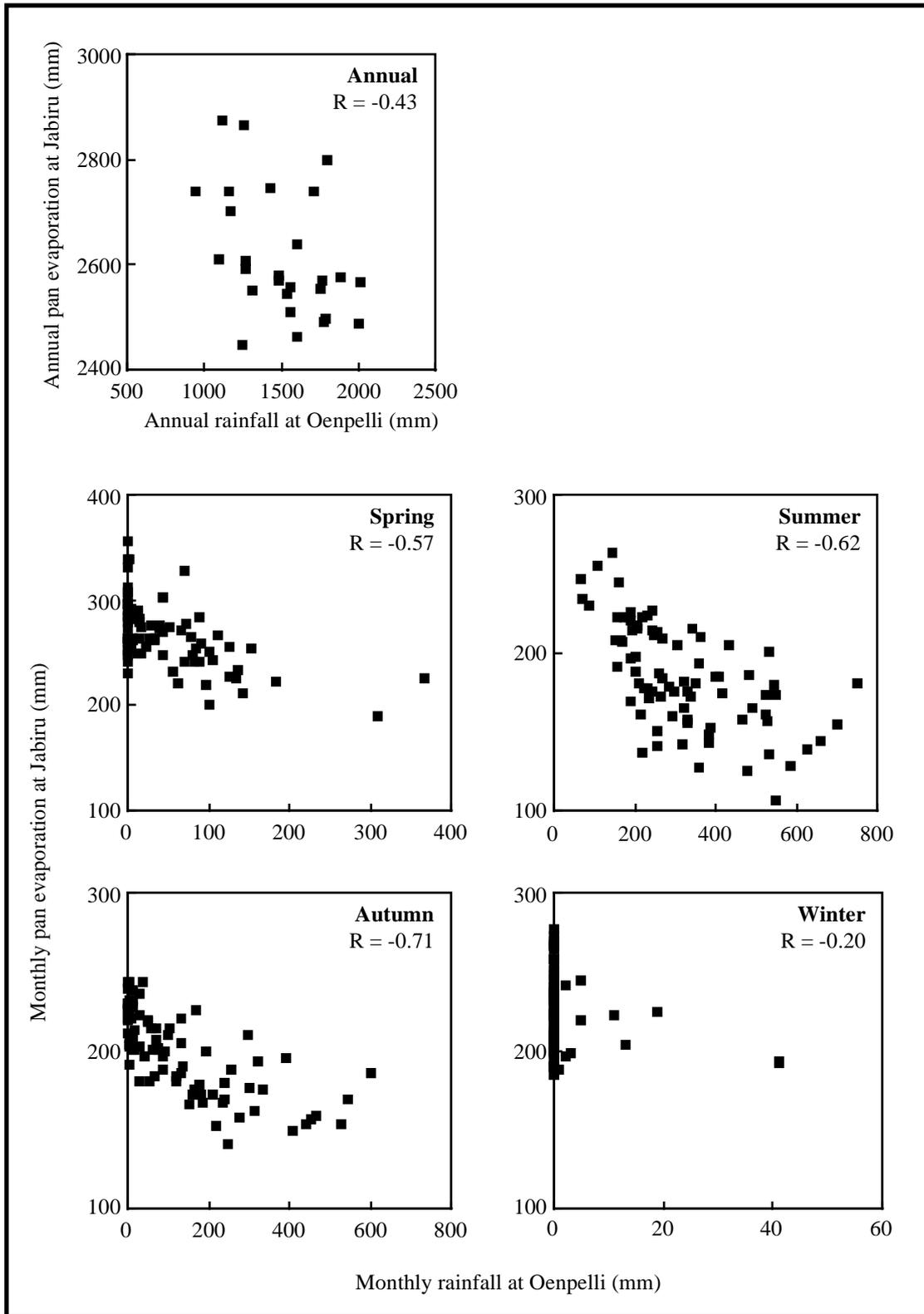
The pan evaporation-rainfall correlations for Oenpelli rainfall versus Jabiru evaporation and for Jabiru rainfall versus Jabiru evaporation are similar (see table 3.1). This is expected because the rainfall characteristics at Jabiru and Oenpelli are similar.

**Table 3.1** Correlation between annual pan evaporation and rainfall and between monthly pan evaporation and rainfall for the four seasons at various locations

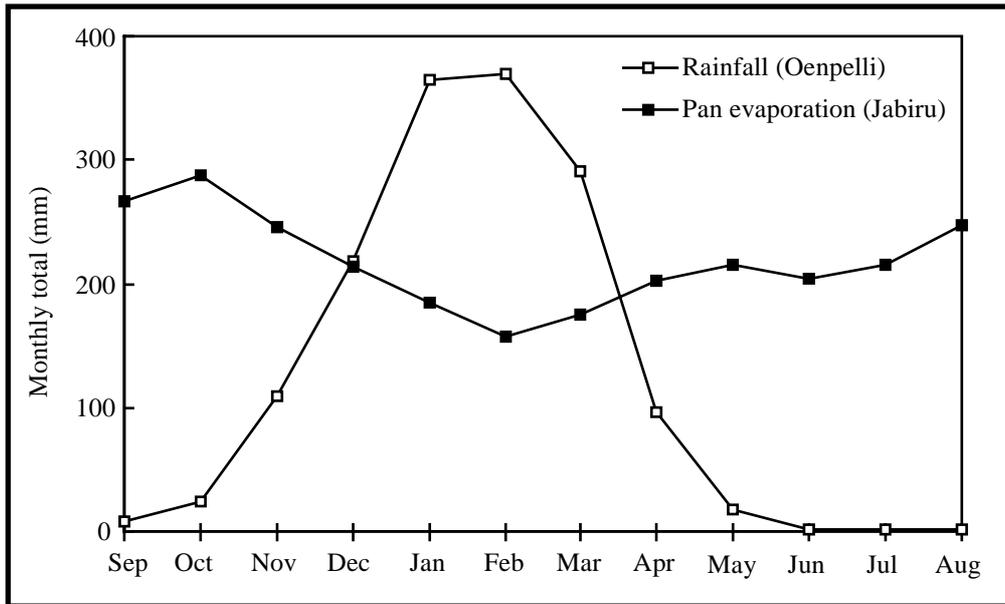
	Jabiru ET versus Oenpelli rain (1972–1998)	Jabiru ET Versus Jabiru rain (1972–1998)	Maningrida ET versus Maningrida rain (1967–1998)	Middle Point ET versus Middle Point rain (1965–1997)	Darwin ET versus Darwin rain (1958–1997)
Annual	- 0.43	- 0.43	- 0.19	- 0.33	+ 0.33
Spring	- 0.57	- 0.55	- 0.05	- 0.56	- 0.32
Summer	- 0.62	- 0.47	- 0.28	- 0.48	- 0.34
Autumn	- 0.71	- 0.71	- 0.34	- 0.48	- 0.54
Winter	- 0.20	- 0.24	- 0.16	- 0.12	+ 0.08

Table 3.1 also shows the correlations between annual and monthly pan evaporation and rainfall at the next three closest locations to the Jabiluka site where there are both pan evaporation and rainfall data. All the data were obtained from the National Climate Centre. All three stations are located close to the coast. Maningrida is about 150 km north-east, Middle Point is about 150 km west and Darwin is about 250 km west of the Jabiluka site. Table 3.1 indicates that the inverse relationship between evaporation and rainfall is strongest at Jabiru.

In summary, the daily rainfall data at Oenpelli (September 1911 to August 1998) and monthly pan evaporation data at Jabiru (September 1971–August 1998) are used here. The inverse relationship between pan evaporation and rainfall is statistically significant (at  $p=0.025$  for the annual values) and this is taken into account in generating the data. Figure 3.6 shows the distribution of monthly rainfall at Oenpelli and pan evaporation at Jabiru through the year.



**Figure 3.5** Inverse relationship between annual and monthly pan evaporation (at Jabiru) and rainfall (at Oenpelli) (data from September 1971 to August 1998)



**Figure 3.6** Mean monthly precipitation at Oenpelli and pan evaporation at Jabiru (data from September 1971 to August 1998)

### 3.3 Stochastic generation of daily rainfall data

The DMM (daily-monthly-mixed) algorithm (Wang & Nathan 1999) described in Appendix A is used to generate daily rainfall data. The advantages of this algorithm are that it has a small number of parameters (six for each month) and is capable of reproducing key characteristic statistics simultaneously at the daily, monthly and annual time periods. The six parameters for each month are estimated from

- the transitional probability of a wet day following a wet day
- the transitional probability of a wet day following a dry day
- the mean rainfall depth in a wet day
- the standard deviation of rainfall depth in a wet day
- the standard deviation of monthly rainfall
- the lag-1 autocorrelation of monthly rainfall.

The Oenpelli daily rainfall record (1911–1998) is used here to obtain these parameters. To evaluate the DMM algorithm for this site, statistics from 1000 years of generated daily rainfall data are compared with the statistics for the observed data. Table 3.2 summarises some of the daily statistics. The table shows that the generated data closely reproduce the observed statistics, including the skewness which is not used in the model fitting.

The monthly statistics are given in table 3.3. The monthly means are closely reproduced except for April, where there is a small difference. The coefficients of variation (standard deviation divided by the mean) are closely reproduced for the wet months. The difference between the observed and simulated coefficients of variation for the dry season will have little effect on the storage water balance simulations because of the small rainfall. In fact, the difference is caused by the small rainfall, because the coefficient of variation is highly sensitive to small deviations from the mean when the mean is small. The skewness, which is

not used in model fitting, is reasonably reproduced in the generated data for the wet months. The difference in skewness for the dry months is again of little consequence because of the small rainfall.

The annual statistics are given in table 3.4. The mean and the coefficient of variation are almost exactly reproduced. A small positive skewness is produced in the generated data while the observed skewness is almost zero.

**Table 3.2** Comparison of key daily rainfall statistics in the generated and observed data

	Mean (mm)		CV		Skewness	
	Observed	Simulated	Observed	Simulated	Observed	Simulated
All data	3.82	3.81	3.02	3.05	5.52	5.52
Jan	10.90	10.95	1.66	1.69	3.28	3.28
Feb	11.33	11.26	1.58	1.60	3.15	3.14
Mar	9.00	9.01	1.90	1.92	3.58	3.58
Apr	2.60	2.63	3.79	3.89	8.02	8.00
May	0.46	0.44	8.08	7.13	12.16	12.15
Jun	0.05	0.04	16.99	16.94	31.60	31.64
Jul	0.08	0.06	20.55	16.15	25.76	25.79
Aug	0.03	0.01	20.79	23.01	40.85	40.10
Sep	0.16	0.11	11.12	12.00	21.52	21.48
Oct	0.86	0.84	5.92	5.37	12.18	12.13
Nov	3.60	3.59	2.68	2.70	4.96	4.96
Dec	7.19	7.21	2.02	2.02	3.65	3.66

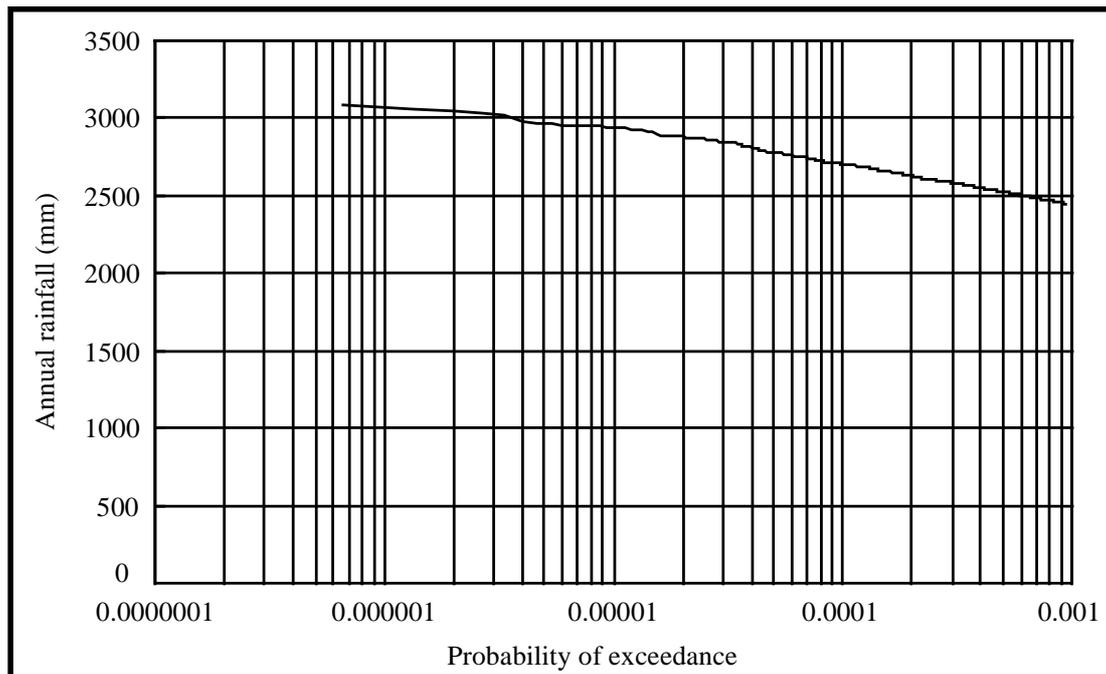
**Table 3.3** Comparison of key monthly rainfall statistics in the generated and observed data

	Mean (mm)		CV		Skewness	
	Observed	Simulated	Observed	Simulated	Observed	Simulated
All data	116.4	116.1	1.31	1.31	1.39	1.36
Jan	337.9	335.2	0.39	0.38	1.11	0.70
Feb	320.2	317.2	0.42	0.42	0.97	0.68
Mar	278.9	281.8	0.50	0.50	0.47	0.83
Apr	77.9	87.1	1.12	1.00	2.03	1.79
May	14.3	15.9	2.04	1.99	3.66	3.14
Jun	1.6	1.7	3.84	4.31	5.15	8.08
Jul	2.4	1.7	3.91	4.65	4.78	6.85
Aug	0.8	0.4	3.69	5.59	4.40	9.92
Sep	4.7	3.2	2.28	3.47	3.66	6.27
Oct	26.7	25.4	1.31	1.33	1.63	2.81
Nov	108.0	104.7	0.65	0.64	1.26	0.98
Dec	222.8	218.6	0.45	0.44	1.16	0.74

**Table 3.4** Comparison of key annual rainfall statistics in the generated and observed data

Mean (mm)		CV		Skewness	
Observed	Simulated	Observed	Simulated	Observed	Simulated
1394	1393	0.21	0.21	0.00	0.37

The DMM algorithm is used to generate 1.5 million years of daily rainfall data for the storage water balance simulations. Figure 3.7 shows the frequency curve of the annual rainfall for the 1.5 million years of generated data. The 1 in 10 000 year annual rainfall is 2702 mm. The estimate of 1 in 10 000 year annual rainfall by the Bureau of Meteorology is 2460 mm with a standard deviation of 85 mm. Although our estimate of 2702 mm is almost three standard deviations higher than the estimate of 2460 mm by the Bureau of Meteorology, the standard deviation given by the Bureau accounts for only sampling error but not for possible model error in extrapolating from 88 years of recorded data to the 10 000 year return period. Our model is implicit in the DMM algorithm, while the Bureau assumed a normal distribution for the annual rainfall, resulting in the difference in the estimates. Given the large uncertainty in modelling the extrapolation to a 10 000 year return period, it is not possible to judge which of the two estimates is more appropriate.



**Figure 3.7** Frequency curve of annual rainfall based on 1.5 million years of generated data

### 3.4 Stochastic generation of monthly pan evaporation data

A simple algorithm is used here to generate monthly pan evaporation data. The daily variation in evaporation within a given month is not considered necessary in the water balance simulation. The monthly pan evaporation data generation algorithm is described in Appendix B. The following monthly statistics are used in the model fitting

- mean of monthly evaporation
- standard deviation of monthly evaporation
- lag-1 autocorrelation of monthly evaporation
- lag-2 autocorrelation of monthly evaporation
- cross correlation between monthly evaporation and monthly rainfall

The Jabiru pan evaporation data (September 1971 to August 1998) are used to obtain the model parameters. To evaluate the algorithm for this site, statistics from 1000 years of generated monthly pan evaporation are compared with the statistics from the observed data. Tables 3.5 and 3.6 summarise the monthly and annual statistics respectively. The generated data closely reproduce the observed mean and coefficients of variation, for all the monthly and annual values. The skewness is not well reproduced because of the large uncertainties in the skewness estimated from only 27 years of observed data. This is reflected in the irregular fluctuation of skewness from month to month. In any case, the skewness in the data is not important because of the relatively small coefficients of variation.

The cross correlation between the generated annual pan evaporation and rainfall is -0.41. This is in close agreement with the observed cross correlation of -0.43.

**Table 3.5** Comparison of key monthly pan evaporation statistics in the generated and observed data

	Mean (mm)		CV		Skewness	
	Observed	Simulated	Observed	Simulated	Observed	Simulated
All data	218	219	0.20	0.19	0.29	0.27
Jan	184	186	0.13	0.12	-0.11	0.01
Feb	156	151	0.15	0.13	0.10	-0.13
Mar	175	175	0.11	0.12	0.61	-0.06
Apr	203	204	0.11	0.11	0.11	-0.20
May	216	216	0.07	0.07	-0.22	-0.41
Jun	204	205	0.07	0.07	0.65	-0.06
Jul	216	216	0.08	0.08	0.72	-0.09
Aug	247	247	0.07	0.07	0.19	-0.03
Sep	268	268	0.08	0.08	1.02	-0.12
Oct	288	287	0.10	0.10	0.59	-0.44
Nov	244	246	0.11	0.11	-0.05	-0.07
Dec	212	214	0.12	0.12	-0.18	-0.11

**Table 3.6** Comparison of key annual pan evaporation statistics in the generated and observed data

Mean (mm)		CV		Skewness	
Observed	Simulated	Observed	Simulated	Observed	Simulated
2615	2617	0.05	0.04	0.77	-0.13

## 4 Estimation of storage capacity

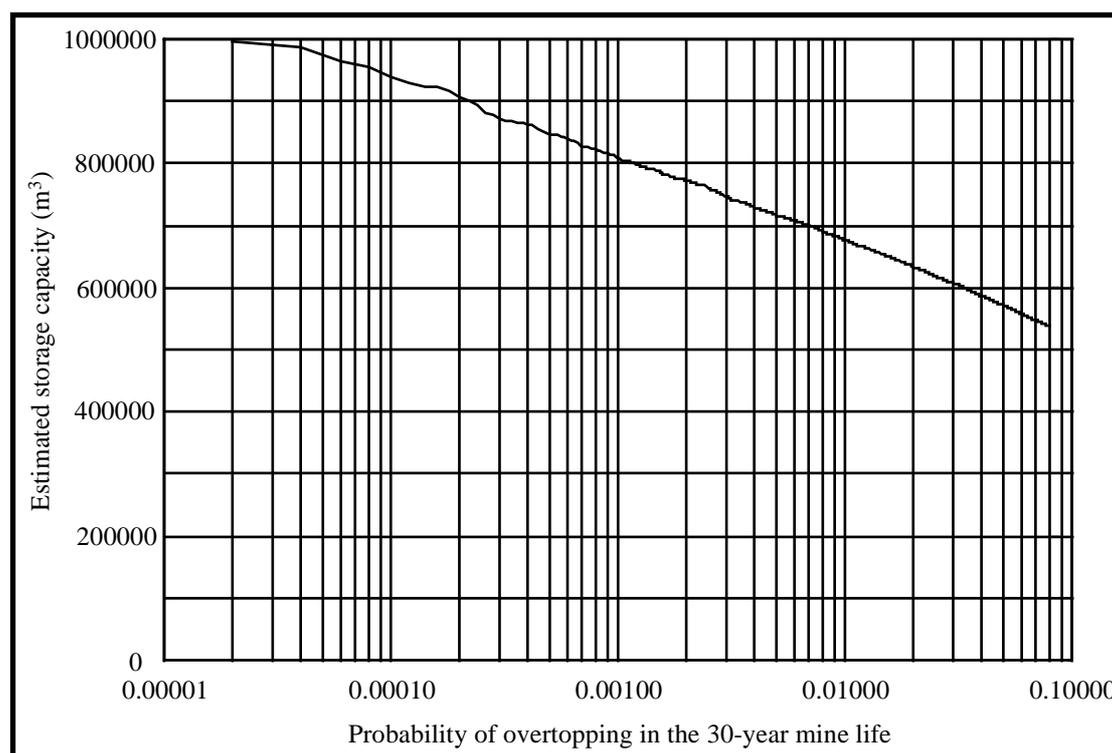
### 4.1 Storage water balance

The equation below describes a daily simulation of the storage water balance

$$S_{t+1} = S_t + \text{Inflows} - \text{Losses}$$

where  $S_t$  is the present storage and  $S_{t+1}$  is storage on the following day. The inflows into (runoff and mine dewatering) and losses from the storage (evaporation, mill requirement, ore wetdown and ventilation loss) are described in section 2. All the losses in table 2.2 are subtracted from the storage as long as there is water in the storage.

Fifty thousand runs are carried out here, with each run simulating the daily storage water balance over a 30-year mine life, starting with an empty storage. The largest storage level in each run gives an estimate of the storage capacity required such that the storage volume will not be exceeded in that run. The largest of these 50 000 values is therefore the estimate of the storage capacity with a 0.00002 (1/50000) probability of being exceeded during the 30-year mine life, the tenth largest of these values is the estimate of the storage capacity with a 0.0002 (or 0.02%) probability of being exceeded in the 30-year mine life, and so on.



**Figure 4.1** Estimated storage capacities for various probabilities of exceedance

Figure 4.1 shows the estimates of storage capacity for various probabilities of exceedance of the design volume. Based on these simulations, the estimate of storage capacity with a 0.01% probability (1 in 10 000) of being exceeded in the 30-year mine life is 939 000 m<sup>3</sup> (an equivalent depth of 10.4 m in the 90 000 m<sup>2</sup> storage area). Figure 4.1 also indicates that the maximum pond storage volume of 706 000 m<sup>3</sup> estimated in the Kinhill-ERA simulations

(page B1–10 and table B8 of the Jabiluka PER Appendices) has a 0.6% probability of being exceeded. The final Kinhill-ERA recommended storage of 810 000 m<sup>3</sup> (based on a 9.0 m depth for the 9 ha pond, see page 4–52 of the PER report) has a 0.08% probability of being exceeded in the 30-year mine life.

## **5 Discussion of storage water balance simulations**

In this section we review the storage water balance simulations carried out here and the Kinhill-ERA simulations described in the Jabiluka PER Appendices. The differences between the two approaches, and the effects of the various parameters on the storage water balance are discussed.

### **5.1 A typical simulation run**

The plots in figure 5.1 show results from a simulation run using the approach described in section 4 and the water balance considerations described in section 2. The simulation is carried out for 26 years using *observed* daily rainfall data at Oenpelli and *observed* pan evaporation data at Jabiru between September 1972 and August 1998. The plots show the annual rainfall and the storage volume at the end of each month of the simulation.

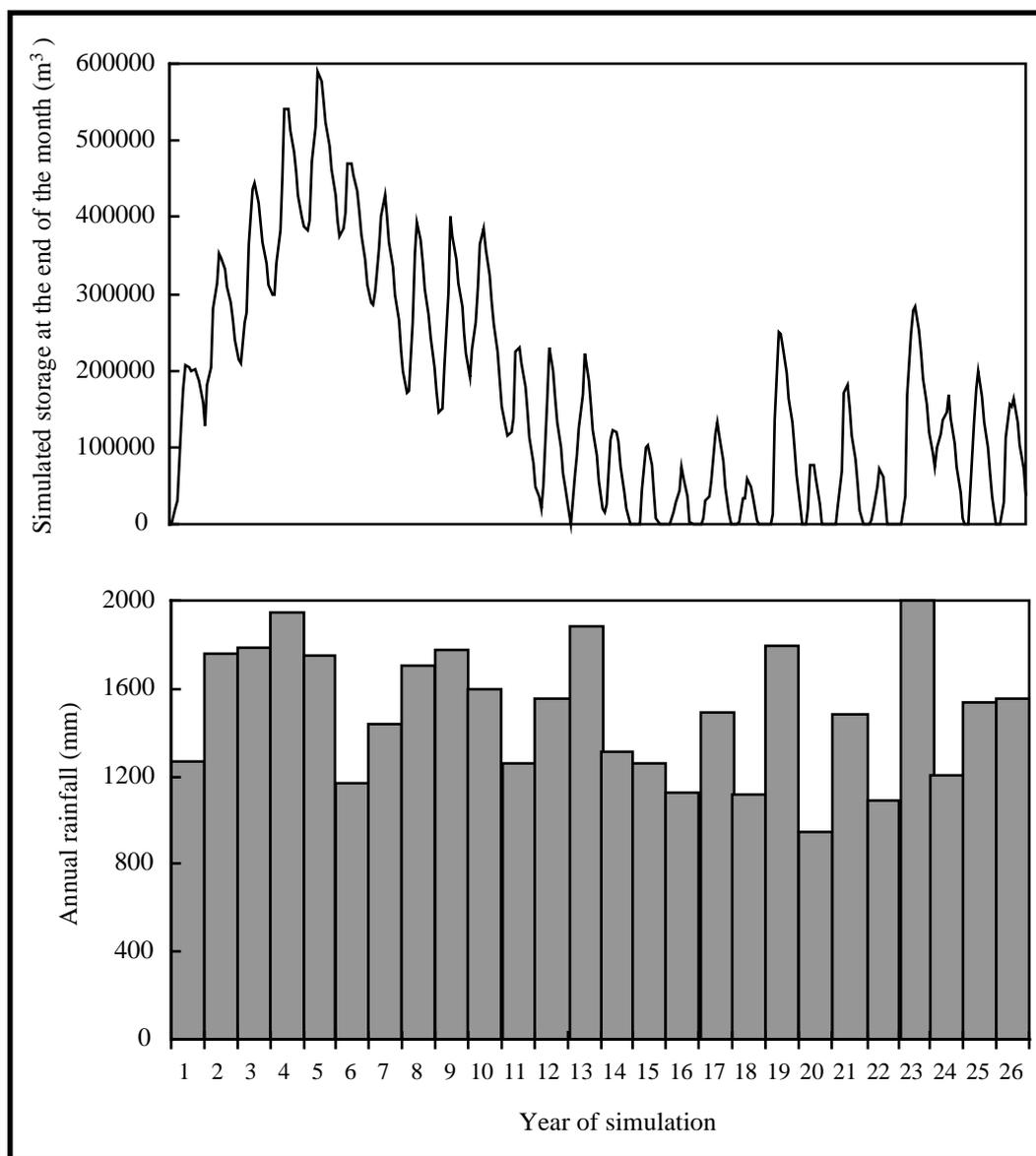
The plot shows that the biggest storage in the simulation is 594 000 m<sup>3</sup>, and to prevent the storage design volume from being exceeded in this 26-year run, a storage capacity of 594 000 m<sup>3</sup> is therefore required. The plot also shows the seasonal fluctuation in the storage volume, with the volumes being highest at the end of the wet season (around March or April).

In this simulation, the biggest storage occurred in the fifth year. The storage design volume is most likely to be exceeded in the first few years of the simulations, because of the lower water usage in the first few years compared to the later years. After the tenth year, about 456 000 m<sup>3</sup>/year can be extracted from the storage (mill water use is 180 000 m<sup>3</sup>/year, ventilation loss is 90 000 m<sup>3</sup>/year, ore wetdown/plant washdown is 10 000 m<sup>3</sup>/year and average evaporation in this simulation is 176 000 m<sup>3</sup>/year), compared to about 392 365 m<sup>3</sup>/year of inflows into the storage (mine dewatering/contaminated laundry is 73 365 m<sup>3</sup>/year and average runoff in this simulation is 319 000 m<sup>3</sup>/year). It should be noted that because the high water usage is likely to prevent the storage from being exceeded in the later years in the simulations, it must be ensured that this high water usage can be sustained during the actual mining operation (particularly the volumes associated with the mill requirement and the ventilation loss).

### **5.2 The approach adopted here versus Kinhill-ERA approach**

There are several differences between this approach and the approach used by Kinhill-ERA (described in Appendix B1 in the Jabiluka PER Appendices) to estimate the storage capacity. In the Kinhill-ERA approach, 10 000 years of annual rainfall data were first generated, using a log-normal distribution (the 1 in 10 000 years annual rainfall estimated by Kinhill-ERA is similar to the estimate here, see section 3.3). A 15-year storage water balance simulation was carried out using a typical 15-year sequence as the base data. Ten simulations were carried out - with the first year of the base data replaced with the wettest of the 10 000 years of the generated rainfall, with the first two years of the base data replaced with the wettest two-year sequence, with the first three years replaced with the wettest three-year sequence, up to the first ten years replaced with the wettest ten-year sequence. The simulations were carried out on a monthly time step, with the annual rainfall data distributed over the 12 months using the

same monthly factors for all years of simulations (see table B1.1 in the Jabiluka PER Appendices). Mean monthly storage evaporation rates were used for the simulations (calculated using Hatton's (1997) pan coefficients times the mean monthly pan evaporation). Except for the ventilation loss, the other water use considerations were similar to those used here.



**Figure 5.1** Annual rainfall (September–August) and storage levels for a simulation using rainfall data at Oenpelli and evaporation data at Jabiru (September 1972 to August 1998)

The Kinhill-ERA simulation of a 15-year storage water balance instead of a 30-year storage water balance is to a large extent reasonable because the storage is most likely to be overtopped in the first few years (as explained in section 5.1). This is probably also the reason why Kinhill-ERA replaced the first years of the base data with the wettest sequence to mimic the extreme wet conditions. However, it is not possible to attribute directly a probability of overtopping to the storage capacity derived by Kinhill-ERA. In fact, the

Kinhill-ERA storage capacity estimate has a probability of exceedance greater than 0.01% for two reasons. First, a wet sequence starting several years after the mining began (for example, in the fifth year of the simulation like in fig 5.1) could result in a bigger storage requirement compared to the wet sequence starting from the first year of mining itself. Second, the Kinhill-ERA approach assumes that typical/average rainfall sequences follow the extreme sequence of wet years, while in reality, the annual rainfalls following the extreme sequence of wet years could be higher than average.

### **5.3 Sensitivity analyses**

It is difficult to directly compare the Kinhill-ERA approach with the approach used here. Therefore in this section we investigate the sensitivity of the differences in the water balance components and the modelling time step between the two approaches.

The simulations carried out for this section utilised the same 26 years of observed data used in section 5.1 (rainfall at Oenpelli and pan evaporation at Jabiru between September 1972 and August 1998). Six storage water balance simulations are carried out using a monthly time step and three simulations are carried out on a daily time step. The biggest storage in each of the nine simulations (which is the storage capacity required to prevent overtopping in the 26-year run) are tabulated in table 5.1 and the results are discussed below. The average annual inflows into and losses from the storage over the 26 years in all the nine simulations are the same, except Runs 5 and 6 which have higher evaporation because of the use of higher pan evaporation coefficients in two months. These simulations therefore investigate how the different considerations affect the estimate of the storage capacity. It should be noted that these simulations only provide some indication of the sensitivity of the storage capacity estimate to various parameters, and the results can be different under extreme conditions. Ideally this investigation should be based on simulations using a long data sequence (as in section 4), but was not possible to do this because of the very high computational requirement.

#### **Interannual variability in evaporation and inverse relationship between evaporation and rainfall**

Runs 1 and 2 are the same except for the use of pan evaporation data. Run 1 uses the actual monthly pan evaporation data while in Run 2, the mean monthly pan evaporation rates, averaged over the 26 years, are used. Run 1 therefore takes into account the interannual variability in evaporation and the inverse relationship between evaporation and rainfall (as is done here), while Run 2 does not (as in the Kinhill-ERA approach). Table 5.1 indicates that for this 26-year run, a 3% higher storage capacity is required when the interannual variability in evaporation and the inverse relationship between evaporation and rainfall are taken into account compared to when they are not.

#### **Actual rainfall versus monthly distribution of annual rainfall**

Runs 1 and 3 differ only in the use of rainfall data. Run 1 uses the actual monthly rainfall data, while in Run 3, the annual rainfall is distributed over the 12 months using the monthly factors in table B1.1 in the Jabiluka PER Appendices. The use of the actual monthly rainfall data would give a higher storage capacity estimate because the rainfall in some months can be significantly greater than the monthly rainfall calculated as a proportion of the annual rainfall using a typical distribution through the year. The biggest storage in this 26-year run is 1.7% greater when the actual rainfall data is used compared to when a typical distribution is used to proportion the annual rainfall to the months.

**Table 5.1** Biggest storages in the sensitivity analyses simulations

	Rainfall data	Pan evaporation data	Ventilation loss	Runoff estimation	Biggest storage (m <sup>3</sup> )
<b>Monthly simulation</b>					
1	Actual data	Actual data	Same for each month	Runoff coefficient	582 944
2	Actual data	Long-term monthly average	Same for each month	Runoff coefficient	565 939
3	Fixed distribution of annual rainfall	Actual data	Same for each month	Runoff coefficient	573 145
4	Actual data	Actual data	Lower in wet season	Runoff coefficient	589 744
5	Actual data	Actual data (Hatton's pan factor)	Same for each month	Runoff coefficient	568 502
6	Actual data	Long-term monthly average (Hatton's pan factor)	Same for each month	Runoff coefficient	540 962
<b>Daily simulation</b>					
7	Actual data	Actual data	Same for each month	Runoff coefficient	591 188
8	Actual data	Actual data	Same for each month	Conceptual storages	587 515
9	Actual data	Actual data	Lower in wet season	Conceptual storages	593 812

**Constant ventilation loss versus smaller ventilation loss in the wet season**

Runs 1 and 4 differ only in the ventilation loss calculations. Run 1 uses a constant ventilation loss throughout the year (as in the Kinhill-ERA approach) while Run 4 considers that the ventilation loss is smaller in the wet season compared to the dry season (as described in section 2.6). The biggest storage simulated in Run 4 would be greater than that in Run 1 because of the smaller potential loss available in the wet months when higher runoff inflows occur. Table 5.1 shows that in this 26-year simulation, the consideration of smaller ventilation loss in the wet season led to a 1.2% greater storage capacity compared to the use of a constant ventilation throughout the year.

**Pan evaporation coefficients**

Run 1 uses the pan evaporation factors in table 2.3, while Run 5 uses the pan evaporation factors given by Hatton (1997) and adopted by Kinhill-ERA. The factors are the same except for two months. The simulations in section 4 uses pan evaporation factors of 0.66 and 0.77 for October and April respectively, while the Kinhill-ERA simulations use higher factors of 0.75 and 0.85 in the two months respectively. For this 26-year simulation, the storage capacity estimated using the smaller pan evaporation factors is 2.5% higher (see table 5.1).

**Daily versus monthly simulation**

Runs 1 and 7 differ only in the simulation time step. The use of a bigger time step would give a smaller storage capacity estimate because of the averaging over a longer period (for example, the storage may be overtopped because of a high rainfall-runoff inflow on a particular day, but a simulation over a monthly time step may not show this because there may be sufficient losses over the month to balance this high inflow). In this 26-year simulation, the biggest storage estimate is 1.4% greater in the daily simulation compared to the monthly simulation (see table 5.1).

### **Runoff coefficient versus conceptual rainfall-runoff modelling**

Runs 7 and 8 are the same except for the method used to estimate surface runoff. In Run 7, surface runoff is estimate as a runoff coefficient multiplied by rainfall while in Run 8, surface runoff is simulated using a conceptual rainfall-runoff model, with the soil capacity parameter optimised to produce the same total runoff as in Run 7 (see section 2.1). Table 5.1 indicates that there is less than 1% difference in the storage capacities estimated in Runs 7 and 8.

### **This approach versus Kinhill-ERA approach**

Although there is only a small difference between each of the different considerations, there can be quite a large difference in the storage capacity estimate when all the differences between the approach used here and the Kinhill-ERA approach are taken into account. The storage capacity estimate in Run 9, where the approach described in sections 2 to 4 is used, is almost 10% greater than the storage capacity estimate in Run 6, where the Kinhill-ERA method and considerations are used within the framework of the storage water balance simulation described in this report. The results here only provide some indication of the sensitivity of the storage capacity estimate to various parameters and considerations, and the 10% difference between the two approaches will be different for other simulations, particularly the extremes. In addition, the differences in the basic methodology of the two approaches (ie, many annual rainfall sequences are used here to mimic the climate possibilities, while the Kinhill-ERA method combines the wettest two or three year sequence with the average climate, see section 5.2) will also result in the Kinhill-ERA storage capacity estimate being smaller than the estimate derived here.

## **References**

- Hatton T 1997. *Review of evaporation estimates Ranger Mine*. A consultancy report to ERA Environmental Services, Research Report 82 BG, Consultancy Report 97-43, August 1997.
- Jabiluka Mill Alternative Public Environment Report 1998. The report is prepared by Kinhill Pty Ltd in association with ERA Environmental Services. ISBN 1 876821 00 0.
- Jabiluka Mill Alternative Public Environment Report Technical Appendices 1998. The report is prepared by Kinhill Pty Ltd in association with ERA Environmental Services. ISBN 1 876821 00 0.
- McQuade CV 1993. *Probabilistic design for risk analysis in mine water resource management systems*. Ph.D. Thesis, University of Queensland.
- Morton FI 1993. Operational estimates of actual evapotranspiration and their significance to the science and practice of hydrology. *Journal of Hydrology* 66, 1-76.
- Vardavas IM 1987. Modelling the seasonal variation of net all-wave radiation flux and evaporation in tropical wet-dry region. *Ecological Modelling* 39, 247-268.
- Wang QJ, McConachy FLN, Chiew FHS, James R, de Hoedt GC & Wright WJ 1999. *Climatic atlas of Australia: Maps of evapotranspiration*. Australian Bureau of Meteorology, In Press.
- Wang QJ & Nathan RJ 1999. *A daily and monthly mixed (DMM) algorithm for stochastic simulation of rainfall data*. Paper to be submitted to Water Resources Research.

Wasson J, White I, Mackey B & Fleming M 1998. *The Jabiluka project – Environment issues that threaten Kakadu National Park*, Submission to UNESCO World Heritage Committee Delegation to Australia.

# Appendices

# Appendix A The DMM algorithm for stochastic generation of rainfall series

The DMM (daily and monthly mixed) algorithm previously developed by (Wang & Nathan 1999) is used in this project. The main advantages of the DMM algorithm are that it has a small number of parameters and is capable of reproducing key characteristic statistics simultaneously at daily, monthly and annual levels.

In the following, we first describe a ‘basic’ algorithm which forms the basis of the DMM algorithm and then describe the DMM algorithm itself.

## A ‘Basic’ algorithm

Daily rainfall generation algorithms usually consist of two components: one to simulate the rainfall occurrence to provide a sequence of dry and wet days, and one to simulate the rainfall amounts in wet days.

One of the widely used models for simulating rainfall occurrence is the two-state first-order Markov chain. A day can be either dry or wet (two states). The probability of being wet in any day depends only on whether the previous day was wet or dry (first order). Thus, the occurrence of rainfall can be described by 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 shown to be

$$\pi = \frac{p_{W|D}}{1 + p_{W|D} - p_{W|W}} \quad (A1)$$

One of the widely used models for simulating the rainfall amount,  $x$ , on a wet day is the gamma distribution. The density function of the gamma distribution is given by

$$f(x) = \frac{(x/\beta)^{\alpha-1} \exp(-x/\beta)}{\beta\Gamma(\alpha)} \quad (A2)$$

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

$$\mu(x) = \alpha\beta \quad (A3)$$

$$\sigma^2(x) = \alpha\beta^2 \quad (A4)$$

Seasonal variation is modelled by allowing the model parameters to vary with each of the 12 months.

Given the two-state first-order Markov chain model for rainfall occurrence and the gamma distribution model for daily rainfall amount, the rainfall total,  $X$ , over a month of  $N$  days has an expected value

$$\mu(X) = N\pi\alpha\beta \quad (A5)$$

and variance

$$\sigma^2(X) \approx N\pi\alpha\beta^2 \left[ 1 + \alpha(1-\pi) \frac{1 + p_{W|W} - p_{W|D}}{1 - p_{W|W} + p_{W|D}} \right] \quad (\text{A6})$$

There are a total of four parameters that need to be estimated for each of the 12 month. These are the two transitional probabilities  $p_{W|D}$  and  $p_{W|W}$  used to define the first order Markov chain for modelling rainfall occurrence, and the shape and location parameters  $\alpha$  and  $\beta$  used to define the gamma distribution for modelling rainfall amount in a wet day. The general approach is to estimate these parameters by using daily statistics:  $p_{W|D}$  and  $p_{W|W}$  are found by directly counting number of the wet and dry day sequences and taking appropriate ratios;  $\alpha$  and  $\beta$  are found from the mean and standard deviation of rainfall amounts in wet days by using (A3) and (A4). The problem with such an approach is that the generated rainfall data often do not reproduce the observed standard deviation,  $\sigma(X)$ , of the monthly rainfall total. While one may use (A5) and (A6) to estimate  $\alpha$  and  $\beta$  to force the matching of the standard deviation of the monthly rainfall total, the generated rainfall data often do not reproduce the observed standard deviation,  $\sigma(x)$ , of the daily rainfall amounts.

In addition, the basic algorithm does not consider the persistence of monthly rainfall total. As a result, the basic algorithm would often be unable to simulate the prolonged dry and wet spells and produce lower than observed coefficient of variation of annual rainfall total.

### The DMM algorithm

The DMM algorithm was developed (Wang and Nathan, 1999) to overcome the problem that the ‘basic’ algorithm is unable to reproduce key statistical characteristics simultaneously at daily, monthly and annual levels and, at the same time, to retain the simplicity of the ‘basic’ algorithm. The DMM algorithm involves the following steps:

- (a) For month  $i$ , generate a sequence of wet and dry days for the whole month using a two-state first-order Markov chain with transitional probabilities,  $\hat{p}_{W|D}$  and  $\hat{p}_{W|W}$ , which are estimated from the countings of dry-wet and wet-wet day sequences.
- (b) For any wet day in that month, generate rainfall amount  $x^d$  from a gamma distribution with parameters  $\hat{\alpha}^d$  and  $\hat{\beta}^d$  which are estimated from the mean and variance of daily rainfall amounts by using (A3) and (A4).
- (c) For the same wet day, generate, using exactly the same cumulative probability as in step (b), a twin rainfall amount  $x^m$  also from a gamma distribution but with parameters  $\hat{\alpha}^m$  and  $\hat{\beta}^m$  which are estimated from the mean and variance of monthly rainfall total by using (A5) and (A6).
- (d) Manipulate the monthly total of the daily rainfall generated in step (c),  $\tilde{X}_i = \sum x^m$ , to produce a new monthly total  $X_i$  by using the following first-order autoregressive equation:

$$\frac{X_i - \mu(X_i)}{\sigma(X_i)} = \rho \frac{X_{i-1} - \mu(X_{i-1})}{\sigma(X_{i-1})} + (1 - \rho^2)^{1/2} \frac{\tilde{X}_i - \mu(X_i)}{\sigma(X_i)} \quad (\text{A7})$$

The subscripts  $i-1$  and  $i$  in (A7) denote the previous and current months respectively. The lag-1 serial correlation  $\rho$  may vary with month and is estimated from monthly rainfall data.

(e) Produce a new daily rainfall series  $x$  for that month by multiplying all  $x^d$  by a constant

$$(X_i / \sum x^d) \tag{A8}$$

To recast the algorithm, Steps (a) and (b) generate a rainfall series that preserves the daily rainfall characteristics. Steps (a) and (c) generate a rainfall series that preserves the mean and standard deviation of monthly rainfall total. Step (d) modifies the monthly rainfall total produced in Step (c) to simulate persistence in monthly rainfall. Step (e) modifies the daily rainfall series produced in Step (b) so that it has the ‘correct’ monthly characteristics.

The final daily rainfall series preserves exactly the mean, standard deviation, and lag-1 autocorrelation coefficient of the monthly rainfall total. How well the final daily rainfall series preserve the daily statistics depends on the degree of adjustment required in Step (e) to the daily rainfall series produced in Step (b) that already preserves exactly the daily rainfall statistics. If the degree of adjustment is small, we would expect that the final adjusted daily rainfall series would preserve reasonably well the daily statistics.

The DMM algorithm was specifically designed to minimize the adjustment required in Step (e). The core of the design is to generate two daily rainfall series which closely 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. The two daily rainfall series have an identical wet and dry day sequence as produced in Step (a). In generating the rainfall amounts for a wet day, the algorithm uses the same non-exceedance probability for both series so that a high rainfall amount in the first series will correspond to a high rainfall amount in the second series, and vice versa. Thus, the two daily rainfall series resemble each other closely. This leads to small adjustment necessary in Step (e).

Note that when obtaining  $\hat{\alpha}^d$  and  $\hat{\beta}^d$  in Step (b), the sample estimate of  $\sigma^2(x)$  in equation (A4) should be adjusted to sample estimate of  $\sigma^2(x) - \sigma^2(X) / (N\pi)^2$  to remove the variance in daily rainfall that has already been accounted for by the inter-annual variation of monthly rainfall total.

The algorithm requires, for each month, sample estimates of two transitional probabilities, mean and variance of rainfall amount on a wet day, and variance of monthly rainfall total, and autocorrelation coefficient of monthly rainfall total. The mean of monthly rainfall total is not another independent statistic as it can be found from the two transitional probabilities, the mean rainfall amount on a wet day, and the number of days in a month. Thus, the algorithm has in total 72 parameters. The number of parameters may be reduced, if necessary, by using a Fourier series approximation to the variation of the parameters with month.

## Appendix B An Algorithm for stochastic generation of monthly evaporation

A simple regression model is used for simulating monthly evaporation. The model reproduces the mean, standard deviation, lag-1 and lag-2 autocorrelation of monthly evaporation, and cross correlation between monthly evaporation and rainfall. The model is given by

$$\frac{E_i - \mu(E_i)}{\sigma(E_i)} = a_i + b_i \frac{R_i - \mu(R_i)}{\sigma(R_i)} + c_i \frac{E_{i-1} - \mu(E_{i-1})}{\sigma(E_{i-1})} + d_i \frac{E_{i-2} - \mu(E_{i-2})}{\sigma(E_{i-2})} + e_i \varepsilon_i \quad (\text{B1})$$

where subscripts  $i$ ,  $i-1$ ,  $i-2$  denote current month, last month, the month before last month respectively;  $E$  and  $R$  are monthly evaporation and rainfall respectively; coefficients  $a$ ,  $b$ ,  $c$ , and  $d$  are found by regression;  $e$  is the standard deviation of regression residues;  $\varepsilon$  is a standard normal variate. All the coefficients are found by regression.