Simulation of Daily Rainfall Sequences Using Markov Chain

HUYNH NGOC PHIEN AND SOMKIAT WARAKITTIMALEE
Asian Institute of Technology, P.O. Box 2754, Bangkok, Thailand

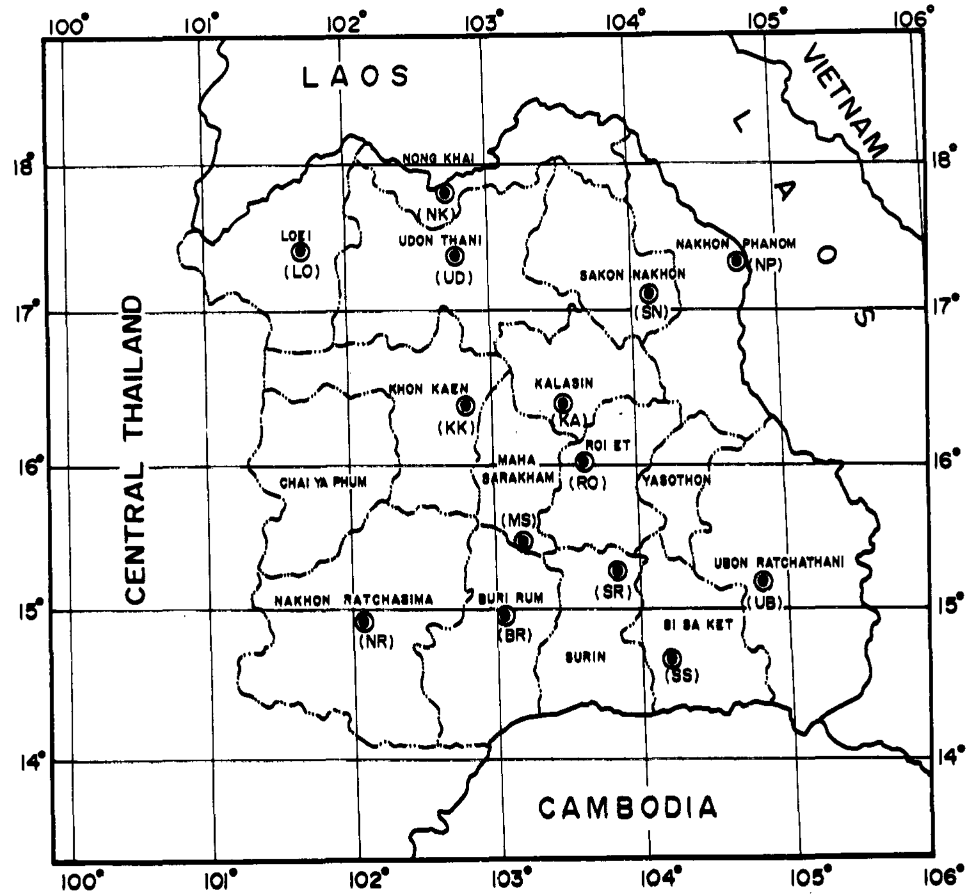

Figure 1
Map of the northeast with the locations of selected rainfall stations

Abstract

Several characteristics of rainfall sequences in Northeast Thailand are incorporated in the development of a first-order Markov chain model for the generation of daily rainfall amounts in that region. The developed model is able to *reproduce* the distribution of monthly rainfall; the mean, standard deviation, and skewness coefficient of maximum amounts of daily rainfall; the first two moments of the number of wet and dry days; and the frequencies of wet and dry spells *of the historical data.*

Introduction

The Northeast of Thailand (Fig. 1) has an area of about 167 300 km^2 and a population of about 15 million. This region often faces severe droughts in many months of the year and also suffers from severe floods in the wettest periods. Since farming is the primary source of income for the inhabitants of the region, this unfavorable situation explains to some extent their poor economic conditions.

In an attempt to improve the economic situation in this part of the country, several development programmes have been formulated. In the water resources section, several studies have been carried out with an aim to arrive at an efficient use of available water and a suitable cropping pattern needed in the region. Important results regarding the characteristics of rainfall sequences and drought conditions which are based upon the existing historical data have been made available (Phien *et al.*, 1980 a,b). In this study, additional characteristics of rainfall sequences in the region are presented. These in turn will be used along with some previous results in the development of a model

TABLE 1
LIST OF RAINFALL STATIONS AND THEIR LOCATIONS

Station	Place (Location)	Lat. N (Location)	Long. E (Location)	Period of Record Used
BR	A. Muang, Buriram	14° 59' 30"	103° 06' 30"	1953 – 1977
KA	A. Muang, Kalasin	16° 25' 50"	103° 30' 40"	1952 – 1977
KK	A. Muang, Khon Kaen	16° 25' 40"	102° 50' 20"	1952 – 1977
LO	Huai Nam Man Weir, A. Muang, Loei	17° 29' 00"	101° 42' 30"	1952 – 1977
MS	A. Phayakkhaphum Phisai, Maha Sarakham	15° 30' 50"	103° 11' 50"	1952 – 1977
NK	A. Muang, Nong Khai	17° 52' –	102° 45' –	1952 – 1977
NP	A. Muang, Nakhon Phanom	17° 24' 35"	104° 47' 00"	1952 – 1973
NR	A. Muang, Nakhon Ratchasima	14° 58' 10"	102° 08' 10"	1952 – 1977
RO	A. Muang, Roi Et	16° 03' –	103° 41' –	1952 – 1977
SN	A. Muang, Skhon Nakhon	17° 10' –	104° 09' –	1952 – 1977
SR	A. Rattanaburi, Surin	15° 19' 33"	103° 51' 24"	1952 – 1975
SS	A. Khukhan, Si Sa Ket	14° 42' 58"	104° 12' 03"	1952 – 1977
UB	A. Muang, Ubon Ratchathani	15° 13' 18"	104° 51' 42"	1952 – 1977
UD	R.I.D. Office, A. Muang, Udon Thani	17° 26' –	102° 46' –	1953 – 1977

(–): not available in the provided data sheets

TABLE 2
RESULTS OF FITTING ANNUAL MAXIMUM SERIES OF DAILY RAINFALL

Station	n	Lognormal	Pearson Type III	Log-Pearson Type III	Exponential	Gumbel	Ev3
BR	25	–	–	–	–	+	+
KA	26	–	–	–	–	+	+
KK	26	–	+	–	–	+	+
LO	26	–	–	–	–	+	+
MS	26	–	–	+	–	+	+
NK	22	–	+	–	–	+	+
NP	26	–	–	–	–	+	+
NR	26	–	–	–	–	+	+
RO	26	–	–	–	–	+	+
SN	26	–	–	–	–	+	+
SR	24	–	+	–	–	+	+
SS	26	–	–	–	–	+	+
UB	26	–	–	–	–	+	+
UD	25	–	–	+	–	+	+

Notes n: number of events (length of record)
–: not fitted
+: fitted

TABLE 3
RESULTS OF FITTING PARTIAL DURATION SERIES BY THE SHIFTED EXPONENTIAL DISTRIBUTION

Threshold Value

Station	60 (1)	60 (2)	60 (3)	65 (1)	65 (2)	65 (3)	70 (1)	70 (2)	70 (3)	75 (1)	75 (2)	75 (3)	80 (1)	80 (2)	80 (3)
BR	0,090	117	0,126	0,048	94	0,140	0,062	80	0,149	0,049	72	0,157	0,031	66	0,164
KK	0,063	61	0,171	0,088	48	0,192	0,074	42	0,205	0,108	31	0,238	0,078	29	0,246
NK	0,103	101	0,135	0,124	81	0,149	0,131	63	0,168	0,065	45	0,198	0,078	38	0,215
NR	0,098	77	0,152	0,091	57	0,177	0,119	49	0,190	0,151	40	0,210	0,158	31	0,238
RO	0,087	85	0,145	0,078	74	0,155	0,094	64	0,167	0,121	48	0,192	0,150	37	0,218
UB	0,060	97	0,138	0,039	72	0,157	0,048	58	0,175	0,061	47	0,194	0,181	40	0,210

Notes: (1) Kolmogorov-Smirnov statistic.
(2) Number of observations.
(3) Critical value at 5% significance level.

for daily rainfall generation. Finally, a simulation model for the evaluation of rainfed irrigation in the region is outlined.

Characteristics of Rainfall Sequences

The distributions of annual and monthly rainfall sequences in the region have been reported by Phien *et al.* (1980a). In order to supply more information for the development of a model for daily rainfall generation, to be discussed later, additional characteristics of rainfall sequences were sought. These are the distributions of maximum amounts of daily rainfall, wet and dry spells. For this purposes, data at 14 stations listed in Table 1 (also shown in Fig. 1) were used. These data sets have been found most representative for the whole region (AIT, 1978).

Distribution of Maximum Daily Rainfall

The maximum daily amount of rainfall investigated in this study comprises two sequences, namely, the annual maximum and the partial duration series. The annual maximum series consists of the largest amounts of rainfall, one in a year, while the other series is formed by taking all the amounts of rainfall which are equal or greater than a threshold value.

There are many distributions which are usually used to fit the maximum values. In this study, six distributions which are most useful in hydrological applications were used. These are the lognormal, Pearson Type III, log-Pearson Type III, shifted exponential, Gumbel (or Type I) and Type III extreme value (EV3) distributions. Their descriptions and the methods of estimating the parameters involved can be found in many publications, especially in Kite (1977).

The results of the fitting are collected in Table 2 for annual maximum series, where the parameters involved were all estimated by the method of maximum likelihood. It is clear that the Gumbel and EV3 distributions can be used to fit them in all cases. However, inspection of the computer outputs reveals that the Gumbel distribution gives the better fit. For more details on this distribution, refer to the work of Phien and Arbhabhirama (1980).

Regarding the partial duration series, the threshold value is normally determined so that the number of elements in the resulting sequence is in between n and 5n, where n is the length of record in years. It was found that the partial duration series consisting of daily rainfall amounts exceeding 60 mm can be fitted by the shifted exponential distribution of the following form:

$$F(R) = 1-\exp[-(R-R_0)/\lambda], \quad R \geq R_0 \quad (1)$$

where R denotes the rainfall amount, R_0 is the threshold value, $\lambda > 0$ is the parameter, and F is the probability distribution function. Typical results of the fitting are given in Table 3, where the computed values of the Kolmogorov-Smirnov statistic are much less than the corresponding critical values at the 5% significance level, indicating that the fitting is satisfactory.

Distributions of Wet and Dry Spells

The number of wet and dry spells of different lengths was computed from historical sequences. In most cases, the number of wet spells is highest at the length of one day and decreases to zero as the length increases. In the wet season, April through October, the number of dry spells has the same variation, but in the dry season, November through March, it fluctuates highly.

In this study, the modifications of the negative binomial distribution, namely the shifted negative binomial distribution, the geometric distribution and the logarithmic series distribution (as described in Buishand, 1977) were employed. Unfortunately, none of them was found satisfactory in fitting these spells in all months. Details of the fitting were presented by Warakittimalee (1980).

Model for Daily Rainfall Generation

For the generation of *amounts* of daily rainfall at a station, Markov chain models have been proven to be very useful. Research work on this problem was reported by Khanal and Hamrick (1974) in which a first-order Markov chain with 14 states was employed. This work was extended by Allen and Haan (1975) for the simulation of daily rainfall at Kentucky. These successful attempts along with the work of Selvalingam and Miura (1978) suggest the use of a Markov chain model in the present study. Chin (1977) pointed out that the proper order of the selected Markov chain must be determined using the identification procedure by Akaike (1974) and Tong (1975). However, such a determination might be necessary when one wants to model the occurrence of wet and dry days rather than to simulate the amounts of daily rainfall. So, the first-order was adopted and this simplifies the determination of the number of states to be used and the estimation of the corresponding transition probabilities.

Due to the rainfall pattern in the region (Phien *et al.*, 1980a) the transition probabilities between different states may be assumed to be stationary within each month as made by Allen and Haan (1975). There remains further consideration of the number of states and the distribution within each class interval corresponding to a state. Since there are no general methods for determining the number of states, one normally relies upon the trial and error procedure. In order to provide a guidance for such a procedure, it is necessary to set up the criteria for the evaluation of the selected model. The model being aimed at in this study should be able to reproduce the following statistics of the historical data at a station:

(1) The distribution of monthly rainfall sequences

(2) The mean, standard deviation and skewness coefficient of maximum amounts of daily rainfall in each month

(3) The mean and standard deviation of the number of wet or dry days in each month

(4) The frequencies of wet and dry spells in each month

For the case of rainfall data used in this study, the minimum amount recorded is 0,1 mm. However, for irrigation consideration, a day with rainfall less than 1 mm is normally treated as a dry day, and this practice is also adopted here. Thus the first state corresponds to the class interval [0, 1].

As seen from the foregoing analysis, the partial-duration series of maximum amounts of daily rainfall exceeding 60 mm can be fitted very well by the shifted exponential distribution. However, since the daily amounts of rainfall in the region fluctuate highly from month to month, the use of only one value for the parameter λ (Eq. 1) throughout the year as in Allen and Haan (1975) is not realistic. Selvalingam and Miura (1978) employed a trial and error procedure in determining the threshold value and the parameter λ, but this introduces many more parameters and is time consuming. To compromise these two situations, this study uses the same threshold value for maximum daily rainfall amounts in all the twelve months. Consequently, the lower limit of the last class interval should be ≥ 60 mm. This implies that for a month (in the dry season) where daily amounts of rainfall do not exceed the threshold value, the last class interval does not apply.

By grouping daily amounts of rainfall in the range 1 – 65 mm into several class intervals and analysing them, it was found that these amounts in some classes are better fitted by the beta distribution than the uniform distribution. However, fitting rainfall amounts in each class interval by the beta distribution would complicate the model in paramter estimation which sometimes is unreliable because of the fact that the frequencies of occurrences in the intervals with higher values for class limits are very low (noting that the maximum length of record is 26 years). Therefore, the uniform distribution was used for all class intervals except the first and the last.

Following the suggestion of Allen and Haan (1975), the increase of class widths as a geometric progression was applied. After several trials seven class intervals with limits obtained from Allen and Haan (1975) by changing 100 x in. to mm as shown in Table 4 were finally adopted. Consequently, the threshold R_o in Eq. 1 is 63 mm. In summary, the selected Markov chain is of the first-order and has 7 states. The first state corresponds to a dry day whereby rainfall is practically considered as zero. States 2 through 6 cover all rainfall amounts between 1 mm through 63 mm, where they are assumed to be uniformly distributed in each corresponding class interval. The last state is reserved for rainfall amounts exceeding 63 mm which are represented by the shifted exponential distribution.

TABLE 4
STATES AND CORRESPONDING CLASS LIMITS

State	Limits (mm)
1	0,0 – 1,0
2	> 1,0 – 3,0
3	> 3,0 – 7,0
4	> 7,0 – 15,0
5	> 15,0 – 31,0
6	> 31,0 – 63,0
7	> 63,0 – + ∞

Generation Procedure

With the above discussions, the generation procedure at a selected station is as follows:

(1) For each month, the transition probabilities, P_{ij}, and the parameter, λ, are estimated from the historical data by the following equations:

$$P_{ij} = f_{ij} \Big/ \sum_{j=1}^{7} f_{ij} \qquad i = 1, \ldots, 7 \qquad (2)$$

$$\lambda = \bar{R} \cdot R_o$$

where f_{ij} is the historical frequency of transition from state i to state j and $\bar{R}$ is the mean of daily rainfall amounts exceeding R_o (= 63 mm) in that month. However, if all daily rainfall amounts are less than R_o, then λ is simply set equal to zero. Then the cumulative transition probabilities, Q_{ij}, are computed by:

$$Q_{ij} = \sum_{k=1}^{j} P_{ik} \qquad i, j = 1, \ldots, 7 \qquad (4)$$

(2) A uniform random number, U, between 0 and 1 is generated.

(3) Knowing the state, i, of one day, the state, j, of the following day is determined by comparing U with Q_{ij}:

$$\text{if} \quad U \leq Q_{i,1}, \qquad \text{then } j = 1 \qquad (5)$$

$$\text{and if } Q_{i,k-1} < U \leq Q_{i,k}, \qquad \text{then } j = k\,(k = 2, \ldots, 7) \qquad (6)$$

When the two conditions k = 7 and λ = 0 are simultaneously met, this step is repeated.

(4) If j = 1, the rainfall amount, R, is set equal to zero. If 2 ≤ j ≤ 7, a new uniform random number, V, on (0,1) is generated and R is computed by

$$R = V(C_{j+1} - C_j) + C_j, \qquad \text{if } 2 \leq j \leq 6 \qquad (7)$$

where C_j and C_{j+1} are the lower and upper limits of the jth class interval, or by

$$R = R_o - \lambda \ln(1 - V), \qquad \text{if } j = 7 \qquad (8)$$

Steps (2) through (4) are repeated until the desired length of the generated sequence is reached.

It should be noted that Eq. 8 is obtained from Eq. 1 because F(R) is a random variable uniformly distributed on (0,1). Since 1 – V is also a uniform random number on (0,1), Eq. 8 can be simply rewritten as

$$R = R_o - \lambda \ln V \qquad (9)$$

where V is used instead of 1 – V. This was adopted by Gordon (1978) in discussing methods for generating exponential variables.

A FORTRAN computer program was developed and run on the IBM 370/145 system of the Regional Computer Center of the Asian Institute of Technology (AIT). Uniform random numbers were generated by the IBM scientific sub-routine RANDU.

To start the generation, the state of the first day must be given. It was found after some test runs that this initial state has no important effects on the generated sequences as long as the stipulated criteria are concerned. Even so, it was assumed here to be the dry state because there is a very high probability that the first day of any month is a dry day, as revealed by all the historical data employed.

Evaluation of the Model

To evaluate the model performance, the stipulated criteria were used. In the following, the generated and historical sequences are compared with regard to these criteria.

Distribution of monthly rainfall sequences

As reported by Phien *et al.* (1980a) monthly rainfall sequences in Northeast Thailand consist of many zero values for the months of the dry season, November through March, and they can be fitted very well by the Leakage law or *loi des fuites* as employed by Buishand (1977). In the wet season, April through October, they can be fitted also by the lognormal and gamma distributions of which the probability density functions were described by Yevjevich (1972). For the leakage law, it is more convenient to use the probability *distribution* function:

$$\Pr(R = O) = \exp(-\Theta) \qquad (10)$$

$$(\Pr(R \leq r) = \exp(-\Theta) + \int_0^r \exp(-\Theta - \varrho x)(\varrho\Theta/x)^{1/2} I_1[2(\varrho\Theta x)^{1/2}]dx, \; r > O \qquad (11)$$

where R denotes rainfall, I_1 is the modified Bessel function of order 1, Pr stands for probability, and ϱ and Θ are the two parameters of the leakage law.

Typical results obtained from a generated sequence of the same length as the historical data (which were selected at random) for the case of Buriram are shown in Table 5. Inspection of this table reveals that the generated sequence in each month has the same distribution as that of the historical data. Consequently, the monthly mean, standard deviation, skewness coefficient, etc., of the historical data are reproduced by the model.

The above observation is also valid for other stations under consideration. To support this statement, Table 6 is prepared where the results of fitting by the leakage law, which is applicable to all the twelve months of the year, are collected for several stations. First, it can be seen that the computed values of the Kolmogorov-Smirnov statistic corresponding to the historical data in all months are much less than the corresponding critical values at the 5% significance level, indicating that monthly historical rainfall sequences are fitted very well by the leakage law. Having established this fact, an attempt was then made to fit generated sequences in each month by the same distribution. This means that for each month, the generated sequences are fitted by the leakage law of which the parameters Θ and ϱ have been estimated from the historical data (when fitting the historical sequence). For simple presentation, only the average value of the Kolmogorov-Smirnov statistic computed from 10 generated sequences, each having the same length as the respective historical record are shown. These values (as well as each individual value obtained from a generated sequence not reported in Table 6) are all smaller than the critical values, and therefore the model preserves the distribution of historical data in each month.

Maximum daily rainfall in each month

For the model to be simple, the threshold value was set equal to 63 mm throughout the year as previously mentioned. This implies that for a month (in the dry season) where daily rainfall does not exceed the threshold value, the last class is not used.

TABLE 5
VALUES OF THE KOLMOGOROV-SMIRNOV STATISTIC(*) IN FITTING MONTHLY RAINFALL SEQUENCES AT BURIRAM (BR)

Month	Historical sequence (I)			Generated sequence(+) (II)		
	(1)	(2)	(3)	(1)	(2)	(3)
April	0,111	0,147	0,121	0,176	0,203	0,175
May	0,089	0,147	0,108	0,139	0,166	0,145
June	0,119	0,146	0,133	0,149	0,179	0,157
July	0,050	0,090	0,061	0,191	0,209	0,195
August	0,111	0,148	0,123	0,161	0,185	0,167
September	0,062	0,093	0,073	0,173	0,188	0,177
October	0,072	0,127	0,092	0,134	0,153	0,140
November	0,072	–	–	0,164	–	–
December	0,027	–	–	0,062	–	–
January	0,065	–	–	0,079	–	–
February	0,046	–	–	0,119	–	–
March	0,128	–	–	0,192	–	–

Notes:
(1) by leakage law
(2) by lognormal distribution
(3) by gamma distribution
(–) not applicable
(*) critical value at 5% significance level = 0,264
(+) the parameters of the distributions involved are also estimated from the historical data as in (I)

TABLE 6
VALUES OF THE KOLMOGOROV-SMIRNOV STATISTIC IN FITTING GENERATED SEQUENCES BY THE LEAKAGE LAW AT SOME SELECTED STATIONS

Station	Critical Value		Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.	Jan.	Feb.	Mar.
			Month											
BR	0,264	(1)	0,062	0,105	0,082	0,094	0,091	0,097	0,061	0,074	0,114	0,045	0,225	0,167
		(2)	0,140	0,190	0,152	0,204	0,149	0,161	0,129	0,139	0,116	0,055	0,118	0,169
KA	0,259	(1)	0,059	0,118	0,111	0,111	0,078	0,090	0,072	0,099	0,037	0,043	0,096	0,066
		(2)	0,165	0,163	0,158	0,208	0,148	0,144	0,164	0,149	0,096	0,125	0,108	0,130
KK	0,259	(1)	0,120	0,048	0,107	0,078	0,102	0,097	0,100	0,188	0,076	0,101	0,191	0,176
		(2)	0,144	0,159	0,119	0,196	0,218	0,150	0,123	0,155	0,100	0,116	0,204	0,159
LO	0,259	(1)	0,053	0,070	0,048	0,106	0,110	0,084	0,079	0,202	0,150	0,102	0,083	0,111
		(2)	0,174	0,135	0,139	0,169	0,146	0,156	0,141	0,203	0,133	0,158	0,153	0,156
MS	0,259	(1)	0,073	0,172	0,088	0,097	0,132	0,086	0,061	0,186	0,049	0,050	0,146	0,162
		(2)	0,142	0,151	0,145	0,170	0,153	0,162	0,128	0,156	0,085	0,086	0,120	0,148
NK	0,281	(1)	0,114	0,111	0,112	0,102	0,174	0,082	0,075	0,150	0,062	0,101	0,152	0,128
		(2)	0,168	0,162	0,158	0,166	0,179	0,145	0,155	0,123	0,094	0,157	0,168	0,147

Notes: (1) from historical data
(2) from 10 generated sequences

TABLE 7
MEAN OF MAXIMUM DAILY RAINFALL IN EACH MONTH (mm)

Station		Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.	Jan.	Feb.	Mar.
		Month											
NP	(1)	37,57	60,88	104,52	80,96	119,37	78,93	21,82	4,19	0,67	5,83	12,88	22,64
	(2)	43,64	61,01	108,20	82,17	111,73	78,33	23,21	2,79	1,07	6,05	12,41	23,28
NR	(1)	35,18	46,13	42,21	39,86	36,95	69,56	48,92	12,95	2,61	2,24	15,45	28,47
	(2)	36,98	45,78	46,24	35,91	33,98	70,69	48,32	15,18	2,82	3,16	15,05	28,71
RO	(1)	35,06	56,42	54,07	16,52	62,47	73,18	30,95	6,22	1,05	2,09	9,00	14,33
	(2)	29,90	57,90	52,53	63,25	60,56	71,84	35,52	6,19	1,37	2,68	9,39	12,74
SN	(1)	33,25	59,66	57,43	49,15	75,15	68,52	29,80	5,34	3,83	4,20	9,38	16,50
	(2)	32,16	60,66	58,46	52,77	73,83	67,33	30,13	4,75	3,74	5,27	9,29	15,42
SR	(1)	99,56	55,67	54,22	56,13	52,23	76,09	38,14	9,73	0,46	1,95	9,30	21,89
	(2)	90,76	54,98	56,72	55,65	53,60	76,22	37,55	9,81	0,81	2,72	11,49	20,80

Notes: (1) from historical data.
(2) *from ten generated sequences.*

Even with this simplification as well as with the use of the uniform distribution in class intervals 2 through 6, the results collected in Tables 7 – 9 show that the historical maximum daily rainfall amount in each month is reproduced by the model in terms of the first three moments because the mean, standard deviation, and skewness coefficient obtained from simulated sequences are quite close to those of the historical record. These statistics were obtained by taking the average of their values from ten generated sequences, each having the same length as the corresponding historical record. Although this number of sequences employed is small, a good agreement has been reached.

Numbers of wet and dry days in each month

It must be recalled that in this study, a day with rainfall amount greater than 1 mm is considered as a wet day, otherwise it is a dry day. As previously seen, all the distributions employed cannot satisfactorily represent the wet and dry spells, therefore, only the numbers of wet and dry days are considered. It suffices then to deal with the number of wet days, for example. The results based on ten generated sequences are shown in Tables 10 and 11, for the mean and standard deviation, respectively. Good agreement between simulated and historical values can be clearly seen.

TABLE 8
STANDARD DEVIATION OF MAXIMUM DAILY RAINFALL IN EACH MONTH

Station		Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.	Jan.	Feb.	Mar.
NP	(1)	24,74	24,50	81,41	33,00	63,07	35,10	20,45	7,18	3,06	15,16	16,93	16,86
	(2)	25,07	21,92	78,40	37,61	56,25	33,86	21,18	6,83	3,83	14,55	16,82	18,12
NR	(1)	21,55	26,85	22,24	21,00	17,05	32,60	26,12	16,25	5,07	4,19	18,81	22,82
	(2)	22,69	23,49	23,48	26,50	16,63	34,45	21,35	19,54	5,51	5,34	19,28	22,08
RO	(1)	21,83	24,78	27,33	31,02	28,65	42,61	19,54	9,12	4,29	5,50	13,92	14,70
	(2)	21,52	23,27	23,10	32,93	25,75	43,54	19,85	10,18	4,13	6,00	12,95	15,01
SN	(1)	17,82	38,70	20,28	35,20	81,27	41,89	28,24	11,87	9,88	7,53	16,52	14,02
	(2)	18,08	38,82	19,67	34,82	79,81	41,91	27,93	10,59	10,75	8,12	14,79	16,93
SR	(1)	18,85	29,43	28,76	24,03	22,90	36,40	22,91	16,16	1,70	4,82	16,23	20,73
	(2)	18,56	26,51	30,86	24,63	23,09	35,31	23,00	15,27	2,47	4,23	17,62	21,69

Notes: (1) from historical data.
(2) from ten generated sequences.

TABLE 9
SKEWNESS COEFFICIENT OF MAXIMUM DAILY RAINFALL IN EACH MONTH

Station		Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.	Jan.	Feb.	Mar.
NP	(1)	1,67	1,11	3,36	1,57	1,82	0,31	2,06	1,97	4,74	3,87	1,07	0,48
	(2)	1,33	0,68	3,77	1,50	0,92	0,89	1,77	1,70	4,82	3,53	1,49	0,64
NR	(1)	0,80	1,35	1,56	1,79	0,32	0,67	1,29	1,12	2,22	2,13	1,02	1,51
	(2)	0,39	1,90	1,60	1,24	0,50	0,66	0,84	1,12	2,24	2,25	1,18	1,10
RO	(1)	0,43	0,50	1,26	0,82	0,56	1,58	0,55	2,00	4,68	3,47	2,18	1,48
	(2)	0,10	0,6	0,97	1,08	0,63	1,21	0,39	2,42	4,62	2,95	2,45	0,97
SN	(1)	0,11	2,87	0,07	2,66	4,19	1,80	1,88	3,16	2,62	1,77	1,92	0,88
	(2)	0,09	2,53	0,09	2,71	3,55	1,20	1,46	3,04	3,36	1,71	1,71	0,95
SR	(1)	0,25	0,75	1,18	0,75	0,94	1,84	1,30	2,08	3,74	2,31	1,92	0,82
	(2)	0,19	0,93	1,00	0,68	0,70	1,11	0,98	2,02	3,99	2,48	1,42	0,70

Notes: (1) from historical data.
(2) from ten generated sequences.

TABLE 10
MEAN OF NUMBER OF WET DAYS IN EACH MONTH (DAYS)

Station	SS		UB		UD	
Month	(1)	(2)	(1)	(2)	(1)	(2)
Apr.	2,77	2,89	5,38	5,46	6,92	6,99
May	5,27	5,52	11,96	12,06	14,16	14,08
Jun.	5,27	5,20	15,00	14,84	11,24	10,81
Jul.	6,85	6,72	16,15	15,91	15,04	14,65
Aug.	8,08	7,83	18,58	18,57	15,64	15,50
Sep.	9,15	8,83	16,54	16,42	12,44	12,98
Oct.	4,77	4,38	7,31	7,38	3,44	3,39
Nov.	1,19	1,08	1,77	1,60	0,44	0,35
Dec.	0	0	0,42	0,47	0,24	0,22
Jan.	0	0	0,19	0,20	0,96	0,94
Feb.	0,15	0,21	0,85	0,77	2,04	2,32
Mar.	1,15	1,17	2,31	2,32	2,92	2,88

Notes: (1) from historical data.
(2) from ten generated sequences.

TABLE 11
STANDARD DEVIATION OF NUMBER OF WET DAYS IN EACH MONTH (DAYS)

Station	SS		UB		UD	
Month	(1)	(2)	(1)	(2)	(1)	(2)
Apr.	2,08	1,81	1,84	2,13	2,80	2,48
May	2,63	2,27	2,64	3,41	2,75	3,19
Jun.	2,34	2,34	4,14	2,96	3,42	3,22
Jul.	3,14	3,60	3,91	3,73	3,60	3,55
Aug.	3,00	2,86	2,40	3,01	4,46	4,36
Sep.	3,40	3,61	2,68	3,16	4,19	3,52
Oct.	2,67	2,62	3,34	3,18	2,53	2,06
Nov.	1,50	1,05	1,63	1,39	0,82	0,69
Dec.	0	0	0,75	0,74	0,59	0,58
Jan.	0	0	0,40	0,44	1,17	1,24
Feb.	0,46	0,56	1,08	0,95	1,67	1,69
Mar.	1,35	1,07	1,96	1,67	1,68	1,94

Notes: (1) from historical data.
(2) from ten generated sequences.

TABLE 12
FREQUENCY OF WET SPELLS AT BURIRAM (BR)

Month		Length of spells (days) 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	>15
Apr.	(1)	49	12	4	0	0	0	0	0	0	0	0	0	0	0	0	0
	(2)	52	12	3	1	0	0	0	0	0	0	0	0	0	0	0	0
Jun.	(1)	97	20	12	4	0	2	0	0	1	0	0	0	0	0	0	0
	(2)	90	27	11	3	1	1	0	0	1	0	0	0	0	0	0	0
Aug.	(1)	91	28	7	3	2	1	0	0	2	0	0	0	0	0	0	0
	(2)	81	32	11	4	2	1	0	0	0	0	0	0	0	0	0	0
Oct.	(1)	56	18	9	5	2	0	0	0	0	0	0	0	0	0	0	0
	(2)	56	21	9	4	2	0	0	0	0	0	0	0	0	0	0	0
Dec.	(1)	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	(2)	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Feb.	(1)	11	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	(2)	11	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Notes: (1) from historical data.
(2) from ten generated sequences (after rounding up).

TABLE 13
FREQUENCY OF DRY SPELLS AT BURIRAM (BR)

Month		Length of spells (days) 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	>15
Apr.	(1)	12	10	6	5	4	7	9	4	4	5	4	1	2	2	2	10
	(2)	10	8	8	9	7	6	5	5	4	4	3	2	2	2	2	10
Jun.	(1)	41	36	27	13	4	5	4	4	3	2	2	1	3	0	0	3
	(2)	38	29	21	14	13	8	6	4	3	3	2	1	2	6	1	1
Aug.	(1)	38	33	18	12	12	6	7	3	3	2	2	1	1	1	1	3
	(2)	37	29	19	14	11	9	6	8	3	3	2	2	1	1	1	2
Oct.	(1)	24	18	11	13	8	2	0	3	5	2	3	1	2	2	1	10
	(2)	18	14	11	12	8	3	2	5	4	3	3	3	2	2	1	6
Dec.	(1)	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	25
	(2)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	25
Feb.	(1)	2	1	0	1	1	2	2	1	1	1	1	0	0	0	0	24
	(2)	1	1	1	1	1	2	1	1	1	1	1	0	1	0	0	23

Notes: (1) from historical data.
(2) from ten generated sequences (after rounding up).

Runs of wet and dry days in each month

The criterion of reproducing the runs of wet and dry days in each month implies criterion 3. However, this implication is valid only when the whole structure of the runs is reproduced. The structure of the sequence of wet and dry days in a month is normally modelled by a Markov chain, and as previously discussed, such a modelling involves the determination of the order of the chain to be used, which is not simple. In this study, only the frequencies of the dry and wet spells at different lengths without regard to the order of occurrence were employed and thus the third criterion is still needed. The frequencies obtained by taking the average values from ten generated sequences are shown in Table 12 for the wet spells, and in Table 13 for the dry spells at Buriram. Inspection of these two tables reveals that the model can reproduce the frequencies of wet and dry spells of the historical record at that station. By examining the related computer outputs, it was found that the above statement (for Buriram) re-

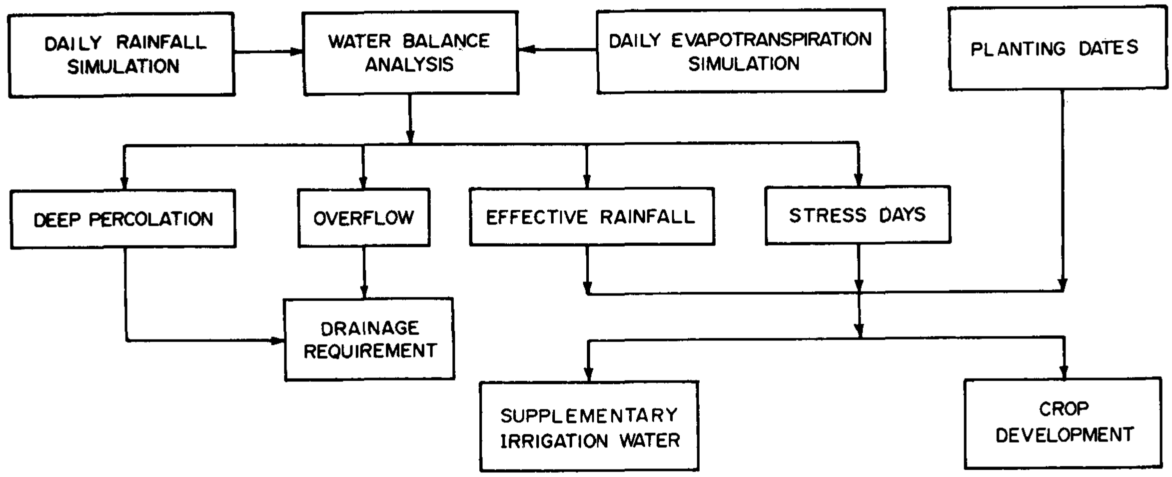

Figure 2
Sketch of the simulation model for rainfed irrigation assessment

mains valid also for all other stations listed in Table 1.

In summary, all the four stipulated criteria are met by the developed model which is rather simple and easy to use.

Simulation Model for Assessment of Rainfed Irrigation

In a previous study (Phien *et al.*, 1980b), a simple mathematical model was developed for the assessment of rainfed irrigation in Northeast Thailand, which was based upon the historical data of rainfall. In order to provide more insight into the potential success of rainfed irrigation in the region, daily evapotranspiration for each type of crop should be taken into account along with the occurrence and amount of daily rainfall. Moreover, the different planting dates which are being practised by the farmers in the region and the growth stages of crops should also be considered so that the effect of dry spells and number of stress days can be assessed as well. For such purposes, a more comprehensive model as sketched in Fig. 2 is being attempted. Since no simple relationships exist between its different components a simulation study would be appropriate. Such a study normally requires more than one sequence for rainfall and evapotranspiration, which are the main components of the model. Consequently, besides historical data, several sequences of daily rainfall amounts have to be generated.

Consider the water balance analysis for a cultivation field of a crop. On *each day,* a rainfall amount is fed into the field and the evapotranspiration is computed using a suitable formula. This rainfall amount is generated by the aforementioned Markov chain model (here the historical record may simply be treated as a generated sequence). Roughly speaking, if it is small, then it evaporates without penetrating the soil. Accordingly, the *effective rainfall* on that day is set equal to zero. If it is high, a portion of it infiltrates into the crop root zone and a smaller protion may eventually go deeper to become *deep percolation.* Another portion may flow to the surrounding fields, giving rise to the *overflow water* or runoff. The remaining portion is retained in the field or evaporated. That part which is stored in the field or in the root zone will be treated as effective rainfall. If it happens that the soil moisture in the root zone drops below the water holding capacity of the soil in a day due to very light or no rains in several preceding days, then that day is considered as a *stress day.*

Having defined these components, the water balance analysis is carried out on a day-after-day manner so as to account for the carry-over effect. Knowing the deep percolation and the overflow, one can estimate the *drainage requirement.* It should be noted that from the determination of the effective rainfall and the computed value of evapotranspiration on each day, the *supplementary irrigation* water for a specified period can also be calculated. Finally the effect of the stress days on the development and production of the crop may be evaluated by comparing its growth stage which is reflected by the number of days reckoned from the planting date with the occurrence and the length of these stress days.

The water balance analysis is carried out at least for the entire growing season of the crop and the whole analysis is repeated until the desired number of years in the simulation study is reached. Results of this simulation study will soon be reported.

Summary and Conclusions

The distributions of the maximum amounts of daily rainfall and the wet and dry spells were investigated for the Northeast of Thailand. It was found that –

- the annual series of maximum daily rainfall can be fitted by the Type III extreme value and the Gumbel (Type I extreme value) distributions, but the latter gives a better fit; and
- the partial-duration series of maximum daily rainfall exceeding 60 mm can be fitted by the shifted exponential distribution.

However, no suitable distributions among the most widely used ones can satisfactorily represent the wet and dry spells in all months of the year.

By using these findings along with the results obtained in previous studies, a first-order Markov chain model was developed for the generation of amounts of daily rainfall. In order to determine the number of states, and the distribution used for each class interval, four criteria were imposed. These are the reproducibility, for each month of the year, of (i) the distribution of monthly rainfall; (2) the mean, standard deviation, and skewness coefficient of the maximum amounts of daily rainfall; (3) the mean and standard deviation of the number of wet (or dry) days; and (4) the frequencies of the wet and dry spells; all these statistics being referred to the historical record at each station.

Acknowledgements

The authors are indebted to the Royal Irrigation Department of Thailand for having kindly provided the rainfall data needed and to the Asian Institute of Technology (AIT) for its valuable support which made this study possible.

References

AIT (Asian Institute of Technology) (1978) Water for the North-east: Drought Analysis Part I: Rainfall Analysis. Report prepared for the Water Resources Planning Subcommittee of Thailand, Bangkok, 167 pp.

AKAIKE, H. (1974) A new look at the statistical model identification *IEEE Trans. Automatic Control* **AC-19** 716–723.

ALLEN, D.M. and HAAN, C.T. (1975) Stochastic Simulation of Daily Rainfall. Report No. 82, University of Kentucky, Water Resources Inst., 112 pp.

BUISHAND, T.A. (1977) *Stochastic Modelling of Daily Rainfall Sequences.* H. Veenman & B.V. Zonen, Wageningen, 210 pp.

CHIN, E.H. (1977) Modelling daily precipitation occurrence process with Markov chain. *Water Resources Research* **13** (6) 949–956.

GORDON, G. (1978) *System Simulation.* Prentice-Hall, Inc., New Jersey, 2nd. ed, 324 pp.

KHANAL, N.N. and HAMRICH, R.L. (1974) A stochastic model for daily rainfall data synthesis. Proc. Symposium on Statistical Hydrology, U.S. Dept. of Agriculture, 172–196.

KITE, G.W. (1977) Frequency and Risk Analyses in Hydrology. Water Resources Publications, Fort Collins, Colorado, 224 pp.

PHIEN, H.N. and ARBHABHIRAMA, A. (1980) A comparison of statistical tests on the extreme-value type-1 distribution. *Water Resources Bulletin* **16** (5) 856–861.

PHIEN, H.N., ARBHABHIRAMA, A. and SUNCHINDAH, A. (1980a) Rainfall distribution in northeastern Thailand. *Hydrological Sciences Bulletin* **25** (2) 167–182.

PHIEN, H.N., EGGERS, H. and SUNCHINDAH, A. (1980b) An assessment of rainfall irrigation for Northeast Thailand. Proc. Int. Conference on Water Resources Development, Taipei, Taiwan, vol. I, 325–336.

SELVALINGAM, S. and MIURA, M. (1978) Stochastic modelling of monthly and daily rainfall sequences. *Water Resources Bulletin* **14** (5) 1105–1120.

TONG, H. (1975) Determination of the order of a Markov chain by Akaike's information criterion. *Journal of Applied Probability* **12** 488–497.

WARAKITTIMALEE, S. (1980) Characteristics of Monthly and Daily Rainfall Sequences in Northeast Thailand. Master Thesis, Asian Institute of Technology, 142 pp.

YEVJEVICH, V. (1972) Probability and Statistics in Hydrology. Water Resources Publications, Fort Collins, Colorado.