Pertanika 4(1), 1–9 (1981)

A Stochastic Model of Daily Rainfall for Universiti Pertanian Malaysia, Serdang.

M. ZOHADIE BARDAIE and AHMAD CHE ABDUL SALAM[1]
Faculty of Agricultural Engineering, Universiti Pertanian Malaysia, Serdang, Selangor, Malaysia.

Key words: Stochastic model; daily rainfall; simulation

RINGKASAN

Pengunaan proses stokastik untuk mengambarkan dan menganalisa corak hujan harian di Universiti Pertanian Malaysia (U.P.M.), Serdang dihuraikan. Suatu model yang berdasarkan kepada rantai Markov peringkat pertama telah dibentuk. Model tersebut mengunakan data hujan harian untuk memperolehi anggaran bagi kebarangkalian peralihan Markov. Model ini membahagikan tahun kepada empat musim. Tiap-tiap satu musim mempunyai matriks kebarangkalian peralihan yang berasingan. Julat hujan harian dibahagikan kepada sebelas kelas. Oleh itu, tiap-tiap satu musim mempunyai matriks kebarangkalian peralihan yang mengandungi 11 × 11 unsur. Model ini mempunyai keupayaan untuk meramalkan rekod hujan harian sapanjang mana yang diperlukan bagi kawasan tersebut. Penilaian model ini dijalankan dengan mengkaji perbezaan di antara rekod ramalan model dan rekod yang sabenar bagi hujan harian di kawasan tersebut untuk satu tahun.

SUMMARY

An application of stochastic process for describing and analysing daily the rainfall pattern at Universiti Pertanian Malaysia (U.P.M.), Serdang, is presented. A model based on the first-order Markov chain was developed. The model uses historical rainfall data to estimate the Markov transition probabilities. The year is divided into four seasons, each is represented by a separate transition probability matrix. The range of rainfall values is divided into eleven states, thus resulting in a 11 × 11 transition probability matrix for each season. The model is capable of simulating a daily rainfall record of any length for the area. It is evaluated by comparing the simulation result with observed data for a one-year period.

INTRODUCTION

The natural systems are so complex that no exact laws have yet been developed that can explain completely and precisely the natural hydrological phenomena. Before such laws are found complicated hydrological systems such as rainfall can only be approximated essentially by modelling.

The design of a water resources system and the analysis of watershed behaviour often require long records of rainfall for simulation studies. For this purpose historical records may be inadequate and a method must be found to synthesize these records. A rainfall model based on daily precipitation is attractive because relatively long and reliable records can be generated, and such a model is frequently sufficient for many practical problems. Such a model is needed in the evaluation of excessive discharges or draughts, and in the development and management of water resources systems.

Stochastic Models of Rainfall

Statistical methods in hydrology consist of two main groups, probabilistic and the stochastic methods. In the probabilistic method, events are treated as time independent. The stochastic method, in addition to its probabilistic nature, treats the sequence of events as time dependent.

[1] Present address: Agricultural Engineer, Kemubu Agricultural Development Authority, Kota Bharu, Kelantan, Malaysia.

Key to authors' names: M. Zohadie and A. Salam

The Russian mathematician, Markov, introduced the concept of a process (later named after him 'a Markov process') in which the probability of the process being in a given state at a particular time may be deduced from knowledge of the immediately preceding state. A Markov chain is a sequence or chain of discrete states in time for which the probability of transition from one state to any given state in the next step in the chain depends on the condition during the previous step. The Markov chain contains a finite number of states and the probabilities associated with transition between states are stationary with time.

Models of rainfall based on stochastic processes have been developed by various workers– Gabriel and Neuman,(1962;) Green, (1964);Wiser, (1966); Feyerham and Bark, (1967). Todorovic and Woolhiser (1975) presented an n-day rainfall model in which the transition from wet to dry day is based on a 2-state Markov chain, and the amount of rain on rainy days is assumed to be exponentially distributed. Haan *et al.*, (1976) describe a 7-state Markov chain model of daily rainfall in which the amount of rain in each state is assumed to be uniformly distributed, except for the last state in which a shifted exponential distribution is used.

Model for U.P.M., Serdang

The daily rainfall model presented here for U.P.M. Serdang is based on a first-order Markov chain with eleven states. It is an extension of the model presented by Haan *et al.* (1976)., The model can be represented as follows.

Let X_t ($t = 0, 1, 2, \ldots\ldots$) be observation of rainfall for day t. Also, let the range of rainfall observable be divided into $m+1$ classes denoted by $C_0, C_1, C_2, \ldots\ldots\ldots\ldots\ldots\ldots, C_m$. If $P[X_{t+1} = C_j \mid X_0 = C_0, X_1 = C_1 \ldots\ldots, X_t = C_i] = P[X_{t+1} = C_j \mid X_t = C_i]$ (where equals in above means belong to the class), then we have a first-order Markov chain with $m+1$ states. If in addition, within a specified period, which we will call a season, $P[X_{t+1} = C_j \mid X_t = C_i]$ does not depend on t, then the transition probabilities may be denoted by p_{ij} and the Markov chain is said to be stationary. p_{ij} is thus the probability that the rainfall on any day in the season will belong to class C_j given it was in class C_i one day earlier. These transition probabilities may then be gathered together into a $m+1$ transition matrix p_{ij}.

Denoting season by the letter k, then we can represent the stochastic transition probability matrix for the kth. season by $P_{ij}(k)$. It represents the probability that the rainfall on any day in the kth. season will belong to class C_j given that it was in class C_i one day earlier.

Transition Probability Matrices

The model for U.P.M. consists of 4 seasons, each with 11 states. The range of rainfall for each state is shown in Table 1. The four seasons used in the model are as follows.

TABLE 1
States of the Rainfall Model for U.P.M.

State	Range of Rainfall, X (centimeters)
1	$X \leqslant 0.025$
2	$0.025 < X \leqslant 0.127$
3	$0.127 < X \leqslant 0.254$
4	$0.254 < X \leqslant 0.381$
5	$0.381 < X \leqslant 0.508$
6	$0.508 < X \leqslant 0.762$
7	$0.762 < X \leqslant 0.889$
8	$0.889 < X \leqslant 1.524$
9	$1.524 < X \leqslant 2.413$
10	$2.413 < X \leqslant 4.445$
11	$4.445 < X$

1. November to March – the North-East Monsoon period.
2. April – transitional period.
3. May to September – the South-West Monsoon period.
4. October – transitional period.

Daily rainfall records of U.P.M. at Serdang for 1968 to 1978 were used to determine the parameters in the model. The method of maximum likelihood was used for estimating the transition probabilities:

$$p_{ij}(k) = \frac{f_{ij}(k)}{\sum_{j=0}^{m} f_{ij}(k)} \qquad i,j = 0,1,2,\ldots\ldots, m. \qquad k = 1,2,\ldots\ldots\ldots, s.$$

where f_{ij} (k) is the historical frequency of transition from C_i to class C_j within season k, and s is the number of seasons.

The computation of p_{ij} (k) for all i,j and k was carried out in a computer (Abdul Salam, 1980). The transition probability matrix obtained for the 4 seasons are shown in Tables 2, 3, 4, and 5.

TABLE 2
Transition Probability Matrix For North-East Monsoon Season (November to March).

ACTUAL DAY

STATE	1	2	3	4	5	6	7	8	9	10	11
1.	0.6309	0.0593	0.0444	0.0196	0.0210	0.0346	0.0185	0.0605	0.0494	0.0358	0.0260
2.	0.3900	0.1400	0.0700	0.0400	0.0600	0.0500	0.0000	0.1200	0.0500	0.0300	0.0500
3.	0.4545	0.0909	0.0649	0.0519	0.0649	0.0390	0.0130	0.0650	0.0390	0.0779	0.0390
4.	0.3421	0.0526	0.0789	0.0526	0.0131	0.0789	0.0526	0.0263	0.1052	0.0263	0.0789
5.	0.4107	0.0357	0.0892	0.0178	0.1071	0.0714	0.0357	0.0714	0.0357	0.0535	0.0714
6.	0.5238	0.0476	0.0317	0.0158	0.0476	0.0317	0.0634	0.0634	0.0634	0.0952	0.0317
7.	0.3823	0.0294	0.0882	0.0294	0.0588	0.0000	0.0882	0.0882	0.1176	0.0588	0.0588
8.	0.3909	0.0818	0.0636	0.0181	0.0545	0.0454	0.0091	0.1272	0.1181	0.0909	0.0091
9.	0.5465	0.0348	0.0465	0.0465	0.0232	0.0581	0.0232	0.0581	0.0465	0.0697	0.0348
10.	0.4756	0.0487	0.0487	0.0000	0.0243	0.0243	0.0365	0.1341	0.0975	0.0975	0.0121
11.	0.3863	0.1136	0.0454	0.0909	0.0681	0.1136	0.0227	0.0454	0.0227	0.0909	0.0000

PRECEDING DAY

TABLE 3
Transition Probability Matrix For First Transitional Period Season (April).

ACTUAL DAY

STATE	1	2	3	4	5	6	7	8	9	10	11
1.	0.5530	0.0682	0.0530	0.0303	0.0151	0.0227	0.0454	0.0606	0.0606	0.0227	0.0682
2.	0.4782	0.0869	0.0434	0.0000	0.0000	0.0434	0.0000	0.0869	0.0434	0.1739	0.0434
3.	0.2307	0.0769	0.0769	0.0000	0.0000	0.0000	0.1538	0.2307	0.0769	0.1538	0.0000
4.	0.1818	0.2727	0.0000	0.0909	0.0000	0.2727	0.0000	0.0000	0.0909	0.0909	0.0000
5.	0.2500	0.1250	0.0000	0.0000	0.1250	0.1250	0.0000	0.0000	0.1250	0.1250	0.1250
6.	0.3571	0.0714	0.1428	0.0000	0.0714	0.1428	0.0714	0.0714	0.0714	0.0000	0.0000
7.	0.5000	0.0833	0.0000	0.0000	0.0000	0.0833	0.0833	0.0000	0.0833	0.1666	0.0000
8.	0.0869	0.0000	0.0434	0.1304	0.0869	0.0434	0.0000	0.2174	0.1739	0.0869	0.0869
9.	0.3888	0.0555	0.1111	0.1111	0.0000	0.0000	0.0000	0.1111	0.0555	0.1666	0.0000
10.	0.5238	0.0952	0.0476	0.0000	0.0476	0.0476	0.0952	0.0476	0.0000	0.0952	0.0000
11.	0.4666	0.0666	0.0000	0.0666	0.0666	0.0666	0.0000	0.1333	0.0000	0.0000	0.1333

PRECEDING DAY

TABLE 4
Transition Probability Matrix For South-West Monsoon Season (May - September)

ACTUAL DAY

STATE	1	2	3	4	5	6	7	8	9	10	11
1.	0.6789	0.0421	0.0358	0.0189	0.0252	0.0389	0.0126	0.0536	0.0316	0.0421	0.0200
2.	0.5542	0.0602	0.0602	0.0240	0.0240	0.0602	0.0120	0.0723	0.0964	0.0361	0.0000
3.	0.6056	0.0704	0.0423	0.0000	0.0423	0.0563	0.0141	0.0282	0.1127	0.0141	0.0141
4.	0.5833	0.0555	0.0555	0.0555	0.1111	0.0555	0.0277	0.0000	0.0277	0.0277	0.0000
5.	0.4889	0.0667	0.0889	0.0444	0.0889	0.0667	0.0000	0.0000	0.0444	0.0667	0.0444
6.	0.5500	0.0833	0.1167	0.0000	0.0167	0.0167	0.0000	0.0833	0.0333	0.0500	0.0500
7.	0.5714	0.0476	0.0000	0.1428	0.0476	0.0000	0.0476	0.0476	0.0476	0.0476	0.0000
8.	0.5294	0.0706	0.0706	0.0353	0.0235	0.0118	0.0118	0.0941	0.1059	0.0235	0.0235
9.	0.5205	0.0685	0.0685	0.0548	0.0274	0.0411	0.0137	0.0548	0.0548	0.0411	0.0548
10.	0.5312	0.0938	0.0313	0.0156	0.0313	0.0625	0.0156	0.0625	0.0625	0.0781	0.0156
11.	0.3750	0.1563	0.1250	0.0313	0.0000	0.0000	0.0000	0.1563	0.0938	0.0625	0.0000

PRECEDING DAY

TABLE 5
Transition Probability Matrix For Second Transitional Period Season (October)

ACTUAL DAY

STATE	1	2	3	4	5	6	7	8	9	10	11
1.	0.5827	0.0551	0.0472	0.0079	0.0394	0.0630	0.0157	0.0787	0.0630	0.0394	0.0079
2.	0.3000	0.0500	0.0500	0.1000	0.0500	0.0500	0.0000	0.1500	0.0500	0.2000	0.0000
3.	0.4286	0.0714	0.0000	0.0714	0.0000	0.0000	0.0714	0.2143	0.1428	0.0000	0.0000
4.	0.1426	0.0000	0.0000	0.0000	0.0000	0.2857	0.0000	0.4286	0.0000	0.1426	0.0000
5.	0.5000	0.0625	0.0625	0.0000	0.0000	0.0000	0.0000	0.1250	0.1250	0.1250	0.0000
6.	0.3158	0.1053	0.0000	0.0000	0.1580	0.0526	0.0000	0.0526	0.2632	0.0000	0.0526
7.	0.1111	0.1111	0.0000	0.0000	0.2222	0.1111	0.0000	0.2222	0.1111	0.1111	0.0000
8.	0.4828	0.0690	0.0690	0.0000	0.0345	0.0690	0.0690	0.0000	0.0690	0.0690	0.0690
9.	0.2414	0.1034	0.0000	0.0690	0.1034	0.0345	0.0690	0.1034	0.1034	0.1034	0.0690
10.	0.3000	0.0500	0.0500	0.1000	0.0500	0.0000	0.0500	0.0500	0.1500	0.0500	0.1500
11.	0.2000	0.0000	0.1000	0.0000	0.0000	0.3000	0.0000	0.0000	0.3000	0.1000	0.0000

PRECEDING DAY

Probability Distribution for Each Class

Probability distributions were then fitted to each class for all seasons. The assumption is made that the same set of distributions are applicable for each season. The distributions were estimated using the transition matrix of each season. Cumulative distribution functions $F_j(k; x)$, for $j = 0,1,2,\ldots\ldots, m$ were determined, where

$$F_j(k; x) = P\,[\text{next day rainfall} < x;\ \text{when rainfall today belongs to class } C_j \text{ for season } k].$$

Upon analysing the transition matrix, it was determined that exponential density function is the most appropriate. Exponential density function can be represented as follows:

$$f_j(k; x) = \lambda_{jk} \exp(-\lambda_{jk}\, x) \text{ for } x > 0 \qquad \lambda_{jk} > 0$$

where,

$\frac{1}{\lambda_{jk}}$ = expected value of the distribution for class j in season k.

The cumulative function is thus given by

$$F_j(K; x) = 1 - \exp(-\lambda_{jk}\, x)$$

The parameter λ_{jk} is the value which defines the specific exponential distribution representing each class for each season. It has the dimension of inverse rainfall (mm^{-1}), and it is the inverse of the expected value of the rainfall for each class in the particular season. Table 6 shows the values of λ_{jk} for each class in the 4 seasons.

Model Evaluation

Having the transition probability matrices and the probability distribution functions for each class in each season, it is thus possible to generate simulated daily rainfall records for any duration for U.P.M. A daily simulation run of a one-year period beginning with 1st. November 1978 was initiated for the purpose of model evaluation. The simulation process was carried out on a computer (Abdul Salam, 1980), where ten simulations were run for each month.

In evaluating the model, it should be remembered that the purpose of the model is for generating synthetic rainfall records which would have broadly the same properties as historical data. These synthetic rainfall records are necessary for water resources planning studies, since in most situations historical data are not easy to come by especially for long periods. In water resources planning studies, the more critical variables are the broad rainfall pattern such as the annual rainfall variation and the number of wet and dry days. Based on this philosophy, the model for U.P.M. will be evaluated as to whether it has broadly the same properties as that of the observed values.

Annual Rainfall

The results of the ten simulations with regard to the annual rainfall value are presented in Table 7.

Daily Rainfall Evaluation

Table 8 shows the frequency of simulated rainfall with values greater than 5mm, 10mm and 20mm, for the simulated period of one year.

TABLE 6
Values of λjk for Each State in the 4 seasons.

State \ Season	North-East Monsoon (Nov. - March)	First Transitional Period (April)	South-West Monsoon (May - Sept.)	Second Transitional Period (October)
1.	0.19258	0.13046	0.21783	0.20324
2.	0.13232	0.09229	0.20481	0.09475
3.	0.13464	0.09107	0.19664	0.15816
4.	0.13232	0.14084	0.33827	0.06585
5.	0.10212	0.06408	0.14259	0.11897
6.	0.10915	0.22452	0.14568	0.09888
7.	0.12572	0.11637	0.23351	0.09455
8.	0.09810	0.06529	0.16501	0.10869
9.	0.11960	0.11495	0.18676	0.08026
10.	0.12627	0.18834	0.15482	0.06529
11.	0.17306	0.09618	0.15778	0.08778

It demonstrates that the model closely approximated the number of days within the year which have rainfall values within the range shown in the table.

TABLE 7
Annual Rainfall Values from Simulation.

Simulation	Predicted Annual Rainfall (mm)
1	2425.8
2	2470.0
3	2402.3
4	2475.7
5	2464.7
6	2375.3
7	2388.7
8	2342.7
9	2493.2
10	2532.1
Average	2437.1
Standard deviation	59.6
Average Observed Value	2221.1
Standard deviation	297.0

Table 9 gives the magnitude of the largest daily rainfall for the ten simulations along with the average observed value. Here again the model seems to be doing a reasonably good job of simulating the rainfall pattern. However, a slight trend toward underestimating the largest rainfall is indicated.

TABLE 9
Maximum Daily Rainfall

Simulation	Maximum Daily Rainfall (mm)
1	120.5
2	88.9
3	105.3
4	63.5
5	48.7
6	91.0
7	73.5
8	82.3
9	50.2
10	116.5
Average	84.1
Standard deviation	25.1
Average Observed Value	121.5

TABLE 8
Frequency of Simulated Rainfall with Various Range of Values

Simulation	No. of days with rainfall > 5mm	No. of days with rainfall > 10mm	No. of days with rainfall > 20mm
1	103	81	26
2	91	74	31
3	109	63	27
4	103	65	29
5	120	70	32
6	109	61	28
7	93	64	33
8	105	69	31
9	111	72	34
10	95	65	29
Average	103.9	68.4	30.0
Average observed value	92	61	41

CONCLUSION

A rainfall model for Universiti Pertanian Malaysia, Serdang, was developed based on a first-order Markov chain. Simulated rainfall was compared with actual rainfall for a one-year period. It was found that the model is doing a reasonably good job of simulating a rainfall pattern for the area. By applying factors for area and shape, it is possible for this model to be used for generating synthetic watershed precipitation data which can be converted to run off data. It can then be used for soil and water engineering studies in the area.

Stochastic and simulation techniques are becoming increasingly popular in hydrological design and analysis. These techniques are powerful tools which are available to present day engineers. This study is an attempt to use these techniques to represent the rainfall pattern at Universiti Pertanian Malaysia, Serdang, which can hopefully be used in future hydrological studies in the area.

REFERENCES

ABDUL SALAM, A. (1980): A Stochastic Model of Daily Rainfall for Universiti Pertanian Malaysia, Serdang. A project report submitted in partial fulfilment for the Degree of Bachelor of Engineering (Agricultural) at the Faculty of Agricultural Engineering. Universiti Pertanian Malaysia.

FEYERHAM, A.M. and BARK, L.D. (1967): Goodness of fit of a Markov chain model for sequences of wet and dry model for sequences of wet and dry days. *J. Appl. Meteorol.* **6,** 770–773.

GABRIEL, R. and NEUMAN, J. (1962): A Markov chain model for daily rainfall occurrence at Tel Aviv, Israel. *Q. Jl. met. Soc.* **88,** 90–95.

GREEN, J.R. (1964): A model for rainfall occurrence *Jl. R. statist. Soc.* **26,** 345–353.

HAAN, C.T., ALLEN, D.M. and STREET, J.O. (1976): A Markov chain model of daily rainfall. *Water Resour. Res.* **12(3),** 443–449.

THEIL, H. (1961): Economic Forecasts and Policy. North - Holland. Amsterdam.

THEIL, H. (1966): Applied Economic Forecasting. Rand McNally. Chicago. Illinois.

TODOROVIC, P. and WOOLHISER, D. (1975): Stochastic model of n-day precipitation. *J. Appl. Meteorol.* **14,** 17–24.

WISER, E.H. (1966): Monte Carlo methods applied to precipitation frequency analysis. *Trans. Am. Soc. agric. Engrs.* **9,** 538–542.

(Received 12 June 1980)