*Hydrological Sciences – Journal – des Sciences Hydrologiques*, **28**, 2, 6/1983

# Daily precipitation occurrences modelling with Markov chain of seasonal order

MARIA MIMIKOU
*Civil Engineering Department, Technical University of Athens and Public Power Corporation, Athens, Greece*

ABSTRACT The seasonality of the Markov chain order for properly modelling daily precipitation occurrences for precipitation stations in Greece is studied. The average seasonal storm duration, being the average seasonal ratio of the number of wet days to the number of storm events and influenced by the meteorology and geography of the station, is proposed as a criterion for determining the proper seasonal Markov order. For the records studied, and on a monthly basis, the latter, minimizing the chain's modelling error, is the closest positive integer to the average storm duration in days reduced by one. Orders greater than 1 are prevalent in wet periods, whereas in dry periods orders of 1 or 0 predominate. The appearance of the period with orders higher than 1 is delayed and its duration decreases towards the east and with decreasing altitude. The monthly sums of wet days are modelled better by a second order autoregressive model than by aggregating daily precipitation occurrences generated from a Markov chain.

*Mise en modèle de l'occurence des précipitations journalières au moyen d'une chaîne de Markov d'ordre saisonnier*
RESUME On étudie le caractère saisonnier qu'on doit attribuer à l'ordre d'une chaîne de Markov pour mettre en modèle correctement l'occurence des précipitations journalières aux stations pluviomètriques de la Grèce. La durée moyenne saisonnière d'averse, rapport moyen pour la saison du nombre de jours humides au nombre d'averses en tenant compte des conditions météorologiques et géographiques des stations, est proposée comme critère pour déterminer l'ordre saisonnier qui convient pour la chaîne de Markov. Pour les relevés qui ont été étudiés sur une base mensuelle, l'ordre pour lequel l'erreur de mise en modèle de la chaîne est minimale est le nombre entier le plus proche de la durée moyenne d'averse en jours réduite d'une unité. Des ordres supérieurs à 1 sont prédominants en périodes humides et des ordres de 1 ou 0 prédominent en périodes sèches. L'apparition de périodes d'ordre supérieire à un est sujette à un certain retard et leur durée décroît en allant vers l'est et avec des altitudes décroîssants. Les sommes mensuelles de jours humides sont mises en modèle de façon plus appropriée par un modèle auto régressif de second ordre plutôt qu'en

ajoutant les occurences de précipitations journalières générées par une chaîne de Markov.

## INTRODUCTION

Precipitation modelling is very important for planning and management of water resources and has many practical applications in engineering and agriculture. The majority of hydrological methods for precipitation modelling try to represent the generating mechanism of the physical process; they are basically mathematical descriptions of the nature of precipitation and of the structure of its sample time sequence. Since the stochastic nature of precipitation is generally accepted, stochastic models are mainly used to describe the process. Two characteristic features of the process are usually distinguished: precipitation occurrences and precipitation depths or intensities, which are treated either separately or together. The present research work deals only with the precipitation occurrence processes, and, more specifically, with modelling daily precipitation occurrences.

The approaches that have been used in developing stochastic models for daily precipitation occurrences have usually subdivided time into wet periods (storm, wet days) and dry periods (inter-storm period, dry days). Since Gabriel & Neumann (1962) proposed that the sequence of wet and dry days be represented by a Markov chain, the latter has been one of the most widely and commonly used models for representing daily precipitation occurrence processes (Todorovic & Woolhiser, 1976; Haan *et al.*, 1976; Buishand, 1978; Waymire & Gupta, 1981).

In this paper some problems for properly modelling point daily precipitation occurrences with a Markov chain are studied with respect to the seasonality of the order of the chain. The purpose is to improve the chain's efficiency for modelling daily precipitation occurrences.

The seasonality of the Markov order of conditional dependence of daily precipitation occurrences and its dependence on the meteorological conditions and geographical location of the precipitation measuring station have been observed (Chin, 1977). The average seasonal storm duration, which is the average seasonal ratio of the number of wet days to the number of storm events and which is influenced by the precipitation station's meteorology and geography, is proposed in this paper as an explicit and general criterion for determining the seasonal order of conditional dependence of daily precipitation occurrences and hence the proper Markov order for representing it. For seven precipitation stations in northwestern and western Greece the error of modelling daily precipitation occurrences with a Markov chain is minimized with chains of seasonally varying order; their proper monthly Markov order is found to be well related to the average monthly storm duration. The influence of the geographical and topographical location of the precipitation measuring station on Markov order's seasonal behaviour is also studied.

Finally, the Markov chain's efficiency for modelling monthly sums of wet days is checked by comparison with a low order autoregressive stochastic model applied to the cyclically standardized values of the variable.

## MARKOV CHAINS OF SEASONAL ORDER

The seasonality of the order of conditional dependence of daily precipitation occurrences, and hence of the proper Markov order for representing it, has been observed and studied (Chin, 1977). This section discusses the problem of selecting the proper seasonal order of a Markov chain used for modelling daily precipitation occurrences. The search for a physical independent variable to serve as a criterion for explicitly determining the seasonal Markov order and the proper seasonal discretization of time, such that all seasonal changes of the true order of conditional dependence of the physical phenomenon are detected, are the main aspects of the problem. The solutions given in this paper make some use of the observations and conclusions of previous research works (Chin, 1977).

### *Criterion for Markov order determination*

Chin (1977) analysing daily precipitation records for 25 years at more than 100 stations in the conterminous United States observed that "the order of conditional dependence of daily precipitation occurrences depends on the season and geographical location. There exists a prevalence of first-order conditional dependence in summer and higher-order conditional dependence in winter." An explanation based on meteorology is proposed and the conclusion is that the proper Markov order for representing the conditional dependence of daily precipitation occurrences has to be determined either by data analysis or by referring to available results of regional type from previous studies and cannot be assumed *a priori*. Based on these observations, an attempt is made in this paper to define an explicit and general criterion sensitive to seasonal changes, which is also influenced by the location of the precipitation measuring stations, to simplify the determination of the seasonal order of conditional dependence of daily precipitation occurrences and therefore of the Markov order for properly representing it. As such the average storm duration in days over a time period is proposed. A storm event is an uninterrupted sequence of wet (rainy) days and the average storm duration is defined as the ratio of the number of wet days to the number of storm events of a certain time period. Storm duration and frequency are strongly influenced by the meteorological conditions and geographical location of the precipitation measuring station. Chin (1977) found that these also influence the order of conditional dependence of daily precipitation occurrences. For instance, the migratory cyclones especially prevailing in winter months have a broad characteristic length scale and a characteristic lifetime of several days. This means that at any station the precipitation occurrences associated with cyclone passage would most likely last several days (long storm durations) exhibiting a conditional dependence with order higher than 1. In the summer months much of the precipitation comes from local thunder-showers which have a characteristic time of a few minutes (individual cells) to a few hours (cell clusters). This may account for the prevalence of relatively short storm durations and of low (including zero) order of conditional dependence. The location of the station also influences the average storm duration, e.g. the existence of a neighbouring lake would increase the atmospheric moisture and would affect the frequency and duration of the storm events and the conditional dependence of the precipitation occurrences in the vicinity. Therefore, the average over a time period storm duration, as defined here, must be related to the order of conditional dependence of daily precipitation occurrences during the same period.

The time period over which the storm duration is averaged must be selected in a way that all seasonal changes of the prevailing meteorological conditions are detected. Consequently a monthly interval seems to be appropriate. The period of a whole season could also be used in cases where the inter-seasonal changes of the monthly average storm durations are not significant. Additionally, if the location of the station remains the same, one can expect no significant yearly changes in the order of conditional dependence of the physical phenomenon; therefore, the average storm duration for each month or season can be also averaged over the years of the precipitation record to form the average monthly or seasonal storm duration.

*Definitions*

A Markov chain is a sequence of discrete random variables $\{X_i\}_{i \geq 0}$ and is said to be of order r, if for each n > r the conditioned random variables satisfy:

$$X_n | X_{n-1}, \ldots, X_{n-r} = X_n | X_{n-1}, \ldots, X_o \quad (1)$$

For daily sequences $X_n$ is representing a precipitation occurrence at day n and is said to be equal to 1 ($X_n = 1$), if the day is wet (rainfall depth above a certain threshold) and equal to 0 ($X_n = 0$), if the day is dry. In the present work the monthly discretization of time for estimating the average storm duration is adopted. Thus, for a N-year record, p = 1, 2,..., N, the sum of $X_i$ over a month t, t = 1,..., 12, is the sum of wet days of the month:

$$(\Sigma_i X_i)_{p,t} = W_{p,t} \qquad i = 1, 2, \ldots, n, \ldots \quad (2)$$

where p denotes the year of the record. Dividing $W_{p,t}$ by k the number of the individual storms (groups of consequent wet days bounded by dry ones) of the month, one can obtain the average storm duration for the month $d_{p,t}$ as:

$$W_{p,t}/k = d_{p,t} \quad (3)$$

In the cases where a storm extends from one month into the next, it is considered to belong to the month in which the majority of daily precipitation occurrences happened. Averaging $d_{p,t}$ over the N years of the record the average monthly storm duration $d_t$ is obtained as follows:

$$N^{-1} \Sigma_{p=1}^{N} d_{p,t} = d_t \quad (4)$$

Therefore for any daily precipitation record there are 12 (or less if another discretization of time is decided) $d_t$ values estimated from the data of the Markov chain's calibration period.

## MODELLING WITH CHAINS OF PROPERLY SELECTED SEASONAL ORDER

The validity of the average seasonal storm duration as a criterion for determining the seasonal order of conditional dependence of daily precipitation occurrences and hence the proper Markov order for representing it is significantly supported in what follows. In trying to improve the Markov chain's efficiency for modelling daily precipitation occurrences for seven precipitation measuring stations in Greece, it is found that the minimization of the error in modelling is achieved with chains of seasonally varying order and that a certain empirical relationship exists between the latter and the average storm duration on a monthly basis. The limitations imposed by the limited number of stations and area of application of the study and by the decision procedure followed for selecting the proper Markov order do not permit a serious analysis of the sensitivity of the average seasonal storm duration to the chain order and thus of its general validity as a sole criterion for choosing the order. Besides, these limitations imply that additional research work is needed for studying the extension to different climatic regimes, geographical locations, Markov chain applications, etc., of the seasonal relationship between the proposed criterion and the proper Markov order of conditional dependence of daily precipitation occurrences.

### *Markov chain's calibration*

Daily precipitation data from seven stations widely dispersed in northwestern and western Greece were used. Details about the stations (catchment, longitude, latitude, altitude, operation) and the precipitation records (time length) are given in Table 1, and

**Table 1** Precipitation measuring stations

| Station | Catchment | Longitude | Latitude | Altitude (m a.s.l.) | Operated since: | Record length, N (years) |
|---|---|---|---|---|---|---|
| Pades | Aoos | 20°55′ | 40°03′ | 1170 | 1967 | 13 |
| Pramada | Aracthos | 21°06′ | 31°32′ | 817 | 1963 | 17 |
| Mikra Gotista | Arachthos | 21°02′ | 39°41′ | 760 | 1959 | 20 |
| Damaskinia | Aliakmon | 20°11′ | 40°20′ | 990 | 1964 | 16 |
| Anemorachi | Aracthos | 21°05′ | 39°19′ | 390 | 1965 | 14 |
| Chalara | Aliakmon | 21°14′ | 40°39′ | 880 | 1962 | 17 |
| Eani | Aliakmon | 21°49′ | 40°10′ | 480 | 1962 | 18 |

the location of the stations is shown in Fig.1. A day was defined as wet, if the water equivalent of precipitation recorded on that day was at least 0.25 mm; otherwise it was defined as a dry day.

Markov chains of four competing orders, $r = 1...4$, have been calibrated for all seven daily records with time lengths N given in Table 1. All chains were found to be stationary on a monthly basis, which means that the transition probabilities:

$$P(X_n = a_r | X_{n-1} = a_{r-1}, \ldots, X_{n-r} = a_o) = P_{a_o}, \ldots, a_r = P_a \qquad (5)$$

with $a_o,\dots,a_r$ the states of the chain defined here to be either 0 (dry day) or 1 (wet day), were independent of n within month t. The conditional transition probability $P_a$ expresses the ratio of the number of $a_r,a_{r-1},\dots,a_o$ sequences to the number of $a_{r-1},\dots,a_o$ sequences for the time interval into which $P_a$ is estimated. As the chain order r increases the number of the transition probabilities required for the description of all possible daily rainfall sequences is also increasing, i.e. for r = 1 four transition probabilities are needed, for r = 2 eight, etc., whereas the estimation procedure becomes more sensitive to sampling and other data errors. The


**Fig. 1** Location of the precipitation measuring stations.

stationarity was verified by checking, for all months of each precipitation occurrences record, the standard normal variate z (Spiegel, 1975):

$$z = \left[P_a^{(t)} - P_a^{(n)}\right]\left[P_a^{(t)}\ (1 - P_a^{(t)})\ \left(\frac{1}{N_1} + \frac{1}{N_2}\right)\right]^{-\frac{1}{2}} \tag{6}$$

where $P_a^{(t)}$ are the transition probabilities of month t, $P_a^{(n)}$ the transition probabilities of each day within month t for the unified sample of all N years of the record, estimated according to formulae given in the literature (Haan, 1977; Chin, 1977), and $N_1,N_2$ the sizes of the samples with which the two types of transition probabilities were estimated correspondingly. At the 5% significance level the stationarity was accepted for $|z| \leq 1.96$. These chains of order r ranging between 1 and 4 and stationary on a monthly basis were then used to synthesize the daily precipitation occurrences during the calibration periods for all seven stations. The synthesis was done by two procedures. First, the r preceding day n states of the chain were taken to be the actual ones, and second, they were supposed to be unknown and taken as estimated from the previous steps of the generation process with the Markov chain. Data generation from Markov chains requires the knowledge of initial state and the transition probability matrix (Haan, 1977). To estimate the state at day n of a month t with a chain of order r, a random number is selected between 0 and 1; if this random number is less than or equal to the monthly transition probability for having the day n dry, i.e.

$$P^{(t)}_{a_o,\ldots a_{r-1},0}$$

day n is said to be dry; if it is between

$$P^{(t)}_{a_o,\ldots,a_{r-1},0}$$

and 1, day n is said to be wet, for $a_o, a_1, \ldots$ zero or one.

*Selection of the proper monthly Markov order*

The decision criterion for selecting the proper monthly Markov order $r_t$ is the minimization of the mean monthly daily precipitation occurrences modelling error $\varepsilon_t\%$. It is defined as the ratio of the number of failures in estimating the actual sequences of the daily precipitation states (a failure is a wet day estimated as dry and *vice versa*) to the total number of days of month t for the unified sample of all N years of each record. Since the error $\varepsilon_t\%$ shows the efficiency of the model for generating the actual daily rainfall occurrence sequences, it is indicative of the suitability of the chain order to describe the true order of the physical phenomenon and thus, it is significant for the model identification. Thus, for each station the proper monthly Markov order $r_t$ is selected among the other competing orders such that the error $\varepsilon_t\%$ is the minimum. The whole procedure has been implemented in a computer code. The $r_t$ values minimizing the errors for both procedures of data synthesis and for all stations are related to the corresponding average monthly storm durations and are found to be equal to:

$$r_t = D_t - 1 \qquad r_t \geqq 0 \tag{7}$$

where $D_t$ is the closest integer to the average monthly storm duration $d_t$ estimated by (4). In other words, the proper monthly Markov order is the closest positive integer to the average monthly storm duration in days reduced by one. The relationship between $r_t$ and $D_t$ becomes apparent from Table 2. From this table one can see that all Markov orders greater than 1 (in the frame traced inside the table) are concentrated in months between October and April, namely in the wet period, a fact also observed by Chin (1977). The period with orders higher than 1 is not the same for all stations. Its appearance, relative to the beginning of the hydrological year, is delayed and its duration decreases going eastwards and with decreasing altitude. Thus, this period at Pramada begins earlier and lasts longer than the corresponding periods at Mikra Gotista and Anemorachi, even

**Table 2** Relationship between $D_t$ and $r_t$ in days

| Month | Pades: $D_t$ | $r_t$ | Pramada: $D_t$ | $r_t$ | Mikra Gotista: $D_t$ | $r_t$ | Damaskinia: $D_t$ | $r_t$ | Anemorachi: $D_t$ | $r_t$ | Chalara: $D_t$ | $r_t$ | Eani: $D_t$ | $r_t$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Oct. | 3 | 2 | 3 | 2 | 2 | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 2 | 1 |
| Nov. | 3 | 2 | 3 | 2 | 3 | 2 | 3 | 2 | 3 | 2 | 2 | 1 | 2 | 1 |
| Dec. | 3 | 2 | 3 | 2 | 3 | 2 | 3 | 2 | 3 | 2 | 2 | 1 | 2 | 1 |
| Jan. | 3 | 2 | 3 | 2 | 3 | 2 | 3 | 2 | 3 | 2 | 3 | 2 | 2 | 1 |
| Feb. | 3 | 2 | 4 | 3 | 3 | 2 | 3 | 2 | 3 | 2 | 3 | 2 | 2 | 1 |
| March | 3 | 2 | 4 | 3 | 3 | 2 | 3 | 2 | 2 | 1 | 2 | 1 | 2 | 1 |
| April | 3 | 2 | 3 | 2 | 3 | 2 | 3 | 2 | 2 | 1 | 2 | 1 | 2 | 1 |
| May | 2 | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 2 | 1 |
| June | 2 | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 1 | 0 | 2 | 1 | 2 | 1 |
| July | 2 | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 1 | 0 | 1 | 0 | 2 | 1 |
| Aug. | 2 | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 1 | 0 | 1 | 0 | 1 | 0 |
| Sept. | 2 | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 2 | 1 |

though the former station lies east of the others. This is due to the higher altitude of Pramada station compared to the Mikra Gotista and Anemorachi stations. A similar situation is observed between Damaskinia and Anemorachi. It is noteworthy that the Eani station, located east of all other stations and with relatively low altitude, has a Markov order equal to 1 in all months except in August where the order is 0, i.e. in this month the precipitation occurrences exhibit no conditional dependence. The latter is also observed in summer months for another two stations with rather short periods with orders higher than 1, namely for the Anemorachi and Chalara stations.

As an example, the monthly errors of modelling daily precipitation occurrences at Pades station and the improvement of the chain's modelling efficiency by using the proper monthly Markov order are presented herein. In Fig.2 the monthly $\varepsilon_t$% values are plotted for both procedures of data synthesis and for all Markov orders $r = 1,\ldots,4$, kept constant throughout the year. From this figure it becomes apparent that in the wet periods lasting from October to April the chain of order 2 gives the minimum monthly errors, whereas from May to September, this happens with Markov order equal to

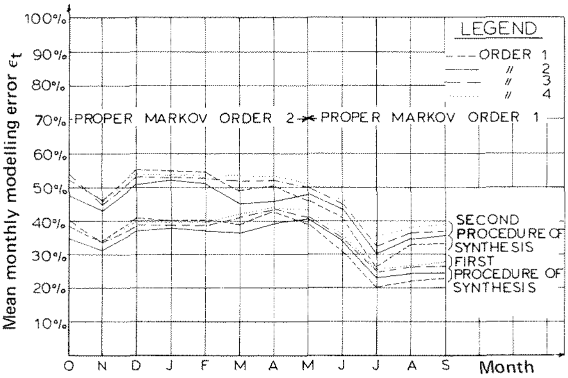


**Fig. 2** Selection of the proper seasonal Markov order: Pades station.

1, for both procedures of synthesis. The average over the former period minimum mean monthly errors are 36% and 48% for the first and the second procedure of synthesis respectively, whereas the corresponding values for the latter period are 27% and 36%. In order to emphasize the improvement of the Markov chain's modelling efficiency by using the proper monthly order, it must be noticed that for the October-April period the average mean monthly errors would be 40% for the first and 51% for the second procedure of synthesis, if the order 1 was used as constant throughout the year, whereas for the May-September period the corresponding error values would be 31% and 39%, if the order 2 was consistently used throughout the year. Similarly, significant improvement of the Markov chain's modelling efficiency by using the proper monthly Markov order has been observed for all studied stations.

*Markov chain's efficiency for modelling monthly wet days sums*

For all seven daily precipitation records under study the monthly sums of wet days $W_{p,t}$ for all years $p = 1, 2, \ldots, N$, with N given in Table 1, have been estimated according to (2). On the other hand, estimates of $W_{p,t}$ were obtained by aggregating the daily precipitation occurrences for each month generated from Markov chains as previously described. A Markov chain's efficiency for modelling $W_{p,t}$ is checked by comparison with another stochastic model fitted on the monthly $W_{p,t}$ series.

The autocorrelation analysis of the $W_{p,t}$ time series has shown that the series are not stationary, because their autocorrelograms were fluctuating beyond the one typical error limit: $\pm 1/\sqrt{(12N)}$. The non parametric cyclic standardization procedure has been used in order to stationarize the series as follows:

$$(W_{p,t} - W_t)/\sigma_t = w_{p,t} \tag{8}$$

where $W_t$ and $\sigma_t$ are the mean and standard deviation values of the monthly series $W_{p,t}$ for month t, estimated as:

$$W_t = N^{-1} \Sigma_{p=1}^{N} W_{p,t} \tag{9}$$

$$\sigma_t = [(N - 1)^{-1} \Sigma_{p=1}^{N} (W_{p,t} - W_t)^2]^{\frac{1}{2}} \tag{10}$$

The autocorrelograms of the stationary series $w_{p,t}$ were found to have a very good resemblance to the autocorrelogram of an AR(2) stochastic process (Box & Jenkins, 1970). Therefore the model for describing the time sequences of the monthly wet days sums for all stations is the following:

$$W_{p,t} = W_t + \sigma_t(\phi_1 w_{p,t-1} + \phi_2 w_{p,t-2} + \xi_{p,t}) \tag{11}$$

where $\phi_1, \phi_2$ are the parameters of the AR(2) model and $\xi_{p,t}$ is a zero mean random residual series with specified statistical characteristics. Based on the skewness test for normality discussed by Hipel *et al.* (1977), the skewness coefficients of the $\xi_{p,t}$ residual series for all seven analysed records were found to be not significantly different from zero at 97.5% confidence level. Therefore, series $\xi_{p,t}$ can be assumed to be normal. Box & Jenkins (1970) give the expression for $\phi_1, \phi_2$ and for the standard deviation $\sigma_\xi$ of $\xi_{p,t}$ as follows:

$$\phi_1 = \rho_1 (1 - \rho_2) (1 - \rho_1^2)^{-1} \tag{12}$$

$$\phi_2 = (\rho_2 - \rho_1^2) (1 - \rho_1^2)^{-1} \tag{13}$$

$$\sigma_\xi = \sigma_w (1 - \rho_1\phi_1 - \rho_2\phi_2)^{\frac{1}{2}} \tag{14}$$

with $\rho_1, \rho_2$ the first and second order autocorrelation coefficients and $\sigma_w$ the standard deviation of the stationary series $w_{p,t}$.

The seven AR(2) models calibrated with all available data of the stations were then implemented in a computer code to estimate the monthly wet days sums during the calibration periods. The modelling error $\alpha$%, defined as:

$$\alpha = \Sigma_{p=1}^{N} \Sigma_{t=1}^{12} (W_{p,t} - \hat{W}_{p,t})^2 / \Sigma_{p=1}^{N} \Sigma_{t=1}^{12} W_{p,t}^2 \ \% \tag{15}$$

with $\hat{W}_{p,t}$ the estimated monthly values, was found to range between 10 and 15% for all records analysed. A comparison was then made for each record between the errors $\alpha$%, of the $\hat{W}_{p,t}$ values estimated by (11) and of the aggregated monthly sequences of the wet days generated from Markov chains of properly selected seasonal order previously presented. The comparison is given in Table 3. From this

**Table 3** Errors in modelling monthly wet days sums

| Station | $\alpha$% Markov chain Procedure of synthesis: 1st | 2nd | AR(2) |
|---|---|---|---|
| Pades | 18 | 23 | 13 |
| Pramada | 16 | 21 | 10 |
| Mikra Gotista | 17 | 23 | 12 |
| Damaskinia | 20 | 24 | 15 |
| Anemorachi | 15 | 20 | 11 |
| Chalara | 20 | 25 | 14 |
| Eani | 19 | 24 | 15 |

table one can estimate that the average overall station error is approximately 13% for the AR(2) model, whereas it is 18 and 23% for the Markov chain with the first and the second procedure of data synthesis, respectively. Therefore, the monthly wet days sums are better modelled by the second order autoregressive model of (11) than by Markov chains.

## CONCLUSIONS

The conclusions drawn from this research are the following:

(a) The proper selection, on a seasonal basis, of the order of a Markov chain significantly improves its efficiency for modelling daily precipitation occurrences.

(b) The average seasonal storm duration, being the average seasonal ratio of the number of wet days to the number of storm events and influenced by the precipitation station's meteorology and geography, is proposed as a criterion for determining the proper seasonal Markov order of a chain for modelling daily precipitation occurrences.

(c) For seven precipitation stations in Greece the error of modelling daily precipitation occurrences with a Markov chain is minimized with chains of seasonally varying order. The proper monthly Markov order is the closest positive integer to the average monthly storm duration in days reduced by one. Additional research work is needed for analysing the sensitivity of the average seasonal storm duration to the order of the chain and for studying the extension to different climatic regimes, geographical locations, Markov chain applications, etc., of the seasonal relationship between the proposed criterion and the proper Markov order of conditional dependence of daily precipitation occurrences.

(d) There is a prevalence of Markov orders higher than 1 in wet periods and of orders 1 or 0 in dry periods. The appearance at each station of the period with orders higher than 1, relative to the beginning of the hydrological year, is delayed and its duration decreases going eastwards and with decreasing altitude.

(e) The monthly sums of wet days are better modelled by a second order autoregressive model than by aggregating daily precipitation occurrences generated from a Markov chain.

*ACKNOWLEDGEMENTS* The author wishes to thank the Environmental Research Centre of the Greek Public Power Corporation for providing the daily precipitation data.

## REFERENCES

Box, G.P. & Jenkins, G.M. (1970) *Time Series Analysis - Forecasting and Control*. Holden-Day, San Francisco.

Buishand, T.A. (1978) Some remarks on the use of daily rainfall models. *J. Hydrol.* 36, 295-308.

Chin, E.H. (1977) Modelling daily precipitation occurrence process with Markov chain. *Wat. Resour. Res.* 13 (6), 949-956.

Gabriel, K.R. & Neumann, J. (1962) A Markov chain model for daily rainfall occurrences at Tel Aviv. *Quart. J. Roy. Met. Soc.* 88, 90-95.

Haan, C.T. (1977) *Statistical Methods in Hydrology*. Iowa State Univ. Press, Ames, Iowa.

Haan, C.T., Allen, D.M. & Street, J.O. (1976) A Markov chain model for daily rainfall. *Wat. Resour. Res.* 12 (3), 443-449.

Hipel, K.W., McLeod, A.I. & Lennox, W.C. (1977) Advances in Box-

Jenkins modelling, 1. Model construction. *Wat. Resour. Res.* 13 (3), 567-576.

Spiegel, M.R. (1975) *Theory and Problems of Probability and Statistics*. Shaums Outline Series, McGraw-Hill, New York.

Todorovic, P. & Woolhiser, D.A. (1976) Stochastic structure of the local pattern of precipitation. In: *Stochastic Approaches to Water Resources*, vol. II (ed. by H.W.Shen), 15.1-15.37. Water Resources Publications, Fort Collins, Colorado.

Waymire, E. & Gupta, V.K. (1981) The mathematical structure of rainfall representations, 1. A review of the stochastic rainfall models. *Wat. Resour. Res.* 17 (5), 1261-1272.

*Received 14 September 1982; revised 24 November 1982.*