WATER RESOURCES RESEARCH, VOL. 23, NO. 3, PAGES 515–518, MARCH 1987

Interpolation of Binary Series Based on Discrete-Time Markov Chain Models

EFI FOUFOULA-GEORGIOU

Department of Civil Engineering, Iowa State University, Ames

TRYPHON T. GEORGIOU

Department of Electrical and Computer Engineering, Iowa State University, Ames

We consider the problem of interpolating missing observations in a time series modeled by a discrete-time Markov chain. The general interpolation scheme involves a finite enumeration of all possible paths (i.e., admissible values for the missing data) and computation of the probability distribution of the paths. Procedures for the selection of a particular path are discussed in terms of a prespecified interpolation objective. In the special case of two-state Markov chains, we investigate an efficient way of enumerating the paths based on the set of sufficient statistics. An example using daily rainfall occurrence series is presented.

1. INTRODUCTION

The hydrologist often faces the problem of incomplete records which must be filled in before the data can be used for operational hydrologic analyses. For instance, streamflow simulation by routing effective rainfall through a conceptual rainfall-runoff model usually requires uninterrupted rainfall records during the simulation period. In the case of nonintermittent time series such as monthly or annual series, one may use parametric models such as autoregressive moving average (ARMA) models to fill in the missing data [cf. *Damsleth*, 1980; *Anděl*, 1979]. However, the identification and fitting of such models to incomplete series may require a nonroutine autocorrelation and spectral analysis [cf. *Marshall*, 1980; *Bloomfield*, 1970].

In this note we study the interpolation of missing data in a time series modeled by a discrete-time Markov chain. Our analysis addresses the case of a two-state Markov chain (MC) or order 1. Higher orders can be approached in a similar fashion even though the computational issues become significantly more involved. Moreover, any finite-state MC of order greater than one can always be brought into the form of a MC of order 1 by appropriately augmenting the set of states. The general interpolation scheme is based on the enumeration of all possible paths, i.e., admissible sequences for the missing values, and the computation of the probability distribution of the paths. A particular path is then selected in accordance with a prespecified interpolation objective.

The objectives of an interpolation scheme can vary depending on the particular application. In simulation studies where preservation of the historical statistical characteristics is desired, a valid approach would be to draw the missing sequences at random from their estimated probability distribution. However, quite often the local features of the record are of central importance. In such cases, estimates of the missing data are usually chosen on the basis of their likelihood of occurrence. Certain negative aspects to that approach are evident when one considers the overall effect this has on the statistics of the record. Alternatively, an interpolation scheme can be based on a prespecified "statistical characteristic" of the gap which is considered of primary interest. For instance, in case the number of occurrences of an event is of interest and not the precise order of events, the very number of occurrences will form the statistical characteristic of the gap to be preserved. A set of sufficient statistics [cf. *Kedem*, 1980] would be another choice. In all cases, we are led to determine the estimates of the missing data on the basis of the likelihood of their statistical characteristic. In our developments we will focus on the set of sufficient statistics because any other statistic "factors" through a set of sufficient statistics and the development would be similar. Apart from the locally optimal nature of this choice, it appears that this scheme has in many cases minimal effect of the global statistical features of the interpolated record.

The general interpolation scheme is presented in section 2, while in section 3 the special case of two-state Markov chains is discussed in detail. Two-state (binary) Markov chains are of particular interest to hydrologic applications. The two states can be thought of corresponding to "success" and "failure," respectively. For instance, presence or absence of measurable rain during a day, exceedence-nonexceedence of daily base flow, and detection-nondetection of a threshold pollutant concentration during a sampling interval are a few examples of binary (zero-one) series. It should be noted that a binary process may arise either naturally or by clipping of a continuous process. In the second case, some information is usually lost by throwing away the continuous data, but this might be unavoidable or even desirable, depending on the nature of the data and the intended use of the outcome. In any case, by making use of the sufficient statistics of binary Markov chains the computational efficiency of our interpolation scheme can be enhanced by enumerating only the number of distinct sets of sufficient statistics instead of the number of all possible paths. An example involving interpolation of daily rainfall occurrence series is presented in section 4 to illustrate the procedure.

2. GENERAL INTERPOLATION SCHEME

Let a random sequence X_k, $k \in Z$ be an n-state stationary Markov chain of order m; i.e.,

$$P\{X_k = x_k | X_{k-1} = x_{k-1}, \cdots\} = P\{X_k = x_k | X_{k-1} = x_{k-1}, \cdots, X_{k-m} = x_{k-m}\}$$

where x_k takes on values from the set of states $E_n = (e_1, \cdots, e_n)$. Consider now the sequence $\{Y_k\}$ formed out of m tuples of the sequence $\{X_k\}$; i.e.,

$$Y_k = (X_k, X_{k-1}, \cdots, X_{k-m+1})$$

Paper number 6W4514.
0043-1397/87/006W-4514$02.00

One can easily show that the random sequence $\{Y_k\}$ is an N-state Markov chain of order 1 where $N = n^m$. Thus with no loss in generality we may begin with Y_k, $k \in Z$, an N-state stationary Markov chain of order 1, taking on values $y_k \in F = (E_n)^m$. Let

$$\mathscr{X}_{(l)} = \{\cdots, y_{-1}, y_0; y_{l+1}, y_{l+2}, \cdots\}$$

be a realization of Y_k with a gap of length l. The function $P\{Y_1 = y_1, \cdots, Y_l = y_l | \mathscr{X}_{(l)}\}$ can be though of as a conditional likelihood function and will be denoted by $L(y_1, \cdots, y_l; \mathscr{X}_{(l)})$. Values of $L(y_1, \cdots, y_l; \mathscr{X}_{(l)})$ can be computed for all possible paths based on estimates for

$$p_i = P\{Y_k = f_i\} \qquad (1a)$$

$$p_{ij} = P\{Y_k = f_j | Y_{k-1} = f_i\} \qquad (1b)$$

which can be obtained from the available observations. Also, if Y_k is derived from an n state, mth order Markov chain X_k, as above, the $N \times N$ transition probability matrix

$$\mathbf{P} = [p_{ij}]_{i,j=1}{}^N$$

will be sparse having at most n nonzero entries per column.

Due to the fact that Y_k is a Markov chain it follows that (for a proof see the appendix)

$$P\{Y_1 = f_{i_1}, \cdots, Y_l = f_{i_l} | \mathscr{X}_{(l)}\} = P\{Y_1 = f_{i_1}, \cdots, Y_l = f_{i_l} | Y_0 = f_\alpha, Y_{l+1} = f_\beta\} \qquad (2)$$

Also

$$P\{Y_0 = f_\alpha, \cdots, Y_{l+1} = f_\beta\} = p_\alpha p_{\alpha i_1} p_{i_1 i_2} \cdots p_{i_l \beta}$$

$$P\{Y_0 = f_\alpha, Y_{l+1} = f_\beta\} = p_\alpha \sum_{j_k \in F} p_{\alpha j_1} p_{j_1 j_2} \cdots p_{j_l \beta}$$

Hence the conditional likelihood of the path $(f_{i_1}, f_{i_2}, \cdots, f_{i_l})$ is

$$L(f_{i_1}, \cdots, f_{i_l}; f_\alpha, f_\beta) = p_{\alpha i_1} \cdots p_{i_l \beta} \Big/ \sum_{j_k \in F} p_{\alpha j_1} p_{j_1 j_2} \cdots p_{j_l \beta} \qquad (3)$$

We wish to classify all admissible paths into path classes according to prespecified statistical characteristics. Ideally, we would like to group them in path classes corresponding to different values of a set of sufficient statistics [cf. *Kedem*, 1963]. It is not known, however, what a set of sufficient statistics is for a general N state Markov chain of order 1. Thus one can possibly employ other statistics, e.g., the number of occurrences of an event, and then approach the problem directly based on (3). That is, after forming path classes, determine their respective conditional probabilities by summing up the probabilities of their elements each given by (3).

It is clear that paths corresponding to the same values of sufficient statistics would have the same likelihood of occurrence. Even though the converse is not true, in general, it suggests an interesting alternative, i.e., to classify path classes according to the likelihood of occurrence of their elements. Thus one could group paths corresponding to the same likelihood of occurrence given by (3) into classes and proceed thereafter to choose first the class with the highest likelihood of occurrence and then a path out of this class drawn at random.

In several interesting cases, as, for instance, binary Markov chains of order equal to or greater than one, a set of sufficient statistics is known [*Whittle*, 1955; *Kedem*, 1980]. In such a case, the likelihood of a path can be determined directly as a function of the sufficient statistics. In the next section we discuss the case of binary Markov chains of order 1.

3. Binary Markov Chains: An Application

Let X_k, $k = 1, 2, \cdots, n$ be a binary Markov chain of order 1, where X_k takes values in $\{0, 1\}$. One hydrologic application of such a process may be the modeling of sequences of wet and dry days. Consider the marginal and conditional probabilities

$$p_1 = P\{X_i = 1\} \qquad (4a)$$

$$p_{11} = P\{X_i = 1 | X_{i-1} = 1\} \qquad (4b)$$

The transition probability matrix is given by

$$\mathbf{P} = \begin{bmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{bmatrix} = \begin{bmatrix} (1 - 2p_1 + p_1 p_{11})/q_1 & (1 - p_{11})p_1/q_1 \\ 1 - p_{11} & p_{11} \end{bmatrix} \qquad (5)$$

where p_{ij} was defined in (1b), and $q_1 = 1 - p_1$. (Note that p_1 and p_{11} must obey the inequality $\max(0, (2p_1 - 1)/p_1) < p_{11} < 1$.)

Now, define the quantities

$$S = \sum_{i=1}^{n} X_i \qquad R_1 = \sum_{i=2}^{n} X_i X_{i-1} \qquad H = X_1 + X_n \qquad (6)$$

Note that S is the number of 1's, R_1 the number of transitions from 1 to 1, and H the "end points" condition of the binary series of length n. The joint distribution $P\{X_1 = x_1, \cdots, X_n = x_n\}$ is given by

$$P\{X_1 = x_1, \cdots, X_n = x_n\} = p_1{}^{s-r_1}(1 - p_1)^{s-h-n+2} p_{11}^{r_1} \cdot (1 - p_{11})^{2s-2r_1-h}(1 - 2p_1 + p_1 p_{11})^{-2s+r_1+h+n-1} \qquad (7)$$

[cf. *Kedem*, 1980, p. 11], where $h \in \{0, 1, 2\}$, $r_1 \in \{0, 1, \cdots, s-1\}$, and $s \in \{h, h+1, \cdots, n\}$. It is interesting to note [*Klotz*, 1973] that (S, R_1, H) is a sufficient statistic for the pair (p_1, p_{11}) under the standing hypothesis that the process is generated by a two-state Markov chain of order 1. That is, all the information (with regard to the actual values of p_1 and p_{11}) that can be extracted from the finite record $(x_1, x_2, \cdots, x_n)$ is contained in the derived statistics (S, R_1, H). The joint distribution of the sufficient statistics (S, R_1, H) is

$$P\{S = s, R_1 = r_1, H = h\} = M_n(s, r_1, h) P\{X_1 = x_1, \cdots, X_n = x_n\} \qquad (8)$$

where $M_n(s, r_1, h)$ is the number of binary sequences having the same (s, r_1, h), and $(x_1, \cdots, x_n)$ is any such particular realization. *Klotz* [1973] gives

$$M_n(s, r_1, h) = \binom{2}{h}\binom{n-s-1}{s-r_1-h}\binom{s-1}{r_1} \qquad (9)$$

with the convention $\binom{-1}{1} = 1$. For second-order Markov chains, similar but more involved results apply [*Kedem*, 1963, 1977], and extensions to higher-order Markov chains can be obtained [*Whittle*, 1955; *Kedem*, 1980].

The above results can now be used to handle the interpolation of binary Markov chains. Let $(y_1, \cdots, y_l)$ be a gap of length l in a series described by a binary Markov chain (BMC) of order 1 and let $y_0 = f_\alpha$, $y_{l+1} = f_\beta$ be the known end points of the gap. Using the above results one can easily see that instead of enumerating all the 2^l possible sequences $(y_1, \cdots, y_l)$ it suffices to enumerate all the possible pairs (s, r_1) given the end point condition $h = f_\alpha + f_\beta$. The number of these pairs, which specifies the number of path classes to be enumerated, depends quadratically (and not exponentially) on the gap length. Using combinatorial arguments, these numbers have been computed for all possible configurations of gap lengths

TABLE 1. Number of Path Classes as a Function of the Gap Length l and the Known End Points Condition h

h	l	Number of Path Classes
2, or 0	even	$(l/2)(l/2 + 1) + 1$
	odd	$([l/2] + 1)([l/2] + 1) + 1$
1	even	$(l/2 + 1)(l/2 + 1)$
	odd	$([l/2] + 1)([l/2] + 2)$

Each path class includes all possible paths with the same likelihood of occurrence.

and end points conditions, and the results are given in Table 1. Notice that in formulae (6)–(9) n is now equal to $l + 2$.

4. Daily Rainfall Occurrences: An Example

We consider the daily rainfall occurrence series during the months April through June at Cedar Falls, Washington, and for the period 1948–1977. In computing the statistics of the series, the "end effects" at the beginning and end of the season have been smoothed out by continuing in the next season until the whole wet period has been considered. For this series we estimated the transition probabilities $p_{00} = 0.667$ and $p_{11} = 0.685$, which result in $p_1 = (1 - p_{00})/(2 - p_{11} - p_{00}) = 0.514$.

Suppose now that there is a gap of 10 values in the series and that the gap starts with a dry day and ends with a wet day. Instead of enumerating the $2^{10} = 1024$ possible paths and computing their conditional likelihood from (3), we only have to enumerate the 36 possible pairs (s, r_1) given $h = 1$. Each of these 36 (s, r_1) pairs defines a path class. The number of paths in each path class conditioned at the known end points is $M_n(s, r_1|h) = M_n(s, r_1, h)/\binom{n}{2}$, and the conditional probability of the path class is $p(s, r_1|h) = M_n(s, r_1|h)p(s, r_1, h)$. The results of these computations are given in Table 2. From that table we also observe that paths corresponding to distinct (s, r_1) pairs have different probabilities of occurrence and thus the admissible values of (s, r_1) characterize different path classes.

In general, interpolation can be based on the probability distribution of the paths. If the results of interpolation are to be used for simulation purposes, then a path can be drawn at random from its probability distribution. On the other hand, if a predictive scheme based on likelihood of occurrence is utilized, the path most likely to have occurred is the one with $s = 11$ and $r_1 = 10$. This corresponds to the trivial path of having all values equal to 1, as was expected intuitively from the values of the conditional probabilities. Clearly, such a choice would alter significantly the statistics of the record. Alternatively, the selection can be based upon the likelihood of a certain statistic. In all cases, the results can be directly computed from Table 2. In particular, on the basis of the likelihood of the sufficient statistic pair (s, r_1), the optimal choice is a path drawn at random from the path class with $s = 6$ and $r_1 = 4$. Any such path has the most likely to have occurred number of wet days and number of transitions from wet to wet day. Furthermore, in the case that only the number of wet days during the missing period is of importance, it can be verified from Table 2 that the optimal path would be a path drawn at random from the probability distribution of the four path classes inside the larger class described by $s = 8$. (This is because the corresponding statistic is now s, and larger classes corresponding to constant s values can be formed and the one with the maximum likelihood selected.)

Note that in cases where interpolation is based on the likelihood of a statistic, the choice amounts to preserving the mode of the associated probability distribution, whereas drawing a path at random amounts to preserving the mean. The above example was merely presented for illustration purposes. It suggests, however, that several interpolation alternatives exist which apart from being locally optimal do not significantly alter the global characteristics of the record. Simulation studies are needed to establish the properties of the different interpolation alternatives.

5. Concluding Remarks

In this note we have suggested and exploited an interpolation scheme for filling in gaps in hydrologic time series modeled by Markov chains. Since an N-state higher-order Markov chain can always be brought into the form of a multistate first-order MC by appropriately augmenting the states, the general ideas of the interpolation scheme have been presented in terms of an N-state first-order MC. The special case of a two-state MC of order 1 has been analyzed in detail in terms of the sufficient statistics. Finally, an example involving daily rainfall occurrences has been presented. Depending on the objective of interpolation, alternative optimality criteria have been discussed, and their implementation has been illustrated.

Although we have concentrated on discrete-valued stochastic processes described by Markov chains, it is clear that the suggested interpolation procedure can also be applied to real-valued Markov processes after an appropriate set of representative states has been selected [cf. *Yakowitz*, 1979].

TABLE 2. Example of Filling in a Ten-Value Gap of a Rainfall Occurrence Series

Class	(s, r_1)	$M_n(s, r_1\|h)$	$p(s, r_1\|h)$
1	(1, 0)	1	0.002816
2	(2, 0)	9	0.005976
3	(2, 1)	1	0.002892
4	(3, 0)	28	0.004396
5	(3, 1)	16	0.010928
6	(3, 2)	1	0.002971
7	(4, 0)	35	0.001295
8	(4, 1)	63	0.010143
9	(4, 2)	21	0.014721
10	(4, 3)	1	0.003052
11	(5, 0)	15	0.000135
12	(5, 1)	80	0.003040
13	(5, 2)	90	0.014850
14	(5, 3)	24	0.017280
15	(5, 4)	1	0.003135
16	(6, 0)	1	0.000002
17	(6, 1)	25	0.000225
18	(6, 2)	100	0.003900
19	(6, 3)	100	0.017000
20	(6, 4)	25	0.018500
21	(6, 5)	1	0.003220
22	(7, 2)	15	0.000135
23	(7, 3)	80	0.003200
24	(7, 4)	90	0.015750
25	(7, 5)	24	0.018240
26	(7, 6)	1	0.003308
27	(8, 4)	35	0.014350
28	(8, 5)	63	0.011277
29	(8, 6)	21	0.016401
30	(8, 7)	1	0.003398
31	(9, 6)	28	0.005152
32	(9, 7)	16	0.012832
33	(9, 8)	1	0.003490
34	(10, 8)	9	0.007416
35	(10, 9)	1	0.003585
36	(11, 10)	1	0.003683

Number of paths in each path class and conditional likelihood of occurrence of all possible path classes are also given.

APPENDIX: PROOF OF STATEMENT (2)

Let $\{Y_k: k \in Z\}$ be a discrete-time, N-state, Markov chain or order 1. We tacitly assume that $\{Y_k\}$ is irreducible and ergodic so that no state or group of states is absorbing, and the unconditional probabilities of the various states are nonzero. In this appendix we will prove the statement that was employed in (2).

Lemma

Under the above hypothesis, $\{\hat{Y}_k = Y_{-k}, k \in Z\}$ is a discrete-time N-state Markov chain of order 1, with state transition matrix given by

$$\hat{\mathbf{P}} = [\hat{p}_{ij}]_{i,j=1}{}^N = \left[p_{ji}\frac{p_j}{p_i}\right]_{i,j=1}^N \tag{A1}$$

where

$$\mathbf{P} = [p_{ij}]_{i,j=1}^N$$

is the state transition matrix of $\{Y_k, k \in Z\}$, and the p_i's are the unconditional probabilities defined in (1*a*).

Proof of the Lemma

We first show that Y_k is time reversible; that is, $\hat{Y}_k$ is an N-state Markov chain of order 1. It suffices to show that

$$P\{Y_0 = \alpha | Y_1 = \beta, Y_2 = \gamma, \cdots\} = P\{Y_0 = \alpha | Y_1 = \beta\}$$

Using Bayes' theorem we obtain that

$$P\{Y_0=\alpha|Y_1=\beta,\cdots\} = \frac{P\{Y_1=\beta,\cdots|Y_0=\alpha\}P\{Y_0=\alpha\}}{P\{Y_1=\beta,\cdots\}}$$

$$= \frac{P\{Y_2=\gamma,\cdots|Y_0=\alpha, Y_1=\beta\}P\{Y_1=\beta|Y_0=\alpha\}P\{Y_0=\alpha\}}{P\{Y_2=\gamma,\cdots|Y_1=\beta\}P\{Y_1=\beta\}}$$

$$= \frac{P\{Y_1=\beta|Y_0=\alpha\}P\{Y_0=\alpha\}}{P\{Y_1=\beta\}} = P\{Y_0=\alpha|Y_1=\beta\}$$

where the step before the final one follows from the Markovian property of $\{Y_k\}$. We now show the validity of (A1). It follows once again from Bayes' theorem that

$$\hat{p}_{ij} = P\{\hat{Y}_{k+1} = f_j | \hat{Y}_k = f_i\}$$

$$= P\{Y_{-(k+1)} = f_j | Y_{-k} = f_i\}$$

$$= P\{Y_{l-1} = f_j | Y_l = f_i\} \qquad \text{where } l := -k$$

$$= \frac{P\{Y_l = f_i | Y_{l-1} = f_j\}P\{Y_{l-1} = f_j\}}{P\{Y_l = f_i\}}$$

$$= p_{ji}\frac{p_j}{p_i}$$

This completes the proof of the lemma. We now proceed with the proof of the following proposition.

Proposition

$$P\{Y_1 = f_{i_1}, \cdots, Y_l = f_{i_l} | \mathscr{Z}_{(l)}\}$$
$$= P\{Y_1 = f_{i_1}, \cdots, Y_l = f_{i_l} | Y_0 = f_\alpha, Y_{l+1} = f_\beta\} \tag{A2}$$

Proof of Proposition

Let $\mathscr{P} = (\cdots, Y_{-1} = f_{\alpha_1}, Y_0 = f_\alpha)$ and $\mathscr{F} = (Y_{l+1} = f_\beta, Y_{l+2} = f_{\beta_1}, \cdots)$ denote the past and future observations, respectively, and $G = (Y_1 = f_{i_1}, \cdots, Y_l = f_{i_l})$ denote the gap in the record. Then,

$$P\{G|\mathscr{P}, \mathscr{F}\} = \frac{P\{G, \mathscr{F}|\mathscr{P}\}}{P\{\mathscr{F}|\mathscr{P}\}}$$

$$= \frac{P\{G, \mathscr{F}|Y_0 = \alpha\}}{P\{\mathscr{F}|Y_0 = \alpha\}}$$

$$= P\{G|Y_0 = \alpha, \mathscr{F}\}$$

as it follows from the Markovian property of $\{Y_k\}$. But

$$P\{G|Y_0 = \alpha, \mathscr{F}\} = \frac{P\{G, Y_0 = \alpha|\mathscr{F}\}}{P\{Y_0 = \alpha|\mathscr{F}\}}$$

$$= \frac{P\{G, Y_0 = \alpha|Y_{l+1} = \beta\}}{P\{Y_0 = \alpha|Y_{l+1} = \beta\}}$$

where the last step follows from the Markovian property of $\{\hat{Y}_k = Y_{-k}\}$. So finally,

$$P\{G|\mathscr{P}, \mathscr{F}\} = P\{G|Y_0 = \alpha, Y_{l+1} = \beta\}$$

and the proof of the proposition is completed.

Remark

It is interesting to note that in the case of a two-state Markov chain $\{Y_k\}$ the corresponding "time-reverse" Markov chain $\{\hat{Y}_k\}$ has the same state transition matrix:

$$\hat{p}_{ii} = p_{ii} \qquad i = 0, 1$$

$$\hat{p}_{01} = p_{10}\frac{p_1}{p_0} = (1 - p_{11})\frac{(1 - p_{00})/(2 - p_{00} - p_{11})}{(1 - p_{11})/(2 - p_{00} - p_{11})}$$

$$= 1 - p_{00} = p_{01}$$

and, similarly,

$$\hat{p}_{10} = p_{10}$$

For Markov chains with more than two states this is not true and, in general, $\hat{\mathbf{P}}$ is given in terms of $\mathbf{P}$ from (A1).

Acknowledgment. We thank the referees for their helpful comments and suggestions.

REFERENCES

Anděl, J., On interpolation of multiple autoregressive processes, in *Contributions to Statistics*, edited by J. Tureckova, D. Reidel, Hingham, Mass., 1979.

Bloomfield, P., Spectral analysis with randomly missing observations, *J. R. Stat. Soc., Ser. B*, *32*, 369–380, 1970.

Damsleth, E., Interpolating missing values in a time series, *Scand. J. Stat.*, *7*, 33–39, 1980.

Kedem, B., Sufficient statistics associated with a two-state second order Markov chain, *Biometrika*, *63*(1), 127–132, 1963.

Kedem, B., Counting the number of 0-1 stationary time series having the same likelihood, *Discrete Math.*, *18*, 285–289, 1977.

Kedem, B., *Binary Time Series, Lecture Notes in Pure and Applied Mathematics*, vol. 52, Marcel Dekker, New York, 1980.

Klotz, J., Statistical inference in Bernoulli trials with dependence, *Ann. Stat.*, *1*, 372–373, 1973.

Marshall, R. J., Autocorrelation estimation of time series with randomly missing observations, *Biometrika*, *67*(3), 567–570, 1980.

Whittle, P., Some distribution and moment formulae for Markov chains, *J. R. Stat. Soc., Ser. B*, *17*, 235–242, 1955.

Yakowitz, S., A nonparametric Markov model for daily river flow, *Water Resour. Res.*, *15*(5), 1035–1043, 1979.

E. Foufoula-Georgiou, Department of Civil and Construction Engineering, Iowa State University, Ames, IA 50011.

T. T. Georgiou, Department of Electrical and Computer Engineering, Iowa State University, Ames, IA 50011.

(Received April 23, 1986;
revised October 31, 1986;
accepted November 11, 1986.)