JOURNAL OF THE AMERICAN WATER RESOURCES ASSOCIATION
VOL. 37, NO. 4 AMERICAN WATER RESOURCES ASSOCIATION AUGUST 2001

SEASONAL ARIMA INFLOW MODELS FOR RESERVOIR SIZING[1]

Dimitris M. Papamichail and Pantazis E. Georgiou[2]

ABSTRACT: The reliable sizing of reservoirs is a very important task of hydraulic engineering. Although many reservoirs throughout the world have been designed using Rippl's mass curves with historical inflow volumes at the dam site, this technique is now considered outdated. In this paper, synthetic series of monthly inflows are used as an alternative to historical inflow records. These synthetic series are generated from stochastic SARIMA (Seasonal Autoregressive Integrated Moving Average) models. The analyzed data refer to the planned Almopeos Reservoir on the Almopeos River in Northern Greece with 19-year monthly inflow series. The analysis of this study demonstrates the ability of SARIMA models, in conjunction with the adequate transformation, to forecast monthly inflows of one or more months ahead and generate synthetic series of monthly inflows that preserve the key statistics of the historical monthly inflows and their persistence Hurst coefficient K. The forecasted monthly inflows would be of help in evaluating the optimal real time reservoir operation policies and the generated synthetic series of monthly inflows can be used to provide a probabilistic framework for reservoir design and to cope with the situation where the design horizon of interest exceeds the length of the historical inflow record.
(KEY WORDS: seasonal ARIMA model; generating; forecasting; persistence; reservoir sizing.)

INTRODUCTION

Reservoir capacity-yield procedures can be classified theoretically into three main groups (McMahon and Mein, 1986) although the distinction between groups is not always clear-cut. The first group (critical period techniques) includes methods in which a sequence of flows for which demand exceeds inflows is used to determine the storage size. The second group of procedures (Probability Matrix Methods) consists of those methods which are based on Moran's Dam Theory (Moran, 1954). Finally, the third group consists of those procedures which are based on inflow data generated by stochastic models. The most widely used method for reservoir sizing is the Rippl's method, which belongs to the first group and is based on the range analysis of the historical inflow volumes at the dam site (Rippl, 1883). The method assumes implicitly that the historical reservoir inflows will be repeated in the future during the operation life of a reservoir, an assumption which is obviously incorrect. The estimation of a unique storage volume that does not allow any risk assessment and its dependency on the length of historical inflow time series are other important drawbacks of the conventional deterministic method. To solve the problems created by the deterministic conventional approach for reservoir sizing, which is based on the mass curve procedure (Rippl Diagram) with the historical monthly inflow time series at the dam site, a stochastic simulation technique was applied. This technique uses stochastically generated inflows and allows the estimation of a probability distribution function of the storage volumes, relating the latter to their risks.

In this paper, the synthetic monthly inflows into a planned Reservoir (Almopeos river basin) in Northern Greece were generated by using a multiplicative SARIMA (Seasonal Autoregressive Integrated Moving Average) model based on the discrete time series of historical monthly inflow values of Almopeos river for the last 19 years. It is increasingly recognized that the stochastic models, known as SARIMA models are of considerable practical use in dealing with generating of natural flows (Box and Jenkins, 1976; Salas *et al.*, 1985; Lungu and Sefe, 1991; Bender and Simonovic, 1994; Hipel and McLeod, 1994). SARIMA

[1]Paper No. 00021 of the *Journal of the American Water Resources Association.* **Discussions are open until April 1, 2002.**
[2]Respectively, Associate Professor and Research Hydrologist, Department of Hydraulics, Soil Science and Agricultural Engineering, School of Agriculture, Aristotle University of Thessaloniki, 54006 Thessaloniki, Greece (E-Mail/Papamichail: papamich@agro.auth.gr).

models became very popular because of their simple mathematical structure, ideal representation of the statistical and the correlation structure of the corresponding observed inflows, convenient representation of data in terms of a relatively small number of parameters and their applicability to stationary as well as nonstationary processes. These models are also well suited to the analysis of time series that are by nature persistent, which is true in the case of flow or inflow time series (Hipel and Mcleod, 1994). Persistence is a long-term statistical property of an inflow series according to which high inflow periods tend to follow other high inflow periods and this cluster phenomenon is likewise behaved with the low inflow periods. When sizing reservoirs, by using stochastically generated inflows, the rescaled adjusted range (RAR), or equivalently the Hurst coefficient K (Hurst, 1951, 1957; Salas and Boes, 1974; Anis and Lloyd, 1976; McLeod and Hipel, 1978; Hipel and McLeod, 1978; Salas *et al.,* 1979; Klemes *et al.,* 1981; Mimikou and Nalbandis, 1987; Vogel and Stedinger, 1988; Salas, 1993), expressing their persistence, must be preserved by the inflow generating model. The analysis of this study demonstrated the ability of SARIMA models, in conjunction with the adequate transformation, to forecast monthly inflows of one or more months ahead and generate synthetic series of monthly inflows which preserve the key statistics of the historical monthly inflows and the Hurst coefficient K. The forecasted monthly inflows would be of help in evaluating the optimal real time reservoir operation policies and the stochastically generated monthly inflows can be used to ensure the reliable sizing of the reservoir.

METHODOLOGY

SARIMA Models

A discrete time series $Z_1, Z_2, Z_3, ..., Z_{N-1}, Z_N$ of measurements at equal time intervals can be simulated by a stochastic SARIMA model (Box and Jenkins, 1976; Salas *et al.,* 1985; Brockwell and Davis, 1987; Montgomery *et al.,* 1990; Lungu and Sefe, 1991; Bender and Simonovic, 1994; Hipel and McLeod, 1994) of the form

$$\varphi(B)\Phi(B^S)(1-B)^d(1-B^S)^D Z_t = \theta(B)\Theta(B^S)e_t \qquad (1)$$

where t is the discrete time; S is the length of seasons; B is the backward shift operator defined by $BZ_t = Z_{t-1}$ and $B^S Z_t = Z_{t-S}$; $e_t = [NID(0,\sigma_e^2)]$ is the normally independently distributed white noise residual with mean zero and variance σ_e^2; $\varphi(B) = 1 - \varphi_1 B - \varphi_2 B^2 - ... - \varphi_p B^p$ is the nonseasonal autoregressive (AR) operator of order p and φ_i, i = 1,2, ..., p are the nonseasonal AR parameters; $(1 - B)^d$ is the nonseasonal differencing operator of order d to produce nonseasonal stationarity of the dth differenced data, usually d = 0, 1, or 2; $\Phi(B^S) = 1 - \Phi_1 B^S - \Phi_2 B^{2S} - ... - \Phi_P B^{PS}$ is the seasonal autoregressive (AR) operator of order P and Φ_i; i = 1,2, ... , P are the seasonal AR parameters; $(1-B^S)^D$ is the seasonal differencing operator of order D to produce seasonal stationarity of the Dth differenced data, usually D = 0, 1, or 2; $\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - ... - \theta_q B^q$ is the nonseasonal moving average (MA) operator of order q and θ_i, i =1, 2, ..., q are the nonseasonal MA parameters; and $\Theta(B^S) = 1 - \Theta_1 B^S - \Theta_2 B^{2S} - ... - \Theta_Q B^{QS}$ is the seasonal moving average (MA) operator of order Q and Θ_i, i = 1, 2, ..., Q are the seasonal MA parameters.

The notation $(p,d,q)(P,D,Q)_S$ is used to represent the SARIMA model. In this paper, the time series of monthly inflows is simulated by a SARIMA model. The application of the SARIMA models requires stationarity of time series data obtained by different transformations. Models in hydrology frequently use the natural logarithm of the measurements rather than working with the historical untransformed data. The logarithmic transformation is chosen to stabilize the variance of the series and to transform the usually skewed inflow distribution into a normal distribution.

Although in practice the most popular transformation is the logarithmic, other transformations such as the family of one parameter transformations as introduced by Box and Cox (1964), is considered in this paper and computed by

$$\frac{Z_t^{(\lambda)} - 1}{\lambda g^{\lambda-1}} \qquad , \lambda \neq 0$$

$$Z_t^{(\lambda)} = g \ln Z_t \qquad , \lambda = 0 \qquad (2)$$

where Z_t is the historical untransformed time series; $Z_t^{(\lambda)}$ is the historical transformed time series; λ is the parameter transformation; t is the discrete time; and g denotes the sample geometric mean. The transformation in Equation (2) is valid for $Z_t > 0$. Since there are no months of zero inflow we use the family of one parameter transformations.

Estimates for the parameters in

$$\varphi(B)\Phi(B^S)(1-B)^d(1-B^S)^D Z_t^{(\lambda)} = \theta(B)\Theta(B^S)e_t \qquad (3)$$

have to be derived. For a trial series of λ values and corresponding $Z_t^{(\lambda)}$ time series, the estimates for the parameters in SARIMA model in Equation (3) can be derived and the particular white-noise sequence e_t and the variance $\sigma^2(e_t,\lambda)$ of each time series are computed respectively, by

$$e_t = \theta^{-1}(B)\Theta^{-1}(B^S)\varphi(B)\Phi(B^S)(1-B)^d(1-B^S)^D Z_t^{(\lambda)} \quad (4)$$

$$\sigma^2(e_t,\lambda) = \sum_{t=1}^{N} e_t^2 / N \quad (5)$$

The maximized log likelihood is, except for a constant, given by

$$L_{max}(\lambda) \propto -\left(\frac{N}{2}\right)\log \sigma^2(e_t,\lambda) \quad (6)$$

where N is the record length.

From the plot of the maximized likelihood function $L_{max}(\lambda)$ for a trial series of λ values the maximizing value λ can be read off. If the maximizing value λ is equal to zero the application of the SARIMA models requires stationarity of time series obtained by logarithmic transformation (Equation 2).

The construction of SARIMA models involves various stages (Box and Jenkins, 1976; Hipel and McLeod, 1994), which are the identification, the estimation, and the diagnostic checking. The purpose of the identification stage is to determine the differencing required to produce stationarity and to estimate the order of both the seasonal and nonseasonal AR and MA operators for the stationary series. The identification is examined by the cumulative periodogram, the autocorrelation function and the partial autocorrelation function. The estimation stage involves the estimation of the time series model parameters. This estimation is obtained by the residuals squares minimization using the Marquardt algorithm (Kuester and Mize, 1973). Finally, the diagnostic checking involves examination of the residuals fitted model and this can or cannot prove the model inadequancy, and it can also inform about the model improvement. This determination can be achieved by using the identification stage tests and, furthermore, the Akaike's information criterion (AIC) (Akaike, 1974), the Kashyap's posterior probability (PP) criterion (Kashyap, 1977) and the Residual variance (RV) criterion (Kareliotis and Chow, 1992). Cline (1981) derived the AIC and PP criteria for logarithmic transformed series and they have been used in this study. The selection rules in the AIC, PP, and RV criteria are to select the model with the lowest AIC, PP, and RV values, respectively. The best model chosen is used for forecasting and generating synthetic monthly inflows into a reservoir.

Persistence

In modeling hydrologic time series for simulation studies of reservoir systems, storage-related statistics may be particularly important. Consider the hydrologic time series Z_t, t = 1,..., N and a subsample Z_1, ..., Z_n with $n \leq N$ the partial sums and the adjusted partial sums are defined, respectively, as (Salas *et al.*, 1979; Salas, 1993)

$$S_i = S_{i-1} + Z_i \quad (7)$$

$$S_i^* = S_{i-1}^* + \left(Z_i - \overline{Z_n}\right) \quad (8)$$

where $S_0 = S_0^* = 0$ and $\overline{Z_n}$ is the sample mean. The crude range R_n, the adjusted range R_n^* and the rescaled adjusted range R_n^{**} are defined by (Salas *et al.*, 1979; Salas, 1993)

$$R_n = \max(0, S_1, S_2, ..., S_n) - \min(0, S_1, S_2, ..., S_n) \quad (9)$$

$$R_n^* = \max\left(0,S_1^*,S_2^*,...,S_n^*\right) - \min\left(0,S_1^*,S_2^*,...,S_n^*\right) \quad (10)$$

$$R_n^{**} = R_n^* / s_n \quad (11)$$

in which s_n is the sample standard deviation. Both R_n^* and $\overline{R_n^{**}}$ have been widely used in the literature as measures of long term dependence and for comparing alternative models of hydrologic series (McLeod and Hipel, 1978; Hipel and McLeod, 1978).

In particular, Hurst (1951, 1957) showed that for a large number of geophysical time series such as stream-flow, precipitation, temperature, and tree-ring series, the mean rescaled adjusted range R_n^{**} is proportional to n^h with $h > 0.5$. The discrepancy between theoretical results stating that $h = 0.5$ and Hurst's empirical finding that $h > 0.5$ has become known as the Hurst phenomenon. Several estimators of h have been proposed and used in stochastic hydrology such as the original Hurst estimator K (Hurst, 1951, 1957) which is considered in this paper and defined as

$$K = \frac{\log\left(\overline{R_n^{**}}\right)}{\log(n/2)} \quad (12)$$

Hurst estimator K was found to range from 0.46 to 0.96 with a mean of 0.73 and a standard deviation of 0.09 (McLeod and Hipel, 1978; Hipel and McLeod, 1978; Salas, 1993).

Reservoir Sizing by Using Historical and Stochastically Generated Inflows

One of the earliest methods for estimating the size of storage to meet a given draft is the Rippl's method which is based on the estimation of the maximum range of the inflow time series at the dam site (Rippl, 1883). The level of development a is defined as follows

$$a = \sum_{i=1}^{N} Q_i \Delta t \,/\, \sum_{i=1}^{N} Z_i \Delta t \tag{13}$$

where Δt is a time interval taken in this paper equal to one year; Q_i is the annual draft or reservoir release; Z_i is the annual inflow; and N is taken equal to the number of years of the inflow time series. For each inflow time series the maximum range R of the cumulative inflow monthly volumes x_t has been estimated as follows

$$R = \max_{S \le t \le T} \Delta(t) - \min_{S \le t \le T} \Delta(t) \tag{14}$$

$$\Delta(t) = \left[x_{S+t} - x_S - \frac{1}{T}\left(x_{S+T} - x_S\right) \right] \tag{15}$$

in which the subscript S denotes time at the beginning of the inflow time series, and T is the time at the end of the inflow time series. The maximum range R of cumulative departures of the inflow volumes from the draft time series for a sequence of inflows covering N years is defined to be the storage capacity V required to meet the draft requirements over N years. Thus, for the n stochastically generated inflow time series the values of storage volume V_i, i = 1,2,...,n, have been estimated. The cumulative probability distribution function of these volumes $F_{V_i}(v)$ has been then estimated according to the usual statistical procedure. This distribution relates the storage volumes to their risks ε%. Risk ε% is defined to be the probability of not meeting the design storage capacity (the reservoir will be empty or will not meet the required water yield), as follows

$$F_{V_i}(v)(\%) = 100 - \varepsilon\% \,. \tag{16}$$

RESULTS AND DISCUSSION

In the present study, the SARIMA model is constructed to simulate the discrete time series of historical monthly data of inflows into a planned reservoir located on the Almopeos river (22°09′, 40°46′N), in Northern Greece. The drainage area of the river basin at the dam site is equal to 1021 km^2. The available historical monthly inflows at the dam site cover a 19-year period (1975 to 1993) and are shown in Figure 1.

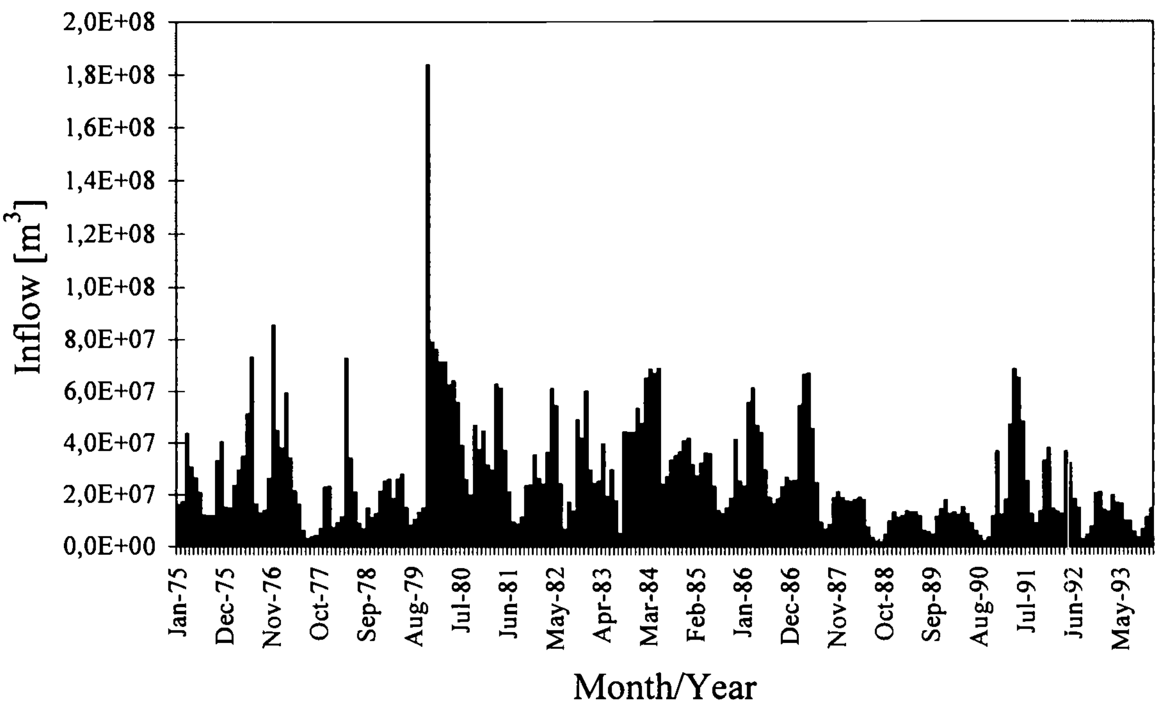

Figure 1. Monthly Inflows Into Almopeos Reservoir.

The logarithmic transformation was chosen to stabilize the variance, as recommended by Ledolter (1978), for this type of series. From the cumulative periodogram, the autocorrelation function and the partial autocorrelation function of the logarithmic transformed time series of monthly inflows (1975 to 1993) into Almopeos Reservoir (Figure 2), the model with the structure (2,0,0) $(0,1,1)_{12}$ was found to be a candidate model. To allow for possible identification errors, a set of nine models (Table 1) with the structures close to that of (2,0,0) $(0,1,1)_{12}$ were considered. The parameters of each of the model structure were estimated using the Marquardt algorithm and the AIC, PP and RV criteria were applied to select the best model. The AIC, PP and RV values for all the models are also listed in Table 1. It can be seen that the model (2,0,0) $(0,1,1)_{12}$ has the lowest AIC, PP and RV values. These indicate that the (2,0,0) $(0,1,1)_{12}$ model is suitable. The estimates of the parameters of the (2,0,0) $(0,1,1)_{12}$ model together with their confidence limits, standard errors, T-values and P-values are given in Table 2.

TABLE 1. Akaike Criterion (AIC), Posterior Probability Criterion (PP), and Residual Variance (RV) for the Different Models of the Logarithmic Transformed Time Series of Monthly Inflows (1975 to 1993) Into Almopeos Reservoir.

Model Structure	AIC	PP	RV
(0,1,1) $(0,1,2)_{12}$	7353.656	7406.80	0.246407
(0,1,1) $(0,1,3)_{12}$	7357.50	7428.38	0.246249
(0,1,1) $(1,1,1)_{12}$	7351.50	7395.79	0.246246
(1,1,1) $(1,1,1)_{12}$	7350.15	7412.16	0.240534
(1,1,2) $(1,1,1)_{12}$	7349.03	7428.76	0.235195
(1,0,1) $(0,1,1)_{12}$	7337.37	7381.67	0.23346
(2,0,0) $(0,1,1)_{12}$	**7337.35**	**7372.78**	**0.23145**
(3,0,0) $(0,1,1)_{12}$	7338.21	7382.50	0.232299
(1,0,0) $(0,1,2)_{12}$	7343.42	7378.86	0.239769
(0,1,1) $(0,1,1)_{12}$	7349.33	7384.77	0.24606

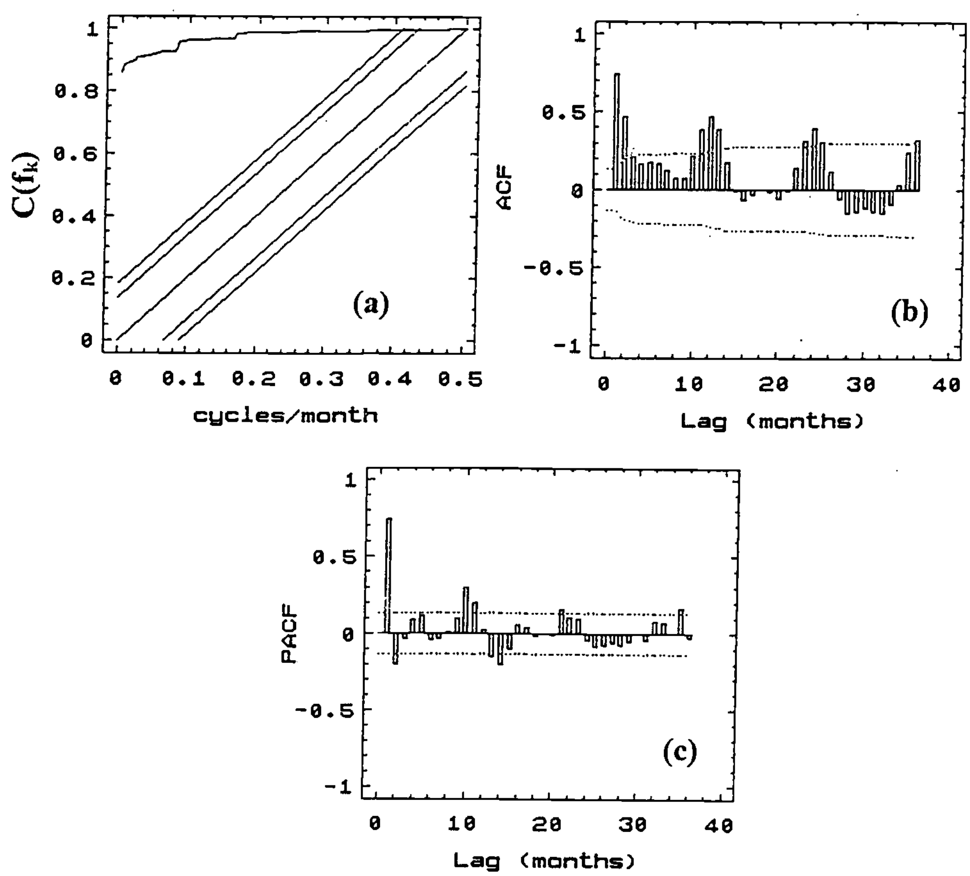

Figure 2. (a) Cumulative Periodogram [$C(f_k)$], (b) Autocorrelation Function [ACF], and (c) Partial Autocorrelation Function [PACF] of the Logarithmic Transformed Time Series of Monthly Inflows (1975 to 1993) Into Almopeos Reservoir.

TABLE 2. Parameter Estimates and Their 95 Percent Confidence Limits, Standard Errors, T-Values, and P-Values for the (2,0,0) $(0,1,1)_{12}$ of the Logarithmic Transformed Time Series of Monthly Inflows (1975 to 1993) Into Almopeos Reservoir.

Parameter	Estimate	95 Percent Confidence Limits	Standard Error	T-Value	P-Value
φ_1	0.60211	(0.46883,0.73539)	0.06762	8.90489	0.00000
φ_2	0.16415	(0.03057,0.29773)	0.06777	2.42234	0.01626
Θ_1	0.76570	(0.67359,0.85781)	0.04673	16.38408	0.00000

The candidate model has to be validated for its suitability through the diagnostic checking before making a final selection. The tools used for diagnostic checking are: (1) overfitting and (2) check for randomness of residuals. The check for overfitting is to examine the confidence limits of the parameters. The hypothesis is that if the confidence limits include the value zero, then the corresponding parameter needs to be discarded, and if they include 1, then replacement of the associated term by a difference operator (1-B) is necessary. For the selected model, the confidence limits of none of the parameters include either 0 or 1 (Table 2), indicating there is no need for a revision of the model. The other hypothesis for checking is that if the model is adequate then the residuals should be white-noise. From the cumulative periodogram, the autocorrelation function and the partial autocorrelation function of the residuals for the (2,0,0) $(0,1,1)_{12}$ model, it was found that all values are well within the 95 percent confidence limits confirming the residuals are white noise.

For a trial series of λ values and corresponding $Z_t^{(\lambda)}$ time series, the estimates for the parameters in (2,0,0) $(0,1,1)_{12}$ model are derived and the particular white-noise sequence e_t (Equation 4), the variance $\sigma^2(e_t,\lambda)$ (Equation 5), and the maximum likelihood estimates L_{max} (λ) (Equation 6) of each time series are computed. The plot of L_{max} (λ) for the trial series of λ values is given in Figure 3 and it can be seen that the maximizing value λ is equal to zero. This confirms that the application of the (2,0,0)$(0,1,1)_{12}$ model requires stationarity of time series obtained by logarithmic transformation.

Using the identified multiplicative SARIMA model (2,0,0) $(0,1,1)_{12}$, the logarithmic monthly inflow Z_t can be forecasted or generated by the equation

$$Z_t = \varphi_1 Z_{t-1} + \varphi_2 Z_{t-2} + Z_{t-12} - \varphi_1 Z_{t-13} - \varphi_2 Z_{t-14} + e_t - \Theta_1 e_{t-12} \qquad (17)$$

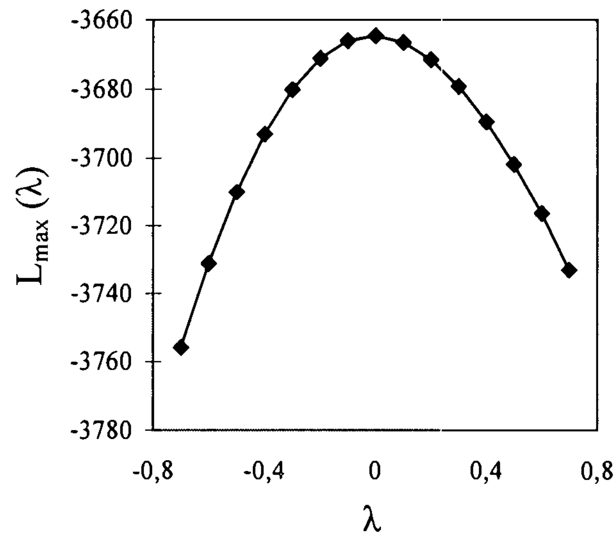

Figure 3. Plot of the Function $L_{max}(\lambda)$ for the (2,0,0) $(0,1,1)_{12}$ Model of the $Z_t^{(\lambda)}$ Time Series of Monthly Inflows (1975 to 1993) Into Almopeos Reservoir.

We investigated the use of Equation (17) for forecasting monthly inflows of one month ahead for the final two years (1992 and 1993) of data and compare these forecasts with inflows that actually occurred. As we choose to use the final two years (1992 and 1993) of data for comparing measured inflows with forecasted inflows, we reestimate the parameters for the (2,0,0) $(0,1,1)_{12}$ model for the time series of monthly inflows shortened by one month each time. The new estimates of parameters for each time series of monthly inflows shortened by one month differ only slightly from estimates for the full records (Table 2). In Figure 4 the one month ahead forecasts of monthly inflows for the final 24 months (1992 and 1993) are shown. This figure clearly demonstrates the ability of the SARIMA model of Equation (17) to forecast well the final 24 months measured values of inflows.

Also, we investigated the use of Equation (17) for generating synthetic monthly inflows into Almopeos Reservoir for 50 years. Using the identified multiplicative SARIMA model (2,0,0) $(0,1,1)_{12}$ of Equation

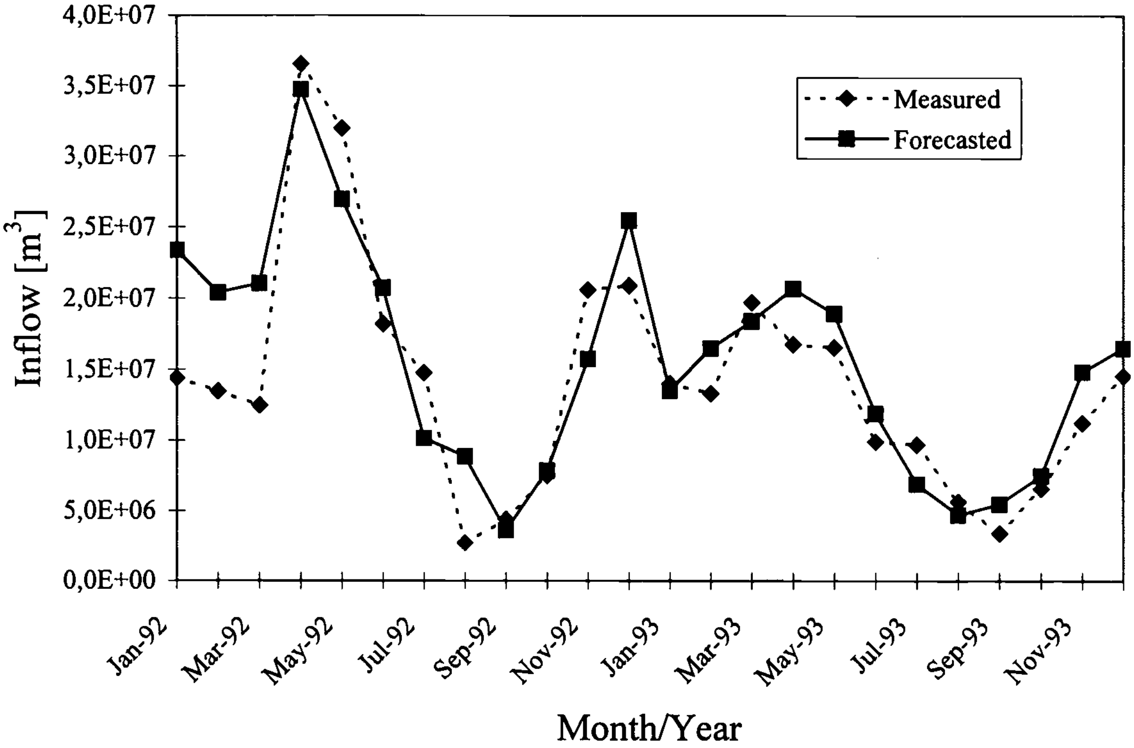

Figure 4. Comparison Between Measured and One Month Ahead Forecasted Values of Monthly Inflows (1992 to 1993) Into Almopeos Reservoir.

(17) with the parameter estimates of Table 2 and the normally independently distributed white-noise residual e_t with mean zero and variance σ_e^2 = 0.23346, the synthetic monthly inflows for 50 years have been obtained. Figure 5 shows the comparison between the monthly means and the monthly standard deviations of the measured inflows (1975 to 1993) into Almopeos Reservoir with that of 100 generated time series (50 years each) of synthetic monthly inflows. It can be seen that the monthly means and the monthly standard deviations of the generated synthetic inflows are close to that of the measured inflows. This confirms the fact that the selected SARIMA model of Equation (17) is suitable for generating synthetic inflows into Almopeos Reservoir.

When sizing reservoirs, by using stochastically generated inflows, the Hurst estimator K, expressing their persistence, must be preserved by the inflow generating model. The Hurst estimator K is estimated for the Almopeos Reservoir historical yearly inflow series of 19-year length and its value is K = 0.8313. The Hurst estimator K for each synthetic yearly inflow time series of 50-year length is estimated according to the persistence procedure described and considering n = 10. The values of K were found to range from 0.6146 to 0.8093, with a mean of 0.7328 and standard deviation of 0.0356. The R_n^{**}, $\overline{R_n^{**}}$, and K values as averages of all subsequences with n = 10 of 100 replicates of the yearly generated inflow time series are given in Table 3. The difference between the estimated K value of the generated yearly inflow series of 50-year length (0.7328) and the estimated K value of the historical yearly inflow series of 19-year length (0.8313) is due to the sample size, as one interpretation of the Hurst phenomenon (Salas *et al.*, 1979) accepts that series exhibiting this phenomenon have K greater than 0.5 for small or moderate values of sample size and their K converges rapidly to 0.5 as sample size $\rightarrow\infty$. Both K values (0.7328 and 0.8313) estimated of the generated and the historical yearly inflow series, respectively, are within the Hurst limits who has found K to range from 0.46 to 0.96 (McLeod and Hipel, 1978). This confirms the preservation of Hurst estimator K and proves that the selected SARIMA model of Equation (17) is suitable for generating synthetic inflows into Almopeos Reservoir.

The annual draft requirement for the Almopeos Reservoir is estimated and taken constant and equal to $150*10^6$ m^3/year which corresponds to a level of development a = 0.46 (46 percent) (Equation 13). Thus, Equation (15) indicates the surplus or deficit of water into the reservoir at time t. The maximum range R (Equation 14) of cumulative departures of the inflow volumes from the draft time series for a

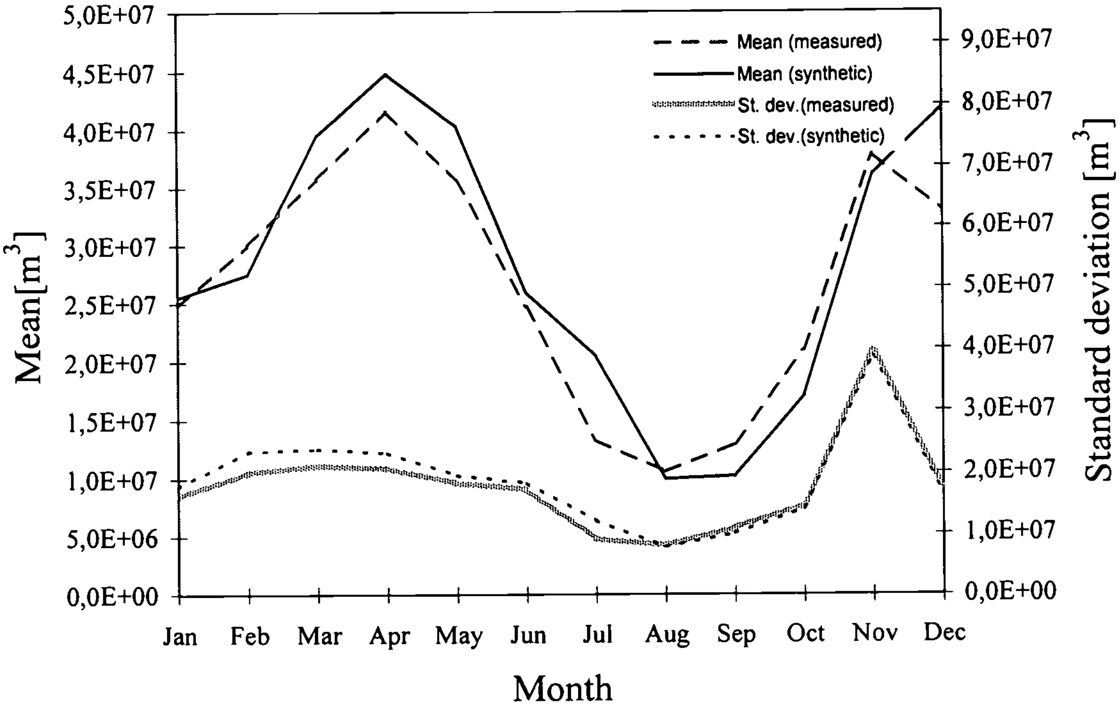

Figure 5. Comparison Between the Monthly Means and the Monthly Standard Deviations of the Measured Inflows (1975 to 1993) Into Almopeos Reservoir With That as Averages of 100 Replicates of the Monthly Generated Inflow Time Series (50 years each).

sequence of monthly values covering 50 years is defined to be the storage capacity V required to meet the annual draft requirement of $150*10^6$ m^3 over 50 years. Thus, 100 values of storage volumes V_i, i =1,...,100, have been estimated. The cumulative probability distribution function of these volumes $F_{V_i}(v)$ has been then estimated according to the log-normal distribution which best describes the sample data of the 100 storage volumes. This cumulative distribution obtained for the Almopeos Reservoir is shown in Figure 6 and relates a range of storage volumes reasonable for the Almopeos Reservoir to the $F_{V_i}(v)$. Thus, instead of a unique volume which is obtained with the Rippl's method, where the maximum range is estimated for the historical monthly inflows sequence only, a probability distribution function of storages which allows the risk assessment and decision making in the design (sizing with a pre-calculated design risk) is obtained. For example, by accepting a risk of 25 percent such that 75 percent of time in 50 years the storage is sufficient to produce the required draft of $150*10^6 m^3$/year, from Figure 6 the design capacity is $180*10^6 m^3$ which is a reasonable storage value for the Almopeos Reservoir.

TABLE 3. Estimation of the Hurst Estimator K as Average of 100 Replicates of the Yearly Generated Inflow Time Series Into Almopeos Reservoir.

n	R_n^{**}	$\overline{R_n^{**}}$	K
10	3.3388		
	3.2804		
	3.3685		
	3.4661		
	2.8079		
	Average	**3.25234**	**0.7328**

CONCLUSIONS

The analysis in this study demonstrates the ability of multiplicative SARIMA models, in conjunction with the adequate transformation, to simulate the discrete time series of historical monthly inflows into a reservoir. These models can forecast monthly inflows of one or more months ahead which would be of help in evaluating the optimal real time reservoir operation policies, in effective utilization of available water, especially in a multipurpose context and in deriving the optimal cropping pattern, especially during the shortage periods. These models can also generate sequences of monthly synthetic inflows into a reservoir of unlimited duration. These stochastically generated inflows preserve the key statistics of the historical monthly inflows and the Hurst estimator K, expressing their persistence, which is true in the case of inflow time series. The stochastic reservoir sizing, which is based on the generated series of monthly inflows can be used as an alternative to the historical inflow records, as it provides a probabilistic framework for decision making in the design (sizing with a pre-calculated design risk) and allows the coping with the situation where the design horizon of interest exceeds the length of the historical inflow record.

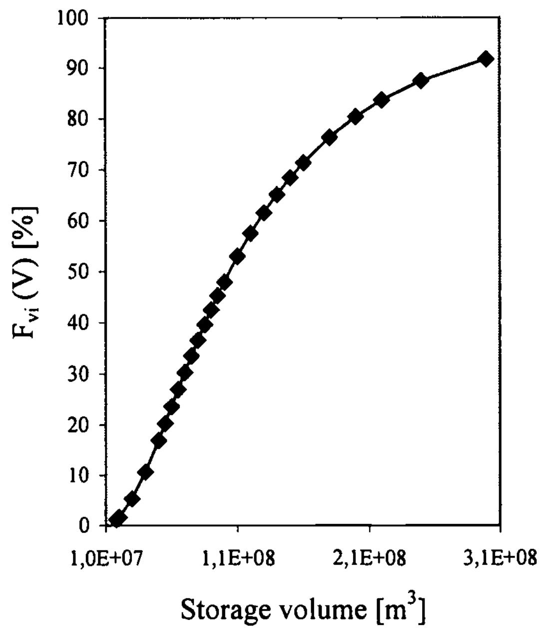

Figure 6. Cumulative Probability Distribution Storage Volumes for the Almopeos Reservoir.

ACKNOWLEDGMENTS

The authors are grateful to the Ministry of Agriculture for the financial support received through the Research Project on Water Balance for Watersheds in Northern Greece which is under the direction of Professor George Terzidis.

LITERATURE CITED

Anis, A. A. and E. H. Lloyd, 1976. The Expected Values of the Adjusted Rescaled Hurst Range of Independent Normal Summands. Biometrika 63:111-116.

Akaike, H., 1974. A New Look at the Statistical Model Identification. IEEE Trans. Autom. Control AC-19(6):716-723.

Bender, M. and S. Simonovic, 1994. Time-Series Modeling for Long-Range Stream-Flow Forecasting. Journal of Water Resources Planning and Management 120(6):857-870.

Box, G. E. P. and D. R. Cox, 1964. An Analysis of Transformations. Journal of Royal Statistics Society 26:211-252.

Box, G. E. P. and G. M. Jenkins, 1976. Time Series Analysis: Forecasting and Control (Revised Edition). Holden Day, Inc., San Francisco, California, 532 pp.

Brockwell, P. J. and R. A. Davis, 1987. Time Series: Theory and Methods. Springer Verlag, Inc., New York, New York, 520 pp.

Cline, T. B., 1981. Selecting Seasonal Streamflow Models. Water Resources Research 17(4):975-984.

Hipel, K. W. and A. I. McLeod, 1978. Preservation of the Rescaled Adjusted Range – 2. Simulation Studies Using Box-Jenkins Models. Water Resources Research 14(3):509-516.

Hipel, K. W. and A. I. McLeod, 1994. Time Series Modelling of Water Resources and Environmental Systems. Developments in Water Science, No 45, Elsevier Science B.V., 1013 pp.

Hurst, H. E., 1951. Long-Term Storage Capacity of Reservoirs. Transaction of American Society of Civil Engineering 166:770-799.

Hurst, H. E., 1957. Suggested Statistical Model of Some Time Series Which Occurs in Nature. Nature (London), 180:494.

Kareliotis, S. J. and V. T. Chow, 1972. Analysis of Residual Hydrologic Stochastic Process. Journal of Hydrology 15(1):113-130.

Kashyap, R. L., 1977. A Bayesian Comparison of Different Classes of Dynamic Models Using Empirical Data. IEEE Trans. Autom. Control 22(5):715-727.

Klemes, V., R. Srikanthan, and T. A. McMahon, 1981. Long-Memory Flow Models in Reservoir Analysis: What Is Their Practical Value? Water Resources Research 17(3):737-751.

Kuester, J. L. and J. H. Mize, 1973. Optimization Techniques With Fortran. McGraw-Hill, New York, New York, 500 pp.

Ledolter, J., 1978. A General Class of Stochastic Models for Hydrological Sequences. Journal of Hydrology 36(2):309-325.

Lungu, E. M. and F. T. K. Sefe, 1991. Stochastic Analysis of Monthly Streamflows. Journal of Hydrology 126:171-182.

McLeod, A. I. and K. W. Hipel, 1978. Preservation of the Rescaled Adjusted Range: 1. A Reassessment of the Hurst Phenomenon. Water Resources Research 14(3):491-508.

McMahon, T. A. and R. G. Mein, 1986. River and Reservoir Yield. Water Resources Publications, Littleton, Colorado, 368 pp.

Mimikou, M. and J. Nalbandis, 1987. Influence of Reservoir Inflows' Persistence on Storage Capacity. International Journal of Modelling and Simulation 7(4):164-167.

Moran, P. A. P., 1954. A Probability Theory of Dams and Storage Systems. Austral. J. Appl. Sci. 5:116-124.

Montgomery, D. C., L. A. Johnson, and J. S. Gardiner, 1990. Forecasting and Time Series Analysis. McGraw-Hill, Inc., New York, New York, 384 pp.

Rippl, W., 1883. The Capacity of Storage Reservoirs for Water Supply. Proc. Inst. of Civil Engin. 71:270-278.

Salas, J. D., 1993. Analysis and Modeling of Hydrologic Time Series. *In:* Maidment D.R. (editor), Handbook of Hydrology. McGraw Hill, Inc., New York, New York, , pp. 19.1-19.72.

Salas, J. D. and D. C. Boes, 1974. Expected Range and Adjusted Range of Hydrologic Sequences. Water Resources Research 10(3):457-463.

Salas, J. D., D. C. Boes, V. Yevjevich, and G. G. S. Pegram, 1979. Hurst Phenomenon as a Pre-Asymptotic Behavior. Journal of Hydrology 44:1-15.

Salas, J. D., J. W. Delleur, V. Yevjevich and W. L. Lane, 1985. Applied Modelling of Hydrologic Time Series. Water Resources Publications, Littleton, Colorado, 484 pp.

Vogel, R. M. and J. R. Stedinger, 1988. The Value of Stochastic Streamflow Models in Overyear Reservoir Design Applications. Water Resources Research 24(9):1483-1490.