Journal of Spatial Hydrology
Spring vol. 5 no. 1

Performances of Stochastic Approaches in Generating Low Streamflow Data for Drought Analysis

Kadri YUREKLI[1] and Ahmet KURUNC[2]

Abstract

This study analyzed the monthly-minimum daily discharge data of each month from three gauge stations on Cekerek Stream for forecasting using stochastic approaches. Initially non-parametric test (Mann-Kendall) was used to identify the trend during study period. The two approaches of stochastic modeling, ARIMA and Thomas-Fiering models, were used to simulate the monthly-minimum daily discharge data of each month. The error estimates (RMSE and MAE) of forecasts from both approaches were compared to identify the most suitable approach for reliable forecast. The two error estimates calculated for two approaches indicate that ARIMA model appear to be slightly better than Thomas-Fiering. However, both approaches were identified as appropriate method for simulating the monthly-minimum daily discharge data of each month from three gauge stations on Cekerek Stream.

Key words: monthly-minimum daily discharge, stochastic model, ARIMA, Thomas-Fiering

Introduction

Water naturally stored in a catchment as soil moisture and in lakes, rivers, aquifers and wetland helps determine the impact of precipitation deficit. Precipitation deficits over a prolonged period that affect surface and subsurface water supply reduce streamflow, groundwater, reservoir, and lake levels. Especially, extreme streamflow droughts are the cause of serious and damaging economic and social impacts, sometimes even loss of life. Therefore, severity and duration of streamflow droughts are taken into consideration for water resource planning, design and management. In fact, the hydraulic structures are designed and built for service under abnormal conditions. During times of streamflow drought, importantance and value of these structures depend upon their performance. In addition, the hydrologic regime of streamflow droughts is of primary consideration in the design of municipal or industrial water systems, hydroelectric plants, navigation development, supplemental irrigation scheme, low flow augmentation systems, recreation facilities and other. Water quality criteria and stream standards are tied to the regime of streamflow droughts, that affect also the preservation of fish and fowl habitat. Evaluation the adequacy of river low flows to supply requirements for disposal of liquid wastes depends also on the characteristics of the process of streamflow drought.

Low streamflow characteristics so highly depend on watershed topography, climate, and land use. The difficulty in estimating low streamflow characteristics reveals that the complexity of low streamflow regime needed to be examined carefully (Chang and Boyer, 1977). Prediction of low streamflow is required for proper management of a given watershed. Sharma et al. (1997) cited that it is very important to generate synthetic streamflow sequences to analyze alternative designs, operation policies, and rules for water resources systems, and that the dependence structure of streamflow sequences is often assumed to be Markovian, that is, dependent on only a fine set of prior values.

Characteristic of many types of hydrologic time series has periodically varying components. Data of this type may be modeled using a linear stochastic model that is commonly referred to as autoregressive integrated moving average (ARIMA) model (Lewis and Ray, 2002). Walls and Bendel (1987) pointed out that the highly popularized ARIMA model could be successively applied in time series forecasting related to many areas, but also as a promising tool for modeling the empirical dependencies between successive times. It also results in satisfactory predictive performances (Ho and Xie, 1998). The popularity of ARIMA model in many areas resulted from having quite flexible of the model, due to the inclusion of both autoregressive and moving average terms. The ARIMA model approach has several advantages over others such as moving average, exponential smoothing, neural network and fuzzy logic, in particular, its forecasting capability and its richer information on time-related changes. In most time series, there is a serial correlation among observations. This

1. Res. Ass. Dr. Gaziosmanpasa Unıversity, Faculty of Agriculture, Department of Agricultural Technology 60250 Tasliciftlik-Tokat/TURKEY. Phone: (+90-0356 252 1479), Fax: (+90-0356 252 1488) (kadriyurekli@yahoo.com)

2. Res. Ass. Dr. Gaziosmanpasa Unıversity, Faculty of Agriculture, Department of Agricultural Technology 60250 Tasliciftlik-Tokat/TURKEY. Phone: (+90-0356 252 1479), Fax: (+90-0356 252 1488) (akurunc@yahoo.com)

characteristic is effectively considered by ARIMA model. This model also provides systematic searching in each stage (identification, estimation and diagnostic check) for an appropriate model (Chatfield, 1996; Zhang, 2003). One of the most important disadvantages for the ARIMA modeling approach is the necessity of large amount of data. Wei (1990) stated that a minimum of 50 observations is need to build reasonable ARIMA model. However, Velicer and Harrop (1983) noted that even this number could not be adequate for accurate model identification. Another disadvantage of this modeling approach is that ARIMA is a complex technique, requires a great deal of experience and although it often produces satisfactory results, those results depend on the researcher's level of expertise (Bails and Peppers, 1993).

Researchers have used this approach for many different scientific and technical applications. Ahlert and Mehta (1981) examined the stochastic structure of flow data for the Upper Delaware River to describe the random component of streamflow time series by ARIMA model. Fernando and Jayawardena (1994) used various ARMA models in forecasting monthly rainfall records. Venama *et al*. (1996) investigated climate change in Senegal River basin via this approach. Nunnari et al. (1998) used ARIMA model to identify models for the prediction of pollutant concentration. Chaloulakou *et al*. (1999) forecasted the daily maximum 1-hour ozone concentrations whereas Ahmad *et al*. (2001) analyzed water quality data by using ARIMA model. See and Openshaw (1998) enchanced flood forecasting on the river Ouse by using ARIMA model. Hsu et al. (1995) used an ARMA model for the prediction of streamflow on a medium sized basin in Mississippi. Yurekli et al. (2005) applied the ARIMA model to monthly data from Kelkit Stream watershed. Yurekli et al. (2005) analyzed the residuals from the ARIMA models fitted to monthly streamflow data for three gauging stations located on Çekerek stream watershed by alternatives methods. Yurekli and Ozturk (2003) used the ARIMA model in simulation of the annual-extreme daily discharge. Kurunc et al. (2005) evaluated performance of two stochastic approaches (ARIMA and Thomas-Fiering) for forecasting water quality and streamflow data from Yeşilırmak River, Turkey.

In this study, the outcome of an attempt to generate synthetic the monthly-minimum daily discharge data of each month (hereafter referred to as monthly low data) for three gauge stations on Cekerek Stream by using ARIMA and Thomas-Fiering which this method was taken into consideration to the ARIMA model as an alternative approaches, is presented.

Material and Method

Study area

In this study, monthly low data from three gauge stations as numbered 1404, 1409 and 1424, which are managed by General Directorate of Electric Power Research Survey and Development Administration (EIE), in Cekerek Stream watershed were used as materials. The approximate locations of the gauge stations were given in Figure 1 and a summary of identification number, names and drainage areas for the gauge stations was presented in Table 1.

Cekerek Stream watershed is bounded 39º 30' and 40º 45' N latitudes, 35º 15' and 36º 15' E longitudes, covering approximately 1165440 ha which is about 1.5% of Turkey’s total area. The study area is located on the north Anatolia fault line that is one of the most effective faults in the world. Therefore, tectonic movement affects this watershed area. Cekerek Stream is formed by joining together of small streams that originate from Kızık, Dinar, Calı and Kavak hills, near the Camlıbel district. Cekerek Stream joins to Yesilirmak River near Kayabası. The stream is approximately 276.0 km in length and water quality of the stream is C_2S_1 for irrigation (Anonymous, 1970).

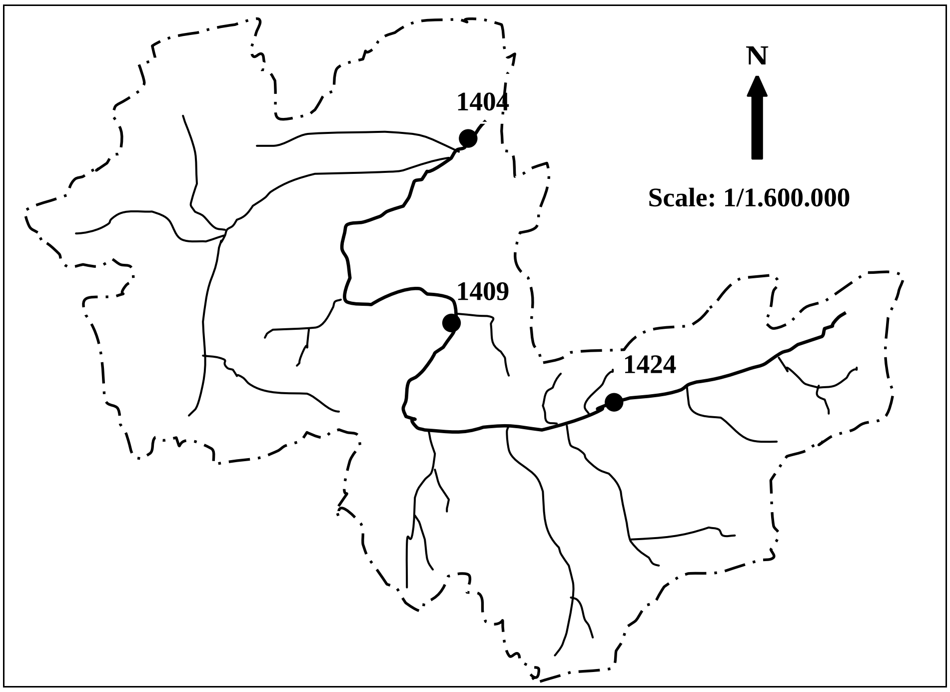

Figure 1. Location of gauge stations on Cekerek Stream

Table 1. Cekerek Stream Gauge Station Identification

Station Number	Station Name	Drainage Area, km^2	Number of years of data
1404	Cekerek-Kayabası	11724.0	13
1409	Cekerek-Akcakecili	5267.6	38
1424	Cekerek-Cırdak Bridge	1032.8	27

Time series analysis for monthly low streamflow data

In order to analyze time series for monthly low data from the three gauge stations, linear stochastic models known as either Box-Jenkins or ARIMA (autoregressive-moving average) and Thomas-Fiering model were used in this study.

ARIMA model

For fitting seasonal ARIMA model to the time series of monthly low streamflow data, three-stage procedure of model identification, estimation of model parameters and diagnostic checking of estimated parameters has been adopted. This seasonal ARIMA model (Hipel et al., 1977) denoted as ARIMA $(p,d,q)(P,D,Q)_s$ is expressed as

$$\varnothing(B)\Phi(B^s)(w_i - \mu) = \theta(B)\Theta(B^s)a_i \quad (1)$$

$$w_i = (1-B)^d (1-B^s)^D x_i \quad (2)$$

In Equation 1, w_i should be taken as z_i if the series is stationary.

Identification stage is purposed to determine the differencing required to produce stationarity and also the order of both the seasonal and nonseasonal AR and MA operators for a given series. By plotting original series (monthly series), seasonality, trends in the mean and variance may be revealed (Box and Jenkins, 1976). The following non-parametric test (Mann-Kendall) can be applied to decide whether trend exists in the monthly low data. The Mann-Kendall test recommended by Hirsch et al. (1982) is given as:

$$u_c = \frac{S + m}{\sqrt{V(S)}} \quad (3)$$

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} z_k \quad (4)$$

$z_k = 1$ if $x_j > x_i$
$z_k = 0$ if $x_j = x_i$
$z_k = -1$ if $x_j < x_i$ (5)

$$V(S) = 18^{-1}(n^2 - n)(2n + 5) - \sum_{i=1}^{t} e_i (e_i - 1)(2e_i + 5) \quad (6)$$

$m = 1$ if $S < 0$
$m = 0$ if $S = 0$
$m = -1$ if $S > 0$ (7)

To determine whether there is a trend, u_c statistic in Equation 3 should be compared to the z-table critical value. If the u_c statistic lies within the 5% significance interval, there is no trend for the data set. The hypothesis of an upward or downward trend cannot be rejected at the α significance level if the absolute value of $u_c > u_{1-\alpha/2}$, where $u_{1-\alpha/2}$ is the 1- α/2 quantile of standard normal distribution.

Autocorrelation function (ACF) and partial autocorrelation function (PACF) should be used to gather information about the seasonal and nonseasonal AR and MA operators for the monthly low series. ACF measures the amount of linear dependence between observations in a time series. Therefore, the most useful device is the autocorrelation function of the time series. In this sense, the identification of the appropriate parametric time series model depends on the shape of ACF. Additional to ACF, a powerful complementary identification tool, the partial autocorrelation function (PACF) can also be used (Janacek and Swift, 1993).

Estimation stage consists of using the data to estimate and to make inferences about values of the parameters conditional on the tentatively identified model. In an ARIMA model, the residuals (a_i) are assumed to be independent, homoscedastic, and usually normally distributed. However, if the constant variance and normality assumptions are not true, they are often reasonably well satisfied when the observations are transformed by a Box-Cox transformation. The transformations can be expressed as either of the following equations (Wei, 1990):

$$z_{i=1}^{n} = \lambda^{-1}\left[\left(x_{i=1}^{n} + c\right)^{\lambda} - 1\right] \qquad \lambda \neq 0 \quad (8)$$

$$z_{i=1}^{n} = \ln\left(x_{i=1}^{n} + c\right) \qquad \lambda = 0 \quad (9)$$

Box and Jenkins (1976) cited that the model should be parsimonious. Therefore, they recommended the need to use as few model parameters as possible so that the model fulfils all the diagnostic checks. Akaike (1974) suggests a mathematical formulation of the parsimony criterion of model building as AIC (Akaike Information Criterion) for the purpose of selecting an optimal model fits to a given data. Mathematical formulation of AIC is defined as:

$$\text{AIC}(M) = n \ln\sigma_a^2 + 2M \quad (10)$$

Where M is the number of AR and MA parameters to estimate. The model that gives the minimum AIC is selected as a parsimonious model.

Shibata (1976) has shown that the AIC criterion trends to overestimate the order of the autoregression. But, Akaike (1978, 1979) has developed a Bayesian extension of minimum AIC procedure, called as BIC. Similar to Akaike's BIC, Schwarz (1978) suggested the following Bayesian criterion for model selection, which has been called Schwarz Bayesian Criterion (SBC):

$$\text{SBC}(M) = n \ln\sigma_a^2 + M \ln n \tag{11}$$

Diagnostic check stage determines whether residuals are independent, homoscedastic and normally distributed. The residual autocorrelation function (RACF) should be obtained to determine whether residuals are white noise. There are two useful applications related to RACF for independence of residual. The first one is the corrologram drawn by plotting r_k (a) against lag k. If some of the RACF are significantly different from zero, this may mean that the present model is inadequate. The second one is Q (r) statistic suggested by Ljung-Box (1978). A test of this hypothesis can be done for the model adequacy by choosing a level of significance and then comparing the value of calculated χ^2 to χ^2-table of critical value. If the calculated χ^2 value is smaller than the χ^2-table critical value, the present model is adequate on the basis of available data. The Q (r) statistic is calculated by using:

$$Q(r) = n(n+2)\sum_{k=1}^{m}(n-k)^{-1} r_k(a)^2 \tag{12}$$

$$r_k(a) = \sum_{i=k+1}^{n} a_i a_{i-k} / \sum_{i=1}^{n} a_i^2 \tag{13}$$

The following test described by Breusch and Pagan (1979) is very useful to determine whether a transformation of the data is needed. If there is a change in variance (heteroscedasticity) of residuals, a transformation is necessary for the data. For the test, the residuals from the model fit to the data are divided into two groups. Then, residual sum of squares (ESS_F, ESS_S) for these group are obtained. Breusch-Pagan test statistic (F_{cal}) is obtained from the following equation. If F_{cal} is smaller than F-table critical value, the residuals are assumed to be homoscedastic.

$$F_{cal} = \frac{ESS_S / (n_S - k_p)}{ESS_F / (n_F - k_p)} \approx F_{table}\left[(n_S - k_p), (n_F - k_p)\right] \tag{14}$$

There are many standard tests available to check whether the residuals are normally distributed. Chow et al. (1988) cited that if a historical data is normally distributed, the graph of the cumulative distribution for the data should appear as a straight line when it is plotted on normal probability paper. Haan (1977) expressed that the other way to check normality of residuals is the Kolmogorov-Smirnov method.

Thomas-Fiering model

Thomas-Fiering model presents a set of 12 regression equations. This linear stochastic model is used for generating synthetic monthly data. The well-known Thomas –Fiering model equation can be given as (Clarke, 1984):

$$\frac{X_{i,j} - \overline{Q_j}}{S_j} = r_j \frac{X_{i,j-1} - \overline{Q_{j-1}}}{S_{j-1}} + a_{ij}\sqrt{(1 - r_j^2)} \tag{15}$$

Comparison of the results

Two error estimates were taken into consideration for comparison of the results from ARIMA and Thomas-Fiering approaches (Antonopoulos et al., 2001). The first of them is the Root Mean Square Error (RMSE), which is given as:

$$\text{RMSE} = \sqrt{n^{-1}\sum_{i=1}^{n}\left\{Q_{obs}(i) - Q_{pred}(i)\right\}^2} \tag{16}$$

The second is the Mean Absolute Error (MAE), which is defined as

$$MAE = n^{-1}\sum_{i=1}^{n}\left|Q_{obs}(i) - Q_{pred}(i)\right| \tag{17}$$

Results and Discussion

To determine whether there is a trend in monthly low streamflow data sequences from 1404, 1409 and 1424 gauge stations, the non-parametric test (Mann-Kendal test) at 5% significance level was applied to monthly low data sequences. Mann-Kendal test results were given in Table 2. The Mann-Kendal statistic (u_c) values of monthly low data from three gauge stations were between z-table critical values (±1.96) at 5% significant level. This suggests that there is no linear trend in monthly low data sequences of each mentioned gauge station.

Table 2. The ARIMA models selected for Cekerek Stream gauge stations

Gauge Station	**ARIMA Model**	**Model Statistics**						
		u_c	**AIC**	**SBC**	**LBQ/P**	**Norm**	**Homosce**	**σ_a^2**
1404	(1,0,0)(0,1,1)	0.021	287.6	293.6	0.449	0.168	0.809	0.369
1409	(2,0,1)(0,1,1)	0.011	655.5	671.8	0.302	0.053	0.930	0.240
1424	(1,0,1)(0,1,1)	0.014	452.8	464.0	0.923	0.182	0.722	0.230

The plots of the auto-correlation function (ACF) and the partial auto-correlation function (PACF) for monthly low data sequences were drawn to gather information about the seasonal and nonseasonal AR and MA operators concerning with the monthly series for the ARIMA model. The ACFs for monthly low data follow an attenuating sine wave pattern that reflects the random periodicity of the data and possibly indicates the need for non-seasonal and/or seasonal AR terms in the model. For these data sequences that have the cyclic seasonal component, seasonal differencing was needed. By taking seasonal differencing operator as one (1), the seasonal wave pattern in the ACFs was removed.

All the ACF graphs of monthly data sequences were significantly different from zero. This implies the existence of linear dependence between amounts of monthly data sequences. Additional to this, Ljung-Box Q statistics estimated also emphasis that the observations related to the monthly low data sequences have linear dependence. The ACFs did not cut off but rather damp out. This may suggest the presence of autoregressive (AR) terms. The PACFs possess significant values at some lags but rather tail off. This may imply the presence of moving average (MA) terms. The ACFs have significant values at lags that are multiples of 12. This may stress that seasonal AR terms are required but these values attenuate. There are peaks on graphs of the PACFs at lags that are multiples of 12 may suggest seasonal MA terms, but these peaks damp out.

Alternative ARIMA models were estimated by considering the ACFs and PACFs graphs from the monthly low data. The SBC was taken into account for obtaining parsimonious model among these alternatives. The model that has the minimum SBC was assumed to be parsimonious. In addition to this, model parameters were analyzed at 5% significant level by using t-test to select the best model fit to the data. If there is any parameter greater than 5%, the model with that parameter was eliminated. The delineated models based on SBC criterion and parameter analysis for three gauge stations was identified as best model.

Diagnostic checks were applied in order to determine whether the residuals from the delineated models were independent, homoscedastic and normally distributed. A Box-Cox transformation was required for monthly low data of all gauge stations as the residuals from the best models were not fulfilled the condition concerning with homoscedasticity and normality. By taking the value of constant (c) as 0.0 for monthly low data sequences from all gauge stations in Equations (6) and (7), a Box-Cox transformation caused the residuals to be homoscedastic and approximately normally distributed.

The models that have the minimum SBC and the parameter(s) less than the critical value of t-test at %5 significance level among the selected models fulfilled all the diagnostic checks were delineated as the best model for monthly low data sequences from the gauge stations. The best model for each gauge station was presented in Table 2. The critical assumption of independence for the RACFs of the residuals was done by using the χ^2 distributed Ljung-Box Q statistic. The probabilities of Q statistics calculated for the best models were given in Table 2. Since the probabilities of Q statistics are greater than 0.05, the residuals from the best models are not significantly different from zero. That is, the residuals were assumed to be independence. Similarly, the RACFs drawn for the best models indicated that the residuals were not significantly different from a white noise series at 5% significance level. Inspection of the RACFs and the residuals integrated periodogram (Figure 2) confirmed a strong model fit.

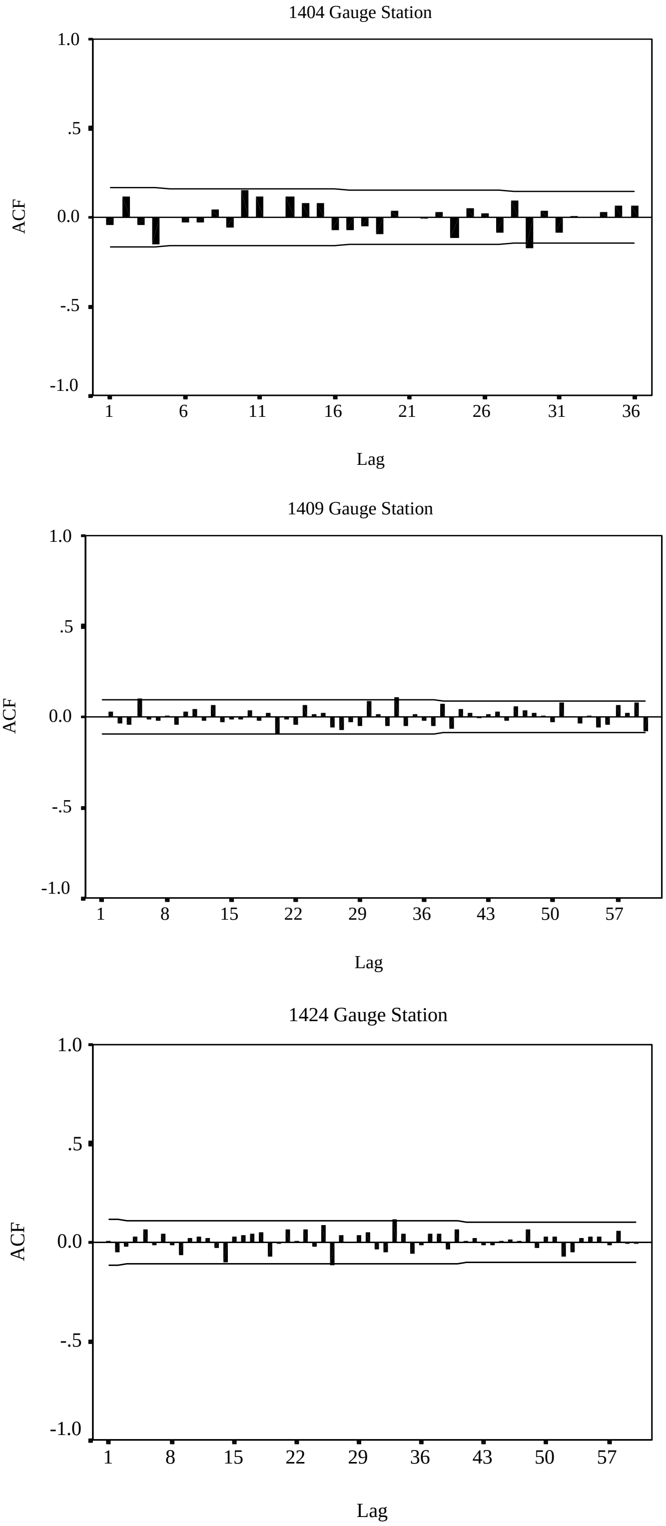

Figure 2. Residual ACF- monthly low data

In Table 2, test results from Kolmogorov-Simirnov method for the normality and test results from Breusch-Pagan approach for homoscedascity of the residuals were also presented. Since the probabilities of the values belonging to normality and homoscedasticity tests are greater than 0.025 and 0.05, respectively, the best models for the gauge stations performed the condition of normality and homoscedasticity.

The value (V) of the parameters concerning with the best models associated the standard errors (SEV), t-ratios and probabilities (<5%) for the standard errors were listed in Table 3. The standard errors calculated for the model parameters were rather small compared to the parameter values. Therefore, all of the parameters are significant and these parameters should be included in the models (Table 3).

Table 4 gives the error estimates obtained for monthly low data of the two different approaches used in the study for forecasting. The two error estimates (RMSE and MAE) obtained for two approaches indicate that ARIMA approach appear to be slightly better than Thomas-Fiering. Figure 3 shows the relationship between five-years of monthly low data at each gauge station and predicted data for the same years by using the best models from ARIMA and Thomas-Fiering approaches for each gauge station. As they were shown in Figure 3, the predicted data obtained from the approaches follow monthly low data very closely for three gauge stations on Cekerek Stream. Therefore, both models seem to be adequate for simulating monthly low data.

Table 3. Statistical analysis for the model parameters

Gauge Station	**Model Parameters**	**Variables in the Model**			
		V	**SEV**	**t-ratio**	**Probability**
1404	ϕ_1	0.751	0.052	14.50	0.000
	Θ_1	0.908	0.114	7.99	0.000
1409	ϕ_1	1.460	0.224	6.53	0.000
	ϕ_2	-0.503	0.182	-2.77	0.006
	θ_2	0.756	0.203	3.72	0.000
	Θ_1	0.895	0.028	32.47	0.000
1424	ϕ_1	0.791	0.049	16.15	0.000
	θ_1	0.194	0.077	2.52	0.012
	Θ_1	0.948	0.059	16.19	0.000

Table 4. Comparison of the results from different approaches

Gauge Station	**ARIMA**		**Thomas-Fiering**	
	RMSE	**MAE**	**RMSE**	**MAE**
1404	0.64	0.44	0.96	0.71
1409	0.00	0.00	0.97	0.69
1424	0.50	0.37	0.98	0.72

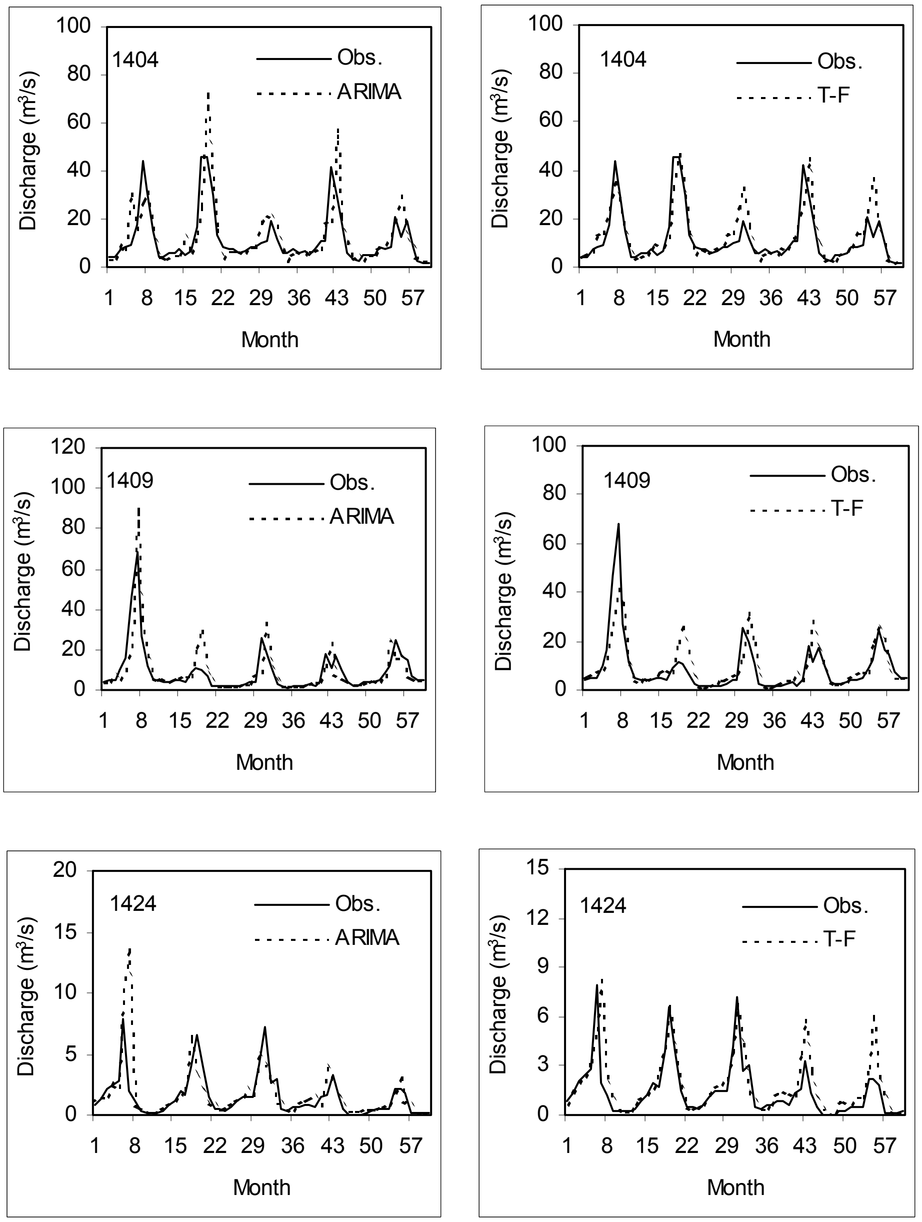

Figure 3. Comparison of observed data to predicted data using different approaches

Comparison of means and standard deviation for observed data and generated data from the used approaches were given in Figure 4. To determine whether there is significant difference between the mean and standard deviation values from the observed and predicted data for each month, z-test for the means and F-test for the standard deviation were applied (Devore and Peck, 1993; Haan, 1977). Since means of each month from all gauge station were between z-critical table values (± 1.96 for two tailed at 5% significance level), the data support the claim that there is no difference between the mean values of observed and predicted data. Similarly, standard deviation values of each month from all gauge station were smaller than F-critical table values at 5% significance level. Thus, the results show that generated data preserve the basic statistical properties of the observed series.

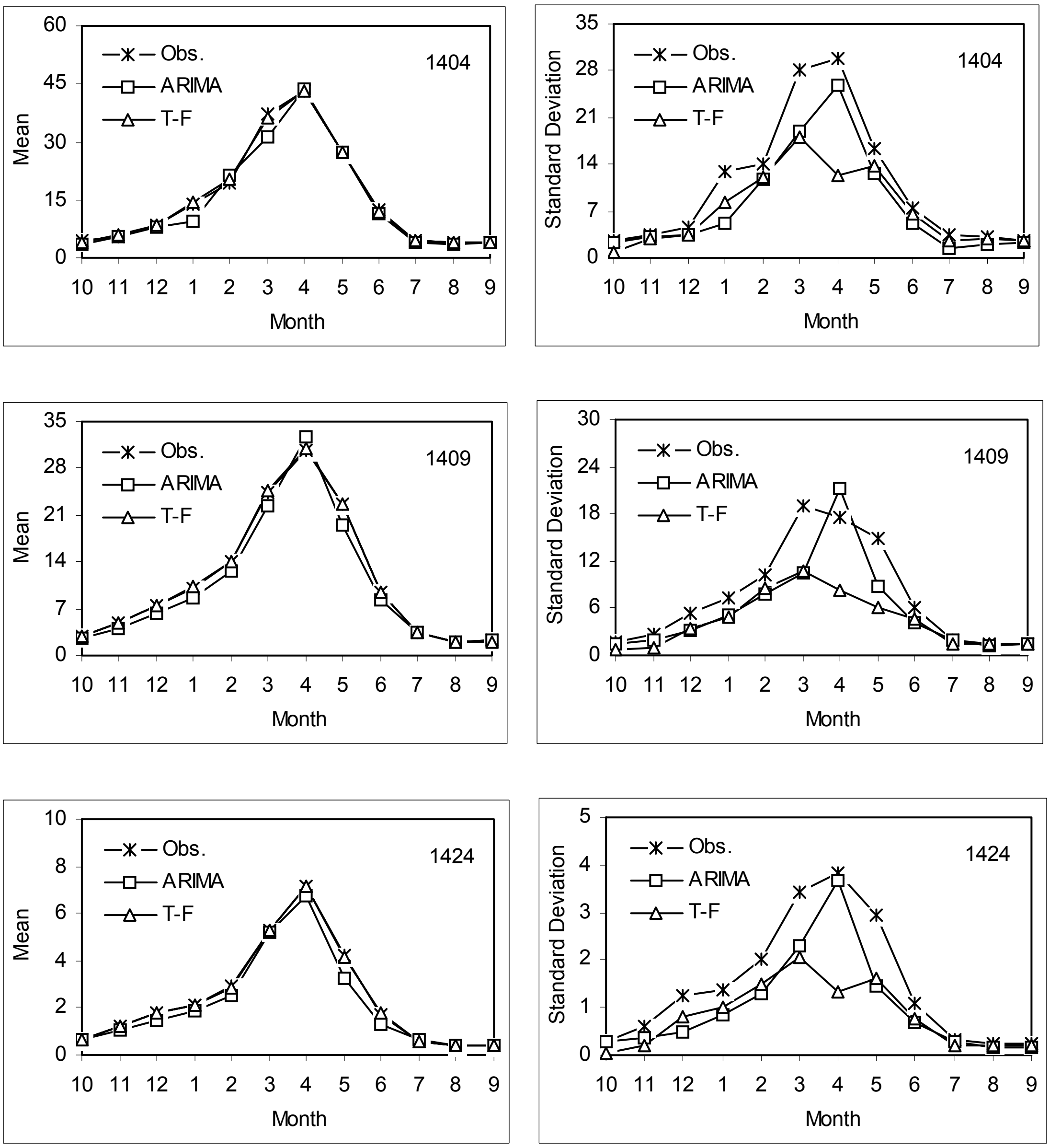

Figure4. Comparison of some statistical properties from observed and predicted data

Nomenclature

a_i	white noise time series value at time i
a_{ij}	independent standard normal variable at time i in the j^{th} month
B	backward shift operator
c	constant for Box-Cox transformation
d	order of the nonseasonal differencing operator
D	order of the seasonal differencing operator
ESS_F	the residual sum of square for first group
ESS_S	the residual sum of square for second group
k_p	degree of freedom
LBQ/P	probability for Q(r)
n	the number of observation
n_F	the number of residuals in the first group
n_S	the number of residuals in the second group
$\overline{Q}_j$	the mean monthly discharges during month j
Q(r)	Ljung-Box statistic at lag m
Q_{obs}	observed discharge

Q_{pred}	predicted discharge
r_j	the serial correlation coefficient for discharge in the j^{th} month from the $(j-1)^{th}$ month
$r_k(a)$	ACF of a_i at lag k
s	seasonal length
S_j	the standard deviation monthly discharges during month j
u_c	Mann-Kendall statistic
x_i	discrete time series value at time i
$X_{i,j}$	predicted discharge for the j^{th} month from the $(j-1)^{th}$ month at time i
w_i	stationary series formed by differencing the x_i
z_i	transformation of x_i series

Greek symbols

λ	exponent for Box-Cox transformation
μ	mean level of the w_i series (if $D+d>0$ often $\mu \approx 0$)
$\varnothing_i$	i^{th} nonseasonal AR parameter
Φ_i	i^{th} seasonal AR parameter
θ_i	i^{th} nonseasonal MA parameter
Θ_i	i^{th} seasonal MA parameter

References

Ahlert, R.C., and Mehta, B.M. (1981) Stochastic Analysis and Transfer Fuctions for Flows of The Upper Delaware River, Ecological Modeling, 14, 59-78.

Ahmad, S., Khan, I.H., and Parida, B.P. (2001) Performance of Stochastic Approaches for Forecasting River Water Quality. Water Resources, 35, 4261-4266.

Akaike, H. (1974) A Look at The Statistical Model Identification, IEEE Transactions on Automatic Control, AC-19, 716-723.

Akaike, H. (1978) A Bayesian Analysis of The Minimum AIC Procedure, Ann. Inst. Statist. Math., 30A, 9-14.

Akaike, H. (1979) A Bayesian Extension of The Minimum AIC Procedure of Autoregressive Model Fitting, Biometrika, 66, 237-242.

Anonymous. (1970) Yesilirmak Havzasi Topraklari. Ankara-TURKEY: Topraksu Genel Mudurlugu Yayinlari (in Turkish).

Antopoulos, V.Z., Papamichail, D.M., and Mitsiou, K.A. (2001) Statistical and Trend Analysis of Water Quality and Quantity Data for The Strymon River in Greece, Hydrology and Earth System Science, 5, 679-691.

Bails, D.G., and Peppers, L.C. (1993) Business Fluctuations: Forecasting Techniques and Applications. New Jersey: Prentice-Hall, Inc.

Box, G.E.P., Jenkins, G.M. (1976) Time Series Analysis Forecasting and Control. San Francisco: Holden-Day.

Breusch, T., and Pagan, A. (1979) A Simple Test of Heteroscedasticity and Random Coefficient Variation, Econometrica, 47, 1287-1294.

Chaloulakou, A., Assimacopoulos, D., and Lekkas, T. (1999) Forecasting Daily Maximum Ozone Concentrations in The Athens Basin, Environmental Monitoring and Assessment, 56, 97-112.

Chang, M., and Boyer, D.G. (1977) Estimates of Low Flows Using Watershed and Climatic Parameters, Water Resources Research, 13, 997-1001.

Clarke, R.T. (1984). Mathematical Models in Hydrology. Rome: FAO of United Nation.

Chatfield, C. (1996) The Analysis of Time Series: An Introduction. New York: Chapman and Hall.

Chow, V.T., Maidment, D.R., and Mays, L.W. (1988) Applied Hydrology. New York: McGraw-Hill Book Company.

Devore, J., and Peck, R. (1993) Statistics: The Exploration and Analysis of Data. California: Second Edition, Duxbury Press.

Fernando, D.A.K., and Jayawardena, A.W. (1994) Generation and Forecasting of Monsoon Rainfall Data. 20th WEDC Conference, Colombo, Sri Lanka, 310-313.

Haan, C.T. (1977) Statistics Methods in Hydrology. Iowa: Iowa State Press.

Hipel, K.W., McLeod, A.I., and Lennox, W.C. (1977) Advances in Box-Jenkins modeling: I. Model Construction, Water Resources Research, 13, 567-575.

Hirsch, R.M., Slack, J.R., and Smith R.A. (1982) Techniques of Trend Analysis for Monthly Water Quality Data, Water Resources Research, 18, 107-121.

Ho, S.L., and Xie, M. (1998) The Use of ARIMA Models for Reliability Forecasting and Analysis, Computers and Industrial Engineering, 35, 213-216.

Hsu, K.L., Gupta, H.V., and Sorooshian, S. (1995) Artifical Neural Network Modeling of The Rainfall-Runoff Process, Water Resources Research, 31, 2517-2530.

Janacek, G., and Swift, L. (1993) Time Series Forecasting, Simulation, Application. New York: Ellis Horwood.

Kurunc, A., Yurekli, K., and Cevik, O. (2005) Performance of Two Stochastic Approaches for Forecasting Water Quality and Streamflow Data from Yeşilırmak River, Turkey, Environmental Modelling & Software, 20, 1195-1200.

Lewis, P.A.W., and Ray, B.K. (2002) Nonlinear Modeling of Periodic Threshold Autoregressions Using TSMARS, Journal of Time Series Analysis, 23, 459-471.

Ljung, G.M., and Box, G.E.P. (1978) On a Measure of Lack of Fit in Time Series Models, Biometrika, 65, 297-303.

Nunnari, G., Nucifora, A.F.M., and Randieri, C. (1998) The Application of Neural Techniques to The Modeling of Time-Series of Atmospheric Pollution Data, Ecological Modeling, 111, 187-205.

See, L., and Openshaw, S. (1998) Using Soft Computing Techniques to Enhance Flood Forecasting on The River Ouse. Proceeding Hydroinformatics'98: 3rd International Conference on Hydroinformatics, Copenhagen, Denmark, 24-26 August, pp. 819-824.

Schwartz, G. (1978) Estimating The Dimension of A Model, Ann. Statist., 6, 461-464.

Sharma, A., Tarboton, D.G., and Lall, U. (1997) Streamflow Simulation: A Nonparametric Approach, Water Resources Research, 33, 291-308.

Shibata, R. (1976) Selection of The Order of An Autoregressive Model by Akaike's Information Criterion, Biometrika, 63, 117-126.

Velicer, W.F., and Harrop, J. (1983) The Reliability and Accuracy of Time Series Model Identification, Evaluation Review, 7, 551-560.

Venema, H.D., Schiller, E.J., and Adamowski, K. (1996) Evidence of Climate Change in The Senegal River Basin, Water Resources Development, 12, 531-546.

Walls, L.A., and Bendel, A. (1987) Time Series Methods in Reliability, Reliability Engineering, 18, 239-265.

Wei, W.W.S. (1990) Time Series Analysis. New York: Addison-Wesley Publishing Company, Inc.

Yurekli, K., Kurunc, A., and Ozturk, F. (2005) Application of Linear Stochastic Models to Monthly Flow Data of Kelkit Stream, Ecological Modelling, 183, 67-75.

Yurekli, K., Kurunc, A., and Ozturk, F. (2005) Testing The Residuals of an ARIMA Model on The Cekerek Stream Wtershed in Turkey, Turkish Journal of Engineering and Environmental Sciences, 29, 61-74.

Yurekli, K., and Ozturk, F. (2003) Stochastic Modeling of Annual Maximum and Minimum Streamflow of Kelkit Stream, Water International, 28, 433-441.

Zhang, G.P. (2003) Time Series Forecasting Using a Hybrid ARIMA and Neural Network Model, Neurocomputing, 50, 159-175.