Analysis of Recorded Inflow Data of Ataturk Reservoir

Prof. Ahmed Mohammed Ali
Department of Water Resources
College of Engineering
Baghdad University
Email: drahmedmali@yahoo.com

Asst. Lect. Zienh Sami Saaed
Department of Water Resources
College of Engineering
Baghdad University
Email: zienhsami@yahoo.com

ABSTRACT

Since the beginning of the last century, the competition for water resources has intensified dramatically, especially between countries that have no agreements in place for water resources that they share. Such is the situation with the Euphrates River which flows through three countries (Turkey, Syria, and Iraq) and represents the main water resource for these countries. Therefore, the comprehensive hydrologic investigation needed to derive optimal operations requires reliable forecasts. This study aims to analysis and create a forecasting model for data generation from Turkey perspective by using the recorded inflow data of Ataturk reservoir for the period (Oct. 1961 - Sep. 2009). Based on 49 years of real inflow data from the Euphrates River recorded at Ataturk, a spilt-sample approach was adopted for testing homogeneity. The autoregressive model of order one [AR(1)] was found to be the best for the forecasting as it accurately reproduced the means, standard deviations, and skewness coefficients observed in the generated records forecast at the Ataturk reservoir. Ten sets of 100 years data have been forecasted.

In Iraq, optimization of the operation of all reservoirs is necessary after operating new reservoirs in Turkey.

Keywords: Time series, Reservoir operation, Euphrates River, Ataturk Dam, and Forecasting.

تحليل معلومات الجريان الداخل الى خزان أتاترك

أ.د.أحمد عبد الصاحب محمد علي
قسم هندسة الموارد المائية
كلية الهندسة / جامعة بغداد

م.م. زينة سامي سعيد
قسم هندسة الموارد المائية
كلية الهندسة / جامعة بغداد

الخلاصة :

منذ بداية القرن الماضي ازداد الصراع على الموارد المائية , وخاصة بين الدول المتشاطئة لعدم وجود اتفاقيات بينهم, كما هو الحال مع نهر الفرات الذي يجري بثلاث دول (تركيا , سوريا , العراق), ويمثل المصدر الرئيسي للمياه لتلك الدول, لذالك أصبح من الضروري التكهن بالمتغيرات الهيدرولوجية لمشاريع الموارد المائية من أجل الادارة المثلى لتلك المشاريع.

ان الهدف من هذه الدراسة هو تحليل بيانات التدفق الداخلة لسد أتاترك والتكهن بها من وجهة النظر التركية, وذلك اعتمادا على التصاريف المرصوده لنهر الفرات عند سد أتاترك خلال 49 سنة مائية للفترة من تشرين الاول 1961 ولغاية ايلول 2009. حيث وجد بان نموذج ((AR(1) Autoregressive يعد الأفضل في هذا المجال لانها أنتجت قيم كل من المعدل والانحراف المعياري ومعامل الالتواء للبيانات المرصودة نفسها او قريبة جدا منها للبيانات الصنيعة .وعليه تم استخدام هذا النموذج للتنبؤ بعشر سلاسل من التصاريف الصنيعة الداخلة لخزان سد أتاترك ولفترة 100 عام لكل منها.

ان التشغيل الامثل لكل الخزانات في العراق ضروري جدا وخاصة بعد تشغيل تركيا لخزاناتها الجديدة .

الكلمات الرئيسية :
البيانات المتسلسلة مع الزمن , تشغيل الخزانات , نهر الفرات , سد اتاترك , التكهن بالبيانات.

1. INTRODUCTION

The Euphrates has its source in eastern Turkey. Euphrates brings water to the Mesopotamian lowlands of Iraq as well as hydropower and irrigation to parts of southeastern Turkey and much of northern and eastern Syria. It is the longest river (2,700 km) in southwest Asia west of the Indus, although its maximum average annual volume (35.9×10^9 m³ at Hit, Iraq) is less than that of the Tigris (70.4×10^9 m³ at Baghdad) or the Karun (48.8×10^9 m³ at Ahwaz), **Cressey, 1958**. The Euphrates River enters Iraq border at Hussaiba town. Because its waters comes from melting snows, maximum flows are in April and May, while minimum flows are in September and October. Many researchers have studied the Euphrates and Tigris rivers basins and the Turkish Great Anatolia Project (GAP). **Mujumdar and Kumar, 1990** investigated 10 candidate models from the autoregressive moving average model (ARMA) family for representing and forecasting monthly and 10-day stream flow in three Indian rivers.

Kolars and Mitchell, 1991 introduced a chart for projected sequential depletion of the Euphrates River for the period 1990-2040.

Al-Tikriti, 2001 used single site multivariate autoregressive models, AR(1), to model seven parameters of average weekly water quality data and discharges at two stations on the Euphrates River (Al-Hindiya and Al-Samawa stations) for the period 1984-1997.

El-Obaidy, 2006 studied the effect of future Turkish projects on the Tigris River. In order to consider the uncertain conditions affecting the future performance of the system, a multisite ARMA (1, 1) model was used to generate a monthly time series over 70 years of inflows to various reservoir sites on the Tigris River in Turkey.

2. AREA OF STUDY

The Ataturk Dam, power station, and irrigation project is located near town of Bozova, 70 km northeast of Urfa and 181 km downstream from Karakaya. It will be the largest dam in Turkey, its filled reservoir capacity as well as its embankment volume will make it the fifth-largest dam in the world, **TDN, 1988**. Constructed in the 1980s on the Euphrates River in semi-arid Southeastern Turkey, it forms the central component of a large-scale regional GAP development project. The dam and its associated hydroelectric power plant went into service in 1992 and today plays an important role in the development of Turkey's energy and agriculture sectors, **Ozcana et al., 2012.** The dam is located at 37˚28`54�much N 38˚ 19` 03‶ E / 37.48167˚ N and 38.31750˚ E, www.marefa,org> **index.php**. It is the third dam built on the Euphrates River with an active capacity of 19×10^9 m^3, within a total capacity of 50×10^9 m^3 and installed power of 2,400 MW. Its height from the river bed is 169 m with a crest length of 1,664 m and a catchment area of 92,240 km^2**, Demir et al., 2009**.

3.THEORIES

3.1 Hydrologic Time Series

A hydrologic time series is defined as a continuous set of sequential observations, usually expressed as an average value over specified intervals of time such as mean daily, mean monthly, or mean annual flows. Hydrologic time series may consist of four components depending on the type of variable and the average time interval**, Yevjevich , 1972**. In seasonal stream flow series four components exist as shown in Eq. (1):-

$$Q = J_e + T_e + P_e + E_e \tag{1}$$

where:

Q is the time series value (actual data) at period t,

J_e is the jump component at period t,

T_e is the trend-cycle component at period t,

P_e is the periodic or cyclic component at period t, and
E_e is the stochastic (or random) component at period t.
The first three components represent the deterministic part of the process while the fourth component represents the non-deterministic part.

3.1.1 Test and Removal of Non-Homogeneity

This test is performed to check the homogeneity of the historical data in the stream flow series. In order to construct a model of a stream flow that remains valid for the future, the hydrologic data series which are used in generating the model should be homogenous. By definition, homogeneity requires at least two conditions, **Yevjevich, 1972**:

1. The hydrological data series must not contain any systematic error.
2. All the hydrological conditions should be constant.

If these conditions are satisfied then the series may be considered homogeneous. These two conditions imply that a homogenous series should be free from both trend and jump components; therefore, homogeneity is enforced at the beginning of an analysis by detection and removal of these components.
A jump component is defined as a sudden slippage (either negative or positive) in the parameters of the historical data, such as in the means or standard deviations of the stream flow data. A sudden increase is termed a positive jump, whereas a sudden decrease is termed a negative jump. The jump component usually results from human activities; for example; the construction of a dam, a reservoir, or an outflow canal upstream of the observation station.
A trend component is defined as a systematic and continuous change over an entire sample in any parameter of the series. A trend can be negative or positive. It may be traced to human causes (such as diversions of flow for irrigation), to natural causes (such as climate changes), or to methodological causes (such as measurement inconsistencies or other systematic errors).

To check for the existence of these components, statistical test methods such as the (t-test or F-test) may be used to detect significant changes in means or standard deviations of two samples at a desired percent probability level of significance. If those tests indicate significant changes, an analyst concludes that the two samples are from different populations and a jump and/or trend component exists. Trend components may also be detected by regression analysis and described mathematically by means of polynomial functions. **Yevjevich, 1972** maintains that the most powerful method for testing homogeneity is carried out by using the split-sample approach. Here, the entire sample is divided into two subsamples, and then means and standard deviations for each subsample are calculated. These are then tested to ascertain whether their differences are significantly nonzero at a 95% confidence level.

3.2 Analysis and Forecasting of ATATURK Recorded Data

Tests for homogeneity require that the data sample be divided into two subsamples. The recorded inflow data at Ataturk dam from Oct.1961 to Sep.2009 were thus divided into two subsamples; the first was 24 years long, spanning Oct.1961 to Sep.1984, while the second was 25 years long, spanning Oct.1985 to Sep.2009.

To remove non-homogeneity, **Yevjevich, 1972** suggests fitting linear regression equations to both annual averages and annual standard deviations according to the following equation:

$$Y_{(j,t)} = Sd_2\, [\, X_{(j,t)} - M_{(j)}\,] / S_{(j)} + Av_2 \tag{2}$$

where:
j,t = the annual and seasonal positions of observations, respectively,
Y = transformed series (homogeneous),
X= historical non-homogeneous series,

Av_2 , Sd_2= the average and standard deviation of the second subsample, respectively,
$M_{(j)}$, $S_{(j)}$ = linear regression of annual historical mean and standard deviation against years (The equations in the upper right corner in **Figs. 1 and 2**).
The trend component of the considered historical data is removed by applying Eq.(3):

$$Y(j,t) = \frac{Y(j,t) - 912.595 + 2.8655 * (i + j/12)}{810.332 - 3.7177 * (i + j/12)} * 199.1 + 813.3 \quad (3)$$

where:
j,t = the annual and seasonal positions of observations, respectively; the constants (813.3 and 199.1) m^3/s are the overall mean and standard deviation of the second subseries, respectively over the 25 years Oct.1985- Sep.2009.

The test is now repeated to check the existence of trend component by using [Y(j,t)] as the new series for the whole dataset.

3.2.1 Detection and Removal of the Periodic Component

The correlogram is useful for the detection of the periodic component. If it reflects periodicity, that means there is a periodic component in the series, otherwise there is not. The serial correlation coefficients of the flow at Ataturk are calculated for lags (1 to 24), using the expression given by Eq.(4).

$$r(k) = \frac{\sum_{j=1}^{N-k} Y(j)\,Y(j+k) - \frac{1}{N-k}\left(\sum_{j=1}^{N-k} Y(j)\right)\left(\sum_{j=1}^{N-k} Y(j+k)\right)}{\left[\sum_{j=1}^{N-k} Y^2(j) - \frac{1}{N-k}\left(\sum_{j=1}^{N-k} Y(j)\right)^2\right]^{\frac{1}{2}} \left[\sum_{j=1}^{N-k} Y^2(j+k) - \frac{1}{N-k}\left(\sum_{j=1}^{N-k} Y(j+k)\right)^2\right]^{\frac{1}{2}}} \quad (4)$$

where:
r(k) = the lag (k) serial correlation coefficient,
N = sample size,
k = lag in time units, and
Y(j) = the homogeneous series value at time t.

For k=0, r (k) =1. In practice, k is limited by [N/4].

The non-parametric method may be used to remove the periodic component from the hydrological time series as follows:

$$Z_{(j,t)} = (Y_{(j,t)} - Avy_{(t)}) / Sdy_{(t)} \quad (5)$$

where:
$Z_{(j,t)}$= the series free from periodic component at year (j) and month (t)
$Y_{(j,t)}$= the homogeneous series
$Avy_{(t)}$ = the sample average of Y(j,t) at month (t)
$Sdy_{(t)}$= the sample standard deviation of Y(j,t) at month (t).

The resulting series, Z(j,t) is called a stochastic series. The application of Eq.(5) is also called standardization as it gives a series [Z(j,t)] with zero mean and unit variance . This series contains a dependent part which may be represented by an autoregressive model ,AR(p), moving average model , MA(q), or an autoregressive moving average model of higher order , ARMA(p, q), and an independent part that can only be described by some probability distribution function.

3.2.2 Data Normalization

Box and Cox transformation has been used to transform the series by applying Eqs. (6) and (7).

$$Z^{*} = (Z^{\lambda} - 1)/\lambda \qquad \text{when} \quad \lambda \neq 0 \tag{6}$$

and $$Z^{*} = \log(Z) \qquad \text{when} \quad \lambda = 0 \tag{7}$$

where (λ) is the transformation coefficient.

The value of the parameter (λ) is found by choosing random values between (-1 to 1) with steps 0.1 and computing the corresponding Cs and Ck values of the transformed series. For normally-distributed data, Cs=0 and Ck≈3.

Where Cs = coefficient of skewness.
Ck = coefficient of kurtosis.

3.3 The Univariate Stochastic Model

The basis of the Box-Jenkins approach for modeling time series consists of three phases: identification, estimation and testing, and application. These three basic stages have been adopted for univariate model building. The input of this analysis is the stochastic series [Z^{*}(j,t)] and the output is the independent stochastic component (ζ p,t).

3.4 Model Identification

The principal tools of model identification are the behavior of the autocorrelation function (ACF) coupled with that of the partial autocorrelation function (PACF).
Values of ACF that fall outside the 95% confidence level were significantly different from zero at the 5% level; the lower and upper limits were found by:

$$\text{Lower Confidence Limit} = \frac{-1-1.96\left[Nj-k-2\right]^{0.5}}{Nj-k-1} \tag{8}$$

$$\text{Upper Confidence Limit} = \frac{-1+1.96\left[Nj-k-2\right]^{0.5}}{Nj-k-1} \tag{9}$$

Where the value 1.96 is the z-tabulated under the normal curve and Nj is the sample size.

3.5 Autoregressive Model [AR(p)]

The general form of this linear model is:

$$E_{p,t} = \sum_{k=1}^{p} a_{k,t} \,.\, E_{p,t-k} + \sigma_{\zeta,t} \,.\, \zeta_{p,t} \tag{10}$$

where:
p = the degree of model.
$\mathbf{E_{p,t}}$ = the dependent stochastic component.

$\zeta_{p,t}$ = the independent stochastic component at year (t) and month (p).
$a_{k,t}$ and $\sigma_{\zeta,t}$ the model parameters.

Yevjevich, 1972 suggests a simplified practical method to express the goodness of fit of an autoregressive model by the determination coefficients (D_i, i=1,2,3,...), which represent the percentage of the total variance of ($E_{p,t}$) explained by the ith order term of an autoregressive equation. The remaining portion of this variance is explained by the ($\sigma_{\zeta,t}$, $\zeta_{p,t}$) term. The criterion used is as follows: the explanatory power of the (i+1)th order term must exceed that of the ith order term by at least a chosen threshold ΔD for the higher order model to be favored. Said another way, if the difference between the percentage of variance explained by the (i+1)th and ith order terms of the model, i.e.,($D_{i+1} - D_i$), is less than ΔD , then the model order (p) taken equal to (i). ΔD is usually set at 0.01, i.e., one percent of the total variance of ($E_{p,t}$).It is expected that the degree of the model (p) will not exceed three; therefore, the determination coefficients D_1, D_2, and D_3, are typically the only ones calculated. The equations of **Yevjevich, 1972** used are:

$$D_1 = r_1^2 \tag{11}$$

$$D_2 = \frac{r_1^2 + r_2^2 - 2 r_1^2 r_2}{1 - r_1^2} \tag{12}$$

$$D_3 = \frac{r_1^2 + r_2^2 + r_3^2 + 2 r_1^3 r_3 + 2 r_1^2 r_2^2 + 2 r_1 r_2^2 r_3 - 2 r_1^2 r_2 - 4 r_1 r_2 r_3 - r_1^4 - r_2^4 - r_1^2 r_3^2}{1 - 2 r_1^2 - r_2^2 + 2 r_1^2 r_2} \tag{13}$$

and the order of the model may be found through using the following steps:

1. if $D_2 - D_1 < \Delta D$ and $D_3 - D_1 < 2\Delta D$ then p=1
2. if $D_2 - D_1 \geq \Delta D$ and $D_3 - D_2 < \Delta D$ then p=2
3. if $D_2 - D_1 \geq \Delta D$ and $D_3 - D_2 > \Delta D$ then p=3

where r_1, r_2, and r_3 are the serial correlation coefficients for lags 1, 2, and 3, respectively. From these results, it can be seen that a first order autoregressive model, AR(1), fits the series, since higher degree models do not account for an increase of the explained variance of 1% or more over that explained by the first order model.

3.6 MARKOV Model (Autoregressive Model) [AR(P)]

This model describes the dependence in a hydrologic stochastic series ($E_{p,t}$) by assuming that each value ($E_{p,t}$) is the combined effect of previous values plus an independent stochastic component ($\zeta_{p,t}$), which occurs at the same time of occurrence of ($E_{p,t}$) . The independent stochastic series ($\zeta_{p,t}$) is a series of random numbers usually with zero mean and unit variance . The formulation of this model is given as, **Makridakis et al., 1998**:

$$E_{p,t} = a_1 E_{p,t-1} + \sigma \zeta_{p,t} \tag{14}$$

Where

$$a_1 = r_{1,t} \text{ or } a_1 = r_1 \quad \textbf{(15)}$$

$$s = \sqrt{1 - r_1^{2}} \quad \textbf{(16)}$$

3.7 Diagnostic Checking

Diagnostic checking means statistically verifying the adequacy of the formulated model. For this checking, the residual series is examined for any lack of randomness.

The effect of using AR (1) may be tested by finding whether the model satisfactorily removes the dependence from the stochastic variables (**$E_{p,t}$**), i.e., whether the resulting (**$\zeta_{p,t}$**) can be considered independent at a 95% confidence level. The independent stochastic series (**$\zeta_{p,t}$**) is found from Eq.(17) with a_1=r_1=0.63 as follows:

$$z_{p,t} = (E_{p,t} - a_1 E_{p,t-1}) / \sqrt{1 - a_1^{2}} \quad \textbf{(17)}$$

To test the independence of the resulting (**$\zeta_{p,t}$**) series, the ACF and PACF of this component are computed up to lag (24).

3. 8 Verification of the Model

To verify the model, 10 new sets of time series were generated. The generation procedure for the first order autoregressive model, AR(1), can be regarded as reversing the analysis procedure with slight differences, as shown by the following steps**:**

1. Generate the independent stochastic component (**$\zeta_{j,t}$**) using a pseudo-random number generator.
2. Generate the dependent stochastic component ($E_{(j,t)}$) using:

$$E_{(j,t)} = a_1 E_{(j,t-1)} + \sqrt{1 - a_1^{2}}\, z_{(j,t)} \quad \textbf{(18)}$$

3. Apply the inverse power transformation:

$$Y(j,t) = (/ * E(j,t) + 1)^{1/ /} \quad \textbf{(19)}$$

4. Standardize the new series using the monthly mean $MY_{(t)}$ and monthly standard deviation $SY_{(t)}$ at month (t):

$$\boldsymbol{Yt_{(i,j)} = (Y_{(i,j)} - MY_{(t)}) / Sy_{(t)}} \quad \textbf{(20)}$$

5. Calculate the normalized flow series X(j,t) using:

$$\boldsymbol{X_{(j,t)} = My_{(t)} + Yt_{(j,t)} * Sy_{(t)}} \quad \textbf{(21)}$$

Where **$X_{(j,t)}$** = the generated flow of month (t) and year (j).

Each time series generated covers 50 years. A comparison between the properties of the observed data and generated data is presented.

3.9 Generation of the Model

The generation procedure for the first order autoregressive model follows the same steps (1 to 5) used above in generating the verification models.

4. CALCULATIONS AND RESULTS

In creating the forecasting model for data generation, a split-sample approach – the most powerful method for testing homogeneity – was adopted**. Figs. 1 and 2** show the annual means and annual standard deviations of these time-series. The lines crossing the data points in these figures represent the averages of the annual means and standard deviations of the series. These figures show that there is no jump in the annual flow data; however, the determination coefficients reveal a trend component indicating non-homogeneity. **Table 1** presents the results of this split sample test for the recorded inflow data at the Ataturk dam. The result of testing the jump by splitting the data are shown in **Figs. 3 and 4**, which present the annual mean and standard deviations of the split data, showing that a jump component does not exist. The trend component of the considered historical data is removed by applying Eq.(3).

The test is now repeated using [Y(j,t)] as the new series for the whole dataset; the results are shown in **Figs. 5 and 6** for the annual means and standard deviations, respectively. The determination coefficients are very small for linear trends, indicating the absence of a trend component. The slopes of the lines representing the linear regression fit are small enough to be attributed to sample fluctuations. Therefore, the series may now be considered homogeneous, i.e., free of jump and trend components .The results of a split-sample test of the data after removing the jump and trend components are shown in **Table 2**.

For the detection of the periodic component, the correlogram is useful. By using Eq.(4) the serial correlation coefficients of the flow at Ataturk reservoir were calculated for lags (1 to 24). The existence of an annual cycle is evident from the occurrence of peaks in the correlogram as shown in **Fig. 7.** The high magnitude of the peak values shows that the deterministic periodic components form a dominant part of the structure of monthly flow time series at Ataturk.

The non-parametric method is used to remove the periodic component from the hydrological time series by using Eq.(5). The resulting series, Z(j,t), which is shown in **Fig. 8,** is called a stochastic series. The application of Eq.(5) is also called standardization as it gives a series [Z(j,t)] with zero mean and unit variance.

Transformation has been done to transform the series of Ataturk monthly inflows by applying Eqs. (6) and (7). **Table 3** shows the value of (λ) at which the series is normally distributed together with other parameters. **Fig. 9** shows the transformed series (or normalized series). As the flow data after removing the periodic component included negative and positive values, a value equal to 3 was added to all of the data to make the computation process easier.

The behavior of the autocorrelation function (ACF) coupled with that of the partial autocorrelation function (PACF) are the principal tools of model identification.

Figs. 10 and 11 illustrate the behavior of the ACF and the PACF of the stochastic component (or normalized flow data), [Z*(j,t)]. The lower and upper limits were found by using Eqs. (8) and (9).

From **Figs. 10 and 11,** it may be concluded that the autoregressive model [AR(1)] shows the best fit for this data set, since the ACF shows an exponential decrease and the PACF shows a cutoff after the first lag at $[\alpha_{1,1} \neq 0]$ and $[a\ k,t = 0 \quad \text{for} \quad k = 2,3,4,\ldots]$.

The determination coefficients D_1, D_2, and D_3 which express the fit of an autoregressive model were found by using, **Yevjevich, 1972** method. The Eqs. used are (11),(12), and (13) with three steps to find the order of the model. The results of the application of the equations and steps are shown below:

$r_1 = 0.63$ $r_2 = 0.41$ $r_3 = 0.37$

$D_1 = 0.39$ $D_2 = 0.39$ $D_3 = 0.41$

$D_2 – D_1 = 0.0004 < 0.01$

D3 – D1 = 0.0179< 0.02

$D_3 – D_2 = 0.0175 > 0.01$

Diagnostic checking means statistically verifying the adequacy of the formulated model. For this checking, the residual series is examined for any lack of randomness. The independent stochastic series ($\zeta_{p,t}$) is found from Eq. (17) with $a_1=r_1=0.63$.

To test the independence of the resulting ($\zeta_{p,t}$) series, the ACF and PACF of this component are computed up to lag (24) as shown in **Figs. 12 and 13**. The results show that all computed values lie inside the 95% confidence range; therefore, the series can be considered to exhibit a white noise term. Hence the diagnostic check on the AR(1)model indicates that it is verifiably adequate.

For the verification of the model, 10 new sets of time series were generated. The generation procedure for the first order autoregressive model, AR(1), can be regarded as reversing the analysis procedure with slight differences by using the five steps and the Eqs. (18) through (21).

Each time series generated covers 50 years. A comparison between the properties of the observed and generated data is presented in **Table 4**. It can be seen that the total mean flow and total standard deviation in the generated data are the same as those in the observed data. It can also be seen that maximum values in the generated flow data are larger than those in the observed data, while the minimum values in the generated data are less than those in the observed data.

Fig. 14 shows a comparison between the monthly mean flow of observed and generated data. As the two coincide, the conclusion can be made that the AR(1) model is perfectly fitted to the observed data.

The generation procedure for the first order autoregressive model follows the same steps (1 to 5) used above in generating the verification models. Ten sets of data were generated for stream flow at Ataturk using the autoregressive model AR(1). Each set was 100 years long. The properties of the generated monthly stream flow series were compared with those of the observed series.

Table 5 shows the general properties of the generated sequences. The model is capable of preserving the general means characteristics of the original series (Av), and standard deviation (S.D.).The skewness coefficient (Cs) and the kurtosis coefficient (Ck) are almost preserved.

Table 6 shows the monthly means of the generated monthly stream flow data series. The tabulated results clearly indicate that the AR(1) model preserved with a high degree of accuracy the basic statistical characteristics of the recorded data. All the monthly means of the generated data pass the (t-test) at the 95% confidence level. **Fig. 16** shows that the monthly means of the generated and observed data are the same.

Table 7 and **Fig. 17** show the monthly standard deviation of the generated data and observed data.

5. CONCLUSIONS

The negative trend in the historical monthly inflow data may due to natural reasons such as dry seasons or to man made reasons such as construction of a new dam on the river or another hydraulic structure .

From **Tables 4** through **7** it can be concluded that the univariate autoregressive model of order one [AR (1)] with constant parameters model accurately reproduced the means, standard deviations, and skewness coefficients observed in the generated records forecast at Ataturk reservoir.

6. RECOMONDATIONS

- Handling of the uncertain behavior of hydrologic variables. Soft computing systems like artificial neural networks (ANNs) or fuzzy inference system (FIS) models could be applied to the historical inflow data used herein to produce alternative reservoir inflow forecasts. The results using those forecasts could then be compared to the results of this study.
- New researches that focus on the operation of Turkish and Syrian future projects implementation on Tigris, Euphrates, Greater Zab and Less Zab River will be required to determine the future inflow and salinity at Iraqi borders.
- Optimization of the operation of all reservoirs in Iraq is necessary after operating new reservoirs in Turkey.

REFERENCES

- Al-Tikriti, H. A. 2001, *Forecasting of Pollution Levels in Accordance with Discharge Reduction in selected Area on Euphrates River*, M.Sc. Thesis, College of Engineering, University of Baghdad, Iraq.

- Cressey, G. B. 1958, *Shatt al- Arab basin-- Geological review*, Journal of Middle East, 12, pp. 448-60.

- Demir, H., Erkan, A. Z., Baysan, N., and Bilgen G. K. 2009, *The possible effects of irrigation schemes and irrigation methods on water budget and economy in Ataturk dam of south-eastern Anatolia region of Turkey*. Technological Perspective for Rational Use of Water Resources in the Mediterranean Region**,** Options Méditerranéennes, A n^o 88, 2009. GAP Şanlıurfa Tünel Çıkış Ağzı, GAP Cankaya, Ankara, Turkey.

- El-Obaidy, A.I.F. 2006, *Effect of Turkish Future Projects Implementation on Tigris River*, Ph. D. Thesis, College of Engineering, University of Baghdad, Iraq.

- Kolars John F., and Mitchell William A. 1991, *The EuphratesRiver and The Southeast Anatolia Development Project*, Southern Illinois University Press, Illinois.

- Makridakis, S., Wheelwright, S. C., and Hyndman, R. J. 1998, *Forecasting: Methods and Applications*. 3rd ed., John Wiley and Sons, Inc.

- Mujumdar, P. P. and Kumar, D. N. 1990, *Stochastic Models of Stream Flow: Some Case Studies*. Hydrological Science-Journal, Vol. 35, No. 4, PP. 395-410.

- O. Ozcana, , B. Bookhagen , N. Musaoglu, 2012, *Impact of the ATATÜRK Dam Lake on Agro-Meteorological Aspects of the Sotheastern Anatolia Region Using Remote Sensing and GIS Analysis*. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XXXIX-B8, 2012, XXII ISPRS Congress, 25 August – 01 September 2012, Melbourne, Australia.

- TDN 1988, *Turkey, Syria, and Iraq to discuss waterways*, Turkish Daily News, TDN, Ankara, 21, Nov.

- Yevjevich, W. M. 1972, *Structural Analysis of Hydrologic Time Series*, Colorado State University, Hydrology Paper No. 56, Fort Collin, Colorado.

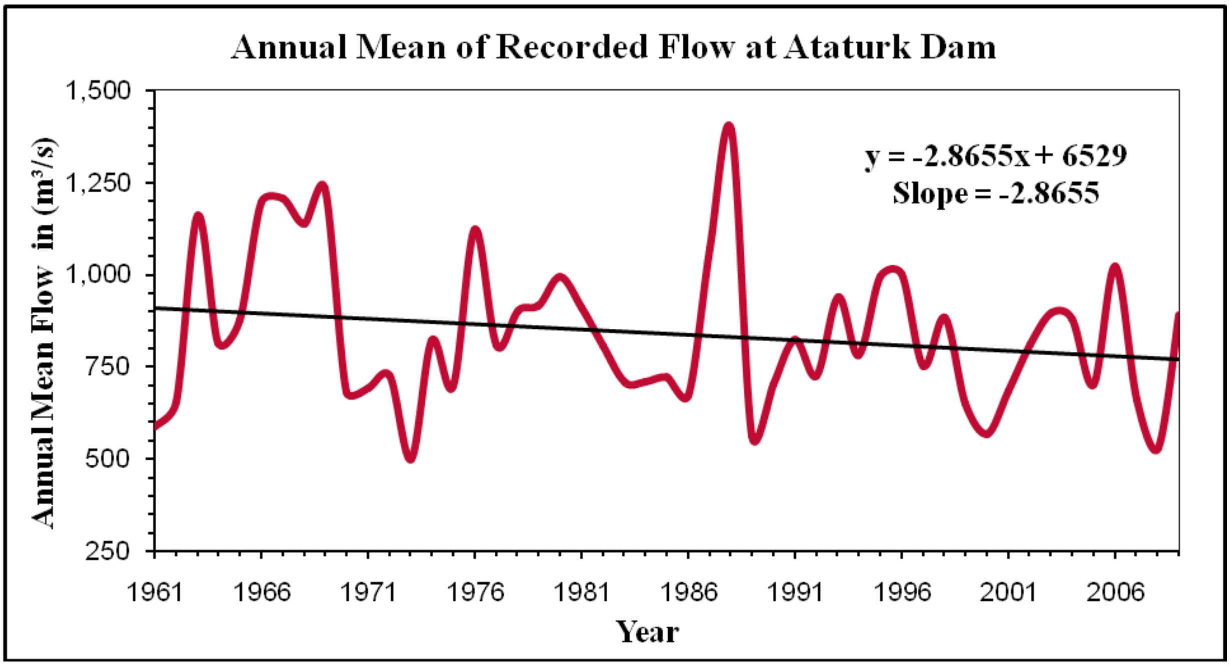

Figure 1. Annual mean flow at Ataturk dam (Oct.1961- Sep.2009), for detecting trend component.

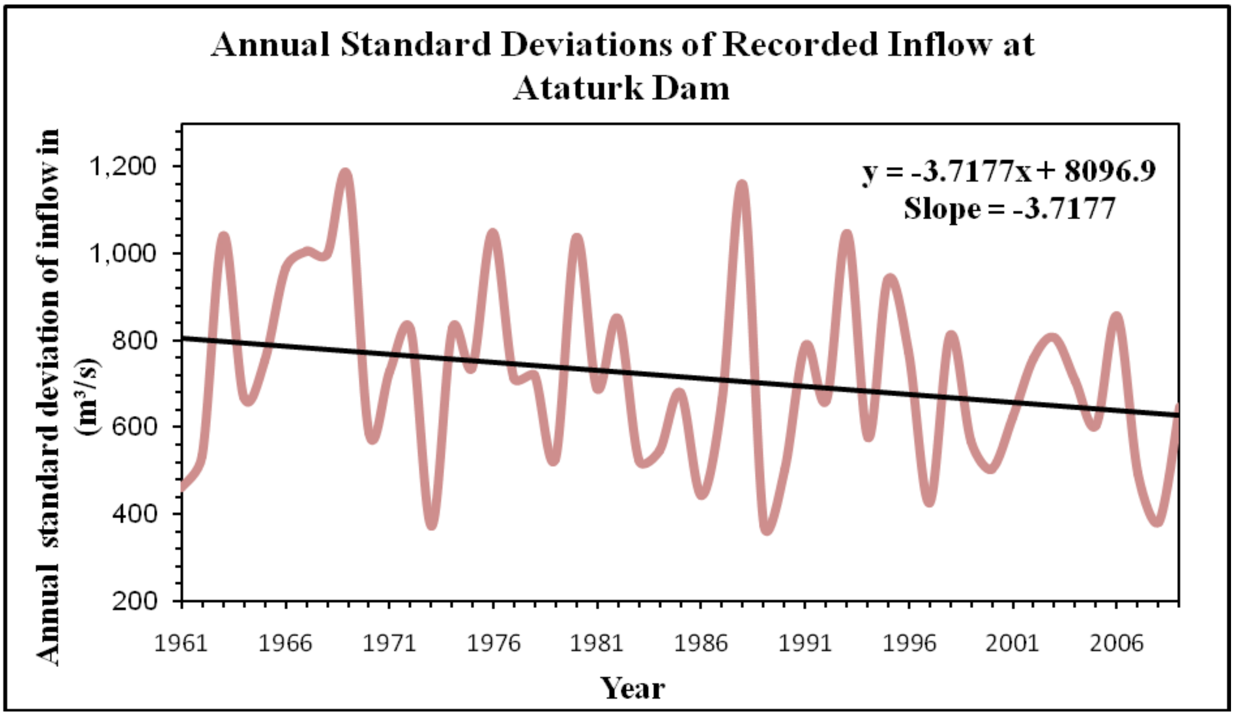

Figure 2. Annual standard deviations of the recorded inflow at Ataturk dam for (Oct.1961- Sep. 2009), for detecting trend component.

Table 1. Result of split-sample test of the Ataturk Dam's monthly recorded inflow (Oct.1961- Sep.2009).

	Statistical Parameters	**Annual Average (m^3/s)**	**Annual standard deviation (m^3/s)**
1st Period (Oct.1961-Sep.1984)	**Number of years**	24	24
	Average (m^3/s)	870	764
	Standard deviation (m^3/s)	213	217
2nd Period (Oct.1985-Sep.2009)	**Number of years**	25	25
	Average (m^3/s)	813	672
	Standard deviation (m^3/s)	192	199
	t-calculate	0.975	1.546
	t-table at 5% significance level	2.013	2.013
	F- calculate	1.229	1.187
	F-table at 5% significance level	2.014	2.003
	Jump component (t-test)	**Not exist**	**Not exist**
	Jump component (F- test)	**Not exist**	**Not exist**

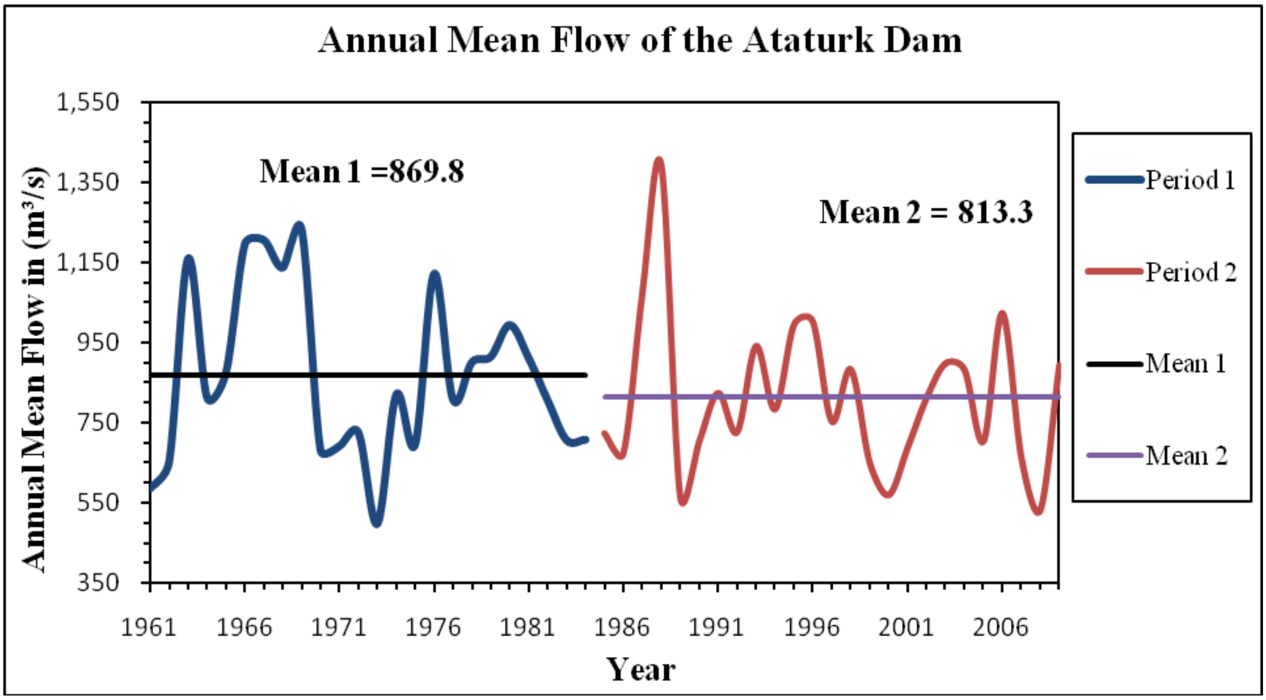

Figure 3. Annual means inflow at the Ataturk Dam (Oct.1961- Sep.2009).

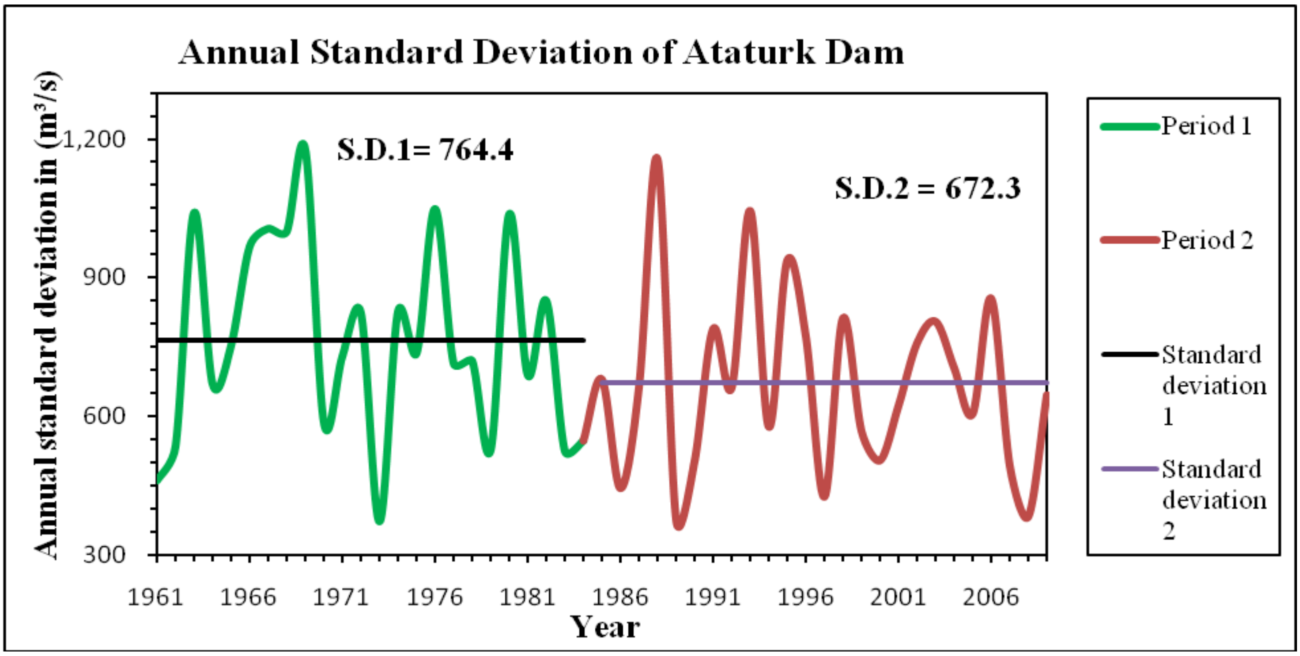

Figure 4. Annual standard deviation of the monthly recorded inflow at Ataturk Dam (Oct.1961- Sep.2009).

Table 2. Result of split-sample test of Ataturk dam's monthly recorded inflow after removing trend component.

	Statistical Parameters	Annual Average (m^3/s)	Annual standard deviation (m^3/s)
1st Period (Oct.1961-Sep.1984)	Number of years	24	24
	Average (m^3/s)	812	200
	Standard deviation (m^3/s)	54	56
2nd Period (Oct.1985-Sep.2009)	Number of years	25	25
	Average (m^3/s)	816	199
	Standard deviation(m^3/s)	55	58
	t-calculate	0.240	0.012
	t-table at 5% significance level	2.013	2.013
	F- calculate	1.041	1.070
	F-table at 5% significance level	2.014	2.003
	Jump component (t - test)	Not exist	Not exist
	Jump component (F- test)	Not exist	Not exist

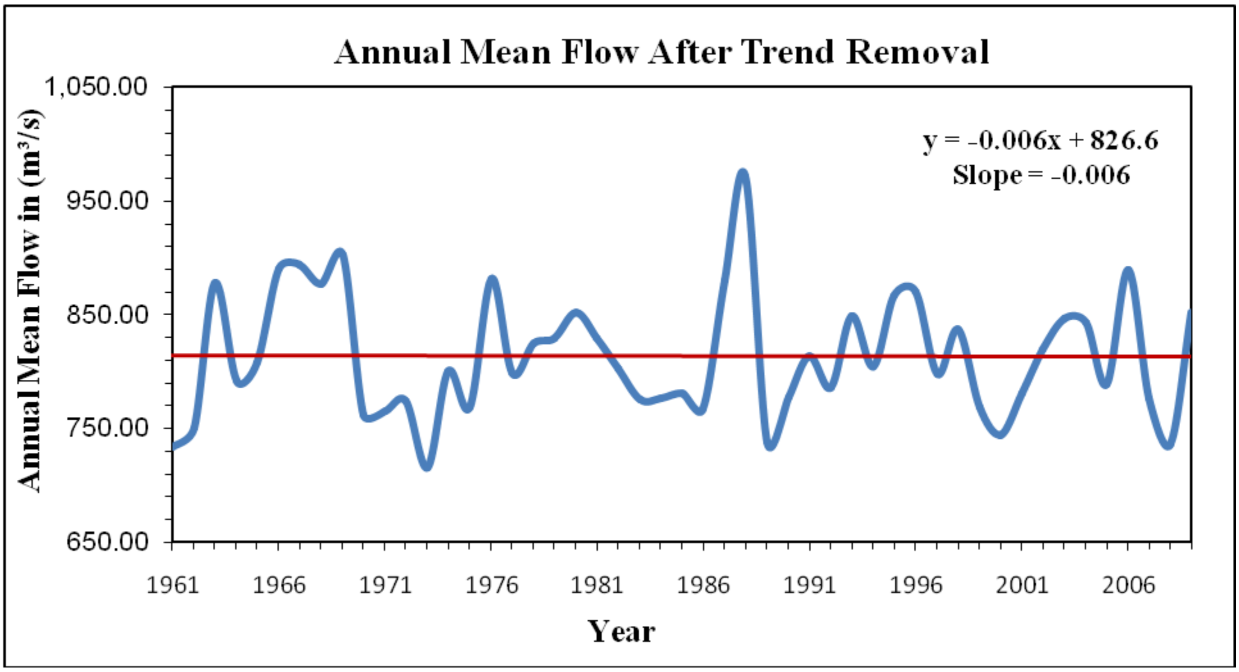

Figure 5 . Annual mean of recorded inflow at Ataturk Dam for Oct.1961- Sep.2009 after removing the trend component.

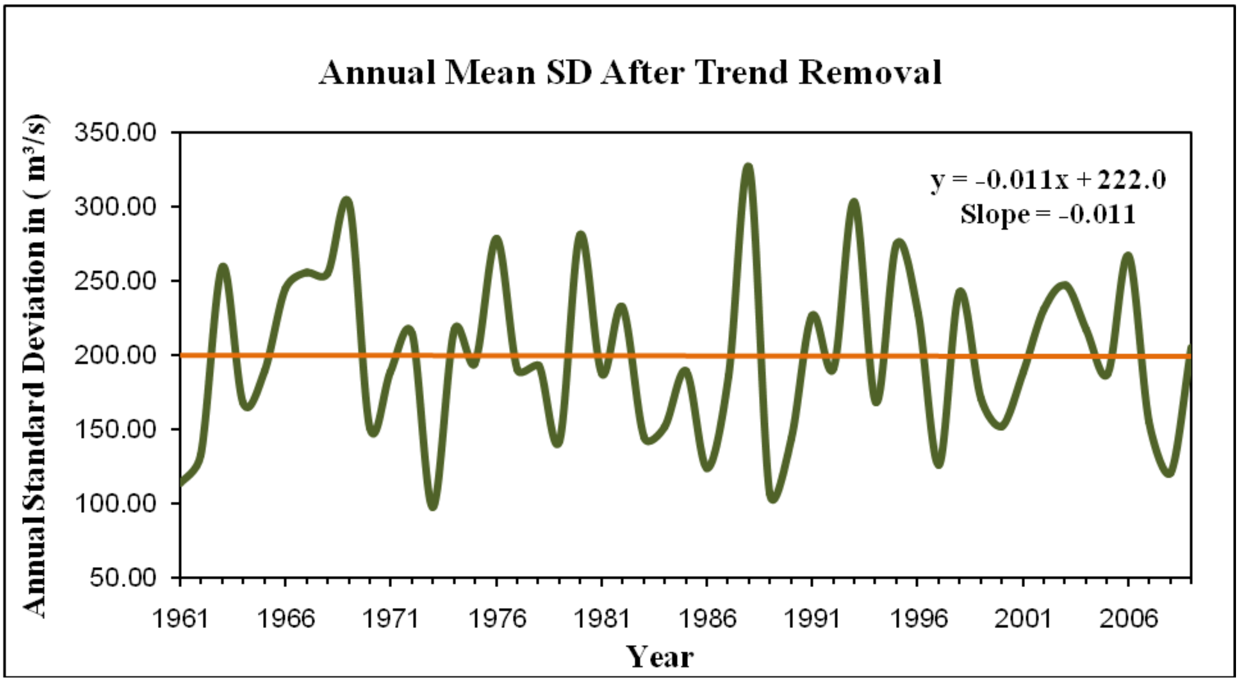

Figure 6 . Annual flow standard deviations of recorded inflow at the Ataturk Dam for Oct.1961- Sep.2009 after removing the trend component.

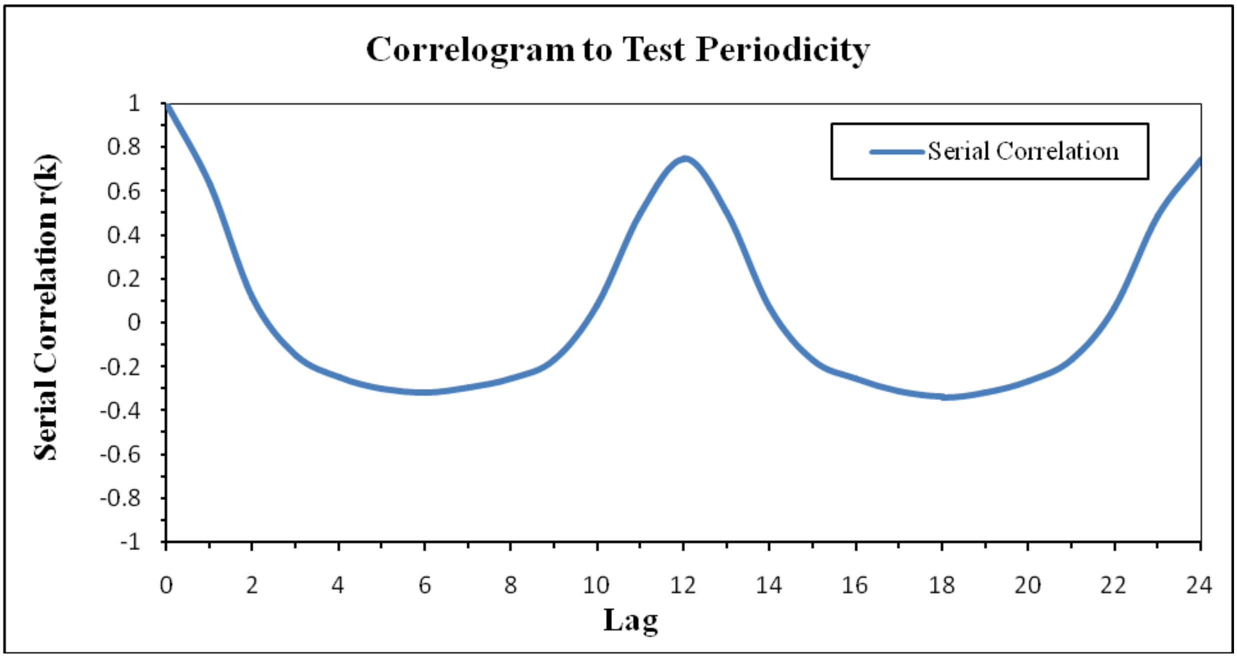

Figure 7. Serial correlation for Ataturk Dam's time series for monthly data after removing jump and trend components.

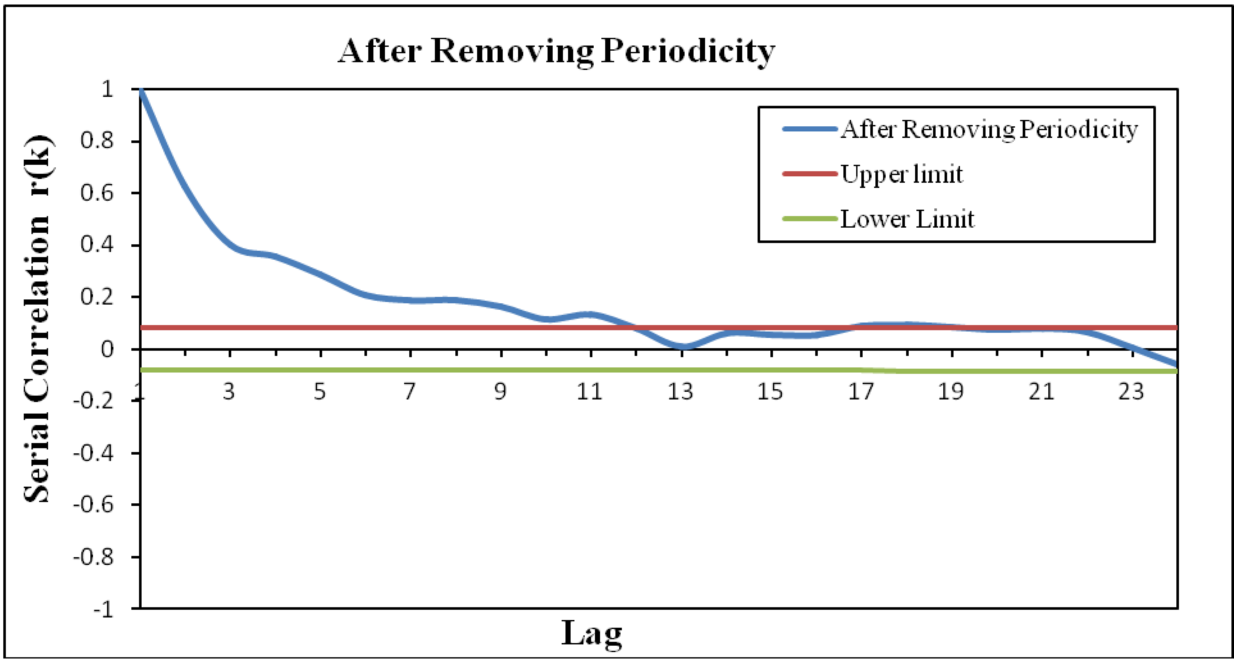

Figure 8. Ataturk correlogram for the mean monthly flow data after removing periodic component.

Table 3. Values of the statistical parameters of the recorded data and transformed data series.

Series	λ	Av.	S.D.	Cs	Ck
Recorded data	-	814	206	1.66	5.24
Transformed data	0.28	1.23	0.44	0.01	3.93

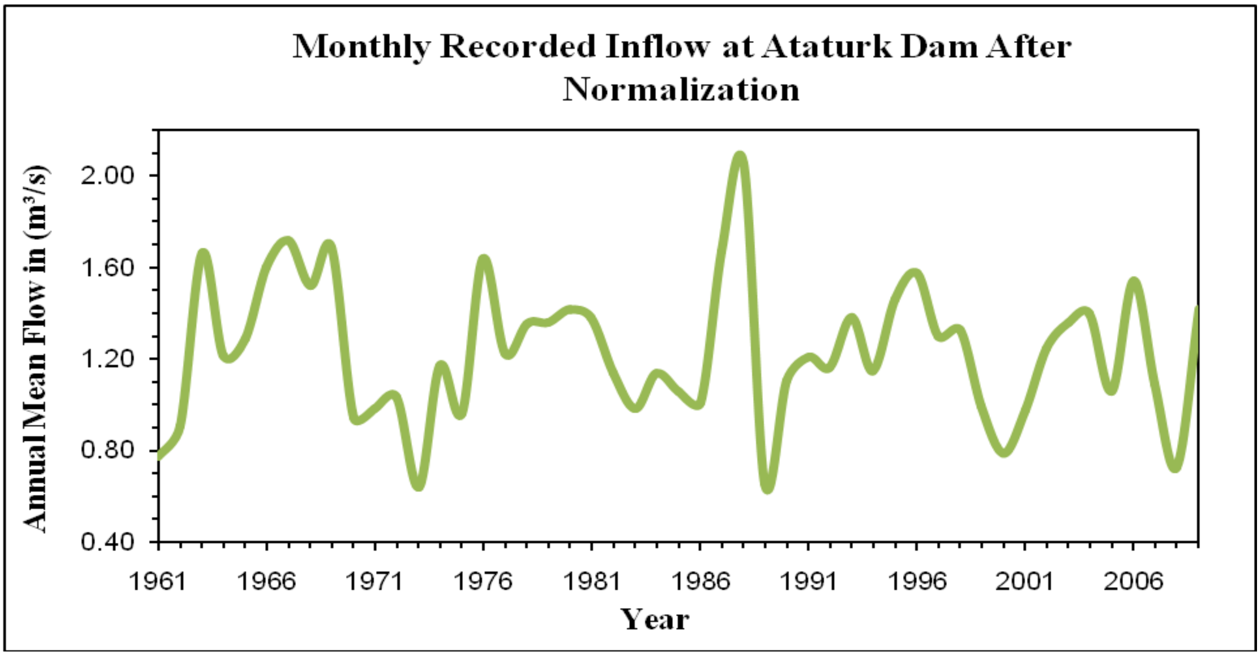

Figure 9. Annual mean flow at the Ataturk Dam after normalization.

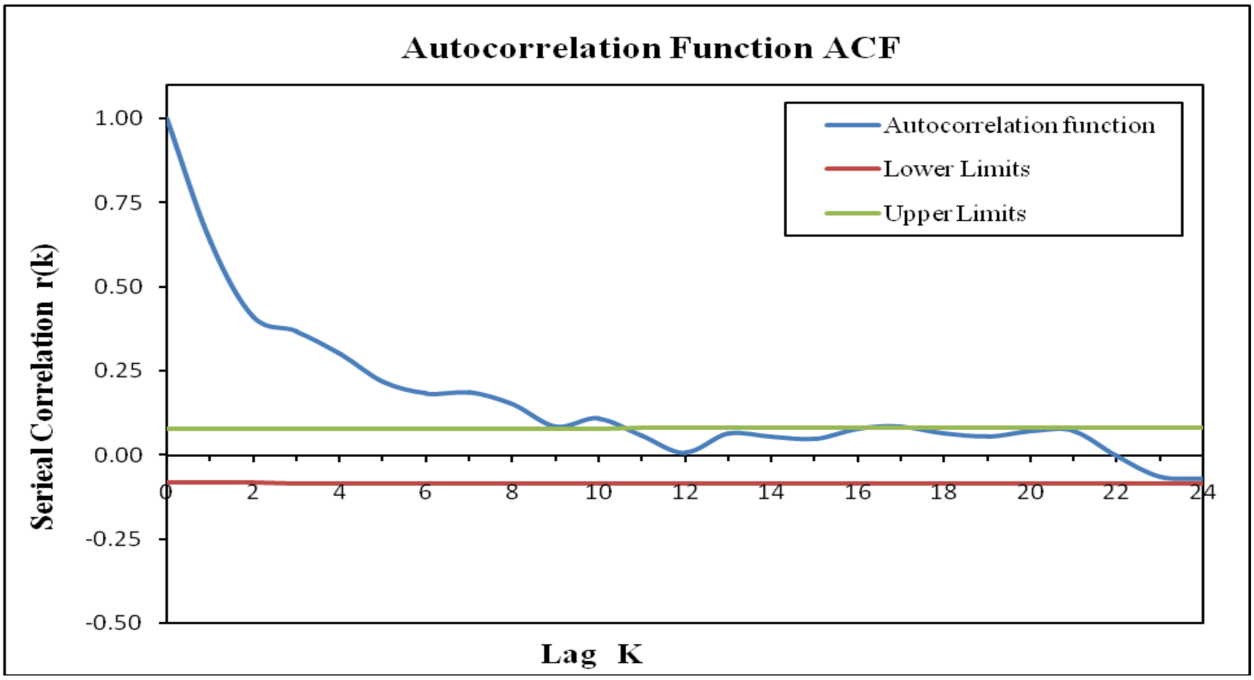

Figure 10. The Autocorrelation Function of monthly recorded inflow at the Ataturk Dam after normalization.

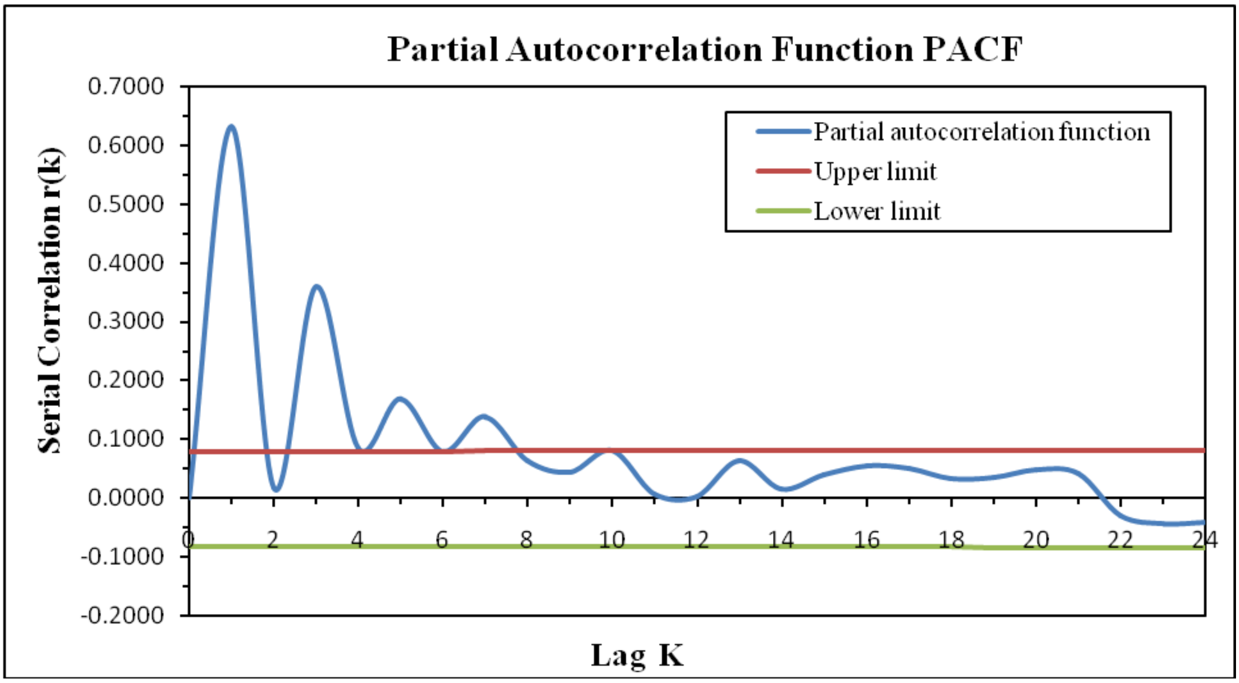

Figure 11. The Partial Autocorrelation Function of monthly recorded inflow of Ataturk Dam after normalization.

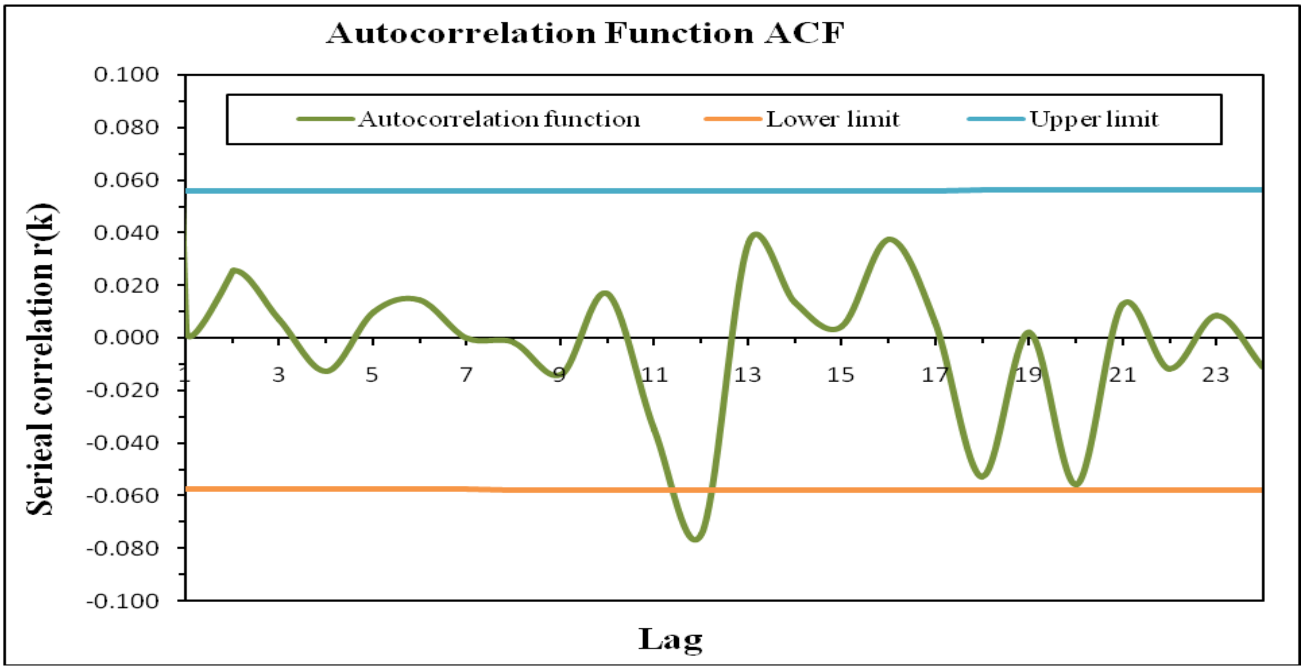

Figure 12. Autocorrelation Function of Ataturk dam for the independent stochastic component obtained by AR(1) model.

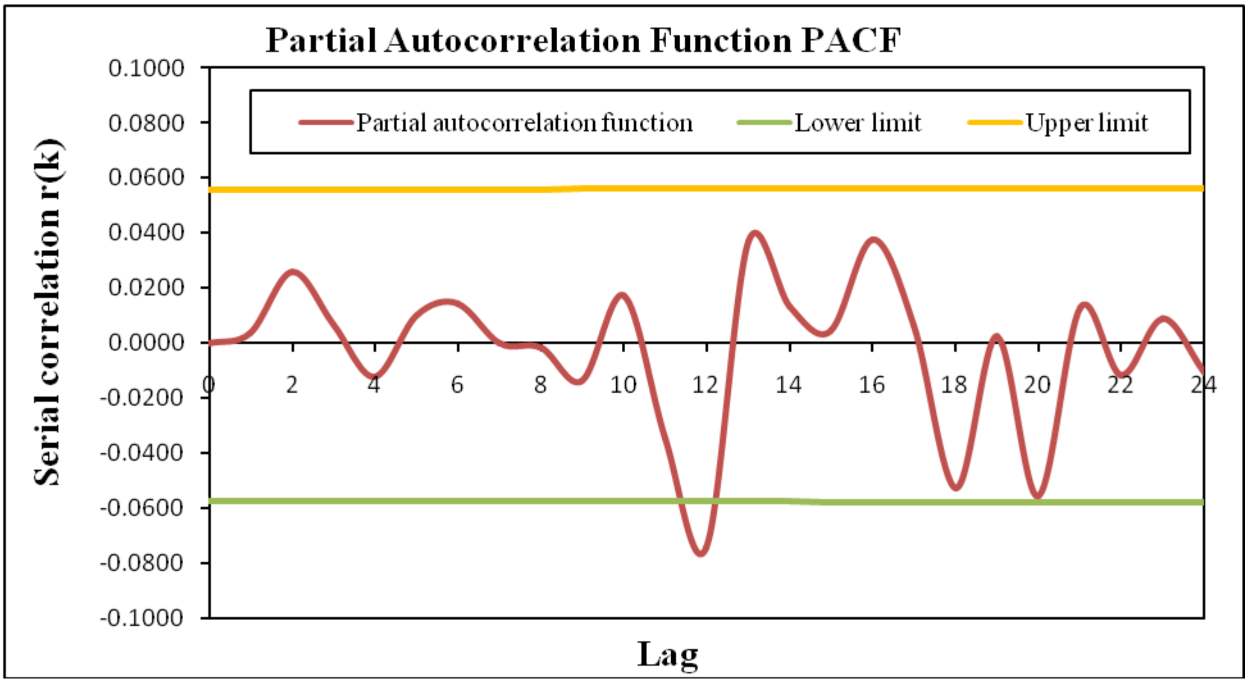

Figure 13. Partial Autocorrelation Function of Ataturk dam for the independent stochastic component obtained by AR(1) model

Table 4. Comparison between the general properties of the observed data and the generated data by AR(1) model for verification.

Series	Set	Meanm³/s	S.D. m³/s	C_s	C_k	Max. m³/s	Min. m³/s
Observed	---	813.61	205.58	1.66	5.24	1621.56	637.88
Generated	1	813.61	205.59	1.77	6.16	1766.82	620.12
Generated	2	813.61	205.59	1.78	6.26	1784.71	615.57
Generated	3	813.61	205.59	1.79	6.27	1844.67	608.82
Generated	4	813.61	205.59	1.78	6.09	1761.47	626.54
Generated	5	813.61	205.59	1.81	6.64	1954.09	603.88
Generated	6	813.61	205.59	1.71	5.69	1688.39	604.54
Generated	7	813.61	205.59	1.86	7.03	2030.68	633.04
Generated	8	813.61	205.59	1.83	6.65	1891.71	604.57
Generated	9	813.61	205.59	1.80	6.32	1806.60	613.84
Generated	10	813.61	205.59	1.95	8.01	2109.59	616.23

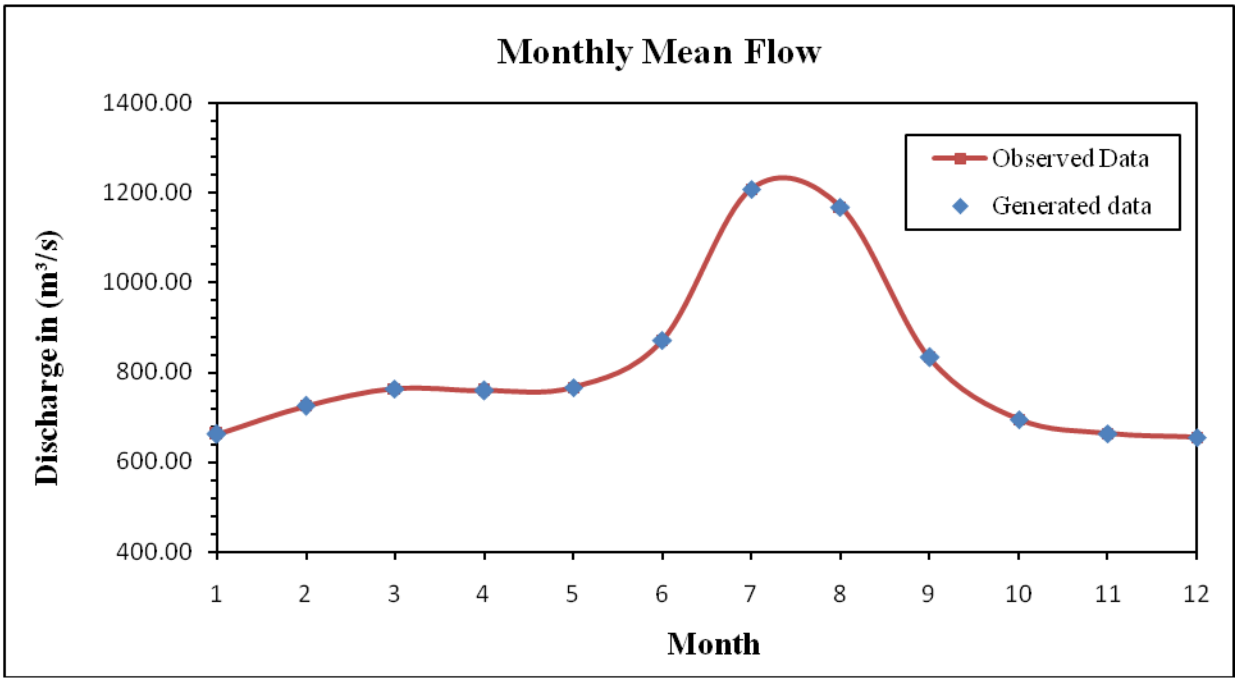

Figure 14. Comparison between the monthly averages of the observed data and the data generated by AR (1) model for verification

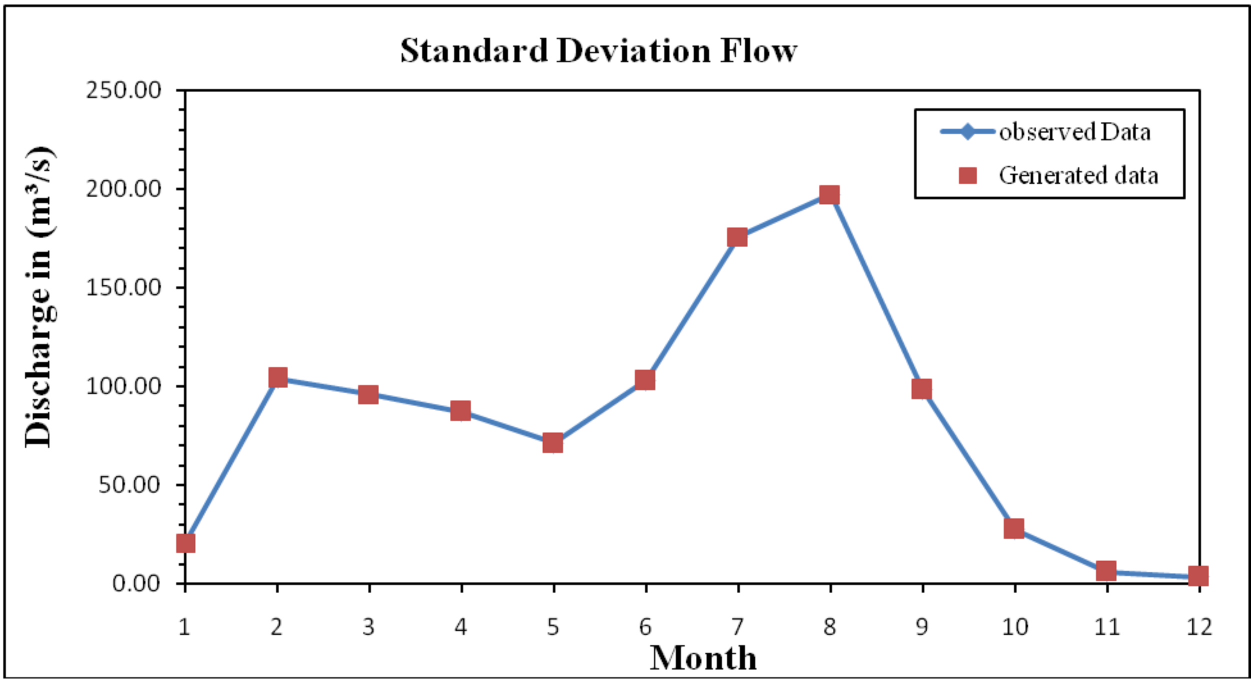

Figure 15. Comparison between the monthly standard deviations of observed and generated data by model AR (1) for verification.

Table 5. Comparison between the general properties of the observed data and the generated data by AR(1) model.

Series	Set	Mean m³/s	S.D. m³/s	C_s	C_k	Max. m³/s	Min. m³/s
Observed	---	813.62	205.58	1.66	5.24	1621.56	637.88
Generated	1	813.62	205.75	1.86	6.82	2017.45	615.12
Generated	2	813.62	205.75	1.89	7.26	2029.42	614.70
Generated	3	813.62	205.75	1.83	6.64	1912.29	612.04
Generated	4	813.62	205.75	1.82	6.55	1935.01	613.31
Generated	5	813.62	205.75	1.81	6.42	1816.65	614.63
Generated	6	813.62	205.75	1.80	6.45	1881.28	605.78
Generated	7	813.62	205.75	1.85	6.88	1993.83	623.58
Generated	8	813.62	205.75	1.78	6.23	1845.81	601.33
Generated	9	813.62	205.75	1.75	5.93	1767.78	619.12
Generated	10	813.62	205.75	1.83	6.77	2051.41	611.11

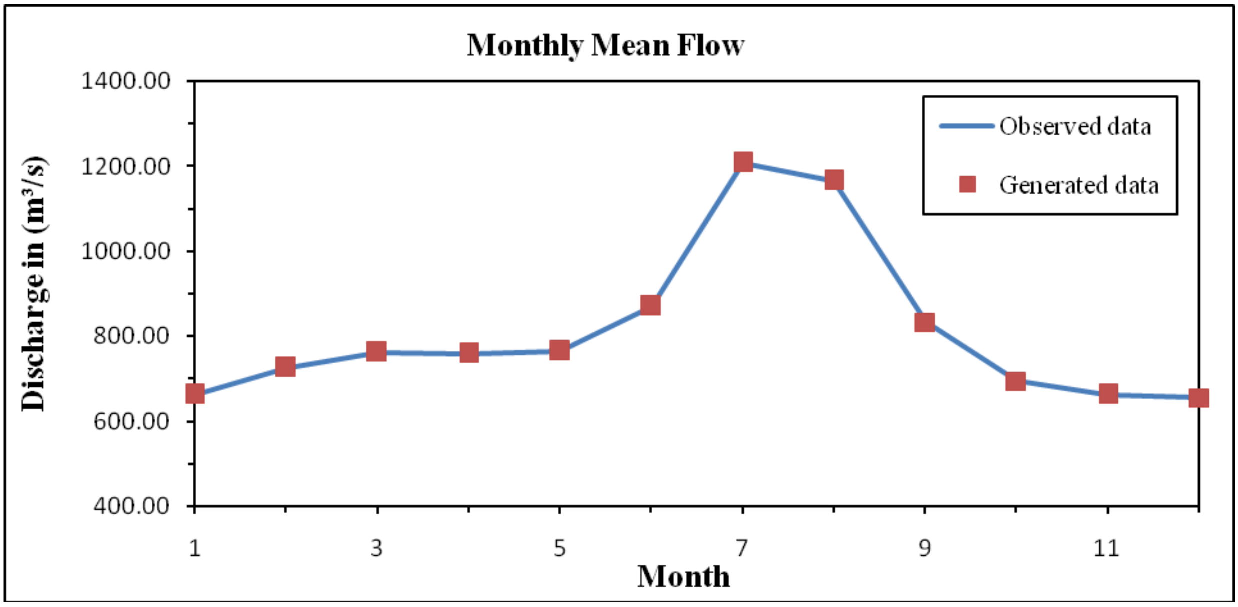

Figure 16. Comparison between the monthly averages of the observed data and the data generated by AR(1) model.

Table 6. Comparison between the monthly averages of the observed data and the data generated by AR(1) model (m^3/s).

Series	Oct	Nov	Dec	Jan	Feb.	Mar	Apr	May	Jun	Jul	Aug	Sep
Obse.	661.75	724.86	762.52	758.64	765.31	870.79	1207.67	1166.83	832.92	694.42	662.98	654.66
Gen.1	661.75	724.86	762.52	758.64	765.31	870.79	1207.67	1166.83	832.92	694.42	662.98	654.66
Gen.2	661.75	724.86	762.52	758.64	765.31	870.79	1207.67	1166.83	832.92	694.42	662.98	654.66
Gen.3	661.75	724.86	762.52	758.64	765.31	870.79	1207.67	1166.83	832.92	694.42	662.98	654.66
Gen.4	661.75	724.86	762.52	758.64	765.31	870.79	1207.67	1166.83	832.92	694.42	662.98	654.66
Gen.5	661.75	724.86	762.52	758.64	765.31	870.79	1207.67	1166.83	832.92	694.42	662.98	654.66
Gen.6	661.75	724.86	762.52	758.64	765.31	870.79	1207.67	1166.83	832.92	694.42	662.98	654.66
Gen.7	661.75	724.86	762.52	758.64	765.31	870.79	1207.67	1166.83	832.92	694.42	662.98	654.66
Gen.8	661.75	724.86	762.52	758.64	765.31	870.79	1207.67	1166.83	832.92	694.42	662.98	654.66
Gen.9	661.75	724.86	762.52	758.64	765.31	870.79	1207.67	1166.83	832.92	694.42	662.98	654.66
Gen.10	661.75	724.86	762.52	758.64	765.31	870.79	1207.67	1166.83	832.92	694.42	662.98	654.66

Table 7. Comparison between the monthly standard deviations of the observed data and the data generated by AR(1) model (m^3/s).

Series	Oct	Nov	Dec	Jan	Feb.	Mar	Apr	May	Jun	Jul	Aug	Sep
Obse.	20.88	104.20	95.93	87.42	71.24	103.03	175.76	197.22	98.25	27.52	6.22	3.59
Gen.1	20.88	104.20	95.93	87.42	71.24	103.03	175.76	197.22	98.25	27.52	6.22	3.59
Gen.2	20.88	104.20	95.93	87.42	71.24	103.03	175.76	197.22	98.25	27.52	6.22	3.59
Gen.3	20.88	104.20	95.93	87.42	71.24	103.03	175.76	197.22	98.25	27.52	6.22	3.59
Gen.4	20.88	104.20	95.93	87.42	71.24	103.03	175.76	197.22	98.25	27.52	6.22	3.59
Gen.5	20.88	104.20	95.93	87.42	71.24	103.03	175.76	197.22	98.25	27.52	6.22	3.59
Gen.6	20.88	104.20	95.93	87.42	71.24	103.03	175.76	197.22	98.25	27.52	6.22	3.59
Gen.7	20.88	104.20	95.93	87.42	71.24	103.03	175.76	197.22	98.25	27.52	6.22	3.59
Gen.8	20.88	104.20	95.93	87.42	71.24	103.03	175.76	197.22	98.25	27.52	6.22	3.59
Gen.9	20.88	104.20	95.93	87.42	71.24	103.03	175.76	197.22	98.25	27.52	6.22	3.59
Gen.10	20.88	104.20	95.93	87.42	71.24	103.03	175.76	197.22	98.25	27.52	6.22	3.59

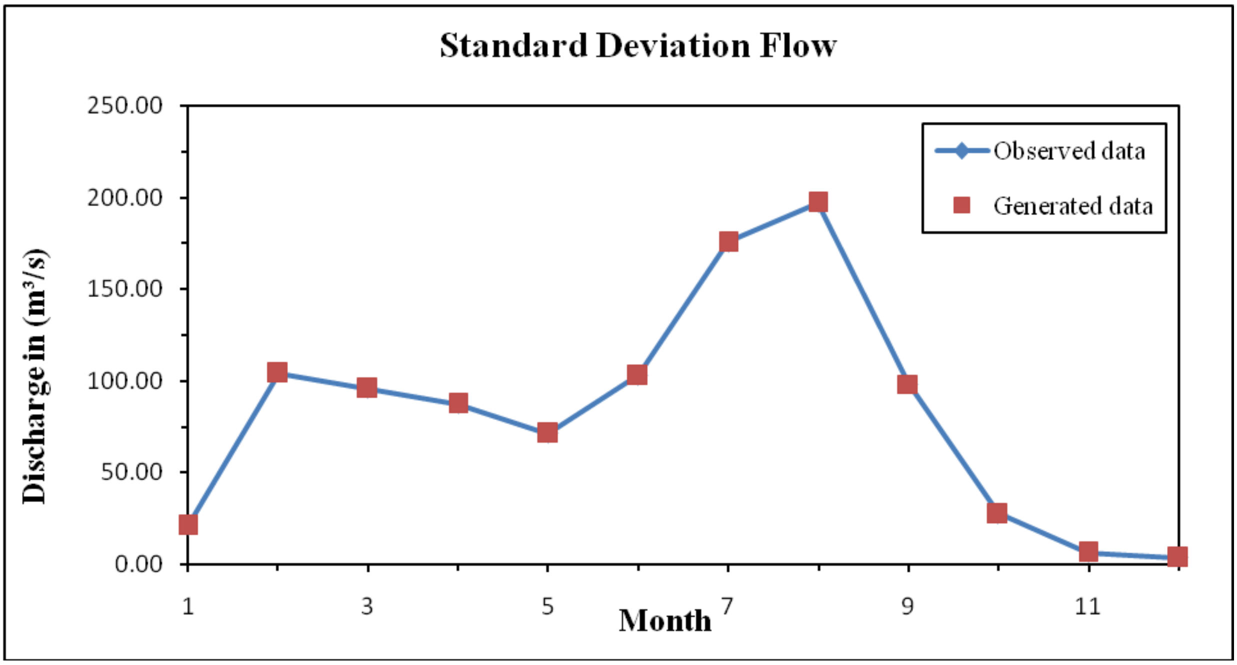

Figure 17. Comparison between the monthly standard deviations of the observed data and the data generated by AR(1) model.