RIVER NILE FLOOD FORECASTING AND ITS EFFECT ON NATIONAL PROJECTS IMPLEMENTATION

Nahla Sadek

Researcher, Nile Research Institute, NWRC, Egypt

ABSTRACT

Since long time in the history, Egyptian civilization has depended on the live–giving waters of the River Nile. It is recognized as the longest river in the world. It gets its water from Lake Victoria and Lake Albert in central Africa through the White Nile and from Lake Tana in Ethiopia through the Blue Nile. Over the last five decades River Nile has experienced various floods that had major effects on water levels and discharges upstream and downstream of Aswan High Dam (AHD). Flood forecasting plays an important role in matching the supply with the estimated demands, taking the necessary actions to avoid high flood damages, and minimizing the impacts of recurrent droughts. The purpose of this study is to predict the incoming flood during the period from 2005 to 2017. This period is selected according to the government plans and after the national projects implementation. One of the national projects is the construction of Toshka Spillway Barrage downstream at km 8.00 of the entrance to increase water management efficiency. Other projects are South Valley Project development and Salem Canal project. In this paper, the statistical forecasting approaches (ARIMA models) were used for flood predicting with analyzing the historical inflow data. It is applied time serious analysis using the historical monthly data. Five successive years were used for this analysis from 1999/2000 until 2004/2005 to prove the model suitability in forecasting the incoming floods to the Lake Nasser upstream the high dam. The analysis results were illustrated during this study.

Keywords: River Nile – Flood forecasting – Statistical models

RIVER NILE

The River Nile has a length of 6,825 km approximately only about 1,530 km lie within the borders of Egypt's territory. Throughout its length in Egypt; it receives no single tributary, Shahin, 1985 [1]. It is considered the main source of water in Egypt. Deep groundwater aquifers in the desert, and scattered rainfall areas are auxiliary water sources. The Nile catchment's area is about three million square kilometers in ten countries with a variety of different geographic and climatic characteristics. The main water supply sources are the equatorial lakes, Bahr El Gazal water shed and the Ethiopian Plateau. The Nile River has under gone a wide variety of the floods over history. The inflow is ranging from a maximum value of 150 billion m^3/y (1878-1879) to a minimum value of 42 billion m^3/y (1913-1914). Statistical analysis for historical annual inflow to study the critical cases of successive floods and their probability of occurrence was analyzed and presented by Sadek and Aziz, 2005 [2]. Flood forecasting is the best tool for management of the incoming flood successfully to avoid the side effects resulting from severe floods and droughts.

FLOOD ANALYSIS

Monthly inflow records at Aswan have been collected in order to get an idea about the high and low flows during the period from 1900/1901 to 2003/2004. The available data consists of 104 years of monthly discharge from 1900 to 1929 the data have been published by, the Ministry of water resources and irrigation (MWRI, 1933 [3] and by the Water Master Plan to 1975 WMP, 1980 [4]. Data recording has continued under the Planning Sector of the MWRI to the present time. Table 1 shows the historical data statistical analysis of monthly inflow at Aswan. This analysis represents in the arithmetic mean, the median (center value), standard deviation (STDV), variance, minimum and maximum values. From this table it can be noticed that the hydrological cycle begins with the first rains of July; the flow then increases until September with an average monthly flow of 19 billion m^3/ month in September. April, May and June are substantially drier. A sharp drop corresponding to the dry season occurs in November and the minimum flow occurs in May with an average inflow of 2.32 billion cubic meters. In addition, it could be inferred that the low flow records may occur from March to June. Also, this table shows the monthly average inflow is almost the same for all the months except that for April and May.

Table 1: Historical Data Statistical Analysis of Flow at Aswan

Statistics	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Mean	3.97	2.94	2.48	2.41	2.32	2.38	5.54	18.04	19.90	13.03	7.33	4.97
Median	3.94	2.93	2.36	1.70	1.73	2.40	5.43	18.21	20.02	12.75	7.21	4.90
STDV	0.77	0.82	0.87	1.15	1.08	0.75	1.79	3.85	4.47	3.65	1.85	0.90
Variance	0.59	0.68	0.75	1.33	1.16	0.57	3.20	14.86	20.00	13.29	3.42	0.82
Minimum	1.72	1.15	1.07	0.95	0.80	0.90	1.74	6.50	7.31	5.97	3.94	2.83
Maximum	6.57	6.04	5.81	5.26	4.72	4.91	11.03	28.28	32.79	22.80	13.30	7.88

(Flows expressed in billion cubic meters)

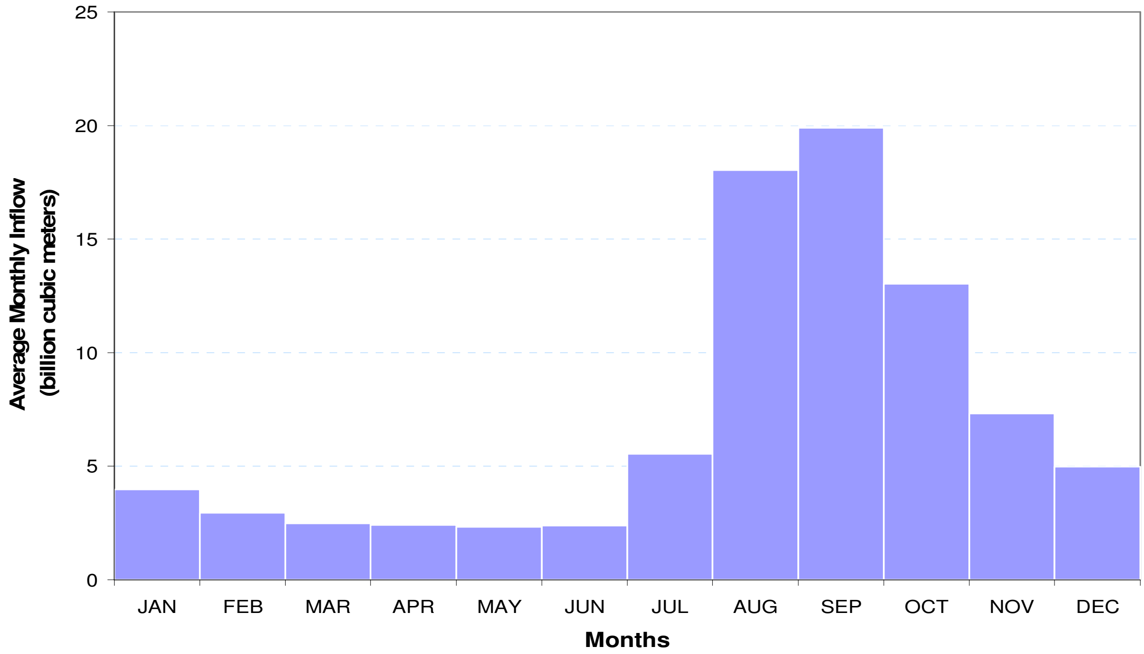

Figure 1: River Nile Average Monthly Inflow at Aswan from year 1900 to 2004

FLOOD FORECASTING STATISTICAL MODEL

Flood forecasting plays an important role in matching the supply with the estimated demands, taking the necessary actions to avoid high flood damages (excessive bank erosion, floodplain inundation and damaging hydraulic structures) and minimizing the impacts of recurrent droughts (navigation bottleneck problems – water requirements).

Several flood forecasting statistical models have been developed for forecasting River Nile stream flow such as Neural Network Models, Auto-regressive Models, Recursive Least Square Model, and many others. Comparing the model results with the actual records for different flood, should be performed to find out which model (or models) can be used in flood forecasting and which parameters should be used for this model to give satisfactory results. Some of these flood forecasting techniques were presented by Aziz et al. [5]. Also, different equations, relating monthly inflow and the annual inflow were proposed by Sadek and Aziz, 2005 [2].

In this study the statistical forecasting approaches (ARIMA models) are used. It is applied time series analysis and forecasting. Time series analysis assumes that the data consists of a (usually a set of deterministic components) and random noise (error). Systematic pattern series can be thought of as consisting of three different components: trend component T; seasonal component S; and cyclical component C. Trend represents a general systematic linear or (most often) non-linear component that monotonically changes over time and does not repeat itself within the time range captured by our data. The difference between a cyclic and a seasonal component is that the latter occurs at regular (seasonal) intervals, while cyclic factors have usually a longer duration that varies from cycle to cycle, Box and Jenkins, 1976 [6].

Most time series analysis techniques typically utilize either an additive or multiplicative formula to decompose the time series into its basic components. The specific functional relationship of the additive model is:

$$X_t = T_t + C_t + S_t + I_t$$

while the multiplicative models

$$X_t = T_t * C_t * S_t * I_t$$

The family of usable models is provided by the theory of stochastic processes and is known under the name of ARIMA models. In such models, the value of the series at a given moment is determined by several components:

First, what is called 'auto-regressive' – the AR of ARIMA – which means that the value is determined by a relation with the previous values of the series. The number of previous values (or lags) used in this relation is said to be the order of the AR part. It is often denoted by the letter p, in the models used, it can vary between 0 and 3.

Second, what is called 'moving average' – or MA. This is a random component. It brings in the value of a noise variable, but this variable can be a structure-less noise, or white noise, which means that the different values of the noise are independent from each other, or, to the contrary, possess a certain structure: the successive random terms are correlated with one another. The temporal extent of this structure is described by the order of the MA part, denoted by q. It indicates the number of lags to use for determining the random part from a white noise. Usually, its value is between 0 and 3.

Third, the I, for integrated, indicates whether the series itself should be used or its variations (or derivatives). The differential order, denoted by d, indicates whether the model simply uses the variable itself (d = 0), if it must contain its derivative (its variations, d = 1) or its second derivative (d = 2, maximum value).

Integrating seasonal variations in ARIMA models is done by building up another seasonal model using the basic model. The basic model deals with expressed lags as a number of observed values, in months for monthly data and in quarters for quarterly data. The seasonal model describes the link between the value of the series at a given moment and its value in the previous year. The lags are now expressed as a number of years.

To sum up, the type of model used to describe the series under analysis is described or specified by the two triples data (p, d, q) (sp, sd, sq). The selection phase of a model is called model identification. Afterwards, the ARIMA methods must estimate the coefficients of this model. The number of coefficients to calculate corresponds exactly to the orders retained, excluding differencing.

One of the difficulties in ARIMA type methods stems from the sensitivity of the method to disturbances. In fact, if a time series is subject to events or breaks, they are going to have a strong influence on the calculations necessary for identifying the model and estimating the parameters. It is therefore very important to set up a fine detection mechanism for such disturbances in order to guarantee the quality of the model selected. So, in this study the monthly inflow behaviour are analyzed. Figure 2 shows the monthly inflow variations with time for April and May months as an example. From this figure, it can be noticed that the flow records before mid sixties could belong to one distribution and after this time the data belong to other distribution. Therefore, mid sixties is considered the point of change in the flow behavior. Many researchers have studied the cause of this change. The most accepted explanation is the increase of Lake Victoria water levels due to rain increasing and climatic changes. This level increase reached about 2.2 m during the period 1961-1964, Saied, 1993 [7], and also stated by Aziz, 2005 [8]. Consequently, the series from 1964 up to 2004 is selected to time series analysis to guarantee the quality of the model selected. The time series analysis was done using DEMETRA software.

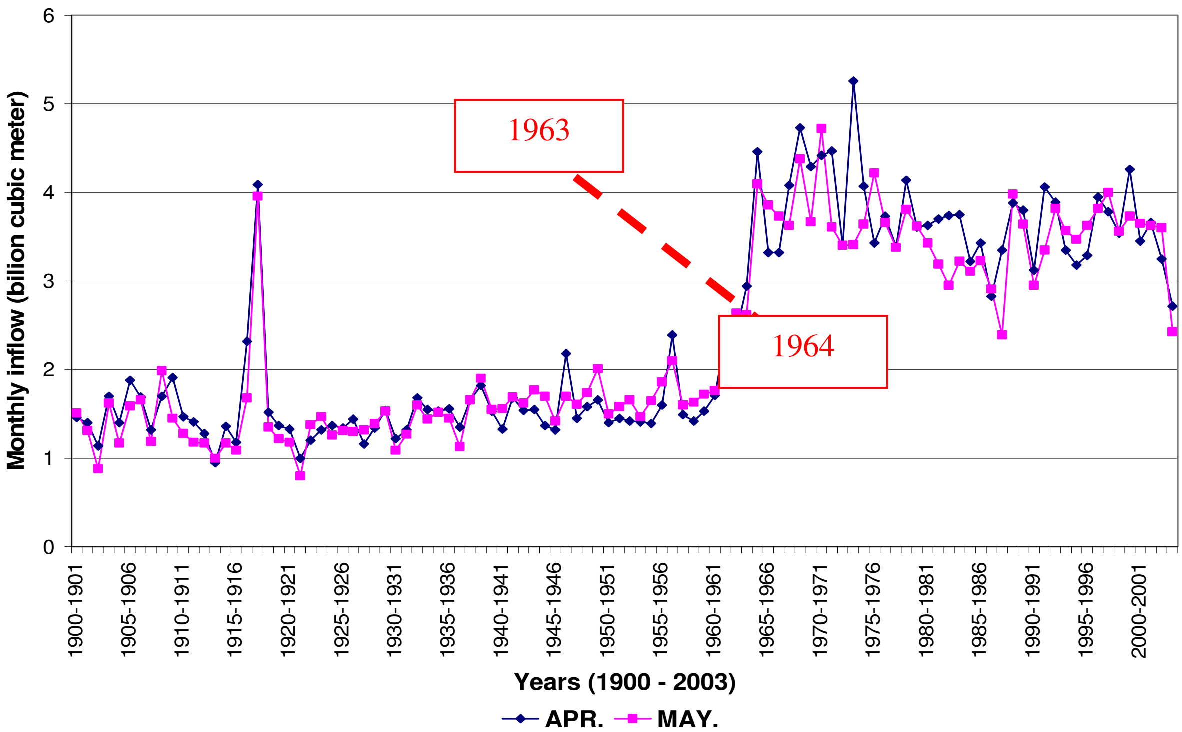

Figure 2: Monthly Inflow Variations with Time

The recorded natural river for the period from 1900 until 2004 were collected and used for reasonably prediction for the future period until 2017. One of the most important tests to determine the ARIMA model; for forecasting the new data is the colerogram figures, (Auto Correlation Function (ACF), Partial Auto Correlation function (PACF)), the main feature of the auto correlation function is to determine the degree of the autoregressive model, while the main function of the PACF is to determine the degree of the moving average model (MA).

Figures 3 and 4 present the order of the autoregressive and moving average models, p = moving average order is equal to 1, while the q = autoregressive order is equal to 1.

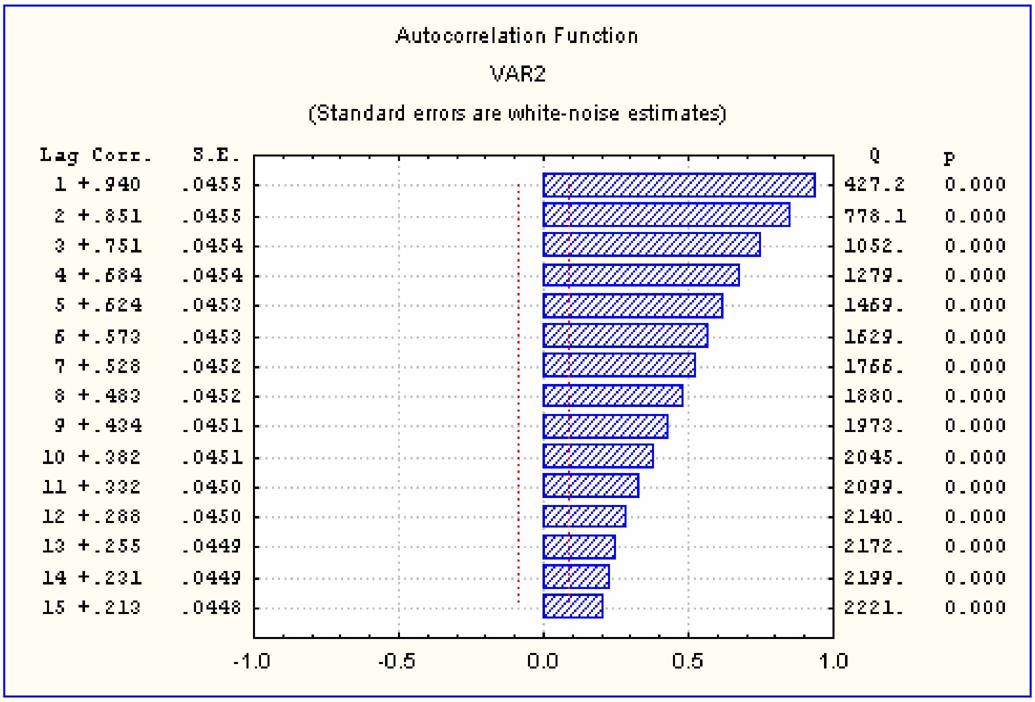

Figure 3: Auto Correlation Function (1964 – 2003)

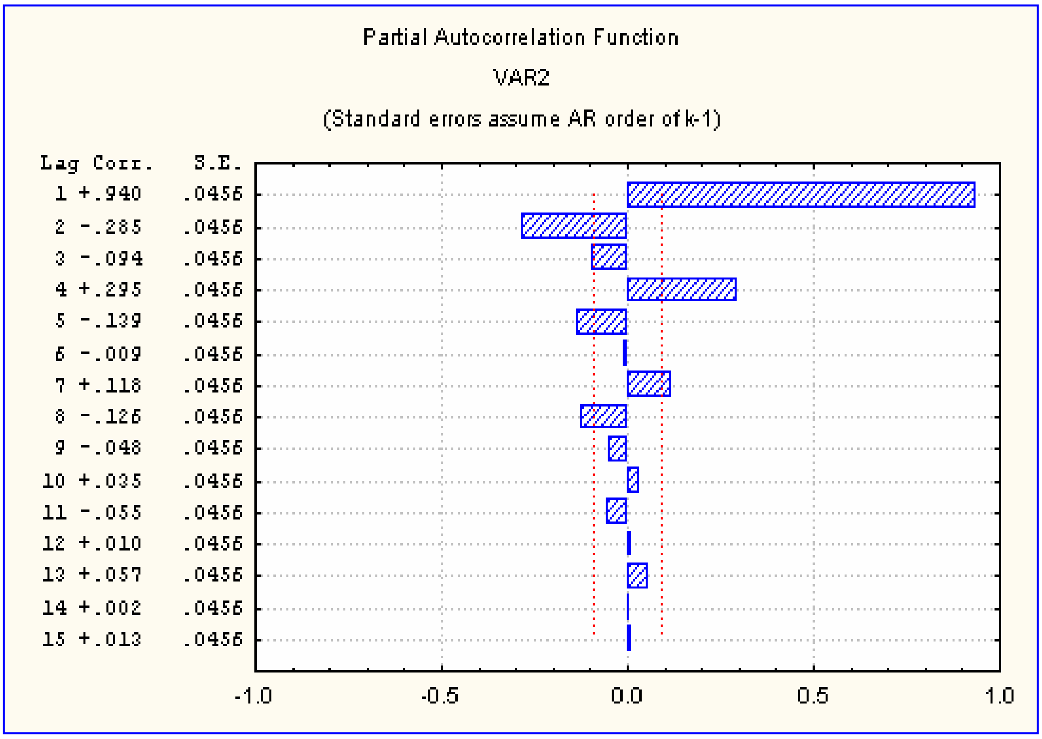

Figure 4: Partial Auto Correlation Function (1964 – 2003)

Figure 5 shows the inflow time series as well as the final seasonality adjusted series after removing the white noise error, and the final trend series. The series span as well as the model spans starts by January 1964 up to December 2003.

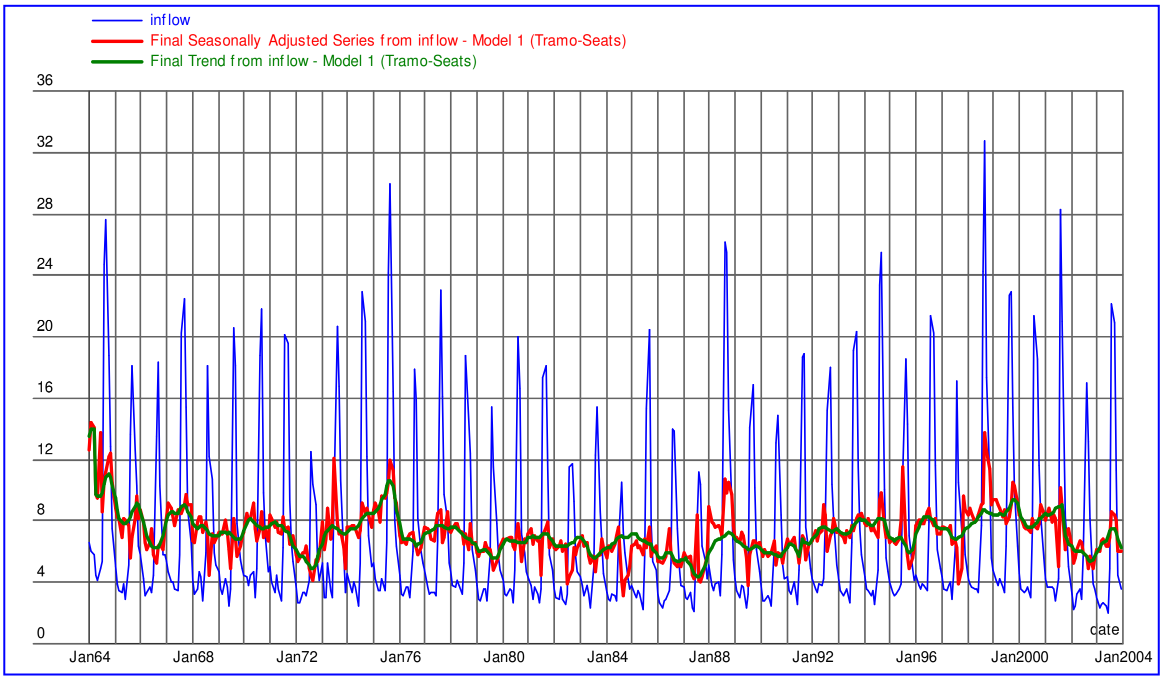

Figure 5: Adjusted Time Series

A full description of the model used is presented in table 2.

Table 2: ARIMA Model Specifications

ARIMA model Specifications	
Pre-Adjustment	
Transformation	Logarithm
Mean Correction	Yes
Mean value	-0.0087
Mean standard error	0.0031
Mean t-value	-2.81 [-1.965, 1.965] 5%
Specif. of the **ARIMA** model	(1 0 0) (0 1 1) (fixed)
Non-seas. AR (lag 1) value (Φ)	-0.5694
Seasonal MA (lag 12) value Seasonal (θ)	-0.8625
Method of Estimation	Exact Maximum Likelihood

The general form for the ARIMA (p,d,q) model can be written as follow:

$$u_t = \sum_{j=1}^{p} \phi_j u_{t-j} - \sum_{j=0}^{q} \theta_j \varepsilon_{t-1}$$

where:

Φ & θ are the model parameters
ε is the error in estimation

MODEL RESULTS

The results of the model were analyzed and evaluated in this study. The minimum – maximum and expected monthly flow is computed during the analysis. The evaluation results include revising the actual recorded monthly flow data with the expected flow through the five consecutive years from 1999/2000 to 2003/2004 which have been presented in the table 3.

From this table it can be concluded that very good agreement occurs between the actual and estimated values of monthly flow and the difference in the extreme and the mean results could be accepted and reasoned to the precision of the field measurements. Some factors indicate the accuracy of the forecasting method is shown in table 4. These factors include Pearson correlation and the rate of change between the actual and estimated annual flow data. Also, the probability value of t critical for two-tail is determined using t-test in this analysis. TTEST use to determine whether two samples are likely to have come from the same two underlying populations that have the same mean. Pearson correlation factor is considered highly correlated because its value is ranging between 96% and 99%. The rate of change between the actual and estimated annual flow is small and it can be neglected except at year 2002/2003. Also the value of t- critical is less 0.05 at the same year.

It is worth noting that year 2002/2003 came after four consecutive years of high flood which started from 1997. These years have relatively higher values in flood months and in most of the cases the results occurred after the sudden change in the historical records, sudden increase or decrease, are affected by the previous data consequently affecting the model results.

The forecasting results for each year from 1999/2000 to 2003/2004 are shown in Figures 6 to 11. Each figure shows the minimum - maximum and predicted flow for the related year and it also shows the actual recorded flow for each month. From these figures it can be concluded that the actual flow is most of the time located between the predicted minimum and maximum flow.

All the previous analyses have proved the model suitability in forecasting the incoming floods to the lake upstream the high dam. Consequently, the statistical ARIMA models are used to achieve the objective in this study.

Table 3: Actual and Estimated Monthly Flow Data

Month	Monthly inflow (actual - estimated bcm/month)									
	1999/2000		2000/2001		2001/2002		2002/2003		2003/2004	
	Act.	Est.	Act.	Est.	Act.	Est.	Act.	Est.	Act.	Est.
Aug.	22.65	22.59	21.36	18.64	28.28	23.15	17.02	18.7	22.10	20.12
Sep.	22.95	26.10	18.54	23.15	21.07	24.14	13.14	16.2	20.91	19.93
Oct.	15.38	13.27	11.58	12.74	8.26	12.97	8.20	11.1	11.24	10.76
Nov.	10.03	8.62	8.42	8.82	6.92	8.43	3.94	6.85	4.50	7.02
Dec.	6.34	6.39	5.56	6.41	5.05	5.98	3.63	4.59	3.61	5.24
Jan.	4.82	5.38	4.58	5.35	4.30	4.96	3.28	3.90	3.05	4.45
Feb.	4.16	4.44	3.61	4.47	3.69	4.10	2.20	3.23	2.32	3.73
March	3.83	4.18	3.33	4.15	3.69	3.84	2.42	2.72	2.58	3.43
April	4.26	4.51	3.45	4.55	3.66	4.13	3.25	3.06	2.72	3.66
May	3.73	4.52	3.65	4.47	3.63	4.11	3.60	3.21	2.43	3.68
June	3.32	3.73	3.07	3.88	2.82	3.55	2.87	3.19	2.05	3.28
July	6.35	8.39	7.38	8.75	4.23	7.65	5.77	7.47	6.86	8.52
Mean monthly for year	8.99	9.34	8.99	7.88	8.78	7.97	5.78	7.02	6.86	8.52

Table 4: Statistical Parameters for the Forecasting Results

Year	1999/2000	2000/2001	2001/2002	2002/2003	2003/2004
Variance (actual annual flow)	53.35	38.67	65.62	22.21	52.97
Variance (estimated annual flow)	56.94	39.75	54.61	30.06	37.92
Actual annual inflow	107.82	94.53	95.60	69.32	82.78
Estimated annual inflow	112.12	105.38	107.00	84.18	91.99
The rate of change (Act.-Est.)	4%	11%	12%	20%	12%
Pearson Correlation	98%	97%	96%	98%	99%
P(T – critical (two-trail) >0.05	0.38	0.07	0.20	0.005	0.05

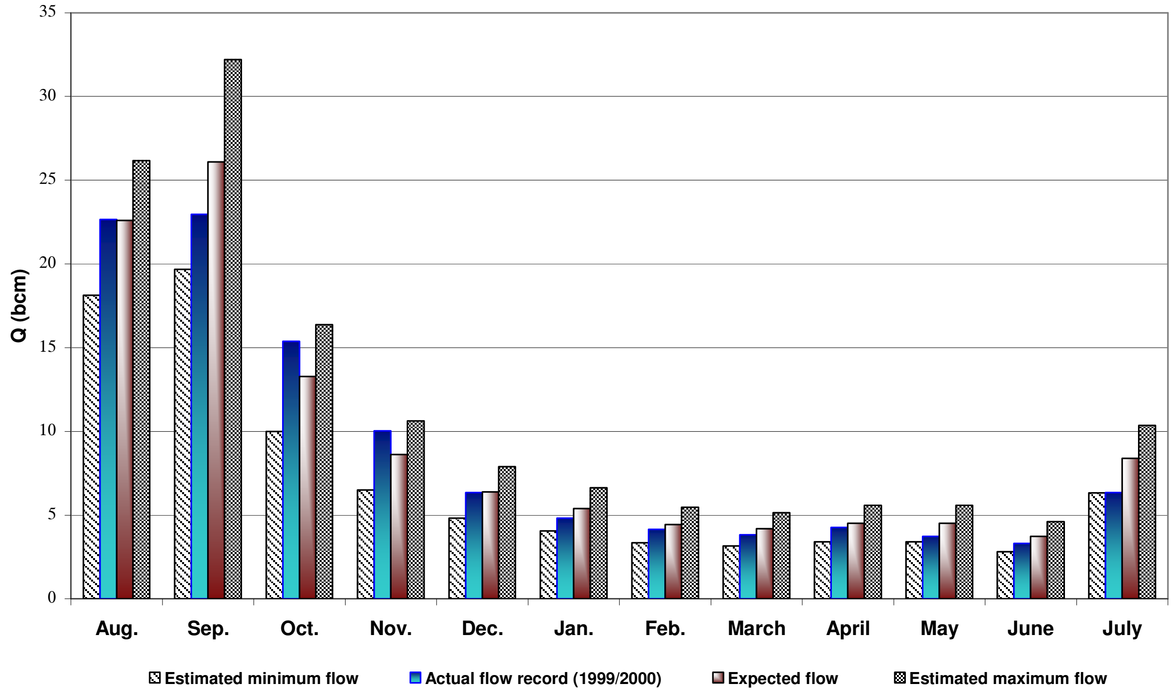

Figure 6: Flood Forecasting For 1999/2000

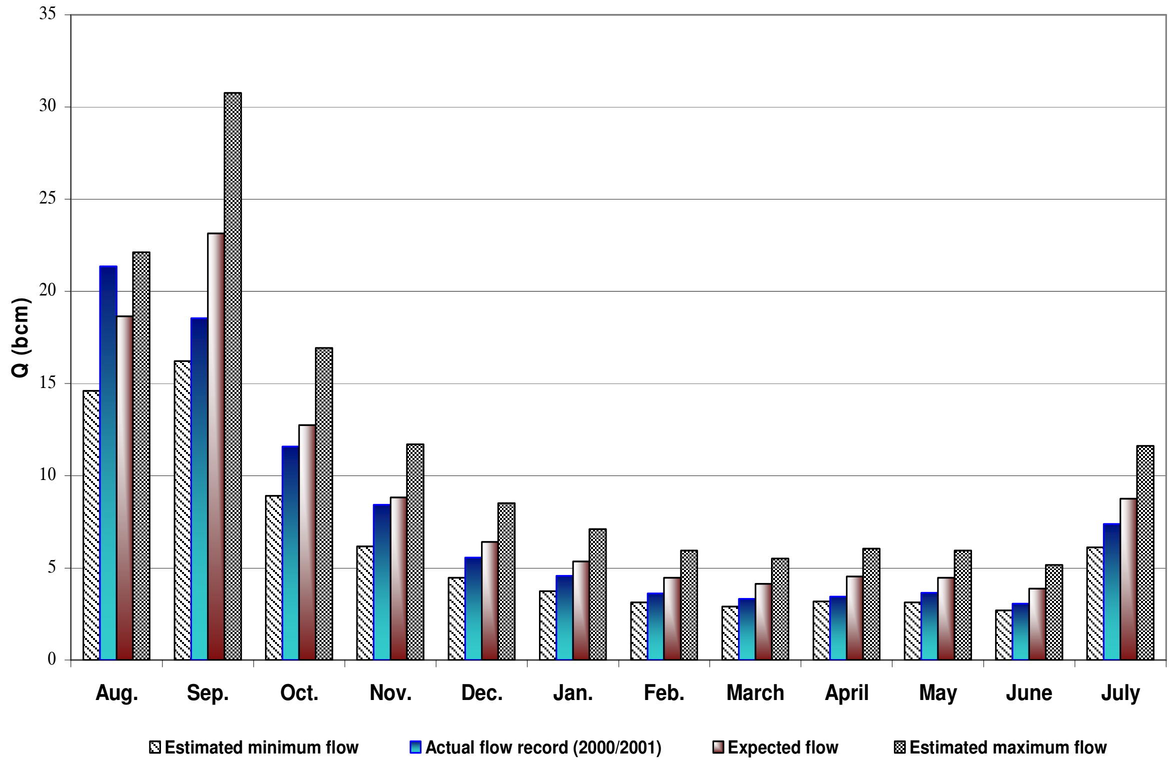

Figure 7: Flood Forecasting For 2000/2001

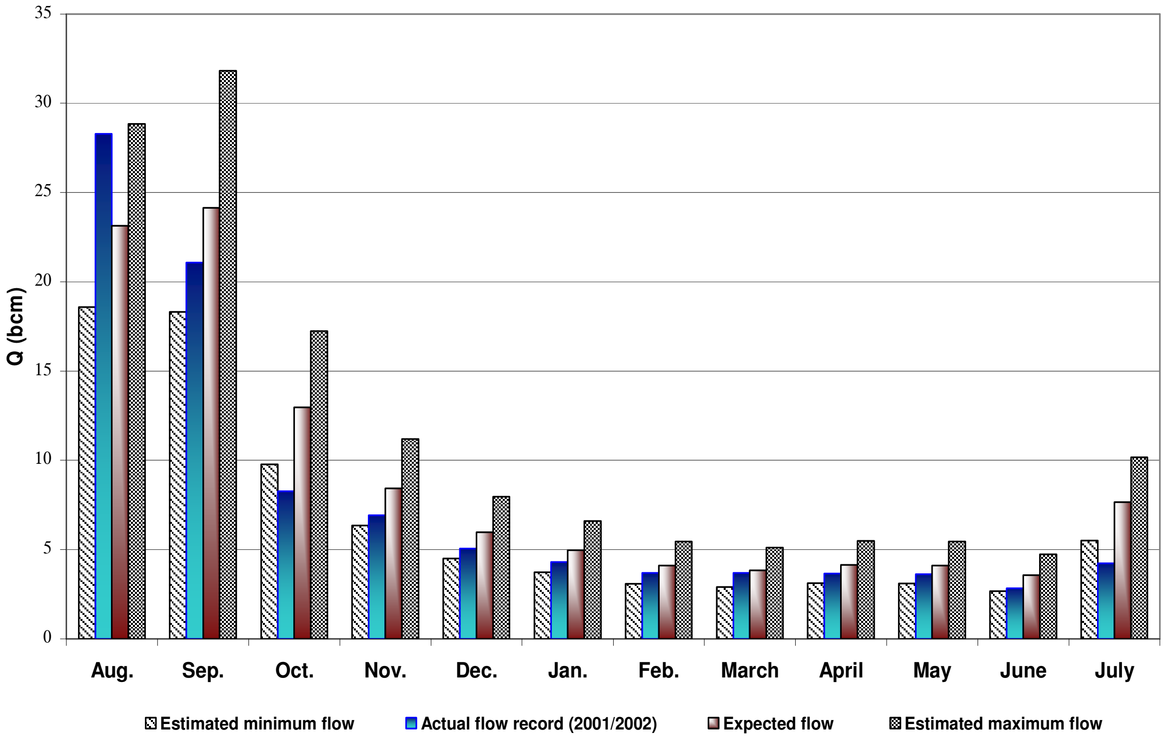

Figure 8: Flood Forecasting For 2001/2002

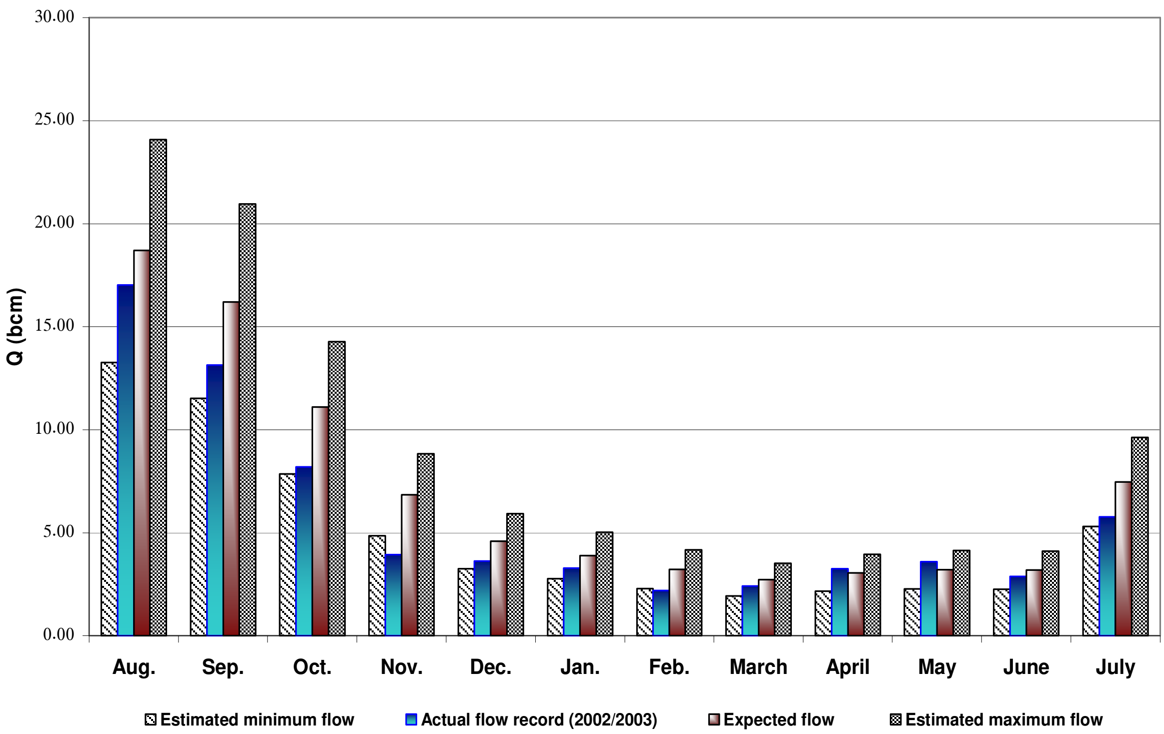

Figure 9: Flood Forecasting For 2002/2003

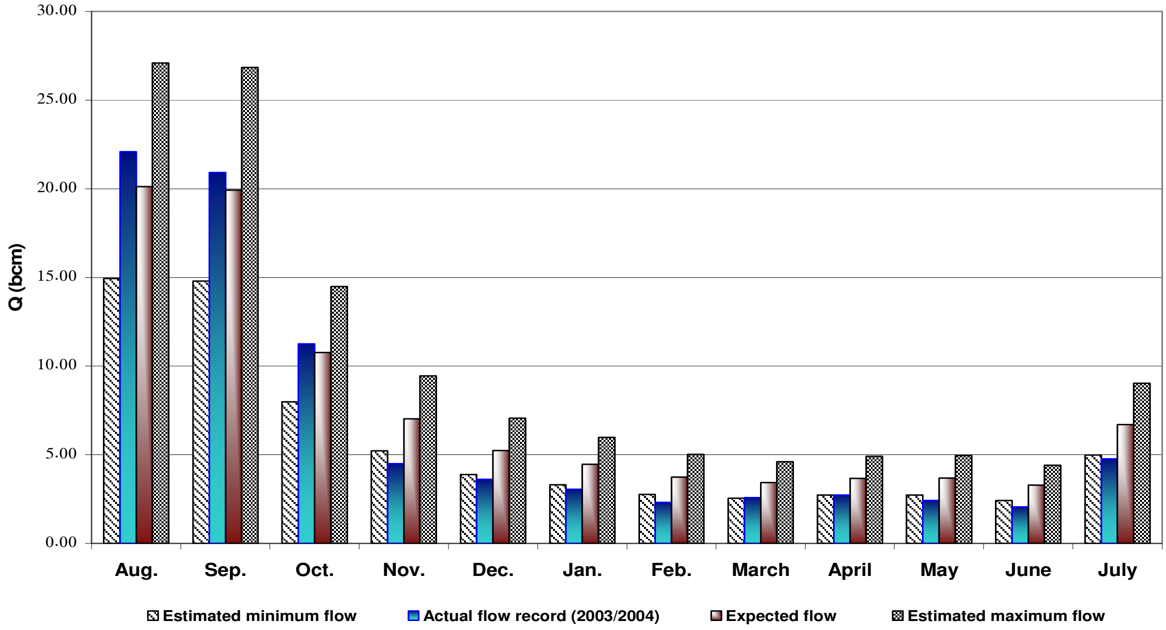

Figure 10: Flood Forecasting For 2003/2004

LONG TERM FLOOD FORECASTING

The future status for water management of Lake Nasser can be deduced during the period from year 2005 to 2017 through a number of important factors such as water requirements for national projects and Sudanese abstractions during this period. In addition, forecasting the incoming flood which is considered one of the main factors.

Flood forecasting will help in achieving better management of the incoming flood successfully to avoid the side effects resulting from a high flood like in the 1988/1999 flood and consider measures in case consecutive low floods arrival. Accordingly the model has been used to forecast the natural river income per the yearly income during the period (2005/2006) until the year (2016/2017) and the results are shown in figure 11. From this figure the minimum and maximum of the predicted flow at year 2004/2005 and 2011/2012 respectively and their values are about 71 and 92 bcm. From these results it is concluded the characteristics of the flood in the study period is recognized low and average flood. So, more studies are required to study the effect of high and low consecutive floods on the Lake Nasser water management especially after national projects implementation.

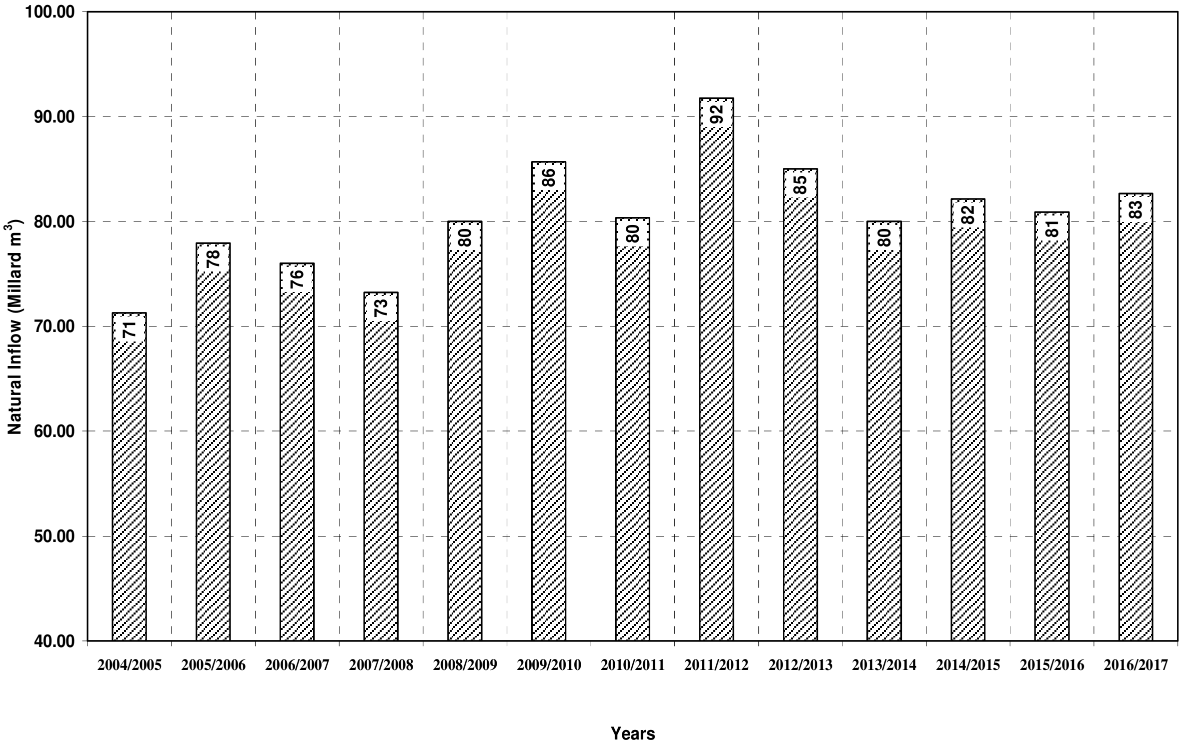

Figure 11: Flood Forecasting From Year 2005 until Year 2017

SUMMARY AND CONCLUSIONS

Flood forecasting plays an important role in matching the supply with the estimated demands, taking the necessary measures to avoid high flood damages, and minimizing the impacts of recurrent droughts hence, manage the flood efficiently. This research work aimed at studying the periodicity of the annual values of natural flow up to the year 2017.

The historical data were used to perform long term flood forecasting by using statistical forecasting methods. In this study the monthly inflow behaviour were analyzed. From these analyses it is clear that the flow records before mid sixties could belong to one distribution and after this time the data belong to other distribution so mid sixties is considered the point of change inflow behavior. Consequently, the series from 1964 up to 2004 was selected to time series analysis to guarantee the quality of the model selected.

The minimum –maximum and expected monthly flow is computed by the ARIMA model and these results were analyzed and evaluated. The evaluation results included revising the actual recorded monthly flow data with the expected flow through the five consecutive years from 1999/2000 to 2003/2004. Also, different factors indicate the accuracy of the forecasting were determined. All the analyses of results have proved the model suitability in forecasting the incoming floods to the lake Nasser upstream the high dam. Consequently, the statistical model could be used to predict the incoming flood during the period from year 2005 until the expected national project implementation at year 2017. The minimum and maximum of the predicted flow at year 2004/2005 and 2011/2012 respectively and their values are about 71 and 92 bcm. From these results it is concluded that the characteristics of the flood in the study period is recognized as low to average floods. Thus, more studies are required to figure out the effect of high and low consecutive floods on the Lake Nasser water management especially after national projects implementation.

REFERENCES

1- Shahin, M. 1985. Hydrology of the Nile Basin. Elsevier, Amsterdam, Holland. P.575.

2- Sadek, Nahla and Aziz, Medhat "Flood Management of Lake Nasser after the New Toshka Barrages Construction" Ninth International Water Technology Conference, IWTC9, Sharm El-Sheikh, Egypt, 2005.

3- Ministry of Public Works (MOPW), "Nile Basin." Ten volumes with supplements Ministry of Public Works, Cairo, Egypt, (1933). Sadek, Nahla. Lake Nasser Flood Analysis. Unpublished Ph.D. Dissertation. Ain Shams University, Egypt, 2002.

4- Water Master Plan (WMP), "Hydrological Simulation of Lake Nasser" Vol. 1, Analysis and Results. Technical Report 14 of the Water Master Plan. Prepared for the Ministry of Water Resources and Irrigation, Cairo, Egypt, 1979.

5- Aziz, Medhat, Hossam Ibrahim, and Nahla Sadek. "Lake Nasser Flood and Sediment Analysis". Eighth International Symposium on River Sedimentation, (ISRS 8), Cairo, Egypt, 2001.

6- Box and Jenkins, "Handbook of Applied Hydrology", 1976.

7- Saied, R. 1993. The Nile River. Dar El-Helal. Egypt. P. 342.

8- Aziz, Medhat, and Ismail, Sherine, "Lake Nasser Flood Analysis". International Conference of UNESCO Flanders Fit Friend/Nile Project", Sharm El-Sheikh, Egypt, 2005.