A Time Series Analysis of the Nile River Low Flows

Edward Aguado

Department of Geography, University of Wisconsin-Madison, Madison, WI 53706

Abstract. Box-Jenkins modeling provides a method for measuring serial autocorrelation. In this example, an 812-year record of the Nile River annual low-flow stages is analyzed and found to be described best by an ARIMA (2, 1, 1) model. The substantial reduction of the variance resulting from the application of this model provides a quantitative measure of the combined effects of basin water retention and climatic persistence.

Key Words: time series analysis, autoregressive-moving-average model, Nile River.

ONE of the major problems in the interpretation of climatic and hydrologic data is the lack of reliable long-term records. Any attempt to analyze such data for their time series characteristics will usually be hampered by the fact that climatic and hydrologic variables (e.g., precipitation and river runoff) will be subject to changes in their statistical parameters over long time periods. Hence a hydrologic record of river flow data extending back even 150 years reflects what may in fact be an inhomogeneous period. Much evidence has recently appeared to support the notion that the past century and a half has been unusual in terms of global precipitation patterns (Lamb 1977). Thus a record that extends for much longer periods of time can be of considerable value.

Records of the Nile River annual maximum and minimum flows have been recorded for as far back as the year 622 A.D. Although the river is gauged at Aswan (24°N, 33°E), its flow is the result of precipitation from two principal sources. The spring flows (observed as annual minima) mostly reflect rainfall in the basin of the White Nile near the equator in the southern Sudan and Lake Victoria region. The summer flows (annual maxima) have their major source in the upper Blue Nile in the Ethiopian Highlands. Although later periods in the record have major gaps in the data, the period between the early seventh century and the fifteenth century is nearly complete, with few missing data values. Recently these data have been used to infer short-term (Riehl and Meitín 1979) and long-term (Riehl, El-Bakry, and Meitín 1979) climatic variations.

This paper will apply the Box-Jenkins (Box and Jenkins 1965) method of time series analysis to the low-flow records of the Nile during the period 622–1451 A.D. in order to identify the most appropriate autoregressive-moving-average (ARMA) model and to determine its coefficients. As with any statistical model, mere calculation of parameters and significance testing of a time series will not provide a causal explanation of runoff phenomena. However, if we operate in the framework of a drainage basin, which we view as a linear filter on a stochastic input (precipitation), and if we take into account our knowledge of the nature of ARMA models, we should be able to make some reasonable interpretations of the data. In a previous paper (Hipel and McLeod 1978), a similar long-term record of the Nile low flows was analyzed along with twenty-two other natural time series. However, there is reason to question the validity of the proposed model owing to its nearly nonstationary nature. This paper will present a more detailed analysis of the data and discuss the problem of stationarity.

Data

This data set was made available more than four decades ago in a paper by C. S. Jarvis (1936). Jarvis did not analyze the data but did provide a discussion of their possible future uses and potential sources of error. The annual high and low flows were originally recorded on the Nilometer, a well five meters by five meters in area in which the stage of the water was etched into the wall.

Because we are dealing with stage rather than discharge, we must exercise some caution in interpreting the data. As the relationship between stage and discharge is nonlinear, we cannot assume a one-to-one correspondence between the two variables. Furthermore, the stage-discharge relationship will not be constant through time as the channel changes its shape and is subject to periods of scour and sedimentation. In fact, over the period of record the river bed has been subject to deposition at the average rate of 0.15 meters per century (Jarvis 1936).

We must also be cautious when making inferences about climatic patterns of the source regions based on this hydrologic record. Minimum and maximum annual flows do not reveal much about the distribution of precipitation in the source regions in any particular year. Any precipitation total for a given year may produce different runoff peaks and minima because of the time distribution of rainfall, and we have no quantitative knowledge of the effect of this variable. Furthermore, although we are examining only the serial autocorrelation of annual minima, it is likely that a given year's minimum stage will be influenced by the previous year's maximum flow. Therefore, some of the observed patterns may result from climatic events in a different region and obscure or falsely imply climatic conditions in the White Nile area.

Only the annual minimum-flow data have been used in this study because the maximum flood record seems subject to greater error. Land taxation in Egypt was based on the height of the annual flood, and it is likely that falsification of this record has occurred in order to increase revenues. Also, the two accounts of river stages for this period, those of Omar Toussoun and Aboul Mehasin, correspond more closely in their data for the minima than for the maxima. Furthermore, the maximum-flood data are of dubious value owing to the unknown magnitude of water diversion for irrigation during high flows.

Time Series Analysis

A time series of any variable is a collection of observations taken sequentially in time, and the analysis of such data can be performed in either the time or frequency domain. When doing a time-domain analysis (as in this paper) we are concerned with fitting a model of the form

$$Y_t = \phi_1 Y_{t-1} + \ldots + \phi_p Y_{t-p} + \theta_1 \epsilon_{t-1} + \ldots \phi_q \epsilon_{t-q} + \epsilon_t + \delta,$$

which best predicts the values of variable Y at time t based on previous observations $Y_{t-1}, \ldots, Y_{t-p}$, previous error terms $\epsilon_{t-1}, \ldots, \epsilon_{t-q}$, and a constant, δ. The ϕ values are collectively referred to as the autoregressive part of the model (of order p), whereas the θ's constitute the moving-average component (of order q). The inclusion of a nonzero δ introduces a deterministic trend in the model. We refer to this stochastic process as an autoregressive-moving-average model (ARMA (p,q)), and we are concerned with identifying the order of the model and estimating its coefficients. Models in which there are no moving-average terms (i.e., $q = 0$) are simply called autoregressive (AR(p)), whereas moving-average models (MA(q)) are those with no autoregressive components.

A series is said to be stationary if its probability characteristics are constant through time so that the joint probability distribution of $Y_{t-1}, \ldots, Y_{t-m}$ depends only on the time intervals between the observations, not their locations within the sequence. Frequently, a series will disclose a trend in the mean through time but will otherwise exhibit apparently homogeneous behavior. Rather than include a deterministic nonzero δ into the model, we may achieve stationarity by applying a filter to the original series. In practice, the differences between successive observations are usually taken to generate a new series, which is then subjected to model identification. We thus obtain an integrated autoregressive-moving-average model (ARIMA $(p,d,q,)$), where d denotes the number of times the series is differenced. Usually $d = 1$ will be sufficient to generate a detrended series with a mean of zero. With the absence of a trend, we may eliminate δ from the model. The model is referred to simply as an integrated moving-average model (IMA (d,q)) if it has no autoregressive components, or an integrated autoregressive model (ARI (p,d)) if no moving-average terms are included.

When a sufficiently long series is purely autoregressive, the process can be treated as an ordinary linear regression, the independent variables being the values of Y observed at previous lag intervals. When we introduce moving-average terms into the model, we may no longer do this. We are then forced to use an iterative scheme such as the Gauss-Newton method to estimate the parameters that minimize the sum of squares for any given order of model. MINITAB is a widely available computer package that offers a routine for the estimation of ARIMA models. All analyses in this paper are based on calculations performed by this package.

A preliminary estimate of the order of the ARIMA model can be made by examination of the autocorrelation function (ACF) and partial autocorrelation function (PACF) of the series. If the autocorrelations for increasingly large lags do not approach zero fairly quickly, we might infer nonstationarity and be forced to use an integrated model. Unfortunately, there is no precise rule for how quickly the ACF should die out before we can assume *stationarity*. If the process appears to be stationary with a steadily decreasing ACF, it often will be described by an autoregressive model; however, if the ACF has one or more large autocorrelations followed by approximately zero-valued autocorrelations at greater lags, there will probably be one or more moving-average terms. ARMA models with both types of terms will usually exhibit ACF's with one or more spikes, followed by a geometric decay. Usually the PACF will have the reverse pattern, with moving-average processes having a geometric decay of partial autocorrelations and autoregressive forms having one or more spikes.

The objective of ARIMA modeling is the identification of a model that yields a purely random series of residuals without the inclusion of unnecessary parameters. Box and Jenkins (1965) demonstrated the technique by analyzing six actual time series. For each of these they compared the ACF's and PACF's of the residuals for models of different order and employed a formal test involving the sum of the residual autocorrelations. If a model is appropriate, the lag autocorrelations of the residuals will be normally distributed about a zero mean with a standard error equal to $(N - d)^{-1/2}$, where N is the number of observations. If an excessive number of significant autocorrelations are found, a higher-order model should be tested. Furthermore, it has been shown that the quantity

$$Q(r) = N\,(N + 2) \sum_{k=1}^{m} (N - k)^{-1} r_k^2$$

will, for large N, be distributed as χ^2_{m-p-q}, where r_k is the estimated residual autocorrelation at lag k (Ljung and Box 1978). Until recently there has been no objective criterion by which one of several apparently adequate models could be chosen to represent the series. For example, the addition of a second autoregressive term to an AR(1) model will reduce the residual error, but a test of significance for that term may have proved inconclusive. The minimization of Akaike's Information Criterion (AIC) has been shown to be a valuable technique in model selection (Ozaki 1977). The use of these procedures will be demonstrated in the analysis of the Nile series. More detailed discussions of these methods are available in the aforementioned references.

Data Analysis

Inspection of the ACF and PACF of the raw data (Figures 1 and 2) suggests the possibility of nonstationarity because the ACF does not rapidly approach zero. Furthermore, a linear regression of the data with time also indicates a gradual but significant trend that can be explained by siltation of the river bed. This trend presents us with an interesting problem. Although the series does not truly meet the condition of stationarity, it comes close.

If we treat this as a stationary series, the ACF and PACF of the raw data indicate the presence of autoregressive terms resulting from the geometric decrease of the ACF and the spiked PACF. Although the presence of a moving-average term along with the autoregressive components cannot be ruled out, it is obvious that we do not have a pure moving-average process at work. Table 1 shows the results of applying various ARMA models to the data. The mixed ARMA models of order (1, 1) and (2, 1) yield superior results to either pure MA or AR forms of Order 1 or 2 as indicated by the lower sum of squares. The ARMA (1, 1) model produces a better fit than either the AR(2) or MA(2) with an equal number of free parameters. In spite of its superiority to the pure AR and MA models, the residual ACF (Figure 3) of the ARMA (1, 1) model does not indicate an adequate fit to the data. With a sample of 812 observations, a residual lag autocorrelation of ± 0.069 indicates significance at the 5 percent level. Four autocorrelations, at lags 19, 20, 23, and 25, exceed this value. More

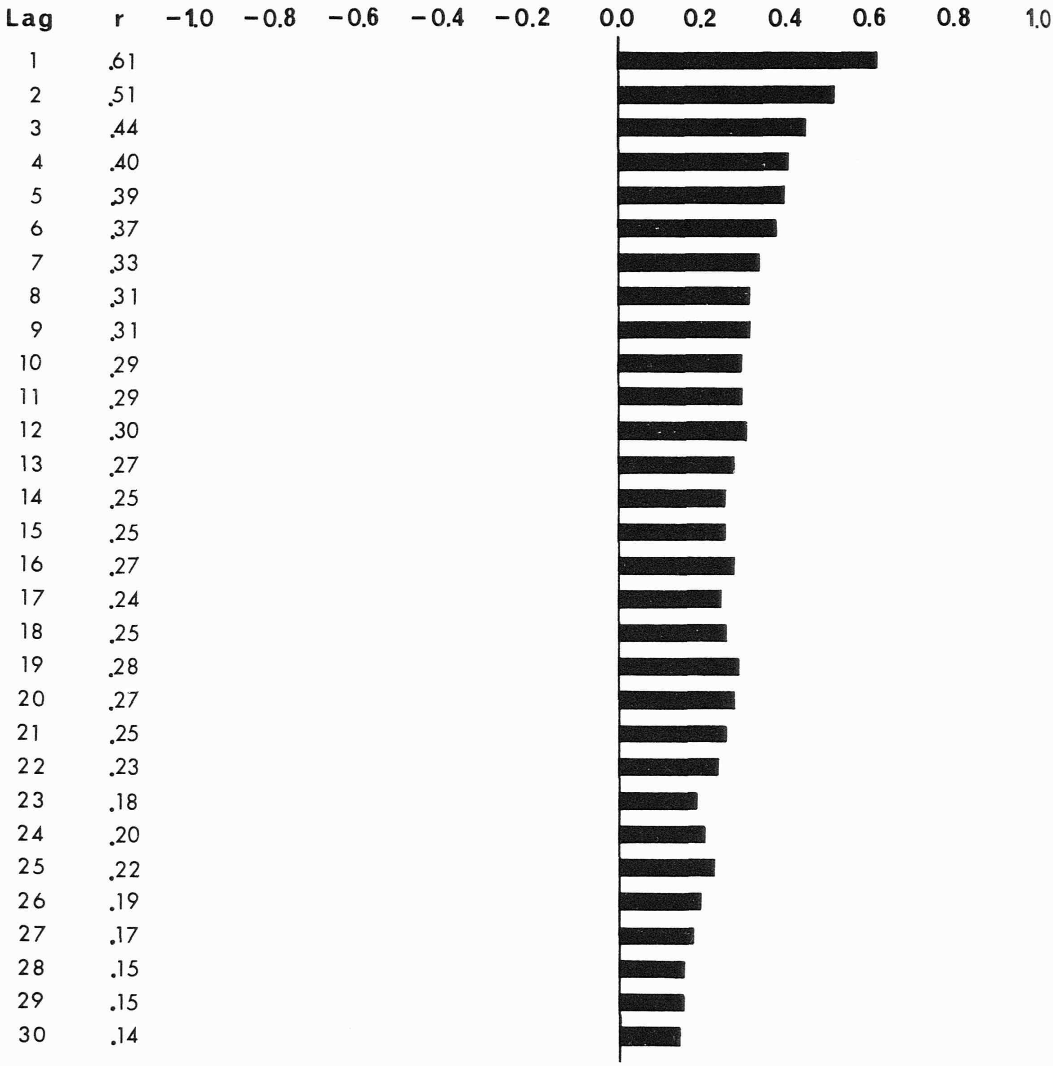

Figure 1. Autocorrelation function (ACF) of raw data.

importantly, for $m = 30$, $Q(r) = 42.42$ and lies outside the 0.05 probability region of the χ^2 distribution.

The residual ACF and PACF (Figures 4 and 5) of the ARMA (2,1) model indicate a better fit than that obtained by the ARMA (1,1). The addition of the second autoregressive term eliminates three of the significant lag autocorrelations and reduces the overall sum of squares. Furthermore, the $Q(r) = 28.18$ is well within the χ^2 distribution, indicating that this model produces a good fit to the data. There is, however, a question regarding the stationarity of this model. A model with two autoregressive terms will be stationary only if both roots of the equation $x^2 - \phi_1 x - \phi_2 = 0$ fall within the unit circle on the complex plane. In this case, the roots are 0.30 and 0.97, with the latter value suggesting a near nonstationarity. A process that truly violates the condition of stationarity will exhibit "explosive" growth or decay, with the local behavior of the series depending on the level of Y_t. A model that is nearly nonstationary will generate episodes where the local level of Y_t departs from the mean (or trend line) but eventually returns to it. This latter condition is consistent with the behavior of the Nile series.

Table 2 compares the results of several models fitted to the differenced series. As was the case with the raw data, the combination of AR and MA terms yields better results than those obtained from purely autoregressive or moving-average components. The ARIMA (1, 1, 1) has a sum of squares of 2316686, substantially lower than the 2562513 and 2350782 associated with the ARI(2, 1) and IMA(1, 2) models with the same number of free parameters. The addition of a second AR term results in an ARIMA (2, 1, 1) model with a residual error of 2299884. Although the ARIMA (1, 1, 1) produces an adequate fit, as indicated by the ACF and PACF (Figures 6 and 7), and $Q(r) = 32.38$, we might wish to sacrifice parsimony in return for a better fit by adding a second AR term. An F-test was used to compare the ARIMA (1, 1, 1) and (2, 1, 1) models. The value of 5.9 indicates that the reduction in the variance is significant at the 0.05 level. Furthermore, the t-test of the second AR term shows it to be significant, and the correlation matrix of the parameters in the higher-order model indicates that they are mutually uncorrelated.

We now have three possible models, a stationary ARMA (2, 1) and the nonstationary ARIMA (1, 1, 1) and (2, 1, 1) processes. The stationary model yields the lowest residual error of the three but requires one more parameter (the constant δ) than does the ARIMA (2, 1, 1). The

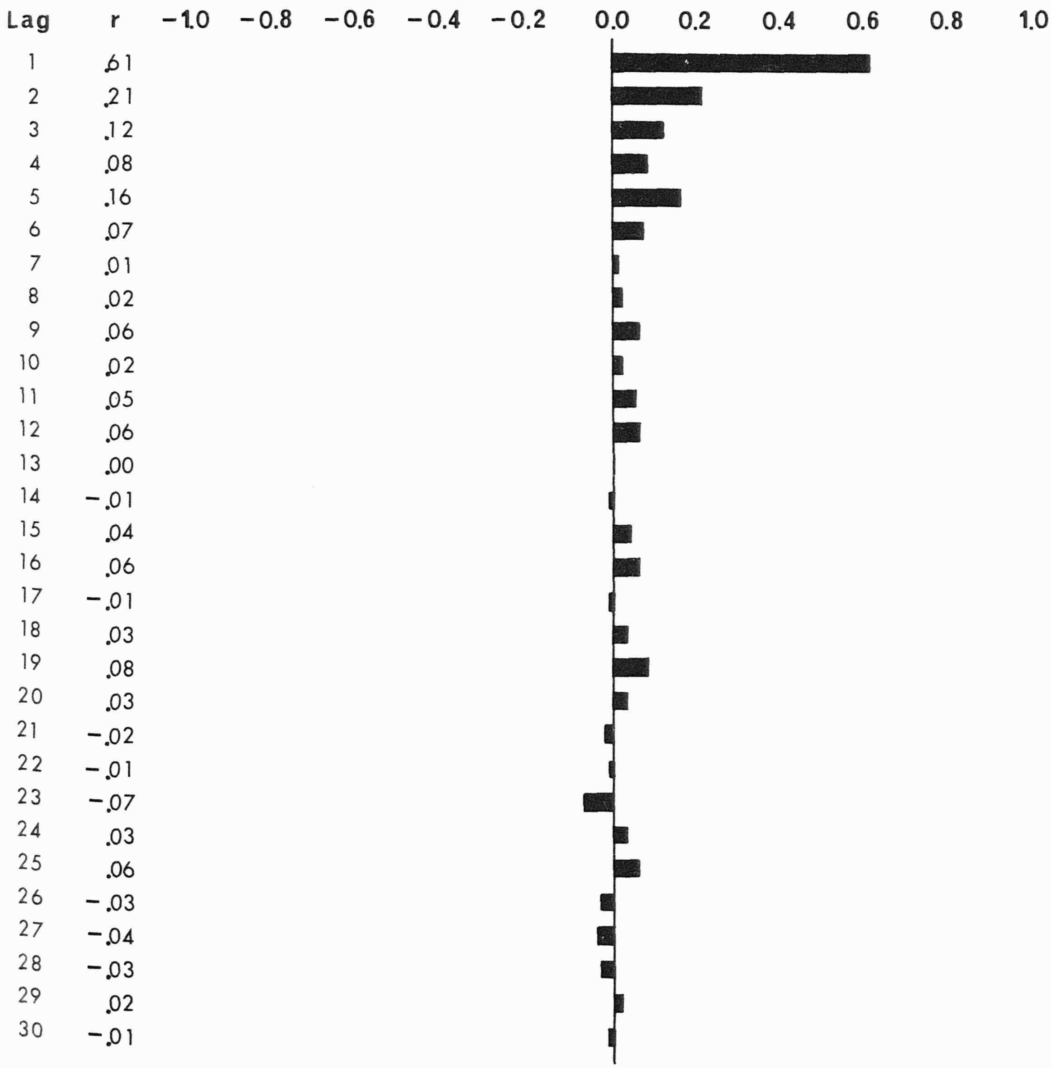

Figure 2. Partial autocorrelation function (PACF) of raw data.

Table 1. ARMA Models for Raw Data

Model	Sum of Squares	Parameters	t-Test
AR (1)	2.50×10^6	$\phi_1 = 0.61$	22.07
		$\delta = 108.58$	55.72
AR (2)	2.39×10^6	$\phi_1 = 0.48$	14.07
		$\phi_2 = 0.21$	6.12
		$\delta = 85.79$	44.99
MA (1)	2.98×10^6	$\theta_1 = 0.45$	14.25
		$\delta = 280.35$	90.92
MA (2)	2.70×10^6	$\theta_1 = 0.47$	14.09
		$\theta_2 = 0.29$	8.53
		$\delta = 280.40$	78.51
ARMA (1, 1)	2.33×10^6	$\phi_1 = 0.88$	37.30
		$\theta_1 = -0.49$	−11.06
		$\delta = 32$	33.75
ARMA (2, 1)	2.29×10^6	$\phi_1 = 1.27$	21.22
		$\phi_2 = -0.29$	−5.60
		$\theta_1 = -0.83$	−18.99
		$\delta = 7$	21.74

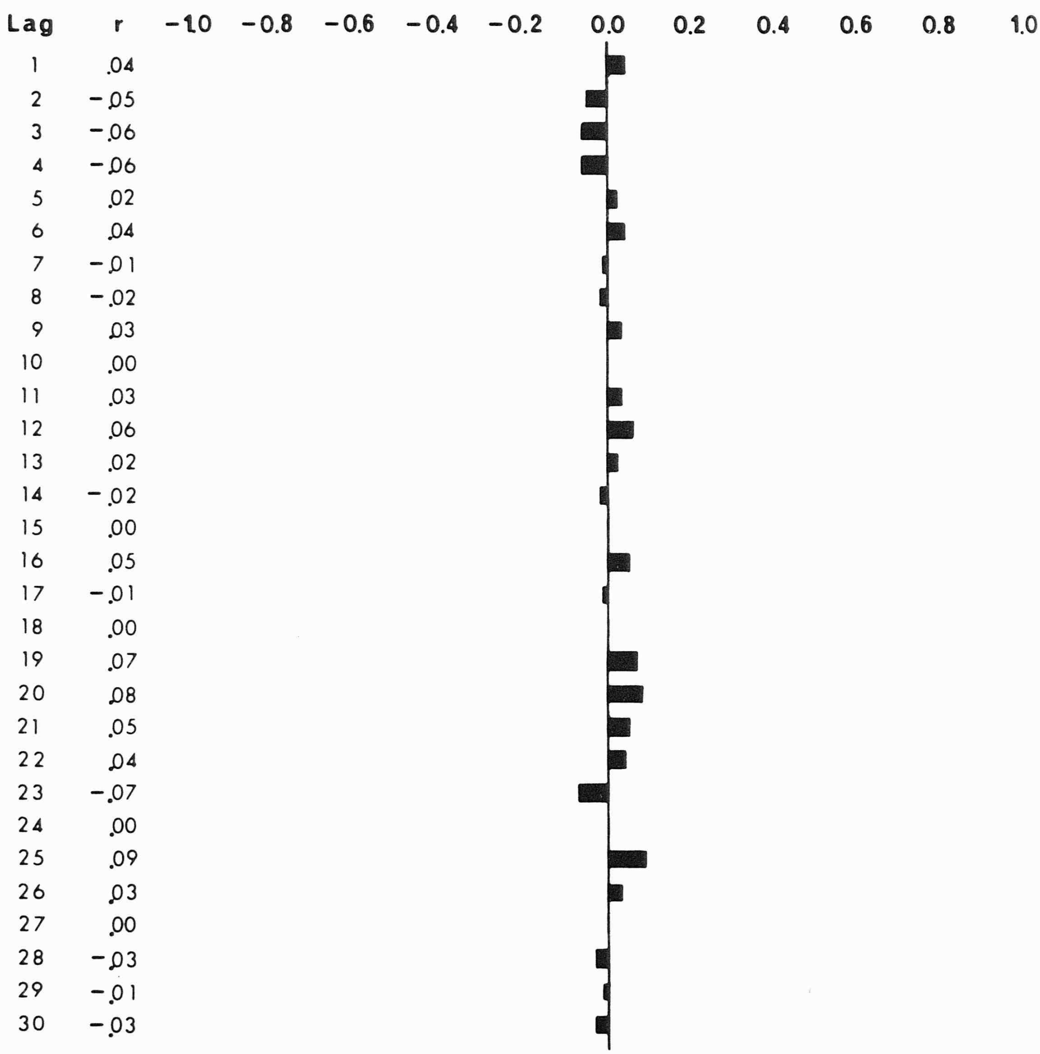

Figure 3. Autocorrelation function of autoregressive-moving-average model (ARMA) (1, 1) residuals.

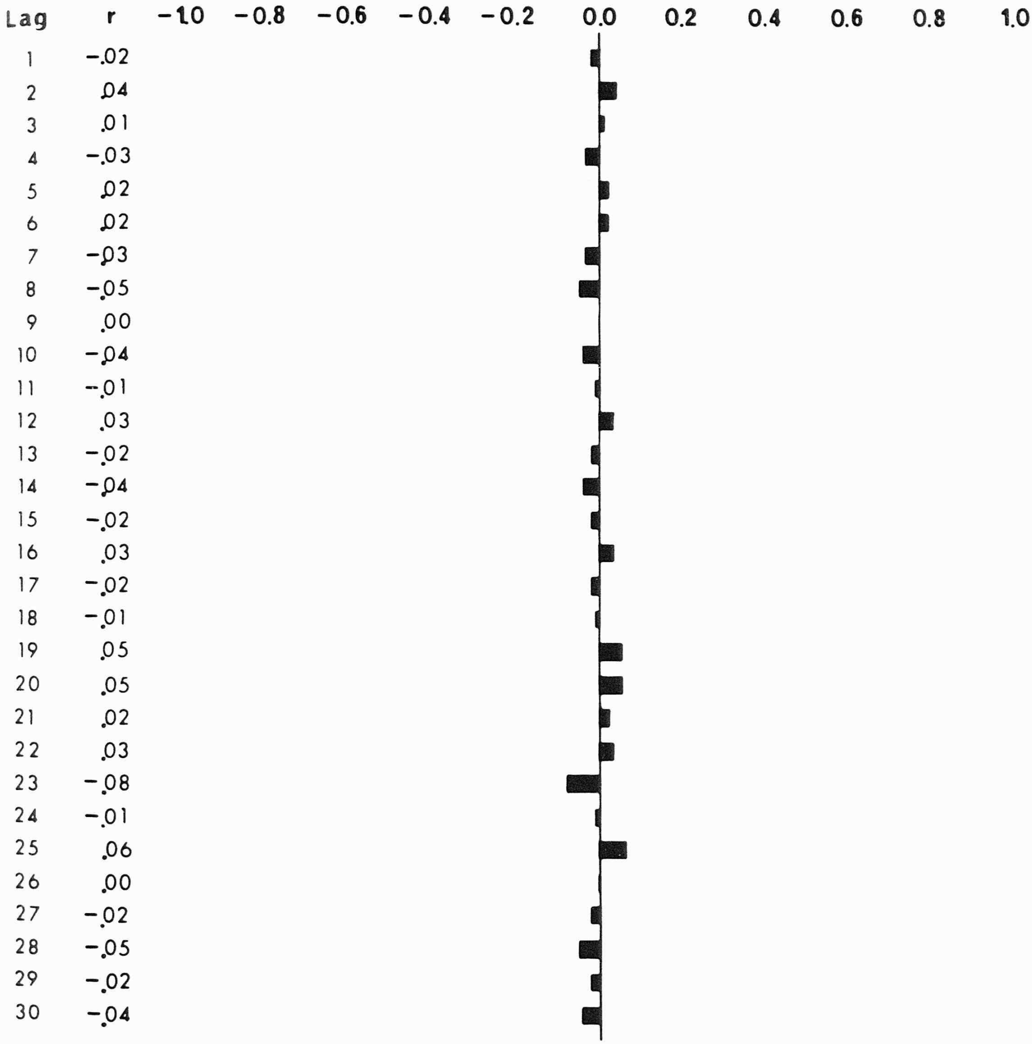

Figure 4. Autocorrelation function of autoregressive-moving-average model (ARMA) (2, 1) residuals.

selection of the appropriate model can be made by determining the one that minimizes Akaike's Information Criterion (AIC). For the purpose of comparing different models, Ozaki (1977) modified the AIC and redefined it as

$$N \log_e \sigma^2 + 2\,(p + q + 2) \qquad \text{for } d = 0,$$

and

$$N \log_e \sigma^2 + 2\,(p + q + 1)\,[N/(N - d)] \qquad \text{for } d \neq 0.$$

The first term in these equations is a measure of the poorness of model fit, which, for any given N, increases with the residual error. The second term is a measure of the unreliability of the model, which increases with the number of free parameters. The AIC values are 6462, 6458, and 6454 for the ARMA (2, 1), ARIMA (1, 1, 1), and ARIMA (2, 1, 1), respectively, thereby supporting the choice of the last of these as the appropriate model. This result is, not surprisingly, in accordance with the results of the F-test used to compare the ARMA (2, 1) model with four free parameters to the ARIMA (2, 1, 1) with three. The value of F = 0.012 does not indicate that the addition of the extra parameter was justified. Thus we obtain the final model:

$$w_t = 0.36 w_{t-1} + 0.10 w_{t-2} - 0.93 \epsilon_{t-1} + \epsilon_t,$$

$$\text{where } w_t = Y_t - Y_{t-1}.$$

Lag	r
1	−.02
2	.04
3	.01
4	−.03
5	.01
6	.02
7	−.03
8	−.05
9	.01
10	−.03
11	−.01
12	.03
13	−.01
14	−.04
15	−.02
16	.04
17	−.02
18	−.02
19	.06
20	.06
21	.01
22	.02
23	−.08
24	−.02
25	.06
26	.01
27	−.02
28	−.05
29	−.01
30	−.03

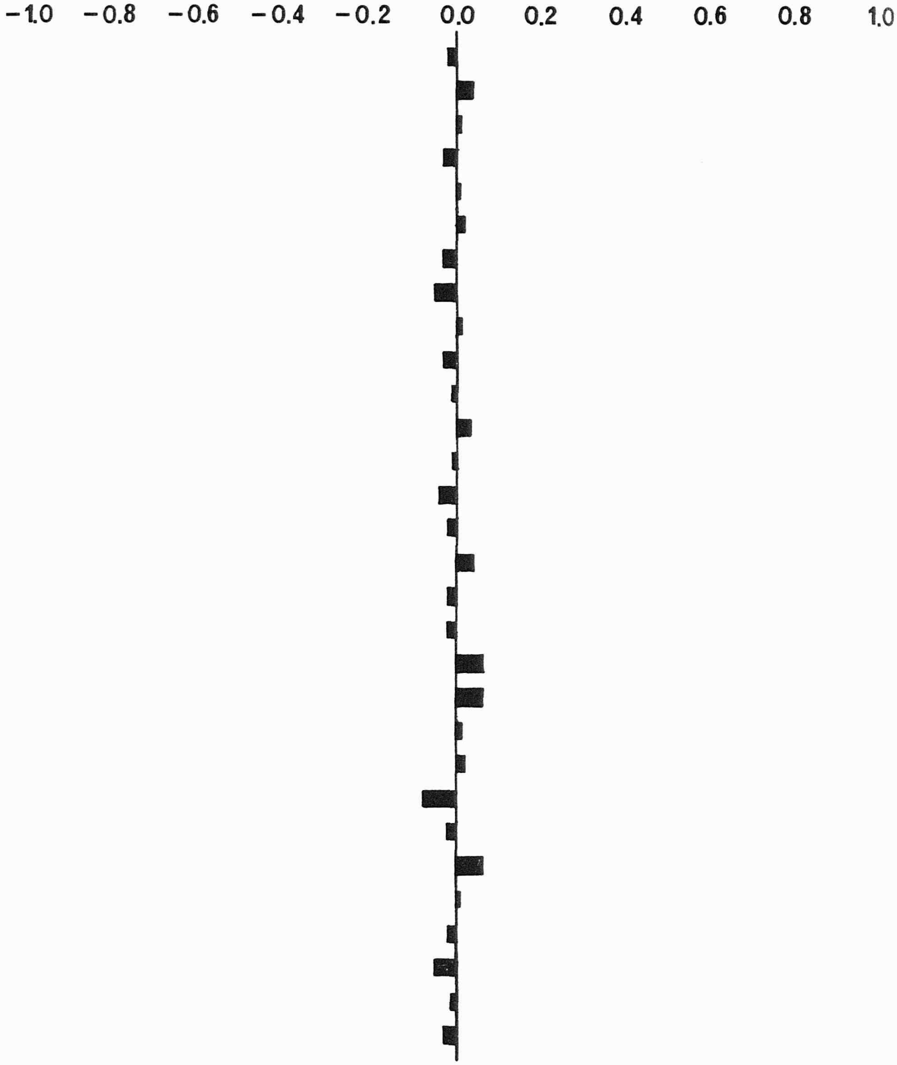

Figure 5. Partial autocorrelation function of autoregressive-moving-average model (ARMA) (2, 1) residuals.

Table 2. ARIMA Models for Differenced Series

Model	Sum of Squares	Parameters	t-Test
ARI (1, 1)	2.69×10^6	$\phi_1 = -0.36$	−11.09
ARI (2, 1)	2.56×10^6	$\phi_1 = -0.44$	−12.85
		$\phi_2 = -0.21$	−6.28
IMA (1, 1)	2.43×10^6	$\theta_1 = -0.64$	−23.89
IMA (1, 2)	2.35×10^6	$\theta_1 = -0.58$	−16.69
		$\theta_2 = -0.20$	−5.67
ARIMA (1, 1, 1)	2.31×10^6	$\phi_1 = 0.36$	8.83
		$\theta_1 = -0.90$	−48.19
ARIMA (2, 1, 1)	2.30×10^6	$\phi_1 = 0.36$	9.60
		$\phi_2 = 0.10$	2.62
		$\theta_1 = -0.93$	−64.30

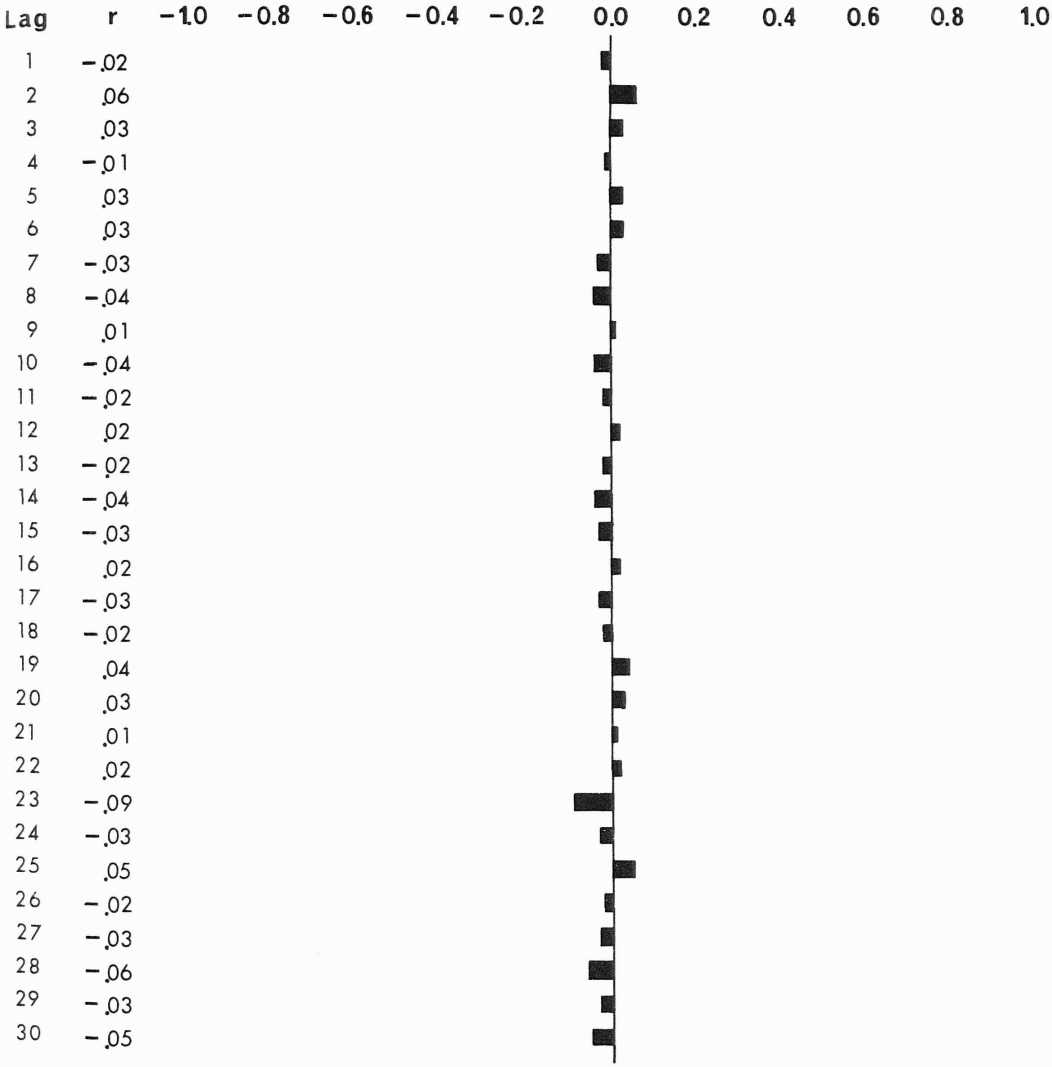

Figure 6. Autocorrelation function of integrated autoregressive-moving-average model (ARIMA) (1, 1, 1) residuals.

Conclusion

This study demonstrated the Box-Jenkins technique of time series analysis and fitted an ARIMA model to an unusually long record of annual minimum stage. The nonstationary model

$$\begin{aligned} Y_t - Y_{t-1} &= 0.36(Y_{t-1} - Y_{t-2}) \\ &\quad + 0.10(Y_{t-2} - Y_{t-3}) \\ &\quad - 0.93\epsilon_{t-1} + \epsilon_t, \end{aligned}$$

or

$$\begin{aligned} Y_t = 1.36\, Y_{t-1} - 0.26\, Y_{t-2} - 0.10\, Y_{t-3} \\ - 0.93\, \epsilon_{t-1} + \epsilon_t, \end{aligned}$$

was shown to be superior to the stationary ARMA (2, 1), as indicated by the F-test and minimization of the Akaike Information Criterion. Hipel and McLeod (1978) previously proposed an ARMA (2, 1) process for the Nile data. They calculated parameter values nearly identical to those obtained for the ARMA (2, 1) model in this paper, but they did not objectively compare the validity of this model to others of different order. The results presented here go beyond their findings. The gradual increase in river stage owing to siltation offers an intuitively acceptable explanation for the nonstationarity. Furthermore, a least-squares regression line fit to the data indi-

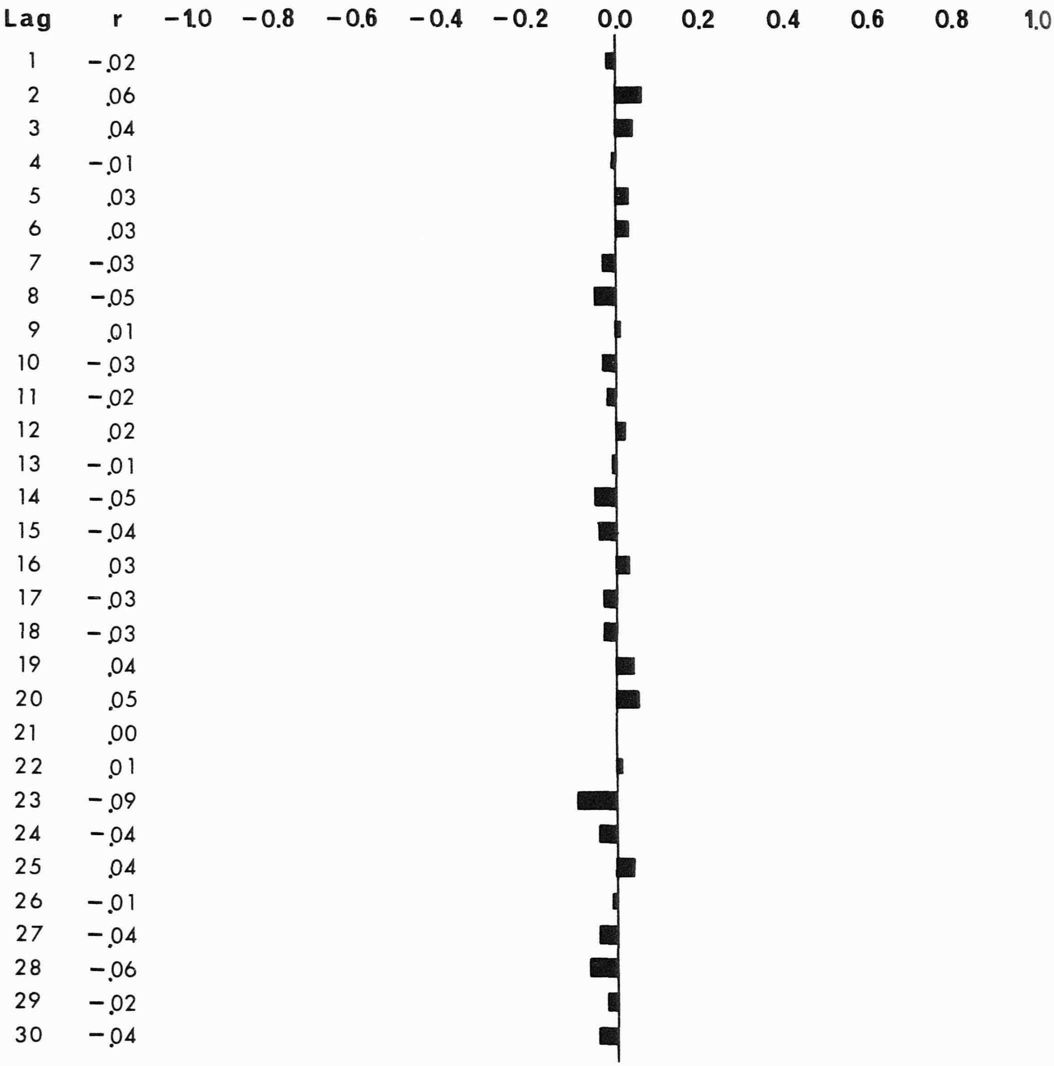

Figure 7. Partial autocorrelation function of integrated autoregressive-moving-average model (ARIMA) (1, 1, 1) residuals.

cates a positive and significant increase in stage with time and supports the notion of a trend in the mean.

It is not surprising that one year's minimum stage is statistically related to previous ones, as the groundwater-flow and surface-runoff processes act as filters on the annual precipitation values over the basin. It is interesting to note, however, the amount of variance "explained" by the two previous years' flow and the preceding error term. The sum of the squared residuals about the regression line accounting for river bed siltation is 3599122. If we fit the ARIMA model to the raw data, we obtain a sum of squares of 2299884 for a reduction of 36 percent. Although ARIMA modeling allows us to measure the serial autocorrelation of a set of data, the physical interpretation of the parameters is not an easy task. ARIMA models of any order need not be expressed in the difference equation form used above; it can be shown that such a process can be expressed in terms of the sum of the current and all previous error terms or a weighted sum of all previous observations Y_{t-i}. The order of the chosen model is that which preserves parsimony while still accounting for a significant reduction of the residual variance. The existence of two autoregressive terms does not necessarily imply that water in the basin is retained for two years. Furthermore, Granger and Morris (1976) have shown mathematically that the sum of two independent series or the sum of an AR process and random noise are likely to result in a mixed ARMA model.

Although this analysis has been performed on data from just one river, it has dealt with problems likely to be encountered in the investigation of any hydrologic system. Comparison of the results obtained here with those from rivers of different sizes and from other environments might be of great interest. Such studies could examine the increasing (or decreasing) importance of serial autocorrelation as models are fit to the flow data for gauges farther downstream within a single river. Finally, one might also wish to analyze this one data set further to determine the change in the serial autocorrelation characteristics during different episodes within this 812-year record. Clearly, this study represents merely a first step in the study of the time-domain behavior of fluvial systems.

Acknowledgment

I would like to thank the anonymous reviewers of this article for their valuable comments and Barbara J. Anderson for preparing the illustrations.

References

Box, G. E. P., and Jenkins, G. M. 1965. *Time series analysis: forecasting and control.* San Francisco: Holden-Day.

Granger, C. W. J., and Morris, M. J. 1976. Time series modelling and interpretation. *Journal of the Royal Statistical Society* 139, Part 2:246–57.

Hipel, K. W., and McLeod, A. I. 1978. Preservation of the rescaled adjusted Range 2: simulation studies using Box-Jenkins models. *Water Resources Research* 14:509–18.

Jarvis, C. S. 1936. Flood storage records of the River Nile. *Transactions of the American Society of Civil Engineers* 101:1012–71.

Lamb, H. H. 1977. *Climate present, past, and future: Vol. 2, climatic history and the future.* London: Methuen.

Ljung, G. M., and Box, G. E. P. 1978. On a measure of lack of fit in time series models. *Biometrika* 65:297–303.

Ozaki, T. 1977. On the order determination of ARIMA models. *Applied Statistics* 26:290–301.

Riehl, H., El-Bakry, M., and Meitín, J. 1979. Nile River discharge. *Monthly Weather Review* 107: 1546–53.

Riehl, H., and Meitín, J. 1979. Discharge of the Nile River: a barometer of short-period climate variation. *Science* 206:1178–79.