TIMCAI: An interactive program to assist in the identification of univariate nonseasonal nonmixed ARIMA time series

ROGER L. BROWN

*Center for Research on Pre-Traumatic Behavior
and Academic Computer Center
University of Wisconsin, Whitewater, Wisconsin 53190*

The need for accurate statistical techniques in the interpretation of time-series data is crucial for many social-psychological studies. While many techniques have been proposed and debated, the technique initially proposed by Box and Tiao (1965) known as autoregressive integrated moving average (ARIMA) modeling seems to show promise. While much has been written about the technique, there seems to be some reluctance on the part of the social-psychological community to use it when appropriate. One of the major reasons for this reluctance may be due to difficulties in the initial step of identifying an appropriate ARIMA model. Other reasons for the reluctance have been discussed elsewhere (Brown, 1980; McCleary & Hay, 1980).

The Box-Tiao approach is an iterative process encompassing several stages of development, with the initial step requiring a tentative model identification. This identification step basically relies on two statistics: (1) the autocorrelation function (AUCF) and (2) the partial autocorrelation function (PACF). Autocorrelation refers to the correlations among successive data points separated by each different observation (lag) in the series. Partial autocorrelation refers to the autocorrelation of a series with the influence of previous time points "partialed out." These two functions are calculated on the time-series data and generally supplied to the users as a correlogram, or a plot of the autocorrelation and partial autocorrelation as a function of K lags.

It is from an ocular analysis of the correlograms that a tentative ARIMA model may be identified. This ocular analytic technique basically compares the estimated AUCF and PACF with theoretical or expected functions, according to a variety of basic ARIMA models. Ambiguity in these estimated functions is then lessened by either placing an approximate 95% confidence band around each function (± 2 standard errors) or by calculating a t -like statistic (Bowerman & O'Connell, 1979). The t -like statistic, for both functions, is simply the value of each AUCF and PACF over their respective standard errors at each lag period.

The purpose of this software is to assist users in their ocular identification of an appropriate ARIMA model by providing statistically based decision support.

Portions of this paper and software were presented at the 90th annual meetings of the American Psychological Association, Washington, D.C., August 1982.

Description. TIMCAI (time-series computer-assisted identification) is a decision support system written in both VAX-11/780 BASIC and HP-2000 BASIC. The software estimates raw score autocorrelations (ρ_k), sometimes called serial correlations between time-series observations (lags), and partial autocorrelations (ϕ_{kk}). The estimate of autocorrelation is provided by the sampled autocorrelation (r_k), calculated as follows:

$$r_k = \frac{\sum_{t=1}^{N-k} (Z_t - \bar{Z})(Z_{t+k} - \bar{Z})}{\sum_{t=1}^N (Z_t - \bar{Z})^2} \times [N/(N-K)],$$

where

$$\bar{Z} = (1/N) \sum_{n=1}^N Z_n,$$

K = the number of lags, and N = the number of time points.

The estimates of partial autocorrelation, being the sampled partial autocorrelation (r_{kk}), require a somewhat more complicated process:

$$r_{kk} = |\hat{P}_k^*|/|\hat{P}_k|,$$

where $\hat{P}_k$ is the K by K estimated AUCF matrix, with $\hat{P}_k^*$ being the $\hat{P}_k$ matrix with the last column replaced by:

$$\begin{matrix} r_1 \\ r_2 \\ r_3 \\ \vdots \\ r_k \end{matrix}$$

So, $r_{11} = r_1$,

$$r_{22} = \frac{\begin{vmatrix} 1 & r_1 \\ r_1 & 1 \end{vmatrix}}{\begin{vmatrix} 1 & r_1 \\ r_1 & 1 \end{vmatrix}},$$

$$r_{33} = \frac{\begin{vmatrix} 1 & r_1 & r_2 \\ r_1 & 1 & r_2 \\ r_2 & r_1 & 1 \end{vmatrix}}{\begin{vmatrix} 1 & r_1 & r_2 \\ r_1 & 1 & r_1 \\ r_2 & r_1 & 1 \end{vmatrix}},$$

... r_{kk} .

Table 1
Nonseasonal Nonmixed Univariate Models

	White Noise	ARIMA (1,d,0)	ARIMA (2,d,0)	ARIMA (0,d,1)	ARIMA (0,d,2)
AUCF	All lag t values < 2.00	Exponential wave decay	Exponential and/or sine wave decay	Lag 1 t value > 2.00	Lags 1 and 2 t values > 2.00
PACF	All lag t values < 2.00	Lag 1 t value > 2.00	Lags 1 and 2 t values > 2.00	Exponential wave decay	Exponential and/or sine wave decay

Output. TIMCAI provides the following output: (1) correlograms for both raw score AUCF and PACF, (2) a differencing routine, (3) examples of theoretical nonseasonal ARIMA model correlograms, (4) a decision support routine for evaluating the stationarity of a series, and (5) a decision support routine outputting a probability-based identification of a model.

Two decision support routines operate in TIMCAI. The first assists the user in recognizing whether a series is stationary or nonstationary, a requirement for using the ARIMA technique. This routine is based on evaluating the number of sequentially marginally significant ($t > 1.8$) autocorrelative t values. The second routine attempts an ARIMA model identification based on the following theoretical model shapes (Table 1).

The identification routine establishes a theoretical exponential decay based on the Lag 1 AUCF/PACF of each realization evaluated. The summations of all lag AUCFs and PACFs are then compared to initialize either an autoregressive or a moving average scan. A parallel variance band, consisting of .01 increments, is then systematically expanded around either of the two decay functions until 90% compliance with the raw AUCF or PACF values is indicated, or no model is identified.

Requirements. The TIMCAI program written in

VAX BASIC is a single program developed and tested on the VAX-11/780 (DEC) minicomputer operating under VMS. The program requires 20 KB of storage. TIMCAI written for the Hewlett-Packard 2000 system is a group of three programs chained together, with a maximum of 15 KB of memory required for execution. Space requirements may be adjusted by the removal of remarks and instruction sets.

Availability. A source listing of either the VAX program or the Hewlett-Packard programs and a user's guide are available without charge from Roger L. Brown, Center for Research on Pre-Traumatic Behavior, University of Wisconsin, Whitewater, Wisconsin 53190.

REFERENCES

- BOWERMAN, B. L., & O'CONNELL, R. T. *Time series and forecasting*. North Scituate, Mass: Duxbury Press, 1979.
- Box, G. E. P., & TIAO, G. C. A change in level of a nonstationary time series. *Biometrika*, 1965, 52, 181-192.
- BROWN, R. L. Software psychology and the data analytic system: Computer assisted identification of correlograms in time-series data analysis. *Proceedings of the Fourth Annual SPSS Users and Coordinators Conference*, 1980, 4, 109-119.
- MCCLEARY, R., & HAY, R. A. *Applied time-series analysis for the social sciences*. Beverly Hills: Sage Publications, 1980.

(Accepted for publication November 27, 1982.)