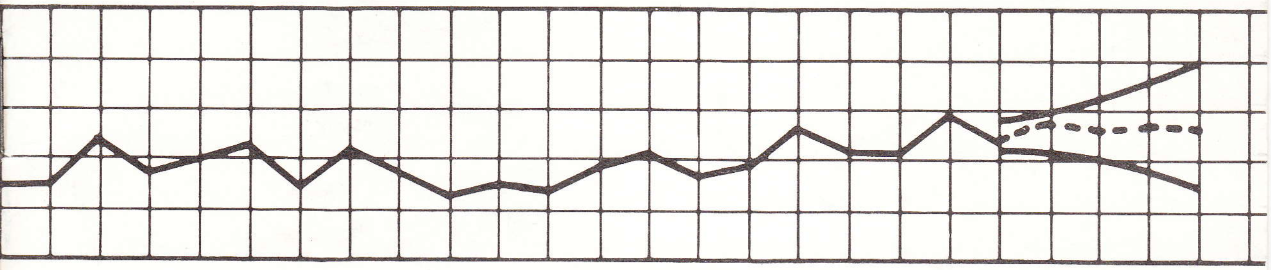
TIMES



A TECHNICAL DESCRIPTION

I. INTRODUCTION

Recently, a comprehensive statistical theory has been developed for deriving, from time series data, models for prediction (forecasting), simulation, and control of stochastic, or random, processes. With particular regard to forecasting, the theory represents a significant improvement over other techniques still in wide use. Understandably, the use of this powerful new technique is not without its difficulties: the computation involved in performing the statistical analysis required to determine the models is considerable and complex. Lambda Corporation has developed a very flexible and general computer program which simplifies the computations required to develop these models and to compute forecasts or simulations based on them. The program is capable of analyzing an almost limitless variety of models, simply by specifying the model structure on a few input control cards. The program not only estimates model parameters, but also computes statistics which indicate the statistical adequacy of the tentative models and guide improved development of the model. These statistics include the autocorrelation function, the partial autocorrelation function, and the spectral density function of the model "residuals."

The time series package, TIMES, consists of an estimation program called ESTIMATE, a forecasting program called FORECAST, and a simulation program called SIMULATE. Using the estimation program, the user develops, through a series of "tentative model" analyses, a model which is a statistically appropriate representation of the process generating the time series. This iterative approach is guided by statistics calculated by the program for each tentative model. This model is then input directly into the forecasting program, which computes a "minimum-variance" forecast, or into the simulation program, which computes a simulation. No computer programming is required of the user. As new observations become available, the forecasts can be readily updated by the computer, using the already-determined model.

The TIMES approach to developing forecasting models is analytically sophisticated, and some familiarity with statistics is needed to use the estimation program most effectively. The TIMES package provides an organization with the means for developing the most advanced forecasting models, without the necessity for making a substantial investment in highly specialized statistical analysis and computer programming. Included with the package is a user's manual describing the statistical background required to use the program effectively. For the statistician, the economist, or the operations research generalist, TIMES represents a significant aid: its flexibility allows rapid analysis of a vast range of tentative models, without the need for time-consuming computations or special-purpose computer program coding.

In this booklet we review the nature of time series analysis, describe the new approach, explain the statistical theory behind TIMES, and discuss the functions that TIMES performs.

II. TIME SERIES ANALYSIS

What is Time Series Analysis?

A time series is a set of observations taken at various points in some time (or space) interval. We shall restrict consideration to discrete time series, whose observations are taken at equally spaced points. Examples of time series include daily prices of a certain stock, monthly sales volumes of a particular product, weekly demands for a particular blood type. A number of mathematical techniques have been developed for describing the essential characteristics of time series and deriving empirical models exhibiting those characteristics. The application of such techniques to study time series is what we mean by time series analysis. In particular, statistical techniques are available for studying time series as stochastic, or probabilistic, processes. Time series analysis may be used to provide efficient quantitative descriptions of the empirical nature of the process generating the observations, and to determine mathematical models of the process which may be used to simulate or forecast its behavior.

The techniques used to analyze time series vary considerably, depending on the purposes of the analysis. On the one hand, we may be interested in a study of the frequency response of an aircraft component to environmental conditions. In such an application, spectral analysis is the appropriate time series analysis technique to use. The statistical techniques of spectral analysis are well developed, widely known, and efficient computer programs are available for performing this type of time series analysis.

The prediction of economic and industrial time series, on the other hand, requires a different type of time series analysis -- model building. In order to determine good forecasters, it is first necessary to determine an appropriate statistical model of the time series. In model building, we seek an efficient parametric representation of the process generating the time series, rather than the nonparametric representation afforded by spectral analysis. Unlike the area of spectral analysis, powerful techniques for building models for prediction are just recently being developed, and have not been available in the form of ready-made computer programs. Lambda's TIMES time series analysis program fills this gap.

Uses of Time Series Analysis Models

It may be desired to predict, or forecast, the expected values of future observations of a time series, using the information contained in observations up to the present time. Such estimates might be used, for example, as a basis for determining inventory levels. In forecasting, it is desirable not only to have a good estimate of a future value, but also some description of the uncertainty associated with that prediction. Also, the number of exogenous variables in the model should be kept to a minimum, since future values of these variates must also be forecast to compute forecasts of the variable of interest. As noted above, in order to determine a "good" forecaster, it is first necessary to determine an adequate statistical representation -- or model -- of the time series.

In some applications, a model of a time series process is sought in order to allow simulation of a process. For example, it could be desired to examine the effect on sales variability on a new inventory policy, starting from the current sales position. In such a case, we seek a mathematical model of the process which exhibits the same statistical properties as the physical process. This model can be used to generate many possible future "realizations," or outcomes, of the time series. Each outcome exhibits variability which is characteristic of the process. Hopefully, the model would involve as few parameters and exogenous variables as possible.

While the need for an adequate statistical model of a time series is apparent if we wish to simulate the time series, this need may not be so apparent if we wish to forecast. For this reason we shall briefly describe the shortcomings associated with forecasters which do not take into account the statistical characteristics of the time series.

Early Approaches to Forecasting

In the past, a number of techniques have been used for forecasting. For a series having a slowly changing deterministic component such as a trend, exponential smoothing has been applied. A drawback associated with the uncritical application of this procedure is that it in fact assumes a particular structure for the statistical model representing the process, and this structure may or may not be appropriate. This procedure in effect misinterprets, or at best, ignores, information contained in relationships between random disturbances of nearby observations of the time series. As a result, the standard deviation, or error, of forecasts is greater than it need be.

Another forecasting procedure which has been employed is the fitting of curves or patterns through recent observations using regression analysis. This technique is difficult to justify, since the "curves" so often seen in time series are often simply manifestations of the stochastic nature of the process. Similarly, the appearance of periodicities having randomly varying phase and amplitude is sometimes inappropriately modeled by using deterministic trigonometric terms (linear combinations of sines and cosines of a number of fixed periods). Stochastic periodicities such as these can arise through simple linear relationships between nearby observations.

Sometimes, as a means of avoiding certain difficulties associated with the statistical analysis, additional exogenous variables are included in the model. Unfortunately, it then becomes necessary to forecast these variables as well as the variable of interest.

Although the above techniques can involve elaborate statistical procedures, the relevance of the statistical analysis to determining a good forecast is questionable. In general, it seems advisable to avoid models which "fit the observed data well," but have little else to suggest predictive value.

A much more appropriate criterion for a forecaster is that it forecast well, rather than fit past data well. The standard measure of performance of a forecaster is the forecast error variance, or mean squared error of prediction. We would like to be able to find forecasting procedures for which the forecast error variance is small. In order to use such a criterion, however, we must be able to determine an appropriate statistical model of the time series. While specialized models of limited applicability have been available for some time, it is not until just recently that a powerful technique of wide applicability has been developed for building time series models.

Forecasters Based on Statistical Time Series Models

Recently, Box and Jenkins* have developed a very general class of time series models, which are capable of describing the statistical characteristics of a tremendous variety of time series. Some of the desirable properties of these models are the following:

- the models can describe time series which are stationary or nonstationary (a series of monthly sales which change level or change trend is an example of a nonstationary time series);
- the models can be used to represent seasonal or non-seasonal time series;
- the models can generally represent time series without the necessity for introducing exogenous variables;
- the models involve very few parameters, thereby avoiding the "curve-fitting" or "seasonal pattern" pitfalls: the possibility of deriving a model that fits the observed data well but is of little predictive value.

---

* G. E. P. Box and G. M. Jenkins, Time Series Analysis, Forecasting, and Control, Holden-Day, Inc., San Francisco, Calif., 1970.

The models studied by Box and Jenkins are technically called mixed autoregressive moving average time series models. Since the models are probabilistic representations of the time series process, they can be used for simulation. Simulation is accomplished very easily, since the model expresses the current observation in terms of past observations and random disturbances (sampled for the simulation).

From the Box-Jenkins model of a time series, it is a straightforward matter to determine the minimum-variance linear forecaster -- of all forecasters which are linear combinations of past observations, this one has the least variance. The standard errors of the forecasts are readily computed, and can be used to define tolerance limits around the forecasts. Since the models are empirical, the forecasters need not depend on exogenous variables. The models have the satisfying property of relating future observations to past observations, rather than identifying "curves" or "seasonal patterns."

An interesting characteristic of the mixed autoregressive moving average forecaster is its adaptive nature. For example, a time series may exhibit trigonometric behavior, but with stochastically varying phase and amplitude. The Box-Jenkins forecaster can handle this type of situation very well, in that the forecast "adapts" to the recent behavior of the data. This property corresponds to the allowance in the model for a continuous distribution of frequencies in the data, rather than a fixed number of specified frequencies. Furthermore, the model can adapt to seasonal periodic behavior which is nontrigonometric.

Lambda's computer program, TIMES, performs the analysis required to develop the Box-Jenkins models, and computes minimum-variance forecasts from those models.

III. THE TIMES MODEL

The basic statistical model analyzed by the TIMES program is the mixed autoregressive moving average model, which in its simplest form can be written as

$$\begin{aligned} z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} + \ldots + \phi_p z_{t-p} \\ + a_t - \theta_1 a_{t-1} - \ldots - \theta_q a_{t-q} + \mu \end{aligned} \tag{1}$$

where $z_t, z_{t-1}, \ldots$ is the observed time series, $\phi_1, \ldots, \phi_p, \theta_1, \ldots, \theta_q, \mu$ are model parameters to be estimated by TIMES, and $a_t, a_{t-1}, \ldots$ is a "white noise" sequence (i.e., a sequence of uncorrelated random variables with zero mean and constant variance σ^2). This model has been found very useful for modeling sales, prices, and other nonstationary phenomena. By including the "θ-terms" in the model, autocorrelated disturbances observed in the data can generally be accounted for without the inclusion of exogenous variables in the model.

The model can be written as

$$\Phi(B)z_t = \Theta(B)a_t + \mu$$

where

$$\Phi(B) = 1 - \phi_1 B - \ldots - \phi_p B^p \quad ,$$

$$\Theta(B) = 1 - \theta_1 B - \ldots - \theta_q B^q \quad ,$$

and B is the backward shift operator, defined by

$$Bz_t = z_{t-1} \quad .$$

Seasonal components can be included in the model by requiring

$$\Phi(B) = \Phi_{s_1}(B^{s_1}) \ldots \Phi_{s_c}(B^{s_c})$$

and

$$\Theta(B) = \Theta_{s_1}(B^{s_1}) \ldots \Theta_{s_c}(B^{s_c})$$

where $s_1, s_2, \ldots, s_c$ are the seasonal periods. Thus seasonal components correspond to multiplicative structure in the autoregressive and moving average polynomials, $\Phi(B)$ and $\Theta(B)$. Nonstationarities are handled by requiring the autoregressive polynomial $\Phi(B)$ to possess zeros on the unit circle, typically through the inclusion of difference factors; i.e., $\Phi(B)$ can be factored as

$$\Phi(B) = \Phi'(B)(1-B)^d$$

$$= \Phi'(B)\nabla^d$$

where ∇, the backward difference operator, is defined by

$$\nabla z_t = (1-B)z_t = z_t - z_{t-1} \quad .$$

In addition to the model shown above, TIMES is capable of simultaneously handling deterministic components such as trigonometric terms and exogenous variables.

IV. WHAT TIMES DOES

Calculating the estimates of parameters in a mixed autoregressive moving average model is quite involved, since the model is, in general, a nonlinear statistical model. It is this complexity that motivated the development of the general purpose time series analysis program, TIMES. The primary function of TIMES is, of course, estimation of the parameters of the model, i.e., the ϕ's, the θ's, μ, and σ^2 (as well as the coefficients of any other variables included in the model). This determination of the model is accomplished by the subroutine ESTIMATE. Once the model has been determined, then it may be used as a basis for either forecasting or simulation. Forecasts are computed by the subroutine FORECAST, and simulation is performed by the subroutine SIMULATE. We shall now describe these three TIMES subroutines in further detail.

The Model Estimation Program "ESTIMATE"

The program ESTIMATE computes estimates of the parameters of the model (1) from the observed time series data. To enable examination of the statistical adequacy of an estimated model, ESTIMATE computes a number of statistics, including the autocorrelation function, the partial autocorrelation function, and the spectral density function of the model residuals. These statistics are useful not only as a test of model validity, but also they are quite helpful, when computed for the raw data or tentative models, in guiding selection of the form of improved models.

There are, of course, many other "time series" programs which compute these statistics. The essential feature of ESTIMATE concerns its very general capability for determining empirical models involving very few parameters. The ESTIMATE program performs the following computations:

1. Estimates of model parameters
2. Covariance matrix of parameter estimates
3. Hotelling's T^2 test of significance of all parameter estimates
4. Model residuals (estimates of a's); plot of residuals; plot of histogram of residuals
5. Mean and variance of residuals
6. Autocorrelation and partial autocorrelation functions of residuals; plot of autocorrelation function (correlogram)
7. Chi-square test of significance of autocorrelations

8. Periodogram and Bartlett estimate of spectrum of residuals; plot of Bartlett spectrum.

We shall now describe each of these items in somewhat more detail.

1. Estimation of Model Parameters

The least-squares technique is used to estimate the parameters (ϕ's, θ's, μ, σ^2) of the model. This method finds estimates that have smaller variance than any other linear estimation procedure. The estimation procedure makes no assumptions concerning the form of the distribution of the a's (such as normality). Whenever parameters from different seasonal components are present in the model, or whenever θ's are present, the model is a nonlinear statistical model, and an iterative procedure is used by ESTIMATE to compute the estimates.

2. Covariance Matrix of Parameter Estimates

The covariance matrix of the parameter estimates is used in testing hypotheses about the significance of the model parameters.

3. Hotelling's T^2 Test of Significance of All Parameters

The estimates of the model parameters are, of course, statistically dependent, and a test of their significance must recognize this fact. The Hotelling T^2 statistic indicates the significance of the model as a whole. Of course, the F-distribution is appropriate for testing the significance of T^2 only if the a's are normally distributed; this latter condition can be tested with a Kolmogorov-Smirnov test.

4. Model Residuals

The model residuals are estimates of the a's of the model equation (1). Statistical examination of the model residuals (through the autocorrelation function or spectrum) indicates the statistical adequacy of a tentative model. Having estimates of the a's enables identification of their probability distribution. Tests of hypotheses concerning the model, and simulation using the model, are then possible. A histogram (estimate of the probability density function of the a's) is plotted by ESTIMATE.

5. Mean and Variance of Residuals

The mean of the residuals should not be statistically significantly different from zero in a correct model. The mean and variance (σ^2) of the residuals are estimated, and a t-statistic (for testing significance of the mean in the case of normally distributed a's) is given.

6. Autocorrelation and Partial Autocorrelation Functions of Residuals

The autocorrelation function indicates whether or not the residuals are stationary, and whether or not they can be regarded as a "white noise" sequence. If, after transforming to achieve stationarity, the residuals are not white, then the nature of the autocorrelation function and partial autocorrelation function indicates how the model should be modified. The program prints out the standard deviation of each autocorrelation and partial autocorrelation estimate. For large samples, the autocorrelations are approximately normally distributed and the t-values provided by ESTIMATE can be used to test the significance of individual autocorrelations. A correlogram (plot of the autocorrelation function) is provided by ESTIMATE.

7. Chi-Square Test of Significance of Autocorrelations

A chi-square test of significance of the autocorrelations as a whole is computed.

8. Periodogram and Bartlett Estimate of Spectrum

By taking the Fourier transform of the autocorrelation function, we obtain the periodogram, which can be used to indicate the presence of deterministic periodicities. Also, the spectrum is estimated using a Bartlett spectral window, and plotted. If the residuals are "white", then the estimated spectral density function should be reasonably "flat."

The Time Series Forecasting Program "FORECAST"

Once a model has been determined by ESTIMATE, the program FORECAST is used to compute forecasts, using this model and all observed past data. Of all forecasters which express the forecast value as a linear combination of previous observations, this one has least standard error. That is, it is the minimum-variance forecaster. This forecaster is in fact the conditional expected value (mean) of the time series in the future, given the past data. The least-squares forecast associated with the TIMES model is obtained by using the model equation (1), inputting observed or previous forecast values for the z's, and inputting least-squares estimates for the past a's. The FORECAST program computes estimates of the past a's. Future a's are replaced by their expected value, zero.

Given an observed time series, the program FORECAST computes the forecast as far into the future as desired. Naturally, the reliability of the forecasts decreases as we move further into the future. One of the advantages of the model (1) is that it is possible to calculate directly the standard error (measure of uncertainty) associated with the forecasts. The FORECAST program computes the standard error of each forecast value.

FORECAST can forecast any model of the form (1), including seasonal components, but excluding exogenous variables. Note that, once the TIMES model has been determined by the ESTIMATE program, the model can be used independently with FORECAST to compute forecasts.

A note of caution is in order here regarding the use of models in which exogenous variables have been included. The conditions under which the data are collected determine the way in which the model can be legitimately used to predict the performance of the system under study. Suppose that we are trying to forecast the level of a variable z (the variable of primary interest), and that we have data on z and another variable, x (an explanatory variable). Generally, such a model can answer only one of the following two questions:

1. What is the predicted change in z corresponding to changes in x, in the case in which we simply observe changes in x, i.e., we do not interfere with the normal operation of the system?

2. What is the predicted change in z corresponding to changes in x, in the case in which we arbitrarily change x, i.e., we interfere with the normal operation of the system?

If the model is to be used to answer the first question, it must be developed from data obtained simply by observing the normal operation of the system. If the model is to be used to answer the second question, it must be developed from data for which independent random changes have been made in the explanatory variable, i.e., we need data from a designed experiment. The reason why a model developed from the first type of data cannot be used to answer the second question is that the observed data may have been influenced by a third, unknown variable which affected both x and z; there may in fact be no real relationship between x and z. In a designed experiment, randomization eliminates this source of confusion, since the changes in x are known to be independently generated.

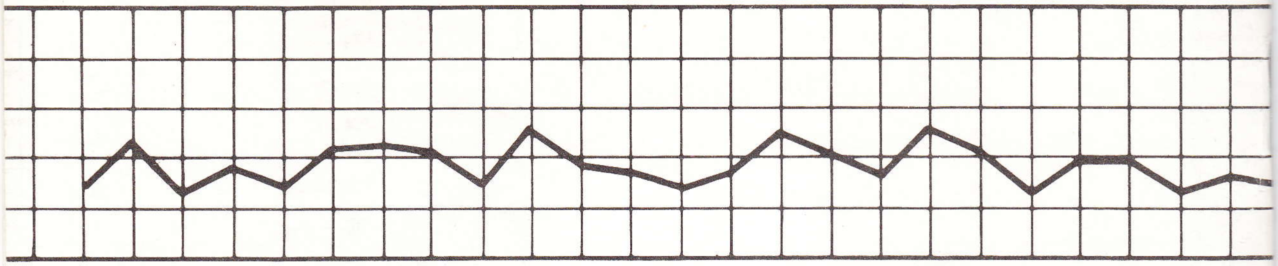
The Time Series Simulation Program "SIMULATE"

The statistical model determined by ESTIMATE can be used to simulate the stochastic process which generates the observed time series data. A simulation can be regarded as a random projection into the future, according to the same probability law associated with the past. It contrasts with a forecast, which is an expected-value projection. The subroutine SIMULATE computes simulations of the time series using the estimated model.

To simulate a possible future realization of the time series using observations up to the current time, all that is required is the parametric model estimated by ESTIMATE, together with the probability distribution of the a's. This distribution may be estimated by the histogram of the model residuals, given by the ESTIMATE program. Alternatively, a parametric form for the distribution, such as normal, may be suggested by the empirical distribution; a Kolmogorov-Smirnov goodness-of-fit test can be used to test the appropriateness of such a parametric distribution. To simulate a future realization of the time series, we sample independent a's from this distribution for each future time period, and use the recursion relation defined by the model to compute the simulated values of the time series variate. To start the simulation, however, we need estimates of past a's. The program SIMULATE computes these estimates from the observed time series data.

To simulate a realization of the time series from an independent start (i.e., a realization which does not start from observed data), it is necessary to sample p dependent starting values for the z's. Once this has been done, the simulation continues as above.

SIMULATE can simulate any model of the form (1), including seasonal components, but excluding exogenous variables. As is the case with FORECAST, SIMULATE may be used independently of the estimation program ESTIMATE once the TIMES model has been determined.



TIMES Office
Lambda Corporation
1501 Wilson Boulevard • Arlington, Virginia 22209
(703) 528-8200

TIMES is a product of Lambda Corporation, an operations research group specializing in developing computer systems for solving problems of critical importance to management. . .resource allocation, production scheduling, sales forecasting, information systems.