RAIRO. Recherche opérationnelle

Warren L. Young

The Box-Jenkins approach to time series analysis and forecasting : principles and applications

RAIRO. Recherche opérationnelle, tome 11, nº 2 (1977), p. 129-143

<http://www.numdam.org/item?id=RO_1977__11_2_129_0>

R.A.I.R.O. Recherche opérationnelle/Operations research
(vol. 11, n° 2, mai 1977, p. 129 à 143)

THE BOX-JENKINS APPROACH TO TIME SERIES ANALYSIS AND FORECASTING: PRINCIPLES AND APPLICATIONS (*)

by Warren L. YOUNG (1)

Abstract. — *In this paper, we first outline Box-Jenkins (B-J) notation and the stages in their approach to time series analysis and forecasting, i. e. identification, estimation and diagnostic checking, and forecasting. We then describe the associated statistical tools used in the choice of B-J models and also discuss seasonal B-J models. We then go on to deal with a specific economic time series, identifying and estimating variant B-J models accordingly. After choosing a model for forecasting purposes, we compare its forecasts with those generated by an econometric model. Finally, we describe extensions of the B-J approach such as "transfer function modelling" and multivariate analysis, and discuss the recently proposed method of "intervention analysis" utilizing the B-J approach.*

INTRODUCTION

At the most general level of analysis, there exist three ways of representing fluctuations in a time series. We can attempt to explain its movements in terms of factors related to the series, that is, by movements in the series itself. Alternatively, we can try to explain variations in one series by movements in another or others, i. e. using bivariate or multivariate methods. Finally, we can attempt to combine these two approaches in one way or another. As for the first approach, there are a number of ad-hoc methods that can be used in this regard, e. g. exponentially-weighted moving averages, while multiple regression techniques based on a priori considerations are usually applied if the second approach is used. With respect to the combination of approaches, this again is usually done on an ad-hoc, ex-post factotem basis. In this paper, however, we focus on a method of time series analysis and forecasting based upon the methodological principle of allowing the data to "speak for itself", that is, the Box-Jenkins approach. The material presented below is divided into two main parts. The first deals with the underlying statistical rationale of and the basic analytical principles involved in the Box-Jenkins approach. In the second part of the paper, we give an example of the Box-Jenkins technique as applied to a specific problem in time series analysis and forecasting — comparing its results to those obtained by using econometric methods — and then go on to briefly discuss some extensions of the Box-Jenkins approach itself [1].

(*) Reçu juillet 1975, révisé juin 1976.
(1) Ministry of Agriculture, Israël.

I. PRINCIPLES

A. Integrated autoregressive moving-average processes and Box-Jenkins notation

Initially, let us take, for example, a non-seasonal time series $\{X\}$; we shall deal with the problem of seasonality below. According to Box and Jenkins, we can explain its movements—in the most fundamental way—by a combination of past movements in the series itself and white noise, i. e. a series of identically distributed uncorrelated deviates. Box and Jenkins propose that there exists an integer (d) which ensures that the equation

$$w_t = (1-B)^d X_t \tag{1}$$

is stationary, where $B^j x_t = x_{t-j}$. Thus, we can express the value of the series at time (t) in the following manner

$$w_t = \frac{\theta_q(B)^q}{\varphi_p(B)^p} a_t \tag{2}$$

where a_t is "white noise", while $\theta_q(B)^q = (1-\theta_1 B \ldots -\theta_q B^q)$ and $\varphi_p(B)^p = (1-\varphi_1 B \ldots -\varphi_p B^p)$ accordingly.

Combining these two equations gives us the general form of Box-Jenkins model which applies to non-seasonal time series, i. e.:

$$\varphi_p(B)^p (1-B)^d X_t = \theta_0 + \theta_q(B)^q a_t, \tag{3}$$

where θ_0 is a constant which differs from zero [2].

In Box-Jenkins notation, then, equation (3) represents an autoregressive integrated moving-average (ARIMA) process of order (p, d, q), where p is the degree of the autoregressive process, d is the degree of differencing, and q is the degree of the moving-average process accordingly. The condition that all roots of the two polynomial equations in B, i. e.: (a) $\varphi(B) = 0$, and (b) $\theta(B) = 0$, lie outside of the unit circle both ensures the stationarity of w_t—in reference to (a)—and satisfies the Box-Jenkins "invertibility requirement"—in reference to (b)—thus guaranteeing that the model as specified is uniquely "representative" [3].

B. Stages in the Box-Jenkins approach

There are three stages in the Box-Jenkins approach to time series analysis and forecasting—identification, estimation and diagnostic checking, and the forecasts themselves. At the identification stage, we first choose a set of temporary values for the parameters p, d, and q based upon an identification procedure which shall be outlined below. We then obtain the initial estimates for the coefficients $\varphi_1, \varphi_2, \ldots, \varphi_p$ and $\theta_1, \theta_2, \ldots, \theta_q$. Diagnostic checks are then made in order to determine the representativeness of the model vis-a-vis the data set. If, as a result of these checks, an alternative form of model is suggested, then the cycle is repeated up to this point. Finally, forecasts are made on the basis of the final model specification as obtained from the estimation process and its associated choice of models criterion [4].

The primary tools in the Box-Jenkins identification procedure consist of the auto-correlation and partial autocorrelation functions. For example, let us take w_t as a stationary process with mean u. An autocorrelation of order k in this case is simply the correlation between w_t and w_{t-k}, that is,

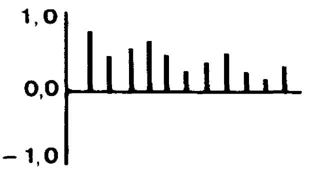
$$\rho_k = E\{(w_t-u)(w_{t-k}-u)\}/E\{(w_t-u)^2\} \qquad \text{with} \quad \rho_{-k} = \rho_k.$$

The partial autocorrelation of order k between w_t and w_{t-k} can be expressed as φ_{kk}, and is given by the equation $\rho_j = \sum_{i=1}^{k} \varphi_{ki}\,\rho_{j-i}$, where $j = 1, 2, \ldots, k$.

It will be noted that under the condition $(1-\varphi_1 B \ldots -\varphi_k B^k)w_t = a_t$ the situation $\varphi_{kk} = \varphi_k$ and $\varphi_{k+j,\,k+j} = 0$ also pertains, for all $j \geqq 1$ [5].

In any event, we how turn to the main points in the Box-Jenkins model identification procedure:

(i) If the series $\{X_t\}$ is not stationary, i. e. $d \neq 0$ in equation (3), the autocorrelations will not decrease quickly for higher values of k (number of lags), and thus differencing is necessary in order to obtain series stationarity. The autocorrelations in this case are shown in graphic form below:

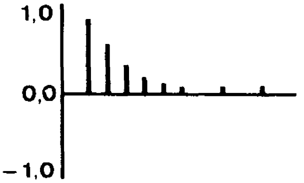


(ii) Let us assume that we have applied a sufficient degree of differencing in order to obtain series stationarity. In this case:

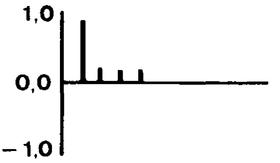
(*a*) Supposing $q = 0$, so that w_t is an autoregressive series of order p, e. g. ARIMA (1, d, 0), then the autocorrelations will die out—for all values of k—following the difference equation

$$\rho_k = \sum_{i=1}^{p} \varphi_i \rho_{k-i},$$

or, in other words, they will damp out in exponential and/or sine wave forms, where the partial autocorrelations will—for all values of $k > p$—themselves take on the value $\varphi_{kk} = 0$. The autocorrelation structure in this situation is illustrated below:



(*b*) If, on the other hand, $p = 0$, in which case we can say that there is a moving-average process involved of order q, e. g. ARIMA $(0, d, 1)$, the autocorrelation structure is of the form $\rho_k = 0$ for all $k > q$, whereas the partial autocorrelations die out according to a mixture of damped exponentials and sine waves. This situation is illustrated in the figure below, which shows the autocorrelation function generated:



(*c*) In the case where $p \neq 0$ and $q \neq 0$, w_t would be an ARIMA process of order (p, d, q). In such a situation, the autocorrelation structure follows the difference equation

$$\rho_k = \sum_{i=1}^{p} \varphi_i \rho_{k-i}$$

for all lags $k > q$, that is, it dampens out by a combination of exponentials and sine waves.

Practically, however, we don't know the true autocorrelation and partial autocorrelation structure and must, therefore, estimate them on the basis of the sample data available to us. Let us take, for example, a sample series w_1, w_2, w_m, ... which is part of the global series w_t. In this case, we can obtain an estimate of ρ_k according to the equation

$$r_k = \frac{1}{n} \sum_{t=k+1}^{n} (w_t - \bar{w})(w_{t-k} - \bar{w}) \Big/ \frac{1}{n} \sum_{t=1}^{n} (w_t - \bar{w})^2,$$

where $\bar{w}$ is the mean of the sample.

The estimate of φ_{kk} is, then, $\hat{\varphi}_{kk}$, which is obtained by solving the following set of equations

$$r_j = \sum_{i=1}^{k} \hat{\varphi}_{k_i} r_{j-i},$$

where j goes from 1 to K. We must, then, rely on the fact that the autocorrelations and partial autocorrelations of the sample are identical to those of the global series. This will hold the larger the sample, and thus we require a sample which is moderately large so as to confidently identify a specific model [6].

We obtain estimates of the coefficients of equation (3) as a result of the minimization of the sum of squares of the errors $\sum a_t^2$ by using a non-linear regression technique. Following from this, standard errors of estimate and confidence intervals for the parameters are derived, and their significance can be tested accordingly. We require, however, starting estimates for the parameters so as to initialize the iterative procedure involved in the non-linear regression technique utilized. These initial estimates are obtained on the basis of the autocorrelations generated at the identification stage [7].

There are a number of tests that can be performed so as to be able to ascertain whether the model specified on the basis of the initial identification and estimation procedures is that which most adequately represents the data. As such, the simplest and best approach in this regard would be to estimate a more general model, i. e. one which encompasses more parameters than that actually identified. Examination of the statistical significance of the additional parameters would then show whether it is necessary to include them in the model itself. But, in any case, it is inadvisable to add both extra autoregressive *and* moving average parameters, since, if the model as originally specified is that which best represents the data, there will be a high correlation between the estimates obtained through adding parameters, in addition to their having high standard deviations, so that the estimation process would not necessarily even converge [8].

Box and Jenkins suggest a number of tests on residuals generated by the model. If we call them $\hat{a}_t$, their autocorrelations $r_k(\hat{a})$, and suppose that the true errors series $\{ a_t \}$ is actually „white noise", then we can say that $r_k(\hat{a})$ has a zero mean and standard deviation approximately equal to $(1/\sqrt{n})$. These autocorrelations represent all deviations from typical white noise behaviour in the residuals and can even suggest an alternative specification for the model itself. A general—but not too powerful—test for the existence of white noise in this case is a comparison of the values of Q and chi-square values for $(M-p-q)$ degrees of freedom, where

$$Q = M \sum_{k=1}^{M} r_k^2(\hat{a})$$

and M is a number greater than or equal to 20. In this test it is possible to use the cumulative periodogram of the residuals to see whether there is regularity in the differences between white noise and the error series itself [9].

Let us suppose that we have constructed an acceptable model for the series $X_1 \ldots X_n$, but that we require forecasts of the future values X_{n+m}, where $m = 1, 2, 3$, etc. In addition, it is not difficult to suppose that if we are at time (n), the optimal forecast—in terms of the minimum forecast errors expected—of X_{n+m} is simply the conditional expected value at time (n). However, the conditional values of X_n, X_{n-1}, etc. are actually their known values at (n), whereas those for a_n, a_{n-1}, etc. are the residuals generated by the model itself, and those for a_{n+1}, a_{n+2}, etc. are equal to zero, while those for X_{n+1}, X_{n+2}, etc. are the forecasts made at (n). The forecast method, if we are given a model of equation form (3) is, therefore, an iterative process, where we take values of $m = 1, 2, 3$, etc. and substitute $(n+m)$ in place of (t) in the equation [10].

Box and Jenkins also suggest a general model which can represent seasonal series in the following manner:

$$\varphi_p(B)^p \Phi_s(B)^s \nabla^d \nabla^s X_t = \theta_q(B)^q \theta_s(B)^s a_t,$$

i. e., an ARIMA model of order (p, d, q) $(P, D, Q)_s$, where the period s is equal to four in the case of quarterly data, and equal to twelve for monthly data, and $\varphi_p = 1-\varphi_1 B-\varphi_2 B^2, \ldots, \Phi_s = 1-\varphi_{1\,s} B, \ldots, \nabla^d = (1-B)^d$, $\nabla^s = (1-B^s)^d$ respectively. Of course, there is no need to use the multiplicative model if the identification procedure suggests this. For example, we may consider two alternative model specifications for quarterly data, viz.

(*a*) $$w_i = (1-B)(1-B^4) X_t = (1-\theta_1 B)(1-\theta_4 B^4) a_t,$$

i.e. ARIMA (0. 1, 1)(0, 1, 1)$_s$

and

(*b*) $$w_i = (1-\mathrm{B})(1-\mathrm{B}^4)\, \mathrm{X}_t = (1-\theta_1 \mathrm{B}-\theta_4 \mathrm{B}^4-\theta_5 \mathrm{B}^5)\, a_t,$$

i.e. ARIMA (0, 1, 1)(0, 1, 2)$_s$.

Now, if $\theta_5 = -\theta_1 \theta_4$, the specifications are identical. If, however, the autocorrelation structure of w_t suggests that such a relationship doesn't exist, then it is preferable to utilize the latter formulation, i. e. (*b*), which is a more general one. Principles relating to fitting models and forecasting in the case of seasonal time series are basically the same as those for non-seasonal series [11].

II. APPLICATIONS

A. Short-term forecasting

In this section, we give an example of the Box-Jenkins technique as applied to a specific problem in short-term forecasting, and compare the results obtained to those generated by econometric methods. The series we take as an example is the number of hire-purchase contracts for new automobiles in England and Wales from 1958-1969 (quarterly). Allard, in his study of hire-purchase made while at the British Treasury, found that the principle variables which explained movements in this series—which gave a general impression of the effects of government policy on the credit purchase of consumer durables—were as follows: (*a*) personal disposable income, (*b*) companies trading profits, (*c*) the rate of purchase tax, and (*d*) a composite variable based upon the rate of interest on hire purchase contracts, the maximum repayment period, and the minimum deposit rate [12].

Below, we consider ex-ante, post-sample forecasts made by Allard's model in comparison to those made by using the Box-Jenkins technique for the same period. But first, let us look at the results of applying the Box-Jenkins method to the series itself. The autocorrelations and partial autocorrelations of the original series (X_t), in addition to regularly differenced, seasonally differenced, and a combination of these two types, i. e. $(1-B)\,X_t$, $(1-B^4)\,X_t$, and $(1-B)\,(1-B^4)\,X_t$, are given in the following tables, and illustrated below:

TABLE I

Original series (X_t)

Lag-k	r_k	Φ_{kk}	Lag-k	r_k	Φ_{kk}
1.....	.67	.67	11.....	.16	.04
2.....	.37	−.14	12.....	.25	−.03
3.....	.46	.51	13.....	.08	.00
4.....	.63	.24	14.....	−.06	−.04
5.....	.38	−.40	15.....	.05	−.04
6.....	.17	.14	16.....	.09	−.08
7.....	.31	.15	17.....	−.07	−.05
8.....	.42	−.11	18.....	−.21	−.12
9.....	.21	−.06	19.....	−.16	−.14
10.....	.03	−.04	20.....	−.14	−.06

mean ($\bar{w}$) = 58.6;
variance = 499.9.

TABLE II

Regularly differenced series $(1-B)\,X_t$

Lag-k	r_k	Φ_{kk}	Lag-k	r_k	Φ_{kk}
1.....	−.04	−.04	11.....	.02	−.13
2.....	−.65	−.66	12.....	.43	−.09
3.....	−.10	−.31	13.....	−.01	.03
4.....	.67	.34	14.....	−.41	−.09
5.....	−.08	−.30	15.....	.04	.00
6.....	−.57	−.22	16.....	.34	−.06
7.....	.04	.08	17.....	.00	.00
8.....	.57	−.03	18.....	−.32	.04
9.....	−.05	.02	19.....	.03	.01
10.....	−.52	−.03	20.....	.25	−.03

mean ($\bar{w}$) = 0.9;
variance = 318.1.

TABLE III

Seasonally differenced series $(1-B^4)\,X_t$

Lag-k	r_k	Φ_{kk}	Lag-k	r_k	Φ_{kk}
1.....	.65	.65	11.....	−.11	−.09
2.....	.18	−.41	12.....	−.08	−.11
3.....	−.17	−.14	13.....	.00	.08
4.....	−.30	−.04	14.....	.05	−.02
5.....	−.21	.10	15.....	.08	.07
6.....	−.04	−.01	16.....	.08	−.07
7.....	.06	−.05	17.....	.08	.08
8.....	−.02	−.22	18.....	.08	.03
9.....	−.12	.04	19.....	.06	.03
10.....	−.13	.03	20.....	−.02	−.18

mean $(\bar{w}) = 3.3$;
variance $= 313.9$

TABLE IV

Combined differenced series $(1-B)\,(1-B^4)\,X_t$

Lag-k	r_k	Φ_{kk}	Lag-k	r_k	Φ_{kk}
1.....	.15	.15	11.....	−.01	−.02
2.....	−.16	−.19	12.....	−.06	−.21
3.....	−.31	−.26	13.....	.04	−.10
4.....	−.33	−.31	14.....	.02	−.18
5.....	−.08	−.14	15.....	.05	−.02
6.....	.10	−.10	16.....	−.02	−.16
7.....	.27	.06	17.....	.01	−.12
8.....	.00	−.22	18.....	.03	−.10
9.....	−.12	−.16	19.....	.09	.12
10.....	−.05	−.03	20.....	.00	−.09

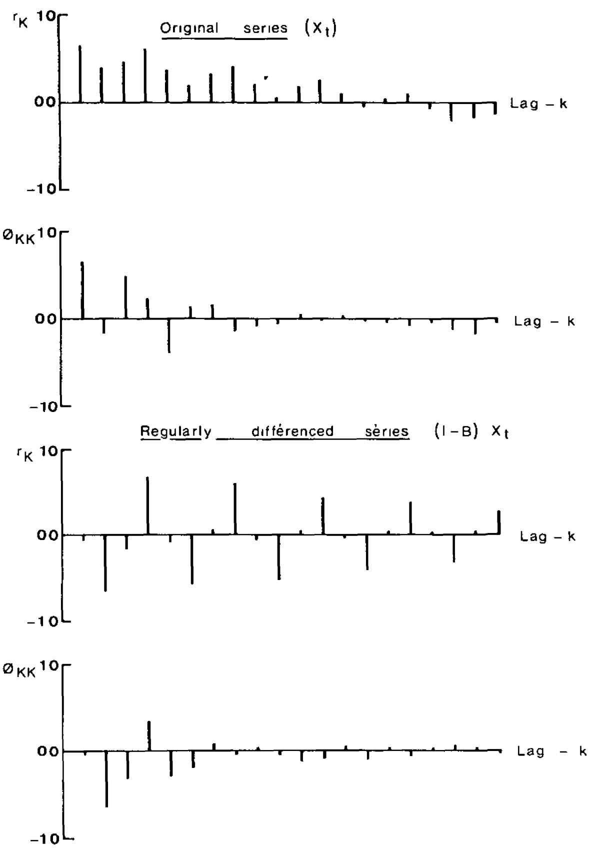
mean $(\bar{w}) = -0.2$;
variance $= 224.5$

We can draw a number of conclusions from the results as presented above:

(*a*) The autocorrelations of X_t, the original series, do not damp out, and thus we must difference the series.

(*b*) The autocorrelations of the regularly differenced series, $(1-B)\,X_t$, do not damp out for multiples of $k = 4$, meaning that some form of seasonal differencing is, therefore, required.

(*c*) The autocorrelations of the regular seasonally differenced series are adequate for model building, since those of the combined differenced series indicate "overfiltering". However, following the principle of "overfitting", we also use this form [13].



Autocorrelation and partial autocorrelation functions.
Original Series (X_t).

In addition, the results suggest that a mixed autoregressive seasonal moving average model of form ARIMA (2, 0, 0) (0, 1, 1)$_s$ can readily be applied, since, in table III, the autocorrelations die out by $k = 2$, and the "kink" at $k = 4$ would be removed by the seasonal moving average filter. One model we fit, therefore, is of the form

$$(1-\varphi_1 B-\varphi_2 B^2)(1-B^4) X_t = (1-\theta B^4) a_t.$$

The equation's parameters were estimated by a non-linear regression technique, and resulted in the following estimates, with standard errors of estimate in parentheses: $\hat{\varphi}_1 = 0.92$ (0.17), $\hat{\varphi}_2 = -0.37$ (0.19), and $\hat{\theta} = 0.21$ (0.21). The error variance as estimated (158.0) represents a decrease of 68 percent in the variance of X_t. The autocorrelations of the residuals are shown in table V below:

TABLE V

Autocorrelations of residuals $r_k(\hat{a})$

Lag-k	$r_k(\hat{a})$	Lag-k	$r_k(\hat{a})$
1......	−0.02	11......	−0.01
2......	−0.04	12......	−0.16
3......	−0.01	13......	0.00
4......	−0.02	14......	−0.02
5......	−0.11	15......	0.05
6......	−0.04	16......	−0.07
7......	0.18	17......	−0.01
8......	−0.03	18......	0.00
9......	−0.13	19......	0.04
10......	−0.04	20......	−0.06

However, 4 of the original 46 data points are "lost" by using the difference filter $(1-B^4)$, thus resulting in 42 "effective" observations in the sample. Now, the Q statistic in our case has the following value:

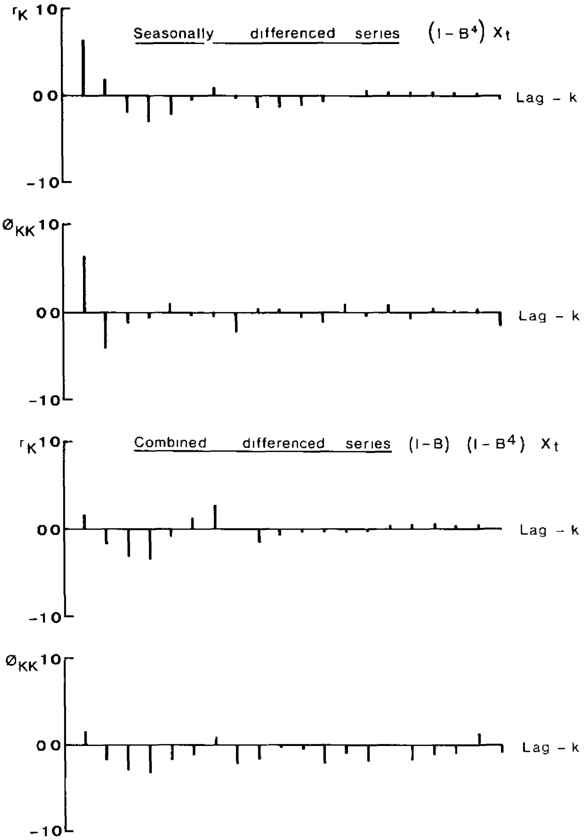
$$Q = 42 \sum_{k=1}^{20} r_k^2(\hat{a}) = 4.49.$$

The test statistic, when compared to the tabulated value of chi-square for 17 degrees of freedom, proved to be less than this value at a significance level of 99 percent. This confirms that there doesn't exist any evidence of significant departure from "white noise" i. e. randomness, as regards the error series, which, in this case, is the residuals, as noted.

But, as indicated previously, the preliminary analysis also suggests a variant model specification. This is of the form ARIMA (0, 1, 1) $(0, 1, 1)_s$, since table IV's autocorrelations die out by $k = 1$, and the partials as a "sine wave", so that the kink around $k = 4$ would be removed by a seasonal moving average filter, given that we are using the combined regularly and seasonally differenced filter $(1-B)(1-B^4)$. In this case, the model is of the form:

$$(1-B)(1-B^4)X_t = (1-\theta_1 B)(1-\theta_2 B^4)a_t,$$

and the estimated parameters were as follows : $\hat{\theta}_1 = 0.01\ (0.16)$, and $\hat{\theta}_2 = 0.56\ (0.15)$. In addition, the error variance was equal to 195.1, a good deal more than that of the first alternative. Thus, the combination of higher



Autocorrelation and partial autocorrelation functions.
Seasonally differenced series (1—B^4) X_t.

variance and insignificance in parametric estimation for one of the two moving average terms leads us to reject this specification and choose the first alternative, i. e. ARIMA (2, 0, 0) (0, 1, 1)$_s$ as preferable for forecasting purposes. An inspection of the autocorrelations shown in table V above does not suggest any alternative specification, so we used ARIMA (2, 0, 0) (0, 1, 1)$_s$ to generate forecasts accordingly.

The table below compares Allard's forecast results with those of the ARIMA model as specified up to nine periods ahead, and gives the forecast errors in percentage terms:

TABLE VI

Actual values and forecast results

	Forecast for: year an quarter, and error, in percent					
	1971 I	%	1971 II	%	1971 III	%
Forecast from and model type:						
1969 II—Box-Jenkins ..	70.3	(+5.4)	69.1	(0.0)	79.9	(−13.7)
1969 II—Allard..	72.5	(+8.7)	70.0	(+1.3)	102.0	(+ 8.4)
1970 IV—Allard..	69.5	(+4.2)	72.0	(+4.2)	106.0	(+12.4)
Actual Value.....	*66.7*	—	*69.1*	—	*93.6*	—

As shown in the table above, the Box-Jenkins forecasts from base 1969 II are better than Allard's *from* the same period in two out of three cases. Furthermore, the 1970 IV Allard forecasts are not that much better than the 1969 II based Box-Jenkins forecasts, while one of the 1969 II based ARIMA forecasts is even better than the 1970 IV based Allard forecast *for* the same period. We can conclude, therefore, that the Box-Jenkins forecasts are, on average, *better than Allard's own for the forecast periods he considers* [14].

B. Extensions of the Box-Jenkins approach

Originally, Box and Jenkins proposed that their method could be extended to encompass bivariate analysis by using a "transfer function-noise model" approach. In this case, we suppose that the variable Y_t is dependent on X_t, so that $Y_t = \alpha + \beta X_t + u_t$, where u_t is the error term. We obtain the estimates $\hat{u}_t$ and $\rho_k(\hat{u}_t)$ for $k = 1, 2, \ldots$ under the condition that $u_t = 1/(1-\varphi B)\, a_t$, where a_t is white noise. We then suppose that $\hat{u}_t$ and u_t have approximately the same autocorrelation structure, so that we obtain, by substitution

$$Y_t = \alpha + \beta X_t + \frac{1}{1-\varphi B} a_t,$$

which, after transformation, yields

$$Y_t = \varphi\, Y_{t-1} + \alpha(1-\varphi B) + \beta(X_t - \varphi X_{t-1}) + a_t.$$

The approach can be extended, therefore, to multivariate analysis, but the estimation of parameters would, in this case, prove to be quite difficult. In any event, Box and Jenkins have outlined a detailed strategy for building bivariate transfer-function type models, while work is presently being done by Granger and others to extend the Box-Jenkins approach to the multivariate case [15].

Recently, the Box-Jenkins approach has also been used to deal with the effects of "interventions" on a response variable given a dependent noise structure. In this regard, Box and Tiao have utilized such an approach—which they call "intervention analysis"—in an attempt to answer questions of the type: "given a known intervention, is there evidence that change in the series of the kind expected actually occured, and, if so, what can be said of the nature and magnitude of the change ?". Glass, who introduced the term "intervention" has, for his part, applied, along with others, the Box-Jenkins approach in order to analyze structural change and non-stationarity in time series. What both groups of researchers have found is that the Box-Jenkins framework is quite applicable to these types of problems. This results from the fact that procedures such as the "*t* test" for estimating mean changes due to an intervention are not applicable to time series, since these are often serially correlated and exhibit non-stationarity, in addition to occassionally reflecting seasonality. As such, parametric or non-parametric tests which depend on normality, constant variance, and independence of the observations are of no use in this case, while those relying on the independence or symmetry of distribution are neither available nor random in nature. Box and Tiao conclude, therefore, that the models based on ARIMA methodology which they propose for use in intervention analysis "may be readily extended to represent many situations of potential interest" [16].

REFERENCES

1. On these and related points, *see* G. Box and G. M. Jenkins, *Time Series Analysis, Forecasting, and Control*, Holden-Day, San Francisco, 1970; C. R. Nelson, *Applied Time Series Analysis*, Holden-Day, San Francisco, 1973; *Seminar on Time Series Analysis*, in: The Statistician, vol. 17, No. 3, 1967; W. L. Young, *Exponential Smoothing, Seasonality, and Projection Sensitivity: the Case of Exports*, Bull. Econ. Res. Yorkshire, vol. 26, No. 1, May 1974, and, by the same author, *Seasonality, Autoregression, and Exponential Smoothing: the Case of Economic Time Series*, Metron, International Statistical Journal, Gini Institute, Rome, vol. 31, n° 1-4, December 1973; C. Chatfield and D. L. Prothero, *Box-Jenkins Seasonal Forecasting: Problems in a Case Study*, J. Roy. Stat. Soc., A, No. 136, 1973, p. 295; T. H. Naylor, T. G. Seaks and D. W. Wichern, *Box-Jenkins Models: an Alternative to Econometric Models*, Int. Stat. Rev., vol. 40, 1972, p. 123; J. M. Bates and C. W. Granger, *The Combination of Forecasts*, Op. Res. Q., vol. 20, 1969.
2. The constant θ_0 can take on a nonzero value, for example, in cases which require "differencing". This introduces a polynomial of degree *d*, which is, by nature, deterministic, into the forecast function eventually generated. On this point, *see* Box and Jenkins, *ibid.*, pp. 91-94 and 194-195.

3. *See* NELSON, *op. cit*, Chapter 2, 3 ff.
4. *Ibid*, Chapter 5 ff.
5. An intuitive explanation regarding the nature of the partial autocorrelation function could be given as follows: if, for example, x_1, x_2, ..., x_t is a time series, then the partial autocorrelation function at, say, lag j would be the autocorrelation of x_t and x_{t+j}, under the condition that we already know the values of the observations x_{t+1}, x_{t+2}, ..., x_{t+j-1}.
6. It should be noted that the *minimum* sample for reliable Box-Jenkins analyses is about fifty observations. In addition, given a time series of length T, the autocorrelation and partial autocorrelation functions should be computed only to about $K \leqq T/4$ lags, due to the fact that for larger values of K, or as it approaches T, these estimates become quite bad. On these and related points, *see* R. L. ANDERSON, *Distribution of the Serial Correlation Coefficient*, Annals Math. Stats., vol. 13, 1942, p. 1; M. S. BARTLETT, *On the Theoretical Specification of Sampling Properties of Autocorrelated Time Series*, J. Roy. Stat. Soc., B, No. 8, 1946, p. 27; J. DURBIN, *The Fitting of Time Series Models*, Rev. Int. Inst. Stats., vol. 28, 1960, p. 223; M. H. QUENOUILLE, *Approximate Test of Correlation in Time Series*, J. Roy. Stat. Soc. B, vol. 11, 1949, p. 68.
7. BOX and JENKINS, *op. cit;* D. W. MARQUARDT, *An Algorithm for Least Squares Estimation of Non-Linear Parameters*, J. Soc. Ind. Appl. Math., vol. 2, 1963, p. 431.
8. C. R. NELSON, *op. cit*, pp. 82 ff.
9. G. BOX and D. A. PIERCE, *Distribution of Residual Autocorrelations in Autoregressive Integrated Moving Average Time Series Models*, J. Amer. Stat. Assn., vol. 65, 1970, p. 1509.
10. For a classification of this type of forecast and forecasting types in general, *see* W. L. YOUNG, *Forecasting Types and Forecasting Techniques: a Taxonomic Approach in Quality and Quantity*, Europ. Amer. J. Method., Elsevier, 1976. Also *see* NELSON, *op. cit*, Ch. 6 ff; P. NEWBOLD and C. GRANGER, *Experience with Forecasting Univariate Time Series and the Combination of Forecasts*, J. Roy. Stats. Soc., A, No. 137, 1974.
11. Practically speaking, however, seasonal modeling using the Box-Jenkins approach can prove difficult. This is due to the fact that at present, we do not know much about the theoretical behaviour of seasonal autocorrelation and partial autocorrelation functions. On these and related points, *see* CHATFIELD and PROTHERO, *op. cit;* NELSON, *Ibid*, Ch. 7 ff; T. F. SMITH, *A Comparison of Some Models for Predicting Time Series Subject to Seasonal Variation*, in: Seminar on Time Series Analysis, *op. cit.*
12. On these and related points, *see* R. J. ALLARD, *An Economic Analysis of the Effects of Regulating Hire Purchase*, H. M. Treasury, Gov't. Economic Service Occasional Papers, 9, London: HMSO, 1974; R. J. ALLARD, *Hire Purchase Controls and Consumer Durable Purchases*, Queen Mary College, Univ. of London, Dept. of Economics Discussion Paper, March 1975 (25-39-75). Also *see* R. J. BALL and P. S. DRAKE, *Impact of Credit Control on Consumer Durable Goods Spending in the United Kingdom*, 1957-1961, Rev. Econ. Stud. vol. 30, No. 3, 1963, for an alternative view of the problem.
13. On the principle of "overfitting" and "parameter redundancy", *see* NELSON, *op. cit*, p. 114.
14. R. J. ALLARD, *op. cit*, Appendix H, 1974, pp. 95 ff. The large forecasting error for 1971 III is due to the fact that hire-purchase credit controls were abolished

in July, 1971, and were also distorted by strikes earlier that year, in addition to the fact that an increasing number of automobiles were purchased with credit from sources other than finance houses, e. g. personal bank loans and overdrafts, etc.

15. BOX and JENKINS, *op. cit*; G. BOX and D. R. COX, *An Analysis of Transformations*, J. Roy. Stat. Soc., B, No. 26, 1964; G. BOX and G. M. JENKINS, *Some Recent Advances in Forecasting and Control*, Applied Stats., 17, 1968; C. GRANGER and P. NEWBOLD, *Forecasting Transformed Series*, J. Roy. Stat. Soc., B, No. 38, 1976.
16. G. V. GLASS *et al.*, *Design and Analysis of Time Series Experiments*, Colorado Associated Univ. Press; Boulder, Colo., 1975; G. GLASS, *Estimating the Effects of Intervention Into a Non-Stationary Time Series*, Amer. Educ. Res. J., vol. 9, No. 3, 1972; G. BOX and G. TIAO, *A Change in a Level of a Non-Stationary Time Series*, Biometrika, vol. 52, June 1965; G. BOX and G. TIAO, *Intervention Analysis with Applications to Economic and Environmental Problems*, J. Amer. Stat. Assn., vol. 70, No. 349, March 1975, pp. 70 ff, 72.