Journal of Forecasting, Vol. 2, 23–35 (1983)

ARIMA Model Building and the Time Series Analysis Approach to Forecasting

PAUL NEWBOLD
University of Illinois, U.S.A.

ABSTRACT

This paper reviews the approach to forecasting based on the construction of ARIMA time series models. Recent developments in this area are surveyed, and the approach is related to other forecasting methodologies.

KEY WORDS ARIMA models Time series model building Multiple time series

The major part of this paper will be devoted to a discussion of the forecasting of time series through the fitting to historical data of autoregressive integrated moving average models. We will consider the use of such models both for the analysis of a single time series, and the exploitation of relationships among series, in the production of superior forecasts. However, before specializing our discussion to these models, we consider the broader issue of the *general approach* to forecasting problems embodied in time series model building, and in the perspective of the time series analyst. In doing so, we are attempting to describe the essentially distinctive features of the time series analyst's approach to model building.

The first, and most important, aspect might seem trivially obvious. Forecasting involves *time!* The prediction 'the dollar price of gold will rise by 20 per cent' is worthless as it stands. Without some time frame, statements of this sort are more or less vacuous. Yet, in recent public policy debates we have been hearing this kind of forecast. Strict control of the money supply, it is urged, will *in the long run* permit non-inflationary economic growth. The well-known rejoinder, that in the long run we are all dead, is pertinent. Without careful consideration of the path along which the economy is supposed to travel to the predicted happy outcome there is no assurance that the patient will not be killed before it is cured. The point about timing is not trivial from a practical perspective. In many fields, particularly in the social sciences, much subject matter theory is essentially static in nature—of the sort 'an increase in X leads, all else equal, to an increase in Y'. Such theory is certainly not irrelevant to the forecaster, but, alone, it is inadequate. If the objective is prediction, then the *timing* of relationships is at least as important as their qualitative nature. I do not wish to imply that time series analysts are unique in having divined this rather obvious point, but simply that it is at the heart of their approach to forecasting problems.

The second ingredient in the time series analysis approach to forecasting is in the construction of a *model* of the historical evolution through time of the phenomenon to be predicted. In doing so we

0277-6693/83/010023-13$01.30

Revised May 1982
Revised September 1982

are implicitly assuming some smooth continuity, though not necessarily stationarity, in behaviour patterns through time. If we were not prepared to accept that history provided some guide as to what to expect in the future, then we may as well resort to the crystal ball for our forecasts. Time series model building is not an attempt to make the data fit a particular model, but rather to make a model that fits the data. In the phraseology of Jenkins (1979) we should 'fall in love with the data', not with the model. In this sense, time series analysis contrasts with many exponential smoothing procedures, where, effectively, a single model form is imposed on the data. Time series analysts have stressed the importance of allowing the data themselves a part to play in determining an appropriate model structure. This emphasis has sometimes led to the misapprehension that we advocate modelling without subject matter theory. This is certainly not so. In constructing a forecasting model, it is necessary to take account of *all* available information—theory and data. Like all good science, statistical model building involves an interaction between theory and empirical observation. The emphasis on the role of data in determining an appropriate model form arose from the recognition that, all too often, subject matter theory turns out to be very vague on the question of timing. Rather than look at one or two simple model structures, the time series analyst would argue that, as far as possible, information gleaned from the data should be used in model specification to fill the gaps left by theory. This point is becoming increasingly well-recognized. For instance, in much of the best econometric work a wide array of model forms is initially contemplated, and data employed to select from among them. A good example of this sort is contained in Hendry and Mizon (1978). Indeed, best practice is now such that it is no longer meaningful to talk about the 'econometric, or regression, approach' and the 'time series approach' as if they were separate entities. Rather, they are now frequently sensibly integrated in applied work. This has by no means always been the case, as discussed in Newbold and Reed (1979). We should therefore not think of time series analysis as an *alternative* to building regression models for forecasting. Rather, modern time series techniques should be incorporated into the model building in order to obtain the most efficient forecasts possible. This point is perhaps best understood if we begin with the notion constructing a regression model to explain the behaviour of, and subsequently predict, a variable of interest, Subject matter theory may well suggest certain relevant explanatory variables. However, rarely, if ever, will theory be sufficiently precise as to demand the use of one particular time lag structure for the relationship, to the exclusion of all other possibilities. Moreover, theory will have nothing to say about the time series properties of the regression errors, even though adequate modelling of these properties is essential for successful forecasting. Time series analytic techniques can be used, through an examination of statistics calculated from the available data, to develop appropriate lag and error structures.

A third important ingredient in time series model building lies in the concept of *parsimony*, urged by Box and Jenkins (1970). Essentially the idea is that, rather than rushing to build over-elaborate models, we should seek relatively simple structures which appear to describe reasonably well the major characteristics of our data. Parsimonious models of this sort should perform relatively well when extrapolated forward for forecasting purposes. The caution implied by the notion of parsimony is against the over-mining of data. For example, given a series of 30 observations on a time series, we know that a perfect *fit* can be obtained with a 29th degree polynomial time trend. It would, however, be foolish to expect such a fitted trend to be very successful in forecasting future observations. A good technical exposition of the importance of the concept of parsimony in time series model building and forecasting is contained in a recent paper by Ledolter and Abraham (1981). For a firm verbal statement I can, again, do no better than quote Jenkins (1979). 'Size and complexity of model structure is not a desirable objective— on the contrary it is often indicative of mediocrity on the part of the model builder.'

In the remainder of this paper, we will specialize our discussion to procedures for building autoregressive integrated moving average forecasting models. Section 1 of the paper briefly reviews the methodology proposed by Box and Jenkins (1970) for fitting such models to a single time series. Section 2 reviews recent developments in this area, whereas Section 3 provides an overview. In Section 4 we briefly discuss the Box–Jenkins procedure for fitting transfer function-noise models, which allow the use of information on a related variable to be incorporated into the forecast-generating mechanism. Section 5 outlines some of the important recent developments in multiple time series model building, whereas the final section summarizes the main points of the paper.

1. BUILDING ARIMA FORECASTING MODELS: THE METHODOLOGY OF BOX AND JENKINS

In a series of articles, and a subsequent book, Box and Jenkins (1970) set out and illustrated a methodology for fitting to time series data a member of the class of autoregressive integrated moving average (ARIMA) models. Their work has had an enormous impact on the practice of time series analysis and forecasting. By now the details of this methodology are well known, so that it is not necessary to set them out at great length here. They are discussed, for example, in the books by Nelson (1973), Granger and Newbold (1977), and Jenkins (1979). In this section, we will simply outline the Box–Jenkins model building approach. Recent developments in this area will be surveyed in the following section.

We should begin by saying there is no such thing as a Box–Jenkins forecasting *method*. This is so in the sense that these authors do not propose a uniform treatment of every forecasting problem. What they do provide is an approach to fitting to data particular types of time series models, which, when used sensibly, should prove useful in attacking a wide range of forecasting problems. The most important aspect of their work lies in the provision of a framework for model building, within which the user is encouraged to think carefully about the behaviour exhibited by the data.

The ARIMA class of models is quite broad. Nevertheless, it does restrict us to considering stationary time series, or those series that exhibit stationarity after differencing. Moreover, the models are linear, in the sense that predictions of future values are constrained to be linear functions of the observations. Now, very often, in practice, linearity and stationarity will be a sufficiently good approximation to reality. However, this will not invariably be so, in which case it would be sensible to contemplate a broader class of models.

Given the assumptions of linearity and stationarity, the ARIMA models have the virtue of parsimony of parameterization. That is, they allow the representation of a wide array of potentially useful predictor functions in models which contain relatively few parameters. For instance, rather than using ARIMA models, it is almost certainly quicker and cheaper to develop linear predictions through the fitting of moderately high order autoregressive models. However, the necessity for estimating a large number of parameters, inherent in that approach, can lead to quite inefficient forecasts. This point is elaborated on in Box and Jenkins (1973) and Davies and Newbold (1980). If autoregressive models are to be used, an attractive possibility is to use either subset or stepwise autoregression, as for example in Parzen (1982) or Kenny and Durbin (1982).

On occasion, the assumption of linearity can be relaxed somewhat by fitting an ARIMA model to a series which has been transformed, perhaps according to a member of the class power transformations of Box and Cox (1964). Empirical evidence on the value of incorporating such power transformations into the time series model is rather mixed. Nelson and Granger (1979), analysing a collection of macroeconomic time series, found little gain in forecast accuracy. On the other hand, Hopwood *et al.* (1981), who examined a collection of series of corporate earnings, found that, on the whole, worthwhile improvements in forecast quality did result when power transformations were employed.

The Box–Jenkins approach to building ARIMA forecasting models involves an iterative three stage cycle of model selection, parameter estimation and model checking.

At the initial selection stage, based on the examination of statistics calculated from the data, a specific model is chosen for further analysis from the general ARIMA class. Box and Jenkins propose that this choice be based on examination of the behaviour of the sample autocorrelations and partial autocorrelations of the original data and its low order differences. The choice of model is, then, based on judgement, rather than a terribly well-defined set of formal rules. Thus, the selection procedure is not deterministic, and certainly it is often not possible to be very sure that the 'right' choice has been made. The analyst is not, however, irrevocably committed to the initially chosen model.

Having selected a specific model from the general class, the parameters of that model are then estimated using efficient statistical techniques. Most computer packages, in common use, base parameter estimation on a non-linear least squares algorithm.

Since the initial model selection is, of necessity, rather tentative, it is important to check whether the fitted model adequately represents the behaviour of the data. Following Box and Jenkins, two types of model check are in common use. First, one can fit to the data a slightly more elaborate model than that originally selected, and assess whether the addition of further parameters yields a significant improvement in fit. Second, the autocorrelations of the residuals from the estimated model can be examined. If the specification is appropriate, these residual autocorrelations should be quite small in magnitude. It frequently happens that, in checking model specifications in this way, an alternative specification is suggested. In that case, the cycle of selection, estimation and checking is iterated until a satisfactory model structure is obtained.

Once an appropriate model has been fitted to the data, that model can be projected forward to obtain forecasts of future values of the time series. Also, given an assumption of normality, it is possible to put confidence intervals around these point forecasts.

2. RECENT DEVELOPMENTS IN ARIMA MODEL BUILDING FOR A SINGLE SERIES

The methodology for constructing forecasting models, very briefly described in the previous section, is now in common use, and is very widely known. There have recently been a number of developments, within this general framework, which may not be quite so well known. In this section we will outline a few of these developments.

In discussing the ARIMA model building methodology, both with students and forecasting practitioners, it is quite clear that the initial model selection stage of the model building causes the most difficulty, and leads to the most complaints about problems of practical implementation. These complaints are of two sorts. First, it is felt that the statistics calculated—the sample autocorrelations and partial autocorrelations—often do not provide enough information on which to base a moderately secure choice of model form. Sometimes, this is a result of inexperience on the part of the user, but nevertheless this point has some validity. Second, many people dislike the necessity of using judgement at this stage, and would prefer some deterministic scheme which yielded a single 'answer'. In one sense, this is understandable. If a large number of time series are to be analysed, model selection can be unduly time-consuming. On the other hand, no careful statistician should deny himself too hastily the right to think about the characteristics of his data.

It must be admitted that, on occasion, the sample autocorrelations and partial autocorrelations do not point as firmly as one would wish in the direction of a specific model from the general ARIMA class. In this context, the paper by Hamilton and Watts (1978), which indicates how useful information about seasonal structure can be extracted from the partial autocorrelations, is valuable. However, certainly one often feels that it would be useful to have other statistics to look at before choosing a model. A number of interesting suggestions have been made in this regard. Cleveland (1972) proposes examination of the inverse autocorrelations, whereas other statistics have been proposed in Gray *et al.* (1978), Beguin *et al.* (1980), and Woodward and Gray (1981). These quantities have, on occasion, been helpful in practical model selection. However, at the present time, my own experience with their use is too limited for a sound judgement of their value.

Another way of looking at the problem of model selection, which does lead to a definite 'answer', is through automatic criteria aimed at estimating model order. Various criteria have been proposed by Akaike (1969, 1974), Parzen (1974), Hannan and Quinn (1979), and Hannan and Rissanen (1982). Perhaps because they are automatic, these approaches are becoming increasingly popular. However, at least within the context of choosing among ARIMA models, experience remains quite limited. Perhaps the most frequently used of these criteria is the AIC criterion of Akaike. This appears to select rather over-elaborate models, as seems to be the case, for example, with the analyses reported by Ozaki (1977). Nevertheless, I feel that it is generally useful to use one of these procedures as well as, if not instead of, judgemental examination of statistics such as the sample autocorrelations.

Recent interest in parameter estimation for ARIMA models has centred on the possibility of using full maximum likelihood. Expressions for the likelihood function of an autoregressive-moving average process have been given by, among others, Newbold (1974), Ansley (1979), and Ljung and Box (1979). Simulation results, reported in Ansley and Newbold (1980), suggest a preference for maximum likelihood over two alternative least squares procedures. In particular, it was found that, on occasion, worthwhile gains in forecast accuracy were obtained when full maximum likelihood estimation was employed. Furthermore, results reported by Ansley and Newbold (1981) suggest that confidence intervals around the forecasts can be a good deal more reliable when maximum likelihood estimation is employed than when estimation is based on the least squares procedures.

In order to check the adequacy of representation of a fitted model to the given data, following Box and Jenkins (1970), the two most frequently employed checks have involved fitting a somewhat more elaborate model than that originally selected, or examining the behaviour of the residual autocorrelations from the initially selected model. Godfrey (1979) has shown how the Lagrange Multiplier test can be employed to check the fitted model against an alternative involving either k additional autoregressive, or k additional moving average, parameters. However, as indicated by Poskitt and Tremayne (1980), these two tests are, in fact, identical. The Lagrange Multiplier tests are particularly convenient, as they can be readily carried out through the estimation of an auxiliary multiple regression equation. Tests based on residual autocorrelations were first discussed by Box and Pierce (1970), whereas further details of the derivation of appropriate tests are contained in McLeod (1978). One proposal, due to Box and Pierce, contemplates a very wide range of possible alternatives to the fitted model. Let $\hat{a}_t$ denote the residuals from the model, fitted to n sample observations, so that the residual autocorrelations are given by

$$\hat{r}_k = \frac{\sum_{t=1}^{n-k} \hat{a}_t \hat{a}_{t+k}}{\sum_{t=1}^{n} \hat{a}_t^2} \qquad (k = 1, 2, 3, \ldots)$$

We can then ask whether, taken as a group, the first M residual autocorrelations are 'too large' in magnitude to support the contention that the model is appropriately specified. Box and Pierce show that, for moderately large M, the 'portmanteau' statistic

$$Q = n \sum_{k=1}^{M} \hat{r}_k^2$$

has an asymptotic chi-square distribution, with degrees of freedom M less the total number of estimated autoregressive and moving average parameters. This portmanteau test is frequently carried out in practice. However, as Davies *et al.* (1977) point out, unless the number of sample observations is very large, the null distribution of the test statistic can differ substantially from the assumed asymptotic distribution. Following Ljung and Box (1978), it seems preferable, for this reason, to use the modified test statistic

$$Q^* = n(n+2) \sum_{k=1}^{M} (n-k)^{-1} \hat{r}_k^2$$

which, of course, has the same asymptotic distribution as Q. Further evidence on the sampling behaviour of this statistic is given by Ansley and Newbold (1979a). Although the portmanteau test can be useful, it should not be employed as the *only* check on model adequacy, since, as shown by Davies and Newbold (1979), it can have very low power in detecting alternatives which represent a serious departure from the assumed model.

In fact, though checks on residual autocorrelations, and the checks against more elaborate models, are often separately carried out, they are, in a sense, not really distinct. Newbold (1980) has shown that the Lagrange Multiplier test against the alternative of k additional autoregressive parameters is identical to the test based on the first k residual autocorrelations. Hosking (1980) has also shown that the portmanteau test can be derived as a Lagrange Multiplier test.

Once a satisfactory ARIMA structure has been fitted to the data, the usual procedure is to project this forward to obtain point and interval forecasts, acting as if the estimated parameters were the true values. For this, and other reasons, the usual standard errors of the forecasts are typically understated. This issue is discussed by Ansley and Newbold (1981), who suggest a modification of the usual procedure for estimating these standard errors.

3. FORECASTING SINGLE SERIES THROUGH ARIMA MODELS: AN OVERVIEW

The great virtue of an approach to forecasting through ARIMA model building lies in the *flexibility* which is allowed. The user is not constrained to force on every time series a single model form, and hence a single specific forecast function, but rather is able to select from a wide class of models. The determination of the particular model structure appropriate in any given instance is based on the characteristics of the available data. Thus, in time series model building, data are used, not only in the estimation of the parameters of some pre-specified model, but also in assessing which model form might be most appropriate. To the extent that successful forecasting is based on an assumption of the continuation of past behaviour patterns, it is difficult to argue on *a priori* grounds with the assertion that this flexibility should be beneficial. It is, after all, well known (see, for example, Ledolter and Box (1978)) that many of the *ad hoc* exponential smoothing predictors are simply special cases of ARIMA predictors. Presumably, if such special prediction functions were appropriate, the available data would so indicate. Thus, while simplicity, or parsimony, is certainly a virtue in a forecasting model, it seems unreasonable to force the same model on every time series if the data strongly suggest otherwise. That is not to say, however, that flexibility in model choice is invariably a virtue. Once the user has been given some freedom, then he or she is, of course, free to make a poor choice. Given adequate experience, this should not happen too often. However, experience with students and other new users suggests that, when difficulties arise, they often do so as a result of attempts to build over-elaborate models. Beginners apparently find it tempting to try to 'model the noise'.

Of course, the virtue of having a broad class of models from which to choose is that, on the average, this should lead to *more accurate* forecasts. Certainly it would be expected that a sensible use of the ARIMA model building approach should, in the aggregate, produce forecasts which are of higher quality than those deriving from the imposition of a single model structure on every time series. Some numerical support for this assertion is contained in Newbold and Granger (1974). However, empirical studies by Makridakis and Hibon (1979), and by Makridakis *et al.* (1982), report little or no improvement in forecast accuracy when using ARIMA models rather than some simpler extrapolative techniques.

So far we have discussed the prediction of a time series from its own current and past values. There is, of course, no good reason for the forecaster to restrict attention to just this information. In particular, the current and past values of other, related, time series may well provide additional useful information. If possible, this information should be incorporated into the forecast-generating mechanism. As we will see in the following two sections, the ARIMA model building approach can be extended to allow the analysis of related time series.

The strengths of the ARIMA model building approach lie in its flexibility, the relative accuracy of the resulting forecasts, and the possibility of extension to the analysis of multiple time series. One cost of this flexibility is that the approach requires the moderately heavy use of skilled manpower. However, given the availability of interactive computer programs, this burden will generally not be intolerable. From time to time some relief has been sought through the development of 'automatic' ARIMA model building programs. However, the comments of Jenkins (1982) suggest that these may often fail to produce satisfactory forecasts.

My own view is that the chief drawback to generating forecasts through the building of single series ARIMA models lies, not in its flexibility, but in its relative *inflexibility*. The restriction, for example, to stationary linear models, while often useful as a good approximation to reality, will not invariably be so. On occasion it may be desirable to allow the forecast to be a non-linear function of past observations. Some recent work on non-linear models is reported in Priestley (1980). Perhaps of even more practical importance is the possibility of an abrupt structural change in a time series. This could involve an outlying observation, a change in level, or a change in trend. If ignored, such changes can have a profound impact on ARIMA model building. In particular, it is quite likely that an inadequate model choice would be made if data were routinely analysed as if the time series under study were stationary. If the structural change occurs at some known point in time, it can be incorporated into the model through the technique of 'intervention analysis', introduced by Box and Tiao (1975). This simply involves the addition of dummy variables to the ARIMA model. A far deeper problem arises when one suspects, *a priori*, the possibility of structural changes in a series, but has no extraneous information as to where these are likely to occur. Within the context of ARIMA model building, little work has been done on this problem, though Abraham and Box (1979) consider the possibility of outliers in autoregressive models. A different approach, known as Bayesian forecasting, is due to Harrison and Stevens (1976). These authors allow, at any point in time, four possible states: 'steady state', 'step change'. 'slope change' and 'outlier'. Based on the data, posterior probabilities for each state are computed. However, their approach does not make it clear how one goes about selecting an appropriate structural form for the underlying, 'steady state' model.

A final difficulty, as we have already noted, with forecasting from a single series ARIMA model is that this might involve the exclusion of important information contained in related time series. To the extent that the information is readily available, concentration on just a single series will yield forecasts which are unnecessarily sub-optimal. In the following two sections we will briefly discuss methods for building multiple time series models.

One can think of moving away from single series ARIMA model building in one of two directions—either towards a less flexible approach, leading perhaps to the imposition of a single model, or one of a small number of models, on every series, or to even greater flexibility. In the latter case, one might want to allow for non-linearity or non-stationarity, or for the building of models relating to two or more time series. The eventual choice will depend on the user's resources, the nature of the problem, the forecasting accuracy required, and the costs involved.

4. FORECASTING FROM THE TRANSFER FUNCTION-NOISE MODELS

We now turn to the case where a time series Y_t is to be predicted, and information is available on a related series, X_t. It will be assumed, in the sense of Granger (1969), that X 'causes' Y, but that there is no 'feedback', so that Y does not also cause X. This concept of causality is phrased in the language of prediction. Suppose that we try to predict Y_{t+1} on the basis of two sets of information:

$$I_0 = [Y_{t-j}; \qquad j \geq 0]$$

and

$$I_1 = [Y_{t-j}, X_{t-j}; \qquad j \geq 0]$$

If the optimal mean squared error linear predictor of Y_{t+1} based on I_1 is superior to that based on I_0, then X is said to 'cause' Y in the Granger sense. In that case, addition of past values of X to past values of Y in the set of information on which forecasts are based leads to better predictions.

Suppose, now, that X_t and Y_t are stationary time series, each with zero mean. (This may be accomplished by suitable transformation, for example differencing, of the original observed time series.) Then a general linear model, linking the current value of Y to current and past values of X, is

$$Y_t = V_0 X_t + V_1 X_{t-1} + V_2 X_{t-2} \cdots + e_t \tag{1}$$

where e_t is an error term that is not necessarily white noise. Employing the back-shift operator B, defined so that $B^j X_t \equiv X_{t-j}$, equation (1) can be written

$$Y_t = (V_0 + V_1 B + V_2 B^2 + \cdots) X_t + e_t \tag{2}$$

Now, the difficulty with the formulations (1) and (2) is that their parametrization may not be terribly parsimonious. On the contrary, if the relationship between the output Y and the input X involves long time lags, an adequate representation would contain a large number of coefficients V_j. One possible solution to this dilemma is to write, approximately

$$V_0 + V_1 B + V_2 B^2 + \cdots = \frac{\omega_0 + \omega_1 B + \cdots + \omega_r B^r}{1 - \delta_1 B - \cdots - \delta_s B^s} \tag{3}$$

where the ω_i and δ_j are parameters, and one hopes that an adequate representation of the relationship between the input and output series can be achieved with relatively small values for the polynomial orders, r and s. Substituting (3) into (2), and allowing the error term e_t to follow an autoregressive-moving average process, leads to the transfer function-noise model

$$Y_t = \frac{\omega_0 + \omega_1 B + \cdots + \omega_r B^r}{1 - \delta_1 B - \cdots - \delta_s B^s} X_t + \frac{1 - \theta_1 B - \cdots - \theta_q B^q}{1 - \phi_1 B - \cdots - \phi_p B^p} a_t \tag{4}$$

where a_t is white noise.

In analysing causal relationships of the sort discussed here, the time series analyst's view is that, for many applications, subject matter theory will be useful in suggesting explanatory variables, X, but is likely to be far less useful in postulating a specific structure for the timing of the relationship—that is, in the determination of appropriate values for r and s in (4). Needless to say, this determination is critically important in the development of a successful forecasting model. Thus, this vital role is generally assigned, in whole or in part, to the data. That is to say, the choice of appropriate values for r and s in (4) is based primarily on information contained in the available data. It is well known that, unless the error structure of a model such as (4) is well specified, very poor forecasts can result. However, this is an area in which subject matter theory has even less of value to contribute. Thus, choice of appropriate values for p and q will typically be based almost exclusively on the data. We can think of the process of building forecasting models, relating a series to be predicted to one or more explanatory variables, or inputs, as involving two stages. Initially, subject matter considerations should suggest possibly relevant explanatory variables. Subsequently, the precise form of the model to be employed in forecasting is likely to be dictated primarily by an appeal to the data.

Box and Jenkins (1970) develop and illustrate a model building strategy, involving selection, estimating and checking, for the fitting to data of transfer function-noise models. Poskitt and Tremayne (1981) provide a discussion of model checks based on the Lagrange Multiplier test.

In principle, the analysis can be readily extended to the case of several input series. An illustration is provided by Jenkins (1979).

A rather different approach, using regression models with lagged dependent variables and autoregressive error structures, is sometimes used in the econometric literature to fit dynamic relationships. This procedure is illustrated by Hendry and Mizon (1978).

5. FORECASTING FROM MULTIPLE TIME SERIES MODELS

The procedures referenced in the previous section are suitable for relating time series when causality runs in one direction only. However, they are not appropriate when the series exhibit feedback. In that case the most useful approach to constructing forecast functions appear to be through the building of vector autoregressive-moving average models.

Let $\mathbf{X}_t' = (X_{1,t}, X_{2,t}, \ldots, X_{m,t})$ be a vector of observations, made at time t, on a set of m series which are jointly stationary, and have zero means. (This may be achieved through some transformation, such as differencing, of the originally observed time series.) Then, the vector autoregressive-moving average model of order (p, q) is written

$$\mathbf{X}_t - \Phi_1 \mathbf{X}_{t-1} - \cdots - \Phi_p \mathbf{X}_{t-p} = \mathbf{a}_t - \mathbf{\Phi}_1 \mathbf{a}_{t-1} - \cdots - \mathbf{\Phi}_q \mathbf{a}_{t-q} \tag{5}$$

where the Φ_i and $\mathbf{\Phi}_j$ are $m \times m$ matrices of coefficients, and the process $\mathbf{a}_t$ is vector zero-mean white noise, so that

$$E(\mathbf{a}_t) = \mathbf{0}$$

and

$$E(\mathbf{a}_t \mathbf{a}_{t-k}') = \begin{matrix} \Sigma(k = 0) \\ 0(k \neq 0) \end{matrix}$$

where Σ is a positive definite matrix.

In the last year or two, a good deal of research on methods for fitting to data models of the form (5) has been reported. In particular, two major papers, by Jenkins and Alavi (1981) and Tiao and Box (1981), have set out and illustrated model building strategies involving an iterative cycle of selection, estimation and checking. These methodologies are extensions of the Box–Jenkins approach in the single series case, outlined in Sections 1 and 2 of this paper.

Even a casual reading of these two articles is enough to suggest that the severest difficulties involved in this extension occur at the initial model selection stage, where tentative values must be chosen for p and q in (5). As we have already noted, this problem is by no means invariably trivial, even in the single series case. However, in the multiple series case, one element of the single series approach is no longer of value. Certainly, in the multivariate situation, we can find matrices of sample autocorrelations and partial autocorrelations, which should prove useful in identifying pure autoregressive or pure moving average behaviour. (The use of partial autocorrelations is discussed by Ansley and Newbold (1979b).) However, in the single series case, one way to recognize particular mixed models is through visual inspection of the patterns of sample autocorrelations. In the multiple series framework, this is a pretty hopeless proposition, as the corresponding patterns would now be followed by *matrices* of order $m \times m$, where m is the number of time series. The question as to how best to go about initial model selection here remains open, and is the subject of much on-going research.

Given an initially chosen model, the remaining steps are rather less difficult to execute. Estimation of the model parameters by maximum likelihood is discussed by Nicholls and Hall (1979) and by Hillmer and Tiao (1979). The Lagrange Multiplier test of model adequacy is considered by Poskitt and Tremayne (1982), whereas Hosking (1981) and Li and McLeod (1981) discuss the extension of the portmanteau test to the multivariate case. Further discussion of these and other checks on model adequacy is contained in Newbold (1982).

6. SUMMARY AND OVERVIEW

In the previous sections of this paper, we have discussed, in broad outlines, a time series analytic approach to model building and forecasting, through the use of ARIMA models. In the cases of single series and transfer function-noise models, there now exists a good deal of practical experience with this approach, and its utility over wide areas of application is now well established. Experience with the multiple series model building techniques of the previous section is also now accumulating. Certainly, it is safe to say that, if moderately long series of observations are available, and only a modest number of series are to be analysed simultaneously, this too provides a valuable addition to the forecaster's armoury.

The time series analyst's approach to model building contemplates a broad class of general models and, as we have already indicated, data are employed to select, estimate the parameters of, and check the adequacy of, an appropriate model from the general class. The procedures employed require the user to exert some judgement, and on occasion this can prove quite challenging. In exercising such judgement, it is important that the principle of parsimony be kept in mind. The objective is to find as simple a model as one can get away with.

Not surprisingly, the more inter-related series we have to simultaneously consider, the less likely will we be to achieve a good model. It is, for example, not possible to take, say, ten time series and hope to achieve anything of value using exclusively the methods of the previous section. The vector ARIMA models will simply contain too many parameters to be parsimonious forecasting models. Rather, it has to be admitted that such a venture is best not undertaken unless we have available subject matter theory which is definite enough to greatly restrict the options to be considered.

Unfortunately, it is often the case that, where such theory apparently exists, it is, on closer inspection, not terribly credible. This dilemma is nicely discussed by Sims (1980). Perhaps the best dictum, in the absence of good theory, is to stick to simple models.

We have seen that, in the last few years, the art of building time series models, on the base outlined by Box and Jenkins (1970), has progressed on many fronts. The topics discussed in this paper continue to provide exciting opportunities for further research. Given what is already available, the technology for building ARIMA models can be extremely valuable in attacking a very wide range of forecasting problems. No doubt, continuing research will further enhance this value.

REFERENCES

Abraham, B. and Box, G. E. P., 'Bayesian analysis of some outlier problems in time series', *Biometrika*, **66** (1979), 229–236.

Akaile, H., 'Fitting autoregressive models for prediction', *Annals of Institute of Statistical Mathematics*, **21** (1969), 243–247.

Akaike, H., 'A new look at the statistical model identification', *IEEE Transactions on Automatic Control*, **AC-19** (1974), 716–723.

Ansley, C. F., 'An algorithm for the exact likelihood of a mixed autoregressive-moving average process', *Biometrika*, **66** (1979), 59–65.

Ansley, C. F. and Newbold, P., 'On the finite sample distribution of residual autocorrelations in autoregressive-moving average models', *Biometrika*, **66** (1979a), 547–553.

Ansley, C. F. and Newbold, P., 'Multivariate partial autocorrelations', *Proceedings of Business and Economic Statistics Section, American Statistical Association* (1979b), 349–353.

Ansley, C. F. and Newbold, P., 'Finite sample properties of estimators for autoregressive-moving average models', *Journal of Econometrics*, **13** (1980), 159–183.

Ansley, C. F. and Newbold, P., 'On the bias in estimates of forecast mean squared error', *Journal of American Statistical Association*, **76** (1981), 569–578.

Beguin, J. M., Gourieroux, C. and Monfort, A., 'Identification of a mixed autoregressive-moving average process: the corner method', in O. D. Anderson (ed.), *Time Series*, Amsterdam: North Holland, 1980.

Box, G. E. P. and Cox, D. R., 'An analysis of transformations', *Journal of Royal Statistical Society B*, **26** (1964), 211–243.

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

Box, G. E. P. and Jenkins, G. M., 'Some comments on a paper by Chatfield and Prothero and on a review by Kendall', *Journal of Royal Statistical Society A*, **136** (1973), 337–345.

Box, G. E. P. and Pierce, D. A., 'Distribution of residual autocorrelations in autoregressive integrated moving average time series models', *Journal of American Statistical Association*, **65** (1970), 1509–1526.

Box, G. E. P. and Tiao, G. C., 'Intervention analysis with applications to economic and environmental problems', *Journal of American Statistical Association*, **70** (1975), 70–79.

Cleveland, W. S., 'The inverse autocorrelations of a time series and their applications', *Technometrics*, **14** (1972), 277–298.

Davies, N. and Newbold, P., 'Some power studies of a portmanteau test of time series model specification', *Biometrika*, **66** (1979), 153–155.

Davies, N. and Newbold, P., 'Forecasting with misspecified models', *Applied Statistics*, **29** (1980), 87–92.

Davies, N., Triggs, C. M. and Newbold, P., 'Significance levels of the Box–Pierce portmanteau statistic in finite samples', *Biometrika*, **64** (1977), 517–522.

Godfrey, L. G., 'Testing the adequacy of a time series model', *Biometrika*, **66** (1979), 67–72.

Granger, C. W. J., 'Investigating causal relations by econometric models and cross-spectral methods', *Econometrica*, **37** (1969), 424–438.

Granger, C. W. J. and Newbold, P., *Forecasting Economic Time Series*, New York: Academic Press, 1977.

Gray, H. L., Kelley, G. D. and McIntire, D. D., 'A new approach to ARMA modelling', *Communications in Statistics B*, **7** (1978), 1–77.

Hamilton, D. C. and Watts, D. G., 'Interpreting partial autocorrelation functions of seasonal time series models', *Biometrika*, **65** (1978), 135–140.

Hannan, E. J. and Quinn, B. G., 'The determination of the order of an autoregression', *Journal of Royal Statistical Society B*, **41** (1979), 190–195.

Hannan, E. J. and Rissanen, J., 'Recursive estimation of mixed autoregressive-moving average order', *Biometrika*, **69** (1982), 81–94.

Harrison, P. J. and Stevens, C. F., 'Bayesian forecasting', *Journal of Royal Statistical Society B*, **38** (1976), 205–247.

Hendry, D. F. and Mizon, G. E., 'Serial correlation as a convenient simplification, not a nuisance: a comment on a study of the demand for money by the Bank of England', *Economic Journal*, **88** (1978), 549–563.

Hillmer, S. C. and Tiao, G. C., 'Likelihood function of stationary multiple autoregressive-moving average models', *Journal of American Statistical Association*, **74** (1979), 652–660.

Hopwood, W. S., McKeown, J. C. and Newbold, P., 'Power transformations in time series models of quarterly earnings per share', *Accounting Review*, **56** (1981), 927–933.

Hosking, J. R. M., 'Lagrange multiplier test of time series models', *Journal of Royal Statistical Society B*, **42** (1980), 170–181.

Hosking, J. R. M., 'Lagrange multiplier tests of multivariate time series models', *Journal of Royal Statistical Society B*, **43** (1981), 219–230.

Jenkins, G. M., *Practical Experiences with Modelling and Forecasting Time Series*, Lancaster: GJP Publications, 1979.

Jenkins, G. M., 'Some practical aspects of forecasting in organizations', *Journal of Forecasting*, **1** (1982), 3–21.

Jenkins, G. M. and Alavi, A. S., 'Some aspects of modelling and forecasting multivariate time series', *Journal of Time Series Analysis*, **2** (1981), 1–47.

Kenny, P. B. and Durbin, J., 'Local trend estimation and seasonal adjustment of economic and social time series', *Journal of Royal Statistical Society A*, **145** (1982), 1–41.

Ledolter, J. and Abraham, B., 'Parsimony and its importance in time series forecasting', *Technometrics*, **23** (1981), 411–414.

Ledolter, J. and Box, G. E. P., 'Conditions for the optimality of exponential smoothing forecast procedures', *Metrika*, **25** (1978), 77–93.

Li, W. K. and McLeod, A. I., 'Distribution of the residual autocorrelations in multivariate ARMA time series models', *Journal of Royal Statistical Society B*, **43** (1981), 231–239.

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

Ljung, G. M. and Box, G. E. P., 'The likelihood function of stationary autoregressive-moving average models', *Biometrika*, **66** (1979), 265–270.

Makridakis, S., Andersen, A., Carbone, R., Fildes, R., Hibon, M., Lewandowski, R., Newton, J., Parzen, E. and Winkler, R., 'The accuracy of extrapolation (time series) methods: results of a forecasting competition', *Journal of Forecasting*, **1** (1982), 111–153.

Makridakis, S. and Hibon, M., 'Accuracy of forecasting: an empirical investigation', *Journal of Royal Stastical Society A*, **142** (1979), 97–145.

McLeod, A. I., 'On the distribution and applications of residual autocorrelations in Box–Jenkins models', *Journal of Royal Statistical Society B*, **40** (1978), 296–302.

Nelson, C. R., *Applied Time Series Analysis for Managerial Forecasting*, San Francisco: Holden Day, 1973.

Nelson, H. L. and Granger, C. W. J., 'Experience with using the Box–Cox transformation when forecasting economic time series', *Journal of Econometrics*, **10** (1979), 57–69.

Newbold, P., 'The exact likelihood function for a mixed autoregressive-moving average process', *Biometrika*, **61** (1974), 423–426.

Newbold, P., 'The equivalence of two tests of time series model adequacy', *Biometrika*, **67** (1980), 463–465.

Newbold, P., 'Model checking in time series analysis', in A. Zellner (ed.), *Proceedings of A.S.A.-CENSUS-NBER Conference on Applied Time Series Analysis of Economic Data*, Washington, D.C., 1982.

Newbold, P. and Granger, C. W. J., 'Experience with forecasting univariate time series and the combination of forecasts', *Journal of Royal Statistical Society A*, **137** (1974), 131–165.

Newbold, P. and Reed, G. V., 'The implications for economic forecasting of time series model building methods', in O. D. Anderson (ed.), *Forecasting*, Amsterdam: North Holland, 1979.

Nicholls, D. F. and Hall, A. D., 'The exact likelihood function of multivariate autoregressive-moving average models', *Biometrika*, **66** (1979), 259–264.

Ozaki, T., 'On the order determination of ARIMA models', *Applied Statistics*, **26** (1977), 290–301.
Parzen, E., 'Some recent advances in time series modeling', *IEEE Transactions on Automatic Control*, **AC-19** (1974), 723–730.
Parzen, E., 'ARARMA models for time series analysis and forecasting', *Journal of Forecasting*, **1** (1982), 67–82.
Poskitt, D. S. and Tremayne, A. R., 'Testing the specification of a fitted ARMA model', *Biometrika*, **67** (1980), 359–363.
Poskitt, D. S. and Tremayne, A. R., 'An approach to testing linear time series models', *Annals of Statistics*, **9** (1981), 974–986.
Poskitt, D. S. and Tremayne, A. R., 'Diagnostic tests for multiple time series models', *Annals of Statistics*, **10** (1982), 114–120.
Priestley, M. B., 'State-dependent models: a general approach to non-linear time series analysis', *Journal of Time Series Analysis*, **1** (1980), 47–71.
Sims, C. A., 'Macroeconomics and reality', *Econometrica*, **48** (1980), 1–48.
Tiao, G. C. and Box, G. E. P., 'Modeling multiple time series with applications, *Journal of American Statistical Association*, **76** (1981), 802–816.
Woodward, W. A. and Gray, H. L., 'On the relationship between the S array and the Box–Jenkins method of ARMA model identification', *Journal of American Statistical Association*, **76** (1981), 579–587.

Author's biography:
Paul Newbold is Professor of Economics at the University of Illinois, Urbana-Champaign. His doctorate is in Statistics, from the University of Wisconsin. Previously he was at the University of Nottingham, and has held visiting appointments at the Universities of California (San Diego), Chicago, and Wisconsin. He is co-author, with C. W. J. Granger, of *Forecasting Economic Time Series* (Academic Press, 1977).

Author's address:
Paul Newbold, Department of Economics, University of Illinois, 1206 South Sixth Street, Champaign, Illinois 61820, U.S.A.