Time Series Modelling and Interpretation

Author(s): C. W. J. Granger and M. J. Morris

Source: *Journal of the Royal Statistical Society. Series A (General)*, Vol. 139, No. 2 (1976), pp. 246-257

Published by: Oxford University Press for the Royal Statistical Society

Stable URL: https://www.jstor.org/stable/2345178

Accessed: 28-02-2025 00:33 UTC

J. R. Statist. Soc. A
(1976), **139**, *Part* 2, p. 246

Time Series Modelling and Interpretation

By C. W. J. Granger and M. J. Morris

University of California, San Diego *University of East Anglia, U.K.*

Summary

By considering the model generating the sum of two or more series, it is shown that the mixed *ARMA* model is the one most likely to occur. As most economic series are both aggregates and are measured with error it follows that such mixed models will often be found in practice. If such a model is found, the possibility of resolving the series into simple components is considered both theoretically and for simulated data.

Keywords: TIME SERIES; MIXED AUTOREGRESSIVE-MOVING AVERAGE MODEL; AGGREGATION; MEASUREMENT ERROR; PARSIMONIOUS MODELS

1. Introduction

The recent publication of the book by Box and Jenkins has greatly increased interest by time series analysts and econometricians in more complicated, even if more "parsimonious", time series models. However, the intuitive interpretation of the models that arise is not always simple and an applied research worker may well ask how such a model could have arisen in practice. In this paper, the possibility that a complicated looking model could have arisen from a simpler situation is explored and some suggestions made about the possible interpretation of some models that appear to arise in practice.

The only series to be considered are those with zero means and generated by models with time-invariant parameters. Thus, possibly after an appropriate linear transformation, the series will be taken to be second-order stationary. Let X_t be such a series, so that

$$E\{X_t\} = 0, \quad \text{all } t.$$

Initially, suppose that X_t is second-order stationary, so that the kth autocovariance

$$\mu_k = E\{X_t X_{t-k}\} \tag{1.1}$$

is time invariant. The series will be said to be a white noise process if

$$\mu_k = 0, \quad \text{all } k \neq 0.$$

Use will be made of the backward operator B, defined so that

$$B^j X_t = X_{t-j}. \tag{1.2}$$

A series with the representation

$$a(B)\, X_t = \varepsilon_t, \tag{1.3}$$

where

$$a(B) = \sum_{j=0}^{p} a_j B^j, \quad a \neq 0, \quad a_0 = 1, \tag{1.4}$$

and ε_t is white noise, will be said to be generated by an autoregressive model of order p, and denoted by $X_t \sim AR(p)$.

Suppose further that

$$a(B) = \prod_{j=1}^{p} (1 - \theta_j B) \tag{1.5}$$

so that θ_j^{-1}, $j = 1, \ldots, p$, are the roots of the equation $a(z) = 0$. If $|\theta_j^{-1}| > 1$, all j, then X_t will be stationary. In this case

$$\mu_k = \sum_{j=1}^{p} \lambda_j \theta_j^k. \tag{1.6}$$

A series will be said to be generated by a moving average of order q if it can be represented by

$$X_t = b(B)\,\varepsilon_t,$$

where

$$b(B) = \sum_{j=0}^{q} b_j B^j \tag{1.8}$$

and ε_t is white noise. This is denoted by $X_t \sim MA(q)$. A necessary but not sufficient condition that such a representation is possible is

$$\mu_k = 0, \quad k > q. \tag{1.9}$$

The fact that this condition is not sufficient may be seen from the result that the maximum value of the first serial correlation coefficient $\rho_1 = \mu_1/\mu_0$ achievable from an $MA(q)$ process is

$$\rho_1(\max) = \cos\left(\frac{\pi}{q+2}\right). \tag{1.10}$$

It follows, for example, that if $\rho_1 = 0{\cdot}8$, $\mu_k = \rho_k = 0$, all $k > 1$, then there is no corresponding $MA(1)$ process which has these values. Davies *et al.* (1974) have found more general inequalities on autocorrelations for MA processes.

A necessary and sufficient condition that there exists an $MA(q)$ model corresponding to a specific set of autocorrelation coefficients u_j, $j = 0, 1, \ldots, q$, has been given by Wold (1953, p. 154).

Let

$$u(x) = 1 + \sum_{j=1}^{q} u_j(x^j + x^{-j}).$$

Applying the transformation $z = x + x^{-1}$, let $u(x)$ be transformed into the qth order polynomial $v(z)$. The required necessary and sufficient condition is then that $v(z)$ has no zero of odd multiplicity in the interval $-2 < z_0 < 2$. Note that

$$v(2) = u(1) = f(0) \geqslant 0 \quad \text{and} \quad v(-2) = u(-1) = f(\pi) \geqslant 0,$$

where $f(\omega)$ is the spectrum of the moving average process. It follows that if $v(z)$ obeys the above condition, then $v(z) \geqslant 0$ for all z in the interval $(-2, 2)$.

An $MA(q)$ process, with $q < \infty$, is always stationary.

A stationary $AR(p)$ process can always be represented by an $MA(\infty)$ process (as can any stationary, purely non-deterministic process). If the roots of the equation $b(z) = 0$ all have absolute value greater than one, then the corresponding $MA(q)$ process may be inverted and be equally represented by an $AR(\infty)$ process.

Ever since their introduction by Yule (1921), the autoregressive and moving average models have been greatly favoured by time series analysts. It is not difficult to convince oneself that both models are intuitively reasonable and could well occur in practice. Simple expectations models or a momentum effect in a random variable can lead to AR models. Similarly, a variable in equilibrium but buffeted by a sequence of unpredictable events with a delayed or discounted effect will give MA models, for example.

A fairly obvious generalization of these models is the mixed autoregressive-moving average process generated by

$$a(B)\,X_t = b(B)\,\varepsilon_t, \tag{1.11}$$

where ε_t is again white noise and $a(B)$ and $b(B)$ are polynomials in B of order p and q respectively. Such a process is denoted by $X_t \sim ARMA(p,q)$ and has an autocovariance sequence given by (1.6) for $k>q$. An $ARMA(p,q)$ process can always be represented by an $MA(\infty)$ model and, with the same invertibility condition for $MA(q)$, it can also be represented in an $AR(\infty)$ form.

In practice, the *ARMA* model is more difficult to fit to data than for example an *AR* model, but in other ways the *ARMA* model may be the statistically more efficient of the two. Box and Jenkins suggest that if one fits both an $AR(p)$ model and an $ARMA(p',q')$ model, then one will generally need fewer coefficients in this second model to get a satisfactory fit, i.e. $p'+q'<p$. As there are good reasons to prefer a model with as few parameters as possible, their principle of parsimony then suggests that the *ARMA* will often be preferable.

Although the *ARMA* model may involve the estimation of fewer coefficients than alternative models, it is certainly more difficult to comprehend, interpret or explain how it might occur in the real world. Chatfield and Prothero (1973) comment on the difficulty in the interpretation of many Box–Jenkins models. In the following sections a number of ways are suggested in which such an *ARMA* model can arise from simpler models. It will be concluded that a mixed model is the one that is most likely to arise in practice.

2. The Sum of Two Independent Series

If X_t and Y_t are two independent, zero-mean, stationary series, then in this section statements of the following kind will be considered:

$$\begin{aligned} \text{if}\quad & X_t \sim ARMA(p,m), \quad Y_t \sim ARMA(q,n) \\ \text{and}\quad & Z_t = X_t + Y_t \\ \text{then}\quad & Z_t \sim ARMA(x,y). \end{aligned}$$

Such a statement will be *denoted* by

$$ARMA(p,m)+ARMA(q,n) = ARMA(x,y).$$

Lemma.

$$MA(m)+MA(n) = MA(y), \tag{2.1}$$

where

$$y \leqslant \max(m,n). \tag{2.2}$$

Necessity follows directly from (1.9) and the inequality from the fact that there may be an inter-reaction between the two *MA* processes which reduces the order of the *MA* process of the sum. This latter point will be discussed in greater detail below.

To prove that the sufficiency condition holds, let $v_1(z), v_2(z)$ be the $v(z)$ functions for $MA(m)$ and $MA(n)$ respectively, as defined after equation (1.10). As $v_1(z)$ and $v_2(z)$ do correspond to moving averages, they will both be non-negative over the interval $-2\leqslant z\leqslant 2$. If $v_s(z)$ is the function for $MA(m)+MA(n)$, it follows that as $v_s(z) = v_1(z)+v_2(z)$ it will also be non-negative in the interval $-2\leqslant z\leqslant 2$ and so $v_s(z)$ cannot have a root of odd multiplicity in this interval, so the truth of the lemma holds. Note that if $m\neq n$ then there must be equality in (2.2).

Basic Theorem.

$$ARMA(p,m)+ARMA(q,n) = ARMA(x,y), \tag{2.3}$$

where

$$x \leqslant p+q \quad \text{and} \quad y \leqslant \max(p+n, q+m). \tag{2.4}$$

(This is a form of the theorem proved by Box and Jenkins in their Appendix to Chapter 4.)

Proof. Let

$$a(B)\, X_t = c(B)\, \varepsilon_t \tag{2.5}$$

and

$$b(B)\, Y_t = d(B)\, \eta_t, \tag{2.6}$$

where a, b, c, d are polynomials in B of order p, q, m, n respectively and ε_t, η_t are two independent, white noise series.

As $Z_t = X_t + Y_t$, it follows that

$$a(B)\, b(B)\, Z_t = b(B)\, a(B)\, X_t + a(B)\, b(B)\, Y_t = b(B)\, c(B)\, \varepsilon_t + a(B)\, d(B)\, \eta_t. \tag{2.7}$$

The first term on the right-hand side is $MA(q+m)$ and the second term is $MA(p+n)$ and so, from the lemma, the right-hand side is $MA(y)$ where

$$y \leqslant \max(p+n, q+m).$$

The order of the polynomial $a(B)\, b(B)$ is not more than $p+q$, and so the theorem is established.

(The assumption of independence may be weakened to allow for contemporaneous correlation between the noise series. The sum would still be of the form (2.3) but the parameters of the representation would be different.)

The need for the inequalities in the expressions for x and y partly arise from the fact that the polynomials $a(B)$ and $b(B)$ may contain common roots and so part of the operator need not be applied twice. For example, if

$$(1-\alpha_1 B)\, X_t = \varepsilon_t, \quad \text{i.e. } X_t \sim AR(1)$$

and

$$(1-\alpha_1 B)(1-\alpha_2 B)\, Y_t = \eta_t, \quad \text{i.e. } Y_t \sim AR(2)$$

then with $Z_t = X_t + Y_t$ we have $(1-\alpha_1 B)(1-\alpha_2 B)\, Z_t = (1-\alpha_2 B)\, \varepsilon_t + \eta_t$, i.e. $Z_t \sim ARMA\ (2, 1)$. In general, if the polynomials $a(B), b(B)$ have just k roots in common, the inequalities (2.4) become

$$x = p+q-k \quad y \leqslant \max(p+n-k, q+m-k). \tag{2.8}$$

That the inequality for y in (2.8) is still necessary is seen from the following example. Suppose

$$(1-\alpha B)\, X_t = \varepsilon_t, \quad \text{i.e. } X_t \sim AR(1)$$

$$(1+\alpha B)\, Y_t = \eta_t, \quad \text{i.e. } Y_t \sim AR(1)$$

and also $\operatorname{var}(\varepsilon) = \operatorname{var}(\eta) = \sigma^2$.

If $Z_t = X_t + Y_t$, then $(1-\alpha B)(1+\alpha B)\, Z_t = (1+\alpha B)\, \varepsilon_t + (1-\alpha B)\, \eta_t$. Denote the right-hand side of the last equation by $\zeta_t = \varepsilon_t + \alpha\varepsilon_{t-1} + \eta_t - \alpha\eta_{t-1}$; then $\operatorname{var}(\zeta_t) = 2(1+\alpha^2)\, \sigma^2$ and $E\{\zeta_t\, \zeta_{t-k}\} = 0$, all $k > 0$, so ζ_t is a white noise process and $Z_t \sim AR(2)$ rather than $ARMA(2, 1)$ which would generally occur when two independent $AR(1)$ processes are added together. Those situations in which a simpler model arises than might generally be expected will be called "coincidental situations".

If X_t and Y_t are both stationary, then any case in which x and y in (2.4) take values less than the maximum ones might well be considered coincidental. However, a more general model proposed by Box and Jenkins for marginally non-stationary processes suggests that common roots might well be expected, this common root being unity. They suggest that many series met with in practice obey a simple kind of non-stationary model:

$$(1-B)^f a(B) X_t = b(B) \varepsilon_t,$$

where $a(B)$ is a polynomial of order p and with all roots of $a(z)=0$ lying outside the unit circle and $b(B)$ is of order m. In this case X_t is said to obey an integrated autoregressive-moving average model, denoted by $ARIMA(p,f,m)$. Newbold and Granger (1974) found such models fitted a large number of economic time series, although possibly with the addition of seasonal terms. If two independent series X_t and Y_t both obey such a model, i.e. $X_t \sim ARIMA(p, f_1, m)$, $Y_t \sim ARIMA(q, f_2, n)$, then their sum will be $ARIMA(x, f', y)$ where $x \leqslant p+q; f' = \max(f_1, f_2); y \leqslant \max(p+n+f_1-f_2, q+m)$, if $f_1 \geqslant f_2$; $y \leqslant \max(p+n, q+m+f_2-f_1)$, if $f_2 \geqslant f_1$.

Similar considerations apply to the method of taking seasonal effects into account suggested by Box and Jenkins.

Given the basic theorem there is, of course, no difficulty in generalizing it to cover the sum of any number of independent series. Using an obvious generalization of the previous notation, the theorem becomes

$$\sum_{j=1}^{N} ARMA(p_j, m_j) = ARMA(x, y),$$

where

$$x \leqslant \sum_{j=1}^{N} p_j \quad \text{and} \quad y \leqslant \max(x - p_j + m_j, j = 1, \ldots, N).$$

3. Series Aggregation and Observational Error Models

Two situations where series are added together are of particular interpretational importance. The first is where series are aggregated to form some total, and the second is where the observed series is the sum of the true process plus observational error, corresponding to the classical "signal plus noise" situation. Most macroeconomic series, such as G.N.P., unemployment or exports, are aggregates and there is no particular reason to suppose that all of the component series will obey exactly the same model. Virtually any macroeconomic series, other than certain prices or interest rates, contain important observation errors. It would be highly coincidental if the "true" series and the observational error series obeyed models having common roots, apart possibly from the root unity, or that the parameters of these models should be such that the cancelling out of terms produces a value of y in (2.4) less than the maximum possible.

It is interesting to consider a number of special cases of the basic theorem, concentrating on those situations in which coincidental reductions of parameters do not occur. In the cases considered the interpretations given will assume that aggregates are of independent components, an assumption almost certainly not true in practice, and that the observational error is white noise and independent of the true process, an assumption of debatable reality but on which little empirical work is available from which to form an opinion.

(i) $AR(p) + \textit{white noise} = ARMA(p, p)$

This corresponds to an $AR(p)$ signal as true process plus a simple white noise observational error series.

(ii) $AR(p) + AR(q) = ARMA(p+q, \max(p, q))$ *and in particular* $AR(1) + AR(1) = ARMA(2, 1)$

This situation might correspond to a series which is aggregated from two independent AR series. A further case of possible interest is: the sum of k $AR(1)$ series is $ARMA(k, k-1)$. Case (ii) with an $AR(p)$ signal and $AR(q)$ noise was the subject of a paper by Bailey (1965).

(iii) *$MA(p)+MA(q) = MA(\max(p,q))$ and in particular $MA(p)+$ white noise $= MA(p)$.*

Thus if a true process follows an MA model then the addition of a white noise observation error will not alter the class of model, although the parameter values will change.

(iv) *$ARMA(p,m)+$ white noise $= ARMA(p,p)$ if $p>m$ $= ARMA(p,m)$ if $p<m$.*

Thus the addition of an observational error may alter the order of an $ARMA$ model but need not do so.

(v) *$AR(p)+MA(n) = ARMA(p, p+n)$.*

Again, this is possibly relevant to the aggregation case, or to the observation error case with noise not being white noise.

If one accepts the belief that time series, particularly in economics, obey either AR or MA models, both of which can be explained fairly easily in terms of the working of the economy, then the above special cases have a number of possible implications.

Cases (ii) and (v) suggest that a series that is an aggregate of several series, some of which are AR series, will be very likely to be an $ARMA$ process, although from case (iii) an MA process is possible. The independence of the components can be relaxed in a realistic fashion with the same conclusion being reached. If each component consists of two parts, the first a factor common to all components representing the influence of the whole economy on them and the second a factor relating just to the particular component and independent of the economy-wide factor, then simple models can be assumed for each factor and on aggregation the more general form of the basic model involving the sum of several series will apply. Such factor models have been successfully used with series of stock market prices and firms' earnings.

The observational error or signal plus noise situation is also likely to produce an $ARMA$ model, from cases (i) and (iv), although case (iii) again suggests that an MA model is possible.

Thus both of the realistic situations considered in this section are quite likely to produce $ARMA$ models. These situations are discussed further in Sections 5 and 6. Other situations that also produce $ARMA$ models are considered in the following section.

4. TIME AGGREGATION, NON-INTEGER LAGS AND FEEDBACK MODELS

Three other situations that might occur in practice lead either exactly, or approximately, to $ARMA$ models. These are:

(i) A variable that obeys a simple model such as $AR(1)$ if it were recorded at an interval of K units of time but which is actually observed at an interval of M units. Details may be found in Amemiya and Wu (1972). The variable considered has to be an accumulated one, that is of the type called a stock variable by economists.

(ii) If X_t is an instantaneously recorded variable, called a flow variable by economists, and suppose that it obeys the model

$$X_t - \alpha X_{t-b} = \varepsilon_t, \quad |\alpha| < 1,$$

where b may be a non-integer multiple or fraction of the observation period, then it is easily shown (see Granger, 1964) that this is equivalent to the $AR(\infty)$ model

$$\sum_{j=0}^{\infty} h_j X_{t-j} = \varepsilon_t,$$

where $h_j = \{\sin(j-b)\pi\}/(j-b)\pi$ and doubtless this model could well be approximated by an $ARMA$ model.

(iii) If X_t, Y_t are generated by the bivariate autoregressive scheme with feedback:

$$a(B)X_t + b(B)Y_t = \varepsilon_t, \quad c(B)X_t + d(B)Y_t = \eta_t,$$

where ε_t, η_t are uncorrelated white noise series and $b(0) = c(0) = 0$, then the model obeyed by X_t alone is found by eliminating Y_t in the equations to be

$$[a(B)\,d(B) - c(B)\,b(B)]\,X_t = d(B)\,\varepsilon_t + b(B)\,\eta_t$$

and so the $ARMA(p,q)$ model occurs once more, and it is easily shown that generally $p > q$.

The results of this and the previous section suggest that a number of real data situations are all likely to give rise to $ARMA$ models. Combining these situations will only reinforce this conclusion. In fact, one might well conclude that the model most likely to be found in practice is the mixed autoregressive moving-average model.

5. Realizability of Simple Models

Given a specific $ARMA(p,q)$ model it is interesting to ask whether or not this could have arisen from some simpler model. On occasions an answer can be found immediately: if $p < q$ then a feedback model is not appropriate for example. To merely illustrate the problem of realizability, a few examples from Section 3 are considered. As will be seen simplifications are not always possible, as conditions on the coefficients of the $ARMA$ model need to be satisfied for a simpler model to be realizable.

(i) *Can* $ARMA(1,1) = AR(1) +$ *white noise?*

Suppose $(1 + aB)\,X_t = \varepsilon_t$ and $Y_t = \eta_t$, where both ε_t and η_t are white noise series, independent of each other, then if $Z_t = X_t + Y_t$ it was seen that Z_t is $ARMA(1,1)$ given by

$$(1 + aB)\,Z_t = (1 + aB)\,\eta_t + \varepsilon_t. \tag{5.1}$$

Now suppose that the given $ARMA(1,1)$ process is

$$(1 + cB)\,Z_t = (1 + dB)\,\zeta_t, \tag{5.2}$$

where ζ_t is a white noise series.

Let μ_0, μ_1 denote the variance and first autocovariance of the $MA(1)$ process on the right-hand side of (5.2). For (5.1) and (5.2) to be equivalent one needs

$$c = a, \tag{5.3(i)}$$

$$\mu_0 = (1 + d^2)\,\sigma_\zeta^2 = (1 + a^2)\,\sigma_\eta^2 + \sigma_\varepsilon^2, \tag{5.3(ii)}$$

$$\mu_1 = d\sigma_\zeta^2 = a\sigma_\eta^2. \tag{5.3(iii)}$$

The realizability conditions may then be written in the form

$$\frac{1}{1 + c^2} > \frac{\rho_1}{c} \geqslant 0, \tag{5.3(iv)}$$

where

$$\rho_1 = \frac{\mu_1}{\mu_0} = \frac{d}{1 + d^2},$$

the first inequality coming from (5.3(ii)) and the condition $\sigma_\varepsilon^2 > 0$, the second inequality from (5.3(iii)) and (5.3(i)).

(ii) *Can* $ARMA(2,2) = AR(2) +$ *white noise?*

Suppose $(1 + a_1 B + a_2 B^2)\,X_t = \varepsilon_t$ and $Y_t = \eta_t$, where both ε_t and η_t are white noise series, independent of each other; then if $Z_t = X_t + Y_t$, Z_t is ARMA(2, 2) given by

$$(1 + a_1 B + a_2 B^2)\,Z_t = (1 + a_1 B + a_2 B^2)\,\eta_t + \varepsilon_t. \tag{5.4}$$

Now suppose that the given $ARMA(2,2)$ process is

$$(1+c_1 B+c_2 B^2) Z_t = (1+d_1 B+d_2 B^2)\, \zeta_t, \tag{5.5}$$

where ζ_t is a white noise series.

Let μ_0, μ_1, μ_2 denote the variance, first and second autocovariance of the $MA(2)$ process on the right-hand side of (5.5). For (5.4) and (5.5) to be equivalent one needs

$$c_1 = a_1, \tag{5.6(i)}$$

$$c_2 = a_2, \tag{5.6(ii)}$$

$$\mu_0 = (1+d_1^2+d_2^2)\,\sigma_\zeta^2 = (1+a_1^2+a_2^2)\,\sigma_\eta^2+\sigma_\varepsilon^2, \tag{5.6(iii)}$$

$$\mu_1 = d_1(1+d_2)\,\sigma_\zeta^2 = a_1(1+a_2)\,\sigma_\eta^2, \tag{5.6(iv)}$$

$$\mu_2 = d_2\,\sigma_\zeta^2 = a_2\,\sigma_\eta^2. \tag{5.6(v)}$$

These realizability conditions may be written in the form

$$\frac{1}{1+c_1^2+c_2^2} > \frac{\rho_2}{c_2} \geqslant 0 \tag{5.6(vi)}$$

and

$$\frac{\rho_1}{c_1(1+c_2)} = \frac{\rho_2}{c_2}, \tag{5.6(vii)}$$

where

$$\rho_1 = \frac{\mu_1}{\mu_0} = \frac{d_1(1+d_2)}{1+d_1^2+d_2^2} \quad \text{and} \quad \rho_2 = \frac{\mu_2}{\mu_0} = \frac{d_2}{1+d_1^2+d_2^2}.$$

As an equality is involved, the realizability of the simple model may be classified as being coincidental.

(iii) *Can $ARMA(2,2) = ARMA(2,1)+$white noise?*

With the same notation as in (ii), the realizability conditions are

$$\frac{1}{1+c_1^2+c_2^2} > \frac{\rho_2}{c_2} \geqslant 0. \tag{5.7}$$

These examples illustrate the fact that some models are not capable of simplification and that the realizability conditions will usually be rather complicated. The conditions will also be coincidental if the number of parameters, or "degrees of freedom", are less than the parameters in the initial *ARMA* model.

The problem considered here is only of theoretical interest and is not realistic. An *ARMA* model will generally have been estimated from some data. A more realistic question, and one of considerable interpretational importance, is whether a simple model would have a level of fit that is not significantly different from that achieved by the *ARMA* model. This question is briefly considered in the next section, using simulated time series.

6. SIMULATION OF OBSERVATION ERROR MODELS

Two examples are considered:

Example 1: Can a given $ARMA(1,1)$ be expressed as $AR(1)+$white noise?
Example 2: Can a given $ARMA(2,2)$ be expressed as $AR(2)+$white noise?

Example 1

Two series, each of 200 terms, are generated from $AR(1)$ and white noise processes.

Series 1: $(1-0{\cdot}8B)\,X_t = \varepsilon_t, \quad t = 1,\ldots,N \quad (N = 200).$
Series 2: $Y_t = \eta_t, \quad t = 1,\ldots,N \quad (N = 200)$
with $\sigma_\varepsilon^2 = 4$, $\sigma_\eta^2 = 1$.

The two series added together gives the series $Z_t = X_t + Y_t$, $t = 1,\ldots,N$, which may be regarded as being generated by the process

$$(1-0{\cdot}8B)\,Z_t = (1-0{\cdot}8B)\,\eta_t + \varepsilon_t$$

and this, from (5.3(i)), (5.3(ii)) and (5.3(iii)) may be written as

$$(1-0{\cdot}8B)\,Z_t = (1-0{\cdot}14B)\,\zeta_t \tag{6.1}$$

with $\sigma_\zeta^2 = 5{\cdot}52$.

Using the maximum likelihood estimation method described in Box and Jenkins (1970) the model

$$(1+cB)\,Z_t = (1+dB)\,\zeta_t$$

fitted to the series Z_t, $t = 1,\ldots,N$, gives estimates $c = -0{\cdot}76$, $d = -0{\cdot}04$.

The residual sum of squares is 1,273·0, giving an estimate of $\sigma_\zeta^2 = 6{\cdot}43$.

Now we take the given $ARMA(1,1)$ series to be

$$(1+cB)\,Z_t = (1+dB)\,\zeta_t$$

with

$$c = -0{\cdot}76, \quad d = -0{\cdot}04, \quad \sigma_\zeta^2 = 6{\cdot}43. \tag{6.2}$$

Can this be expressed as $AR(1)$+white noise? Realizability conditions (5.3(iv)) are satisfied since

$$\frac{1}{1+c^2} = 0{\cdot}63$$

and

$$\frac{\rho_1}{c} = \frac{d}{c(1+d^2)} = 0{\cdot}05.$$

The realizability conditions (5.3(i)), (5.3(ii)) and (5.3(iii)) may be solved to give

$$a = -0{\cdot}76, \quad \sigma_\eta^2 = 0{\cdot}34, \quad \sigma_\varepsilon^2 = (1+d^2)\,\sigma_\zeta^2 - (1+a^2)\,\sigma_\eta^2 = 5{\cdot}90.$$

Thus, the series (6.2) is exactly equivalent to the series

$$(1-0{\cdot}76B)\,X_t = \varepsilon_t, \quad Y_t = \eta_t, \quad Z_t = X_t + Y_t \tag{6.3}$$

with $\sigma_\eta^2 = 0{\cdot}34$, $\sigma_\varepsilon^2 = 5{\cdot}90$, both (6.2) and (6.3) fitting the given series equally well.

It is therefore seen that if one were given the datum Z_t, one would not know if it were generated by an $ARMA(1,1)$ model of if it were the sum of two components, one of which is $AR(1)$ and the other is white noise. This is, of course, true of any $ARMA(1,1)$ model for which the realizability condition holds.

Example 2

Two series, each of 200 terms, are generated from $AR(2)$ and white noise processes.

Series 1: $(1-0{\cdot}8B+0{\cdot}4B^2)\,X_t = \varepsilon_t, \quad t = 1,\ldots,N \quad (N = 200).$
Series 2: $Y_t = \eta_t, \quad t = 1,\ldots,N \quad (N = 200)$
with $\sigma_\varepsilon^2 = 4$, $\sigma_\eta^2 = 1$.

The two series added together give the series $Z_t = X_t+Y_t$, $t = 1, \ldots, N$ which may be regarded as being generated by the process,

$$(1-0{\cdot}8B+0{\cdot}4B^2)Z_t = (1-0{\cdot}8B+0{\cdot}4B^2)\,\eta_t+\varepsilon_t$$

and this, from (5.6(i)) to (5.6(v)), may be written as

$$(1-0{\cdot}8B+0{\cdot}4B^2)Z_t = (1-0{\cdot}19B+0{\cdot}07B^2)\,\zeta_t \tag{6.4}$$

with $\sigma_\zeta^2 = 5{\cdot}59$.

The model $(1+c_1B+c_2B^2)Z_t = (1+d_1B+d_2B^2)\,\zeta_t$ fitted to the series Z_t, $t = 1, \ldots, N$, gives estimates $c_1 = -1{\cdot}01$, $c_2 = 0{\cdot}59$, $d_1 = -0{\cdot}51$, $d_2 = 0{\cdot}12$. The residual sum of squares is $988{\cdot}5$ giving an estimate of $\sigma_\zeta^2 = 5{\cdot}04$.

Now we take the given $ARMA(2,2)$ series to be

$$(1+c_1B+c_2B^2)Z_t = (1+d_1B+d_2B^2)\,\zeta_t \tag{6.5}$$

with $c_1 = -1{\cdot}01$, $c_2 = 0{\cdot}59$, $d_1 = -0{\cdot}51$, $d_2 = 0{\cdot}12$, $\sigma_\zeta^2 = 5{\cdot}04$.

Can this be expressed as $AR(2)$+white noise? This example is more interesting than Example 1, because it involves a reduction in the number of parameters. $ARMA(2,2)$ contains five parameters, whereas $AR(2)$ and white noise contain a total of four.

The realizability conditions (5.6(vi)) and (5.6(vii)) are not both satisfied.

Equation (5.6(vi)) is satisfied because $(1+c_1^2+c_2^2)^{-1} = 0{\cdot}42$ and

$$\frac{\rho_2}{c_2} = \frac{d_2}{c_2(1+d_1^2+d_2^2)} = 0{\cdot}16$$

but the equality (5.6(vii)) is not satisfied because

$$\frac{\rho_1}{c_1(1+c_2)} = \frac{d_1(1+d_2)}{c_1(1+c_2)(1+d_1^2+d_2^2)} = 0{\cdot}28.$$

It is therefore not possible to solve the equations (5.6(i))–(5.6(v)) to obtain a model of the form $AR(2)$+white noise which is exactly equivalent to (6.5) and which fits the data equally well. However, because the given $ARMA(2,2)$ series has been constructed as the sum of $AR(2)$ and white noise, one would hope to find a model of the form $AR(2)$+white noise which is almost as good.

We are seeking a model

$$(1+a_1B+a_2B^2)\,X_t = \varepsilon_t, \quad Y_t = \eta_t, \quad Z_t = X_t+Y_t \tag{6.6}$$

which may also be written as

$$(1+a_1B+a_2B^2)Z_t = (1+a_1B+a_2B^2)\,\eta_t+\varepsilon_t = (1+b_1B+b_2B^2)\,\zeta_t, \quad \text{say.}$$

We may proceed as follows:

Let $\gamma_{Zk} = \operatorname{cov}(Z_t, Z_{t-k})$ and $\gamma_{Xk} = \operatorname{cov}(X_t, X_{t-k})$. Because X_t and Y_t are independent, $\gamma_{Z0} = \gamma_{X0}+\sigma_\eta^2$ and $\gamma_{Zk} = \gamma_{Xk}$, $k>0$. In particular,

$$\gamma_{Z0} = \gamma_{X1}+\sigma_\eta^2, \quad \gamma_{Z1} = \gamma_{X1}, \quad \gamma_{Z2} = \gamma_{X2}.$$

The given $ARMA(2,2)$ series may be used to estimate γ_{Zk}, $k = 0,1,2$, by

$$c_{Zk} = \frac{1}{N}\sum_{t=1}^{N-k}(Z_t-\bar{Z})(Z_{t+k}-\bar{Z}).$$

Then, given an estimate of σ_η^2, γ_{Xk}, $k = 0,1,2$, may be estimated by

$$c_{X0} = c_{Z0}-\sigma_\eta^2, \quad c_{X1} = c_{Z1}, \quad c_{X2} = c_{Z2}.$$

If $r_1 = c_{X1}/x_{X0}$, $r_2 = c_{X2}/c_{X0}$, the Yule–Walker equations may be solved to give estimates of a_1, a_2,

$$a_1 = \frac{r_1(1-r_2)}{1-r_1^2}, \quad a_2 = \frac{r_1^2 - r_2}{1-r_1^2}$$

and σ_ε^2 may be estimated by

$$\sigma_\varepsilon^2 = c_{X0}(1 + r_1 a_1 + r_2 a_2),$$

b_1 and b_2 may then be estimated using the realizability conditions (5.6(iii))–(5.6(v)) by solving the equations

$$\frac{b_2}{b_1(1+b_2)} = \frac{a_2}{a_1(1+a_2)}$$

$$\frac{b_2}{1+b_1^2+b_2^2} = \frac{a_2\sigma_\eta^2}{(1+a_1^2+a_2^2)\sigma_\eta^2+\sigma_\varepsilon^2}.$$

These equations may be written as

$$b_1 = \frac{Ab_2}{1+b_2}, \tag{6.7(i)}$$

$$1+(2-C)b_2+(2-2C+A^2)b_2^2+(2-C)b_2^3+b_2^4 = 0, \tag{6.7(ii)}$$

where

$$A = \frac{a_1(1+a_2)}{a_2}, \quad C = \frac{(1+a_1^2+a_2^2)\sigma_\eta^2+\sigma_\varepsilon^2}{a_2\sigma_\eta^2}$$

choosing the solution of (6.7(ii)) which lies between 0 and a_2 (from 5.6(v)).

The model

$$(1+a_1B+a_2B^2)Z_t = (1+b_1B+b_2B^2)\zeta_t$$

is then fitted to the given series, and the residual sum of squares obtained.

By considering a range of values of σ_η^2 it is possible to determine the parameter values which minimize the residual sum of squares.

For the given series, these values are $a_1 = -0{\cdot}91$, $a_2 = 0{\cdot}60$, $b_1 = -0{\cdot}40$, $b_2 = 0{\cdot}20$ with $\sigma_\eta^2 = 1{\cdot}62$, $\sigma_\varepsilon^2 = 2{\cdot}40$. The residual sum of squares is 994·5, giving an estimate of $\sigma_\zeta^2 = 5{\cdot}07$.

So the model (6.6) with $a_1 = -0{\cdot}91$, $a_2 = 0{\cdot}60$, $\sigma_\eta^2 = 1{\cdot}62$, $\sigma_\varepsilon^2 = 2{\cdot}40$ is almost as good as the $ARMA(2,2)$ model (6.5).

It is, of course, no surprise that the particular Z_t series here considered although it would be identified as an $ARMA(2,2)$ process it could equally well be explained as being the sum of an $AR(2)$ plus white noise, since this is how it was generated. However, many series identified to be $ARMA(2,2)$ might also be equally explained by the simple sum, which is more parsimonious in the sense that it involves fewer parameters to be estimated. If a series can be so decomposed into two or more components this could help with interpretation of its generating process. One might, for example, suggest that the true series might be considered as a true $AR(2)$ signal plus a white noise observation error series. If true, this decomposition would suggest the importance of the observation error measured in terms of the contribution of this error towards the variance of the optimum one-step forecasts. If this contribution is large enough it might be worth spending more to reduce observation error and thus decrease prediction error.

The technique used in our second example can easily be adapted to deal with other models and might allow the analyst to suggest the importance of the noise contribution to prediction error variance in a signal plus noise model.

REFERENCES

AMEMIYA, T. and WU, R. Y. (1972). The effect of aggregation on prediction in the autoregressive model. *J. Amer. Statist. Ass.*, **67**, 628–632.

BAILEY, M. G. (1965). Prediction of an autoregressive variable subject both to disturbances and to errors of observation. *J. Amer. Statist. Ass.*, **60**, 164–181.

BOX, G. E. P. and JENKINS, G. M. (1970). *Time Series Analysis*. San Francisco, Holden-Day.

CHATFIELD, C. and PROTHERO, D. L. (1973). Box–Jenkins seasonal forecasting: problems in a case study (with Discussion). *J. R. Statist. Soc.* A, **136**, 295–336.

DAVIES, N., PATE, M. B. and FROST, M. G. (1974). Maximum autocorrelations for moving average processes. *Biometrika*, **61**, 199–200.

GRANGER, C. W. J. and HATANAKA M. (1964). *Spectral Analysis of Economic Time Series*. Princeton University Press.

NEWBOLD, P. and GRANGER, C. W. J. (1974). Experience with forecasting univariate time series and the combination of forecasts (with Discussion). *J. R. Statistical Soc.* A, **137**, 131–164.

WOLD, H. (1953). *A Study in the Analysis of Stationary Time Series*, 2nd ed. Stockholm: Alqvist and Wiksell.