

RESEARCH REPORTS DIVISION
NAVAL POSTGRADUATE SCHOOL
MONTEREY, CALIFORNIA 93943

NPS55-84-022

NAVAL POSTGRADUATE SCHOOL

Monterey, California



TECHNICAL

HIGHER ORDER RESIDUAL ANALYSIS FOR NONLINEAR TIME
SERIES WITH AUTOREGRESSIVE CORRELATION STRUCTURES

BY

P.A.W. Lewis
&
A. J. Lawrance

September 1984

Approved for public release; distribution unlimited

Prepared for:

Chief of Naval Research
Arlington, VA 22217

FEDDOCS
D 208.14/2:
NPS-55-84-022

NAVAL POSTGRADUATE SCHOOL
Monterey, California

Commodore R. H. Shumaker
Superintendent

David A. Schradly
Provost

This work was supported by the Office of Naval Research under
Grant NRO42-469.

Reproduction of all or part of this report is authorized.

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER NPS55-84-022	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) HIGHER ORDER RESIDUAL ANALYSIS FOR NONLINEAR TIME SERIES WITH AUTOREGRESSIVE CORRELATION STRUCTURES		5. TYPE OF REPORT & PERIOD COVERED Technical
7. AUTHOR(s) A. J. Lawrance P. A. W. Lewis		6. PERFORMING ORG. REPORT NUMBER
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Postgraduate School Monterey, CA 93943		8. CONTRACT OR GRANT NUMBER(s)
11. CONTROLLING OFFICE NAME AND ADDRESS Naval Postgraduate School Monterey, Ca 93943		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61153N; RR014-05-01 N0001484WR 24004
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) Chief of Naval Research Arlington, VA 22217		12. REPORT DATE September 25, 1984
		13. NUMBER OF PAGES 24
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited.		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number)		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) The paper considers nonlinear time series whose second order autocorrelations satisfy autoregressive Yule-Walker equations. The usual linear residuals are then uncorrelated, but not independent, as would be the case for linear autoregressive processes. Two such types on nonlinear model are treated in some detail: random coefficient autoregression and multiplicative autoregression. The proposed analysis involves crosscorrelation of the usual linear residuals and their squares. This function is obtained for the two types of model considered, and allows differentiation between models with the same auto-		

correlation structure in the same class. For the random coefficient models it is shown that one side of the crosscorrelation function is zero, giving a useful signature of these processes. The non-zero features of the crosscorrelations are informative of the higher order dependency structure. In applications this residual analysis requires only standard statistical calculations, and extends rather than replaces the usual second order analysis.

**HIGHER ORDER RESIDUAL ANALYSIS FOR NONLINEAR TIME SERIES WITH
AUTOREGRESSIVE CORRELATION STRUCTURES**

A J LAWRENCE

Department of Statistics
University of Birmingham
Birmingham, England

P A W LEWIS

Department of Operations Research
Naval Postgraduate School
Monterey, California, USA

ABSTRACT

The paper considers nonlinear time series whose second order autocorrelations satisfy autoregressive Yule-Walker equations. The usual linear residuals are then uncorrelated, but not independent, as would be the case for linear autoregressive processes. Two such types of nonlinear model are treated in some detail: random coefficient autoregression and multiplicative autoregression. The proposed analysis involves crosscorrelation of the usual linear residuals and their squares. This function is obtained for the two types of model considered, and allows differentiation between models with the same autocorrelation structure in the same class. For the random coefficient models it is shown that one side of the crosscorrelation function is zero, giving a useful signature of these processes. The non-zero features of the crosscorrelations are informative of the higher order dependency structure. In applications this residual analysis requires only standard statistical calculations, and extends rather than replaces the usual second order analysis.

1. INTRODUCTION

This paper is concerned with the statistical assessment of dependency beyond autocorrelation in the context of nonlinear time series models. The central theme is that residuals from models fitted according to inapplicable linearity assumptions can also profitably be used for further analysis beyond linearity. A well known property of linear time series residuals is that they are uncorrelated; in the context of linear models they should also be independent, apart from the effects of parameter estimation with short series. However, in the context of nonlinear models it is not often recognised that uncorrelated residuals also hold information concerning higher order dependence in the data. Developments of such a higher order residual analysis will be explored here, in particular for two types of nonlinear autoregressive model which have the usual linear Yule-Walker autoregressive correlation structure. Some higher order dependency correlations will be obtained. In a companion paper, the suggestion of reversed residuals will be made, and the analysis given in this paper will be extended to encompass these reversed residuals.

2. AUTOREGRESSION AND LINEAR AUTOREGRESSIVE RESIDUALS

2.1 Autoregression for Nonlinear Models

The standard form of autoregression needs widening for use with nonlinear models; we consider first the standard form which is explicitly autoregressive in a linear additive way and then several weaker variants. A stationary time series $\{X_t\}$ of mean μ is assumed. Under the *linear autoregressive model*, of order p , the $\{X_t\}$ satisfy the equation

$$X_t - \mu = \alpha_1(X_{t-1} - \mu) + \alpha_2(X_{t-2} - \mu) + \dots + \alpha_p(X_{t-p} - \mu) + \epsilon_t, \quad t=0, \pm 1, \pm 2, \dots \quad (2.1)$$

where the ϵ_t are independent and identically distributed and $\mu, \alpha_1, \alpha_2, \dots, \alpha_p$ are *fixed* parameters. A more general definition of autoregression of order p , could be the *linear conditional expectation* requirement that

$$E(X_t - \mu | X_{t-1}, X_{t-2}, \dots, X_{t-p}) \\ = \alpha_1(X_{t-1} - \mu) + \alpha_2(X_{t-2} - \mu) + \dots + \alpha_p(X_{t-p} - \mu), \quad t=0, \pm 1, \pm 2, \dots \quad (2.2)$$

The definition (2.1) implies (2.2) but not vice-versa. Thus this definition could apply to models which are not of the linear form (2.1), either because the ϵ_t are dependent, but still with $E(\epsilon_t | X_{t-1}, X_{t-2}, \dots, X_{t-p}) = 0$, or because the model has some other structure altogether. For instance, there are the random coefficient models of Nicholls and Quinn (1982), the exponential distribution random coefficient models of Lawrance and Lewis (1981, 1984), the discrete distribution random coefficient models of Jacobs and Lewis (1983), and the gamma-beta random coefficient models of Lewis (1981).

Random coefficient autoregressive models of order p take the general form

$$X_t = A_t^{(1)} X_{t-1} + A_t^{(2)} X_{t-2} + \dots + A_t^{(p)} X_{t-p} + B_t, \quad t=0, \pm 1, \pm 2, \dots \quad (2.3)$$

where the vector of coefficients $\{A_t^{(1)}, A_t^{(2)}, \dots, A_t^{(p)}, B_t\}$ is a stationary vector sequence of independent random variables, and sometimes in addition, B_t is independent of the random coefficients at time t . It is easy to see that such models satisfy the linear conditional expectation definition of p th order autoregression, but are nonlinear. A first order case of the type (2.3) will be used to illustrate the proposed method of residual analysis of autoregressive nonlinearity.

A further and weaker definition of autoregression is the requirement that the autocovariances of the $\{X_t\}$, denoted by $\{\gamma_k\}$, just satisfy Yule-Walker linear difference equations of the form,

$$\gamma_r = \alpha_1 \gamma_{r-1} + \alpha_2 \gamma_{r-2} + \dots + \alpha_p \gamma_{r-p}, \quad \gamma_r = \gamma_{-r}, \quad r=1, 2, \dots \quad (2.4)$$

for suitable constants $\alpha_1, \alpha_2, \dots, \alpha_p$; this will be referred to as *Yule-Walker autoregression*. It is true for processes which satisfy (2.1) and (2.2), as may be verified in the usual manner by multiplying X_t by X_{t-r} and taking expectations. The reverse is not true, (2.4) does not imply (2.1) or (2.2). A case in which (2.4) holds but in which (2.1) and (2.2)

do not, is the product autoregression model of McKenzie (1982), where, with $p=1$, $E(X_t|X_{t-1})$ is a fractional power of X_{t-1} ; this model will also be used in Section 4.3 as an illustration of the proposed residual methods.

A variety of other models can satisfy the Yule-Walker definition of autoregression, and yet not satisfy the linear conditional expectation definition. Amongst these are first order Markov chains under fairly general conditions, and others such as the semi-Markov generated processes discussed in Cox and Lewis (Chapter 7, 1966) and Lewis (1980). Notice that we do not define nonlinear autoregressive models in a constructive way. The class is so wide as to make this impossible; one such class has been studied by Jones (1978). Rather, we require that the autocorrelations should satisfy linear equations, similar in structure to those satisfied by the autocorrelation of linear autoregressive models. In view of this, our suggested analysis extends, rather than superceeds, conventional methods.

2.2 Definition and Discussion of Linear Autoregressive Residuals

For the analysis of time series data involving models satisfying (2.1), or (2.2) or minimally (2.3), the use of *linear autoregressive residuals* of order p , defined as

$$R_t^{(P)} = (X_t - \mu) - \alpha_1(X_{t-1} - \mu) - \alpha_2(X_{t-2} - \mu) - \dots - \alpha_p(X_{t-p} - \mu) \quad (2.5)$$

is suggested. This suggestion is based on the following theorem, which is a generalization of a result given in Lawrance and Lewis (1984, Section 7.2) for $p=2$.

Theorem. Let the stationary process $\{X_t\}$ satisfy the Yule-Walker type equations (2.4). Then the linear autoregressive residuals $\{R_t^{(P)}\}$ defined at (2.5) are uncorrelated (although not necessarily independent).

Proof. The autocovariances of the residuals (2.5) are

$$\begin{aligned} \text{Cov}[R_t^{(P)}, R_{t+r}^{(P)}] &= \text{Cov}[(X_t - \mu), R_{t+r}^{(P)}] - \alpha_1 \text{Cov}[(X_{t-1} - \mu), R_{t+r}^{(P)}] - \dots \\ &\quad - \alpha_p \text{Cov}[(X_{t-p} - \mu), R_{t+r}^{(P)}] \end{aligned} \quad (2.6)$$

$$\begin{aligned} &= \text{Cov}[(X_t - \mu), R_{t+r}^{(P)}] - \alpha_1 \text{Cov}[(X_t - \mu), R_{t+r+1}^{(P)}] - \dots \\ &\quad - \alpha_p \text{Cov}[(X_t - \mu), R_{t+r+p}^{(P)}]. \end{aligned} \quad (2.7)$$

Equation (2.7) follows because the $\{X_t\}$ process is stationary and consequently the $\{R_t^{(P)}\}$ process is stationary. The covariances in (2.7) need only be considered for positive lag since the autocovariance is an even function of r . Then the crosscovariances on the right-hand side of (2.7) are all of the same type and given by

$$\begin{aligned} \text{Cov}[(X_t - \mu), R_{t+r}^{(P)}] &= \text{Cov}[(x_t - \mu), \{(x_{t+r} - \mu) - \alpha_1(x_{t+r-1} - \mu) - \dots - \alpha_p(x_{t+r-p} - \mu)\}] \\ &= \gamma_r - \alpha_1 \gamma_{r-1} - \dots - \alpha_p \gamma_{r-p} \quad r=1, 2, \dots \end{aligned} \quad (2.8)$$

Now by the Yule-Walker equations (2.4), the expression (2.8) is zero. Thus using (2.8) in (2.7)

$$\text{Cov}[R_t^{(P)}, R_{t+r}^{(P)}] = 0 \quad r=\pm 1, \pm 2, \dots, \quad (2.9)$$

as was to be proven. The proof is immediate for the linear autoregressive model (2.1), since $R_t^{(P)} = \epsilon_t$, and ϵ_t and $R_{t+r}^{(P)}$ are by definition independent.

Note that the linear autoregressive residuals $\{R_t^{(P)}\}$ will generally still be dependent though uncorrelated in nonlinear modelling of the $\{X_t\}$; with the linear autoregressive model (2.1) the $\{R_t^{(P)}\}$ will not only be uncorrelated but also independent. *It is this difference which will be exploited to explore nonlinearity in p th order nonlinear autoregressive processes.* The dependency attributable to parameter estimation is taken to be small in the large scale applications we have in mind; indeed, nonlinear modelling of short series of data may well be hard to justify.

The quantities $R_t^{(P)}$ are autoregressive residuals in the sense that $R_t^{(P)}$ is the residual of X_t after subtracting off $\mu_t^{(P)}$, its best linear least squares predictor in terms of $X_{t-1}, X_{t-2}, \dots, X_{t-p}$, given by

$$\mu_t^{(P)} = \mu + \alpha_1(X_{t-1} - \mu) + \alpha_2(X_{t-2} - \mu) + \dots + \alpha_p(X_{t-p} - \mu). \quad (2.10)$$

Thus the residuals $\{R_t^{(P)}\}$ give the basic way of taking out the linear

correlation component in models with p th order autoregression minimally of the 'Yule-Walker' form (2.4). Note too that if the process is autoregressive in this sense, linear autoregression on terms beyond $X_{t-p-\mu}$, for example $(X_{t-p-k-\mu})$, will yield $\alpha_{p+k}=0$, $k=1,2,\dots$; this is a property of the linear component being taken out so as to minimize the expected mean square of the residual $R_t^{(p+k)}$.

A further point worth noting about the residuals $R_t^{(p)}$ concerns their crosscovariances with the X_t 's. In the proof of the theorem it is seen that $\text{Cov}[X_{t-\mu}, R_{t+r}^{(p)}] = 0$ for $r = 1, 2, \dots$. However, the other half of these crosscovariances is non-zero.

Example: The nonGaussian linear AR(1) model

Taking (2.1) with $p=1$, and ρ instead of α_1 , the crosscorrelation of $X_{t-\mu}$ and $R_{t-r}^{(p)}$ in the AR(1) model is given by

$$\text{Corr}(X_{t-\mu}, R_{t-r}) = (1-\rho^2)^{1/2} \rho^r \quad \text{for } r=1, 2, 3, \dots; \quad (2.11)$$

the superscript has been dropped from R_{t-r} , as it will be in similar future use. Note further, that in this case R_{t+r} is independent of X_t for $r=1, 2, \dots$. Further use of (2.11) will be made in Sections 3 and 5.

For the use of $\{R_t^{(p)}\}$ in data analysis, the order p of the linear aspect of the autoregression needs to have been chosen; any of the available standard methods may still be used. In addition, of course, the $\mu, \alpha_1, \alpha_2, \dots, \alpha_p$ need to be estimated; there are at least two convenient possibilities: (1) the assumption of a linear autoregressive model like (2.1) and the use of least squares estimation, and (2) a non-model based approach to estimation employing the first p Yule-Walker type equations (2.4). The latter is suggested here; however, Tjostheim and Paulsen (1983) recommend (1) when dealing with modest sized samples from linear models, on account of serious estimation bias with the Yule-Walker estimates. Since the use envisaged here is primarily with nonlinear models and substantial sets of data, the superiority of (1) over (2) is not established or

crucial.

3. ASSESSMENT OF HIGHER ORDER DEPENDENCE

Since the linear autoregressive residuals $\{R_t^{(p)}\}$ of (2.5) have zero covariances when (2.4) holds, but for nonlinear processes need not be independent, a residual analysis of this nonlinearity can be based on an assessment of their higher order dependence. To consider what form this might take, we note that the use of $\{X_t^2\}$ has been suggested by Granger and Andersen (1978,p.63) for bilinear models, for which many of the simple cases have zero autocorrelations themselves. The corresponding initial suggestion here is, as was briefly illustrated in Lawrance and Lewis (1984), to use the residuals $\{(R_t^{(p)})^2\}$. Displays can then easily be made of the associated autocorrelation functions, scatter plots, periodograms, cumulative periodograms, etc, using standard (second order) time series software.

However, the autocorrelations of $\{(R_t^{(p)})^2\}$ are fourth order quantities in the original series $\{X_t\}$, which is a double jump from the second order autocorrelations of the series $\{X_t\}$. Such quantities will be very difficult to handle theoretically with most types of nonlinear model. The crosscorrelation function of $\{R_t^{(p)}\}$ and $\{(R_t^{(p)})^2\}$ is essentially 3rd order, but by involving the variance of $\{(R_t^{(p)})^2\}$, also needs some fourth order joint moments of $\{X_t\}$ up to lag p . The necessary calculations of all these quantities for two nonlinear models will be given in Section 4.

The behaviour of the crosscorrelation function of $R_t^{(p)}$ and $\{(R_t^{(p)})^2\}$ may be judged against the fact that the $\{R_t^{(p)}\}$ are *independent* for the linear autoregressive model of order p , and hence it will be zero except at lag zero. For the random coefficient autoregressive processes considered in Section 4, the suggested crosscorrelation function will be shown to possess a useful cut-off property; this generalizes the NEAR(2) result given

in Lawrance and Lewis (1984).

Two rather simpler crosscorrelation quantities can be proposed from consideration of the zero covariance result of (2.8), and the additional fact that when the pth order linear autoregressive model holds, X_t and $R_{t+r}^{(P)}$ ($r=1,2,\dots$) will be independent, and not just uncorrelated. Working in terms of $\{X_t\}$ adjusted for its mean μ , which is better computationally and often nicer theoretically, the following may be considered

$$\text{Corr}[(X_t - \mu)^2, R_{t+r}^{(P)}], \quad \text{for } r=0, \pm 1, \pm 2, \dots \quad (3.1)$$

$$\text{Corr}[X_t - \mu, (R_{t+r}^{(P)})^2], \quad \text{for } r=0, \pm 1, \pm 2, \dots \quad (3.2)$$

These autocorrelations are not equivalent in the aspects of higher order dependency of the $\{X_t\}$ process which they assess. To see this, note that the covariance corresponding to (3.1) involves only third order joint moments of the form $E[(X_t - \mu)^2(X_{t+r-i} - \mu)]$, $i=0, \pm 1, \dots, \pm p$ while the covariance corresponding to (3.2) involves additional joint moments of the form $E[X_t - \mu)(X_{t+r} - \mu)(X_{t+r-i})]$, $i=0, \pm 1, \dots, \pm p$. Further, the denominator of (3.2), by needing $\text{var}\{(R_t^{(P)})^2\}$, is more complicated in the higher order moments it involves relative to the denominator of (3.1) which needs $\text{var}\{R_t^{(P)}\}$. However, both correlations, and particularly (3.1), are more tractable than the autocorrelations of $\{(R_t^{(P)})^2\}$ or the crosscorrelations of $\{R_t^{(P)}\}$ and $\{(R_t^{(P)})^2\}$.

Example (Continued from (2.11)): The nonGaussian linear AR(1) model.

As an illustration of the use of (3.1) and (3.2), for the linear AR(1) model, ((2.1) with $p=1$ and ρ instead of α_1) there are the results

$$\text{Corr}[(X_t - \mu), (R_{t-r})^2] = \begin{cases} 0 & r=-1, -2, \dots \\ \text{skew}(X) \frac{\text{var}(X)}{\text{var}(R^2)} (1 - \rho^3) \rho^r, & r=0, 1, \dots \end{cases} \quad (3.3)$$

$$\text{Corr}[(X_t - \mu)^2, R_{t-r}] = \begin{cases} 0 & r=-1, -2, \dots \\ \text{skew}(X) \frac{\text{Var}(X)}{[\text{Var}\{(X - \mu)^2\}]^{1/2}} \left[\frac{1 - \rho}{1 + \rho} \right]^{1/2} \rho^{2r}, & r=0, 1, \dots \end{cases} \quad (3.4)$$

Note the faster geometric decrease in (3.4) relative to (3.3); (3.4)

contains more higher moment information. The other halves of both these

crosscorrelation functions are of course zero. Behaviour of this type for these functions - zero for $r < 0$ and geometric decay for $r > 0$ - would suggest linear models with skewed marginal distributions or nonlinear models. Discussion of parallel results for two types of nonlinear model are given in section 4.

The use of squaring in the construction of these higher order dependency measures is recognized as being pragmatic and somewhat arbitrary; it does however lead to expressions involving selected types of simpler higher order cross moments. The end use of the higher order dependency measures can either be exploratory for a given data set, to ascertain whether there is appreciable nonlinearity present, or constructively, to provide evidence for fitting suitable types of nonlinear model which can match the observed higher order dependency.

Earlier discussion of higher order dependence in nonGaussian *Linear* models is given by Rosenblatt (1980) in terms of the bispectrum; aspects of nonlinearity and higher order spectra are briefly considered in Rosenblatt (1979).

4. HIGHER ORDER DEPENDENCY FOR TWO TYPES OF NONLINEAR MODEL

In this section we obtain properties of the measures of higher order dependency discussed in the previous section, for two specific types of nonlinear models. The models considered are autoregressive in the Yule-Walker sense of (2.4), but not in the linear sense of (2.1).

4.1 Random Coefficient Autoregressive Models

As already remarked in Section 1, a generalization of the linear autoregressive model (2.1) is to let the coefficients $\alpha_1, \alpha_2, \dots, \alpha_p$ be random variables. One general class of such models has been discussed by Nicholls and Quinn (1982) who cite Andell (1976) and articles in the economic literature. Other classes of models with random coefficients include the discrete distribution models of Jacobs and Lewis (1983) and the exponential models of Lawrance and Lewis (1981, 1984). The class of random coefficient autoregressive processes to be of interest here is given by

$$X_t = A_t^{(1)}X_{t-1} + A_t^{(2)}X_{t-2} + \dots + A_t^{(p)}X_{t-p} + B_t \quad t=0, \pm 1, \pm 2, \dots, \quad (4.1)$$

where $\{A_t^{(1)}, A_t^{(2)}, \dots, A_t^{(p)}, B_t\}$ is a stationary vector sequence of independent random variables with $E(A_t^{(j)}) = \alpha_j$ for $j=1, \dots, p$. The components of the vectors are not necessarily independent. For example the discrete distribution models of Jacobs and Lewis (1983) can be written in this form, and have dependent coefficients, as do the exponential models of Lawrance and Lewis (1981, 1984). It is easily verified that the process (4.1) satisfies conditional expectation autoregression (2.2) and thus also the weaker Yule-Walker definition (2.4); it will also clearly be stationary. Note also that the standard linear autoregressive model (2.1) is a special case of (4.1) in which the random coefficients are actually constant.

We now give a characteristic result for this type of process when higher order dependency is measured by $\text{Corr}[R_t^{(p)}, (X_{t+r} - \mu)^2]$ or $\text{Corr}[R_t^{(p)}, (R_{t+r}^{(p)})^2]$, assuming that the $R_t^{(p)}$'s are uncorrelated.

Theorem. With the random coefficient model (4.1),

$\text{Corr}[R_t^{(P)}, (X_{t-r} - \mu)^2]$ and $\text{Corr}[R_t^{(P)}, (R_{t-r}^{(P)})^2]$ are equal to zero for $r=1, 2, \dots$

Proof: Using the definitions (2.5) and (4.1),

$$\begin{aligned} R_t^{(P)} &= (X_{t-\mu})^{-\alpha_1} (X_{t-1-\mu})^{-\alpha_2} \dots - \alpha_p (X_{t-p-\mu}) \\ &= (A_t^{(1)})^{-\alpha_1} X_{t-1} + (A_t^{(2)})^{-\alpha_2} X_{t-2} + \dots + (A_t^{(P)})^{-\alpha_p} X_{t-p} \\ &\quad + B_t^{-(1-\alpha_1 \dots - \alpha_p)} \mu \end{aligned} \quad (4.2)$$

On multiplying (4.2) by $(R_{t-r}^{(P)})^2$ and taking expectations, thus obtaining $\text{Cov}[R_t^{(P)}, (R_{t-r}^{(P)})^2]$, we have a sum of p terms given by

$$\sum_{j=1}^p E\{(A_t^{(j)})^{-\alpha_j} X_{t-j} (R_{t-r}^{(P)})^2\} \quad (4.3)$$

and a last term involving B_t which is clearly zero. Now $(A_t^{(j)})^{-\alpha_j}$ is independent of both the X_{t-j} and $(R_{t-r}^{(P)})^2$, for $r=1, 2, \dots$, which may nevertheless be themselves dependent. Thus the j th term in (4.3) becomes

$$E(A^{(j)})^{-\alpha_j} E\{X_{t-j} (R_{t-r}^{(P)})^2\} = 0 \quad \text{for } r=1, 2, \dots,$$

since $E(A^{(j)})^{-\alpha_j} = \alpha_j$. This completes the proof which clearly includes the first cross correlation mentioned in the theorem. *The proof highlights the fact that it is the independence of the vector of coefficients $\{A_t^{(1)}, \dots, A_t^{(P)}, B_t\}$ on previous X_t 's which creates the effect in this type of model.*

The results of the theorem can be used to help validate random coefficient autoregressive models; also useful in this respect are the non-zero higher order residual crosscorrelations, eg for positive r $\text{Corr}[R_t^{(P)}, (R_{t+r}^{(P)})^2]$. These have been obtained for the second order autoregressive exponential process studied in Lawrence and Lewis (1984); similar results for any first order random coefficient model of the type (4.1) are given in the next subsection.

It is worth noting that with G_{t-r} defined as any reasonable function of $(X_{t-r}, X_{t-r-1}, \dots)$, a similar argument to that given in the proof shows that

$$\text{Corr}[R_t^{(P)}, G_{t-r}] = 0, \quad r=1, 2, \dots \quad (4.4)$$

A result of this type is not, however, sufficient to establish, for instance, that $R_t^{(p)}$ and $R_{t-r}^{(p)}$ are independent; they are dependent.

The random coefficient autoregressive structure of (4.1) is not a necessary condition for the results of this section to hold; this may be illustrated by noting that $\text{Corr}[R_t^{(p)}, (X_{t-r}-\mu)^2]$ is zero when $C_{12}(r) = E\{(X_t-\mu)(X_{t-r}-\mu)^2\}$ satisfy the equations

$$C_{12}(r) = \alpha_1 C_{12}(r-1) + \alpha_2 C_{12}(r-2) + \dots + \alpha_p C_{12}(r-p), \quad r=1, 2, \dots \quad (4.5)$$

The similarity of these equations to standard Yule-Walker equations (2.4) will be apparent. Thus any process with this property will have $\text{Corr}[R_t^{(p)}, (X_{t-r}-\mu)^2]$ equal to zero for non-negative r .

4.2 Higher Order Dependency for First Order Random Coefficient Autoregressive Models

The model to be considered is the first order ($p=1$) case of (4.1), now to be denoted as

$$X_t = A_t X_{t-1} + B_t, \quad t=0, \pm 1, \pm 2, \dots, \quad (4.6)$$

in which A_t and B_t are independent within and between each t ; we also write

$$a = E(A_t), \quad \mu = E(X_t) = (1-a)^{-1} E(B_t).$$

Many basic mathematical and probabilistic properties of this equation have been studied by Vervaat (1979). Interest here is restricted mainly to the residual crosscovariances of (R_t^2, R_{t-r}) and $[(X_t-\mu)^2, R_{t-r}]$ where R_t is the first order residual given by

$$R_t = X_t - \mu - a(X_{t-1} - \mu). \quad (4.7)$$

It has been established in Section 4.1 that $\text{Cov}(R_t, R_{t-r}^2) = \text{Cov}(R_t^2, R_{t+r}) = 0$ for $r=1, 2, \dots$, and so now $\text{Cov}(R_t^2, R_{t-r})$ for $r=0, 1, 2, \dots$ is obtained explicitly.

The calculation begins by writing the required covariance as

$$\begin{aligned} \text{Cov}(R_t^2, R_{t-r}) &= E\{[(X_t-\mu) - a(X_{t-1}-\mu)]^2 [(X_{t-r}-\mu) - a(X_{t-r-1}-\mu)]\} \\ &= C_{21}(r) - aC_{21}(r+1) - 2a\{C_{111}(r) - aC_{111}(r+1)\} \\ &\quad + a^2\{C_{21}(r-1) - aC_{21}(r)\} \end{aligned} \quad (4.8)$$

where, as defined just before (4.5),

$$C_{21}(r) = E\{(X_t - \mu)^2(X_{t-r} - \mu)\}, \quad C_{111}(r) = E\{(X_t - \mu)(X_{t-1} - \mu)(X_{t-r} - \mu)\}.$$

The higher order triple moment $C_{111}(r)$ is easily seen to be equivalently given by $aC_{21}(r-1)$, in the present instance, and hence (4.8) becomes,

$$\text{Cov}(R_t^2, R_{t-r}) = C_{21}(r) - aC_{21}(r+1) - a^2\{C_{21}(r-1) - aC_{21}(r)\}, \quad r=1, 2, \dots \quad (4.9)$$

For $r=0$, with $\mu_3 = E[X_t - \mu]^3$, there is the separate result

$$E(R_t^3) = (1 + 2a^3)\mu_3 - 3aC_{21}(1). \quad (4.10)$$

The calculation of $C_{21}(r)$ is effected by writing the defining equation (4.6) in the form

$$X_t - \mu + \mu = A_t(X_{t-1} - \mu) + (\mu A_t + B_t).$$

Squaring both sides, multiplying by $X_{t-r} - \mu$, and taking expectations then gives

$$C_{21}(r) + 2\mu C(r) = a_2 C_{21}(r-1) + 2(\mu a_2 + E(A_t B_t))C(r-1) \quad (4.11)$$

where

$$C(r) \equiv \text{Cov}(X_t, X_{t-r}) = a^r \sigma^2, \quad \sigma^2 \equiv \text{var}(X_t), \quad a_2 \equiv E(A_t^2).$$

Further simplifications of (4.11) using $C(r) = a^r \sigma^2$ gives the recursive equation

$$C_{21}(r) = a_2 C_{21}(r-1) + b_3 a^{r-1}, \quad r=1, 2, \dots, \quad (4.12)$$

where

$$b_3 = 2\sigma^2\{\mu \text{var}(A_t) + \text{Cov}(A_t, B_t)\}.$$

Equation (4.12) has explicit solution

$$C_{21}(r) = a_2^r \mu_3 + b_3(a_2^r - a^r)/(a_2 - a), \quad r=0, 1, \dots \quad (4.13)$$

Going back to (4.9) for $\text{Cov}(R_t^2, R_{t-r})$ and using (4.12), we have finally

$$\begin{aligned} \text{Cov}(R_t^2, R_{t-r}) &= (a_2 + a^3 a_2 - a a_2^2 - a^2) C_{21}(r-1) \\ &\quad + (1 - a^2 + a^3 - a a_2) b_3 a^{r-1}, \quad r=1, 2, \dots \quad (4.14) \end{aligned}$$

Thus, (4.14) and (4.13) constitute the explicit solution for the crosscovariances of the residuals (R_t^2, R_{t-r}) ; notice that when A_t is the constant a , the case of the standard first order linear model, these crosscovariances are correctly zero. The NEAR(1) model of Lawrance and Lewis (1981) specializes (4.6), by having independence within each (A_t, B_t) and particular forms for A_t and B_t .

If the covariances (4.14) are to be converted to correlations, then $\text{Var}(R_t)$ and $\text{Var}(R_t^2)$ must be obtained. This requires the following

calculation,

$$\text{Var}(R_t^2) = E(R_t^4) - [E(R_t^2)]^2 \quad (4.15)$$

where

$$E(R_t^2) = (1 - a^2)\sigma^2 = \text{Var}(R_t), \quad (4.16)$$

$$E(R_t^4) = (1+a^4)\mu_4 - 4aC_{31}(1) + 6a^2C_{22}(1) - 4a^3C_{13}(1), \quad (4.17)$$

with

$$\mu_4 = E[(X_t - \mu)^4], \quad C_{ij}(1) = E[(X_t - \mu)^i (X_{t-1} - \mu)^j], \quad (i, j) = (3, 1), (3, 2), (1, 3). \quad (4.18)$$

The joint moments in (4.17) were obtained in terms of the first four moments about the mean of X_t , the first three moments about zero of A_t and the first two moments about zero of B_t . These are not reproduced here. The explicit result for (4.15) was checked numerically against a slightly different method of computations while being used to obtain the illustrations given in Figure 1 and 2 for NEAR(1) models.

It was remarked in Section 3, following (3.2) that the crosscorrelations of $(X_t - \mu)^2$ and $R_{t+r}^{(p)}$ can also be useful in assessing higher order dependence. In the present case of first order autoregression there are the results,

$$\text{Cov}[(X_t - \mu)^2, R_{t-r}] = C_{21}(r) - a C_{21}(r+1), \quad (4.19)$$

$$\text{Var}(R_t) = (1-a^2)\sigma^2, \quad \text{Var}[(X_t - \mu)^2] = E(X_t - \mu)^4 - \sigma^4. \quad (4.20)$$

These formula can all be applied to the NEAR(1) model quite simply, noting that in this case

$$\begin{aligned} E(A_t^r) &= \alpha\beta^r, \quad E(B_t^r) = r! [p + (1-p)b]^r \\ b &= (1-\alpha)\beta, \quad p = (1-\beta) / [1 - (1-\alpha)\beta]. \end{aligned} \quad (4.21)$$

Figures 1 and 2 give the results of computations of $\text{Corr}[R_t^2, R_{t-r}]$ and $\text{Corr}[(X_t - \mu)^2, R_{t-r}]$ for the NEAR(1) model. The top left frame of Figure 1 is the linear FAR(1) case for which all cross-correlations apart from lag zero are zero; the other three cases in Figure 1 each have zero crosscorrelations at negative lags, in agreement with the theoretical results in Section 4.1, but have some non-zero values at the zero and positive lags. It is evident that the lag zero cross correlations contain

NEAR(1) RESIDUAL CROSSCORRELATIONS ($R(T) \cdot 2, R(T-R)$)

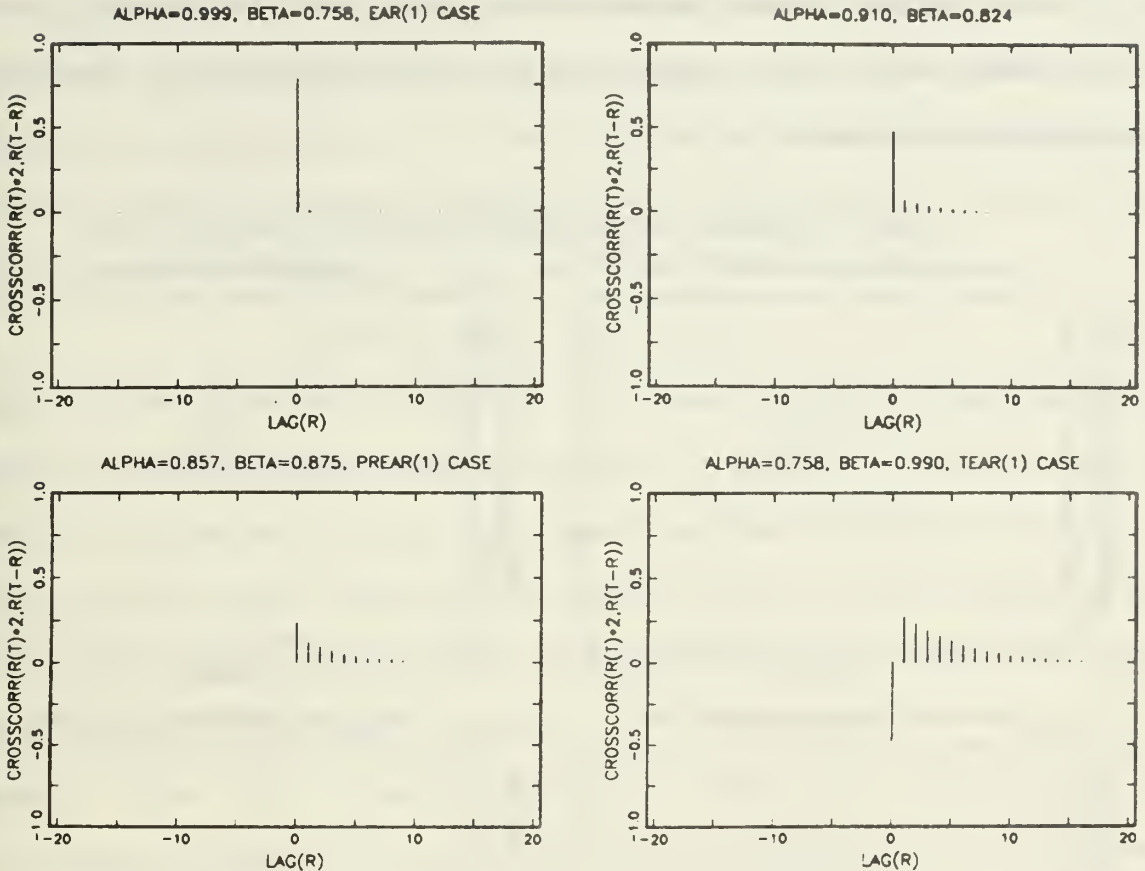


Figure 1. Four computation of the crosscorrelations for various lags, between the linear autoregressive residual, $R_t(P)$, and $\{R_t(P)\}^2$ for the NEAR(1) process with $\rho(1)=\alpha\beta$ held constant at 0.75; in effect the remaining free parameter is being varied through its allowable range.

much of the discriminating information between the four cases; this will be so for first order autoregressive models in which much of the higher order dependency is at lag one, and the lag zero cross correlation involves both third and fourth order aspects of the lag one dependency. For further information on the TEAR(1) and PREAR(1) cases see Lawrance and Lewis (1981).

Figure 2 gives $\text{Corr}[(X_t - \mu)^2, R_{t-r}]$ for the same four cases used in Figure 1. The negative lags are again zero for all cases, and hence this

property does not discriminate the linear from the nonlinear cases. Rather it is the strength of the crosscorrelations at positive lags which performs this task, albeit less clearly than the crosscorrelations of the residuals and squared residuals.

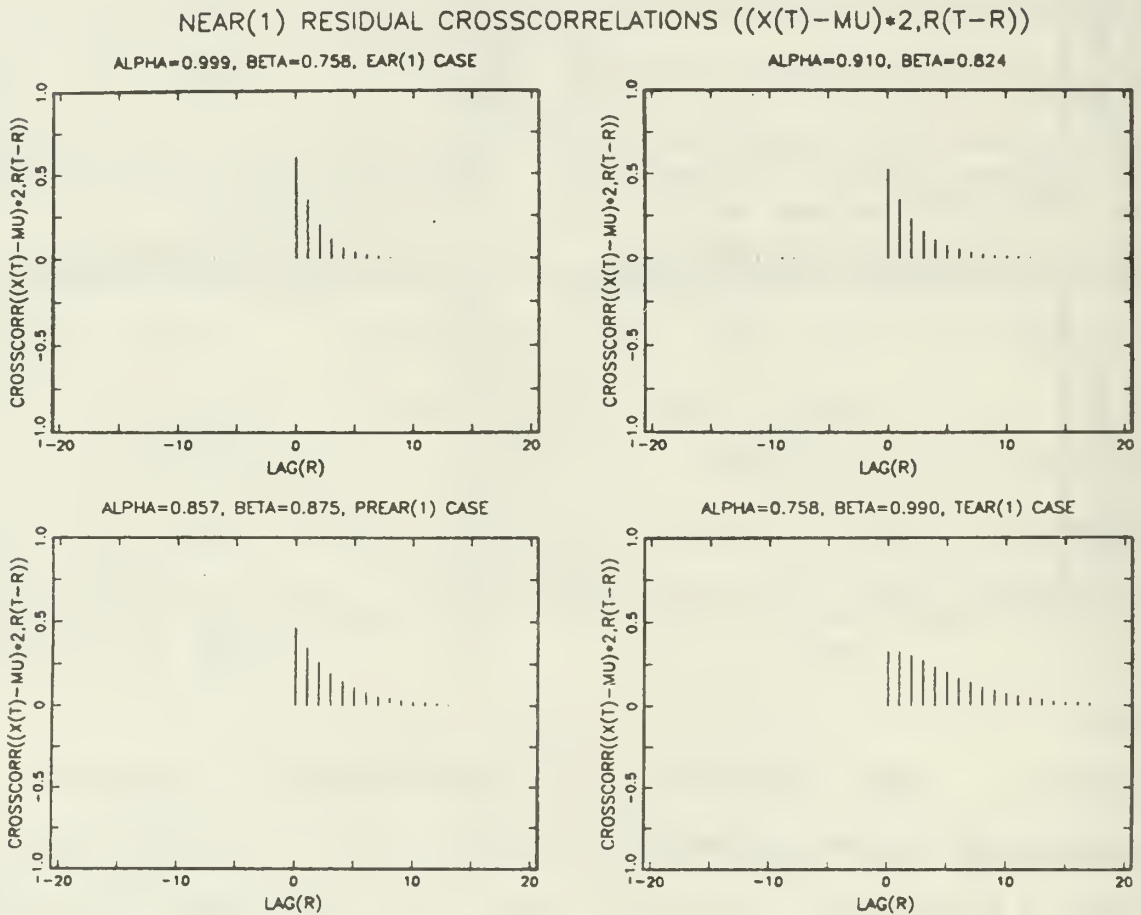


Figure 2. Four computations of the crosscorrelations, for various lags between the linear autoregressive residual, $R_t(P)$, and $(X_t - \mu)^2$ for the NEAR(1) process with $\rho(1) = \alpha\beta$ held constant at 0.75; in effect the remaining free parameter is being varied through its allowable range.

4.3 Higher Order Dependency for the Product Autoregression Model

Another form of nonlinear autoregressive model, called PAR(1), was introduced by McKenzie (1982); the basic idea of its construction is to consider the exponentiation of the standard AR(1) equation. Then its additive structure becomes multiplicative, and the general form of the PAR(1) product autoregressive model is

$$X_t = (X_{t-1})^\rho B_t, \quad t=0, \pm 1, \pm 2, \dots, \quad (4.22)$$

where ρ is the dependency parameter, $0 \leq \rho < 1$, and $\{B_t\}$ is an independent and identically distributed innovation sequence. McKenzie studies the model when B_t is chosen so that $\{X_t\}$ has a gamma marginal distribution, and obtains several basic results; for instance, that ρ still represents the lag one autocorrelation and that the autocorrelations in general satisfy the Yule-Walker first order equations ((2.4) with $p=1$). It is apparent, however, that the linear conditional expectation definition of autoregression (equation (2.4)) is not satisfied because of the power form implied by (4.22).

As with the first order random coefficient models in Section 4.2, we consider the first order residual R_t , given by (4.7), and will likewise determine $\text{Cov}(R_t^2, R_{t-r})$ for $r=0, \pm 1, \pm 2, \dots$, noting that for this model these correlations are non-zero for all lags. We will use the general expression (4.8) in terms of the third order central moments $C_{21}(r)$ and $C_{111}(r)$, but this time there is no simple relation between them, and both are needed for all lags. Also, calculation of $C_{21}(r)$ and $C_{111}(r)$ must be in terms of their uncentered components, since these are the quantities which can immediately be determined from the PAR(1) equation (4.22). The required uncentered moments will be written as

$$\begin{aligned} e_2 &= E(X_t^2), & e_3 &= E(X_t^3) \\ e_{21}(r) &= E(X_t^2 X_{t-r}), & e_{111}(r) &= E(X_t X_{t-1} X_{t-r}) \end{aligned} \quad (4.23)$$

for $r=0, \pm 1, \pm 2, \dots$, and there is need to note the special cases,

$$e_{21}(0) = e_3, \quad e_{111}(0) = e_{21}(1), \quad e_{111}(1) = e_{21}(-1).$$

All these quantities will be calculated.

First, it is necessary to note the following relations between the centred and uncentred moments,

$$C_{21}(r) = e_{21}(r) - \mu^3 - 2\mu C(r) - \mu\sigma^2 \quad (\text{all } r)$$

$$C_{111}(0) = C_{21}(1), \quad C_{111}(1) = C_{21}(-1) \quad (4.24)$$

$$C_{111}(r) = e_{111}(r) - \mu^3 - \mu\{C(r) + C(r-1) + C(1)\}, \quad (r \neq 0, 1).$$

The calculations of $e_{21}(r)$ and $e_{111}(r)$ need to be treated separately for positive and negative lags, but follow in the same general manner and will be illustrated by that for $e_{21}(r)$ for positive lags. By iterating the PAR(1) equation (4.22) r steps backward,

$$X_t = X_{t-r} \rho(r) \prod_{i=0}^{r-1} B_{t-i} \rho(i)$$

where $\rho(r) = \rho^r, r=0, 1, \dots$. Squaring this equation, multiplying it by X_{t-r} and taking expectations, gives

$$e_{21}(r) = E\{X_{t-r}^{2\rho(r)+1} \prod_{i=0}^{r-1} E\{B_{t-i}^{2\rho(i)}\}\}. \quad (4.25)$$

To obtain the expectations in the repeated product, taking the $2\rho(i)$ th power of (4.22), leads to

$$E\{X_t^{2\rho(i)}\} = E\{X_{t-1}^{2\rho(i)}\} E\{B_t^{2\rho(i)}\},$$

and hence dropping the unnecessary suffix t , to the result

$$E\{B^{2\rho(i)}\} = E\{X^{2\rho(i)}\} / E\{X^{2\rho(i+1)}\}.$$

Now (4.25) can be expressed purely in moments of X , as

$$e_{21}(r) = E\{X^{2\rho(r)+1}\} E\{X^2\} / E\{X^{2\rho(r)}\}. \quad (4.26)$$

To proceed further, invoke a gamma marginal distribution for X with density

$$f(x) = \theta \beta x^{\beta-1} e^{-\theta x} / \Gamma(\beta), \quad \theta, \beta > 0, \quad x \geq 0; \quad (4.27)$$

this has mean β/θ , variance β/θ^2 , third central moment $2\beta/\theta^3$, and there is the k th moment result $E\{X^k\} = \Gamma(\beta+k)/\theta^k \Gamma(\beta)$. From (4.26) and similar expressions we then have

$$e_{21}(r) = \begin{cases} \beta(\beta+1)(\beta+2\rho^r)/\theta^3, & r=1,2,\dots \\ \beta(\beta+\rho^{|r|})(\beta+\rho^{|r|+1})/\theta^3, & r=-1,-2,\dots \end{cases} \quad (4.28)$$

$$e_{111}(r) = \begin{cases} \beta(\beta+\rho)(\beta+\rho^{r-1}+\rho^r)/\theta^3, & r=2,3,\dots \\ \beta(\beta+\rho^{|r|})(\beta+\rho+\rho^{|r|+1})/\theta^3, & r=-1,-2,\dots \end{cases} \quad (4.29)$$

with $e_{21}(0)$, $e_{111}(0)$ and $e_{111}(1)$ being given by the special cases of

(4.23). Use of (4.24) gives finally the required expressions

$$C_{21}(r) = \begin{cases} 2\beta\rho^r/\theta^3 & r=0,1,\dots \\ \beta(\rho^{|r|}+\rho^{2|r|})/\theta^3, & r=-1,-2,\dots \end{cases} \quad (4.30)$$

$$C_{111}(r) = \begin{cases} \beta(\rho^r+\rho^{r+1})/\theta^3, & r=2,3,\dots \\ \beta\{(1+\beta)\rho^{|r|+1}+\rho^{2|r|+1}-\beta\rho^{|r|-1}\}/\theta^3 & r=1,2,\dots \end{cases} \quad (4.31)$$

All these may be used in (4.8) to obtain the desired $\text{Cov}(R_t^2, R_{t-r})$ for $r=0, \pm 1, \pm 2, \dots$; also the simpler covariance or correlation of $(X_t - \mu)^2$ and R_{t-r} is easily available from (4.19) and (4.31).

Conversion of these covariances to correlations requires $\text{Var}(R_t^2)$ as at (4.15); the required intermediate results are

$$\mu_4 = 9\beta^3/\theta^4,$$

$$C_{31}(1) = 3\beta(\beta+2)\rho, \quad C_{13}(1) = \beta\{3\beta+(1+\rho)(2+\rho)\}\rho,$$

$$C_{22}(1) = \beta^2+2\beta\rho+2\beta(\beta+2)\rho^2. \quad (4.32)$$

These then give the explicit expression

$$\begin{aligned} \text{Var}(R_t^2) = & \{(9\beta-1)\beta^2 - 4\beta(\beta+6)\rho^2 + 12\beta\rho^3 \\ & + \beta(9\beta^2-\beta+16)\rho^4 - 12\beta\rho^5 - 4\beta\rho^6\}/\theta^4 \end{aligned} \quad (4.33)$$

Together with $\text{Var}(R_t)$ which is simply $(1-\rho^2)\beta/\theta^2$, $\text{Cov}(R_t^2, R_{t-r})$ can be converted into $\text{Corr}(R_t^2, R_{t-r})$.

Figures 3 and 4 illustrate the residual and squared residual cross correlation for the PAR(1) model. Figure 3 shows how these residual cross correlations vary over the range of ρ values from $\rho=0$ to $\rho=0.9$ for an exponential marginal distribution. At $\rho=0$, the PAR(1) is an IDD process and the residuals are trivially independent, resulting in the only non-zero crosscorrelation at lag zero. As the ρ value increases the dependence spreads out, most strongly at lag minus one and positive lags. An interesting feature is the lag zero crosscorrelation which changes from

being strongly positive at $\rho=0$ to approximately zero at $\rho=0.75$, to moderately negative at $\rho=0$.

PAR(1) RESIDUAL CROSSCORRELATIONS ($R(T)+2R(T-R)$)

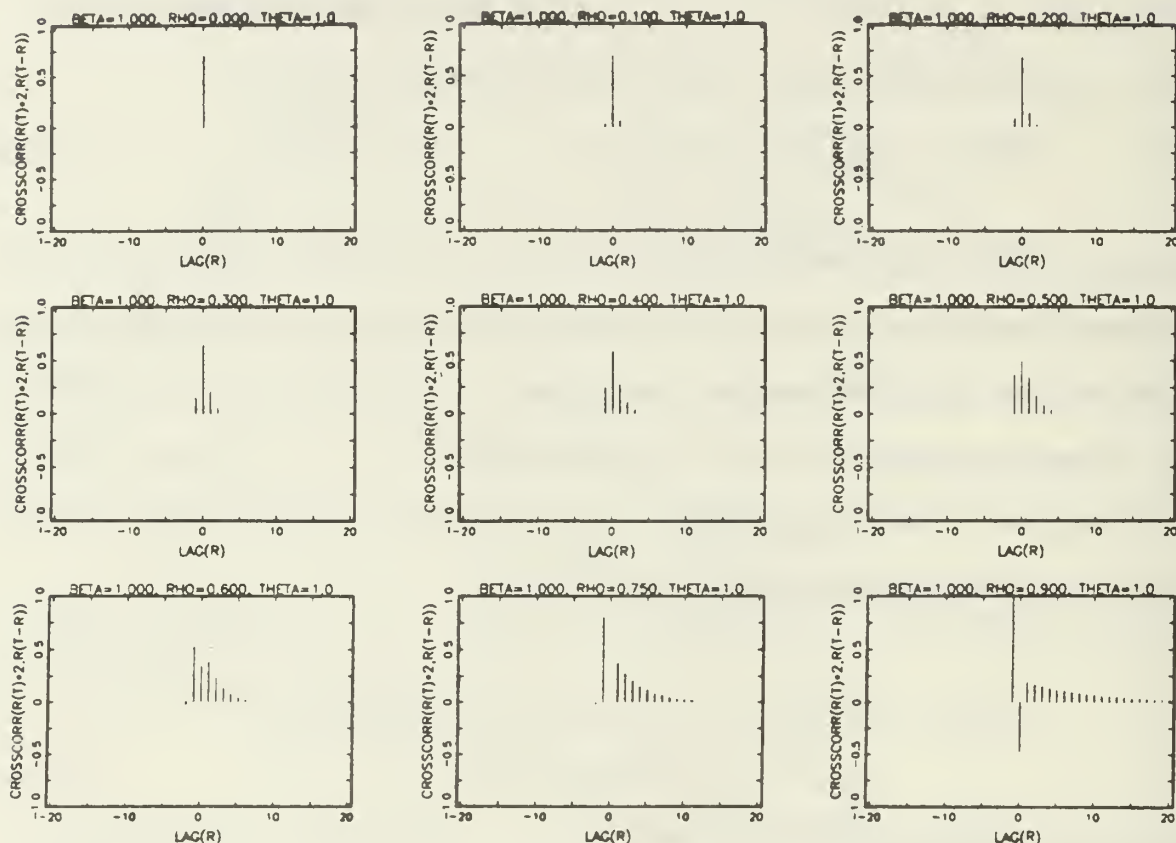


Figure 3. Computations of the residual crosscorrelations, $Corr(R_t^2, R_{t-r})$ for the PAR(1) model. Since $\beta=1$ the marginal distribution is exponential. The lag one serial correlation is increased from $\rho(1)=0.0$ (upper left) to $\rho(1)=0.9$ (lower right).

Figure 4 gives four different gamma cases of the PAR(1) residual cross correlations, all with $\rho=0.75$. The gamma shape parameter β takes the values 0.5, 1.0, 2.5 and 10.0; these cases indicate that changes in the gamma parameter cause only modest changes in detail of the cross correlations.

PAR(1) RESIDUAL CROSSCORRELATIONS ($R(T) \neq 2, R(T-R)$)

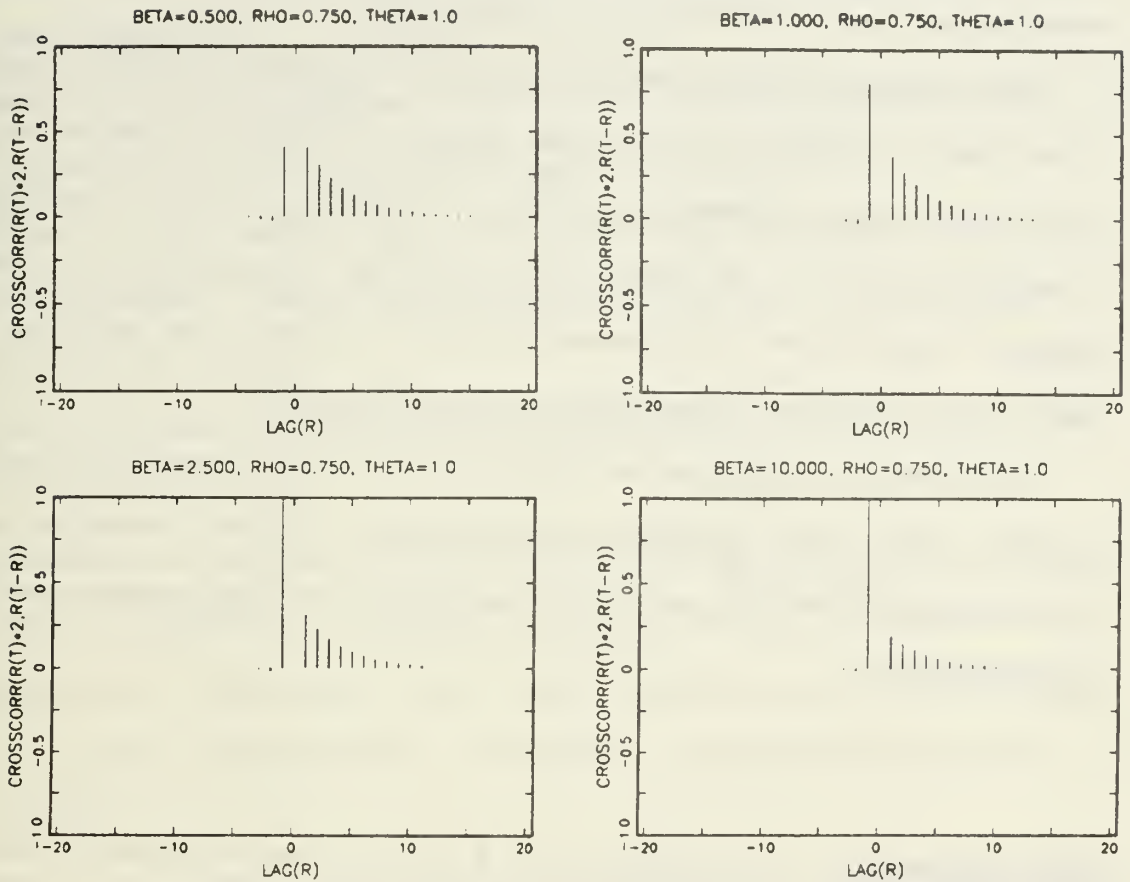


Figure 4. Computations of the residual crosscorrelations, $\text{Corr}(R_t^2, R_{t-r})$ for the PAR(1) model. The ρ for each case of 0.75, and the figures illustrate the effect on the residual crosscorrelations of changing the index β of the gamma distribution through 0.5, 1.0, 2.5 and 10.0.

5. CONCLUSIONS

A methodology for analyzing higher order dependence in nonlinear time series with p th order autoregressive correlation structure has been proposed. It utilizes standard uncorrelated linear autoregressive residuals, and the crosscorrelation function of these residuals and their squares. The behaviour of this crosscorrelation function has been utilized for two rather different types of nonlinear model: random coefficient

autoregression and multiplicative autoregression: the behaviour has been shown to allow discrimination between models in the same class with the same marginal and autocorrelation structures.

The residuals crosscorrelation function provides a partial analysis of third order information in the time series; it does not attempt to capture all third order information, which is the aim of such techniques as bispectral analysis and which will often be intractable with nonlinear models. Being based on standard linear residuals, the analysis extends rather than replaces conventional residual analysis.

Developments of the analysis which focus on the directionality implicit in many time series are being investigated; reversed residuals assume a reversed directionality and allow exploration of the consequences of such an assumption.

Acknowledgements The research of P.A.W. Lewis was supported by the Office of Naval Research under Grant NR-042. The graphs were produced by an experimental APL package from IBM which the Naval Postgraduate School is using under an agreement with the IBM Research Centre, Yorktown Heights, N.Y.

6. REFERENCES

- Andell, J. (1976) Autoregressive series with random parameters, *Math. Operationsforsch. u. Statist.*, 7, 735-741.
- Granger, C.W.J. and Andersen, A.P. (1978) *An Introduction to Bilinear Time Series Models*. Vandenhoeck and Ruprecht, Gottingen.
- Jacobs, P.A. and Lewis P.A.W. (1983) Stationary discrete autoregressive-moving average time series generated by mixtures, *J. Time Series Analysis*, 4, 18-36.
- Jones, D. (1978) Non-linear autoregressive processes. *Proc. Roy. Soc. London A*, 360, 71-95.
- Lawrance, A.J. and Lewis, P.A.W. (1981) A new autoregressive time series model in exponential variables. *Adv. Appl. Prob.*, 13, 86-845.
- Lawrance, A.J. and Lewis, P.A.W. (1984) Modelling and residual analysis of nonlinear autoregressive time series in exponential variables. *Preprint*.
- Lewis, P.A.W. (1980) Simple models for positive-valued and discrete-valued time series with ARMA correlation structure. In *Multivariate Analysis-V* (Ed. P.R. Krishnaiah), 151-166. North Holland, Amsterdam.
- Lewis, P.A.W. (1981) Simple multivariate time series for simulations of complex systems. In *Proc. Winter Simulation Conference* (Eds. T. I. Oren, C. M. Delfosse and C. M. Shub), 389-390. IEEE Press, New York.
- McKenzie, E. (1982) Product autoregression: a time series characterization of the gamma distribution. *J. Appl. Prob.*, 19, 463-468.
- Nicholls, D.F. and Quinn, B.G. (1982) *Random Coefficient Autoregressive Models: An Introduction*. Springer-Verlag, Heidelberg.
- Rosenblatt, M.R. (1979) Linearity and nonlinearity in time series: prediction. *Proc. 42nd Session I.S.I. (Manila)*, Vol. 1, 423-434.

Rosenblatt, M.R. (1980) Linear processes and bispectra. *J. Appl. Prob.* 17, 265-270.

Tjøstheim, D., and Paulsen, J. (1983) Bias of some commonly-used time series estimates. *Biometrika*, 70, 389-400.

Vervaat, W. (1979) On a stochastic difference equation and a representation of non-negative infinitely divisible random variables. *Adv. Appl. Prob.* 11, 750-783.

Weiss, G. (1975) Time reversibility of linear stochastic processes. *J. Appl. Prob.*, 12, 831-836.

DISTRIBUTION LIST

	NO. OF COPIES
Defense Technical Information Center Cameron Station Alexandria, VA 22314	2
Library Code 0142 Naval Postgraduate School Monterey, CA 93943	2
Research Administration Code 012A Naval Postgraduate School Monterey, CA 93943	1
Library Code 55 Naval Postgraduate School Monterey, CA 93943	1
Professor P.A.W. Lewis Code 55Lw Naval Postgraduate School Monterey, California 93943	150

U215870

DUDLEY KNOX LIBRARY - RESEARCH REPORTS



5 6853 01060232 9

~~U215870~~