Journal of Official Statistics
Vol. 4, No. 2, 1988, pp. 113–124

On Autoregressive Model Identification

Ette Harrison Etuk[1]

Abstract: Since Cleveland (1972) introduced the inverse autocorrelation function, it has been recognized as a competitor to the partial autocorrelation function as a time series model identification tool. By using simulated and real data, we have demonstrated that neither of these is consistently more powerful than the other for identification of autoregressive (AR) models. However when the underlying AR process is of full order, the partial autocorrelation function invariably is the superior. But when a subset order AR model generates the data, the inverse autocorrelation function is generally more informative. On the whole the partial autocorrelation function exhibits better performance. For instance, in two of the three cases of real series used it clearly outperforms the inverse autocorrelation function.

Key words: Autoregressive model identification: partial autocorrelation function; inverse autocorrelation function.

1. Introduction

Since Yule (1927) introduced autoregressive modelling, it has played a significant role in the analysis of data recorded sequentially in time or space. The basic component of all time series models is the white noise process defined as a sequence $\{\varepsilon_t\}$ of uncorrelated zero mean and constant variance random variables. A mean corrected time series $\{X_t\}$ is said to be an autoregressive process of order p (designated AR(p)) if it is a stationary solution of the following difference equation.

$$X_t + \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \ldots + \alpha_p X_{t-p} = \varepsilon_t \,, \qquad (1.1)$$

where the α_i 's are constants. For stationarity, the characteristic polynomial $\sum_{i=0}^{p} \alpha_i z^i$, $\alpha_0 = 1$, must have zeros outside the unit circle.

Any stationary time series can be expressed as an infinite-order autoregression. Autoregressive modelling therefore involves the approximation of the autoregression by a finite order one of the form (1.1). Thus the order, in essence, constitutes an additional parameter to be estimated. This order determination represents a major obstacle since underfitting increases the residual variance while overparametrization decreases the reliability of the model. Model identification involves more than order determination; the relative contribution of the parameters to the model structure should also be estimated.

A diagnostic aid for autoregression order determination is the partial autocorrelation function (PACF), advocated for this purpose by Box and Jenkins (1976). It is known to cut off at lag p for (1.1). Hence its estimate

[1] Department of Mathematics and Statistics, University of Calabar, Calabar, Nigeria.

indicates a possibility of an AR(p) component of the underlying model, if it fails to be significant after lag p. A similar tool, the inverse autocorrelation function (IACF), has been introduced by Cleveland (1972).

The question of which is better in AR model identification has engaged the attention of many researchers since the introduction of the IACF. Cleveland (1972), Chatfield (1979), Hipel, McLeod, and Lennox (1977), and Oyetunji (1985) to mention a few, believe that the IACF is the better. However, McLeod, Hipel, and Lennox (1977) observe that in certain time series applications, the PACF is comparatively better in specifying the model. Abraham and Ledolter (1984) have demonstrated that the PACF is more powerful in identifying purely AR processes.

It is the objective of this work to further investigate the relative merits of the two methods. Like Abraham and Ledolter (1984), we simulate AR models and observe the frequency with which the functions detect certain features of the model. We observe the effect of the nature of roots of the characteristic equation of the underlying model, the sample size, and the distance of the model from the boundary of stationarity on their comparative performance. Another factor of variation of interest is whether the model is full order or subset order. Inspired by the work of Bhansali (1983), we use the intuitively appealing autoregressive estimates of the IACF, having generated AR models. Since the estimation of the functions, especially that of the PACF, is inextricably tied to that of the model, the nature of their estimates depends on the mode of model estimation. Our Monte Carlo study uses the Yule-Walker approach to autoregression estimation. However in Section 6 we use also Burg's estimates of the IACF and PACF, to illustrate our results on some real series.

2. Partial Autocorrelation Function (PACF)

Suppose we write (1.1) more specifically as

$$X_t + \alpha_{p1}X_{t-1} + \alpha_{p2}X_{t-2} + \ldots + \alpha_{pp}X_{t-p} = \varepsilon_t \quad (2.1)$$

where α_{ij} is the jth coefficient of an AR(i) model. The last coefficient α_{pp} is called the partial autocorrelation of lag p.

The sequence $\{\alpha_{pp}\}$, regarded as a function of p, is the PACF of $\{X_t\}$.

Let $\mu = E(X_t)$, $\gamma_k = E[(X_t-\mu)(X_{t-k}-\mu)] = \gamma_{-k}$, and $\varrho_k = \frac{\gamma_k}{\gamma_0}$, $k = 0, \pm 1, \pm 2, \ldots$ be the mean, autocovariance of lag k, and autocorrelation of lag k, respectively, of $\{X_t\}$. They can be estimated by

$$\bar{X} = \frac{1}{N}\sum_{i=1}^{N} X_t, \; C_k = \frac{1}{N}\sum_{t=k+1}^{N} (X_t-\bar{X})(X_{t-k}-\bar{X}),$$

$$\text{and } r_k = \frac{C_k}{C_0}, \; k = 0, 1, 2, 3, \ldots$$

respectively, from a realization, $X_1, X_2, \ldots, X_N$, of $\{X_t\}$.

Contemporary techniques of AR modelling use automatic order determination criteria whose minimum within a specified order range gives the optimum order.
Criteria like

$$\text{FPE}(p) = \hat{\sigma}_p^2\left(1 + \frac{p}{N}\right), p = 0,1,2,\ldots,$$
(Akaike (1969))

$$\text{FPE}\alpha(p) = \left(1 + \frac{\alpha p}{N}\right)\left(1 - \frac{p}{N}\right)^{-1}\hat{\sigma}_p^2,$$

$$\alpha > 0, p = 0,1,2,\ldots,$$
(Bhansali and Downham (1977))

$$\text{AIC}(p) = N\ln\tilde{\sigma}_p^2 + 2p, p = 0,1,2,\ldots,$$
(Akaike (1977))

$$\text{BIC}(p) = N\ln\tilde{\sigma}_p^2 - (N-p)\ln\left(1-\frac{p}{N}\right) + p\ln N$$

$$+ pln\{p^{-1}\left(\frac{C_o}{\tilde{\sigma}_p^2} - 1\right)\}, p = 0,1,2,\ldots,$$
(Akaike (1977))

$$\mathrm{SIC}(p) = Nln\tilde{\sigma}_p^2 + plnN, p = 0,1,2,\ldots,$$
(Schwarz (1978))

$$\mathrm{CAT}_2(p) = \begin{cases} \frac{1}{N}\left(\sum_{j=1}^{p} \frac{1}{\hat{\sigma}_j^2}\right) - \frac{1}{\hat{\sigma}_p^2}, p = 1,2,\ldots \\ -\left(1 + \frac{1}{N}\right), p = 0, \end{cases}$$
(Parzen (1977))

$$\mathrm{CAT}_3(p) = \frac{1}{N}\left(\sum_{j=o}^{p} \frac{1}{\hat{\sigma}_j^2}\right) - \frac{1}{\hat{\sigma}_p^2},$$
$p = 0,1,2,\ldots,$ (Tong (1977))

$$S_N(p) = (N + 2p)\,\hat{\sigma}_p^2, p = 0,1,2,\ldots,$$
(Shibata (1980))

$$\phi(p) = ln\tilde{\sigma}_p^2 + N^{-1}2pc\,lnln\,N, c>1, p = 0,1,2,\ldots$$
(Hannan and Quinn (1979))

are to mention a few, where $\hat{\sigma}_p^2$ and $\tilde{\sigma}_p^2$ are the least squares and maximum likelihood estimates of the residual variance, respectively.

An approximate least squares estimate (Yule-Walker estimate) of (2.1) can be obtained by the use of the recursive formula:

$$\phi_{k+1} = \hat{\alpha}_{k+1,k+1} = \frac{r_{k+1} - \sum_{j=1}^{k} \hat{\alpha}_{kj}\, r_{k+1-j}}{1 - \sum_{j=1}^{k} \hat{\alpha}_{k,j}\, r_j},$$

$$\hat{\alpha}_{k+1,j} = \hat{\alpha}_{k,j} - \phi_{k+1}\,\hat{\alpha}_{k,k-j+1}, j = 1, 2,\ldots, k.$$

Another method of autoregression estimation is the maximum entropy method proposed by Burg and formalized by Andersen (1974).

Under the hypothesis of an AR (p) model, it is well known that the estimated partial autocorrelations of order $p+1$ and higher are approximately independently distributed with $E(\hat{\alpha}_{kk}) = 0$, and

$$\mathrm{var}(\hat{\alpha}_{kk}) \cong \frac{1}{N}, k \geqslant p + 1 \qquad (2.2)$$

as $N \to \infty$.

3. Inverse Autocorrelation Function (IACF)

For a stationary stochastic process $\{X_t\}$ with spectral density function $f(\omega)$, Cleveland (1972) has defined the inverse autocovariance of lag k as

$$\gamma i_k = \int_{-\pi}^{\pi} e^{i\omega k} fi(\omega)\, d\omega = \gamma i_{-k}, k = 0, 1, 2,\ldots$$

where $fi(\omega) = \frac{1}{f(\omega)}$. The inverse autocorrelation function is defined as

$$\varrho i_k = \frac{\gamma i_k}{\gamma i_o}, k = 0, 1, 2, \ldots .$$

An equivalent but time-domain definition of γi_k is given by Chatfield (1979).

Using the duality between AR and moving average processes, it has been shown (Cleveland (1972); Chatfield (1979)) that for the AR (p) process (1.1), ϱi_k is given by

$$\varrho i_k = \begin{cases} \left(\alpha_k + \sum_{j=1}^{p-k} \alpha_j\,\alpha_{j+k}\right) / \left(1 + \sum_{j=1}^{p} \alpha_j^2\right) \\ k = \pm 1,\ldots, \pm p \\ 0, |k| > p \end{cases} \qquad (3.1)$$

and

$$\gamma i_o = \left(1 + \sum_{j=1}^{p} \alpha_j^2\right) \sigma_\varepsilon^2, \qquad (3.2)$$

where

$$\sigma_\varepsilon^2 = \mathrm{var}(\varepsilon_t).$$

Cleveland (1972) has suggested two methods of estimating ϱi_k stemming from two methods of spectrum estimation, viz., the autoregressive and window methods. The autoregressive method, which is quite popular, consists of approximating the process by an AR model of sufficiently high order p for a good fit, estimating the parameters of the model and using the estimates in equations (3.1) and (3.2) to obtain an estimate of IACF.

The subjectivity introduced in choosing p poses problems. Hipel, McLeod, and Lennox (1977) suggest trying four values of p between 10 and 40 (where $p < N/4$) and choosing the value of p which gives the most representative graph of the resultant inverse autocorrelation estimate ri_k against lag k. Chatfield (1979) advises against the use of automatic criteria like the AIC, BIC, etc. since for determining p optimal parametric parsimony is not of interest. He however warns that p must not be so large as to make the variance of the estimates too high; the choice of p must be such that for high lags, the estimates of the inverse autocorrelations approach zero. He suggests some form of trial and error until the foregoing criteria are met. Hosking (1980) suggests that p should vary with the sample size N. In their work Abraham and Ledolter (1984) use the values of 5 and 10 for p since their simulated models had lower orders and report the results for $p = 10$ after observing the results for the two values to be similar.

The window method involves smoothing the periodogram. A good exposition on this approach is given by Priestley (1981). Other methods of estimation of IACF include those proposed by McClave (1978) and Chatfield (1979).

Hosking (1980) has shown that for a stationary time series, $N^{1/2}(ri_k - \varrho i_k)$ is normal with mean zero and covariances given by

$$\text{cov}(ri_k, ri_{k+t}) \simeq \frac{1}{N}\sum_{v=-\infty}^{\infty}\{\varrho i_v \varrho i_{v-t} + \varrho i_{v-k-t}\varrho i_{v+k} - 2\varrho i_k \varrho i_{v-k-t} - 2\varrho i_{k+t}\varrho i_v \varrho i_{v-k}\} \text{ as } N \to \infty.$$

Thus for an AR (p) process, asymptotically,

$$\text{var}(ri_k) \simeq \frac{1}{N}\{1+2\sum_{v=1}^{p}\varrho i_v^2\}, k > p. \quad (3.3)$$

4. Simulated Data

Eight AR(2) series with (α_1, α_2) equal to (–1.68, 0.70), (0.00,–0.78), (–0.66, 0.10), (0.00,–0.15), (–1.08, 0.77), (0.00, 0.89), (–0.46, .0.08) and (0.00, 0.10) are simulated independently twenty times each. In the sequel, we shall refer to them as series I trough VIII, respectively. Table 1 shows which of them have characteristic equations with real or complex roots and those close to or far from the boundary of stationarity.

Table 1. Categories of simulated series

	Real roots	Complex roots
Close	I, II	V, VI
Far	III, IV	VII, VIII

The motivation for this choice of models is the need to cover the parameter surface. We also use sample sizes of 50, 150, and 250 for each series.

The white noise process for each simulation is a sequence of pseudorandom numbers generated using the RAN function of the FORTRAN 77 language. The sequence is made approximately standard normal.

5. Results

We, as did Abraham and Ledolter (1984), use the critical region $\pm 2\ N^{-1/2}$ for assessing the performance of both criteria. We also use $p=5$ and $p=10$ and compare the results.

Like that of Abraham and Ledolter (1984) our observation is that the PACF detects the significance of the lag two coefficient more than the IACF for all the models. This is especially true for full-order models, those with characteristic equations having real roots and those close to the boundary of stationarity. Both the performance of each criterion and the relative performance of the IACF increase with sample size. IACF also performs better for $p = 5$ than for $p = 10$. We have observed also that both functions are good at detecting zero third-order components.

To examine their relative power in order determination, we note the frequency with which each has a significant second-lag value and a non-significant third-lag value. Coefficients at higher lags were observed to be non-significant. Table 2 shows our results. Clearly, the PACF excels for all the models and for both lags. The larger the sample size, the better the comparative performance of the IACF.

Comparing their potential for correctly detecting α_1, we find that while the IACF excels for subset-order models, the PACF excels for full-order models. Our result for the subset-order models supports the claims of Cleveland (1972) and Chatfield (1979), to mention a few. The details of the results are in Table 3. The IACF is more powerful especially when the underlying subset-order

Table 2. Frequencies of the events $\{(\hat{\alpha}_{11}, \hat{\alpha}_{22}, \hat{\alpha}_{33}): |\hat{\alpha}_{22}| > 2N^{-1/2}, |\hat{\alpha}_{33}| \leq 2N^{-1/2}\}$ *and* $\{ri_1, ri_2, ri_3\} : |ri_2| > 2N^{-1/2}, |ri_3| \leq 2N^{-1/2}\}$

		$N=50$		$N=150$		$N=250$		Total	
Series I	IACF	0	0*	1	1*	5	5*	6	6*
	PACF	12		13		10		5	
Series II	IACF	18	20*	18	20*	15	20*	51	60*
	PACF	19		20		18		57	
Series III	IACF	0	0*	0	0*	4	5*	4	5*
	PACF	2		2		11		15	
Series IV	IACF	3	0*	4	1*	10	11*	17	12*
	PACF	3		3		11		17	
Series V	IACF	10	5*	20	20*	20	19*	50	44*
	PACF	19		19		20		58	
Series VI	IACF	20	20*	16	20*	19	19*	55	59*
	PACF	19		20		20		59	
Series VII	IACF	2	1*	5	4*	5	4*	12	9*
	PACF	3		7		5		15	
Series VIII	IACF	0	0*	7	8*	7	5*	14	13*
	PACF	1		6		7		14	
Total		53	46*	71	74*	85	88*		
		78		90		99			

* For $p = 10$.

Table 3. Frequencies of the events $\{(\hat{\alpha}_{11}, \hat{\alpha}_{22}, \hat{\alpha}_{33}) : |\hat{\alpha}_{11}| > 2N^{-1/2}\}$ *and* $\{(ri_1, ri_2, ri_3) : |ri_1| > 2N^{-1/2}\}$

		$N = 50$		$N = 150$		$N = 250$		Total	
Series I	IACF	20	20*	20	20*	20	20*	60	60*
	PACF	20		20		20		60	
Series II	IACF	0	1*	0	0*	1	0*	1	1*
	PACF	9		8		9		26	
Series III	IACF	20	20*	20	20*	20	20*	60	60*
	PACF	20		20		20		60	
Series IV	IACF	2	0*	0	0*	0	0*	2	0*
	PACF	3		2		1		6	
Series V	IACF	20	19*	20	20*	20	20*	60	59*
	PACF	20		20		20		60	
Series VI	IACF	0	0*	0	0*	0	0*	0	0*
	PACF	0		0		0		0	
Series VII	IACF	8	8*	20	20*	20	28*	48	56*
	PACF	12		20		20		52	
Series VIII	IACF	2	0*	0	0*	0	0*	2	0*
	PACF	3		0		0		3	
Total		72	68*	80	80*	81	88*		
		87		90		90			

* For $p = 10$.

model is close to the boundary of stationarity or has a characteristic equation with real roots and for $p = 10$.

Assessing the overall relative performance of the two functions in correctly determining the subset-order models, we find as evident

Table 4. Frequencies of the events $\{(\hat{\alpha}_{11}, \hat{\alpha}_{22}, \hat{\alpha}_{33}): |\hat{\alpha}_{11}| \leq 2N^{-1/2}, |\hat{\alpha}_{22}| > 2N^{-1/2}, |\hat{\alpha}_{33}| \leq 2N^{-1/2}\}$ *and* $\{(ri_1, ri_2, ri_3): |ri_1| \leq 2N^{-1/2}, |ri_2| > 2N^{-1/2}, |ri_3| \leq 2N^{-1/2}\}$ for the subset-order models

		$N = 50$		$N = 150$		$N = 250$		Total	
Series II	IACF	18	18*	18	20*	15	19*	51	57*
	PACF	10		12		10		32	
Series IV	IACF	3	0*	4	1*	10	12*	17	13*
	PACF	3		3		11		17	
Series VI	IACF	20	20*	16	20*	19	19*	55	59*
	PACF	19		20		20		59	
Series VIII	IACF	0	0*	7	8*	7	5*	14	13*
	PACF	1		6		7		14	
Total		41	38*	45	49*	51	55*		
		33		41		48			

* For $p = 10$.

from Table 4 that the IACF is the better for all the models and for all the sample sizes.

Abraham and Ledolter (1984) base their use of $\pm 2N^{-1/2}$ for the assessment of both functions on the hypothesis of independence of observations. They also see the rationale for using $\pm 2N^{-1/2}\{1 + 2(ri_1^2 + \ldots + ri_{k-1}^2)\}^{1/2}$ for testing the significance of ri_k (on the basis of (3.3)), but rightly argue that its use would worsen its performance for determining the order of their models, all being full order. However, when the underlying subset model has some non-significant ri_j's,the IACF is likely to detect the significant non-zero lag coefficients more often than otherwise. In particular, for our subset-order models, the relative power of the IACF is higher with the non-null critical region than with $\pm 2N^{-1/2}$. The integral $\pm 2N^{-1/2}\{1+2(ri_1^2+\ldots+ri_{k-1}^2)\}^{1/2}$ is not adequately wider than the null critical region. An effect of this is that IACF is not much worse than the PACF in detecting the significance of the lag two coefficients. In addition, the IACF correctly suggests a zero value for the lag three coefficient more often.

It even supersedes the PACF in this sense, especially for $p = 10$ (see Table 5).

Table 5. Frequencies of the events $\{(\hat{\alpha}_{11}, \hat{\alpha}_{22}, \hat{\alpha}_{33}) : |\hat{\alpha}_{11}| \leq 2N^{-1/2}, |\hat{\alpha}_{22}| > 2N^{-1/2}, |\hat{\alpha}_{33}| \leq 2N^{-1/2}\}$ *and* $\{(ri_1, ri_2, ri_3) : |ri_1| \leq 2N^{-1/2}, |ri_2| > 2N^{-1/2}[1 + 2ri_1^2]^{1/2}, |ri_3| \leq 2N^{-1/2}[1 + 2(ri_1^2 + ri_2^2)]^{1/2}\}$ *for the following subset-order models*

		$N = 50$		$N = 150$		$N = 250$		Total	
Series II	IACF	19	20*	20*	20*	17	20*	56	60*
	PACF	10		12		10		32	
Series IV	IACF	2	0*	4	2*	11	12*	17	14*
	PACF	3		3		11		17	
Series VI	IACF	20	20*	20	20*	20	20*	60	60*
	PACF	19		20		20		59	
Series VIII	IACF	0	1*	7	8*	7	5*	14	14*
	PACF	1		6		7		14	
Total		41	41*	51	50*	55	57*		
		33		41		48			

* For $p = 10$.

6. Practical Examples

For the Yule-Walker (Y-W) approach, full order and subset order AR models were fitted using the above outlined order determination criteria. For Burg's maximum entropy (ME) approach we fitted only full-order models. For the criterion ϕ we used $c = 1.50$ and for FPEα $\alpha = 4$. Diagnostic checks were made by the use of Box-Jenkins (1976) portmanteau test statistic, which we denote by R.

6.1. Series A (Box and Jenkins (1976, pp. 525))

For the Y-W approach, BIC, ϕ, SIC, S and FPE4 recommend the full order AR(2)

$$X_t - 0.427X_{t-1} - 0.252X_{t-2} = \varepsilon_t, \qquad (6.1)$$

$$\hat{\sigma}^2 = 0.1002$$

with $R = 100.03$. The Box-Jenkins (1976) portmanteau test rejects the model. AIC, FPE, CAT_2, and CAT_3 choose the AR (7):

$$X_t - 0.373X_{t-1} - 0.197X_{t-2} - 0.020X_{t-3} - 0.014X_{t-4} + 0.015X_{t-5} - 0.062X_{t-6} - 0.156X_{t-7} = \varepsilon_t, \quad (6.2)$$

$$\hat{\sigma}^2 = 0.0950$$

which, with $R = 22.44$, is not discredited by the portmanteau test.

For the ME approach, BIC, ϕ, SIC, and FPE4 choose an AR(2) very close to (6.1) and for which R is also significant. FPE, AIC, S, CAT_2, and CAT_3, however, choose an AR(7) very close to (6.2) and which is also recommended by the R-test.

Applying the subset AR modelling algorithm of Haggan and Oyetunji (1984), for a maximum lag of 15 gave the following subset models: BIC, S, SIC, ϕ, and FPE4 recommend the model:

$$X_t - 0.381X_{t-1} - 0.216X_{t-2} - 0.188X_{t-7} = \varepsilon_t, \quad (6.3)$$

$$\hat{\sigma}^2 = 0.0955,$$

$$R = 24.31.$$

CAT_2 and CAT_3 pick (6.1). AIC chooses:

$$X_t - 0.388X_{t-1} - 0.220X_{t-2} - 0.174X_{t-7} - 0.126X_{t-14} + 0.122X_{t-15} = \varepsilon_t, \hat{\sigma}^2 = 0.0934 \quad (6.4)$$

$$R = 20.15.$$

FPE selects the full order AR(15). Both (6.3) and (6.4) were found adequate by the R-test. To confirm the result of the R-test, Etuk (1987) has shown that the models (6.2), (6.3), and (6.4) have spectra which closely agree with an estimate of the raw spectrum; (6.1) does not. The model (6.3) is the most adequate used here since it is the most parsimonious amongst the adequate models (6.2), (6.3), and (6.4). Cleveland (1972) has also suggested the model

$$X_t + \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \alpha_7 X_{t-7} = \varepsilon_t.$$

Table 6. A comparison of the PACF and IACF for series A

Lag	Y-W PACF	Y-W AR(10) IACF	Burg's AR(10) IACF	Burg's PACF
1	0.57	−0.31	−0.30	0.57
2	0.25	−0.16	−0.15	0.25
3	0.07	−0.02	−0.02	0.08
4	0.07	−0.01	−0.02	0.09
5	0.07	0.01	0.02	0.07
6	0.12	−0.06	−0.07	0.14
7	0.15	−0.14	−0.16	0.19
8	−0.03	0.02	0.03	−0.04
9	0.01	−0.02	−0.02	0.01
10	−0.02	0.01	0.01	−0.01

With $2/\sqrt{N} = 0.14$, we see that the IACF not only correctly recommends an order of 7, but also better suggests zero parameter values for lags 3, 4, 5, and 6, especially for lag 6.

6.2. *Canadian lynx numbers (1821 – 1934) (Campbell and Walker (1977, pp. 430))*

The well-analyzed logarithm transformation has been used. For the Y-W method for full-order modelling, BIC, ϕ, and SIC pick the order 2 which though recommended by the *R*-test has a spectrum which does not tally with the estimated raw spectrum. However, the AR(11) selected by FPE, AIC, S, CAT_2, CAT_3, and FPE4 is found adequate by both diagnostic checking criteria. For the ME technique, BIC and ϕ pick an order 2 also found to be inadequate by the spectrum test, although its *R* value is nonsignificant. SIC, S, CAT_2, CAT_3, and FPE4 pick the AR(11) found to be adequate by both tests. FPE and AIC pick the likely overparametrized AR(12). For the full-order modelling order 11 has been found to be the best (Etuk (1987)).

For subset-order modelling, BIC, ϕ, SIC, and FPE4 recommend an order 11 with significant lags 1, 2, 4, 10, and 11. CAT_2 and CAT_3 select the model with significant lags 1, 2, and 4. FPE chooses the full order AR(15). Etuk (1987) has shown that the subset model with lags 1, 2, 4, 10, and 11 is the best model.

Table 7 gives the PACF and IACF. With $N=114$, $2/\sqrt{N} = 0.19$. Therefore, the PACF suggests an order of 4, 7, or 11. The IACF, however, suggests an order of 1. The PACF is significant at lags 1, 2, 4, 7, 10, and 11 which tallies better with the identified model. Evidence here is therefore in favour of the PACF.

Table 7. A comparison of the PACF with the IACF for the log transform of the lynx data

Lag	Y–W PACF	Y–W AR(15) IACF	Burg's AR(15) IACF	Burg's PACF
1	0.79	−0.40	−0.39	0.79
2	−0.72	0.17	0.17	−0.74
3	−0.14	−0.07	−0.09	−0.12
4	−0.21	0.10	0.12	−0.21
5	0.12	−0.05	−0.06	0.14
6	0.08	0.05	0.06	0.07
7	0.21	−0.03	−0.03	0.23
8	0.12	0.03	0.03	0.13
9	0.10	−0.06	−0.06	0.12
10	−0.19	−0.04	−0.05	−0.22
11	−0.31	0.06	0.07	−0.35
12	−0.10	0.08	0.06	−0.13
13	0.10	−0.04	−0.00	0.05
14	−0.04	−0.00	−0.03	−0.01
15	−0.02	0.01	0.01	−0.04

6.3. *Wolfer's sunspot series (1700 – 1955)*

Data on sunspots are available from 1700 onwards (see Waldmeier (1961)). We used 256 values from 1700 to 1955.

For the Y-W approach for full-order model selection FPE, AIC, and CAT_3 recommend an AR order of 9; FPE4, ϕ, and S recommend 8; SIC recommends 3, CAT_2 zero, and BIC 2. For the ME technique, FPE and AIC pick an AR(18); BIC, SIC, and ϕ pick AR(8); CAT_3 and FPE4 choose the AR(9); CAT_2 selects the AR(0) and S the AR(10).

For subset modelling, BIC, SIC, ϕ, and FPE4 recommend a model with lags 1, 2, and 9. AIC pick the lags 1, 2, 3, 4, 5, and 9. CAT_3 choose the lags 1, 2, and 3; S choose the lags 1, 2, 3, and 9. The BIC subset model has been shown to be the most adequate model (Etuk (1987)).

The critical value for both functions is 0.125. With PACF significant lags are 1, 2, 3, 6, 7, 8, and 18, so that the suggested orders are 3, 8, and 18. The IACF indicates an order of 2 (see Table 8).

We therefore observe an agreement of the PACF with BIC, SIC, and ϕ in choosing an order of 8, with AIC in choosing an order 18. The IACF agrees with BIC in the choice of order 2. The two diagnostic checking methods used here do not discredit an order of 8. The order of 18 is therefore an overestimation. The orders 2 and 3 are too low for the model.

Thus we find that PACF is the better model identifier.

Table 8. Comparison of IACF and PACF for sunspot series

Lag	Y-W PACF	Y-W AR(15) IACF	Burg's AR(20) IACF	Burg's PACF
1	0.81	0.41	0.41	0.82
2	−0.66	0.15	0.14	−0.67
3	−0.15	0.07	0.07	−0.14
4	0.04	−0.08	−0.09	0.05
5	−0.07	0.08	0.07	−0.07
6	0.17	−0.05	−0.03	0.18
7	0.14	0.03	0.01	0.19
8	0.22	−0.04	−0.04	0.23
9	0.10	−0.05	−0.05	0.12
10	0.01	0.04	0.03	0.03
11	0.07	−0.06	−0.06	0.07
12	−0.07	0.02	0.02	−0.07
13	0.00	0.04	0.05	0.00
14	0.06	−0.05	−0.06	0.07
15	−0.09	0.03	0.04	−0.10
16	−0.06	−0.01	−0.00	−0.07
17	−0.10	−0.02	−0.02	−0.11
18	−0.13	0.05	0.06	−0.16
19	−0.01	−0.02	−0.02	−0.00
20	0.03	0.01	0.01	−0.02

7. Conclusion

Evidence here is not conclusive as to the better criterion. However, we have seen that the PACF is definitely the better in identifying non-zero lag coefficients and determining order. Invariably, the PACF outperforms the IACF for full-order AR models. The reason for this is that though, as observed by Abraham and Ledolter (1984), chances are that $|\hat{\alpha}_{jj}| > |\hat{\gamma}_{jj}| \; \forall j = 1, 2, \ldots$, the variance of the PACF is the smaller, for any given lag. For the same reason, the IACF more often correctly detects zero coefficients, especially where they are intervening.

We have observed that the PACF outperforms IACF for two of the three real series we used. Even for the lynx data for which a subset AR model is suggestive, PACF fares the better. However, our Monte Carlo study shows that the IACF is generally the better for subset-order modelling.

We have also noticed a tendency of overestimation with the PACF and of underestimation with the IACF. Exclusive preference of one to the other may therefore not be advisable.

The IACF and PACF of an AR model is the autocorrelation function (ACF) and inverse partial autocorrelation function (IPACF) respectively of the inverse moving average (MA) model. Comparision of the IACF and PACF in AR model identification is tantamount to that of the ACF and IPACF in MA modelling. Therefore, we suggest the application of all four functions: ACF, IACF, PACF, and IPACF in autoregressive moving average (ARMA) modelling.

8. References

Abraham, B. and Ledolter, J. (1984): A Note on Inverse Autocorrelations. Biometrika, 71, pp. 609–614.

Akaike, H. (1969): Fitting Autoregressive Models for Prediction. Annals of the Institute of Statistical Mathematics, 21, pp. 243–247.

Akaike, H. (1977). On Entropy Maximization Principle. Proceedings of the Symposium on Applied Statistics. (P.R. Krishnaiah, ed.) Amsterdam, North – Holland.

Andersen, N. (1974): On the Calculation of Filter Coefficients for Maximum Entropy Spectral Analysis. Geophysics, 39, (1), pp. 69–72.

Bhansali, R.J. (1975): Fitting Time Series Models in the Frequency Domain. Bulletin of the International Statistical Institute, 46, Book 3, pp. 98–104.

Bhansali, R.J. (1983): A Simulation Study of Autoregressive and Window Estimators of the Inverse Correlation Function. Applied Statistics, 32, pp. 141–149.

Bhansali, R.J. and Downham, D.Y. (1977): Some Properties of the Order of an Autoregressive Model Selected by a Generalization of Akaike's FPE criterion. Biometrika, 64, pp. 547–551.

Box, G.E.P. and Jenkins, G.M. (1976): Time Series Analysis Forecasting and Control. Holden-Day, San Francisco.

Campbell, M.J. and Walker, A.M. (1977). A Survey of Statistical Work on the Mackenzie River Series of Annual Canadian Lynx Trappings for the Years 1821–1934 and a New Analysis. Journal of the Royal Statistical Society, Series A, 140, pp. 411–431.

Chatfield, C. (1979): Inverse Autocorrelations. Journal of the Royal Statistical Society, Series A, 142, pp. 363–377.

Cleveland, W.S. (1972): The Inverse Autocorrelations of a Time Series and Their Applications. Technometrics, 14, pp. 277–297.

Etuk, E.H. (1987): On the Selection of Autoregressive Moving Average Models. Ph.D. Thesis. Department of Statistics, University of Ibadan, Nigeria.

Haggan V. and Oyetunji, O.B. (1984): On the Selection of Subset Autoregressive Time Series Models. Journal of Time Series Analysis, (2), pp. 103–113.

Hannan, E.J. and Quinn, B.G. (1979): The Determination of the Order of an Autoregression. Journal of the Royal Statistical Society, Series B, 41, pp. 190–195.

Hipel, K.W., McLeod, A.I., and Lennox, W.C. (1977): Advances in Box-Jenkins Modelling 1. Water Resources Research, 13, pp. 567–575.

Hosking, J.R.M. (1980): The Asymptotic Distribution of the Sample Inverse Autocorrelation of an Autoregressive- moving

Average Process. Biometrika, 67, pp. 223–226.

McClave, J. (1978): Estimating the Order of Moving Average Models: The Max χ^2 Method. Communications in Statistics, A (73), pp. 259–276.

McLeod, A.I., Hipel, K.W., and Lennox, W.C. (1977): Advances in Box-Jenkins Modelling 2. Water Resources Research, 13, pp. 577–586.

Oyetunji, O.B. (1985): Inverse Autocorrelations and Moving Average Time Series Modelling. Journal of Official Statistics, 1, pp. 315–322.

Parzen, E. (1977): Non-parametric Statistical Data Science: A Unified Approach Based on Density Estimation and Testing for "White Noise". Statistical Science Division Report No. 47. State University of New York at Buffalo.

Priestley, M.B. (1981): Spectral Analysis and Time Series. Academic Press, London.

Schwarz, G. (1978): Estimating the Dimension of a Model. Annals of Statistics, 6, pp. 461–464.

Shibata, R. (1980): Asymptotically Efficient Selection of the Order of the Model for Estimating Parameters of a Linear process. Annals of Statistics, 8, pp. 147–164.

Tong, H. (1977): On Some Alternative Approaches to Autoregressive Order Determination. Technical Report 80, Department of Mathematics, University of Manchester, Institute of Science and Technique, U.K.

Waldmeier, M. (1961): The Sun-Spot Activity in the Years 1610–1960. Schulthess, Zurich.

Yule, G.U. (1927): On a Method of Investigating the Periodicities in Disturbed Series, with Special Reference to Wolfer's Sunspot Numbers. Philosophical Transactions of the Royal Society London (A), pp. 226–267.

Received September 1987
Revised April 1988