Jour. Ind. Soc. Ag. Statistics
45 (1), 1993 ; 144-162

Extending AR Models for Seasonal Series : A Case Study

A. Sarkar and B. Kartikeyan
IIT, Kharagpur
(Received : July, 1985)

Summary

A subset AR model has been considered along with seasonal multiplicative model for forecasting a particular seasonal series. The forecasting performance of both models are compared. For estimating parameters of the models working procedure of Marquardt's algorithm has been outlined for easy computer coding. The amplitudes of the Fourier frequencies within periods of seasonal variation are obtained and interpreted. Subset AR model for seasonal series is suggested when the task of identifying a multiplicative model for such series is not easy.
Key words: Subject AR models; Multiplicative ARIMA models; Seasonal index; Marquadrt's algorithm; AIC.

Introduction

An attempt is made to forecast monthly quantity arrivals of a particular cotton variety in Raichur district of a major cotton producer state Karnataka in India. Variations of quantity arrivals in market can be expected to follow a periodic pattern with period s, less than or equal to twelve months. A usual procedure to model such seasonal variations is to employ a multiplicative ARIMA model proposed by Box & Jenkins [1]. In such a modelling a series is fitted in two ways: first observations between periods are linked with an ARIMA (P,D,Q) model and then observations within periods are linked with another ARIMA (p, d, q). Then on combining these two ARIMA models the seasonal multiplicative model of order (p,d,q) x (P,D,Q) is obtained. On the other hand "the general linear model

$$\tilde{Z}_t = \sum_{j=1}^{\infty} \pi_j \tilde{Z}_{t-j} + a_t$$

with suitable values for the coefficients π_j is entirely adequate to describe many seasonal series" (Box & Jenkins [1], pp. 301-302). A sensible way of such a nonparsimonious representation is possible when condition of invertibility is imposed. This means that the contribution of the remote history of the process beyond some time t–p (p > 0) becomes negligible to its current time (t) output. Thus the model reduces to the usual Autoregressive (AR) form. Considering the fact that the techniques of identification and parameter estimation of Autoregressive model are simpler and straight forward compared to the multiplicative model one would desire to fit an Autoregressive model to such a seasonal series. However, such a model, in particular when fitted to a seasonal series is expected to involve a large number of parameters and as such loses its usefulness. To circumvent this problem Akaike's Information Criterion, AIC is used to make the model a parsimonious one.

An attempt is made here to forecast with both types of models the quantity of cotton arrivals in the market based on monthly data collected over a period of six years. A comprehensive study of the series is carried out with the multiplicative model. It is also shown how a parsimonious (or subset) AR model can be obtained using AIC. Finally, the forecast performances of the two models are compared.

2. *The Multiplicative Model*

An ARIMA (p, d, q) model has a difference equation representation (Box and Jenkins, [1]

$$\varphi_p (B)\, \nabla^d\, Z_t = \theta_q (B)\, a_t \qquad (2.1)$$

where Z_t is the time series, $\{a_t\}$ is the white noise process, and $\varphi_p(B)$ and $\varphi_q(B)$ are polynomials of degree p and q respectively in B, the backward shift operator. When $\varphi_q(B) = 0$ and $d = 0$, equation (2.1) reduces to an AR(P) model.

In analysing a seasonal series we link observations Z_t, Z_{t-s}, Z_{t-2s}, . . . by an ARIMA $(P, D, Q)_s$ model which is of the form

$$\varphi_P (B^s)\, \nabla_s^D\, Z_t = \theta_Q (B^s)\, \alpha_t \qquad (2.2)$$

where $\varphi_P (B^s)$ and $\theta_Q (B^s)$ are polynomials of order P and Q respectively in B^s. The α_t which is the residual of the above model would not in general be uncorrelated. The series $\{\alpha_t\}$ thus can be represented by an ARIMA (p, d, q) model

$$\psi_p(B)\, \nabla^d\, \alpha_t = \nu_q (B)\, a_t \qquad (2.3)$$

where a_t is now a white noise process and $\psi_p(B)$ and $\nu_q(B)$ are polynomials of order p and q respectively in B. Combining the models (2.2) and (2.3) one can finally obtain the multiplicative model

$$\psi_p(B)\, \varphi_P(B^s)\, \nabla_s^D\, \nabla^d Z_t = \theta_Q (B^s)\, \nu_q(B)\, a_t \qquad (2.4)$$

3. *Identification and Parameter Estimation*

I. *The multiplicative model:*

To identify the period or periods associated with the seasonals, employ the technique of periodogram analysis (Bloomfield [2]). Plot the intensities $I(f_i)$, corresponding to Fourier frequencies $f_i = i/N$, $i=1, 2, \ldots, [N/2]$, N being the number of observations. The plot exhibits distinct peaks at those frequencies f_i, which correspond to periods of the series. The intensities $I(f_i)$ corresponding to frequencies f_i for any given series of N observations of (Z_t, t = 1, 2, . . ., N) may be obtained by using the expression:

$$I(f_i) = \frac{N}{2}\left(a_i^2 + b_i^2\right) \tag{3.1}$$

where $$a_i = \frac{2}{N}\sum_{t=1}^{N} Z_t \operatorname{Cos}(2\pi f_i t)\ ; \qquad i = 1, 2, \ldots, N/2$$

and $$b_i = \frac{2}{N}\sum_{t=1}^{N} Z_t \operatorname{Sin}(2\pi f_i t)\ ; \qquad i = 1, 2, \ldots, N/2$$

After identifying the seasonal components, inspect the pattern of autocorrelation function of the series to identify P, Q, p and q of equation (2.4). In this identification procedure, characteristic patterns for autoregressive operators of some frequently encountered seasonal models fitted to series can be consulted in Box and Jenkins [1] (pp. 329-333).

II. *The AR(P) model* :

In identifying the model of the form (2.1) with q = 0, first carry out differencing operation of the series with respect to the period s to make it deseasonal. When stationarity is achieved with appropriate differencing, inspect the partial autocorrelation function of the differenced series to get a clue of the full order of the model. The highest lag for which the partial autocorrelation is significant is tentatively taken as the order (p) of the model. For seasonal series with period s=12 the most likely value for p is 12. The value of p is tentative because the parsimonious AR form of the seasonal series must not contain all terms up to lag p and often only a subset of few elements of the set of lags {1, 2, . . ., p} is required to specify the model. The best model is then selected from the possible subsets S using AIC as follows. For the model

$$\tilde{Z}_t = \sum_{r \in S} \varphi_r \tilde{Z}_{t-r} + a_t, \quad S \subset (1,2,\ldots,p), \quad \tilde{Z}_t = Z_t - \bar{Z}$$

the AIC is defined as

$$\mathrm{AIC}(S) = \frac{2 * C(S)}{n - C(S)} + \ln\left[\sum_{t=1}^{n} \frac{a_t^2}{n}\right] \tag{3.2}$$

where C(S) is the cardinality of S and n is the number of observations used. The subset S corresponding to which the AIC is least is chosen as the best model. Call such a model as best subset AR model for a seasonal series. The search for the best model need not proceed through all the $2^{N(S)}$ subsets of S, and some iterative technique may be adopted to save computation time.

Now describe the parameter estimation procedure of the models. This procedure is essentially same for both the models except for the appropriate expression of a_t.

A usual procedure in obtaining the values of the parameters of the model (2.4) is to calculate the sum of a_t^2, where $a_t = \theta_q^{-1}(B^s)\, v_q^{-1}(B)\, \psi_p(B)\, \varphi_p(B^s)\, \nabla^D \nabla^d Z_t$, for different sets of values of parameters in the admissible region of parameter space. The values of the parameters for which $\sum a_t^2$ is minimum give the least square estimates (LSE) of the parameters. Since a_t's are independent Gaussian, the LSEs are asymptotically MLEs. However, an efficient search technique to obtain the optimum parameter estimates is by using Marquardt's algorithm (Marquardt, [5]). This procedure, adopted for estimating the parameters of both the models, is mentioned here, in brief.

It is a combination of the Gauss-Newton iterative scheme and the method of steepest descent. The sum of squares of error (SS of a_t) is a function of the parameters of the model and this function is expressed in a Taylor's series form near an initial value of the parameter. The first derivative of SS with respect to each of the parameter is evaluated. Using these values, the direction of steepest descent and the optimum change of parameter values is calculated to obtain new values of the parameter. Then the SS for the new values of parameters is calculated and compared with old SS. If there is a significant reduction in SS then the new values of the parameters is taken as the initial values and using the same derivative values of SS, the direction and magnitude of steepest descent is calculated to get new values of parameters. If instead, there is an increase in SS then the values of the derivatives are freshly evaluated at the old values of the parameters before calculating the direction of steepest descent. The iteration terminates when there is no significant reduction in SS. The algorithm form of this procedure is given in section 6.

4. *Forecasting*

As forecasting with AR model is a straight forward case, the aspect of forecasting only with multiplicative ARIMA model is discussed to gain some insight in the behaviour of eventual forecast function.

At time t+l, the ARIMA (p, d, q) × (P,D,Q)s model (2.4) may be written as

$$Z_{t+l} = \lambda_1 Z_{t+l-1} + \lambda_1 Z_{t+l-2} + \ldots + \lambda_{\bar{p}} Z_{t+l-\bar{p}} + a_{t+l}$$
$$-\mu_1 a_{t+l-1} - \mu_2 a_{t+l-2} \ldots - \mu_{\bar{q}} a_{t+l-\bar{q}}$$

where $\bar{p}$ = (P+D) s+p+d and $\bar{q}$ + Q. s+q and

λ_i's and μ_i's are function of {ψ, φ, θ, ν}.

Taking conditional expectation at time t, we have

$$\hat{Z}_t (l) = \lambda_1 \hat{Z}_t (l-1) + \lambda_2 \hat{Z}_t (l-2) + \ldots + \lambda_p \hat{Z}_t (l-\bar{p})$$
$$- \mu_1 [a_{t+l-1}] - \mu_2 [a_{t+l-2}] - \ldots - \mu_{\bar{q}} [a_{t+l-q}] \qquad (4.1)$$

where conditional expectation of a_{t+k}, i.e.,

$$[a_{t \cdot k}] = \begin{cases} \varphi & \text{for } k \geq 0 \\ a_{t+k} & \text{for } k < 0 \end{cases}$$

$\hat{Z}_t(-j) = Z_{l-j}$, j>0 and $\hat{Z}_t(j)$, j>0 is the forecast for Z_{t+j}, made at origin t.

Equation (4.1) provides a forecast function. Since $Z_t - \hat{Z}_{t-1}(1) = a_t$, updating the forecasts is straight forward when observations are available.

4.1. *Eventual forecast function:*

The eventual forecast function of a seasonal series provides an insight into its seasonal aspects as can be seen in the following.

For l > $\bar{q}$ equation (4.1) reduces to

$$\hat{Z}_t (l) = \lambda_1 \hat{Z}_t (l-1) + \lambda_2 \hat{Z}_t (l-2) + \ldots + \lambda_p \hat{Z}_t (l-\bar{p}) \qquad (4.2)$$

The difference equation (4.2) has a solution of the form

$$\hat{Z}_t(l) = b_o^{(t)} f_o(l) + b_l^{(t)} f_l(l) + \ldots + b_{\bar{p}-1}^{(l)} f_{\bar{p}-1}(l) \quad l > \bar{q} \tag{4.3}$$

where $f_o(l), f_1(l), \ldots f_{\bar{p}-1}(l)$ are functions of lead time l which in general could include polynomials, sines and cosines and products of these functions (Feller, [3] depending on whether the roots of $\lambda(B) = 0$, $\lambda(B)$ being the stationary autoregressive operator, are real or complex, distinct or repeated. For a given origin t, the coefficients $b_j^{(l)}$ are constants applying to all lead times l; but they change from one origin to the next, adapting themselves appropriately to the particular part of the series being considered. The right hand side of equation (4.3) is known as the eventual forecast function as it gives us the pattern of long term behaviour of the series.

Now for example if $\lambda(B)$ is of the form

$$\lambda(B) = (1-B^{12})$$

then the roots of $\lambda(B) = 0$ are

(i) $B = 1$,

and *(ii)* eleven distinct complex roots

$$B = e^{i2\pi k/12}, \quad k = 1, 2, \ldots, 11.$$

The complementary function is of the form

$$C_o + C_k (e^{i2\pi l/12}), \quad k = 1, 2, \ldots, 11.$$

So the eventual forecast function takes the form

$$\hat{Z}_t(l) = \sum_{k=0}^{11} C_k \; e^{i2\pi l/12}, \quad l > 12 \tag{4.4}$$

where Ck, k = 0, 1, . . ., 11 are twelve adjustable coefficients.

Let $l = 12r + m$, $\quad 0 \leq m \leq 11$

In general, eventual forecast function takes the form

$$\hat{Z}_t(l) = r.b_1^{(t)} + \left(m.A + \sum^{11} C_k \; e^{i2\pi km/12} \right)$$

where A is some constant,

$$= r.b_1^{(t)} + b_{o.m}^{(t)} \qquad 0 \leq m \leq 11 \tag{4.5}$$

where $b_1^{(t)}$ and $b_{0.m}^{(t)}$ are called eventual annual trend and monthly trend respectively. Comparing (4.4) and (4.5) it is noted that eventual annual trend in this case is zero.

$$\text{Thus} \quad \hat{Z}_t(m) = \sum_{k=0}^{11} C_k\, e^{i2\pi km/12} = D_m, \text{ (say)} \qquad (4.6)$$

$$m = 0, 1, 2, \ldots, 11$$

Equation (4.6) reflects and fact that $\{D_m\}$ is the discrete inverse Fourier transform of $\{C_k\}$, and $\{C_k\}$ is the discrete Fourier transform of $\{D_m\}$, that is,

$$C_k = \frac{1}{12} \sum_{m=0}^{11} D_m\, e^{-i2\pi km/12}, \qquad k=0, 1, \ldots, 11$$

Interpret $\{C_k\}$ as the amplitude of the k-th Fourier frequency within the season. Presence of significant amplitude corresponding to some k entails intraseasonal periodic component.

5. *Model Diagnostic Checks*

In order to ensure that the periodic characteristics of the series has adequately taken into account by multiplicative model, cumulative periodogram test is employed.

The cumulative periodogram test :

The periodogram of the residual series $\{a_t\}$ as defined earlier is

$$I(f_i) = \frac{2}{N}\left[\left(\sum_{t=1}^{N} a_t.\, \text{Cos}(2\pi f_i t)\right)^2 + \left(\sum_{t=1}^{N} a_t.\, \text{Sin}(2\pi f_i t)\right)^2\right] \qquad (5.1)$$

For a white noise series, $E\,[I\,(f)] = 2\sigma_a^2$. It follows that $(1/N) \sum_{i=1}^{J} I(f_i)$ provides an unbiased estimate of the integrated spectrum $P(f_j)$ and

$$C(f_j) = \frac{\sum_{i=1}^{J} I(f_i)}{N\, s^2} \qquad (5.2)$$

is an estimate of $P(f_j)/\sigma_a^2$ where s^2 is an estimate of σ_a^2. We refer to $C(f_j)$ as the normalized cumulative periodogram. Now if the model fitted were adequate, and the parameters known accurately, then the a_t's computed from the data would yield a white noise series. For a white noise series, the plot of $C(f_j)$ against f_j would be scattered about a straight line joining the points (0.0, 0.0) and (0.5, 1.0). Marked deviations from this straight line would show model inadequacy. Using kolmogorov-Smirnov test to set confidence limits, the confidence limit boundaries for a given level are a pair of straight lines parallel to the aforementioned line (joining (0,0) and (0.5,1)) at distances $\pm K \frac{\varepsilon}{\sqrt{q}}$ above and below the theoretical line, where $q = (N-2)/2$ for N odd. $K\varepsilon$ for different levels of significance ε are listed in the table 1 (taken from Box & Jenkins, [1]).

Table 1. Values of $K\varepsilon$

ε	0.01	0.05	0.10	0.25
$K\varepsilon$	1.63	1.36	1.22	1.02

6. *Numerical Results*

I. *Fitting of the multiplicative model:*

The cotton series under study was recorded over each month. A total of 72 number of observation is available with us. For fitting and estimation procedure we have used first 60 observations and the remaining 12 observations are kept for comparing forecasts (see Fig. 4, 5, & 6). Since the data are widely scattered the logarithmically transformed series is taken for fitting procedure.

From the plot of autocorrelation function (see Fig. 1 (a)) of original series (Z_t) we note that at lags 1, 2. 11, 12, 13, 14 and 24 autocorrelations are highly significant. This allows to consider that the series is stationary and a periodic component with respect to period s = 12 is present. The fact is also corroborated from the calculated values of $I(f_i)$ of equation (3.1) (see Fig.2). It may be noted here that the identification of periodic component of a series does not necessarily call for using window unlike in the methodology of estimating spectrum of a process. Fig. 2 clearly exhibits one prominent peak corresponding to frequency $f_j = 0.083$, that is, at the period $1/f_j = 12$ time units. We now obtain the autocorrelation function for the differenced series $\nabla^{12} Z_t$ (see Fig. 1 (b)). Here the significant values are seen at 1, 5 and 12. This suggests that moving average operator $\nu(B)$ may tentatively be considered as

$$\nu(B) = (1-\nu_1 B - \nu_2 B^5 - \nu_3 B^{12})$$

The multiplicative model can now be considered as

$$(0,0,12) \times (0,1,0)_{s=12}$$

Thus the identified model for the cotton series takes the form

$$Z_t = Z_{t-12} + a_t - \nu_1 a_{t-1} - \nu_2 a_{t-5} - \nu_s a_{t-12} \qquad (6.1)$$

Having been identified the model (6.1) we now estimate the values of the parameters ν_1, ν_2 and ν_3. Following analogously the algorithm (see at the end of this section) we obtained the estimates of the parameters as

$\nu_1 = -0.4498$, $\nu_2 = 0.1719$ and $\nu_3 = 0.2994$
with $\sigma_a^2 = 1.5536$ and AIC = 0.5739.

Now obtain two sets of forecasts made at time origin 60, as given in table 2 (see also Fig. 4 & Fig. 5) from the model

$$Z_t = Z_{t-12} + a_t + .4498\, a_{t-1} - .1719\, a_{t-5} - .2994\, a_{t-12} \qquad (6.2)$$

Fig. 4 shows one-step-ahead forecasts and Fig. 5 exhibits forecasts for a lead time $l = 1, 2, \ldots, 12$ made at time origin 60 in both cases. As can be seen from these plots forecasts are very close to the observed values. The mean squared error (MSE) as a measure of one-step-ahead forecast accuracy is obtained and is given in table 2.

II. *Fitting of the subset AR model* :

From the plot of partial autocorrelation function (see Fig. 3) we note that the full order of the model is p=12. Now using AIC as given in equation (3.2), obtain the best subset model for the cotton series as given by

$$W_t = .313\, W_{t-1} + .233\, W_{t-6} + .226\, W_{t-10} - .355\, W_{t-12} + a_t \qquad (6.3)$$

where $W_t = \nabla^{12} Z_t$

The value of AIC = 0.0374 and $\sigma_a^2 = 0.8655$

Fig. 6 exhibits one-step-ahead forecast values obtained at time origin 60 from the model (6.3). Table 2 gives the comparative forecast performance of both the models.

Now employ equation (5.2) to check the adequacy of the model (6.2). Corresponding to the values of $f_j = \frac{1}{60}, \frac{2}{60}, \ldots, \frac{90}{60}$ $C(f_j)$'s are obtained and are plotted in Fig. 7. The two thick boundary lines in Fig. 7 are Kolmogorov-Smirnow 95% confidence limit boundaries. Since no points have crossed the limit boundaries, the residuals a_t's are in fact white noise series.

Seasonal index:

Finally, to detect the presence of intraseasonal periodicity, calculate the value of $\{C_k\}$ using equation (4.6). It may be noted that equation (4.6) gives the seasonal index corresponding to each month. An inverse Fourier transform of equation (4.6) gives the values of $\{C_k\}$. The plot of $\{C_k\}$ (see Fig. 8) against $k = 0,1,2,\ldots,11$, exhibits no prominent peak for any k and hence there exists no other intraseasonal periodicity which has not been accounted for.

Table 2. Forecasts comparison

Time t	observed values	Forecasts by seasonal model		Forecast by AR model (one-step)
		one-step	for l time unit	
61	9.07	8.31	8.31	8.06
62	7.85	7.50	7.16	7.24
63	7.99	7.14	6.98	7.32
64	7.47	6.47	6.09	6.66
65	8.59	9.03	8.57	8.77
66	11.27	9.47	9.79	10.07
67	11.63	10.80	10.05	10.63
68	11.26	10.52	10.30	11.63
69	11.89	10.81	10.65	11.21
70	11.98	11.73	11.17	12.17
71	11.76	10.56	10.75	11.34
72	11.04	10.74	10.35	11.02
	MSE	0.907	–	0.696

Algorithm

Objective : To obtain the final estimates of parameters by Marquardt's Algorithm

Input : N → Length of the series

Z(N) → Array containing the series

k- Number of parameters

(k) - Array containing the initial estimates of parameters. Initial estimates may be obtained by some general method (see Box & Jenkins, 1976)

Output : φ(k) → Array containing final estimates of parameters

OSS → Estimated sum of squares of error

Temporary variables:

φN(k) → Array to store modified values of parameters

NSS → Value of error SS corresponding to φN(k)

T(k,K) → Matrix to store products of derivatives of SS with respect to the parameters

A(N), AN(N) → Arrays to store error terms for φ(K) and φN(k) respectively

G (k), D(k) → Arrays for solving the increment of parameter value

H(k) → Array to store optimum increment in parameter values at each iteration

X(k,N) → Matrix to store derivatives of A(N) with respect to φ(k)

it → Iteration number

Del, eps, P1 → Small constants between 0 and 1

F2 → Constant between 1 and 2

Steps : 0 : {initialise} it = 0;

I : Calculate error SS corresponding to φ(k) as

$$OSS = \sum_{t=1}^{N} A^2(t), \text{ where } A(t) = Z(t) - \sum_{l=1}^{k} \varphi(l) Z(t-S_l)$$

II : {calculation of the derivatives}
1 = 1;

IIa : $\varphi N(j) = \begin{cases} \varphi(j) & \text{if } j \neq 1 \\ \varphi(j) + Del & \text{if } j = 1 \end{cases}$

IIb : $AN(t) = Z(t) - \sum_{j=1}^{k} \varphi N(j) Z(t-s_j)$, t = 1, 2, . . ., N

IIc : X (1,t) = {A (t) - AN(t)}, t = 1, 2, . . ., N

IId : l = l+1; if (l ≤ k) go to IIa

III : {form matrix T, array G and scaling factors D}

$$T(i,j) = \sum_{t=1}^{N} X(i,t).X(j,t), \quad i=1, 2, ..., k, \quad j=1, 2, ..., k;$$

$$G(i) = \sum_{t=1}^{N} X(i,t).A(t), i=1, 2, . . ., k;$$

$$D(i) = \sqrt{T_{ii}}, \quad i = 1, 2, . . ., k;$$

IV : {modify matrix T and array G}

$$T(i,j) = \frac{T(i,j)}{D(i).D(j)}, \quad i, j = 1, 2, . . ., k$$

$$G(i) = \frac{G(i)}{D(i)}, \quad i = 1, 2, . . ., k$$

IVa : T(i,i) = 1 + P1, i = 1, 2, . . ., k

V : Solve the set of linear simultaneous equations

T.H = G

for H

VI : {scale back H to get optimum increments}

$$H(i) = \frac{H(i)}{D(i)}, \quad i = 1, 2, . . ., k$$

VII : {get new parameter values}

φN(1) = φ(1) + H(l), l = 1, 2, . . ., k

VIII : Calculate the new error SS as

$$NSS = \sum_{t=1}^{N} AN^2 (t),$$

where $AN(t) = Z(t) - \sum_{l=1}^{k} \varphi N(1) Z(t-S_l)$

IX : If (NSS < OSS) go to X

P1 - P1 + F2; it = it + 1; go to IVa

x : If (|H(1)| <eps, 1 = 1, 2, . . ., k) go to XI

$P1 = \frac{P1}{F2}$; φ(1) = φN(1),

l=1, 2, . . ., k ; it = it + 1; go to I

XI : Stop. {$\varphi(l)$, l=1, 2, . . ., k are the optimum parameter values}

Conclusion :

Results of Kolmogorov-Smirnov test to the residuals of the series as well as the absence of intraseasonal periodicity have revealed that the model (6.2) is adequate. On the otherhand, from table 2 we note that for the concerned case the subset AR model (6.3) gives better forecasts. In order to obtain a general conclusion which of the two competing models should be employed for seasonal series, it is further investigated the suitability of a subset AR model to the well known series G (International Airline Passengers) of Box & Jenkins [1]. In this case the best subset AR model is obtained with only two parameters as given by

$$X_t = -0.357\, X_{t-1} - 0.415\, X_{t-12} + a_t$$

where $X_t = \nabla\ \nabla^{12}\ Z_t$.

The value of AIC = −6.400 ($\sigma_a^2 = 1.61 \times 10^{-3}$) as obtained for this model is slightly larger than the value of AIC = −6.580 ($\sigma_a^2 = 1.34 \times 10^{-3}$) for the well known $(0,1,1) \times (0,1,1)_{12}$ model. Nevertheless the subset AR model is no way inferior to the multiplicative ARIMA model as can be seen when forecasts are generated from any arbitrary origin. Since the identification procedure of subset AR model is simple we suggest the use of such models when there arises complicacy in the identification of multiplicative ARIMA models. If subset AR modes are used for seasonal series one can then go for tractable nonlinear models like Volterra type models (see Sarkar and Kartikeyan, [6] and Kartikeyan and Sarkar, [4] for highly accurate forecasts.

ACKNOWLEDGEMENT

The authors are thankful to Dr. S. Satish, Ex.-Research Scholar, R.D.C., IIT, Kharagpur for providing us the crop series data.

REFERENCES

[1] Box, G.E.P. and Jenkins, G.M., 1976. *Time Series Analysis Forecasting and Control*, Holden-Day Inc.

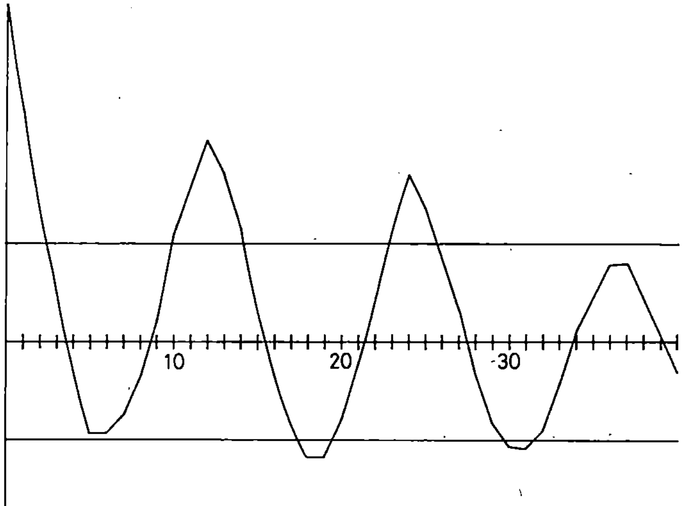
[2] Bloomfield, P., 1976. Fourier Analysis of Time Series. John Wiley and Sons,

[3] Feller, W., 1972. *An Introduction to Probability Theory and Its Applications*, Wiley Eastern Pvt. Ltd.,

[4] Kartikeyan, B and Sarkar, A., 1989. Shape Description by Time Series, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **11**, (9), 977-984.

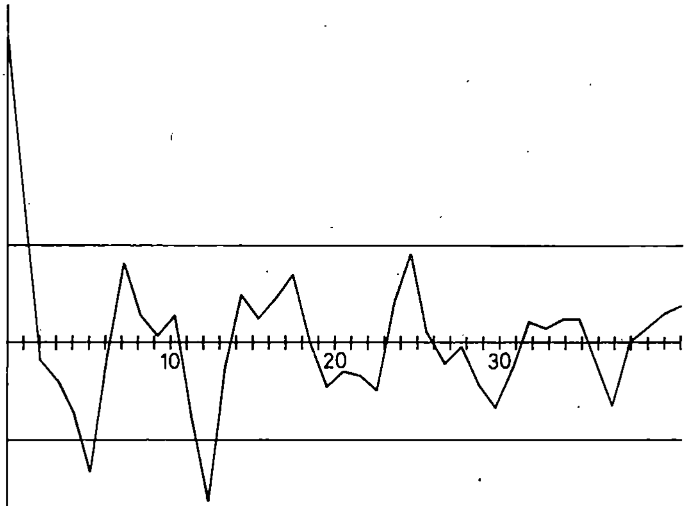
[5] Marquardt, D.W., 1963. An algorithm for least squares estimation of nonlinear parameters, *Journal of SIAM*. **11**, 431-441.

[6] Sarkar, A and Kartikeyan, B., 1987. Forecasting by Volterra type models, *Journal of Statistical Computation and Simulations*, **28**, 245-260.



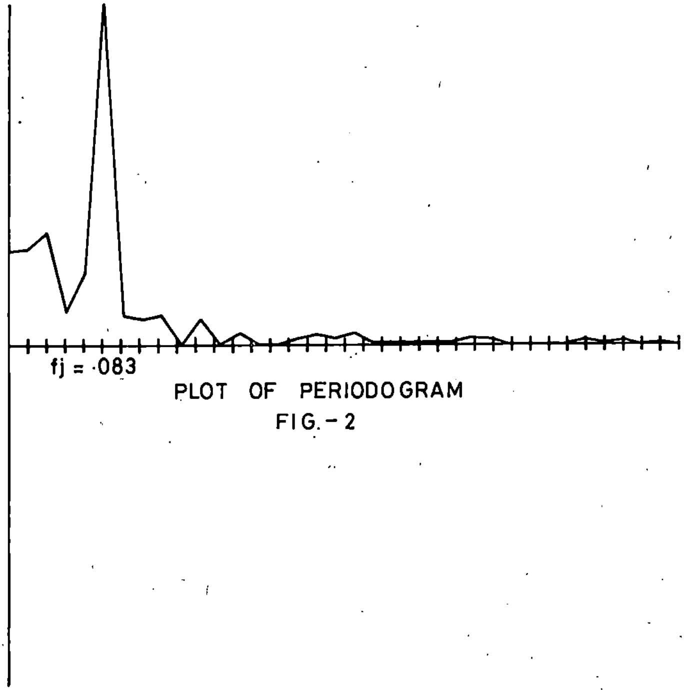
CORRELATION PLOT OF Zt AT LAG K

FIG.- 1(a)

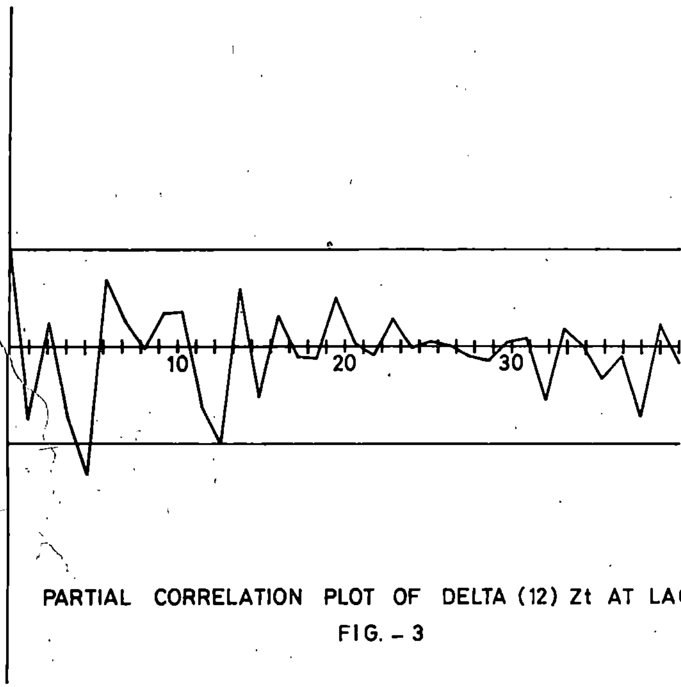


CORRELATION PLOT OF DELTA (12) Zt AT LAG K

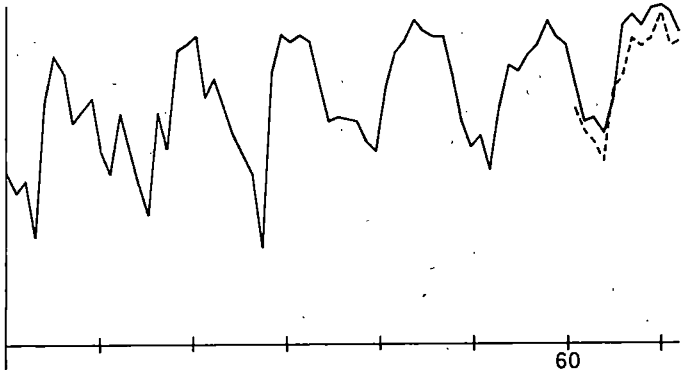
FIG.- 1 (b)



PLOT OF PERIODOGRAM
FIG. – 2

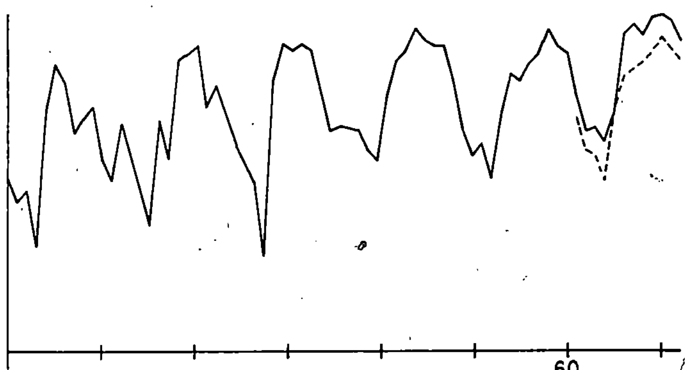


PARTIAL CORRELATION PLOT OF DELTA (12) Zt AT LAG K
FIG. – 3



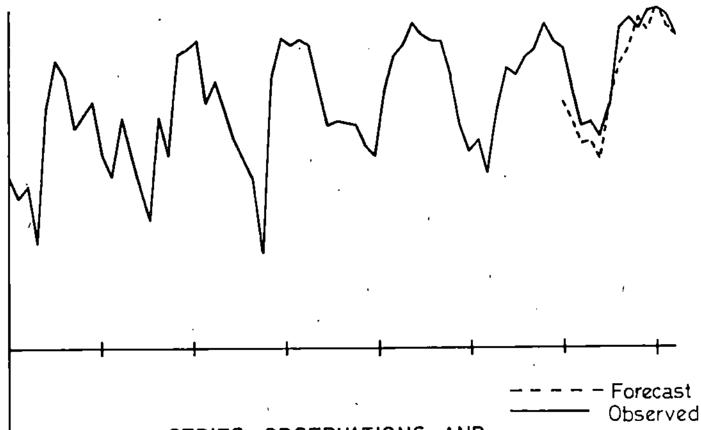
COTTON SERIES AND ONE-STEP-AHEAD FORECASTS MADE AT TIME ORIGIN 60

FIG. - 4



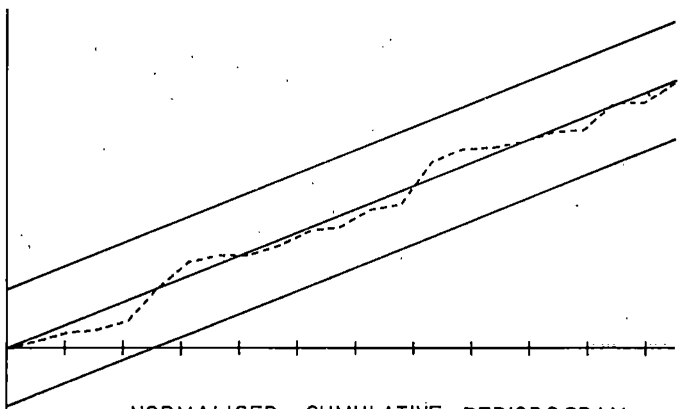
COTTON SERIES AND FORECASTS MADE AT TIME ORIGIN 60 FOR LEAD TIMES 1-12

FIG.- 5



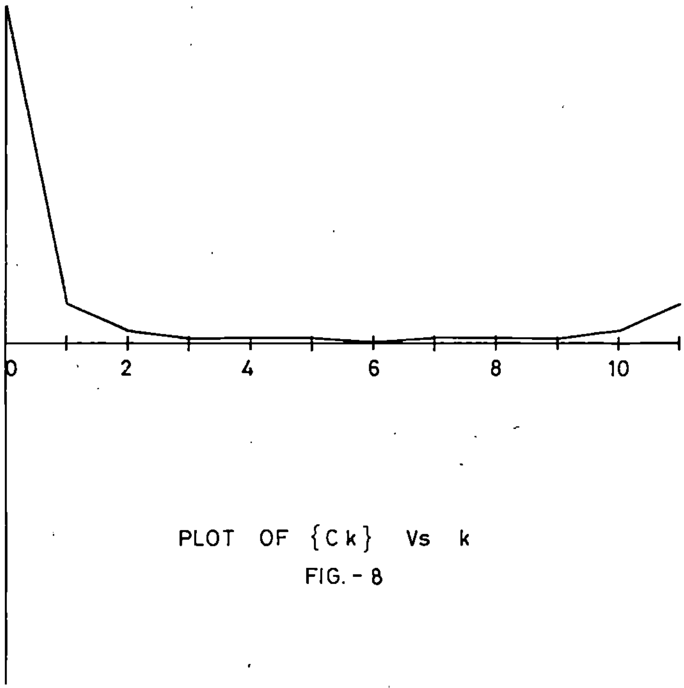
SERIES OBSERVATIONS AND ONE-STEP-AHEAD FORECASTS

FIG. - 6



NORMALISED CUMULATIVE PERIODOGRAM

FIG. - 7



PLOT OF {C k} Vs k

FIG. - 8