Asian Journal of Scientific Research

ISSN 1992-1454

science alert
http://www.scialert.net

ANSInet
an open access publisher
http://ansinet.com

Asian Journal of Scientific Research 1 (5): 481-491, 2008
ISSN 1992-1454

Forecasting in Subsets Autoregressive Models and Autoprojective Models

[1]J.F. Ojo, [2]T.O. Olatayo and [2]O.O. Alabi
[1]Department of Statistics, University of Ibadan, Ibadan, Nigeria
[2]Department of Mathematical Science, Olabisi Onabanjo University, Ago-Iwoye, Nigeria

Abstract: Full autoregressive models are always characterize by many parameters and this is a problem. Some of these parameters are redundant that is close to zero and there is the need to eliminate these parameters through the concept of subsetting. Subsets autoregressive models are free from redundant parameters thereby lowering the residnal variance and forecasting with such models will always give a better forecast. Likewise auto projective models calculate on the basis of current knowledge what the errors would have been which gives us some guide to errors of the future. It is in the light of the above we considered the subsets autoregressive models and auto projective models, to see how these models will perform with regard to forecast. Exponential smoothening was used to forecast the future value in auto projective models while conditional least square predictor was used to forecast the future value in subset autoregressive models. An algorithm was proposed to eliminate redundant parameters from the full order autoregressive models and the parameters were estimated. To determine optimal models, residnal variance, Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) were adopted. Results revealed that the residual variance attached to the subset autoregressive models is smaller than the residual variance attached to the auto projective models. We conclude that the forecast for subset autoregressive is preferred to the forecast for auto projective.

Key words: Forecast, autoregressive, auto projective, residual variance, conditional least square, algorithm

INTRODUCTION

It may be said that the era of linear time series models began with such linear models as Yule's autoregressive (AR) models (1927), first introduced in the study of sunspot numbers. In the past five decades or so, we have seen remarkable successes in the application of linear time series models in diverse fields for example Box and Jenkins (1970), Hannan (1970), Chatfield (1980), Priestely (1978) and Nottingham International Time Series Conference in March (1979). These successes are perhaps rather natural in view of the significant contributions of linear differential equation in all branches of science. In particular, as far as a one-step-ahead prediction is concerned, a linear time series model is often quite adequate.

Forecasting counote an attempt to see into the future. There are two words, which are used to denote numerical forecasting methods namely forecasting and prediction. Forecasting is the process of estimation in unknown situations. Prediction is a similar, but more general term and usually refers to estimation of time series, cross-sectional or longitudinal data. Risk and uncertainty are central to forecasting and prediction. In more recent years, forecasting has evolved into the practice of demanding planning in every day business forecasting for manufacturing companies. The discipline of demand planning, also sometimes referred to as supply chain forecasting, embraces both statistical forecasting and consensus process. Forecasting is commouly used in discussion of time-series data.

Corresponding Author: J.F. Ojo, Department of Statistics, University of Ibadan, Ibadan, Nigeria

Therefore, forecasting is a powerful useful instrument in planning and making a wise decision about future. In this research we considered and compared two methods of forecasting, subset autoregressive and auto-projective.

MATERIALS AND METHODS

Autoregressive Model

A time series $\{X_t\}$ is said to follow an autoregressive process of order P, if it satisfies the difference equation:

$$X_t - \mu = K_1(X_{t-1}-\mu) + K_2(X_{t-2}-\mu) + \ldots + K_p(X_{t-p}-\mu) + e_t$$

A finite stationary stochastic process $\{X_t\}$ is defined as:

$$X_t = \Phi_1 X_{t-1} + \Phi_2 X_{t-2} + \ldots + \Phi_p X_{t-p} + e_t \quad (1)$$

Where:

(i) e_t is a Gaussian process
(ii) $\Phi_1, \Phi_2, \ldots, \Phi_p$ is a finite set of weight parameter
(iii) $E(X_t) = \mu = 0$

using backward shift operator B, we can write (1) as $\Phi(B)\, X_t = e_t$ where

$$\Phi(B) = 1-\Phi_1(B)-\Phi_2(B)^2- \ldots -\Phi_p(B)^p$$

and the equation $\Phi(B) = 0$ is called the characteristic equation. To ensure stationary, the roots of the characteristic equation $\Phi(B) = 0$ must lie outside the unit circle.

The estimate of the parameter Φ_i, i = 1, 2, …, p can be obtained by Yule Walker method.

Subset Autoregressive Model

A zero-mean stationary stochastic process $\{X_t\}$ is said to be generated by an autoregressive model of order K, denoted by AR(k), if it satisfies the difference equation

$$X_t = \Phi_1 X_{t-1} + \Phi_2 X_{t-2} + \ldots + \Phi_k X_{t-k} + e_t \quad (2)$$

where, $\{e_t\}$ is a white noise process with variance σ^2. Here, $\{e_t\}$ will be assumed to a Gaussian process.

When a model of the form (2) is fitted to a set of observations on a stationary time series $\{X_t\}$, the fitted model will include all the terms $\{X_{t-i}; i = 1,2, \ldots,k\}$.

In many situations, in particular where there may be evidence that a time series may have some form of a seasonal behaviour, this may lead to models, which include many more parameters than are strictly necessary to describe its behaviour. It is often desirable to use models of the form (2) where some of the $\{X_t\}$ are set equal to zero. Such models are referred to as subset auto regressive time series models.

A major problem in fitting autoregressive time series models, even of full order, has always been the choice of the order of the model. Consequently, many researchers as Akaike (1973), Box and Jenkins (1970), Hannan (1970), Parzen (1974), Quenonille (1947), Walker (1952) and Whittle (1951) to mention a few, have paid special attention to this problem. The choice of the order of the model goes

further back in classical statistics. For example, choosing the best linear multiple regression has always posed a problem to statisticians.

In time series modelling, subset models are often desirable, especially when the data exhibits some form of periodic behaviour. In such cases, fitting full order models often results in the fitted coefficients of some lags being close to zero.

Furnival (1971) has given an algorithm for including and excluding variables in a multiple linear regression. By applying matrix operations to the augmented matrix associated with the multiple regressions, estimates of the regression coefficients and residual variance of the corresponding regression model are obtained. Furnival pointed out that the application of Gaussian elimination to the augmented matrix has the advantage of saving a great deal of computer time, although it yields ouly the maximum likelihood estimate of the residual variance and not the regression coefficients. This is not a disadvantage, since the coefficients of any model of interest may easily be recomputed.

Algorithm for 2^k-1 Possible Subsets in Time Series Models

The 2^k-1 subsets of time series models make use of the properties of permutation and combinatorial analysis and the algorithm goes thus:

The k-value, which is an integer and a maximum lag, is identified in our model.

The first sets of numbers are one digit number 1, 2, 3, up to k. The second sets of numbers are 2-digit numbers arranged in such a way that the first digit 1 is taking and combine with the next number until k is reached; the next digit 2 is picked and combines with the next number until k is reached, the next digit with the next number until k is reached. This is continuing until (k-1) and k is reached.

The third sets of numbers are three digits and the second digit is operated on to produce third digit. Our guide is that the 2-digit must have the next forward number until k is reached and the next two digits must have the next forward digit until k is reached. This is continuing until we have (k-2)(k-1) and k. This is continuing until we have the maximum digit, which is the digit k that is (k-n)(k-s)(k-u)(k-z)(k-m)(k-y)……k.

where, n<k by 1, s<k by 2, u<k by 3……k. Suppose we have our k to be 4, there are 2^4-1 possible subsets that is 15 subsets. Following our algorithm we shall have the following:

1, 2, 3, 4, 12, 13, 14, 23, 24, 34, 123, 124, 134, 234, 1234

The equation for the following are

$$X_t = b_1x_{t-1} + e_t \quad (3)$$

$$X_t = b_2x_{t-2} + e_t \quad (4)$$

$$X_t = b_3x_{t-3} + e_t \quad (5)$$

.

.

.

.

$$X_t = b_1x_{t-1} + b_2x_{t-2} + b_3X_{t-3} + b_4x_{t-4} + e_t \quad (6)$$

where, $1 = b_1$, $2 = b_2$, $3 = b_3$, $4 = b_4$. If the minimum AIC occurred in (5) that model is the Subset Time Series model.

Conditional Mean Square Predictor

The minimum mean square error predictor is given by

$$X_t(m) = E(X_{t+m}/X_s: -\infty < s \leq t)$$

We shall show as follows that the conditional mean square predictor has the smallest mean square error:

Let M (m) denote the mean square error for the m-step ahead predictor $X_t(m)$ i.e.,

$M(m) = E(X_{t+m} - X_t(m))]$, where, $X_t(m)$ is any predictor.

$M(m) = E\{[(X_{t+m} - X_t(m)) + (X_t(m) - X_t(m))]^2\}$

$= E(X_{t+m} - X_t(m))^2 + E(X_t(m) - X_t(m))^2 + 2E[(X_{t+m} - X_t(m)) + (X_t(m) - X_t(m))]\}$

$M(m) = T_1 + T_2 + T_3$

But $T_3 = E[E(X_{t+m} - X_t(m)(X_t(m) - X_t(m) \mid X_s: -\infty < s \leq t] X_s: -\infty < s \leq t$

Since $X_t(m)$ and $X_t(m)$ are based on $X_s: -\infty < s \leq t$, hence given $X_s: -\infty < s \leq t$ $X_t(m)$ and $X_t(m)$ are mixed.

Therefore $T_3 = E_{xs}(E(X_t(m) - X_t(m) E[X_{t+m} \mid X_s : -\infty < s \leq t - X_t(m)] = 0$

Thus $M(m) = E(X_{t+m} - X_t(m))^2 + E(X_t(m) - X_t(m))^2$ and for M(m) to be minimum $X_t(m) = X_t(m)$ that is, $X_t(m) = E[X_{t+m}/X_s : -\infty < s \leq t]$ has the smallest mean square error.

Autoprojective Models

The method to be considered is that for which we extrapolate forwards an existing series, without regard to other series, which may be concomitant. Auto-projective method calculates on the basis of current knowledge what the errors would have been which gives us some guide to errors of the future. Exponential smoothening is a simple method of auto-projective or adaptive forecasting. It is an effective way of forecasting when you have ouly few observations on which to base your forecast. Uulike regression models, which use fixed coefficients, forecasts from exponential smoothing method adjust based upon past forecast error.

Autoprojective Forecasting Models

The exponential smoothing method is appropriate for series that move randomly above and below a constant mean with no trend and seasonal patterns. The smoothed series $\hat{X}_t$ of X_t is computed recursively by $\hat{X}_t = \alpha X_t + (1-\alpha)\hat{X}_{t-1}$ where, $0<\alpha<1$ is the damping or smoothing factor. The smoother the α the smoother the X_t. By repeated substitution, we can rewrite the recursion as $\hat{X}_t = \alpha\sum(1-\alpha)\hat{X}_{t-s}$. This shows why this method is called exponential smoothing. The forecast of X is a weighted average of the past value of X, where the weight decline exponentially with time.

Holt Winters Method of Exponential Smoothing

This method is appropriate for series with linear time trend and no seasonal variation. This method is a two-parameter method; the smoothed series $\hat{X}_t$ of X_t is given by:

$$\hat{X}_{t+k} = \alpha + bk$$

Where:

α = Parameter component (intercept)

b = Trend

These two coefficient are defined by the following recursions

$\alpha(t) = \alpha Xt + (1- \alpha)\, \alpha(t-1) + b(t-1)$
$b(t) = \beta(\alpha(t)-\alpha(t-1)) + (1-\beta)b(t-1)$

where, $0< \alpha, \beta<1$ are damping factors. This is an exponential smoothing method with two parameters. Forecast are computed by $\hat{X}_{T+k} = \alpha(T) + b(T)k$. These forecast lie in a linear trend with intercept $\alpha(T)$ and slope b(T).

Consider also a process which is a constant α_0 plus a random residual ε with mean zero. We construct the prediction at time t.

$$\alpha_0(t) = (1-\beta)[X_t + \beta X_{t-1} + \beta^2 X_{t-2} + ...] \tag{7}$$

$$= (1-\beta)\sum_{j=0}^{\infty} \beta^j X_{t-j} \tag{8}$$

where, due to $|\beta<1|$ the weights diminish exponentially

$$\alpha_0(t-1) = (1-\beta)[X_{t-1} + \beta X_{t-2} + ...]$$
$$\alpha_0(t) = (1-\beta)X_t + \beta\alpha_0(t-1)..... \tag{9}$$

The formular bears some resemblance to a markoff process. If then we know the constant β, we can construct an estimator of $\alpha_0(t)$ at time t from the estimate made at time t-1 plus (1-β) times the actual observation at time t. This could provide an estimate of the future values for the series.

Let the error at time t-1 in forecasting X_t one unit ahead be e_t that is to say

$$X_t - \alpha_0(t-1) = e_t$$

Substituting for $\alpha_0(t-1)$ in (9) $\alpha_0(t) = X_t - \beta e_t = \alpha_0(t-1) + (1-\beta)e_t$. Thus our forecast at time t is the forecast made at time t-1.

Test for Linearity

A zero-mean stationary stochastic process $\{X_t\}$ is said to be generated by an autoregressive model of order k, denoted by AR(k), if it satisfies the difference equation

$$X_t = \Phi_1 X_{t-1} + \Phi_2 X_{t-2} + \ldots + \Phi_k X_{t-k} + e_t \tag{10}$$

where, $\{e_t\}$ is a white noise process with variance σ^2. Here, $\{e_t\}$ will be assumed to be a Gaussian process.

Suppose, in a multiple linear regression, the response variable is given by Y and there is a set of explanatory variables, say $\{X_1, X_2, \ldots.X_k\}$. The full linear regression model is given a set of derivations on $\{Y, X_1, X_2, \ldots.X_k\}$,

$$Y_i = a_1 X_{1i} + a_2 X_{2i} + \ldots + a_k X_{ki} + e_i: i= 1, 2,\ldots,N \tag{11}$$

where, $\{e_t\}$ are usually assumed to be independently distributed as $N(0, \sigma^2)$. The problem is to search for that subset of explanatory variables, which best explains the variation in Y.

Comparing the AR(k) given by (10) with the multiple regression model given by (11), it may be seen that they are similar, except that for the AR model the X's are not deterministic. However, it is well known that; asymptotically, the results obtained by assuming the X's are deterministic can be directly extended to time series modeling.

With this knowledge, we discussed the test for linearity or test of significance of departure from linear regression (Table 1).

The Deviations Mean Square and the F-test

In the multiple regression models, the deviations of the Y's from the population regression plane have mean 0 and variance σ^2. An unbiased estimate of σ^2 is:

$$S^2 = \sum (Y - \hat{Y})^2 / (n - k)$$

where, n is the size of sample and k is the number of parameters that have been estimated in fitting the regression.

The multiple regression model with two independent variable is given by

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + e \tag{12}$$

Given a sample of n values of (Y, X_1, X_2) the sample regression-the prediction equation is:

$$\hat{Y} = a + b_1 X_1 + b_2 X_2 \tag{13}$$

The values of a, b_1 and b_2 are chosen so as to minimize $\sum (Y - \hat{Y})^2$ the sum of squares of n the differences between the actual and the predicted Y values. With our model, theory shows that the resulting estimates a, b_1, b_2 and $\hat{Y}$ are unbiased and have the smallest standard errors of any unbiased estimates that are linear expressions in the Y's. The value of a is given by the equation

$$a = \bar{Y} - b_1 X_1 - b_2 X_2$$

By substituting for a in 2 the fitted regression can be written as:

$$\hat{Y} = \bar{Y} + b_1 x_1 + b_2 x_2$$

where, $x_1 = X_1 - \bar{X}$, as usual.

Since the sample means of x_1 and x_2 are both zero, the sample mean of the fitted values $\hat{Y}$ is $\bar{Y}$. Write $\hat{y} = \hat{Y} - \bar{Y}$ and $d = Y - \hat{Y}$, so that d represents the observed deviation of Y from the fitted regression at this point. It follows that

$$y = Y - \bar{Y} = (\hat{Y} - \bar{Y}) + (Y - \bar{Y})$$
$$= \hat{y} + d$$
$$\sum y^2 = \sum \hat{y}^2 + \sum d^2$$

Sum of square due to regression $= \sum \hat{y}^2$

$$= b_1 \sum x_1 y + b_2 \sum x_2 y$$

Sum of square deviations from curved regression $= \sum d^2 = \sum y^2 - \sum \hat{y}^2$

Sum of squares of deviations from a linear regression $= \sum y^2 - \frac{(\sum xy)^2}{\sum x^2}$

Table 1: Analysis of variance table for the testing of significance of departure from linear regression

Source of variation	df	Sum of squares	Mean square
Deviations from linear regression	n-2	$\sum y^2 - \sum \hat{y}^2$	$(\sum y^2 - \sum \hat{y}^2)/n-2$
Deviations from curved regression	n-3	$\sum y^2 - \frac{(\sum xy)^2}{\sum x^2}$	* = $(\sum y^2 - \frac{(\sum xy)^2}{\sum x^2})/(n-3)$
Reduction sum of squares	1	Difference	** = Difference/1

F = **/*

Residual Variance

Residual variance or unexplained variance is part of the variance of any residual. In analysis of variance and regression analysis, residual variance is that part of the variance which cannot be attributed to specific causes. The unexplained variance can be divided into two parts. First, the part related to random, everyday, normal, free will differences in a population or sample. Among any aggregation of data these conditions equal out. Second, the part that comes from some condition that has not been identified, but that is systematic. That part introduces a bias and if not identified can lead to a false conclusion.

Akaike Information Criteria (AIC)

The Akaike information criterion (AIC) (pronounced ah-kah-ee-keh), developed by Hirotsugu Akaike in 1971 and proposed in Akaike (1974), is a measure of the goodness of fit of an estimated statistical model. It is grounded in the concept of entropy. The AIC is an operational way of trading off the complexity of an estimated model against how well the model fits the data.

In the general case, the AIC is AIC = 2k-2ln(L)

where, k is the number of parameters and L is the likelihood function.

Over the remainder of this entry, it will be assumed that the model errors are normally and independently distributed. Let n be the number of observations and RSS be the residual sum of squares. Then AIC becomes

$$AIC = 2k + nln(RSS/n)$$

Increasing the number of free parameters to be estimated improves the goodness of fit, regardless of the number of free parameters in the data generating process. Hence AIC not only rewards goodness of fit, but also includes a peualty that is an increasing function of the number of estimated parameters. This peualty discourages overfitting. The preferred model is the one with the lowest AIC value. The AIC methodology attempts to find the model that best explains the data with a minimum of free parameters. The AIC penalizes free parameters less strongly than does the Schwartz.

Bayesian Information Criterion

In statistics, the Bayesian Information Criterion (BIC) is a statistical criterion for model selection. The BIC is sometimes also named the Schwarz criterion, or Schwarz Information Criterion (SIC). It is so named because prof. Gideon E. Schwarz (1978) gave a Bayesian argument for adopting it.

Let:

n = The number of observations, equivalently, the sample size

k = The number of free parameters to be estimated. If the estimated model is a linear regression, k is the number of regressors, including the constant

RSS = The residual sum of squares from the estimated model

L = The maximized value of the likelihood function for the estimated model

The formula for the BIC is:

$$BIC = \left(\frac{-2\ln L + k\ln(n)}{n}\right)$$

Under the assumption that the model errors or disturbances are normally distributed, this becomes:

$$BIC = \ln\left(\frac{RSS}{n}\right) + k\left(\frac{\ln n}{n}\right)$$

Given any two estimated models, the model with the lower value of BIC is the one to be preferred. The BIC is a decreasing function of RSS, the goodness of fit and an increasing function of k. The BIC penalizes free parameters more strongly than does the Akaike Information criterion.

RESULTS AND DISCUSSION

Table 2: Analysis of variance table for the testing of significance of departure from linear regression

Source of variation	df	Sum of squares	Mean square
Deviations from linear regression	185	19.6000	0.1059
Deviations from curved regression	184	19.6044	0.1070
Reduction sum of squares	1	0.0350	0.0350

F = 0.035/0.1070, = 0.33

From Table 2, the test revealed non departure from linearity and as a result the data used for this research is linear in nature.

Fitting of Full and Subset AR to Real Series

For our illustration, we have considered a well known series that is chemical process concentration readings every 2 h with 187 observations. We fitted full AR for 187 observations and forecast for 15 observations. Using the Yule Walker estimation procedure, the linear model of all orders up to AR (30) is fitted. The choice of the order is made on the basis of Akaike Information Criteria (AIC). It is found that AIC is minimum when p = 7. The fitted model is:

$$X_t = 0.361362x_{t-1} + 0.242756x_{t-2} + 0.012948x_{t-3} + 0.045449x_{t-4} + 0.004248x_{t-5} + 0.088649x_{t-6} + 0.245123x_{t-7} + e_t$$

The Algorithm described earlier was employed in fitting of best subset AR model. There are 2^7-1 possible subsets, that is, 127 subsets. The choice of the best subset is made on the basis of minimum AIC and BIC and having considered the 127 possible subsets, it is found that AIC and BIC are minimum in the model

$$X_t = 0.371775x_{t-1} + 0.265296x_{t-2} + 0.109497x_{t-6} + 0.253978x_{t-7} + e_t$$

Auto-Projective Model

$X_t = 0.136840(t-1) = e_t$

Exponential smoothening

Included observations: 187

Method: Holt-Winters No Seasonal
Original Series: CHEMICALF
Forecast Series: CHEMICSM

Parameters:	Alpha	1.0000
	Beta	0.0000
Sum of squared residuals		0.079346
Root mean squared error		0.020172
End of period levels:	Mean	17.26274
	Trend	0.001300

Table 3: Performance of subset autoregressive and autoprojective models

Subset autoregressive model				Autoprojective model			
Mean	RV	AIC	BIC	Mean	RV	AIC	BIC
0.00088	0.08950	0.46317	0.53412	4.30758	55.5150	7.15	7.16

Rv = Residual Variance, AIC = Akaike Information Criterion, BIC = Bayesian Information Criterion

Table 4: Forecast for subset autoregressive models

1	NA	NA	NA	NA	NA
6	NA	NA	16.86460	16.72765	16.59578
11	16.56912	16.75631	16.75708	16.78870	16.80208
16	16.76622	16.72002	16.70706	16.73761	16.74919
21	16.77109	16.78178	16.77740	16.76545	16.75990
26	16.76369	16.76897	16.77867	16.78591	16.78876
31	16.78810	16.78761	16.78879	16.79151	16.79609
36	16.80066	16.80423	16.80654	16.80836	16.81024
41	16.81262	16.81567	16.81898	16.82218	16.82504
46	16.82761	16.83007	16.83260	16.83533	16.83821
51	16.84113	16.84399	16.84675	16.84943	16.85211
56	16.85482	16.85759	16.86039	16.86320	16.86598
61	16.86873	16.87147	16.87421	16.87697	16.87974
66	16.88252	16.88529	16.88806	16.89082	16.89358
71	16.89634	16.89911	16.90188	16.90465	16.90742
76	16.91019	16.91296	16.91573	16.91850	16.92127
81	16.92404	16.92681	16.92959	16.93236	16.93513
86	16.93791	16.94068	16.94345	16.94623	16.94900
91	16.95178	16.95456	16.95733	16.96011	16.96289
96	16.96567	16.96844	16.97122	16.97400	16.97678
101	16.97956	16.98234	16.98513	16.98791	16.99069
106	16.99347	16.99626	16.99904	17.00182	17.00461
111	17.00739	17.01018	17.01296	17.01575	17.01854
116	17.02132	17.02411	17.02690	17.02969	17.03248
121	17.03527	17.03806	17.04085	17.04364	17.04643
126	17.04922	17.05201	17.05481	17.05760	17.06039
131	17.06319	17.06598	17.06878	17.07157	17.07437
136	17.07716	17.07996	17.08276	17.08556	17.08835
141	17.09115	17.09395	17.09675	17.09955	17.10235
146	17.10515	17.10795	17.11076	17.11356	17.11636
151	17.11917	17.12197	17.12477	17.12758	17.13038
156	17.13319	17.13599	17.13880	17.14161	17.14442
161	17.14722	17.15003	17.15284	17.15565	17.15846
166	17.16127	17.16408	17.16689	17.16970	17.17251
171	17.17533	17.17814	17.18095	17.18377	17.18658
176	17.18940	17.19221	17.19503	17.19784	17.20066
181	17.20348	17.20629	17.20911	17.21193	17.21475
186	17.21757	17.22039	17.22321	17.22603	17.22885
191	17.23167	17.23449	17.23732	17.24014	17.24296
196	17.24579	17.24861	17.25144	17.25426	17.25709
201	17.25991	17.26274			

Table 5: Forecast for auto-projective model

1	NA	NA	NA	NA	NA
6	NA	NA	16.86460	16.86590	16.72895
11	16.59708	16.57042	16.75761	16.75838	16.79000
16	16.80338	16.76752	16.72132	16.70836	16.73891
21	16.75049	16.77239	16.78308	16.77870	16.76675
26	16.76120	16.76499	16.77027	16.77997	16.78721
31	16.79006	16.78940	16.78891	16.79009	16.79281
36	16.79739	16.80196	16.80553	16.80784	16.80966
41	16.81154	16.81392	16.81696	16.82028	16.82348
46	16.82634	16.82891	16.83137	16.83390	16.83663
51	16.83951	16.84243	16.84529	16.84805	16.85073
56	16.85341	16.85612	16.85889	16.86169	16.86450
61	16.86728	16.87003	16.87277	16.87551	16.87827
66	16.88104	16.88382	16.88659	16.88936	16.89212
71	16.89488	16.89764	16.90041	16.90318	16.90595
76	16.90872	16.91149	16.91426	16.91703	16.91980
81	16.92257	16.92534	16.92811	16.93089	16.93366
86	16.93643	16.93921	16.94198	16.94475	16.94753
91	16.95030	16.95308	16.95586	16.95863	16.96141
96	16.96419	16.96697	16.96974	16.97252	16.97530
101	16.97808	16.98086	16.98364	16.98643	16.98921
106	16.99199	16.99477	16.99756	17.00034	17.00312
111	17.00591	17.00869	17.01148	17.01426	17.01705
116	17.01984	17.02262	17.02541	17.02820	17.03099
121	17.03378	17.03657	17.03936	17.04215	17.04494
126	17.04773	17.05052	17.05331	17.05611	17.05890
131	17.06169	17.06449	17.06728	17.07008	17.07287
136	17.07567	17.07846	17.08126	17.08406	17.08686
141	17.08965	17.09245	17.09525	17.09805	17.10085
146	17.10365	17.10645	17.10925	17.11206	17.11486
151	17.11766	17.12047	17.12327	17.12607	17.12888
156	17.13168	17.13449	17.13729	17.14010	17.14291
161	17.14571	17.14852	17.15133	17.15414	17.15695
166	17.15976	17.16257	17.16538	17.16819	17.17100
171	17.17381	17.17663	17.17944	17.18225	17.18507
176	17.18788	17.19070	17.19351	17.19633	17.19914
181	17.20196	17.20478	17.20759	17.21041	17.21323
186	17.21605	17.21887	17.22169	17.22451	17.22733
191	17.23015	17.23297	17.23579	17.23862	17.24144
196	17.24426	17.24709	17.24991	17.25274	17.25556
201	17.25839	17.26121			

From Table 3, we could see the performance of the two models. Residual variance attached to subset autoregressive model is smaller than the residual variance attached to the autoprojective model. This is suggesting to us that autoregressive model perform better than the autoprojective model. With this result it is now wise for us to forecast with the best model. Autoprojective model is as well adequate if we forecast with it but the subset autoregressive model should be preferred. The forecast of the two models are as given in Table 4 and 5.

CONCLUSION

In this study, we have seen clearly the estimation of the parameters of subset autoregressive models and autoprojective models. The linearity test was performed on our data for standardization that is ascertaining that the data is linear in nature. As we have noted earlier that forecasting with best model is the appropriate thing to do. As a result we have seen the performance of subset autoregressive model over autoprojective model through the residual variance. Though it is good to forecast with the two models which we have done but the forecast attached to the subset autoregressive model should be preferred. Conclusively, after performing linearity test on our data and we want to go ahead and forecast, subset autoregressive model should be preferred.

REFERENCES

Akaike, H., 1973. Maximum likelihood identification of gaussian autoregressive moving average models. Biometrika, 60: 255-265.

Akaike, H., 1974. A new look at the statistical model indentification. IEEE Trans. Automatic Conrol, 19 (6): 716-723.

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

Chatfield, C., 1980. The Analysis of Time Series-An Introduction. Chapman and Hall.

Furnival, G.M., 1971. All possible regressions with less computation. Technometrics, 13: 403-408.

Hannan, E.J., 1970. Multiple Time Series. John Wiley, New York.

Parzen, E., 1974. A central limit theorem for multi-linear stochastic processes. Ann. Math. Stat., 28: 252-256.

Priestely, M.B., 1978. Non-linear models in time series analysis. Statistician, 27: 159-176.

Quenonille, M.H., 1947. A large sample test for goodness of fit of autoregressive schem. J. R. Stat. Soc. Ser., A110: 123-129.

Schwarz, G., 1978. Estimating the dimension of a model. Ann. Stat., 6 (2): 461-464.

Walker, A.M., 1952. Some properties of asymptotic power functions of goodness of fit tests for linear autoregressive schemes. J. R. Stat. Soc. Ser., B14: 117-134.

Whittle, P., 1951. Hypothesis Testing in Time Series Analysis. University of Uppsala Publication.

Yule, G.U., 1927. On the method of investigating periodicities in disturbed series with special reference to wolfer's snnspot numbers. Phil. Trans. R. Soc. London Ser., (A) 226-267.