APPLICATION OF AUTOREGRESSIVE FORECASTING TECHNIQUE TO PETROLEUM PRODUCTION IN NIGERIA

By

J. OLORO AND J. A .AJIENKA

ABSTRACT

The decision problem concerning forecast of gas and crude oil productions in Nigeria oil industry is the main focus of this research study. Thirty and twenty three years respectively time series of gas and oil production were obtained from NNPC records and analyzed by the autoregressive forecasting technique and multiple linear regressions.

A correlogram developed enabled us to build predictive models for gas and oil (Equation 7 and 8 respectively) for forecasting future production.

The actual gas produced for thirty years fall within range of 12996MCM and 43636MCM while forecasted value fall within the range of 27262.3MCM and 94701.3MCM.

The actual oil produced for twenty-three years fall within range of 450.9Mbbl and 855.7Mbbl while forecasted value fall within the range of 64.1Mbbl and 984Mbbl.

The mean of actual gas produced for thirty years was 26896.6MCM and the forecasted value was 30575.3MCM.

The mean of actual oil produced for twenty-three years was 634.4Mbbl and the forecasted value was 620.6Mbbl

Durbin Watson tests for significance carried out suggests that the models developed are reliable. Each model can be use to predict at 0.5 level of significance.

INTRODUCTION

The need for petroleum and its products has been on the increase over the years. Petroleum has contributed to a High percentage of total exports revenue in the nation economy between 1965 and 1974. It accounted for 80% by 1997.From 1997 to 2000,there was instability in average oil and gas production as a result of the following factors; price, political instability, OPEC quota and clashes between two communities[1]

The increase in the production of oil and gas necessitated the need to look into how much oil and gas are actually being produced.

Since the future is full of uncertainties, production of crude oil and gas in the near future cannot be easily ascertained[1].

Consequently in this work models will be developed, which can be used to forecast what oil and gas production will be in future.

The aim of the work is to develop a forecasting model for the set of data collected.

The objectives of the work are:

1 To apply the auto-regressive technique of forecasting and multiple regression analysis to develop a forecasting model for annual production of crude oil, and gas in the country.

2 To use this model to generate forecast values for oil and gas production in order to be able to test for the authenticity of the models developed.

METHODOLOGY

In this work, thirty and twenty three years gas and oil production respectively time series were analyzed. For these time sets of data, the autocorrelations between the time series and its past values were obtained. With these values a correlogram was graphed[2].

By carefully inspecting the correlogram the lagged variables, which appear to contribute more to the prediction of the time series, were determined. The lagged variables were plugged into the multiple linear regression model and the equations to this model were solved to obtained values for the regression coefficients. The regression parameters were then substituted into the linear regression model to obtain the required predictive equation[3].

The forecasting equation would take the form;

$$Y_t = B_o + B_1 Y_{t-1} + B_2 Y_{1-2} + \ldots\ldots\ldots\ldots\ldots + B_p Y_{t-p} \quad \ldots\ldots 1.0$$

Where Y_t = dependent variable

$Y_{t-1}, Y_{t-2}.. . Y_{t-p}$ = the predictor variable and $B_{0,}\ B_1, B_2 \ldots\ldots B_p$ are the least squares regression estimates.

To determine the correlation between Y_t and Y_{t-k} for any lagk, the equation below was used.

$$\frac{\sum_{t=1}^{N=K} (Y_t - y)(Y_{t_{t-}} - y)}{\sum (Y - y)^2} \quad \ldots\ldots\ldots\ldots 2.0$$

$$y = \frac{\sum Y}{N} \quad \ldots\ldots\ldots\ldots 3.0$$

Were k = lag under consideration

y = is the average of the observation

r_k = is the auto correlation for lag k

Y_t = total observation

N = number of observations

The auto-correlation was plotted against the lags. This plot is known as the correlogram[3]. The correlogram is significant in that by inspecting it, the lagged variables that would contribute to the prediction of the time series variables can be identified

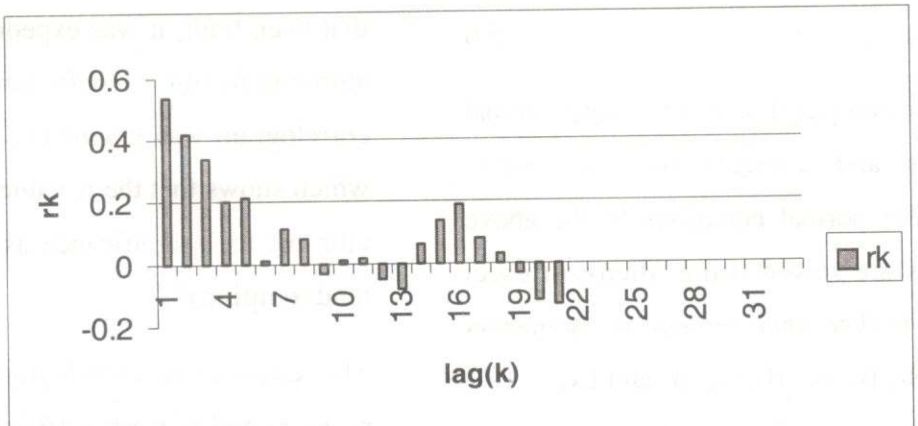

fig. 1 gas production correlogram

Namely y_{t-1}, Y_{t-2}, Y_{t-3}, Y_{t-4}, Y_{t-5} and Y_{t-1}, Y_{t-2}, Y_{t-3} Y_{t-4}, Y_{t-5} for gas and oil respectively. These were the lagged variable that exhibited or wielded highest auto-correlation. Since five lagged variables (y_{t-1}, Y_{t-2}, Y_{t-3}, Y_{t-4}, Y_{t-5}) were considered, the

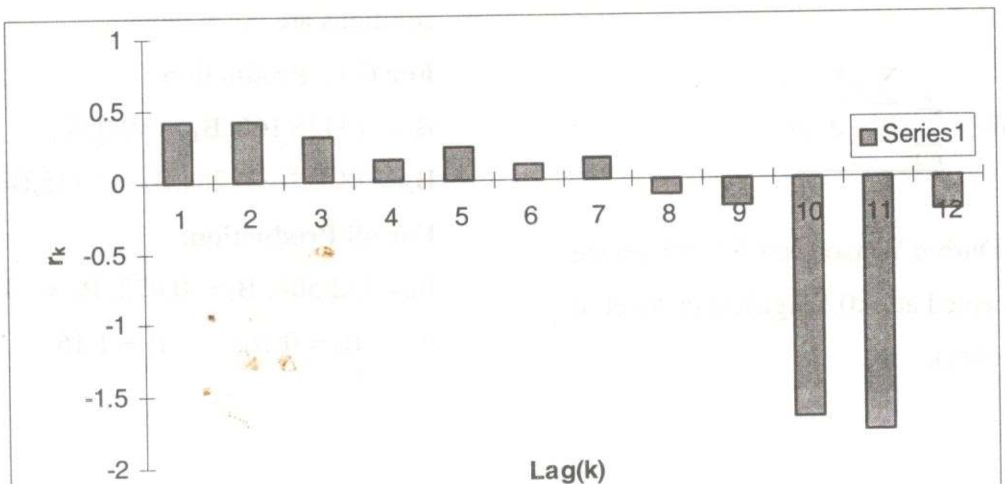

Fig 2 oil production correlogram

equation would be a five variables equation.

FOR GAS PRODUCTION is given as $Y_t = B_0 + B_1 Y_{t-1} + B_2 Y_{t-2} + B_3 Y_{t-3} + B_4 Y_{t-4} + \ldots B_5 Y_{t-5}$ 4.0

The normal equations to the above equation are solved using Microsoft excel program to determine regression estimates.

FOR OIL PRODUCTION

$Y_t = B_0 + B_1 Y_{t-1} + B_2 Y_{t-2} + B_3 Y_{t-3} + B_4 Y_{t-4} + B_5 Y_{t-5}$ 5.0

The data computed are fitted into normal equations and arranged to have matrix forms. The normal equations to the above equations are solved using Microsoft excel program to determine regression parameters namely B_0, B_1, B_2, B_3, B_4, B_4 and B_5, .

MODEL TESTING

The model developed was used to generate values given by Y_t. The residuals were computed and Durbin Watson test for autocorrelation is given as

$$D.W. = \frac{\sum_{t-2}^{n} (e_t^{-1} - e_{t-1})^2}{\sum^{n} e_t^2}$$

In The Durbin Watson test for sigificance was accepted at $p<0.5$ sigificance level or (probability).

equation 6.0 above, e_t = resideuals which is the difference between the actual value and that obtained with the model.

$(e_t - e_{t-1})$ is the difference between the residuals and a prevoius residual[4].

RESULTS A N D DISCUSSION

In order to ascertain which lag variables that would offer predictive values in the model that been built, it was expedient to compute auto-correlation coefficients (r_k). A correlogram was graphed (See fig 1 and 2) which shows that the r_k values oscillates and tails off to insignificance as the lag values tend to infinity[5].

The results of the correlogram points that r_1, r_2, r_3, r_4 and r_k were sufficiently significant to describe the equation that can predict future trend.

The matrixes formed was solved by using Microsoft excel programme and the final solutions are given below.

For Gas Production :

B_0= 13426.166 ,B_1 =0.038 ,

B_2=0.001 ,B_3=0.264,B_4=12.115,B= -11.979

For oil Production:

B_0= 332.504, B_1= -0.072, B_2 = -0.102, B_3 = -0.12,B_4 =-0.166 B_5= 1.16

The regression parameters are shown above. With these, the models for predicting gas and oil production are given as follows:

MODEL DEVELOPED FOR GAS PRODUCTION

$Y_t = 13426.166 + 0.038\ Y_{t-1} + 0.001\ Y_{t-2} + 0.264\ Y_{t-3} + 2.115\ Y_{t-4} - 11.979\ Y_{t-5}$ ……………………. ……………………7.0

Unit for gas is in Million cubic metres

MODEL DEVELOPED FOR OIL PRODUCTION

$Y_t = 332.504 - 0.072\ Y_{t-1} - 0.102 Y_{t-2} - 0.12 Y_{t-3} - 0.166 Y_{t-4} + 1.16\ Y_{t-5}$ …………………………….............. …...8.0

Unit for oil is in Million Barels.

The model developed were used to generate values given by Y_t .The residuals (See table 1 and 2) were computed and Durbin Watson test for autocorrelation is given in equation 6.0 was used for the caculation.In generating the values the models were used as follow;

$Y_t = 13426.166 + 0.038\ Y_{t-1} + 0.001\ Y_{t-2} + 0.264\ Y_{t-3} + 12.115\ Y_{t-4} - 11.979\ Y_{t-5}$ …………………………………….. 9.0

Unit for Gas is million cubic metres(MCM). For instance if we are to generate a value for t= 5, we would have. $Y_5 = 13426.166 + 0.038\ Y_4 + 0.001 Y_3 - 0.264\ Y_2 + 12.115 Y_1 = 1.5048$ MCM

Using the Durbin – Watson statistic given in equation 6.0

$$D.W. = \frac{\sum_{t-2}^{n}(e_t - e_{t-1})^2}{\sum_{t-1}^{n} e_t^2} = \frac{7.25 \times 10^{10}}{6.879 \times 10^{10}}$$

D.W. = 1.05

Note that

H_0: There exist no autocorrelation among residuals

H_1: There exist auto correlation among residuals.

From the table

d_L = 1.071

d_n = 1.833

Hence H_o is rejected

Since Durbin – Watson computed value D.W. = 1.05 < than critical value from the statistical table, we reject the null hypothesis that no autocorrelation exist in error function. We thus assert that the residual i.e. the error function between the actual and forecasted co-vary that is, they bear some functional relationship.

In order to ascertain which lag variables would offer predictive values in the models that were built, It was expedient to compute auto correlation coefficients r_k. The results of the computation were used to draw the correlogram, as shown in fig.1 and fig.2, for gas and oil production respectively. The actual gas produced for thirty years fall within the range of 12996MCM and 43636MCM while forecasted value fall within the range of 27262.3MCM and 94701.3MCM.

The actual oil produced for twenty-three years fall within range of 450.9Mbbl and 855.7Mbbl while forecasted value fall within the range of 64.1Mbbl and 984Mbbl.

The mean of actual gas produced for thirty years was 6896.6MCM and the forecasted value was 30575.3MCM.

The mean of actual oil produced for twenty three years was 634.4Mbbl and the forecasted value was 620.6Mbbl.

Durbin Watson tests for significance carried out .The calculated was 1.05 which was less than the values from the statistical table(1.071 and 1.833) for gas production.For oil production,the calculated value was 0.865,which was also less than the critical values from the statistical table[6]. This suggests that the models developed are reliable. This is because the residuals values obtained are so insiginifcant that the probability of having values greater than the observed would be less than 10%. Although significant point is that our models suggest that a minimum of five years adequate record maintenance is sufficient and reliable for making future predictions on productions.Perhaps, it is pertinent to explain why the correlogram oscillates to zero value as the number of lags increases as shown in fig 1 and fig 2. This is because it would appear that there is complete absence of seasonality. If there had been seasonality then, the months of the year for which seasonal index is highest would have corresponded to the point of highest autocorrelation coefficient r_k. However in this particular study, the production data were given on annual basis and so it was not possible to decompose them in order to detect seasonal index[6].

CONCLUSION

The objective of this study is to apply the autoregressive technique of forecasting and multiple regression analysis to develop a forecasting models for future annual oil and gas productions and this has been realized.

The autoregressive technique perform reasonably well for time series that (1) is not extremely volatile and does not contain extreme amount of random movement and (2) requires a short term or medium range forecast.The fact that the autoregressive procedure does not perform well on a time series containing a great deal of irregular activity is not a serious disadvantage,practically all forecasting techniques perform poorly in this situation.

This technique is better than other techniques like; Box-Jenkins,ARIMA,Adaptive fittering and decline curve.For instance Box-Jenkins require very large time series(60years), also decline curve method which is been used by most oil companies require flow rate and other parameters for one to use it for prediction[7].

RECOMMENDATIONS

Base on the conclusion reached, I wish to recommend

1. Further research work might be necessary to improve in the results of this study.

Table 1: The Durbin – Watson Test For Auto Correlation In Gas Production In Nigeria (1971 –2000)

S\N	YEAR	GAS PRODUCTION Million cubic metres	FORECASTED VALUES Y_t	RESIDUAL e_t	$(e_t - e_{t-1})$	$(e_t - e_{t-1})^2$	e_t^2
1	1971	12996	13426.1	-430.2			1.85×10^5
2	1972	17122	13920.0	3202.0	-3632.2	1.319×10^7	1.02×10^7
3	1973	21882	14089.8	7792.2	-4590.2	2.106×10^7	6.07×10^7
4	1974	27170	10843.9	16326.1	-8533.9	7.283×10^7	2.66×10^8
5	1975	60435	15048.7	-42265.7	58591.8	3.433×10^9	1.78×10^9
6	1976	18656	60120.9	-41464.9	-800.8	6.41×10^5	1.72×10^9
7	1977	21815	66976.4	-45161.4	3696.5	1.366×10^7	2.04×10^9
8	1978	20486	76516.9	-56030.9	10869.5	1.181×10^8	3.14×10^9
9	1979	27450	-6035.3	123488.6	-1795.2	3.22×10^{10}	1.52×10^{10}
10	1980	24551	17089.6	7461.4	116027	1.34×10^7	5.56×10^7
11	1981	17113	49786.7	-3267.8	10729.2	1.15×10^8	1.07×10^7
12	1982	15382	-5701.6	21083.6	-24351	5.93×10^8	4.44×10^8
13	1983	15192	94701.3	-79509.3	100592	1.0×10^{10}	6.32×10^9
14	1984	16251	21887.1	38138.2	-117647	1.38×10^{10}	1.45×10^9
15	1985	18569	90200.8	108769.5	-70631.3	4.98×10^9	1.18×10^{10}
16	1986	18738	-8506.3	27244.3	81525.2	6.6×10^9	7.42×10^8
17	1987	17170	9656.6	7513.4	19730.9	3.89×10^8	5.64×10^7
18	1988	20250	24089.2	-3839.2	11352.6	1.288×10^8	1.47×10^7
19	1989	25129	39556.8	-14427.8	10588.6	1.12×10^8	2.08×10^8
20	1990	28430	14441.3	13988.7	-28416.5	8.07×10^8	1.95×10^8
21	1991	31460	-27262.3	38722.3	-24733.6	6.11×10^8	1.49×10^9
22	1992	32084	47663.6	-15579.6	54301.9	2.94×10^9	2.43×10^8
23	1993	33680	69034.4	-35354.4	19774.9	3.91×10^8	1.25×10^9
24	1994	33680	49859.5	-16179.5	-19174.9	3.67×10^8	2.61×10^8
25	1995	35200	46844.0	-11744.4	-4435.1	1.97×10^7	1.38×10^8
26	1996	35450	29661.1	5848.9	17593.3	3.09×10^8	3.42×10^7
27	1997	37150	37090.7	59.3	5789.6	3.35×10^7	3516.5
28	1998	37039	10187.4	26851.6	-26792.3	7.18×10^8	7.21×10^8
29	1999	43636	27295.8	16340.2	-16340.2	2.67×10^6	2.67×10^8
30	2000	42732	14776.4	27955.6	-11615.4	1.35×10^8	7.815×10^8

Table 2: The Durbin – Watson Test For Auto Correlation In Oil Production In Nigeria

S/N	YEAR	OIL PRODUCED Million barrels	FORECASTED	RESIDUAL e_t	e_{t-1}	$(e_t - e_{t-1})$	$(e_t - e_{t-1})^2$	e_t^2
1	1980	752.2	332.5	19.7		419.7	176148.1	176148.1
2	1981	525.3	278.3	247	19.7	-172.7	29825.3	60988.5
3	1982	470.6	217.9	252.6	247	56	31.4	63827.9
4	1983	450.9	273.2	177.7	252.6	-74.5	5610	31587.2
5	1984	507.9	64.1	443.8	177.7	266.1	70809.2	196925.8
6	1985	547	646.3	-99.3	443.8	-543.1	118680.3	9864.4
7	1986	535.7	649.4	-113.7	-99.3	-213	45369	12926.4
8	1987	482.8	648.2	-165.4	-113.7	-279.1	77896.8	27369.8
9	1988	490.4	616.2	-125.9	-165.4	291.3	84855.7	15824
10	1989	626.4	682	-55.6	-125.9	-181.5	32942.3	3094.4
11	1990	630.2	725	-94.8	-55.6	-150.4	22620	8994.7
12	1991	690.9	705.7	-14.8	-94.8	-109.6	12012.2	217.7
13	1992	716.2	622	94.2	-14.8	109	6304.4	8882.6
14	1993	695.3	599.7	95.6	94.2	14	196	9134.9
15	1994	646.2	748.5	-102.3	95.6	-197.9	39164.4	10463.8
16	1995	672.5	745.5	-73	-102.3	29.3	858.5	5322.5
17	1996	681.8	817.3	-135.5	-73	-62.5	3906.3	18357.6
18	1997	855.7	840.1	15.6	-135.5	151.1	22831.2	242.5
19	1998	766.5	819.9	-53.4	15.6	-69	4761	2854.6
20	1999	767.5	746.2	21.3	-53.4	74.7	5580.1	454.7
21	2000	855.7	763.3	92.4	21.3	71.1	5055.2	8538.1
22	2001	450.9	749.5	-298.6	92.4	-391	152881	89144.3
23	2002	773	984	-213.0	-298.6	85.6	7327.4	45369
					-		**925666.8**	**1164360.7**

REFERENCES

1. Udosen J. Udosen (2000): Project Work on Problems Prospect of Petroleum Production in Nigeria. (A Case Study of NPDC and Dubri Oil Company) University of Benin. Pg. 10 -23

2. Bunn, D. W. (1982): "Short Term Forecasting": a review of Procedures in the Electricity Supply Industry. Journal of operational Research Society, Vol. 33, pg. 533 - 545 Great Britain.

3. Iver B. and John B. Sought (1981): "Introduction Statistics for Business and Economics". Dryden Press Hinsdale, Illinois.pg20.

4. Hill G. W. and Woodworth D. (1980): Automatic Box-Jenkins Forecasting: Operational Research Society. Vol. 31, pg. 413-422, Pergamon Press Ltd., Great Britain.

5. Kontsoyiannis .A.(1984) Theory of Econometrics,second edition, pg. 200, Palgrave Publisher

6. Damodar .N. Gujarati (2001) Basic Econometric, Fouth edition,pg. 200,467-469, Teta Mcgraw-Hill publisher company limited.

7. Wheel Wrigth S. C. and Makridakis S. (An Examination of the Use of Adaptive Filtering in Forecasting. Operational Research Quarterly Vol. 24.