ISSN No:-2456-2165

An Empirical Investigation of Efficiency Measurement in Financial and Economic Time Series Data

[1]AKINTUNDE M. O., AMUSAN, A.S.,[1]MASOPA, N.A. AND [1]ADESIYAN, A.A.
[1]Department of Statistics, Federal Polytechnic, Ede, Osun State

Abstract:- This paper investigates the efficiency measurement in financial and Economic data. Efficiency measurement has contributed significantly to the reduction in the volume of error encountered in the day-to-day human endeavour. The most unfortunate thing is that little or no attention is directed towards the publications addressing this problem. This paper therefore serves as a gap filling study aimed at addressing the problem arising in the direction of efficiency measures. The data employed in the study is Nigerian Crude oil data (2009-2018) analyzed through the use of Econometric View software (E-view). The efficiency measures indices used are Mean Absolute Error (MAE), Root Mean Square Error $(RMSE)$, Mean Absolute Deviation (MAD), Mean Absolute Precision Error (MAPE) and THEIL U inequality. From the results the performance measures indices that produced the most efficient measurement is THEIL U inequality and its components, Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) in that order.

***Keywords**:- Efficiency, MAE, $(RMSE)$, MAD, RMSE and MAPE.*

I. INTRODUCTION

Researchers see the use of performance measure indicators as tools to monitor, control and improve industrial processes and systems for a long time. These indices are needed in all facets of human endeavour, this may be government, academic, business, investment and industry. The efficiency measure embraces all aspects of life. However; the die needs of these indices have not been justified. But rather, solutions so provided are not as needed.

The first notable contribution to the efficiency measurement was traced to the classic 1939 work on statistical process control by Achor et.al. (2018). No doubt there are certainly scholarly focus on specific topics. For instance, United Kingdom has received much attention in her 'league tables' for educational institutions Alim and Hand (2017), Barber (2017). In academia, 'bibliometrics' are commonly used in ranking academic institutions, in awarding research grants, and for promotion assessment. Biener et. al. (2014) used the indices as Performance indicators so also in the area of health Bialowolski et.al. (2018) and Weziak-Bialowolski et.al. (2018), have used performance indices severally and in several articles by Vander-Weele (2017) and his colleagues in the context of implementing Six Sigma programmes.

Fisher (2013) looked into the measurement of innovations as response to the ineffective charity of Nesta (nesta.org.uk) and two UK Government departments embraced the evolution of an innovation index in a way that they will be able to rank firms and governments relative to innovation performances.

II. MATHEMATICAL PRELIMINARY

2.1 Mean square error (MSE)

The Mean-Square Error (MSE) is an invaluable measure of accuracy used to measure the size of the differences between predicted values and the actual values obtained for a series. These difference between the predicted and actual values are referred to as errors. It is use as a sum total of the sizes of the errors in predictions at various times into a single measure of predictive power. It is a constantly used as error index in statistics (Chu and Shir Mohammadi, 2004; Singh et al., 2004; Vasquez-Amábile and Engel, 2005). The lower the MSE the better the model performance. MSE is given as;

$$: \quad MSE = N^{-1}\sum_{t=1}^{T}\left(Y_t - \hat{Y}_t\right)$$

2.2 Mean absolute error (MAE)

The most frequently used and easiest measure of forecast accuracy is called Mean Absolute Error (MAE). The absolute error refers to the deviation between the predicted value and the actual value. It shows the size of an error obtainable from the forecast average. This is expressed mathematically as: $MAE = N^{-1}\sum_{t=1}^{T}\left|Y_t - \hat{Y}_t\right|$

2.3 Mean absolute precision error (MAPE)

Ahlburg, 1995; Campbell 2002; Hyndman and Koehler 2006; Isserman, 1977; Miller2001; Murdock et al, 1984 Rayer; 2007; Sink 1997; Smith 1987; Smith and Sincidi1990, 2001; Taymanet et. al, 1998; Wilson 2007 affirmed that MAPE is always used to examine the cross-sectional forecasts. Its universality note is often found in software

ISSN No:-2456-2165

packages; another remarkable statistical features of MAPE is that it makes use of all observations in set of data and possesses smallest variation from one sample to other. It is a beautiful way of reporting because of its ability to expressed results in percentage terms that are readily understandable to different categories of users. Its popularity is as a result of simplicity in calculating it and equally very easy to understand.

It is given as: $MAPE = N^{-1}\sum_{t=1}^{T}\left|\frac{Y_t - \hat{Y}_t}{Y_t}\right| X \ 100$

2.4 Mean absolute deviation about median (MAD median) shows a direct measure of the scale of a random variable relative to its median and is severally used in many field of human endeavour Babu, C.J. and Roa, C.R., (1992), TUKEY, J. W. (1960) and Pham-Gia, T. and Hung, T.L.(2001).The MAD median is far more better compared to the standard deviation in real life situations where little errors will occur in observation and measured Ghurye, W. H, (2004),Huber, P. (1981).

It is given as:- $MAD = \frac{\sum_{i=1}^{n}\lfloor x_i - \bar{x} \rfloor}{n}$

2.5 Theil's U inequality coefficient

This is a useful measure to determine the forecast accuracy of a given series. Theil's U inequality coefficient as explained by (Pindyck and Rubinfeld, 1998), measures the root mean square error in relative terms, and is defined as

$$U = \frac{\sqrt{\frac{1}{n}\sum_{t=1}^{n}(Y_t^s - Y_t^a)^2}}{\sqrt{\frac{1}{n}\sum_{t=1}^{n}(Y_t^s)^2} + \sqrt{\frac{1}{n}\sum_{t=1}^{n}(Y_t^a)^2}}$$

The denominator of Theil U is in two regions, it is bounded above by 1 and bounded below by 0, that is, $0 \leq U \leq 1$. It is particularly useful as it gives yardstick used in evaluating the accuracy of a model and could be compared to other models. The closer this value to 0, the better is the coefficient and the more accurate the model is. forecast performance is considered bad, if its coefficient equal 1. The U coefficient can be further divided into three proportions which give additional information on the usefulness and the performance of the model.

Bias, $$U^M = \frac{(\bar{Y}^s - \bar{Y}^a)^2}{\frac{1}{n}\sum_{t=1}^{n}(Y_t^s - Y_t^a)^2}$$

Variance, $$U^S = \frac{(\sigma_s - \sigma_a)^2}{\frac{1}{n}\sum_{t=1}^{n}(Y_t^s - Y_t^a)^2}$$

Covariance, $$U^C = \frac{2(1-\rho)\sigma_s\sigma_a}{\frac{1}{n}\sum_{t=1}^{n}(Y_t^s - Y_t^a)^2}$$

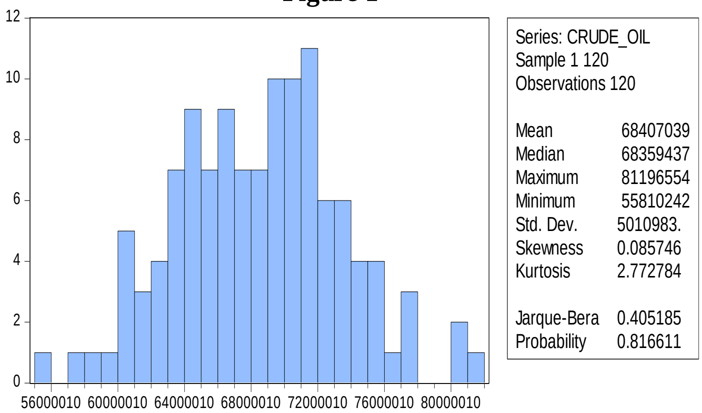
The bias proportion revealed the size of the systematic error of the forecast; it brings together the various share of the simulation errors emanating from bias, that is, the disparity between the predicted and the actual series. The variance proportion revealed the wellness of forecast of the series under study. The covariance proportion gives a measure of the unsystematic error attributed to the series. The acceptable magnitude of ideal distribution of any series must lies within inequality coefficient of $U^m = U^s = 0$ and $U^c = 1$.for any good forecast, the bias and variance proportions should be small, in a way that most of the bias should be feasible in the covariance proportions.

III. DATA ANALYSIS AND INTERPRETATIONS

The data used for the data was extracted from central bank of Nigeria Nigerian National Petroleum website covering the periods of (2009-2018). It was analysed using Econometrics view (E-view) software. Data analysis proceeds as follows:

3.1 Descriptive analysis

Figure 1



From Figure 1, he standard deviation computed is on the high side, revealing that the series under study poses high degree of fluctuations, Skewness in the data aet is tending to zero (0.085746) showing that the series is almost symmetric. From the histogram, Kurtosis is peaked in distribution and so far, the kurtosis is (2.772784) less than 3, then the series has lighter tails than a normal distribution. Jarque-Bera test p-value obtained is not less than zero indicating that the series is fairly normal so that the hypothesis of normality cannot be rejected.

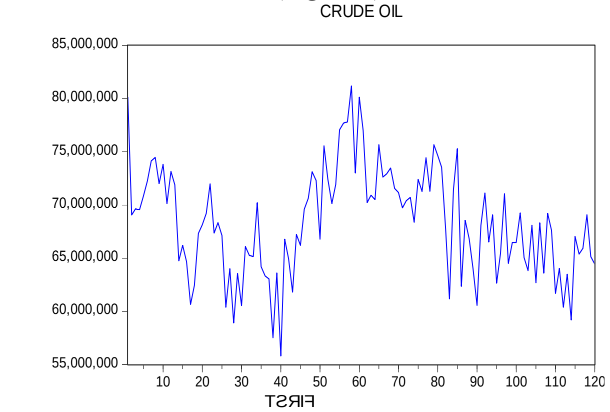
3.2 Stationarity test

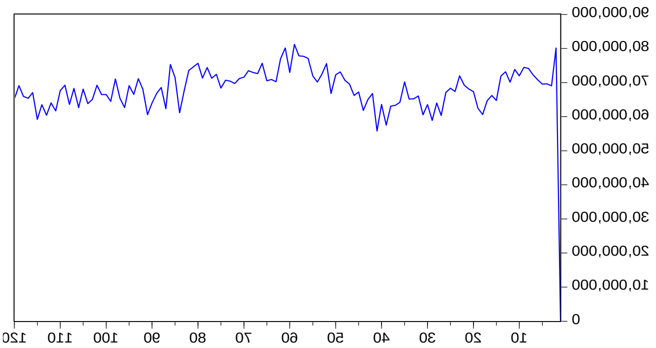
Two approaches were used for the examination of stationarity of the series under study, these are graphical and unit root (ADF). Without examining the stationary properties of the series, we cannot proceed to data analysis stage.

3.2.1 UNIT ROOT TEST

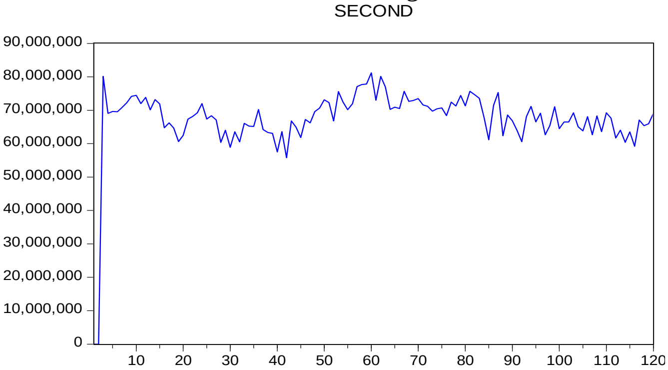
(LEVEL) Figure 2 First difference

(Figure 3)





Second difference Figure 4



Interpretation:

The time plot of the series was shown using Figures 2, 3 and 4. At level and first difference, the series was not stationary as there are evidences of noisy and chaoticness in the series, .that is a very low coefficient of determination at level $\left(R^2 = 0.243\right)$ while at the first difference, the data behaves fairly stationary as it produces a coefficient of determination of $\left(R^2 = 0.762\right)$ and at the second difference, the mean of series became stable as its coefficient of determination accounted for $\left(R^2 = 0.942\right)$, at this point, the series is totally stationary.

3.2.2 UNIT ROOT TEST OF THE ORIGINAL SERIES

Table 1

Null Hypothesis: Ctude oil data has a unit root				
Exogenous: Constant				
Lag Length: 2 (Automatic - based on AIC, maxlag=15)				
			t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic			-4.978454	0.0230
Test critical values:	1% level		-3.454174	
	5% level		-2.871922	
	10% level		-2.572375	
*MacKinnon (1996) one-sided p-values.				

Table 2 UNIT ROOT TEST OF THE FIRST DIFFERENCE

Null Hypothesis: D(FIRSTDIIFF) has a unit root				
Exogenous: Constant				
Lag Length: 12 (Automatic - based on AIC, maxlag=15)				
			t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic			-9.061143	0.0110
Test critical values:	1% level		-3.455289	
	5% level		-2.872413	
	10% level		-2.572638	
*MacKinnon (1996) one-sided p-values.				
Table 3				

UNIT ROOT TEST OF THE SECOND DIFFERENCE

Null Hypothesis: D(SECONDDIFF,2) has a unit root				
Exogenous: Constant				
Lag Length: 15 (Automatic - based on AIC, maxlag=15)				
			t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic			-11.06776	0.0000
Test critical values:	1% level		-3.455786	
	5% level		-2.872630	
	10% level		-2.572754	
*MacKinnon (1996) one-sided p-values.				

Interpretation:

ISSN No:-2456-2165

The p value of the ADF obtained is lesser than 5%, at this level the test procedure only accounted for 24.3% fitness and cannot guarantee proceeding with the data analysis. Since at first difference, the value computed for value of the ADF is lesser than 5% and accounted for 76.2% which is adjudged not to be too good because of the level of significance and at the second difference, the value obtained for the ADF is lesser than 5% and accounted for 94.4% coefficient of determination which is adjudged to be a good fit, hence suggesting that the the crude oil data is stationary.

3.2.3 Model fitting

The descriptive statistic shown in figure 1 shows the crude oil distribution of the crude oil data as normal in nature, implying that there is presence of volatility in the series. The conditional mean was modeled using autoregressive process AR(1) and moving average MA(1) processes. To go about this, the greed search table was constructed which denoted various stages of Akaike information criterion (AIC) from where optimal AIC was selected for forecasting as revealed by table 4 below: -

Table 4GRID SEARCH TABLE (**Table 4**)

AR / MA	1	2	3	4	5	6
1	2.0192	2.1403	2.1574	2.1609	2.1515	2.1386
2	2.4188	3.1525	3.1388	3.1562	3.1519	3.1304
3	2.6840	3.1293	2.6840	3.1496	3.1431	3.1226
4	2.7191	3.1500	3.1475	3.1431	3.1511	3.1285
5	2.7193	3.1507	3.1448	3.1566	3.1333	3.1304
6	2.7172	3.1458	3.1387	3.1504	3.1491	3.1243

Interpretation:

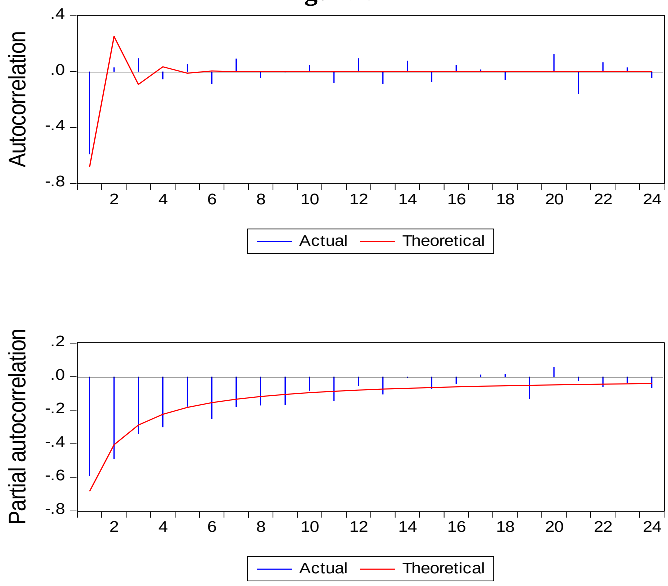
In **Table 4** above, AR and MA were combined to search for the model adequacy, for the combinations at different levels, $ARMA(1,1)$ produced the best fit on the basis of Akaike Information Criterion. The results of $ARMA(1,1)$ are shown in table 5 below:

Table 5

Dependent Variable: SECONDDIFF				
Method: Least Squares				
Convergence achieved after 15 iterations				
MA Backcast: 1992M03				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(1)	-0.366709	0.056505	-6.489885	0.0000
MA(1)	-0.995374	0.004664	-213.3965	0.0000
R-squared	0.680677	Mean dependent var		0.003663
Adjusted R-squared	0.679498	S.D. dependent var		5.212752
S.E. of regression	2.951088	Akaike info criterion		5.009524
Sum squared resid	2360.117	Schwarz criterion		5.035967
Log likelihood	-681.8000	Hannan-Quinn criter.		5.020138
Durbin-Watson stat	2.201807			
Inverted AR Roots	-.37			
Inverted MA Roots	1.00			

The correlogram of the above $ARMA(1,1)$ residuals is given in Figure 5 below: -

Figure 5



Interpretation:

From figure 5 above, is the combined figures of correlogram of both ACF and PACF of the series under study. Form this figure, it is clear that the residuals estimated are random, and that further search of another parsimonious model of ARMA is of no need.

3.2.4 Forecast analysis

Table 6

RMSE	4.989259
MAE	3.89432
MAPE	86.27991
MAD	23.1021
Theil U	0.9243
BIAS PRO	0.0027
VAR. PRO	0.0032
COVAR. PRO	0.99608

Interpretation

Theil-U value of 0.9243 computed revealed excellent model fit, so also the bias and variance proportions computed (0.0027) and (0.0032) are nearer zero showing that the series being student contain a very small error and indicated a measure of goodness of fit and can be used for forecasting and lastly the covariance proportion (0.99608) computed is approximately one, and is regarded as a good value that will enhance accurate forecasting.

IV. CONCLUSION

The study dwells on empirical investigation of efficiency measurement in financial and economic time series data Nigerian crude oil data was used for empirical illustration. Having Subjected the series to stationarity test using graphical and unit root (Augmented Dickey Fuller) approaches, the series was not stationary at both level and first difference. However, at second difference it was stationarity as the chaotic, noisy and volatility was brought to the stable level. Thereafter, the data was analysed and model fitted, from which $ARMA(1,1)$ was selected as it produced the least Akaike Information Criterion (AIC) value. The forecast evaluation revealed that Theil-U and its components produced the best measure, closely followed by Mean Absolute Error (MAE) and Root Mean Square Error $(RMSE)$ successively.

To would be analyst and forecaster the use of Theil-U and its components are strictly recommended but, in its stead, Mean Absolute Error (MAE) and Root Mean Square Error $(RMSE)$ could be used in that order.

REFERENCES

[1]. Achor, S., Reece, A.,Kellerman,G. R. and Robichaux, A. (2018) 9 out of 10 people are willing to earn less money to do more-meaningful work. *Harv. Bus. Rev.*, Nov.

[2]. Ahlburg, D. (1995). Simple versus Complex Models: Evaluation, Accuracy, and Combining. Mathematical Population Studies 5: 281-90.

[3]. Allin, P. and Hand, D. (2017) New statistics for old?—measuring the wellbeing of the UK (with discussion). *J. R. Statist. Soc.* A, **180**, 3–43

[4]. Babu, C.J. and Roa, C.R., (1992). Expansions for statistics involving the mean absolute deviation. Annals of the Institute of Statistical Mathematics, 44, 387-403.

[5]. Barber, M. (2017) Delivering better outcomes for citizens: practical steps for unlocking public value. (Available from https://www.gov.uk/government/publications/delivering-better-outcomes-for citizens- practical-steps-for-unlocking-public-value.)

[6]. Białowolski, P., McNeely, E., VanderWeele, T. J. and We̜ziak-Białowolska, D. (2018) Costs and drivers for productivity loss—turning around ill health and distraction at work in a US manufacturing company. *Technical Report*. Harvard University, Cambridge.

[7]. Biemer, P., Trewin, D., Bergdahl, H. and Japec, L. (2014) A system for managing the quality of official statistics. *J. Off. Statist.*, **30**, 381–415.

[8]. Campbell, P. (2002). Evaluating Forecasting Error in State Population Projections Using Census 2000 Counts. Population Division Working Paper Series, no. 57. Washington DC: US Bureau of the Census

[9]. Chu, T. W., and A. Shirmohammadi. 2004. Evaluation of the SWAT model's hydrology component in the Piedmont physiographic region of Maryland. Trans. ASAE 47(4): 1057−1073.

[10]. Fisher, N. I. (2013) *Analytics for Leaders: a Performance Measurement System for Business Success*. Cambridge: Cambridge University Press.

[11]. Ghurye,W. Hoeffding, W. G. Madow and H. B. Mann, eds.) 448–485. Stanford Univ. Press.

[12]. Hoaglin, D.C., Mosteller, F. and Tukey, J.W. (2000) Understanding Robust and Exploratory Data Analysis. 1st Ed., John Wiley and Sons, New York.

[13]. Hyndman, R. and Koehler, A. (2006). Another Look at Measures of Forecast Accuracy. International Journal of Forecasting 22: 679-688.

[14]. Isserman, A. (1977). The Accuracy of Population Projections for Sub- County Areas. Journal of the American Institute for Planners 43: 247-59.

[15]. Marona, R.A., Martin, D.R. and Yohai, V.J., (2006). Robust statistics: Theory and methods. John Wiley & sons

[16]. Miller, E. R. (2001). Evaluation of the 1990 School District Level Population Estimates Based on the Synthetic Ratio Approach. Population Division Working Paper No. 54. Washington D.C.: US Census Bureau.

[17]. Murdock, S., L. Leistritz, R. Hamm, S. Hwang, and B. Parpia. (1984). An Assessment of the Accuracy of a Regional Economic-Demographic Projection Model. Demography 21: 383-404.

[18]. Pham-Gia, T. and Hung, T.L. (2001) The mean and median absolute deviations. Mathematical and Computer Modelling, 34, 921-936.

[19]. Rayer, S. (2007). Population Forecast Accuracy: Does the Choice of Summary Measure of Error Matter?" Population Research and Policy Review 26: 163-184.

ISSN No:-2456-2165

[20]. Sink, L. (1997). Race and Ethnicity Classification Consistency between the Census Bureau and the National Center for Health Statistics. Population Division Working Paper No. 17. Washington, D.C.: U.S. Census Bureau.

[21]. Smith, S. and Sincich T. (1990). On the Relationship Between Length of Base Period and Population Forecast Errors. Journal of the American Statistical Association 85: 367- 75.

[22]. Smith, S. and Sincich T. (2001). Evaluating the Forecast Accuracy and Bias of Alternative Projections for States. International Journal of Forecasting 8: 495-508.

[23]. Tayman, J. E. Schafer, and L. Carter. 1998. The role of population size in the determination of population forecast errors: An evaluation using confidence intervals for subcounty areas. Population Research and Policy Review 17: 1-20.

[24]. Vander-Weele, T. J. (2017) On the promotion of human flourishing. *Proc. Natn. Acad. Sci. USA*, **31**, 8148–8156.

[25]. Vazquez-Amabile, G.G., and B.A. Engel. 2005. Use of SWAT to Compute Groundwater Table Depth and Streamflow in the Muscatatuck River Watershed. Trans. of the ASAE 48(3):991-1003.

[26]. Weziak-Białowolska, D., Koosed,T., Leon,C. and McNeely, E. (2018) A new approach to the well-being of factory workers in global supply chains: evidence from apparel factories in Mexico, Sri Lanka, China and Cambodia. In *Measuring the Impacts of Business on Well-being and Sustainability* (ed. M. Durand), pp. 130–154. Paris: Organisation for Economic Co-operation and Development.

[27]. Wilson. T. (2007). The Forecast Accuracy of Australian Bureau of Statistics National

[28]. Population Projections. Journal of Population Research 24 (1): 91-117.