On the Predictive Ability of Time-Domain Modeling of Long Memory Data

N. O. Adeboye[1]; O. N. Ogunnusi

Department of Mathematics & Statistics,
Federal Polytechnic, Ilaro, Nigeria.
E-mail: nureni.adeboye@federalpolyilaro.edu.ng[1]

***Abstract* — The study of long memory data required the fitting of an appropriate time-domain model(s) which can be used to achieve a high level of precision in the forecast. To this end, Autoregressive Fractional Integrated Moving Average (ARFIMA) and Seasonal Autoregressive Integrated Moving Average (SARIMA) models were considered with special focus on the comparative predictive ability of the two techniques, using the Nigerian Stock Exchange All-share Index (ASI) as a study. The ASI series was subjected to a Unit Root test using Augmented Dickey-Fuller (ADF) approach and cross-examination of the ACF showed the presence of a long memory structure, which was confirmed using the Hurst exponent test. The Geweke and Porter-Hudak (GPH) method of estimation was used to obtain the long memory parameter of the ARFIMA model while the SARIMA model was also fitted for the ASI. However, based on minimum AIC and Maximum log-likelihoods, ARFIMA(4,0.204,1) and SARIMA (4,1,1)x(1,1,1)12 were found to be the best from several iterated models. Forecast evaluations of the best-fitted models were carried out using MAE, RMSE, and MAPE respectively. Results indicated that the SARIMA model was much better in prediction as against most established literature of ARFIMA's superiority in the modelling of long memory data.**

Keywords-*All-Share Index, ARFIMA, SARIMA, Autocorrelation Function, Predictive Ability, Long-memory.*

I. Introduction

The presence of long memory is very obvious in the modeling of stock market time-series data such as the All-share index (ASI). The behaviors of the investors are influenced, which can make their decisions depend on different investment horizons. According to Baillie (1996), stock market data has been found to exhibit characteristics that are more consistent with long memory. Long memory in time series is described as autocorrelation at long lags (Robinson, 2003). One of the key points explained by Peters (1991) is the fact that most financial markets have a very long memory property. In other words, what happens today affects the future forever. This indicates that current data correlated with all past data to varying degrees. This long memory component of the market cannot be adequately explained by systems that work with short memory parameters.

The empirical evidence of the long memory processes goes back to Hurst (1951) in the field of hydrology. However, interest in long memory models for economic series arises from the works of Granger and Joyeux (1980), who noted that many of such series are apparently not stationary in mean, and yet, the differentiated series usually present clear evidence of over-differencing. The property of long memory is usually related to the persistence that is shown by the sample autocorrelations of certain stationary time series, which decrease at a very slow rate, but finally converge towards zero, indicating that the innovations of these series have transient effects but last for a long time. This behavior is not compatible neither with the stationary models, which have an exponential decrease in autocorrelations and therefore in effects of the innovations1 nor with the integrated models, where innovations have permanent effects.

Previous authors such as Granger and Joyeux (1980) and Hosking (1981, 1984) have modeled time series in the presence of long memory, using the Autoregressive fractionally integrated moving average (ARFIMA) model. ARFIMA (p, d, q) model, a class of long memory models are time series models that generalize ARIMA models by allowing non-integer values of the differencing parameter, where d is a real number. By allowing the order of integration, d, to be a non-integer number, these models act as a "bridge" between the processes with ARIMA unit-roots ($d = 1$) and stationary ARMA processes ($d = 0$). When $0 < d < 0.5$, the ARFIMA processes are stationary, that is, its mean level is constant, but deviations from the series over this level have a longer duration than when $d = 0$. The ARFIMA model searches for a non-integer parameter, d, to differentiate the data to capture long memory. The useful entry points to the literature are the surveys by Robinson, (2003) and Baillie (1996), who have described the development in the modeling of long memory on financial data, and of Bearn, (1995) who have reviewed long memory modeling in other areas. The existence of non-zero d is an indication of long memory and its departure from zero measures the strength of the long memory.

Similarly, from the ARIMA scheme's perspective of forecasting the Nigerian stock market returns, Ojo and Olatayo (2009) studied the estimation and performance of subset autoregressive integrated moving average (ARIMA) models. The estimated parameters for ARIMA and subset ARIMA processes using numerical iterative schemes of Newton-Raphson and the Marquardt Levenberg algorithms. The performance of the models and their residual variance was examined using AIC and BIC. The results of their study showed that the SARIMA model outperformed the ARIMA model with a smaller residual variance.

In addition, Emenike (2010) studied the NSE market returns series using monthly data of the All-Share-Index for the periods January 1985 through December 2008. In his study, an ARIMA (1,1,1) model was selected as a tentative model for predicting index points and growth rates. The results revealed that the global meltdown destroyed the correlation structure existing between the NSE All-Share-Index and its past values. In the same vein, Adeboye and Fagoyinbo (2017) equally modeled the index using the SARIMA model and forecasted future values using the best-fitted model of order $(2,1,1)x(0,1,1)_{12}$.

The main objective of this paper is to explicitly account for persistence to incorporate the long-term correlations in the series using ARFIMA and SARIMA techniques. ARFIMA model has been widely acclaimed as a useful tool in modeling time series with a long memory. That is a series in which deviations from the long-run mean decay more slowly than an exponential decay. SARIMA model which is a generalization of the Seasonal Autoregressive Integrated Moving Average (SARIMA) model, will also be employed as an alternative model in this study. This model is applied in some cases where data show evidence of non-stationarity, where an initial differencing step can be applied to remove stationarity. The model is generally referred to as ARIMA (p, d, q) x(P, D, Q)s model with seasonal period s, where p, d and q are non-negative integers that refer to the order of the autoregressive, integrated and moving average parts of the model and P, D, Q are also non-negative integers that refer to the order of the Seasonal Autoregressive, seasonal integrated and Seasonal moving average parts respectively.

It has been established that previous studies have not employed the ARFIMA technique to model NSE ASI, hence the need for this study.

II. Materials And Methods

In this section, the procedures for building ARFIMA(p,d,q) and SARIMA models for this research are discussed. Monthly data between January 1991 to December 2018 of the NSE ASI are used in the study. The data was sourced from Central Bank of Nigeria statistical bulletin. The data were divided into test and train. The trained data made up about 96.3% of the entire series with the test capturing 3.7%.

The general expression for ARFIMA processes X_t is defined by:

$$\Phi(B) = \theta(B)(1-B)^{-d}\varepsilon_t \quad (1)$$

where

$$\Phi(B) = 1 - \Phi_1 B - \cdots \Phi_p B^p \quad (2)$$

and

$$\theta(B) = 1 + \theta_1 B + \cdots + \theta_q B^q \quad (3)$$

p and q are the autoregressive (AR) and moving average (MA) operators respectively; B is the backward shift operator and $(1-B)^{-d}$ is the fractional differencing operator given by the binomial expression;

$$(1-B)^{-d} = \sum_{j=0}^{\infty} \frac{\Gamma(j+d)}{\Gamma(j+1)\Gamma(d)} B^j = \sum_{j=0}^{\infty} n_j B^j \quad (4)$$

Short memory system such as the generalized Seasonal ARIMA model for univariate analysis is defined by the equation

$$\emptyset(B)\Phi(B)(1-B)(1-B^{12})X_t = \theta(B)\Theta(B^{12})\varepsilon_t \quad (5)$$

where B denotes the backward shift operator, $\emptyset, \Phi, \theta$, and Θ are polynomials for order p, P, q, and Q respectively. X_t is the observed time series and ε_t represent an unobserved white noise series with zero mean and constant variance σ_ε^2.

The presence of a long memory process is tested on the data using the Hurst exponent. When the integration parameter d in an ARIMA process is fractional and greater than zero, the process exhibits long memory (Granger and Joyeux, 1980). In the work of McLeod and Hipel (1978), the process possess long memory if the quantity $\lim_{n\to\infty} \sum_{-n}^{n} |\rho_j|$ is nonfinite.

In modeling ARFIMA and SARIMA models, the variables are first examined for stationarity. The Augmented Dickey-Fuller (ADF) test is used for this purpose. This preliminary test is necessary in order to determine the order of non-stationarity of the data.

The ADF regression equation due to Dickey and Fuller (1979), Said and Dickey (1984) is given by:

$$\Delta_1 X_t = \rho X_{t-1} + \alpha_1 \Delta_1 X_{t-1} + \varepsilon_t \quad (6)$$

where Δ_1 denotes the differencing operator i.e.

$\Delta_1 X_t = X_t - X_{t-1}$

The relevant null hypothesis is ρ = 0 i.e. the original series is non-stationary and the alternative is ρ > 0 i.e. the original series is stationary. Usually, differencing is applied until the ACF shows an interpretable pattern with only a few significant autocorrelations.

Testing for Long Memory

The Hurst exponent was produced by a British hydrologist Harold Hurst in 1951 to test the presence of long memory using the rescaled range analysis approach. The main idea behind the R/S analysis is that one looks at the scaling behavior of the rescaled cumulative deviations from the mean. The R/S analysis first estimates the range R for a given n, where:

$$R_{(n)} = \max \sum_{j=1}^{n} (X_j - \bar{X}) - \min \sum_{j=1}^{n} (X_j - \bar{X}) \qquad (7)$$

$R_{(n)}$ is the range of accumulated deviation of $X_{(t)}$ over the period of n and $\bar{X}$ is the overall mean of the time series. Let $S_{(n)}$ be the standard deviation of $X_{(t)}$ over the period of n. This is an indication that;

$$R/S = \frac{R_{(n)}}{S_{(n)}} \qquad (8)$$

whereas, as n increases, equation 7 holds:

$$\log\left(\frac{R_{(n)}}{S_{(n)}}\right) = \log\alpha + H\log n \qquad (9)$$

From equation (9), the estimate of the Hurst exponent (H) is the slope, where H is a parameter that relates mean R/S values for subsamples of equal length of the series to the number of observations within each equal length subsample (Omekara *et al,* 2016). H is usually greater than 0. However, when 0 < H < 1, the long memory structure exists. If H ≥ 1, the process has infinite variance and is non-stationary. If 0 < H < 0.5, anti-persistence structure exists. If H = 0.5, the process is white noise.

In estimating the long memory parameter, Geweke Porter-Hudak (1983) hereinafter referred to as GPH was used in this present investigation. This method is based on approximated regression equation obtained from the logarithm of the spectral density function. Illustrating this method, the spectral density function of a stationary model $X_t, t = 1, \ldots T$ is written as

$$f_x(\lambda) = \left[4sin^2\left(\frac{\lambda}{2}\right)\right]^{-d} f_\varepsilon(\lambda) \qquad (10)$$

where $f_\varepsilon(\lambda)$ is the spectral density of ε_t, which is assumed to be a finite and continuous function on the interval $[-\pi, \pi]$. The log spectral density function is written as:

$$\log[f_x(\lambda)] = \log[f_\varepsilon(0)] - d\log\left[4sin^2\left(\frac{\lambda}{2}\right)\right] + \log\frac{f_\varepsilon(\lambda)}{f_\varepsilon(0)} \quad (11)$$

where $\log[f_x(\lambda)]$ is a constant; $\log\left[4sin^2\left(\frac{\lambda}{2}\right)\right]$ is the exogenous variable; and $\log\frac{f_\varepsilon(\lambda)}{f_\varepsilon(0)}$ is a disturbance error term. The estimate of d is given by equation 12 below

$$d_{GPH} = \frac{\sum_{j=1}^{m}(x_j - \bar{x})\log[Ix(\lambda_j)]}{\sum_{j=1}^{m}(x_j - \bar{x})^2} \qquad (12)$$

where, $\bar{x} = \frac{\sum_{j=1}^{m} x_j}{m}$

GPH has been analyzed in detail by Tanaka (1999); Hurvich (2002); Robinson (2003), and It has been proved in detail that the estimate is consistent and asymptotically normal under the assumption of series x_t.

The model's residual diagnostic checks employed are Ljung-Box Test, The Akaike Information Criterion (AIC), Root Mean Square Error (RMSE).

III. Analysis and Results

The pattern of time plot is shown in fig.1 below affirmed the series not to be from a normal distribution and clear evidence of the presence of unit root and seasonal variation.

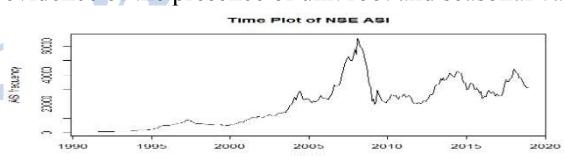

Fig.1: Time Plot of ASI

Table 1: Descriptive Statistics of NSE ASI

Descriptive Stat.	NSE ASI
Minimum	528.7
1st Quarter	5873.3
Median	21486.6
Mean	19625.8
3rd Quarter	29257.9
Maximum	508.2
Skewness	0.54978
Kurtosis	2.60476
**JB statistics:	χ^2 = 19.113; df=2; P-value = 0 .000

*****represent Jarque Bera normality test***

Source: R-Studio Output

Descriptive statistics of the series which can be evidenced from table 1 showed the minimum, mean and maximum all-share index series computed for the periods. It can also be seen that the skewness, kurtosis and JB statistics does not fully correspond to independent samples from a normal distribution but are serially correlated.

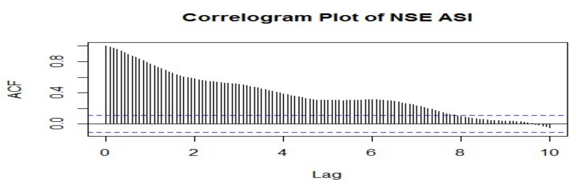

Fig.2: ACF Plot of ASI

Based on the fact that the autocorrelation of the series has exponential decay towards zero as shown in figure 2, the series was subjected to long memory test and a précised order of differencing was determined due to the presence of unit root.

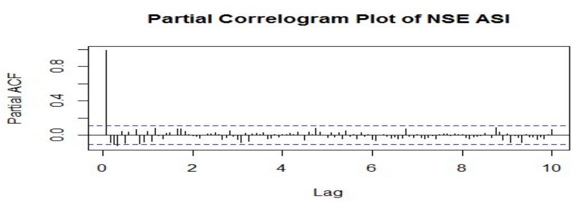

Fig.3: PACF Plot of ASI

Fig.3 described the behavior of the partial autocorrelation as it also showed that it is almost zero from lag to lag. Hence, adjudged the presence of unit root.

Table 2: Long memory test and "*d*" parameter estimation results

Coefficient	Estimate	Asymp.Std. Dev	Std. error deviation
fdGPH "*d*"	0.203885	0.1952754	0.2783841

Hurst Exponent R/S estimation = 0.852844

Hurst exponent (H) produced by the Rescaled range analysis was obtained to be 0.852844 indicating that the All-share index data has a long memory structure since 0.5<H<1. Taking the estimated value of the parameter "d" into consideration, its asymptotic deviation value and regression standard deviation values are reported in table 3 which indicated an estimate of 0.203885.

After estimating the long memory parameter *d*, the degree of autocorrelation in the fractionally differenced ASI is examined using the autocorrelation and partial autocorrelation function as shown in figure 4 and 5.

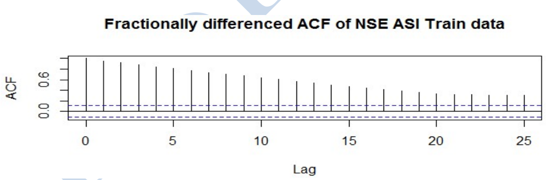

Fig.4: ACF Plot of Fractionally Differenced ASI

Fig.4 showed evidence of MA from lag 1 to 25 since the spikes of autocorrelation are above the upper bound and was found to be significant.

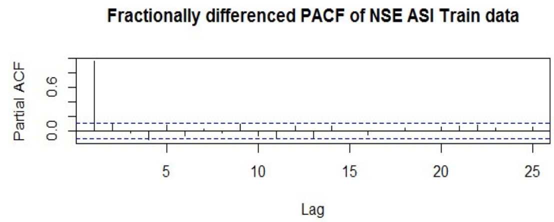

Fig.5: PACF Plot of Fractionally Differenced ASI

The partial ACF of the fractionally differenced ASI in fig.5 also indicated the presence of Autoregressive terms in the model since some of the identified PACF were above the upper and lower bound.

Table 3: Train Data Iterated ARFIMA Models *with* σ^2, AIC, BIC and Log Likelihood results

Model	σ^2	AIC	BIC	Log likelihoods
ARFIMA (2, 0.204, 2)	3001967	4846.09	4872.56	-2416.05
ARFIMA (3, 0.204, 2)	2858426	4831.73	4861.98	-2407.87
ARFIMA (2, 0.204, 1)	3024911	4847.52	4870.20	-2417.76
ARFIMA (4, 0.204, 1)	**2832308**	**4827.43**	**4858.68**	**-2406.22**
ARFIMA (3, 0.204, 2)	2837933	4830.09	4864.11	-2406.04

Source: R-Studio Output

The iterated ARFIMA models in table 3 indicated that ARFIMA(4, 0.204, 1) was found to be the best from the several fitted models with the lowest variance, AIC, and BIC of 2832308, 4827.45, and 4858.68 respectively, and a higher log-likelihood of -2406.22.

Table 4: Results of the estimated ARFIMA (4, 0.204, 1) model parameters

Parameters	Estimates	Z-value	Pr(>|z|)
AR(1) ≡ ϕ_1	0.281097	2.45595	0.0141
AR(2) ≡ ϕ_2	0.070767	6.87034	0.0000
AR(3) ≡ ϕ_3	0.201889	3.68046	0.0002
AR(4) ≡ ϕ_4	-0.24186	-5.74358	0.0000
MA(1) ≡ θ_1	-0.630882	-5.74718	0.0000

Source: R-Studio Output

The estimated coefficients of the identified model are fitted thus:

$$X_t = \varepsilon_t - \theta_1\varepsilon_{t-1} + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \phi_3 X_{t-3} + \phi_4 X_{t-4} \quad (13)$$

$$X_t = \varepsilon_t + 0.630882\varepsilon_{t-1} + 0.281097X_{t-1} + 0.070767X_{t-2} + 0.201889X_{t-3} + 0.24186X_{t-4} \quad (14)$$

Having fitted the ASI series using the fractionally integrated moving average technique, it can be seen that the coefficients of the best-identified model were found to significantly contribute in predicting the out-of-sample data of the series.

Model Identification Using Generalized Box Jenkins SARIMA (p,d,q)X(P,D,Q)12

A graphical inspection of the plot of the original series as conducted in fig.1, fig.2 and fig. 3 does not show any evidence of stationarity. Differencing at order 1 yielded a stationary series as shown in Figs. 6, 7, and 8.

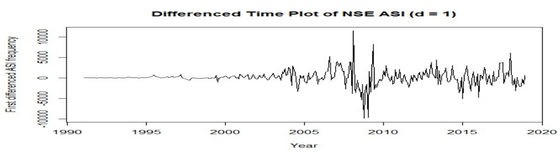

Fig.6: Time Plot of 1st Order Differenced ASI

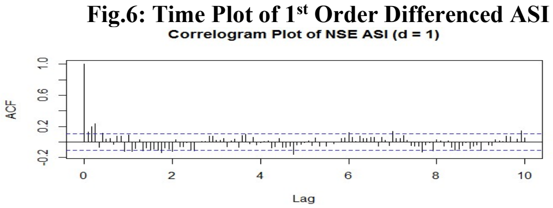

Fig.7: ACF Plot of 1st Order Differenced ASI

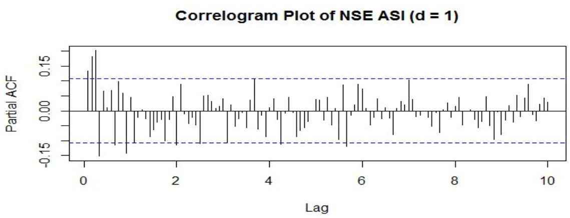

Fig.8: PACF Plot of 1st Order Differenced ASI

Table 5: ADF test for unit root result

Test	*Lag order*	*p-value 5%*	*Test Statistic*
@Level	6	0.1652	-2.9753
@1st Difference	6	0.01	-5.2938

Source: R-Studio Output

Table 5 contains the summary of the results of the ADF test at level and at first differencing which is also a confirmation that the series is stationary at first order differencing because of its p-value of $0.01 < \alpha = 0.05$. Since the series is stationary at order 1 integration, the order of MA is determined from the ACF plot, order of AR at the PACF plot.

Table 6: Train data Iterated SARIMA Models with σ^2, AIC, and Log Likelihood results

Model	σ^2	AIC	Log likelihood
SARIMA (3,1,1)x(1,1,0)12	4211916	5640.69	-2814.34
SARIMA (1,1,3)x(1,1,0)12	4287296	5646.03	-2817.02
SARIMA (4,1,1)x(1,1,1)12	**2770919**	**5542.52**	**-2763.26**
SARIMA (2, 1,2)x(1,1,0)12	4367528	5651.70	-2819.85
SARIMA (1, 1,1)x(1,1,0)12	4392972	5649.64	-2820.64

Source: R-Studio Output

Table 7: Parameter Results of the estimated SARIMA (4, 1, 1)x(1,1,1)12 Model

Parameters	Estimates	Z-value	Pr(>|z|)
AR(1) $\equiv \phi_1$	-0.4140	1.56522	0.0541
AR(2) $\equiv \phi_2$	0.2244	3.67267	0.0000
AR(3) $\equiv \phi_3$	0.2975	4.28057	0.0000
AR(4) $\equiv \phi_3$	-0.0256	0.27033	0.3532
MA(1) $\equiv \theta_1$	0.5471	2.11073	0.0000
SAR(1) $\equiv \Phi_1$	0.0819	1.34925	0.0865
SMA(1) $\equiv \Theta_1$	-1.0000	17.6366	0.0000

From the iterated SARIMA models in table 6, the confirmatory analysis indicated that SARIMA (4, 1, 1)x(1,1,1)12 was found to be the best from several iterated models. The model has the lowest variance and AIC of 2770919 and 5542.52 respectively with the highest log-likelihood of -2763.26.

The model specification for identified SARIMA(4,1,1)x(1,1,1)12 in table 7 is written as;

$$Y_t = \varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2} + \Theta_1\varepsilon_{t-12} - \phi_1 Y_{t-1} - \phi_2 Y_{t-2} - \phi_3 Y_{t-3} - \phi_4 Y_{t-4} - \Phi_1 Y_{t-1} \quad (15)$$

Substituting the coefficients, we have;

$$Y_t = \varepsilon_t + 0.5471\varepsilon_{t-1} - \varepsilon_{t-1} + 0.4140Y_{t-1} - 0.2244Y_{t-2} - 0.2975Y_{t-3} + 0.0256Y_{t-4} - 0.0819Y_{t-12} \quad (16)$$

The adequacies of the two identified models were subjected to a diagnostic check using Ljung-Box test as shown below:

Table 8: Summary of Diagnostic Check for the Two Best Identified Models

Model Type	Model	Chi-Squared	DF	P-value
ARFIMA	ARFIMA (4, 0.204, 1)	0.00167	1	0.9969
SARIMA	SARIMA (4, 1,1)x(1,1,1)12	0.0002	1	0.9674

Source: R-Studio Output

The small values of Chi-square statistic with corresponding larger p-values > 5% significance level in the Ljung-Box test for both ARFIMA and SARIMA models suggested the failure in rejecting the null hypothesis that all of the autocorrelation functions are zero. In other words, we can conclude that there is no evidence for non-zero autocorrelations in the residuals of the fitted ARFIMA and SARIMA models. Hence, the ARFIMA and SARIMA models are adjudged to be parsimoniously fitted. This also indicates that the models captured better the dependence in the series and can be confirmed to be adequate for predicting the Nigerian Stock Exchange All-share index.

Evaluation of Predictions

Predicted values of the fitted ARFIMA(4,0.204,1) and SARIMA(4,1,1)x(1,1,1)12 were studied using the test data after an adequacy check of the models have been done.

From the prediction performances results of the best-fitted ARFIMA and SARIMA models above, the predicted values of ARFIMA(4,0.204,1) are not closer to the observed values from January to April 2018 compared to the predicted values of SARIMA(4,1,1)x(1,1,1)$_{12}$. Other subsequent months showed a closer and better performance of the ARFIMA model in predicting the All-share index but the Seasonal ARIMA model outperforms the former based on their prediction accuracy.

Table 9: Prediction performance of the best fitted ARFIMA and SARIMA models

Date	**Test Data (Observed values)**	**ARFIMA** (4,0.204,1)	**SARIMA** (4,1,1)x(1,1,1)$_{12}$
Jan 2018	44343.65	38293.51	38100.85
Feb 2018	43330.54	38449.59	38951.21
Mar 2018	41504.51	38185.25	38776.91
Apr 2018	41268.01	38088.60	39435.58
May 2018	38104.54	37842.75	40759.14
Jun 2018	38278.55	37575.28	41270.76
Jul 2018	37017.78	37338.84	41578.04
Aug 2018	34848.45	37031.21	40939.09
Sep 2018	32766.37	36760.54	40976.87
Oct 2018	32466.27	36464.35	40828.70
Nov 2018	30874.17	36167.39	40730.00
Dec 2018	31430.5	35878.35	41228.12
	MAE	*3219.33*	*956.1848*
	RMSE	*3727.423*	*1630.871*
	MAPE	*0.088388*	*4.110131*

Source: R-Studio Output

Performances of the best-fitted models were further evaluated using Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE). The lower values of these accuracy statistics (MAE and RMSE) indicate a better fit considering SARIMA(4,1,1)x(1,1,1)$_{12}$. RMSE is a good measure of how accurately the model predicts the response and is an important criterion for fit if the main purpose of the model is prediction. Considering the RMSE, results showed that the Seasonal ARIMA model predicted better than the ARFIMA model due to its lower RMSE of 1630.871 and higher MAPE of 4.11% compared to its ARFIMA counterpart with a higher RMSE of 3727.423 and lower MAPE of 0.089%

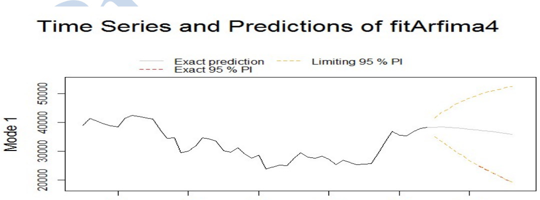

Fig.9: ARFIMA forecasts

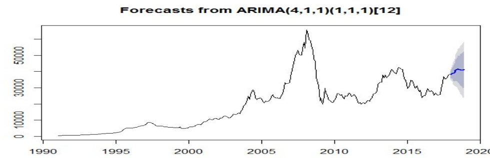

Fig.10: SARIMA forecasts

It is also evident from the predicted series shown in fig. 9 and fig.10 prediction intervals that the upward All-share growth was captured better in the Seasonal ARIMA forecasts compared to its counterpart ARFIMA model.

IV. CONCLUSION

In this paper, efforts to compare the predictive ability of ARFIMA and SARIMA models for long memory data, have been established with dexterity. The autocorrelation function of the All-Share Index (ASI) data showed persistence characteristics with exponential decay towards zero which is one of the features of a long memory process. This led to the fitting of a suitable ARFIMA(4,0.204,1) model. A suitable SARIMA model was equally fitted to capture the presence of seasonal effects in the researched data. The ARFIMA model was not found to be better than the Seasonal ARIMA model as indicated by the model diagnostic tools. The estimated forecast values from ARFIMA model were not of better precision compared to that of SARIMA, as indicated by the forecast evaluation tools applied on models.

Theoretically, it is also not clear if the fractional differenced type of models captures the long-memory tendencies better than the models where the differencing parameter is an integer. Thus, the result of this research is in line with that of Ray (1993); Adeboye and Fagoyinbo (2017) works where the former made comparison between ARFIMA models and standard ARIMA models and the later used SARIMA of order $(2,1,1)x(0,1,1)_{12}$ to model and forecast NES ASI. The results of Ray study showed that higher-order AR models are capable of forecasting the longer term well when compared with ARFIMA models.
The research, therefore, concludes that NSE ASI is better modeled using the generalized SARIMA model than the Autoregressive Fractional Integrated Moving Average.

REFERENCES

Adeboye, N. O., Fagoyinbo, I. S. (2017). Fitting of Seasonal Autoregressive Integrated Moving Average to the Nigerian Stock Exchange Trading Activities. *Edited Proceedings of 1st International Conference, Professional Statistical Society of Nigeria. 2017; 1:12-16.*

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic control, 19(6): 716-723.*

Baillie R. (1996). Long Memory Processes and Fractional Integration in Econometrics. *Vol. 73, issue 1, 5-59*

Beran, J. (1995) *Statistics for Long Memory processes*. Chapman and Hall Publishing Inc., New York.

Box, G.E.P. and Pierce, D.A. (1970). Distribution of the Autocorrelations in Autoregressive Moving Average Time Series Models, *Journal of American Statistical Association*, 65: 1509-1526.

Dickey, D. and Fuller, W. (1979) Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74: 427-431.

Granger C.W.J and Joyeux .R, 1980. An introduction to long memory time series models and fractional differencing. *Journal of Time series Analysis, 15-29*.

Emenike (2010). Modelling stock returns volatility in Nigeria using GARCH Models

Hoskin J.R.M (1981). Fractional Differencing. Biometrika, *68(1), 165 – 176. Doi: 10.2307/2335817*

Hosking, J. (1984). "Modelling Persistence in Hydrological Time Series using Fractional Differencing". Water Resources Research, 1984.

Hurvich, C.M., 2002. Multi-step forecasting of long memory series using fractional exponential models. *International Journal of Forecasting*. 18: 167-179.

Geweke, J. and Porter-Hudak, S. (1983) The estimation and application of long-memory time series models. *Journal of Time series Analysis, 4: 221-238*.

Jarque, C. M., & Bera, A. K. (1980). Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Economics Letters 6*, 255–259.

McLeod, A., & Hipel, K. (1978). Preservation of the rescaled adjusted range, Part 1, A reassessment of the Hurst phenomenon. *Water Resources Research 14*, 491-508

Ojo J.F and Olatayo T.O. (2009). On the Estimation and Performance of Subset Autoregressive Integrated Moving Average models. *Eur. J. Sci. Res., 28: 287 – 293*

Omekara C. O. et al (2016): Forecasting Liquidity Ratio of Commercial Banks in Nigeria. *Journal of Microeconomics and Macroeconomics 16, 4(1): 28-36*

Peters E. E. (1991). Chaos and Order in the Capital Markets. *A Wiley Finance Edition. John Wiley & Sons, New York.*

Ray, B. (1993). Modeling long-memory processes for optimal long-range prediction. *Journal of Time Series Analysis 14*, 511–525.

Robinson P.M. (2003). Time Series with Long Memory. Advanced Text in Econometrics. *Oxford University Press, Oxford England.*

Said, S.E. and Dickey, D. (1984) Testing for unit roots in autoregressive moving average models with unknown order. *Biometrika*, 71: 599-607.

Tanaka, K. (1999) *The nonstationary fractional unit root*. Econometric Theory, 15: 549-582.