Journal of Sustainable Development in Africa (Volume 13, No.5, 2011)
ISSN: 1520-5509
Clarion University of Pennsylvania, Clarion, Pennsylvania

ON THE AUTOREGRESSIVE FRACTIONAL UNIT INTEGRATED MOVING AVERAGE (ARFUIMA) PROCESS

Olanrewaju I. Shittu and OlaOluwa S. Yaya
Department of Statistics, University of Ibadan, Nigeria

ABSTRACT

This paper presents a nonstationary fractional unit integrated moving average process that can model time series data that are not stationary after the first difference. This is autoregressive fractionally unit integrated moving average, $ARFUIMA(p,d,q)$ process with $0.5 < d < 2.5$. The classical unit differencing of Box-Jenkins is combined with the semi-parametric approach to estimate the fractional difference parameter. The model when applied on quarterly Nigerian gross domestic products (GDP) series indicates that ARFUIMA model is better than the corresponding autoregressive integrated moving average (ARIMA) model when compared based on diagnostic tests and forecast.

Keywords: Fractional unit integration, ARFIMA, semi-parametric estimation.

INTRODUCTION

It has been observed that several economic and financial data have been modelled on the assumption that differencing parameter is usually an integer. Although models obtained from integer differencing have been found to be fairly adequate and reliable, however, better models with higher forecast performance can be obtained if appropriate fractional difference parameter is used.

When fractional differencing parameter is non-zero, non-stationarity is suspected. This implies there is strong dependence between distant observations. A number of studies have been carried out on this subject and these include studies on real national product (Diebold et al., 1989); consumer and wholesale price (Geweke and Porter-Hudak, 1982) and stock market prices (Lo, 1991).

The motivation for this paper and presentation is derived from Fatoki, et al., (2010) who investigated the annual GDP from 1980 to 2007. The series was found to be integrated of order 2 $\left[I(2)\right]$ and autoregressive integrated moving average, $ARIMA(1,2,1)$ was fitted using the usual model identification and order determination tools. We are of the opinion that further study could reveal a possibility of the series being integrated of a fractional order and hence could lead to better forecast performance. It is observed that classical ADF unit root test may not give conclusive remark on fractional difference but a test designed for this purpose such as Kwiatkowski, Phillips, Schmidt and Shin (KPSS) test could. Detail about KPSS test can be found in Kwiatkowski, Phillips, Schmidt and Shin (1992). Following Shittu and Yaya (2009), an improved model in terms of parameter estimates and forecasts can be obtained by modelling the remaining long memory in a series after the first or second difference. This paper therefore proposes a class of nonstationary $ARFUIMA(p,d,q)$ where $0.5 < d < 2.5$ process which cater for series that may still be fractionally differenced after first or second differences. The remaining part of this paper is structured as follows: Section 2 presents the model; Section 3 discusses the estimation method and forecast evaluation; Section 4 presents the data analysis results and discussion and Section 5 concludes the work.

THE ARFUIMA PROCESS

Let X_t be any time series process. Fractionally Unit Integrated (FUI) be defined as

$$y_t = (1-L)^{d_0} X_t \tag{1}$$

where X_t is the nonstationary time series, y_t is the covariance stationary process and $(1-L)^{d_0}$ is the difference operator with the fractional unit difference parameter, d_0.

Proposition 1

Suppose a series is nonstationary and can be expressed as $y_t = (1-L)^{d_0} X_t$. If an ADF test of unit root confirmed that the series is $I(u)$, if further fractional differencing of $I(d)$ for $(-0.5 < d < u)$, then the resulting series of $I(d_0)$ where $d_0 = d + u$ is a stationary series.

A fractionally integrated series is invertible and stationary when d is in the interval $(-0.5 < d < 0.5)$ and nonstationary when $d > 0.5$ (Sowell, 1992). When the Augmented Dickey Fuller (ADF) test indicates stationarity at $d = 1$ or $d = 2$, it implies that $d > 2$. When this situation arises, Robinson (1995) and Velasco (1999) indicated that the fractional difference parameter can be obtained from the differenced series. Therefore, setting

$$d = d_0 - u \tag{2}$$

where u is the unit difference parameter, then
equation (1) can be re-written as:

$$y_t = (1-L)^{d_0} X_t \tag{3}$$

$$= (1-L)^{u+d} X_t \tag{4}$$

$$= (1-L)^{u} (1-L)^{d} X_t \tag{5}$$

where u is the unit difference parameter $(u = 1, 2, ...)$ and d is the fractional difference parameter ($-0.5 < d < 1$). From (3), additivity of difference parameters holds and binomial theorem applied on the difference operator results into complex function. Using the Binomial expansion, the fractional difference operator in equation (1) becomes,

$$(1-L)^{d_0} = \sum_{k=0}^{\infty} \binom{d_0}{k} (-1)^k L^k$$

$$= \sum_{k=0}^{\infty} \binom{u+d}{k} (-1)^k L^k ,$$

$$= \sum_{k=0}^{\infty} \frac{\Gamma\left[-(u+d)+k\right]}{\Gamma\left[-(u+d)\right]\Gamma(k+1)} L^k \qquad (6)$$

The above expression will hold under the assumption that the fractionally differenced series is still stationary in the interval $(-0.5 < d_0 < u+d)$. The process in (4) can then be re-written as,

$$y_t = \sum_{k=0}^{\infty} \frac{\Gamma\left[-(u+d)+k\right]}{\Gamma\left[-(u+d)\right]\Gamma(k+1)} X_{t-j} \qquad (7)$$

When $u = 1$,

$$y_t = \sum_{k=0}^{\infty} \frac{\Gamma\left[-(1+d)+k\right]}{\Gamma\left[-(1+d)\right]\Gamma(k+1)} X_{t-j} \qquad (8)$$

When $u = 2$,

$$y_t = \sum_{k=0}^{\infty} \frac{\Gamma\left[-(2+d)+k\right]}{\Gamma\left[-(2+d)\right]\Gamma(k+1)} X_{t-j} \qquad (9)$$

Thus u can take values 1,2,3,4,..., however it rarely gets beyond 2.

METHODOLOGY

We consider estimation of fractional difference parameter, d_0 in the $ARFUIMA(p, d_0, q)$ model using the frequency domain approach described in Geweke and Porter-Hudak (1983) (GPH) and applied in Yaya and Shittu (2010). The definition used in (4) above still applies as 'differencing and adding back' method (Velasco, 2005). The nonstationary series, X_t is then differenced, u times in order to guarantee that the true d is in $-0.5 < d < u$ according to proposition 1.

Once the value of d is determined, y_t is approximated by using $\hat{d}_0 = \hat{u} + \hat{d}$ in,

$$\tilde{y}_t = \sum_{j=t}^{\infty} \frac{\Gamma\left[-\left(\hat{u}+\hat{d}\right)+j\right]}{\Gamma\left[-\left(\hat{u}+\hat{d}\right)\right]\Gamma(j+1)} X_{t-j}$$

$$= \sum_{k=0}^{\infty} \frac{\Gamma\left(-\left(\hat{u}+\hat{d}\right)+t+k\right)}{\Gamma\left(-\left(\hat{u}+\hat{d}\right)\right)\Gamma(t+k+1)} X_k \qquad (10)$$

Forecasts Evaluation

Evaluation of forecast performance of the models will be judged on Meese and Rogoff (1988), MR statistic. The criterion uses the ratio of the root mean square prediction error (RMSPE) of one of the models to the model to the given base model to check the statistical significance. The MR statistic defined as:

$$MR = \frac{\bar{s}_{UV}}{\sqrt{\frac{1}{n^2}\sum_{j=1}^{n} u_j^2 v_j^2}} \quad (11)$$

is asymptotically normally distributed with mean zero and variance one where n is the number of forecasts generated, u and v are transformed functions of forecast errors of the two models; $\bar{s}_{UV}$ is the sample covariance of the means of U and V, approximated by $\bar{s}_{UV} = \frac{1}{n}\sum_{j=1}^{n}\left(u_j - \bar{u}\right)\left(v_j - \bar{v}\right)$ where $\bar{u} = \frac{1}{n}\sum_{j=1}^{n} u_j$ and $\bar{v} = \frac{1}{n}\sum_{j=1}^{n} v_j$ with $u_j = e_{1j} - e_{2j}$ and $v_j = e_{1j} + e_{2j}$ in which e_{2j}, $i = 1,2$ is the j^{th} forecast error of the model i and n is the number of forecasts. The null hypothesis of MR statistic is cov (U, V) = 0. When $MR > Z_{\alpha/2}$, null hypothesis is rejected. This implies that forecast accuracy in the first model is significantly better than the second model. The test is most reliable when n is large.

RESULTS AND DISCUSSION

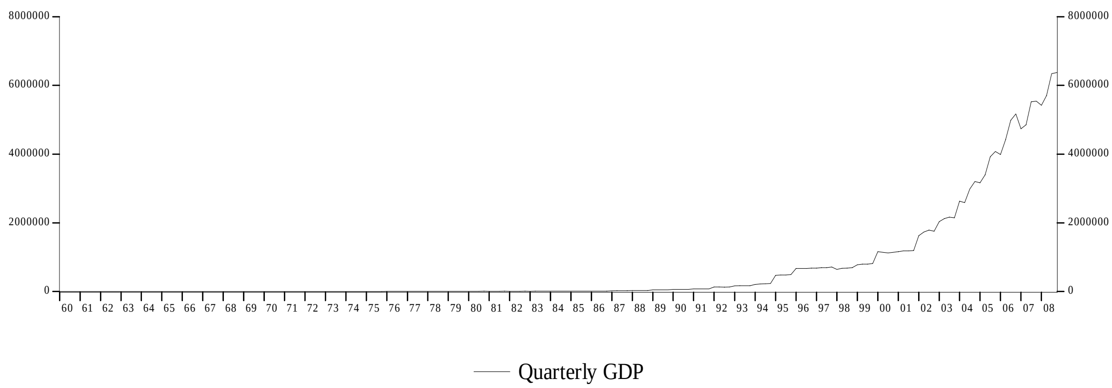
Quarterly Nigeria Gross Domestic Products (GDP) data were used to illustrate the proposed model. The data span from 1960 to 2008 with 196 data points. Even though Fatoki, et al., (2010) used annual data on GDP (1980-2007) i.e. 28 data points. Preliminary analysis given in Table 1 below shows that GDP is positively skewed (Skewness=2.5249) and heavy tailed (Kurtosis= 8.5559).

Table 1: Descriptive analysis on GDP series

Mean	Median	Maximum	Minimum	Std. Dev.	Skewness	Kurtosis	JB	Prob.
667758.1	15006.6	6376225	497.7	1407810	2.524947	8.555905	460.351	0.0000

The time plot presented in Figure 1 shows that there has been astronomical increase in the Nigeria GDP series.

Figure 1. Time Plot of Quarterly Nigeria Gross Domestic Product in Million Naira (1960-2008)



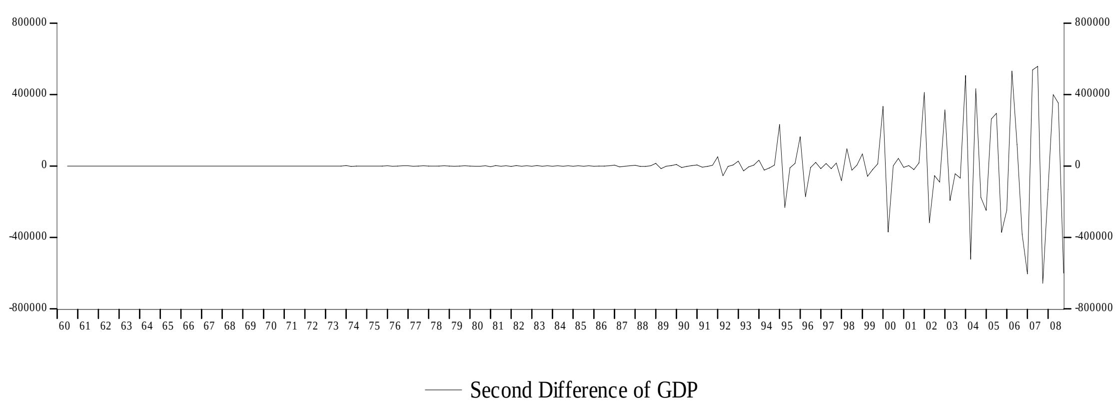
The data will be modelled in two scenarios. First the data will be modelled as ARIMA(p, d, q) and secondly as ARFUIMA(p, d_0, q) model. Then their forecast performance will be examined with a view to determining the better model on the basis of the RMS forecast error.

Table 2: Stationarity Tests on GDP Series

	Original Series			First Difference			Second Difference		
Test	ADF		KPSS	ADF		KPSS	ADF		KPSS
Statistic	4.5581		0.9573	0.2090		0.9786	-5.0126		0.0407
1%	-3.4646		0.7390	-3.4666		0.7390	-3.4666		0.7390
5%	-2.8765	(1.0000)	0.4630	-2.8774	(0.9725)	0.4630	-2.8774	(0.0000)	0.4630
10%	-2.5748		0.3470	-2.5753		0.3470	-2.5753		0.3470

From Table 2, the hypothesis of unit root of the augmented Dickey Fuller (ADF) test is not significant at level and first differenced series of the Nigerian GDP at 5% level but after second difference, the series is stationary which implies that Nigerian GDP is $I(2)$ series. The second difference series is further subjected to KPSS test of long memory and at 5% level of significance, null hypothesis of series stationarity is accepted again the alternative of long memory.

Figure 2. Plot of Second Difference of Gross Domestic Product Series



The GDP series attains stationarity after second difference, $I(u=2)$. Using the model identification tools of ACF and PACF and other diagnostic tools, the most appropriate model for the data ARIMA (4,2,0) model with minimum AIC of 25.53 as shown below.

Table 3: Estimation of Parameters of Subset of ARIMA (4, 2, 0) Model

Estimators	Coefficient	Standard Error	t-probability
$\hat{u}$	2		
$\hat{\phi}_1$	-0.986847	0.0709	0.0000
$\hat{\phi}_2$	-1.109800	0.0813	0.0000
$\hat{\phi}_3$	-0.863167	0.0836	0.0000
$\hat{\phi}_4$	-0.166294	0.0743	0.0260

$$(1-L)^2 X_t = \underset{(0.0000)}{-0.986847}(1-L)^2 X_{t-1} \underset{(0.0000)}{-1.10980}(1-L)^2 X_{t-2} \underset{(0.0000)}{-0.863167}(1-L)^2 X_{t-3} \underset{(0.0260)}{-0.166294}(1-L)^2 X_{t-4} + \varepsilon_t$$

Log-likelihood	-2471.46903	Skewness	0.48381
AIC	25.5306085	Excess Kurtosis	12.111
Variance of Residuals	6.68222E+09	Normality test	278.06 [0.0000]**
Portmanteau test	32.597 [0.0002]**	ARCH test	7.2796 [0.0076]**

The nonlinear estimation with PcGive version 10.0 following the GPH approach estimated $d = 0.7539$ with standard error of 0.06120. The value of d is within the confidence interval of [0.633987, 0.873891]. The estimated fractional unit integrated parameter is obtained using (2) as $d_0 = 1.7539$ and the resulting series is assumed to be in the stationary fractionally integrated range according to proposition 1. Rescaled Statistics (RS) approach employed on the second unit difference $I(u=2)$ series of the GDP data computed d as -0.1350, which implies $d_0 = 1.8650$.

The estimated ARFUIMA (3, 1.7539, 0) model is presented below:

Table 4: Estimation of Parameters of ARFUIMA (3, 1.7539) Model

Estimators	Coefficient	Standard Error	t-probability
$\hat{u}$	1		
$\hat{d}$	0.753939	0.1050	0.0000
$\hat{\phi}_0$	25218.1	0.0000	0.0000
$\hat{\phi}_1$	-0.789660	0.0611	0.0000
$\hat{\phi}_2$	-0.916117	0.0486	0.0000
$\hat{\phi}_3$	-0.662994	0.0601	0.0000

$$(1-L)^{1.7539} X_t = \underset{(0.0000)}{25218.1} - \underset{(0.0000)}{0.789660}(1-L)^{1.7539} X_{t-1} - \underset{(0.0000)}{0.916117}(1-L)^{1.7539} X_{t-2} - \underset{(0.0000)}{0.662994}(1-L)^{1.7539} X_{t-3} + \varepsilon_t$$

Log-likelihood	-2540.0045	Skewness	1.1124
AIC	25.460045	Excess Kurtosis	11.661
Variance of Residuals	6.32124E+09	Normality test	183.86 [0.0000]**
Portmanteau test	30.623 [0.0007]	ARCH test	8.9772 [0.0000]**

Table 5: Forecasts of ARFUIMA model based on Modified GPH estimation Approach of Fractional Difference Parameter, d.

Horizon	ARIMA			ARFUIMA		
	Forecasts (in Millions Naira)	Std. Error (in Millions Naira)	Confidence Interval (in Millions Naira)	Forecasts (in Millions Naira)	Std. Error (in Millions Naira)	Confidence Interval (in Millions Naira)
2009Q1	6289141.0	81745	[6128921, 6449361]	6300711.0	74983	[6153744, 6447678]
2009Q2	6618196.8	114850	[6393091, 6843303]	6467956.8	112430	[6247594, 6688320]
2009Q3	7132262.6	115380	[6906118, 7358407]	6849802.6	112730	[6628852, 7070753]
2009Q4	7207918.4	119200	[6974286, 7441550]	6892758.4	116750	[6663928, 7121588]
2010Q1	7172164.2	125370	[6926439, 7417889]	6946767.2	120480	[6710626, 7182908]
2010Q2	7504010.0	134430	[7240527, 7767493]	7187996.0	132300	[6928688, 7447304]
2010Q3	7944395.8	135800	[7678228, 8210564]	7504708.8	132440	[7245126, 7764291]
2010Q4	8038781.6	140060	[7764264, 8313299]	7671331.6	137000	[7402812, 7939852]
2011Q1	8055389.4	141920	[7777226, 8333553]	7818344.4	140540	[7542886, 8093803]

Using the Meese and Rogoff (MR) (1988), who developed the MR-statistic as reviewed in Section 4:

H_0: Forecast Errors are the same (Models are identical)

H_1: Forecast Errors are not the same (Models are not identical)

Since the computed MR=2.186 is greater than $Z_{0.025}$=1.96, we reject the null hypothesis of equality of forecast errors for the two models. Thus, the estimated models are not identical. The ARFUIMA (10,2.27,0) is considered considering the smaller MAD of 571110.1 against 677692.7 of ARIMA (1,2,1) model.

CONCLUSION

We have considered in the paper autoregressive fractionally unit integrated moving average and (ARFUIMA) model. Its performance was compared with the ARIMA(p,d,q) model in terms of model adequacy and forecast performance. The ARFUIMA model performed better when applied to GDP data than the ARIMA model as indicated by smaller mean squared error of forecast.

REFERENCES

Diebold, F. X. and Rudebusch, G. D. (1989). “Long Memory and Persistence in Aggregate Output", Journal of Monetary Economics, 24: 189-209.

Geweke, J. and Potter-Hudak, S. (1983): “The Estimation and Application of Long memory Time series models”, Journal of Time Series Analysis, 4: 221-238.

Fatoki, O., Ugochukwu, M. and Abass, O.(2010). “An Application of ARIMA Model to the Nigeria Gross Domestic Product (GDP)”. International Journal of Statistics and Systems, 5(1): 63-72.

Lo, A.W. (1991). Long term memory in Stock market prices", Econometrica 59: 1279-1313.

Kwiatkowski, D., Phillips, P. C.B., Schmidt, P, and Shin, Y (1992), "Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root," Journal of Econometrics, 54: 159-178.

Hassler, U. (1993). Fraktional integrierte Prozesse in der Okonometrie, Frankfurta m Main: Haag & Herchen.

Meese, R. A. and Rogoff, K. (1988). “Was Is Real? The Exchange Rate-Interest Differential Relation over the Modern Floating Rate Period”. Journal of Finance. 43: 933 – 948.

Robinson, P.M. (1995). “Log-Periodogram Regression for Time Series with long Range Dependence". Annals of Statistics 23: 1048-1072.

Sowell, F. (1992). “Maximum Likelihood Estimation of Stationary univariate Fractionally Integrated time Series Models,” Journal of Econometrics 53: 165-188.

Shittu, O.I. and Yaya, O.S. (2009). “Measuring Forecast Performance of ARMA and ARFIMA Models: An Application to US Dollar/UK Pound Foreign Exchange Rate”. European Journal of Scientific Research 32: 167-176.

Velasco, C. (1999) “Non-stationary log-periodogram regression”. Journal of Econometrics, 91: 325-371.

Velasco, C. (2005). “Semiparametric Estimation of Long-memory Models”. Department of Economics Universidad Carlos III de Madrid, Spain.

Yaya, O.S. and Shittu, O.I. (2010). “Long Memory and Estimation of Memory Parameters: Nigerian and US Inflation Rates”. International Journal of Physical Sciences, 2(3): 120-131.

ABOUT THE AUTHORS:

Olanrewaju I. Shittu and OlaOluwa S. Yaya: Department of Statistics, University of Ibadan, Nigeria