Journal of the Nigerian Association of Mathematical Physics
Volume 30, (May, 2015), pp237 – 242

Residual Analysis of Generalized Autoregressive Integrated Moving Average Bilinear Time Series Model

J. F. Ojo

Department of Statistics, University of Ibadan, Nigeria.

Abstract

In this study, analysis of residuals of generalized autoregressive integrated moving average bilinear time series model was considered. The adequacy of this model was based on testing the estimated residuals for whiteness. Jarque-Bera statistic and squared-residual autocorrelations were used to test the estimated residuals for whiteness. Generalized autoregressive integrated moving average bilinear time series model was fitted using non linear and non stationary series and the residuals were estimated. The independent test on estimated residuals showed that the residuals were independently distributed. The normality test on the estimated residuals also showed that the residuals followed a normal distribution. The tests on estimated residuals for whiteness were satisfied.

Keywords: Normality test, Residuals, Bilinear model, Jargue-Bera Statistic, Independent test

1.0 Introduction

Bilinear time series model have been studied extensively in literature[1-8]. Recently, generalized bilinear time series model that was robust in achieving stationary for all non-linear series have been studied [9, 10]. The standard large sample estimation theory [11, 12] requires that the residuals be independent and identically distributed with finite variance. Squared residual autocorrelations have been found useful in detecting non linear types of statistical dependence using the residuals of fitted autoregressive moving average (ARMA) models [1, 13]. Jarque-Bera statistic and squared-residual autocorrelations with its associated portmanteau statistic were used to test the estimated residuals for whiteness [14]. In this study, adequacy of a fitted generalized autoregressive integrated moving average bilinear (GARIMABL) model would be based on testing the estimated residuals for whiteness using Jarque-Bera statistic and squared-residual autocorrelations.

2.0 Theoretical Analysis

2.1 Generalized Autoregressive Integrated Moving Average Bilinear Time Series Model

We define generalized autoregressive integrated moving average bilinear time series model as

$$\phi(B)X_t = \alpha(B)\nabla^d X_t + \beta(B)e_t + \sum_{k=1}^{r}\sum_{l=1}^{s} b_{kl} X_{t-k} e_{t-l}\,, \qquad (1)$$

denoted as GARIMABL (p, d, q, r, s) and where

$$\alpha(B) = 1 - \alpha_1 B - \alpha_2 B^2 .... - \alpha_p B^p,\ \beta(B) = 1 - \beta_1 B - \beta_2 B^2 .... - \beta_q B^q \qquad (2)$$

$$X_t = \phi_1 X_{t-1} + ... + \phi_{p+d} X_{t-p-d} + e_t - \beta_1 e_{t-1} - ...... - \beta_q e_{t-q} + b_{11} X_{t-1} e_{t-1} + ... + b_{rs} X_{t-r} e_{t-s} \qquad (3)$$

$\alpha_1,...,\alpha_p$ are the parameters of the autoregressive component; $\beta_1,...\beta_q$ are the parameters of the associated error process; $b_{11},.........,b_{rs}$ are the parameters of the non-linear component and $\beta(B)$ is the moving average operator; *p* is the order of the autoregressive component; *q* is the order of the moving average process; *r, s* is the order of the nonlinear component and $\phi(B) = \nabla^d \alpha(B)$ is the generalized autoregressive operator; ∇^d is the differencing operator and *d* is the degree of consecutive differencing required to achieve a stationary[10].

Corresponding author: J. F. Ojo, E-mail: jfunminiyiojo@yahoo.co.uk, Tel.: +2348033810739

2.2 Model Estimation

Suppose that X_t are generated by equation (3), the sequence of random deviates $\{e_t\}$ could be determined from the relation

$$e_t = X_t - \phi_1 X_{t-1} - \ldots\ldots - \phi_{p+d} X_{t-p-d} + \theta_1 e_{t-1} + \ldots\ldots + \theta_q e_{t-q} - b_{11} X_{t-1} e_{t-1} - \ldots\ldots - b_{rs} X_{t-r} e_{t-s} \tag{4}$$

To estimate the unknown parameters in equation (4), we make the following assumptions:

(i) The errors $\{e_t\}$ are independent and identically distributed with mean zero and variance σ^2

(ii) The values of $|\phi_i's| < 1, |\theta_i's| < 1$ and $|\beta_j's| < 1$ ensure that stationary and invertibility conditions required of the generalized autoregressive integrated moving average bilinear model are satisfied.

We minimize the likelihood function

$$Q(G) = \sum_{i=1}^{n} e_t^2 \tag{5}$$

with respect to the parameter where

$$\mathbf{G} = (\phi_1, \ldots, \phi_{p+d}; \theta_1, \theta_2, \ldots, \theta_q; B_{11}, \ldots, B_{rs}) \tag{6}$$

Newton-Raphson method is employed for parameter estimation [10].

2.3 Squared Residual Autocorrelations

Mcleod and Li [15] obtained the distribution of the residual autocorrelations function as follows

$$\hat{\rho}_e(k) = \sum_{k=1}^{n} \hat{e}_t \hat{e}_{t+k} / \sum_{t=1}^{n} \hat{e}_t^2 \tag{7}$$

and suggested the portmanteau statistic

$$Q_a = n \sum_{k=1}^{M} \hat{\rho}_e^2(k) \tag{8}$$

for testing the whiteness of the residuals. Under the assumption of model adequacy, Q_a is approximately $\chi^2_{(M)}$ provided M and n are large enough [16]. The autocorrelation function of squared residual is estimated by

$$\hat{\rho}_{ee}(k) = \sum_{k=1}^{n} (\hat{e}_t - \hat{\sigma}^2)(\hat{e}_{t+k}^2 - \hat{\sigma}^2) / \sum_{t=1}^{n} (\hat{e}_t^2 - \hat{\sigma}^2)^2 \tag{9}$$

where

$$\hat{\sigma}^2 = \sum_{t=1}^{n} \hat{e}_t^2 / n \tag{10}$$

A significance test is provided by the portmanteau statistic

$$Q_{aa} = n(n+2) \sum_{k=1}^{M} \hat{r}_{ee}^2(k) / (n-k) \tag{11}$$

which is asymptotically $\chi^2_{(M)}$

2.4 Jarque-Bera Test

In statistics, the Jarque-Bera (JB) test is a goodness of fit test of whether sample data have the skewness and kurtosis matching a normal distribution. The test statistic JB is defined as

$$JB = \frac{n}{6}\left[S^2 + \frac{(k-3)^2}{4} \right] \tag{12}$$

where n is the number of observations (or degrees of freedom in general); S is the sample skewness, and K is the sample kurtosis:

$$S = \frac{\hat{\mu}_3}{\hat{\sigma}^3} = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^3}{(\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2)^{3/2}} \tag{13}$$

$$K = \frac{\hat{\mu}_4}{\hat{\sigma}^4} = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^4}{(\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2)^2} \tag{14}$$

where $\hat{\mu}_3$ and $\hat{\mu}_4$ are the estimates of third and fourth central moments, respectively, $\bar{x}$ is the sample mean, and $\hat{\sigma}^2$ is the estimate of the second central moment, the variance. If the data comes from a normal distribution, the JB statistic asymptotically has a chi-squared distribution with two degrees of freedom, so the statistic can be used to test the hypothesis that the data are from a normal distribution. The null hypothesis is a joint hypothesis of the skewness being zero and the excess kurtosis being zero. Samples from a normal distribution have an expected skewness of 0 and an expected excess kurtosis of 0 (which is the same as a kurtosis of 3). As the definition of JBshows, any deviation from this increases the JB statistic. From tables, critical value at 5% level for 2 degrees of freedom is 5.99. If JB>χ^2 critical, reject the null hypothesis of residuals[17, 18].

3.0 Experimental Work

The data used for this study was monthly goals by Manchester United Football Club (League Match) between 1992 and 2012. The data set represents a non-stationary and non linear series, and so, fitting generalized autoregressive integrated moving average bilinear model may be applied. The goals data set is available in [19].

3.1 Fitted GARIMABL Model

$$\hat{X}_t = 0.132925X_{t-1} - 0.115677X_{t-2} - 0.019656X_{t-3} - 0.018766X_{t-4} - 0.249863X_{t-5}$$
$$- 0.175653X_{t-6} - 0.175364X_{t-7} + 0.991371e_{t-1} - 0.117669X_{t-1}e_{t-1} - 0.013815X_{t-2}e_{t-1}$$
$$- 0.009635X_{t-3}e_{t-1} - 0.050570X_{t-1}e_{t-2} - 0.003539X_{t-2}e_{t-2} + 0.059762X_{t-3}e_{t-2} + 0.050255$$
$$X_{t-1}e_{t-3} - 0.073395X_{t-2}e_{t-3} - 0.028319X_{t-3}e_{t-3}$$

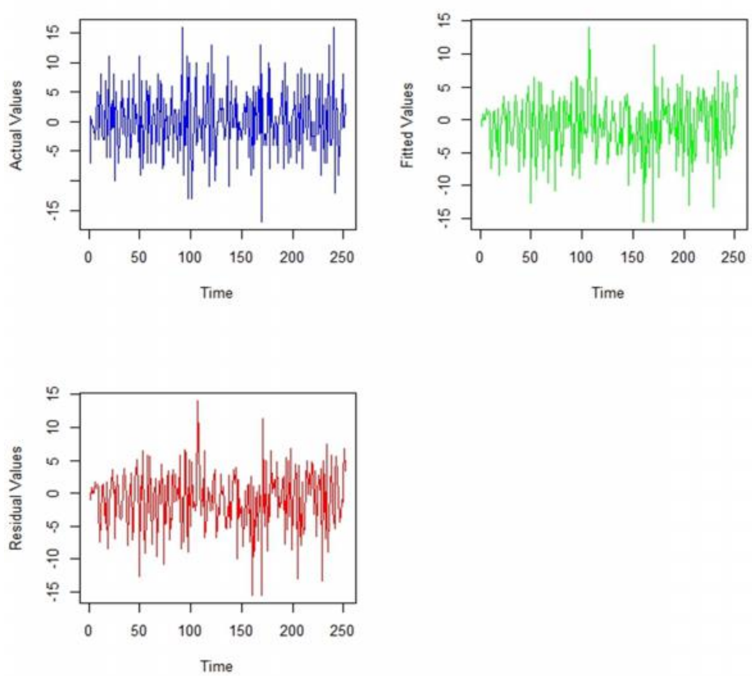

Figure 1: Graph of Actual, Fitted and Residuals of Generalized Autoregressive Integrated Moving Average Bilinear Time Series Model

3.2 Residual Autocorrelation Test of GARIMABL Model

A test of hypothesis is a rule, which, on the basis of relevant statistic, leads to a decision to accept or reject the null hypothesis. The hypothesis, H_0 is called the null hypothesis while H_1 is called the alternative hypothesis [20]. Residuals of GARIMABL will be tested for independence as follows

H_0**:** Residuals of GARIMABL are independently distributed.

H_1**:** Residuals of GARIMABL are not independently distributed.

Table 1: Squared residuals autocorrelation test of GARIMABL Model

Lag k	$\hat{\rho}_{ee}(k)$	**Q-Stat**	**Prob.**
1	0.099	2.4868	0.115
2	0.006	2.4962	0.287
3	0.036	2.8240	0.420
4	0.076	4.3027	0.367
5	-0.038	4.6853	0.455
6	0.073	6.0889	0.413
7	0.016	6.1532	0.522
8	0.005	6.1600	0.629
9	0.068	7.3605	0.600
10	-0.052	8.0845	0.621
11	0.011	8.1164	0.703
12	-0.032	8.3936	0.754
13	-0.029	8.6236	0.801
14	-0.114	12.139	0.595
15	-0.042	12.612	0.632
16	-0.018	12.703	0.694
17	-0.032	12.985	0.737
18	-0.051	13.693	0.749
19	-0.120	17.623	0.548
20	-0.022	17.753	0.604

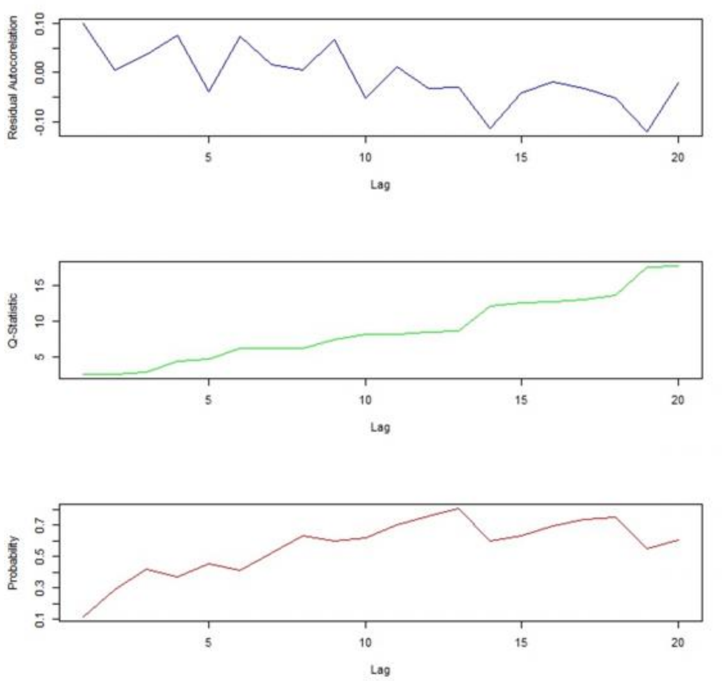

Figure 2: Graph of Squared Residual Autocorrelations, Q–Statistic and Probability Level

3.4 Normality Test

In statistics, normality tests are used to determine if a data set is well modelled by a normal distribution [21]. Residuals of GARIMABL will be tested for normality as follows

H_o: Residuals of GARIMABL follow a normal distribution

H_1: Residuals of GARIMABL do not follow a normal distribution

Table 2: Summary and Jarque-Bera Statistics of Residuals of GARIMABL Model

Series	Residuals
Period	1992 Month 01 – 2012 Month 12
Number of Observations	252
Mean	0.893294
Median	0.384630
Maximum	23.61279
Minimum	-17.73668
Standard Deviation	8.094252
Skewness	0.277986
Kurtosis	2.694341
Jarque-Bera	4.226594
Probability	0.120839

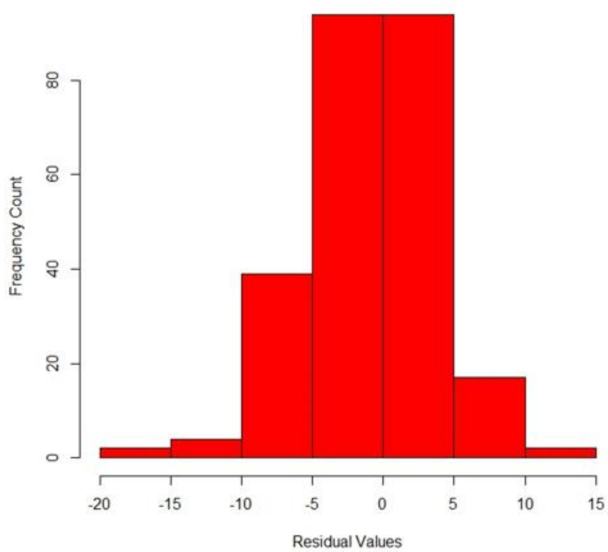

Figure 3: Normal Plot of Residuals

4.0 Results and Discussion

The estimated generalized autoregressive integrated moving average bilinear model is presented in section 3.1. This model is a type that handles all non linear and non stationary time series. The estimated model satisfied the stationary and invertibility conditions. The actual values, the fitted values and estimated residuals were presented on a graph which is depicted as Figure 1. The estimated residuals are the difference between the actual values and the fitted values. Residuals were tested for independence. Squared-autocorrelations values, Q-statistic and probability values at different lags were presented in Table 1where inferences were drawn on the satisfaction of independence of residuals. The probability values at different lags were greater than p = 0.05 which indicated that the null hypothesis of residuals is independently distributed is accepted. Summary and Jarque-Bera statistics were presented in Table 2 while Figure 2 contained normal plot. Jargue-Bera statistic in Table 2 tested residuals for normality and since $\chi^2_{0.05,2} = 5.99$ was greater than 4.23; it indicated that the null hypothesis cannot be rejected. Also the probability value of 0.121 which was greater than p = 0.05 confirmed also that the null hypothesis cannot be rejected and as a result, the residuals of GARIMABL model followed a normal distribution

5.0 Conclusion

We have considered the adequacy of generalized autoregressive integrated moving average bilinear model when the data exhibited non stationary and non linearity [2]. Tests on the residuals of this model have been considered using squared–residual autocorrelations and Jarque-Bera statistic to ascertain the adequacy of this model. The tests on estimated residuals for whiteness are satisfied. Therefore, when faced with non stationary and non linear series and the interest is to generalize the variable under study, the adequate model to be fitted is generalized autoregressive integrated moving average bilinear time series model.

6.0 References

[1]. Granger, C. W. J. and Anderson, A. P. (1978). *Introduction to Bilinear Time Series Models*. Gottingen, Germany: Vandenhoeck and Ruprecht.

[2]. Subba Rao, T. and Gabr, M. M. (1981). *The Estimation and Prediction of Subset Bilinear Time Series Models with Application.*Journal of Time Series Analysis 2(3):89-100.

[3]. Subba Rao, T. (1981).*On theory of Bilinear time Series Models.*Journal of the Royal Statistical Society B 43:244-255.

[4]. Rao, M. B., T. S. Rao and Walker, A. M. (1983). *On the Existence of some Bilinear Time Series Models*.Journal of Time Series Analysis 4(2):60-76.

[5]. Liu, J. (1992). *On stationary and Asymptotic Inference of Bilinear time series models*.Statistica Sinica 2(2):479-494.

[6]. Gonclaves, E., Jacob, P. and Mendes-Lopes, N. (2000). *A decision procedure for bilinear time series based on the Asymptotic Separation.* Statistics 33(4):333-348.

[7]. Shangodoyin D. K. andOjo, J. F (2003). *On the Performance of Bilinear Time Series Autoregressive Moving Average Models.* Journal of Nigerian Statistical Association Volume 16, Pages 1-12.

[8]. Ojo, J. F. (2009). *The Estimation and Prediction of Autoregressive Moving Average Bilinear Time Series Models with Applications*.Global Journal of Mathematics and StatisticsVolume 1, No2, Pages 111-117.

[9]. Ojo, J. Fand Shangodoyin D. K.(2010). *Subsetting and Identification of Optimal Models in Generalized Bilinear Time Series Modelling*.Jordan Journal of Mathematics and Statistics Volume 3, No 1, Pages 1-20.

[10]. Ojo, J. F. (2011).*Generalized Integrated Autoregressive Moving Average Bilinear Time Series Models.*International Journal of Statistics and systems 6(3): 365-374.

[11]. Whittle, P. (1961). *Gaussian estimation in Stationary Time Series.*Bull. Int. Statist. Inst. 33, 105-129.

[12]. Hannan, E. J. (1970) *Multiple Time Series*, Wiley: New York.

[13]. Miller, R. B. (1979) *Book Review on 'An Introduction to Bilinear Time Series Models', by Granger C. W. and Andersen A. P*.Journal of American Statistical Association 74, 927.

[14]. Jarque, Carlos M. and Bera, Anil K.(1987).*A Test for Normality of Observations and Regression Residuals*. International Statistical Review 55 (2): 163–172.

[15]. McLeod, A. I. and Li, W. K. (1983). *Diagnostic Checking ARMA Time Series Models using Squared-Residual Autocorrelations.* Journal of Time Series Analysis Volume 4, No. 4, 269-273.

[16]. McLeod, A. I. and Hipel, K. W. (1978). *Simulation Procedures for Box-Jenkins Models.*Water Resources Research 14, 969-975.

[17]. Jarque, Carlos M. and Bera, Anil K.(1980).Efficient Tests for Normality, Homoscedasticity and Serial Independence of Regression Residuals. Economics Letters 6 (3): 255–259.

[18]. Jarque, Carlos M. and Bera, Anil K.(1981).Efficient Tests for Normality, Homoscedasticity and Serial Independence of Regression Residuals: Monte Carlo Evidence. Economics Letters **7** (4): 313–318.

[19]. www.11v11.com/teams/manchester-united/tab/stats/option/scores. Date Visited, 14/07/2014

[20]. Ojo**,** J. F. (2007).Statistical Inference. Distance Learning Centre Publication, University of Ibadan, Ibadan. ISBN 978-021-274-4.

[21]. en.wikipedia.org/wiki/Normality_test. Date Visited, 26/02/2015.

[22]. Ojo, J. F. (2008). Linear and Non Linear Series in Linear *and Non Linear Time Series Models.*Nigerian Statistical Association Conference Proceedings Pages 266-273.