Journal Homepage: - www.journalijar.com

INTERNATIONAL JOURNAL OF ADVANCED RESEARCH (IJAR)

Article DOI: 10.21474/IJAR01/17635
DOI URL: http://dx.doi.org/10.21474/IJAR01/17635

ISSN NO. 2320-5407

RESEARCH ARTICLE

DEVELOPMENT OF SEASONAL AUTOREGRESSIVE INTEGRATED MOVING AVERAGE MODEL WITH ERROR PROCESS

Ibrahim Maihaja[1], Afolabi Wasiu Babayemi[2] and G. I. Onwuka[2]

1. Department of Mathematics and Statistics, Abdu Gusau Polytechnic, TalataMafara, Zamfara, Nigeria.
2. Department of Mathematics, Kebbi State University, Aleiru, Kebbi, Nigeria.

Manuscript Info

Manuscript History
Received: 28 July 2023
Final Accepted: 31 August 2023
Published: September 2023

Key words:-
Auto Covariance Fuctions, Autoregressives Moving Average, Error Process, Sarima

Abstract

This research work investigated $SARIMA(p,d,q)\times(P,D,Q)^S$ with error process. The auto covariance function maximum likelihood methods, iterative method and chi-squares test statistics is used to develop SARIMA model corrupted with error process that is used to estimate the true parameter of the SARIMA model. The forecast performance measurement and properties of error with different value are also investigated. Test of seasonal unit roots are also carried out in the work. The simulation, and real data of Zamfara state monthly Rain fall from 1998 to 2022 is used to validate the results with R - Statistical software version 4.1.1 and mini tab statistical software version 14. The result showed a significance p- value of 0.000, the proposed model provides a generalization and more flexible specification than the existing models of AR (1) error and ARMA (1,1) error in fitting time series processes in the presence of errors . Hence, the studies showed that, the finding is closely to the true parameter of the process and would be useful to researchers in the prediction and handling of natural calamities that disturb the otherwise stability of a system.

Introduction:-

In the study on structural relationships, a point of contention of practical interest to researches is how to describe a relationship when concerned variable are measured with errorsKomolafeet al (2019). This investigation started with early influential studies that include Lindley (1947), Madansky, (1959), Kendall and Stuart (1961)and Barnett(1967)more recent studies in this area discussed in this study includes: Rudelson and Zhou worked with errors in variable models with dependent measurement, analysed the convergence rates of the gradient descent methods for solving the no convex programs, and show that the composite gradient, descent algorithm is guaranteed to converge at a geometric rate to a neighbourhood of the global minimizes. The size of the neighbourhood is bounded by the statistical error in the ℓ2 norm. The result reveals interesting connections between computational and statistical efficiency and the concentration of measure phenomenon in random matrix theory.

Madanskyworked on fitting of straight lines when both variables are subjected to errors**.** The studyconsidered the situation where X and Y are related by $Y = \alpha + \beta X$, where α and β are unknown and observed X and Y with errors, i.e. observed $x = X + u$ and $y = Y + v$. Assume that $Eu = Ev = 0$ and that the errors (u and v) are uncorrelated with the true values (X and Y). He surveyed and commented on the solutions to the problem of obtaining consistent

Corresponding Author:- Ibrahim Maihaja
Address:- Department of Mathematics and Statistics, Abdu Gusau Polytechnic, Talata - Mafara, Nigeria.

estimates of α and β from a sample of (x, y)'s. Firstly, when one makes various assumptions about properties of the errors and the true values other than those mentioned above, secondly when one has various kinds of "additional information" which aids in constructing these consistent estimates. And problems of obtaining confidence intervals for β and of testing hypotheses about β are not discussed, though approximate variances of some of the estimates of β are the moving average model in essentially a finite impulse response applied to white noise, with some additional interpretation placed on it. The role of the random shocked in the MA model differ from their role in the AR model in two ways, first they are propagated to future values of the time series directly. Second, the MA model shock affects the series values only for the current periods and period in to the future.

Komolafe developed an Integrated Moving Average (IMA) model with a transition matrix for error resulting in a convex combination of two ARMA errors. The basic tools they used are the autocovariance function, maximum likelihood methods, Raphson iterative method and Kolomogorov Smirnov test statistic. The result showed that the proposed model provided a generalization and more flexible specification than the existing models of AR and ARMA errors in fitting time series processes in the presence of error.

Eni (2013) worked on parameter estimation of first order IMA model in the presence of ARMA (1, 1) errors using simulation method and showed that the error was uniformly AR(1) correlated. He used autocovariance functions to estimate the variances of the white noises that characterize the IMA (1) models corrupted with ARMA (1, 1) errors. He developed an iteration formula that can be used to estimate the parameters of the IMA (1) models and ARMA (1, 1) errors using simulation studies to demonstrate the findings. The results showed that the method vary closely to the estimated true parameters of the process. This work demonstrated the use of autocovariance function in the isolation and measurement of correlated shocks.

Eni and Mahmud worked on the parameter estimation of first order IMA model corrupted with AR (1) error using maximum likelihood method, the result showed that error pattern varied between AR and ARMA process within a specified period arising from the varying dynamic process to be observed.

Process involving lags of error terms can be estimated using maximum likelihood estimate. Some of the pioneers of this work are discussed accordingly.

Ansley worked on finite sample properties of estimators for Autoregressive Moving Average Models.He analyzed by simulation the properties of three estimators frequently used in the analysis of autoregressive moving average time series models for both non seasonal and seasonal data. The estimators considered are exact maximum likelihood, exact least squares and conditional least squares. For samples of the size commonly found in economic applications, the estimators are compared in terms of bias, mean squared error, and predictive ability. The reliability of the usually calculated confidence intervals is assessed for the maximum likelihood estimate.

The Markov switching model is a popular type of regime-switching model which assume that un-observed states are determined by an underlying stochastic process known as Markov-chain.

Ayodeji worked on three states Markov-modulated switching Model for exchange rate. He examined the long swings hypothesis in exchange rates using a two-state Markov switching model. This study developed a model to investigate long swings hypothesis in currencies which may exhibit ak-state ($k \geq 2$) pattern, his model was then applied to euro, British pounds, Japanese Yen, and Nigerian Naira. Specification measures such as AIC, BIC, and HIC favoured a three-state pattern in Nigerian Naira but a two-state one in the other three currencies. For the period January 2004 to May 2016, empirical result shows the presence of asymmetric swings in naira and yen and long swings in Euros and Pounds. In addition, taking 0.5 as the benchmark for smoothing probabilities, choice models provided a clear reading of the cycle in a manner that is consistent with the realities of the movements in corresponding exchange rate series.

Lindley worked on regression lines and linear functional relationship. Using least square method and maximum likelihood estimation method for fitting a straight line, $Y = \alpha + \beta X$. All these methods led to the same result, a quadratic equation which can be solved to give an estimate of β.

The aim of this research paper is to develop and estimate SARIMA Model corrupted with combination of AR (1) and AMRA (1, 1) error process. The paper also wishes to achieve the following goals: (i) to examine the properties of error and variation with different values of the parameters (ii) to test the seasonal, unite root on simulated data to investigate the forecast performance measures and (iii) to validate the result with real data.This research tends to help researchers and Government official in making decision on crucial areas under similar studies.

Methodology:-

With reference to Brockwell and Davis (2016), Neusser (2016), Schneeweiss and Shalabh (2007) and Hamilton (1994).Suppose the error b_t is a Markov modulated mixture of AR (1) and ARMA (1, 1), so that when b_t is AR (1) correlated.

$$b_t = \frac{\in_t}{1-\alpha_2 L} \tag{1}$$

And when it is ARMA (1, 1) correlated

$$b_t = \frac{\left(1-\beta L\right)\in_t}{1-\alpha_2 L} \tag{2}$$

Consider SARIMA model corrupted with combination of AR (1) and ARMA (1, 1) errors

$$w_t = (1-\phi L)(1-\lambda L^{12})\varepsilon_t + (1-L)(1-L^{12})b_t$$

(3)

$$w_t = \left(1-\phi L\right)(1-\lambda^{12})\varepsilon_t + (1-L)(1-L^{12})\left[\frac{et}{1-\alpha_1 L} + \frac{\left(1+\beta L\right)}{\left(1-\alpha_2 L\right)}et\right]$$

$$w_t(1-\alpha_1 \mathrm{L})(1-\alpha_2 \mathrm{L}) = (1-\alpha_1 \mathrm{L})(1-\alpha_2 \mathrm{L})(1-\phi \mathrm{L})(1-\lambda \mathrm{L}^{12})\varepsilon_t + (1-\mathrm{L})(1-\mathrm{L}^{12})(1-\alpha_2 L)e_t + (1-L)(1-L^{12})(1-\alpha_1 L)e_t$$

(4)

$$w_t(1-\alpha_2 \mathrm{L}-\alpha_1 L+\alpha_1\alpha_2 \mathrm{L}^2) = (1-\alpha_2 \mathrm{L}-\alpha_1 L+\alpha_1\alpha_2 \mathrm{L}^2)(1-\phi \mathrm{L})(1-\lambda \mathrm{L}^{12})\varepsilon_t + (1-L-L^{12}+L^{13})(1-\alpha_2 L)e_t + (1-L-L^{12}+L^{13})(1-\alpha_1 L)\mathrm{e}_t$$

$$w_t = \alpha_1 w_{t-1} + \alpha_2 w_{t-1} + \alpha_1\alpha_2 w_{t-2} + \varepsilon_t - \phi_{\varepsilon-1} - \lambda\varepsilon_{t-12} + \phi\alpha\varepsilon_{t-13} - \alpha_2\phi\varepsilon_{t-1}$$

$$+\alpha_2\phi\varepsilon_{t-2} + \alpha_2\lambda\varepsilon_{t-13} - \alpha_2\phi\varepsilon_{t-14} - \alpha\varepsilon_{t-1} + \alpha_1\phi\varepsilon_{t-2} + \alpha_1\lambda\varepsilon_{t-13} + \alpha_1\phi\lambda\varepsilon_{t-14} + \alpha_1\alpha_2\varepsilon_{t-2}$$

(5)

$$-\lambda_1\lambda_2\phi\varepsilon_{t-3} - \alpha_1\alpha_2\lambda\varepsilon_{t-14} + \alpha_1\alpha_2\phi\lambda\varepsilon_{t-15} + e_t + \mathrm{e}_{t-1} + e_{t-12} + e_{t-13} - \lambda_2 e_t + \lambda_2 e_{t-2} + \alpha^2 e_{t-13}$$

$$-\alpha_2 e_{t-14} + e_t - e_{t-2} - e_{t-12} + e_{t-13} - \alpha_1\ e_{t-1} + \alpha_1 e_{t-2} + \alpha_1 e_{t-13} - \alpha_1$$

$$w_t = (\alpha_1+\alpha_2)w_{t-1} + \alpha_1\alpha_2 w_{t-2} + \varepsilon_t - (\phi+\alpha_2+\alpha_1)\varepsilon_{t-1} + (\alpha_2\phi\alpha_1\phi + \alpha_1\alpha_2)\varepsilon_{t-2} - \alpha_1\alpha_2\varepsilon_{t-3}$$

$$-\lambda\varepsilon_t + (\phi\lambda + \alpha_2\lambda + \alpha_1\lambda)\varepsilon_{t-13} + (\alpha_1\phi\lambda - \alpha_2\lambda - \alpha_1\alpha_2\lambda)\varepsilon_{t-14} + \alpha_1\alpha_2\phi\lambda\varepsilon_{t-15} + 2e_t \tag{6}$$

$$+(1-\alpha_2-\alpha_1) + (\alpha_1+\alpha_2-1)e_{t-2} + (1-1)\varepsilon_{t-12} + (1+\alpha_2+1+\alpha_1)\varepsilon_{t-13} - (\alpha_1+\alpha_2)e_{t-14}$$

$$Let\quad \alpha_1+\alpha_2 = \beta_1$$

$$\alpha_1\alpha_2 = \beta_2 \tag{7}$$

$$w_t = Z_t$$

$$Z_t = \beta_1 Z_{t-1} + \beta_2 Z_{t-2} + \varepsilon_t - (\phi + \beta_1)\varepsilon_{t-1} + (\phi\beta_2 + \beta_2)\varepsilon_{t-2} - \beta_2\varepsilon_{t-3} - \lambda\varepsilon_{t-12} + (\phi\lambda + \lambda\beta_1)\varepsilon_{t-13}$$

$$- (\beta_1\phi\lambda - \beta_2\lambda)\varepsilon_{t-14} + \beta_2\phi\lambda\varepsilon_{t-15} + 2e_t + (1-(\beta_1))e_{t-1} + (\beta_1 - 1)e_{t-2} + 0e_{t-12}(2+\beta_1)e_{t-13} - \beta_1 e_{t-14} \quad (8)$$

Grouping

$$u_t = \varepsilon_t + 2e_t$$

$$\Omega_1 u_{t-1} = (1-\beta_1)\,\mathrm{e}_{t-1} - (\phi + \beta_1)\varepsilon_{t-1}$$

$$\Omega_2 u_{t-2} = (\phi\beta_2 + \beta_2)\varepsilon_{t-2} + (\beta_1 - 1)\varepsilon_{t-2}$$

(9)

$$\Omega_3 u_{t-3} = (-\beta_2)\varepsilon_{t-3}$$

$$\Omega_4 u_{t-12} = -\lambda\varepsilon_{t-12}$$

$$\Omega_5 u_{t-13} = (\phi\lambda + \lambda\beta_1)\varepsilon_{t-13} + (2+\beta_1)e_{t-13}$$

$$\Omega_6 u_{t-14} = -(\beta_1\phi\lambda - \beta_2\lambda)\varepsilon_{t-14} - \beta_1 e_{t-14}$$

(10)

$$\Omega_7 u_{t-15} = \beta_2\phi\lambda\varepsilon_{t-15}$$

The developed model is $SARIMA\ (2,0,3)\times(0,0,1)_{12}$

$$Z_t = \beta_1 Z_{t-1} + \beta_2 Z_{t-2} + U_t + \Omega_1 U_{t-1} + \Omega_2 U_{t-2} - \Omega_3 U_{t-3}$$

$$- \Omega_4 U_{t-12} + \Omega_5 U_{t-13} - \Omega_6 U_{t-14} + \Omega_7 U_{t-15} \quad (11)$$

Data Analysis And Results:-

Table I:- Shows the results of $SARIMA\ (2,0,3)\times(0,0,1)_{12}$ corrupted with combination of AR (1) and ARMA (1, 1) Error Process.

	Estimate	**Stand. Error**	**Z-value**	**P-value**
AR (1)	0.2560	0.4867	0.5260	1.3500
AR (2)	-0.2800	0.3406	-0.8221	0.4110
MA (1)	-0.3767	0.4867	-0.7740	0.4216
MA (2)	0.2645	0.3743	0.7067	0.1264
MA (3)	-0.0449	0.0445	-1.0090	0.0502
SMA(1)	-0.0268	0.0142	-1.8873	0.0000
Mean	0.0046	0.0104	0.4423	$1.8442\,e^{-31}$

Sigma^2 estimated as 0.8469 log likelihood = -6675.89

AIC = 13367.78 AICC = 13367.8 BIC = 13419.91

The results in TABLE II showed significant results of seasonal moving average of 0.0000
P-values, when an error process switch to AR (1) error process and ARMA (1,1) error process

Table IIII:- Training set error measures.

	ME	**RMSE**	**MAE**	**MPE**	**MAPE**	**MASE**	**ACFI**
Forecasts Training set	$-0.4836\,e^{-06}$	0.9197	0.7328	125.2706	155.0039	0.6918	0.0013

The results in TABLE IIV showed the forecast performance measures and properties of error variation with different values in the training set error measurements.

Table IVI:- $SARIMA\ (2,0,3)\times(0,0,1)_{12}$ for Monthly Rain Fall of Zamfara Sate from 1998 to 2020 Estimates at each iteration

Iteration	**SSE**	**Parameters**						
0	2824812	0.100	0.100	0.100	0.100	0.100	0.100	0.100
1	2134433	0.206	0.127	0.187	0.143	0.103	-0.050	48.479
2	1674369	0.325	0.197	0.294	0.231	0.107	-0.200	34.753
3	1397899	0.445	0.322	0.415	0.381	0.117	-0.320	16.895
4	1259704	0.595	0.180	0.563	0.259	0.115	-0.407	16.221
5	1185908	0.744	0.036	0.713	0.131	0.105	-0.460	15.827
6	1138718	0.892	-0.101	0.863	0.008	0.090	-0.495	15.024
7	1054676	1.034	-0.249	1.013	-0.107	0.086	-0.599	15.254
8	1031567	1.018	-0.387	1.015	-0.120	0.078	-0.741	25.453
9	987287	1.028	-0.404	1.020	-0.142	0.077	-0.674	26.708
10	971784	1.172	-0.487	1.170	-0.228	0.028	-0.677	22.412
11	967101	1.194	-0.515	1.173	-0.235	0.033	-0.676	23.109
12	965833	1.195	-0.514	1.178	-0.234	0.033	-0.675	22.859
13	965334	1.196	-0.513	1.178	-0.233	0.034	-0.675	22.720
14	964954	1.196	-0.511	1.179	-0.231	0.035	-0.675	22.617
15	964826	1.196	-0.511	1.179	-0.231	0.036	-0.675	22.575
16	964816	1.195	-0.510	1.179	-0.232	0.036	-0.675	22.570
17	964802	1.195	-0.510	1.179	-0.232	0.036	-0.675	22.569
18	964787	1.195	-0.509	1.179	-0.232	0.037	-0.675	22.569
19	964771	1.195	-0.509	1.180	-0.233	0.037	-0.675	22.571
20	964756	1.194	-0.509	1.180	-0.233	0.037	-0.675	22.573
21	964740	1.194	-0.509	1.180	-0.234	0.037	-0.675	22.575
22	964724	1.194	-0.509	1.180	-0.234	0.037	-0.675	22.577
23	964709	1.194	-0.509	1.180	-0.234	0.038	-0.675	22.580
24	964693	1.194	-0.509	1.181	-0.235	0.038	-0.675	22.582
25	964678	1.194	-0.509	1.181	-0.235	0.038	-0.675	22.584

** Convergence criterion not met after 25 iterations **

The results in TABLE IVII showed the parameter estimate at each iteration for SARIMA model corrupted with combination of AR (1) and ARMA (1,1) error process.

Table IV:- Final Estimates of Parameters.

Type	**Coef**	**SE Coef**	**T**	**P**
AR (1)	1.1939	0.0856	13.95	0.000
AR (2)	-0.5088	0.0826	-6.16	0.000
MA (1)	1.1808	0.0004	2791.59	0.000
MA (2)	-0.2352	0.1137	-2.07	0.040
MA (3)	0.0381	0.1139	0.33	0.738
SMA (12)	-0.6754	0.0899	-7.51	0.000
Constant	22.5844	0.3685	61.28	0.000
Mean	71.718	1.170		

Number of observations: 144

Residuals: SS = 938575 (back forecasts excluded); MS = 6851 DF = 137

The results in TABLE IV showed the final parameter estimate for SARIMA model corrupted with combination of AR (1) and ARMA (1, 1) error process.

Table V:- Modified Box-Pierce (Ljung-Box) Chi-Square statistic.

S/N	Lag	Chi-Square	DF	P-Value
1.	12	20.3	5	0.001
2.	24	90.7	17	0.000
3.	36	121.8	29	0.000
4.	48	149.2	41	0.000

The results in TABLE V shows the Chi-square statistic for the SARIMA model corrupted with combination of AR(1) and ARMA (1, 1) error process at lag 12, 24, 36 and 48 respectively, for the rainfall data in Zamfara.

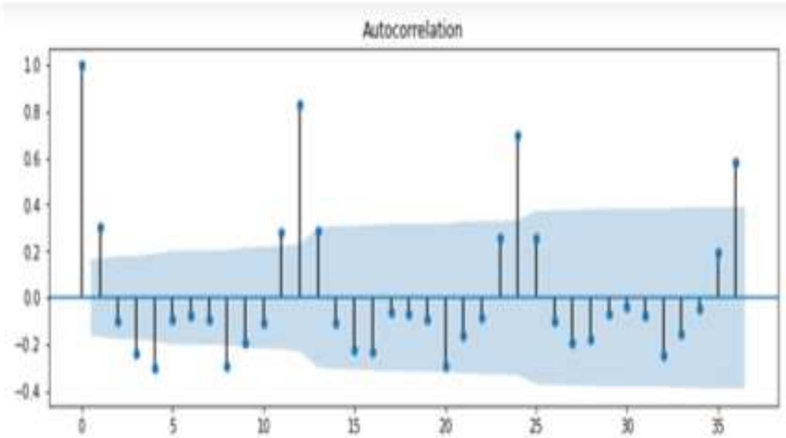

Fig. 1:- Shows ACF at lag 35.

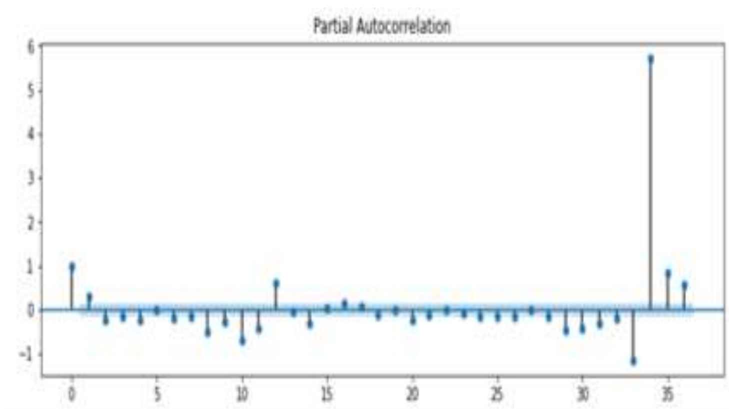

Fig. II:- Shows PACF at lag 35.

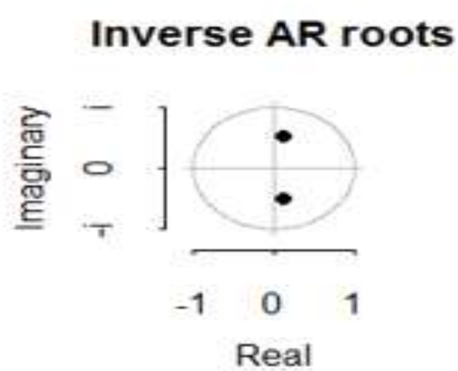

Fig. III:- Inverse AR Roots.

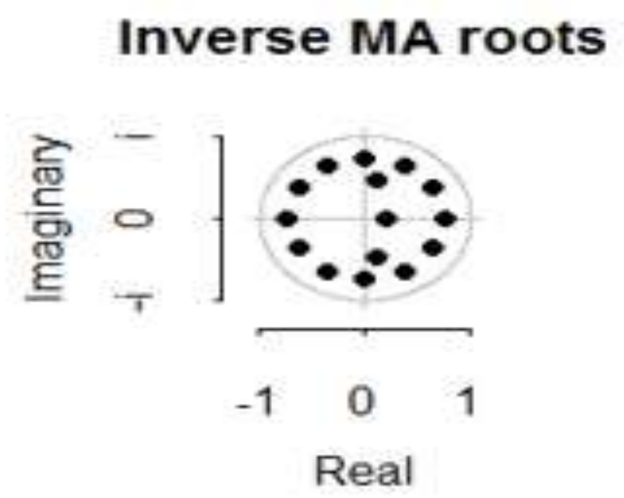

Fig. IV:- Inverse MA Roots.

Figure III is the pictorial representation of SARIMA model for the period of twelve months with double AR (2) root with represent the inner points. Figure IV is the pictorial representation of SARIMA MODEL for the period of twelve months with MA (3) roots inside the unit circle.

Recommendations:-

Finally, this study discussed SARIMA models corrupted with AR(1) and ARMA(1,1) errors it will certainly enhance research of other version of time series like multivariate time series, Autoregressive Conditional Heteroskedasticity (ARCH) and Generalized Autoregressive Conditional Heteroskedasticity (GARCH) are considered for volatility. More data will be required for better results.

Conclusion:-

This study formulated a credible flexible and versatile models capable of accounting for errors from different sources and has been able to apply basic tools such as the auto-covariance function, maximum likelihood method, R-Statistical Software, Minitab and Chi-Square test to estimate the parameter of the models that fit the formulated specification to data. Based on the theoretical results and data application, we concluded that the proposed models are significances and is a generalization of the existing models on IMA(1) corrupted with error, AR(1) error and ARMA(1,1) corrupted with error.

Acknowledgment:-

Firstly, we would like to thank our supervisors for their valuable time, guidance and insightful comments throughout the entire research process. We would like to express our heartfelt gratitude to all those who have contributed to this research. Finally, we extend thanks and appreciation to the Abdugusau Polytechnic Talata - Mafara and the TETFUND, this research would not have been possible without their support.

References:-

[1] Lindley.D. V.Regression Lines and the Linear Functional Relationship,The Journal of the Royal Statistical Society, 1947, vol. 9, pp. 218 – 44.
[2] Madansky.A.The Fitting of Straight Lines when Both Variable are Subject to Error, Journal of the American Statistical Association, 1959, vol. 54, pp. 173 – 205.
[3] KendallM.G. and Stuart .A. The Advance Theory of Statistics. London, Griffin, 1961.
[4] Barnett.V. D. A note on linear structural relationship when both residual variance are known,Biometrica, 1967, vol. 63, Pp39 – 50.
[5] RudelsonM. and Zhou S. Errors in Variables Models with Dependent Measurement, Electronic Journal of Statistics, 2017, vol. 11, pp. 1699 – 1797.

[6] KomolafeS. A. Obilade. T. O. AyodejiI. O. and Babalola.A. R. Development of a First Order Integrated Moving Average Model Corrupted with a Markov Modulated Convex Combination of Autoregressive Moving Average Errors, Journal of statistical theory, 2019, vol. 3, issue 1, pp. 48 – 58.
[7] Eni. D. Parameter Estimation of First Order IMA Model in the Presence of ARMA (1, 1) Errors,'' Research journal in Engineering and applied Science, 2013, 2(4), pp.246 – 50.
[8] Eni .D. and Mahmud S. A.On the Parameter Estimation of First Order IMA Model Corrupted with AR (1) Errors,'' Global Journal of Pure and applied Science, 2008, 14(1), pp. 115 – 20.
[9] AnsleyC.F. and New bold P. Finite Sample Properties of Estimators for Autoregressive Moving Average Models,'' Journal of Econometrics,1980, 13(2):159-183.
[10] Ayodeji.I. O. A three State Markov-Modulated Switching Model for Exchange Rates, Journal of Applied mathematics, 2016, 1 – 9.
[11] Brockwell.P. J.and Davis.R. A.Introduction to Time Series and Forecasting, 3rd edition. Springer: International Publishing, 2016.
[12] Neusser. K.Time Series Econometrics. Springer, New York, 2016.
[13] Schneeweiss. H. andShalabh. H. ''On the Estimation of the Linear Relation when the Error Variances are known,'' Computational Statistics & Data Analysis, 2007, vol. 52, pp. 1143 – 1148.
[14] Hamilton.J. D. Time Series Analysis, Princeton Univ. Press, Princeton, New Jersey, 1994.