ANNALS of Faculty Engineering Hunedoara – International Journal of Engineering
Tome XVIII [2020] | Fascicule 2 [May]

ANNALS of Faculty Engineering Hunedoara – International Journal of Engineering
ISSN 1584 - 2665 (printed version); ISSN 2601 - 2332 (online); ISSN-L 1584 - 2665

[1.]Adedayo F. ADEDOTUN, [1.]Abass I. TAIWO, [1.]Timothy O. OLATAYO

# A REPARAMETRISED AUTOREGRESSIVE MODEL FOR MODELLING GROSS DOMESTIC PRODUCT

[1.]Department of Mathematical Sciences, Olabisi Onabanjo University, Ago-Iwoye, NIGERIA

**Abstract:** This paper is used to propose a reparametrised autoregressive model that is capable of analyzing time series data that follows a non-Gaussian marginal distribution. The Anderson Darling Statistics was used to identify that Nigerian Gross domestic product followed a Gamma distribution. The proposed Gamma autoregressive (GAR) and classical autoregressive models were fitted using a Maximum Likelihood Estimation (MLE) method. The Akaike Information Criteria (AIC) was used to select AR(2) and GAR(2) as the optimal models but GAR(2) was chosen because it has the least value of AIC. The comparison between AR(2) and GAR(2) models based on the values of Mean absolute error (MAE), Mean absolute prediction error (MAPE) and Root mean square error (RMSE) indicted that GAR(2) will yield a more accurate forecast than AR(2). In essence, GAR model is a viable alternative and better model for analyzing GDP growth rate
**Keywords:** Non-Gaussian model, Classical autoregressive, Economic growth, reparameterization

## 1. INTRODUCTION

A key economic instrument used in determining how favorable the economy of a nation behaves is the gross domestic product (GDP) (George and Ibiok, 2015). The gross domestic product is mostly seen as the worth of all finished products and services (monetary) made inside a country at a specific era (Abdulraheem, 2011). If the GDP of a country is relatively high, the economy of the country is presumed to be good. Gross national product data consist of observations taken at equally spaced time intervals. The Nigeria gross national product is a good example. It is computed and documented on yearly quarterly basis. Therefore, the GDP should be a time series data of interest to individual, cooperate organisations and Government at all levels (Okereke and Bernard, 2014). Thus, modelling and analysing GDP requires the use of a suitable model that takes cognisant of the of the characteristic the gross domestic product data exhibits. This is so because of the variations in the data which may often makes the data non-normal in nature (Wei, 2006).

It is a general practice to take the log of GDP before it is analysed by using the Box and Jenkins method (Gujarati, 2004). This is so because of the assumption of normality. A statistical model that is parametric in nature usually obeys a probability distribution that is normally distributed but for a set of observed data like gross domestic product that is assumed to deviate from normality, this leads to scenarios in which a time series data can be obviously non-Gaussian. Ma and Leijon (2011) recommended that a new time series model should be developed to handle this situation. In essence, there are various studies on modelling of time series data with Poisson, gamma, beta marginal distributions in the recent past. For instant, Langberg and Stoffer (1987) developed a bivariate moving average (MA) processes that followed an exponential and geometric marginal distribution. Liseo and Loperfido (2006) proposed priors for scalar skew-normal distribution. Popovici (2010) discussed stochastic properties and parameter estimation for non-normal innovations. Xu (2014) modeled a non-Gaussian time series with the nonparametric Bayesian model while Johannesson et al., (2015) worked on generalized Laplace marginal distribution time series model.

Liu et al., (2018) proposed an efficient framework for the parameter estimation from incomplete heavy-tailed time series based on the stochastic approximation expectation maximization (SAEM) coupled with a Markov Chain Monte Carlo (MCMC) procedure. While as well Hajrajabi and Maleki (2019) proposed a data generating structures which can be represented as the nonlinear autoregressive models with single and finite mixtures of scale mixtures of skew normal innovations. Other methodologies for modelling non-normal time series includes Bayesian forecasting models by (Harvey and Fernandes, 1989), State Space models by (Kashiwagi and Yanagimoto, 1992; Fahrmeir, 1991) and the transformation approach by (Swift and Janacek, 1991).

Several researchers as well have used different distribution to model non-normal time series data. Despite the range of studies that focused on the specification and estimation of time series data, most research work did not indicate the model differently for time series data that are non-normally distributed. Therefore, in the research work, firstly the distribution of the dataset will be identified using Anderson Darling Statistics and a non-normal autoregressive model based on the reparametrised classical autoregressive model will be proposed. The coefficients of the proposed model will be estimated using Maximum likelihood estimation method. The optimal models will be obtained using Akaike Information Criteria and the efficiency and performance of the proposed model will be determined by comparing the values of forecast evaluations of the classical autoregressive and the proposed models with respect to forecast accuracy capabilities.

## 2. MATERIALS AND METHODS

### — Generalized gamma (GG) distribution

Let $G(x, \alpha, \gamma, k)$ be the cdf of generalized gamma distribution written as

$$f(t) = \frac{\lambda\rho\gamma}{\alpha\Gamma(k)}\left(\frac{t}{\alpha}\right)^{\gamma k-1} \exp\left[-\left(\frac{t}{\alpha}\right)^{\gamma}\right]\left\{\gamma, \left[k, \left(\frac{t}{\alpha}\right)\right]\right\}^{\lambda-1}\left(1-\left\{\gamma_1, \left[k, \left(\frac{t}{\alpha}\right)^{\gamma}\right]\right\}^{\lambda}\right)^{\rho-1} \quad (1)$$

If $\rho = 1$ in equation (1), the GG distribution reduces to

$$f(t) = \frac{\lambda\gamma}{\alpha\Gamma(k)}\left(\frac{t}{\alpha}\right)^{\gamma k-1} \exp\left[-\left(\frac{t}{\alpha}\right)^{\gamma}\right]\left\{\gamma, \left[k, \left(\frac{t}{\alpha}\right)\right]\right\}^{\lambda-1} \quad (2)$$

and this is an exponential GG distribution.

### — Parameter estimation for generalized gamma distribution

In order to estimate the parameters of the generalized gamma distribution, the log-likelihood function of equation (2) is taken to obtain

$$L_{gg}(\alpha, \kappa, \gamma)$$
$$= \ln\gamma - \gamma\alpha\ln\kappa - \ln\Gamma(\kappa) + 1/n\sum_{\iota=1}^{n}[(\gamma\alpha-1)\ln X_\iota - (X_\iota/\kappa)^{\gamma}] \quad (3)$$

The partial derivatives of $L_{gg}$ with respect to $\alpha$, $\kappa$ and $\gamma$ are respectively;

$$0 = -\psi(\alpha) - \gamma\ln\kappa + \gamma/n\sum_{\iota=1}^{n}\ln X_\iota \quad (4)$$

$$= -\alpha + 1/n\sum_{\iota=1}^{n}(X_\iota/\kappa)^{\gamma} \quad (5)$$

$$= 1/\gamma + \alpha/n\sum_{\iota=1}^{n}\ln(X_\iota/\kappa) - 1/n\sum_{\iota=1}^{n}(X_\iota/\kappa)^{\gamma}\ln(X_\iota/\kappa) \quad (6)$$

where $\psi(.) = d\ln\Gamma(x)/dx$ is the digamma function solving the system of the equation gives the Maximum likelihood (ML) estimators of $(\alpha, \kappa, \gamma)$.

From equation (5)

$$K = 1/n\textstyle\sum_{\iota=1}^{n}(X_\iota/\kappa)^{\gamma} \quad (7)$$

$$K = \frac{\sum_{\iota=1}^{n}X_\iota}{(\alpha n)^{1/\gamma}} \quad (8)$$

which can be written as;

$$K = \left(\frac{\sum_{\iota=1}^{n}X_\iota{}^{\gamma}}{\alpha n}\right)^{1/\gamma} \quad (9)$$

Substituting equation (9) into equation (6) to attain

$$0 = 1/\gamma + \alpha/n\sum_{\iota=1}^{n}\ln\left(\frac{X_\iota}{\left[\frac{\sum X_\iota^\gamma}{\alpha n}\right]^{1/\gamma}}\right) - 1/n\sum_{\iota=1}^{n}\left(\frac{X_\iota}{\left[\frac{\sum X_\iota^\gamma}{\alpha n}\right]^{1/\gamma}}\right)\ln\left(\frac{X_\iota}{\left[\frac{\sum X_\iota^\gamma}{\alpha n}\right]^{1/\gamma}}\right) \quad (10)$$

Multiply both sides of equation (10) by nγ to obtain

$$0 = n + \alpha\gamma\sum_{\iota=1}^{n}\ln\left(\frac{X_\iota}{\left[\frac{\sum X_\iota^\gamma}{\alpha n}\right]^{1/\gamma}}\right) - \gamma\sum_{\iota=1}^{n}\left(\frac{X_\iota}{\left[\frac{\sum X_\iota^\gamma}{\alpha n}\right]^{1/\gamma}}\right)\ln\left(\frac{X_\iota}{\left[\frac{\sum X_\iota^\gamma}{\alpha n}\right]^{1/\gamma}}\right)$$

$$\alpha = \frac{n - \gamma\sum_{\iota=1}^{n}\left(\frac{X_\iota}{\left[\frac{\sum X_\iota^\gamma}{\alpha n}\right]^{1/\gamma}}\right)\ln\left(\frac{X_\iota}{\left[\frac{\sum X_\iota^\gamma}{\alpha n}\right]^{1/\gamma}}\right)}{\gamma\sum_{\iota=1}^{n}\ln\left(\frac{X_\iota}{\left[\frac{\sum X_\iota^\gamma}{\alpha n}\right]^{1/\gamma}}\right)} \quad (11)$$

By rationalisation, equation (11) becomes

$$\alpha = \frac{n\sum_{\iota=1}^{n}X_\iota^\gamma}{n\gamma\sum_{\iota=1}^{n}X_\iota^\gamma\ln X_\iota - \gamma\sum\ln X_\iota\sum X_\iota^\gamma} \quad (12)$$

— **The Gamma autoregressive model**

The classical AR(1) process with the assumption of normality is given as

$$X_t - \mu = \phi(X_{t-1} - \mu) + e_t \quad (13)$$

where $e_t \sim N(0, \sigma^2)$ is independent identically distributed and $|\phi| < 1$

Suppose $X_t \sim GG(\alpha, k.\gamma)$ where,

$$E(X) = \frac{\alpha\Gamma\left(k+\frac{1}{\gamma}\right)}{\Gamma(k)} \quad (14)$$

$$Var(x) = \frac{\alpha^2\Gamma\left(k+\frac{1}{\gamma}\right)}{\Gamma(k)} - \left(\frac{\alpha\Gamma\left(k+\frac{1}{\gamma}\right)}{\Gamma(k)}\right)^2 \quad (15)$$

Then, by equating (14) to

$$\frac{E(e_t)}{1-\phi_1}$$

this leads to

$$\frac{\alpha\Gamma\left(k+\frac{1}{\gamma}\right)}{\Gamma(k)} = \frac{E(e_t)}{1-\phi_1} \quad (16)$$

Hence

$$\hat{\phi}_1 = \frac{\alpha\Gamma\left(k+\frac{1}{\gamma}\right) - \Gamma(k)E(e_t)}{\alpha\Gamma\left(k+\frac{1}{\gamma}\right)} \quad (17)$$

Therefore, the Gamma Autoregressive model of order (1), GAR(1) is then obtained as,

$$X_t = \frac{\alpha\Gamma\left(k+\frac{1}{\gamma}\right) - \Gamma(k)E(e_t)}{\alpha\Gamma\left(k+\frac{1}{\gamma}\right)}X_{t-1} \quad (18)$$

For $X_t \sim GG(\alpha, k, 1)$ equation (18) becomes,

$$X_t = \frac{\alpha\Gamma(k+1) - \Gamma(k)E(e_t)}{\alpha\Gamma(k+1)}X_{t-1} \quad (19)$$

Using Stirling's formula (Jacques, 1991), equation (19) becomes

$$X_t = 1 - \delta a k^{1/\gamma}X_{t-1} + \varepsilon_t \quad (20)$$

— **Anderson darling statistics**

Anderson Darling Statistics (Stephens, 1974) will be used to determine the distribution that the data-set follows and the test Statistic is given as

$$A^2 = -N - S \quad (21)$$

$$S = \sum_{I=1}^{N} \frac{(2i-1)}{N}\left[\ln F(Y_i) + \ln\left(1 - F(Y_{N+1-i})\right)\right] \quad (22)$$

where F is the Cumulative distribution function of the specified distribution, $Y_i$ are the ordered data, N is the sample size and A is the calculated values.

— **Autoregressive model**

The AR(p) model is defined as

$$X_t = \varphi_1 X_{t-1} + \varphi_2 X_{t-2} + \cdots + \varphi_p X_{t-p} + \varepsilon_t \quad (23)$$

where $\varphi_1, \varphi_2, \dots, \varphi_p$ are the parameters of the model and $\varepsilon_t$ is white noise.

— **Akaike information criteria**

In this research work, Akaike information criteria (AIC) defined as

$$AIC = 2k - 2(\text{Log-Likelihood}) \quad (24)$$

where k is the number of model parameters and Log-likelihood is a measure of model fit. This will be used to identify the optimal model and as well as determine the better model. This will be done by choosing the model with the smallest value of AIC.

— **Forecast evaluations**

The forecast evaluations used in the research article to measure the likely accuracy of the forecast are Mean absolute error (MAE) defined as

$$MAE = \frac{1}{h+1}\sum_{t=s}^{h+s}\left(\widehat{X}_t - X\right)^2 \quad (25)$$

Root mean square forecast error (RMSE) defined as

$$RMSE = \sqrt{\frac{1}{h+1}\sum_{t=s}^{h+s}\left(\widehat{X}_t - X_t\right)^2} \quad (26)$$

and the Mean absolute percentage error is defined as

$$MAPE = \frac{100}{h+s}\sum_{t=s}^{h+s}\left|\frac{\hat{y}_t - y_t}{\hat{y}_t}\right| \quad (27)$$

where $t = s, 1+s, \dots, h+s$. The actual and predicted values for corresponding t values are denoted by $\hat{y}_t$ and $y_t$ respectively.

## 3. RESULTS AND DISCUSSIONS

The time plot of Nigerian Gross domestic product (GDP) obtained for Central Bank of Nigeria Bulletin from 1962 to 2019 is presented in Figure 1. The time plot revealed that the data is not stationary. This means that the series is not normally distributed and the mean and variance are not constant over time. The Descriptive Statistics summary of Nigerian GDP growth is presented in table 1. The result showed the mean of the variable to be 3.99123, the skewness is 2.52599 and this implied that the variable is positively skewed. The kurtosis value is 5.18358 and this implied the series is not normally distributed.

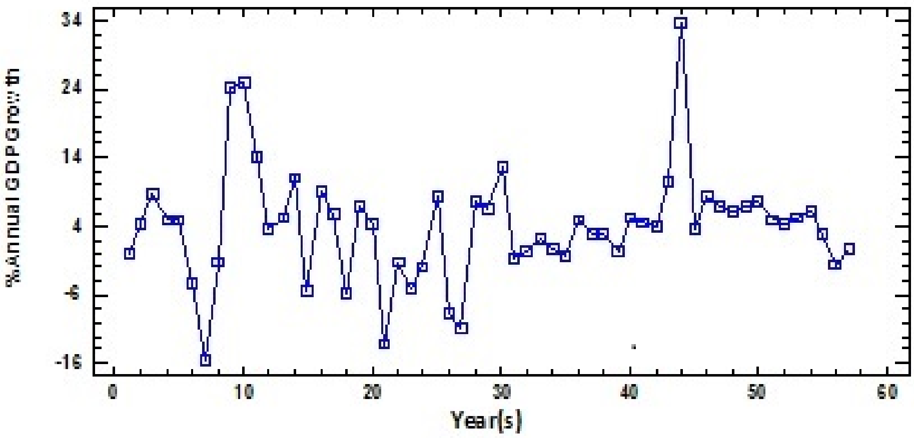


Figure 1: Time plot of Nigeria gross domestic product from 1962 to 2019

Table 1: Descriptive statistics for gross domestic rate

| Count | Average | Standard Deviation | Minimum | Maximum | Skewness | Kurtosis | Range |
|---|---|---|---|---|---|---|---|
| 57 | 3.99123 | 8.23033 | -15.7 | 33.7 | 2.52599 | 5.18358 | 49.4 |

The Anderson Darling statistics was carried out based on non-Gaussian distributions; Gamma, Chi-square, Geometric, Exponential and Normal distribution. The result of the test in table 2 revealed that the series followed Gamma distribution, this is so because it has the least value in Anderson Darling Statistics. The summary of the parameters of the fitted Gamma distributions for GDP series is given in table 3. The parameters of the fitted gamma distribution were $(\text{Shape} = k = 6.831 \text{ and } \text{Scale} = \alpha = 0.730)$.

Table 2: Anderson darling statistics for GDP growth

| Distribution | Gamma | Poisson | Chi-Square | Geometric | Exponential | Normal |
|---|---|---|---|---|---|---|
| GDP | 0.377 | 0.565 | 0.743 | 0.490 | 0.422 | 0.621 |

Table 3: Parameters of fitted Gamma distribution

| Parameters | Estimate |
|---|---|
| Shape | 6.831 |
| Scale | 0.730 |

Akaike Information Criteria (AIC) obtained in table 4 for fitted AR(1), AR(2), AR(3), AR(4), AR(5), AR(6) and AR(7) model was used to determine the optimal autoregressive model as AR(2) since AR(2) has the least AIC value. Hence, the most suitable model for GDP series is AR(2). The AR(2) model obtained is given as:

$$X_t = 0.22788X_{t-1} + 0.086333X_{t-2}$$

The values of Akaike Information Criteria (AIC) obtained and displayed in table 5 for fitted GAR(1), GAR(2), GAR(3), GAR(4), GAR(5), GAR(6) and GAR(7) was used to show GAR(2) as the optimal model. The GAR(2) model obtained is:

$$X_t = -0.0169903X_{t-1} + 0.924414X_{t-2}$$

It is quite obvious based on the values of AIC in table 4 and five that AR(2) and GAR(2) are the optimum models. Comparatively, the AIC value for AR(2) model is 0.5405 and it is higher than the corresponding AIC value for GAR(2) model as shown in table 6. Thus, this indicated GAR(2) model is more stable than AR(2) model. The forecast precision for both classical autoregressive model AR(2) and Gamma autoregressive model GAR(2) were tested based on forecast evaluations. Based on the results in table 7, that is, the value of MAE (0.2497), RMSE (0.2267) and MAPE (1.9502) for GAR(2) is lower than that of AR(2) and this indicated that GAR(2) is the better model for modelling Nigerian gross domestic product.

Table 4: Model selection criteria for classical autoregressive models

| Autoregressive models | Akaike information criteria (AIC) |
|---|---|
| AR(1) | 0.7823 |
| AR(2) | 0.5405 |
| AR(3) | 0.6627 |
| AR(4) | 0.5805 |
| AR(5) | 0.5972 |
| AR(6) | 0.7823 |
| AR(7) | 0.6250 |

Table 5: Model selection criteria for Gamma autoregressive model

| Gamma autoregressive model | Akaike Information Criteria (AIC) |
|---|---|
| AR(1) | 0.4298 |
| AR(2) | 0.2285 |
| AR(3) | 0.4673 |
| AR(4) | 0.3850 |
| AR(5) | 0.3972 |
| AR(6) | 0.5461 |
| AR(7) | 0.4899 |

Table 6: Comparison of AR(2) and GAR(2) models based on values of AIC

| Model | AIC |
|---|---|
| AR(2) | 0.5405 |
| GAR(2) | 0.2285 |

Table 7: Comparison of AR(2) and GAR(2) Models based on forecast evaluations

| Model | GAR(2) | AR(2) |
|---|---|---|
| Mean Absolute Error | 0.5129 | 0.4049 |
| Root Mean Square Error | 0.4269 | 0.3287 |
| Mean Absolute Percentage Error | 112.3074 | 189.8035 |

## 4. CONCLUSIONS

Nigerian gross domestic product from 1962 to 2019 obtained from Central Bank of Nigeria Bulletin was shown to follow a Gamma distribution based on the values of the Anderson Darling Statistics. A reparameterisation method was used to derive a Gamma Autoregressive (GAR) model and this was used to analyse and mode Nigerian gross domestic product. The coefficients of classical and Gamma autoregressive models were estimated using the Maximum likelihood method. The Akaike Information Criteria (AIC) was used to select AR(2) and GAR(2) as the optimal models but GAR(2) was chosen the better model because it has the least value of AIC. The comparison between AR(2) and GAR(2) models based on the values of Mean absolute error (MAE), Mean absolute prediction error (MAPE) and Root mean square error (RMSE) indicted that GAR(2) will yield a more accurate forecast than AR(2). In essence, GAR model is a viable alternative and better model that can be used to model and forecast gross domestic product growth.

**References**

[1] Abdulraheem, A.: Sectorial contributions to gross domestic product in Nigeria, 1971- 2005, Nigeria Forum, 1 – 12, 2011.
[2] Fahrmeir, L.: State Space Modeling and Conditional Mode estimation for Categorical time series, In New directions in time series analysis, Part 1, New York, Springer, 1992.
[3] George, U. and Iboik E. U.: Sectoral contributions to Nigeria gross domestic product using a VAR approach, Global Journal of Pure and Applied Sciences, 21, 137 – 143, 2015.
[4] Gujarati, D. N. (2004). Basic Econometrics, 4th ed., McGraw Hill, 2004.
[5] Hajrajabi , A. and Maleki, M.: Nonlinear semiparametric autoregressive model with finite mixtures of scale mixtures of skew normal innovations, Journal of Applied Statistics, 46(11), 2010 – 2029, 2019.
[6] Harvey, A. C. and Fernandes, C.: Time series models for count or qualitative observations (with discussion), Journal of Business and Economic Statistics, 7, 407 – 422, 1989.
[7] Johannesson, P., Padgoski, K., Rychlik, I. and Shariati, N.: AR(1) Time series with Autoregressive Gamma Variance for Road Topology Modeling, Working Paper in Statistics, No. 2015:5, Department of Statistics, Lund University, 2015
[8] Kashiwagi, N. and Yanagimoto, T.: Smoothing serial count data through a state-space model, Biometrics, 48, 1187 – 1194, 1992.
[9] Langberg, N. A. and Stoffer, D. S.: Moving average models with bivariate exponential and geometric distributions, Journal of Applied Probability, 24, 48 – 61, 1987.
[10] Liseo, B. and Loperfido, N.: A note on Reference Priors for the Scalar Skew Normal Distribution, Journal of Statistical Planning and Inference, 136, 373 – 389, 2006.
[11] Liu, J., Kumar, S. and Palomar O. P.: Parameter estimation of Heavy tailed AR model with Missing data via Stochastic EM, arXiv:1809.07203 [stat.AP], 1 – 18, 2018.
[12] Ma, Z. and Leijon, A.: A Variational Bayesian NMF model for bounded support data, IEEE Transactions on Pattern Analysis and Machine Intelligence, 37(4), 876 – 889, 2011.
[13] Okereke, O. E. and Bernard C. B.: Forecasting Gross Domestic product in Nigeria using Box-Jenkins methodology, Journal of Statistical and Econometric methods, 3(4), 33 – 46, 2014.
[14] Popovici, G.: On Three Classes of Time Series Involving Exponential Distribution, Mathematical Reports, 12, 45 – 57, 2010.
[15] Stoffer, D. S.: Walsh-Fourier analysis of Discrete-valued Time series. Journal of Time Series Analysis, 8, 449 – 467, 1987.
[16] Stephens, M. A.: EDF Statistics for Goodness of fit and some comparisons, Journal of American Statistical Association, 69, 730 – 737, 1974.
[17] Swift, A. and Janacek, G. J.: Forecasting Non-normal time series, Journal of Forecasting, 10, 501–520, 1991.
[18] Xu, Z.: Modeling Non-Gaussian Time-correlated Data Using Nonparametric Bayesian Method, Unpublished Master thesis, Ohio State University, U.S.A, 1 – 140, 2014
[19] Wei, W. S.: Time Series Analysis Univariate and Multivariate Methods. Pearson Addison Welsley, 2006

**ANNALS of Faculty Engineering Hunedoara – International Journal of Engineering**
**ISSN 1584 - 2665 (printed version); ISSN 2601 - 2332 (online); ISSN-L 1584 - 2665**