International Journal of Progressive Sciences and Technologies (IJPSAT)
ISSN: 2509-0119.

http://ijpsat.ijsht-journals.org
Vol. 29 No. 2 November 2021, pp.424-429

Forecasting Food Inflation with ARIMA and ELM Techniques

Saminu Umar and M. K. Aminu

Department of Mathematics and Statistics
Umaru Ali Shinkafi Polytechnic, Sokoto
Sokoto, Nigeria

Abstract— This study focuses on modelling monthly food inflation rate in Nigeria using Autoregressive integrated moving average (ARIMA) model and Extreme Learning Machine (ELM) algorithm. ARIMA (2, 1, 2) was found to be the best fit among the considered ARIMA models based on its minimum value of AIC and BIC selection criterions. Forecast of Nigerian food inflation rate obtained from the selected ARIMA model and ELM were compared in terms of, MAE, MSE and RMSE, the computational results on the data demonstrate that the ELM provides better forecasts than the RIMA model. The study, therefore, recommends that stakeholders and authorities should think in the line of using ELM in conducting food inflation forecasts in Nigeria given its higher predictive and forecasting ability.

Keywords—Time Series, Inflation, ARIMA, feedforward neural networkss, Extreme learning machine.

I. INTRODUCTION

Inflation is a persistent rise in the price level of goods and services in an economy over a period of time, it is mainly measured as inflation rate, which is the percentage change in a price index over time. Maintenance of price stability is one of the macroeconomic challenges facing the Nigerian government in its economic history, as the country experience an unbearable hike in prices especially that of food commodities. The skyrocketing food prices in major cities across the country have had huge effects on the budget of households and firms. This development is not only worrisome but it also creates uncertainties in the economy, weakens investors' confidence and affects expectations. Inflation is undeniably one of the leading and most dynamic macro-economic issues confronting almost all economies of the world. The importance of inflation rate to the economic growth of a country makes imperative to consider and study its dynamism, [13]. For this reason, many researchers and economists have applied various methods in an attempt to model and forecast inflation rates.

Forecasting inflation can be done using time series analysis; there are a number of approaches available for modelling and forecasting univariate economic time series such as inflation rate. One approach, is the conventional techniques of Box-Jenkins, specifically the Autoregressive integrated moving average (ARIMA) model which assume the variable to be stochastic and apply statistical theory to model, usually in a linear frame expressed in terms of past values of itself (the autoregressive component) plus current and lagged values of a 'white noise' error term (the moving average component). Another approach is the artificial neural networks (ANN) or simply neural networks (NNs) which regards the time series process as a low dimensional non-linear deterministic system and exercises a parameterized nonlinear function that can approximate the system for prediction purpose. [10] presented an algorithm for a single layer feed forward NN, identified as Extreme Learning Machine (ELM) capable of solving problems caused by gradient descent-based algorithms like back propagation which are applied to artificial neural networks (ANNs).

The variables in the conventional approach are, in many cases, transformed and differenced. In the non-linear modelling, such preprocessing of the data is not done, as they would remove crucial information about the evolution of the complex system that generates the observed data. Therefore, it should be borne in mind that the data being modelled is different under the two approaches. The main purpose of this study is to understand, model and forecast food inflation dynamics in Nigeria with ARIMA and ELM.

Numerous researches have been conducted associated with utilizing ARIMA and ELM algorithm efficiently for determining problems in various fields including prediction of Inflation markets. The adaption of ARIMA and ANNs enabled [9] to forecast of short-run inflation rate in Nigeria and compared the predictive ability of the two forecasting models. The study found that ANNs has higher predictive or forecasting ability compared to ARIMA (1, 1, 1). In addition, the results of the forecasted monthly inflation for 2018 showed that inflation rate would be slightly lower in the early months.

ANN was applied by for short-term forecasting of agriculture commodity price in Indian based on time series data by [2], the researchers compared forecast performance of ANNs with the result of ARIMA model and observed that the ANNs yields better forecasts than the ARIMA model by considering two forecasting performance parameters MAPE and RMSPE. The study by [7] used a fixed scheme pseudo-out-of-sample forecasting approach to compared the univariate autoregressive model and neutral networks models based on their success at forecasting inflation rate using U.S GDP deflator data, with the aid of some model selection approach, the researcher found out that the NNs outperform AR.

The application of ELM algorithm for chaotic time series was explore by [15] through experiments using sigmoid, sin and hardlim activation functions, the researchers demonstrated that the ELM algorithm can achieve high prediction accuracy, they conclude that ELM is a promising method for time series prediction problems. The study by [5] used E-ELM to analyze and forecast the trend component of time series of total accumulative displacement, the inputs of E-ELM were with respect to selected component based on the combination of GRA and expert knowledge. Performances of the models was evaluated using real data from Baishuihe landslide in the three Gorges Reservoir of China, result obtained provided a good representation of the measured slide displacement behavior. The researchers successfully predicted the obvious deformation of landslide, which was able to provide early warnings. Thus, the method has a good perspective in application and further development. There is perhaps no known attempt to model food inflation in Nigeria explicitly by ARIMA and ELM models.

II. MATERIAL AND METHODS

The data for this study is of the official monthly food inflation in Nigerian. The data is obtained from Central Bank of Nigeria (CBN) with sample period ranging from January 2003 to December 2020 consisting of 222 observations.

2.1 *Autoregressive Integrated Moving Average (ARIMA) Model*

The Autoregressive Integrated Moving Average (ARIMA) model is a combination of two univariate time series models, which are the Autoregressive (AR) model and the Moving Average (MA) model. A general notation for ARIMA (popularly known as Box-Jenkins Methodology [4] is (ARIMA (p,d,q)). Where p, d, and q are integers greater than or equal to zero denoting the number of autoregressive terms, the number of moving average terms and the number of times a series must be differenced or integrated to induce stationarity respectively. As this methodology is applicable only to a differenced stationary data series, that is where the mean, the variance, and the autocorrelation function of the series remain constant through time.
An autoregressive (AR) model is one in which Y_t depends only on its lags or t past values. That is, a pth-order autoregressive process expresses a dependent variable as a function of its own lags or past values is expressed as:

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots \phi_p Y_{t-p} + \varepsilon_t \quad (1)$$

Where Y_t is the response variable being forecasted at time t

$Y_{t-1}, Y_{t-1}, \ldots\ Y_{t-p}$ is the response variable at time t-1, t-2…t-p respectively

$\phi_1, \phi_2, \ldots\ \phi_p$ are the parameter estimates

ε_t is the error term at time t

A moving average (MA) model is one in which Y_t depends only on the past values of t the white noise disturbance (error) terms. That is, a qth-order moving-average process expresses a dependent variable q error terms is expressed as:

$$Y_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots \theta_q \varepsilon_{t-q} \quad (2)$$

Where Y_t is the response variable being forecasted at time t

μ is the constant mean of the process

$\varepsilon_{t-1}, \varepsilon_{t-1}, \ldots\ \varepsilon_{t-q}$ is the response variable at time t-1, t-2…t-q respectively

$\theta, \theta_2, \ldots\ \theta_q$ are the parameters estimates

ARIMA model of a time series Y_t that is integrated of order d, is represented in the form:

$$\phi(\mathrm{L})\Delta^d Y_t = \mu + \theta(L)\varepsilon_t \quad (3)$$

With L representing lag operator, $\nabla^d = (1 - L)^d$ is the itegrated of order d, $\phi\ and\ \theta$ are standard AR and MA polynomials of order p and q respectively and ε_t is independently and normally distributed with zero mean and constant variance. In the event a series Y_t is stationary in levels, the ARIMA model is reduced to ARMA.

2.2 *Model Selection Criteria*

The best ARIMA (p, d, q) model will be selected using the Akaike Information Criterion (AIC) [1] and Bayesian Information Criterion (BIC) [16]. These information criterion attempts to measure of the goodness of fit of an estimated statistical model by judging a model with how close its fitted values tend to be to the true values in terms of the expected values. They also includes a penalty that is an increasing function of the number of estimated parameters and discourages over fitting.

Given a data set, several competing models may be ranked according to their AIC or BIC with the one having the lowest information criterion value being the best. In the general case, the AIC and BIC take the form as shown below:

$$AIC = 2k - 2\log(L) \tag{4}$$

$$BIC = -2\log(L) + K\ln(n) \tag{5}$$

2.3 *Single Hidden Layer Feed-Forward Neural Networks*

Standard SLFNs with L hidden nodes with the activation function of $g(x): R \to R\ G(a_i, b_i, x)$ are mathematically modelled as:

$$\sum_{i=1}^{L} \beta_i g_i(x_j) = \sum_{i=1}^{L} \left(\beta_i g\ (w_i . x_j + b_i\right) = o_j \tag{6}$$

Where $w_i = [w_{i1}, w_{i2}, \dots, w_{in}]^T$ is the weight vector connecting the ith hidden node and the input nodes, $\beta_i = [\beta_{i1}, \beta_{i2}, \dots, \beta_{in}]^T$ is the weight vector connecting the ith hidden node and the output nodes, b_i is the threshold of the ith hidden node, $w_i . x_j$ denotes the inner product of the w_i and x_j and $o_j = [o_{j1}, o_{j2}, \dots, o_{jm}]^T$ is the jth output vector of the SLFN.

The standard SLFNs with L hidden nodes can approximates these N samples with zero error mean that $\sum_{j=1}^{N} \|o_j - t_j\| = 0$, i.e., there exist β_i, w_i and b_i such that

$$\sum_{i=1}^{L} \left(\beta_i g\ (w_i . x_j + b_i\right) = t_j, \qquad j = 1,2, \dots N \tag{7}$$

Equations (7) rewritten in the following matrix form as:

$$\mathbf{H}\beta = \mathbf{T} \tag{8}$$

Where

$$\mathbf{H}(\boldsymbol{w_1}, \dots, \boldsymbol{w_L}, b_1, \dots, b_L, x_1, \dots, x_n) = \begin{bmatrix} g(w_1 . x_1 + b_1) & \cdots & g(w_L . x_1 + b_L) \\ \vdots & \ddots & \vdots \\ g(w_1 . x_N + b_1) & \cdots & g(w_L . x_N + b_L) \end{bmatrix}$$

$$\beta = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_L^T \end{bmatrix} \quad and \quad T = \begin{bmatrix} t_1^T \\ \vdots \\ t_N^T \end{bmatrix}$$

Note that H is the hidden layer output matrix of the SLFN, where the i^{th} column of H represents the i^{th} hidden node output with respect to inputs $x_1, x_2, \dots, x_n$. Further, observe that the matrix H need not be a square matrix [8].

2.4 *Extremely learning machine (ELM)*

Extreme learning machine proposed [10] based on a single-hidden layer feedforward networks (SLFNs) randomly choose hidden nodes and analytically determines the output weights of SLFNs. The proposed algorithm performs better and speeds up the learning process thousands of times faster with higher generalization ability than the conventional learning algorithms for the feedforward neural networks. Like the unsupervised ANNs, ELM is able to analyze all the networks functions, which minimizes human involvement. This method presents an efficient algorithm with various important features such as convenience, prompt learning speed, effective performance, capability to use most nonlinear stimulation and kernel functions and ability to avoid difficulties, such as the stopping criteria [11].

The ELM is based on the following theorems:

- Theorem 1: In this theorem, [12] presented a standard SLFN with L hidden nodes as well as an activation function $g: R \to R$ that is extremely differentiable in any interval, for N distinct samples input vectors $\{x_i |\ x_i \epsilon R^n$, and$\{t_i | t_i \epsilon R^m\}$, for any w_i and b_i randomly chosen from any intervals of R^n and R^m, respectively, according to any continuous probability distribution. Furthermore, with probability one, the hidden layer output matrix H of the SLFN is invertible and $||H\beta' T|| = 0$.
- Theorem 2: In this theorem, [3] conferred for any small positive value ε >0 and activation function $g(x): R \to R$ that is infinitely differentiable in every specification; there exists $L \le N$ such that for N arbitrary distinct input vectors $\{x_i |\ x_i \epsilon R^n, \}and\{t_i | t_i \epsilon R^m\}$ for any w_i and b_i randomly chosen from any intervals of R^n and R^m, respectively, according to any continuous probability distribution, then with probability one $\left||H_{N \times L}\ \beta_{L \times m} - T_{N \times m}\right|| < \varepsilon$

ELM Algorithm: given a training set $\aleph = \{x_i | \, x_i \in R^n,\}and\{t_i | t_i \in R^m, i = 1, \dots, N\}$, L the number of hidden neurons and the activation function $g(x)$.

- Step 1: Randomly assign input weights w_i and bias b_i, $\quad i = 1, \dots, L$
- Step 2 Determine the matrix H defined
- Step 3: Calculate $H^{\dagger}$
- Step 4: Calculate the output weight β given by:

$$\beta = H^{\dagger} T$$

Where $T = [t_1, \dots, t_N]^T$ is the target value matrix and $H^{\dagger}$ is the Moore-Penrose generalized inverse of the hidden layer output matrix H which could be estimated by numerous methods, including orthogonal projection, orthogonalization, iterative methods, Singular Value Decomposition (SVD) etc. [17]. The least-square (LS) solution of $\beta = H^{\dagger} T$ is unique with the smallest norm of weights and minimum training error among all the LS solutions

2.5 *Forecasting and evaluation*

Monthly forecasts of Nigerian food inflation were generated from ARIMA (2, 1, 2) model and ELM for the purpose of model cross-validation. To achieve this objective, the quality of the obtained forecasts was is tested using mean squared error (MSE) and root mean square error (RMSE) and Mean Absolute Error defined as;

$$MSE = \frac{1}{n} \sum_{t=1}^{n} (Y_t - \hat{Y}_t)^2 \qquad (9)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^{n} (Y_t - \hat{Y}_t)^2} \qquad (10)$$

$$MAE = \frac{1}{n} \sum_{t=1}^{n} \left| (Y_t - \hat{Y}_t) \right| \qquad (11)$$

Where Y_t is the actual observed value for the period t and $\hat{Y}_t$ is the forecasted value for the same period and n is length of the test set

III. Data Analysis

To investigate the order of integration of the time series, Augmented Dickey-Fuller (ADF) and Phillips-Perron (PP) unit root/Stationarity tests were conducted. The results are presented in Table I.

TABLE I. UNIT ROOT TEST BEFORE AND AFTER DIFFERENCING

Test	Before differencing		After differencing	
	t-statistics	p-value	t-statistics	p-value
ADF Test	-0.799	0.393	-6.00	0.01
PP test	-2.52	0.362	-152	0.01

At 5% level, the (ADF) test and the (PP) test fails to reject the null hypothesis of unit root and thus it can be concluded that the rat e of inflation is not stationary, for this reason, a first order lagged difference from the original series is obtained and tested for stati onarity. After first differencing, the p-value for ADF and PP test indicate that the series is now stationary.

3.1 *Model selection on ARIMA*

TABLE II. DIFFERENT ARIMA (P, D, Q) MODEL FITTED

ARIMA	AIC	BIC	log likelihood
(1, 1, 1)	974.729	984.756	-484.36
(1, 1, 2)	971.3497	984.7191	-481.67
(2, 1, 1)	968.186	981.5553	-480.09
(2, 1, 2)	962.5382	979.2498	-476.27
(3, 1, 2)	968.6165	985.3282	-479.31

Based on the model selection criteria results presented in Table II, ARIMA (2, 1, 2) was selected.

TABLE III. RESULTS OF ARIMA (2 1 2) MODEL

Parameter	Coefficients	s.e	t.value	p.value
ar1	1.000	0.0308	32.50	0.00e+00
ar2	0.247	0.0312	7.92	1.49e-13
ma1	-0.939	0.0303	-30.98	1.76e-79
ma2	-0.113	0.0275	-4.09	6.09e-05

The results of the ARIMA (2, 1, 2) model presented in Table III indicates that, all the AR component and MA component are statis tically significant, implying that the choice of ARIMA (2, 1, 2) is appropriate and ideal.

3.2 *Diagnostic checking*

The Ljung-Box test and ARCH-LM tests test was conducted to test the autocorrelation. The following results were obtained:

TABLE IV. DIAGNOSTIC TESTS

Tests	t-statistic	p value
ARCH-LM	30.0	2.15e-04
Ljung-Box	4.65	3.25e-01

The two different tests that have been conducted are to check for the adequacy of the model, as shown in table 3 above, the p-value in both of the tests is less than the test statistic. Thus, it can be concluded that there is no any evidence of autocorrelations in the forecast errors of the fitted ARIMA (2, 1, 2), model hence the model is adequate to fit the data.

3.3 *Comparison of model performance*

To examine the fitness of the ARIMA model and theELM model, all the models have been used to make out of sample forecast and their performance errors results are reported and summarized in Table V

TABLE V. FORECAST PERFORMANCE COMPARISON FOR ARIMA AND ELM MODELS

Model	MAE	MSE	RMSE
ELM	1.991725	5.806122	2.409589
ARIMA(2, 1, 2)	2.046667	6.106	2.471032

IV. CONCLUSION

The basic objective of this study was to investigate the forecast performance of the ARIMA model and ELM algorithm in forecasting monthly food inflation for Nigeria based on historical monthly data obtained from the CBN. Results obtained from the study indicated ARIMA (2, 1, 2) as the best ARIMA model based on its minimum values of AIC and BIC. The study also validated and statistically tested that the successive residuals in the fitted ARIMA were not correlated, also, the residuals appeared to be normally distributed with the mean zero and constant variance. Hence, was concluded that the selected ARIMA (2, 1, 2) is a satisfactory predictive model for the food inflation rate in Nigeria. Out of sample forecast of the Nigerian food inflation rate obtained from the ARIMA and ELM indicated an increase in the food inflation. Computational results on the forecasts from the models in terms of RMSE, MAE and MSE demonstrate that the ELM provides better forecasts than the RIMA model. The study, therefore, recommends that stakeholders and authorities should think in the line of using ELM in conducting food inflation forecasts in Nigeria given its higher predictive and forecasting ability.

ACKNOWLEDGMENT

The researchers would like to thank the authors whose works were cited in this paper and the International Journal of Progressive Sciences and Technologies.

REFERENCES

[1] Aikake, H., "A new look at the statistical model identification", IEEE Transactions on Automatic Control, 1974, vol. 19(6). pp. 716–723

[2] Anil Kumar Mahto, M. Afshar Alam, Ranjit Biswas, Jawed Ahmed, and ShahImran Alam. Short-Term Forecasting of Agriculture Commodities in Context of Indian Market for Sustainable Agriculture by Using the Artificial Neural Networks. Hindawi Journal of Food Quality Volume 2021, Article ID 9939906, 13 pages

[3] Annema, A.; Hoen, K.; Wallinga, H. Precision requirements for single-layer feedforward neural networkss. In Proceedings of the Fourth International Conference on Microelectronics for Neural Networkss and Fuzzy Systems, Turin, Italy, 26–28 September 1994; pp. 145–151

[4] Box, G. E. P., and Jenkins, G. M., Time Series Analysis: Forecasting and Control, revised edition, San Francisco: Holden Day, 1976

[5] Cheng Lian, Zhigang Zeng, Wei Yao and Huiming Tang., Ensemble of extreme learning machine for landslide displacement prediction based on time series analysis, Neural Comput & Applic, 2014, 24:99–107

[6] Dickey, D. A., and Fuller, W. A., „Distribution of the Estimators for Autoregressive Time Series with a Unit Root", Journal of the American Statistical Association, (1979), 74,(Pg 427–31).

[7] Emi Nakamura., Inflation forecasting using a neural networks. Economics letter 86. 2005, 373-378

[8] G. B. Huang, Learning capability and storage capacity of two-hidden-layer feedforward networks. IEEE Trans Neural Netw ((2003), 14(2):274–281

[9] Gideon G. Goshit, Paul Terhemba Iorember, short-term inflation rate forecasting in Nigeria: an autoregressive integrated moving average (ARIMA) and artificial neural networks (ANN) models. Journal of economic and financial issues, (2018), volume 4, number

[10] Huang, G.-B.; Zhu, Q.-Y.; Siew, C.-K. Extreme learning machine: A new learning scheme of feedforward neural networkss. In Proceedings of the 2004 IEEE International Joint Conference on Neural Networkss, Budapest, Hungary, 25–29 July 2004; pp. 985–990

[11] Huang, S.; Li, C. Distributed extreme learning machine for nonlinear learning over networks. Entropy 2015, 17, 818–840.

[12] Liang, N.-Y.; Huang, G.-B.; Saratchandran, P.; Sundararajan, N. A fast and accurate online sequential learning algorithm for feedforward networkss. IEEE Trans. Neural Netw. 2006, 17, 1411–1423.

[13] Olatunji G.B., Omotesho O. A., Ayinde O. E. and Ayinde K., Determinants of inflation in Nigeria: A co-integration approach, Contributed paper presented at the Joint 3rd African Association of Agriculture Eonomists (AAAE) and 48th Agricultural Economists Association of South Africa (AEASA) Conference, Cape Town, South Africa, September 19 – 23, 2010.

[14] Phillips, P. C. B., Perron, P., Testing for a unit root in a time series regression. Biometrika, (1988) Vol. 75, No. 2, pp. 335–346.

[15] Rampal Singh, and S. Balasundaram.., Application of Extreme Learning Machine Method for Time Series Analysis. International Journal of Electrical and Computer Engineering (2007) 2:8

[16] Schwarz, G. E., "Estimating the dimension of a model." Annals of Statistics (1978), 6(2): 461– 464.

[17] Singh,R.;Balasundaram,S.Applicationofextremelearningmachinemethodfortimeseriesanalysis. Int. J. Intell. Technol. 2007, 2, 256–262.