Volume XIV Number 4, 2022

Comparative Analysis of ARIMA and Artificial Neural Network Techniques for Forecasting Non-Stationary Agricultural Output Time Series

Olushina Olawale Awe, Ronaldo Dias

Institute of Mathematics, Statistics and Scientific Computing, University of Campinas, Brazil

Abstract

With the vast popularity of the deep learning models in the engineering and mathematical fields, Artificial Neural Networks (ANN) have recently attracted significant research applications in agriculture, economics, informatics and finance. In this paper, we use a deep learning method to capture and predict the unknown complex nonlinear characteristics of agricultural output based on autoregressive artificial neural network, using Nigeria as a case study. Using the proposed model, shocks in agricultural output is analyzed and modeled using data obtained for a period of forty years (1980-2019), and compared with analyses obtained from the autoregressive integrated moving average model (ARIMA). This result is significant because it justifies the superiority of the hybrid ANN model over the traditional Box-Jenkins methodology for forecasting non-stationary time series. The empirical results show that the proposed autoregressive ANN model achieves an improved forecasting accuracy over the traditional Box-Jenkins ARIMA method. It is further proposed that various types of artificial neural networks would be useful in forecasting and solving relevant tasks and problems widely defined in global agricultural production.

Keywords

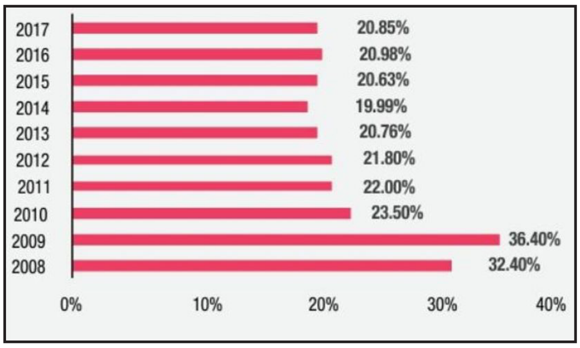
Artificial neural network, ARIMA, agricultural output, Nigeria, forecasting.

Awe, O. O. and Dias, R. (2022) "Comparative Analysis of ARIMA and Artificial Neural Network Techniques for Forecasting Non-Stationary Agricultural Output Time Series", *AGRIS on-line Papers in Economics and Informatics*, Vol. 14, No. 4, pp. 3-9. ISSN 1804-1930. DOI 10.7160/aol.2022.140401.

Introduction

Agricultural output has been said to be affected by various factors including land availability and fertility, temperature, rainfall, population, active labour force participation, exchange rate, inflation rate, oil prices and much more. Most importantly, in Nigeria, an oil-dependent country, agricultural output tends to be majorly affected by the oil market (Awe et al., 2018). This study deals with finding a suitable model for forecasting Agricultural Output, using Nigeria as a case study. Nigeria is an enormous agricultural nation which is endowed with large human and natural resources which comprising of 68 million hectares of arable land, fresh water resources covering about 12.6 million hectares, 960 km of coastline and an ecological diversity which enables the country to produce a wide variety of crops, livestock, forestry and fishery products (Ewetan et al., 2017). Agricultural output helps in revealing the productivity of the agricultural sector of the economy.

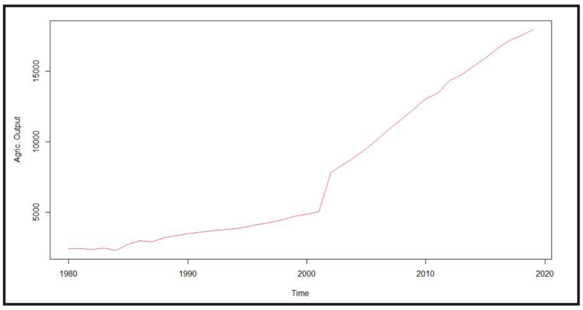
From historical perspectives, agricultural output in Nigeria has been volatile and erratic with average productivity on a downward slope (Akinkumi, 2017). With available statistics from the Central Bank of Nigeria (CBN), the agricultural sector's share of GDP rose from 28% in 1985 to 32% in 1988, went down to 31% in 1989, went up to 37% in 1990 but decreased significantly to 24% in 1992, it then rose again to 37% in 1994. It was 32% in 1996 and increased to 40% in 1998, decreased to 27% in 2000, went up to 37% and went downward to 31% in 2002 and 2006 respectively. The percentage contribution of the agricultural sector to GDP decreased persistently from 0.37 in 2009 to 0.22 in 2012 and to 0.20 in 2014 (Falola et al., 2008). It has continued to increase considerably since 2008 (see Figure 1). However, Nigeria's real GDP growth rate has been lower than those of many other African countries in recent times, including smaller countries in the Northern and Eastern parts of Africa (Awe and Gil-Alana, 2019). Hence, a study on agricultural output is necessary because agriculture is the backbone of the Nigerian economy, after crude oil exploration (Matthew and Mordecai, 2016).



Source: Nigerian Bureau of Statistics (NBS)

Figure 1: Contribution of Agricultural Output to Nigeria's GDP from 2008-2017.

The poor performance of the agricultural sector in Nigeria has been hinged on the oil glut and its consequences on several occasions, resulting in low agricultural output and productivity due to negligence of the agricultural sector. (CBN, 2014; Christiaensen et al, 2007). Agriculture is broadly divided into four sectors in Nigeria – crop production, fishing, livestock and forestry. Crop production remains the largest segment and it accounts for about 87.6% of the sector's total output (Anderu and Omotayo, 2020). This is followed by livestock, fishing and forestry at 8.1%, 3.2% and 1.1% respectively (Igbokwe, 2005). A chart depicting the trajectory of agricultural output from 1980-2019 in Nigeria is shown in Figure 2.



Source: Nigerian Bureau of Statistics (NBS)

Figure 2: Trajectory of Agricultural Output in Nigeria from 1980-2019.

Between 1960 and the early 1970s, agriculture was the primary source of revenue for the Nigerian economy, with revenue from other sources considered a bonus. The growth rate of agricultural output in Nigeria increased from an average of about 3% in the 1990s to about 7 % in mid 2000s and has continued to rise majorly due to the fall in oil price which made the succeding governments to concentrate on improving agricultural output (Igbokwe 2005).

The government of Nigeria has enacted various policies and embarked on many developmental programmes with the goal of increasing agricultural output in order to overhaul the agricultural sector and diversify the Nigerian economy (Awe et al., 2018). With the much that have been expended on these policies and programmes, there is an urgent need to steadily measure and forecast the corresponding progress of agricultural output in Nigeria using modern deep learning tools (Rakhmatuiln et al., 2021). Several time series methodologies have been applied to analyze agricultural data in recent times but have been seen to be incapable of capturing non-linear dynamics like the modern methods (see for instance, Mensi et al., 2017; Hloušková et al., 2018; Awe et al., 2018; Kharin, 2018 and Ayinde et al., 2015). The results and predictions obtained from the use of these modern tools would be useful for policy makers act on.

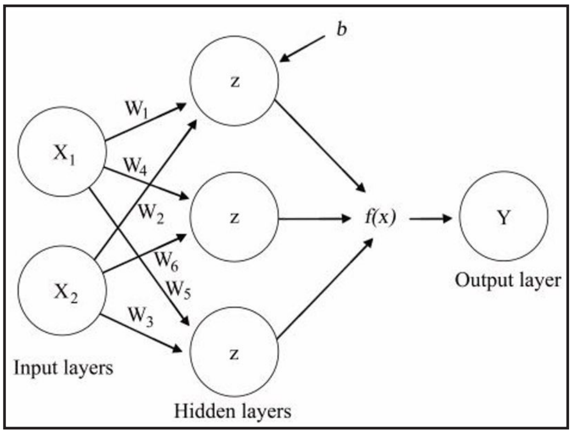
Therefore, this paper examines the forecasting of agricultural output in Nigeria using a modern approach. More so, the use of artificial neural network technique for analyzing and forecasting agricultural output is scanty in literature. The remaining sections of this paper is organized as follows: following this non-exhaustive introductory aspect, the methodology of artificial neural network is discussed in section two. The third section is on empirical analysis and results, while the fourth section dwells on the discussion of results, and finally, the last section is on conclusion and recommendations.

Materials and methods

Artificial neural networks

Artificial neural networks (ANNs) have gained tremendous popularity and use as a promising alternative technique for forecasting climate and agricultural time series because of their several distinguishing features (Abhishek et al., 2012). Similar to the biological structure of neurons, artificial neural networks define the neuron as a central processing unit, which performs a mathematical operation that generates an output from a set of inputs (Rodrigues et al., 2020). Artificial neural networks are one of the most important elements of machine learning and artificial intelligence. They are inspired by the human brain structure and function as if they are based on interconnected nodes in which simple processing operations take place. The spectrum of neural networks application is very wide, and it also includes agriculture (Kujawa and Niedbała, 2021).

Artificial neural networks have been increasingly used by food producers at every stage of agricultural production and in efficient agricultural management (Li and Chao, 2020). Examples of their applications include: forecasting of food production in agriculture on the basis of a wide range of independent variables, classification of diseases and pests, intelligent weed control, and classification of the quality of harvested crops (Niedbała et al., 2020). Artificial intelligence methods support decision-making systems in agriculture, help optimize storage and transport processes, and make it possible to predict the costs incurred depending on the chosen direction of management. An ANN is based on a collection of connected units or nodes called artificial neurons, which loosely model the neurons in a biological brain. An artificial neuron receives a signal then processes it to yield an output which is computed by some non-linear function which is the sum of its inputs. The output of a neuron is a function of the weighted sum of the inputs plus the bias (Rodrigues et al., 2020). The scheme of the artificial neural network used in this study is shown in Figure 3. A hybrid neural network model for agricultural output time series was adopted in this study with the aid of the R package *forecast*, through the *nnetar* function, which generates a feed-forward neural network with a single hidden layer and lagged inputs for forecasting univariate output time series (agricultural output).



Source: Rodrigues et al. (2020)

Figure 3: A feed-forward artificial neural network scheme.

Results and Discussions

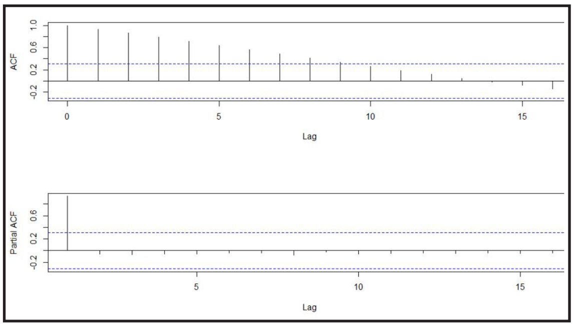
Data

The data used in this study are annual Agricultural Output of Nigeria from 1980-2019 obtained from the database of the World Development Index of the World Bank website (https://data.worldbank.org). The line plot of this data is shown in Figure 1.

Stationarity test

It is vital to test if the time series of agricultural output is stationary. To test for stationary of the data, we use the Augmented Dickey-Fuller (ADF) test. ADF test is a statistical test for finding out if a time series contains a unit root. The null hypothesis for the ADF test is that a time series has a unit root (that is, time series is not stationary). The choice of value of d depends on the number of time a non-stationary time series must be differenced to attain stationarity.

Figure 3 presents the correlograms of autocorrelation functions at some lags for agricultural output of Nigeria over the years. It can be seen from the figure that the data series are not stationary.



Source: Author

Figure 4: Corelogram of Nigerian agricultural output.

A formal test based on ADF test is employed to investigate if the data is stationary. The null hypothesis for this test states that the data series are not stationary while the alternative hypothesis states that the data series are stationary. At 5% level of significance, it can be seen that the data is not stationary because the ADF test fails to reject the null hypothesis (p-value = 0.6944). In addition, each correlogram shows that each data series does not exhibit long – range dependence. As a result, fractional integration of the data series to achieve stationarity may not produce an optimal model as the autocorrelation function of Nigerian agricultural output is seen to experience a fast decay. The outcome of this is that recent methods such as autoregressive fractionally integrated moving average (ARFIMA) and its seasonal version (SARFIMA) may produce sporadic results when applied to analyze the data. This is because the use of ARFIMA and SARFIMA requires a long – range dependence series of data (Awe et al., 2021; Awe and Gil-Alana, 2019), hence it is reasonable to explore the performance of a modern method like ANN.

ARIMA model

The ARIMA methodology of Box & Jenkins was adopted in this study with a view to identifying the optimal model for Nigerian agricultural output and comparing its forecasting performance with that of the proposed autoregressive artificial neural network model described above.

The ARIMA (p, d, q) model on a time series Y_t is defined as (1)

$$\Delta^d Y_t = \emptyset_1 \Delta^d Y_{t-1} + \emptyset_2 \Delta^d Y_{t-2} + \cdots + \\ + \emptyset_p \Delta^d Y_{t-p} + \epsilon_t + \theta_1 \epsilon_{t-1} + \\ + \theta_2 \epsilon_{t-2} + \cdots + \theta_q \epsilon_{t-q} \qquad (1)$$

where p, d and q are orders of autoregressive, integrated and moving average parts respectively, ϵ_t is the residual of the estimated Y_t, which is assumed uncorrelated. Δ is the backward shift operator, $\emptyset_1$, $\emptyset_2$,…, $\emptyset q$ are the parameters of the autoregressive part of the model, $\theta_1,\theta_2,\ldots,\theta_q$ are parameters of the moving average part of the ARIMA model (Awe et al., 2020). The choice of optimal values of p and q are based on the ARIMA (*p, d, q*) model with the least Akaike information criterion and root mean square of error. The parameters of the ARIMA model are estimated by minimizing sum of the square of ϵ_t using maximum likelihood estimation (see Table 1).

ARIMA Model	AIC
ARIMA(2,2,2)	584.5697
ARIMA (0,2,0)	594.9623
ARIMA (1,2,0)	586.755
ARIMA (0,2,1)	577.3763
ARIMA (1,2,1)	579.6402
ARIMA (0,2,2)	579.6185
ARIMA (1,2,2)	582.0142

Source: Author

Table 1: ARIMA model selection

From the Table 1, it shows that ARIMA (0,2,1) performs best because it has the lowest Akaike information criteria of 577.3763.

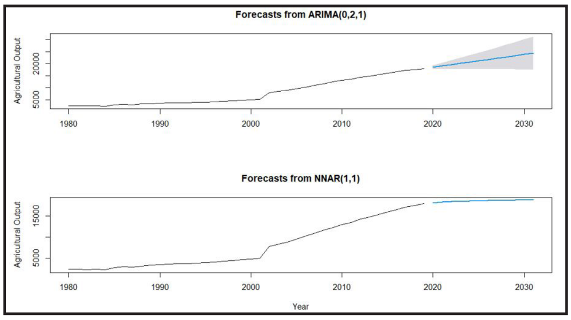
Point forecasts

Shocks in agricultural output is analyzed and modeled with ARIMA using data obtained for a period of forty years (1980-2019), and compared with analyses obtained from the hybrid autoregressive artificial neural network model (AANN). The empirical results show that the proposed deep learning model achieves an improved forecasting accuracy over the traditional Box-Jenkins ARIMA method. Point forecasts from ARIMA and AANN algorithms are shown in Table 2. Notice that predictions from the artificial neural network model are lower than those from the ARIMA model. The unit of measurement of agricultural output is in Billion Naira. Figure 4 depicts a diagrammatic representation of the forecasts from the two techniques. They both depict a steady upward slope.

Year	ARIMA	AANN
2020	18490.55	18140.15
2021	19022.52	18283.31
2022	19554.49	18395.30
2023	20086.46	18482.35
2024	20618.43	18549.69
2025	21150.40	18601.59
2026	21682.37	18641.45
2027	22214.34	18672.01
2028	22746.31	18695.40
2029	23278.28	18713.27
2030	23810.25	18726.91

Source: Author

Table 2: Eleven year point forecasts from ARIMA (0, 2, 1) and AANN (1, 1) models



Source: Author

Figure 5: Point forecasts from ARIMA (0,2,1) and AANN(1,1).

Model evaluation metrics

The Root Mean Squared Error (RMSE) and Mean Percentage Error (MPE) were used in evaluating model performance in this study. The **root mean square error (RMSE)** (also often referred to as the root mean square deviation, RMSD) is a frequently used measure of the difference between values predicted by a model and the values actually being observed. These individual differences are also called residuals, and the RMSE serves to aggregate them into a single measure of predictive power. The RMSE of a model prediction with respect to the estimated variable X_{model} is defined as the square root of the mean squared error (2):

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(X_{obs,i} - X_{model,i})^2}{n}} \quad (2)$$

where X_{obs} is observed values and X_{model} is modelled values at time/place i with sample size n.

It is one of the most commonly used measures in literature for evaluating the quality of predictions. The **mean percentage error (MPE)**, which is the computed average of percentage errors by which forecasts of a model differ from actual values of the quantity being forecasted is also used to corroborate model evaluation in this study. To properly validate the predictive accuracy of the models from the training data, a k-fold cross-validation was adopted in this study. In k-fold cross-validation, model validation is done by dividing the data into k subsets of approximately equal sizes, the model is thereby trained k times (Bergmeir and Benitez, 2012). In each of the k times of model training, one of the subsets is randomly set aside which in turn is used to evaluate the model's performance (Bergmeir et al., 2018). In spite of being computationally intensive and time consuming, this method ensures a more accurate prediction. In this study, a 10-fold cross-validation (k=10), which has been used in several related studies, was adopted. Results of model evaluation, showing the average RMSE and MPE are in Table 3.

Model	RMSE	MPE
ARIMA	437.97	1.21
AANN	416.70	-0.82

Source: Author

Table 3: Model comparison via RMSE and MPE.

Table 3 shows that the autoregressive artificial neural network (AANN) proposed performs better than the traditional ARIMA model because it exhibits the lowest RMSE and MPE respectively.

This work is relevant because it does not only use ANN and ARIMA for prediction of agricultural output but also confirms and justifies the suitability of ANN for prediction of non-stationary time series over the traditional Box-Jenkins methodology, especially for the Nigerian data. The autoregressive neural network proposed generates a feed-forward neural network with a single hidden layer and lagged inputs. Further more, this result is significant because it also addresses the concerns of Faraway and Chatfield (1998) who querried the forecasting ability of the neural network model for predicting non-stationary seasonal monthly time series. This study has shown that ANN are not as poor as portrayed by the authors for modeling such data. More advanced variants of this model would be explored for scientific computing and forecasting of agricultural output in our future studies.

Conclusion

In this paper, we have deduced some salient results on the forecasting of agricultural output using autoregressive integrated moving average and autoregressive artificial neural network models. Comparison between these two methods showed that the autoregressive artificial neural network performs better for forecasting agricultural output data. Hence, deep learning methods are recommended for use in Agricultural forecasting because of their rich computational suitablility especially for non-linear time series data. Other deep learning/artificial neural network methods that would be explored in our future studies include the Long Short-Term Memory (LSTM) which is a type of recursive neural network that learns a mapping from input to output over time as reviewed recently in Rakhmatuiln et al. (2021).

Acknowledgements

The author hereby thank FAPESP, Sao Paulo, Brazil for financial support during this research.

Corresponding authors
Olushina Olawale Awe, Ph.D.
Department of Statistics (IMECC), University of Campinas
Rua da Reitoria, 121 Barão Geraldo, Campinas– SP, Sao Paulo, 13083-872, Brazil
E-mail: oawe@unicamp.br

References

[1] Abhishek, K., Singh, M. P., Ghosh, S. and Anand, A. (2012) "Weather forecasting model using artificial neural network", *Procedia Technology*, Vol. 4, pp. 311-318. E-ISSN 2212-0173. DOI 10.1016/j.protcy.2012.05.047.

[2] Akinkunmi, M. A. (2017) "Nigeria's economic growth: Past, present and determinants", *Journal of Economics and Development Studies*, Vol. 5, No. 2, pp. 31-46. E-ISSN 2334-2390, ISSN 2334-2382. DOI 10.22437/ppd.v8i2.9106.

[3] Anderu, K. S. and Omotayo, E. O. (2020) "Agricultural output and government expenditure in Nigeria", *Jurnal Perspektif Pembiayaan dan Pembangunan Daerah*, Vol. 8, No. 2, pp. 101-110. E-ISSN 2355-8520.

[4] Awe, O. O. and Gil-Alana, L. A. (2019) "Time series analysis of economic growth rate series in Nigeria: structural breaks, non-linearities and reasons behind the recent recession", *Applied Economics*, Vol. 51, No. 50, pp. 5482-5489.ISSN 00036846. DOI 10.1080/00036846.2019.1613513.

[5] Awe, O., Okeyinka, A. and Fatokun, J. O. (2020) "An Alternative Algorithm for ARIMA Model Selection", In *2020 International Conference in Mathematics, Computer Engineering and Computer Science (ICMCECS),* IEEE, pp. 1-4. DOI 10.1109/ICMCECS47690.2020.246979.

[6] Awe, O. O., Akinlana, D. M., Yaya, O. S. and Aromolaran, O. (2018) "Time series analysis of the behaviour of import and export of agricultural and non-agricultural goods in West Africa: A case study of Nigeria", *AGRIS on-line Papers in Economics and Informatics*, Vol. 10, No. 2, pp. 15-22. ISSN 1804-1930. DOI 10.7160/aol.2018.100202.

[7] Awe, O. O., Ayeni, O. C., Sanusi, G. P. and Oderinde, L. O. (2021) "A comparative time series analysis of crude mortality rate in the BRICS countries", *BRICS Journal of Economics*, Vol. 2, No. 2, pp. 17-32. ISSN 2712-7508. DOI 10.38050/2712-7508-2021-2-2.

[8] Ayinde, O. E., Ilori, T. E., Ayinde, K. and Babatunde, R. O., (2015) "Analysis of the behaviour of prices of major staple foods in West Africa: A case study of Nigeria", *AGRIS on-line Papers in Economics and Informatics*, Vol. 7, No. 4, pp. 3-17. ISSN 1804-1930.

[9] Bergmeir, C. and Benítez, J. M. (2012) "On the use of cross-validation for time series predictor evaluation", *Information Sciences*, Vol. 191, pp. 192-213. ISSN 00200255. DOI 10.1016/j.ins.2011.12.028.

[10] Bergmeir, C., Hyndman, R. J. and Koo, B. (2018) "A note on the validity of cross-validation for evaluating autoregressive time series prediction", *Computational Statistics & Data Analysis*, Vol. 120, pp. 70-83. ISSN 0167-9473. DOI 10.1016/j.csda.2017.11.003.

[11] Central Bank of Nigeria (2014) "*Statistical Bulletin*". [Online]. Available: http://statistics.cbn.gov.ng/cbnonlinestats [Accessed: Sept. 21, 2022].

[12] Christiaensen, L, Demery, L. and Kuhl, J. (2007) "The role of agriculture in poverty reduction: An empirical perspective", World Bank Policy Research Working Paper, No. 4013.

[13] Ewetan, O., Fakile, A., Urhie, E. S. and Odunatan, E. (2017) "Agricultural output and economic growth in Nigeria", *Journal of African Research in Business & Technology* ". E-ISSN 2165-9443.

[14] Faraway, J. and Chatfield, C. (1998) "Time series forecasting with neural networks: a comparative study using the airline data", *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, Vol. 47, No. 2, pp. 231-250. ISSN 00359254. DOI 10.1111/1467-9876.00109.

[15] Falola, T. and Heaton, M. (2008) "A history of Nigeria", New York: Cambridge University Press; 2008, 370 p. ISBN 052168157X. DOI 10.1017/CBO9780511819711.

[16] Hloušková, Z., Ženíšková, P. and Prášilová, M. (2018) "Comparison of Agricultural Costs Prediction Approaches", *AGRIS on-line Papers in Economics and Informatics*, Vol. 10, No. 1, pp. 3-13. ISSN 1804-1930. DOI 10.7160/aol.2018.100101.

[17] Igbokwe, E. M. (2005) "Concepts in rural and agricultural sociology", *Agricultural Extension in Nigeria*, pp. 91-100. ISSN 1119944X.

[18] Kharin, S. (2018) "Price Transmission Analysis: The case of milk products in Russia", *AGRIS on-line Papers in Economics and Informatics*, Vol. 10, No. 1, pp. 15-23. ISSN 1804-1930. DOI 10.7160/aol.2018.100102.

[19] Kujawa, S. and Niedbała, G. (ed.) (2021) "Artificial Neural Networks in agriculture", *Agriculture*, Special issue, Vol. 11, No. 6, p. 497. ISSN 2077-0472. DOI 10.3390/agriculture11060497.

[20] Li, Y and Chao, X. (2020) "ANN-Based Continual Classification in Agriculture", *Agriculture*, Vol. 10, No. 5, p.178. ISSN 2077-0472. DOI 10.3390/agriculture10050178.

[21] Matthew, A. and Mordecai, B. D. (2016) "The impact of public agricultural expenditure on agricultural output in Nigeria (1981-2014)", *Asian Journal of Agricultural Extension, Economics & Sociology*, Vol. 11, No. 2, pp. 1-10. E-ISSN 2320-7027. DOI 10.9734/AJAEES/2016/25491.

[22] Mensi, W., Tiwari, A., Bouri, E., Roubaud, D. and AI-Yahyaee, K. H. (2017) "The dependence structure across oil, wheat, and corn: a wavelet-based copula approach using implied volatility indexes", *Energy Economics*, Vol. 66, pp. 122-139. ISSN 0140-9883. DOI 10.1016/j.eneco.2017.06.007.

[23] Niedbała, G., Kurasiak-Popowska, D., Stuper-Szablewska, K. and Nawracała, J. (2020) "Application of Artificial Neural Networks to Analyze the Concentration of Ferulic Acid, Deoxynivalenol, and Nivalenol in Winter Wheat Grain", *Agriculture*, Vol. 10, p. 127. ISSN 2077-0472. DOI 10.3390/agriculture10040127.

[24] Rakhmatuiln, I., Kamilaris, A. and Andreasen, C. (2021) "Deep Neural Networks to Detect Weeds from Crops in Agricultural Environments in Real-Time: A Review", *Remote Sensing*, Vol. 13, No. 21, p. 4486. ISSN 2072-4292. DOI 10.3390/rs13214486.

[25] Rodrigues, P. C., Awe, O. O., Pimentel, J. S. and Mahmoudvand, R. (2020) "Modelling the Behaviour of Currency Exchange Rates with Singular Spectrum Analysis and Artificial Neural Networks", *Stats*, Vol. 3, No. 2, pp. 137-157. ISSN 2571-905X. DOI 10.3390/stats3020012.