GAU Research Journal **47** (1) : 01-11 (January 2022)

Hybrid SARIMA-Facebook Prophet model for prediction and forecasting of the staple food prices in Tanzania

Godfrey Joseph Saqware [1,3*] and Ismail, B.[2]

[1] Department of Statistics, Mangalore University, Karnataka, India.
[2] Department of Statistics, Yenepoya (Deemed to be University), Karnataka, India.
[3] University of Dar es Salaam, Tanzania
Email: [1,3] godjose70@yahoo.com; [2] prof.ismailb@gmail.com

Received: 04th February, 2021 Accepted: 24th February, 2021

ABSTRACT

The country's stable development requires sufficient food production and reserves for its people. However, the global climatic change and other factors, including political instability, resulted in the shortage and volatility of staple food prices. Thus, the best prediction and forecasting models of staple food prices become essential. The advancement of technology gave room for the development of hybrid models. Therefore, in this paper, we develop the SARIMA-Facebook (FB) Prophet model to predict and forecast staple food prices in Tanzania. Based on the Root Mean Squared Error (RMSE) metrics, the proposed hybrid model out performs the SARIMA and other hybrid models in predicting staple food prices. Moreover, the forecasts for the next two years have shown increasing trends in staple food prices. Since most neighbouring countries depend on Tanzania to complement the food crisis in their countries, initiatives should be taken into account to stabilize prices. Finally, the proposed hybrid SARIMA-FB Prophet may be applied to predict and forecast other prices to see the stability of its accuracy. The construction of the theoretical hybrid SARIMA-FB Prophet model may be done in future studies.

Key words : ARIMA,SARIMA-FB Prophet, Staple food, Time series, RMSE.

INTRODUCTION

Global growth in food production has increased from 70 per cent to 110per cent to meet the demand of the rapid population by 2050 is the fundamental debate. The agricultural sector employment in Africa accounts for 65%, and 61% of the rural households depend on agriculture for livelihood in Sub-Saharan Africa. In Tanzania, the future production of staple food crops remains uncertain due to over-dependence on rain-fed, poor agricultural practices, and climate change and variability. Maize experiences higher seasonality compared to rice among staple food crops. The seasonality significantly varies between one market to another. Policymakers should work harder to stabilize the seasonality of food prices and demand to achieve the Sustainable Development Goal of zero hunger by 2030 (Gilbert *et al.*, 2017). The regional markets, rather than global markets, are the source of external pressures on domestic prices. Furthermore, interactions with both external market shocks and domestic weather shocks play a role in how trade policies influence markets. Price transmission and spatial price fluctuations for staple foods vary from one place to another as a factor of distance from the markets. The improved roads and market infrastructures may reduce the price fluctuation between producers and wholesale markets across the country. Moreover, direct price incentives resulting from border protection, government involvement in domestic markets, and border price shocks encourage farmers' supply and price fluctuations– .Climate change potentially makes the communities vulnerable due to decreased agricultural productivity, food insecurity, and water availability" .Supporting market linkage and infrastructure and enforcing transparent and non-restrictive food marketing rules would improve market access' .Tanzania experienced various times of food scarcity owing to limited food production. Even though the country has an undeniable food production capacity, the government and development partners must play a more significant role in achieving food self-sufficiency. Based on

[1]Scientist, Department of Statistics, Manglore University, Karnataka
[1]Scientist, Department of Statistics, Yenepoya (Deemed to be University), Karnataka
[2,3]Professor, University of Dar Es Salaam,pandeyhs13@gmail.com

policy, climate, and price changes, the accurate future prediction and forecasting of the staple food prices in Tanzania is becoming important. The Auto regressive Integrated Moving Average (ARIMA) model gains much application in modelling and forecasting the time series data . The ARIMA model has got immerse applications in the fields of business, economics, meteorology, agriculture, health to mention a few. The model still performs well in predicting and forecasting agricultural production . The ARIMA and SARIMA models contribute to the market forecasting prices. In the education sector, enrolment of students has also been predicted using ARIMA and SARIMA models. In the health sector, the ARIMA models have an application in predicting new cases of diseases. Using ARIMA and SARIMA models in energy sector prediction is essential in forecasting energy load and demand. Despite the numerous applications of the Box and Jenkins methodology in recent years, researchers claimed that using single models in time series forecasting is insufficient. The hybrid time series models have caught many researchers' attention and have more excellent performance.

Therefore, this paper is divided into five sections. Section two discusses related literature on hybrid models and proposes a suitable hybrid model for predicting and forecasting staple food prices; Section three describes the material and hybrid methodologies for the study. Section four will discuss the results obtained. Finally, section five will provide a conclusion and recommendations for further studies.

RELATED LITERATURE

Technology and advanced computer knowledge have brought changes in modelling and algorithm development. Several pieces of literature talk about the hybrid time series models. The hybrid ARIMA–ANN model out performs single models in the time series multi-step ahead forecasts with the highest accuracy. The comparison model on the generated values of the observed data showed that the wavelet-SARIMA-ANN model out performed the wavelet-ANN and wavelet-SARIMA models in terms of forecasting accuracy. The comparative performance analysis of the proposed methodology with ARIMA, ETS, the Multilayer Perceptron (MLP), and some existing hybrid ARIMA–ANN models shows that the proposed hybrid model statistically gave optimistic predictions results. The hybrid methodology that integrates empirical mode decomposition (EMD) and the ARIMA model applied to forecast the monthly rice prices in Malaysia has increased the accuracy. Experimental results with real data sets indicate that the combined model can effectively improve forecasting accuracy instead of using models separately. A hybrid model incorporating ensemble empirical mode decomposition (EEMD), ARIMA, and Taylor expansion with a tracking differentiator was proposed. The EEMD decomposes the financial time series into subseries. The ARIMA and Taylor expansion models forecast the linear and non-linear portions. The findings on the real financial time series data revealed that the hybrid model outperforms the benchmark models.The combination of ANN and ARIMA in forecasting the future exchange prices for US Dollar (USD) / Indian Rupees (INR) reveals that the hybrid ANN-ARIMA can closely forecast the foreign exchange rates. A hybridization methodology based on integrating the SARIMA and ANN models to predict the number of inspections indicates that the hybrid SARIMA-ANN model out performs single models. The hybrid SARIMA model that integrates Support Vector Regression (SVR) with Gaussian white noise applied to forecast the future statistical indicators in the aviation industry was proposed. The results of the empirical study suggest that the proposed hybrid SARIMA-SVR increases forecasting accuracy. Furthermore, the hybrid ARIMA–ANN model to forecast electricity prices improved prediction accuracy. Based on the literature review, the proposed hybrid model works better than single models.

Therefore, a hybrid model can effectively improve the forecasting accuracy of the traditional methods using appropriate strategies. In this paper, we will combine ARIMA or SARIMA model and the FB Prophet models and apply the best hybrid model in forecasting the staple food prices in Tanzania. Furthermore, the proposed hybrid model will be compared with the single and hybrid machine learning models.

MATERIALS AND METHODS

3.1 Data specification

This paper considers monthly recorded data of the staple food prices collected from the Ministry of Industry and Trade, the government of the United

Republic of Tanzania. The data covers the time between January 2005 and December 2021. The analysis of the two data sets is important due to climate change, war, outbreaks of pests and diseases, and the global Corona virus pandemic which affects global food production and ultimately hunger. The data was divided in 80% and 20% training and testing sets respectively.

PREDICTION MODELS

3.2.1 Autoregressive Integrated Moving Average Model (ARIMA)

The Auto regressive Moving Average model (AIRMA) is the combination of the Auto regressive (AR) and Moving Average (MA) Models. The auto regressive model predicts the current value based on the previous observations $w_{t-1}, w_{t-2}, \ldots, w_{t-p}$ while the Moving Average model predicts the current value using the residuals of the previous values $\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots, \varepsilon_{t-p}$

The AR model of order p is given by: -

$$w_t = \psi_1 w_{t-1} + \psi_2 w_{t-2} + \ldots + \psi_p w_{t-p} + \varepsilon_t \quad (1)$$

The model in Eqn. (1) can be re-written as: -

$$w_t - \psi_1 w_{t-1} - \psi_2 w_{t-2} - \ldots - \psi_p w_{t-p} = \varepsilon_t \quad (2)$$

By applying the backward shift operator $B^j w_t = w_{t-j}\ \ j = \{1,2,3,\ldots,\} p$, the model in Eqn. (2) can be re-written as: -

$$\begin{aligned} w_t - \psi_1 B w_t - \psi_2 B^2 w_t - \ldots - \psi_p B^2 w_t &= \varepsilon_t \\ (1 - \psi_1 B - \psi_2 B^2 - \ldots - \psi_p B^2) w_t &= \varepsilon_t \\ \psi(B)\, w_t &= \varepsilon_t \end{aligned} \quad (3)$$

The MA model of order q is given by: -

$$w_t = \varepsilon_t + \phi_1 \varepsilon_{t-1} + \phi_2 \varepsilon_{t-2} + \ldots + \phi_q \varepsilon_{t-q} \quad (4)$$

By applying the backward shift operator $B^j \varepsilon_t = \varepsilon_{t-j} \{j = 1,2,3,\ldots,\} q$, the model in Eqn. (4) can be re-written as: -

$$\begin{aligned} w_t &= \varepsilon_t + \phi_1 B \varepsilon_t + \phi_2 B^2 \varepsilon_t + \ldots + \phi_q B^q \varepsilon_t \\ w_t &= (1 + \phi_1 B + \phi_2 B^2 + \ldots + \phi_q B^q) \varepsilon_t \\ w_t &= \phi(B)\, \varepsilon_t \end{aligned} \quad (6)$$

By applying the backward shift operator $B^j \varepsilon_t = \varepsilon_{t-j} \{j = 1,2,3,\ldots,\} q$, the model in Eqn. (4) can be re-written as: -

$$\begin{aligned} w_t &= \varepsilon_t + \phi_1 B \varepsilon_t + \phi_2 B^2 \varepsilon_t + \ldots + \phi_q B^q \varepsilon_t \\ w_t &= (1 + \phi_1 B + \phi_2 B^2 + \ldots + \phi_q B^q) \varepsilon_t \\ w_t &= \phi(B)\, \varepsilon_t \end{aligned} \quad (6)$$

The *ARMA (p, q)* model is used for forecasting a stationary time series, and it's given in Eqn. (7) below.

$$\begin{aligned} w_t &= \psi_1 w_{t-1} + \psi_2 w_{t-2} + \ldots + \psi_p w_{t-p} + \varepsilon_t \\ &+ \phi_1 \varepsilon_{t-1} + \phi_2 \varepsilon_{t-2} + \ldots + \phi_q \varepsilon_{t-q} \end{aligned} \quad (7)$$

Now, supposed $w_t \square I(d)$ is the stationary series after d[th] the order of differencing, we say $z_t = (1) - B^{\,d} w_t$ and *ARIM(A,d,q)* can be written as: -

$$\begin{aligned} (1-B)^d w_t = z_t &= \psi_1 z_{t-1} + \psi_2 z_{t-2} + \ldots \\ &+ \psi_p z_{t-p} + \varepsilon_t + \phi_1 \varepsilon_{t-1} + \phi_2 \varepsilon_{t-2} + \ldots + \phi_q \varepsilon_{t-q} \end{aligned} \quad (8)$$

We are introducing the backward shift operator as in Eqn. (3 and 6), Eqn. (8) can be written as: -

$$\begin{aligned} [1 - \psi_1 B - \psi_2 B^2 - \ \ldots\ - \psi_p B^p][1-B]^d w_t &= \\ [1 + \phi_1 B + \phi_2 B^2 + \ \ldots\ + \phi_q B^q] \varepsilon_t & \\ \psi(B)[1-B]^d w_t &= \phi_1(B) \varepsilon_t \end{aligned} \quad (9)$$

The above-postulated model in Eqn. (9) works only when there is no seasonal variation. If the variable contains a seasonal component, then it is called Seasonal ARIMA or SARIMA model.The seasonal ARIMA model incorporates both non-seasonal and seasonal components in a multiplicative model. The *SARIMA (p,d,q)*(P,D,Q)s* model is given by: -

$$\begin{aligned} \psi(B)\Psi(B^S)(1-B)^d(1-B^S)^D w_t &= \\ \phi(B)\Phi(B^S)\varepsilon_t & \end{aligned} \quad (11)$$

Where, Ψ and ψ are the AR parameters for the seasonal and non-seasonal components, respectively, while Φ are ϕ the MA parameters for the seasonal and non-seasonal components respectively, B is the backward operator, D and d are the differencing terms.

Infitting the ARIMA/SARIMA models, it involves three main stages.

a) Model identification - This stage involves the analysis of autocorrelations, partial autocorrelations, and cross-correlations. To assess if differencing is required, stationarity tests is done.

b) Model estimation and diagnostic tests - The model's appropriateness is determined using diagnostic statistics generated by the estimate. The white noise residuals tests determine whether the residual series has additional information that a more complex model might use.

c) Forecasting - Here we forecast the future values of the time series and generate confidence intervals for these forecasts.

3.2.2 Facebook FB Prophet Model

The FB Prophet is a developed time series prediction model that is both simple and practical without a lot of data preprocessing. The FB Prophet model can fit very quickly and automatically fill in the missing values. Furthermore, the FB Prophet model allows for flexible adjustment of periodicity, making it suited for various applications. The FB Prophet time series forecasting model comprises trend, cyclic, festival, and error items. The model's composition is as follows: -

$$y(t)=g(t)+s(t)+h(t)+e(t) \quad (12)$$

Where, $g(t)$ is the trend function, $s(t)$ is the periodic term function, $h(t)$ is the impacts from the holidays, and the $(e)t$ is the error term.

Where: -

i) The $g(t)$ can either take the form of the logistic growth model or the piece-wise regression model as follows: -

a) For the logistics growth model, $g(t)=\dfrac{c}{1+e^{(-k(t-m))}}$

C is the saturation value, k is the growth rate, and m is the bias parameter.

b) For the piece-wise regression,

$$y=\begin{cases}\beta_0+\beta_1 x & x\le c\\ \beta_0-\beta_2 c+(\beta_1+\beta_2)x & x>c\end{cases}$$

where c is the trend change point.

ii) The periodic term function $s(t)$ is based on the Fourier series, which results in periodic effects on model flexibility (Harvey & Shephard, 1993). With a conventional Fourier series, the approximate arbitrary smooth seasonal effects is as follows: -

$$s(t)=\sum_{n=1}^{N}\left[a_n\cos\left(\frac{2\pi nt}{P}\right)+b_n\sin\left(\frac{2\pi nt}{P}\right)\right]$$

trepresent periods and P is the regular periodic length of the time series; it can be daily, monthly, or yearly.

iii) The impacts from the holidays, $h(t)=Z(t)K_i$; $K_i\sim N(0,\sigma)$ $Z(t)=[I(t\subset D_1),\ldots,I(t\subset D_i),\ldots,I(t\subset D_L)]$ i represents holidays; D represents the collection of past and future holidays; K represents the impact of each holiday on the forecast.

3.2.3 Hybrid SARIMA-FB Prophet Model.

Suppose $\{Y_t,\ t=1,2,3,\ldots,N\}$ which represent linear SARIMA model and $\{FBt\ t=,1,2,3,\ldots,N\}$ which represent non-linear FB Prophet model, as presented in Eqn. (8) below.

$$Y_t=SARIMA_t+FB_t \quad (13)$$

Where S_t and FB_t denotes the linear and non-linear components, respectively. First, the linear part is obtained by fitting the best SARIMA model. Then, the non-linear part is derived from the best fitted linear model, and then we extract the residuals to fit the hybrid SARIMA-FB Prophet model. Finally, the residual of the best model is given in Eqn. (14) below.

$$e_t=Y_t-SARIM\hat{A}_t \quad (14)$$

Then, the extracted SARIMA model for residual with n input nodes is given in Eqn. (15) below.

$$e_t=f(e_{t-1},e_{t-2},\ldots,e_{t-n})+\varepsilon_t \quad (15)$$

Finally, we can find $\hat{}_tY$ (The final forecasted result of the hybrid SARIMA-FB Prophet model) given in Eqn. (16) below.

$$\hat{Y}_t=SARIM\hat{A}_t+F\hat{B}_t \quad (16)$$

3.2.4 Performance Metrics

To determine the best model, Mean Absolute Percentage (MAPE) and the Root Mean Squared Error (RMSE) criterion were applied to measure model prediction accuracy. The MAPE and RMSE measures the differences between actual and predicted values– .Then, Eqn. (17) and (18) are the MAPE and RMSE formulas.

$$MAPE=\frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_t-\hat{y}_t}{y_t}\right| \quad (17)$$

$$RMSE=\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_t-\hat{y}_t\right)^2} \quad (18)$$

Where the y_t is the actual value, $\hat{y}_t$ is the forecasted value, and n is the forecasting horizon.

EMPIRICAL RESULTS AND DISCUSSION

4.1 The plot of the Staple food prices from 2005 to 2021

The prices of beans and rice have increased compared to other staple food. The round potatoes, maize, and sorghum prices are low for all the recording duration. This shows that rice and beans are the most preferred food; thus, their demand increases their prices. Despite the fact that Tanzania has been suffering from a food deficit for decades, a robust agricultural policy is required to stabilize agricultural production and prices '(Mkonda & He, 2017).Institutional and policy-induced factors influence farmers' decisions on which market to sell their primary goods. As a result, there is a lack of motivation to participate in the production and commercialization of agricultural commodities. The factors influence farmers' decisions on which basic food market to buy in Tanzania. Farmers' integration into staple food markets is poor; according to the findings, female-headed households face more challenges' (Kangile *et al*.., 2020). Fig. 1 below shows the staple food prices in Tanzania from 2005 to 2021.

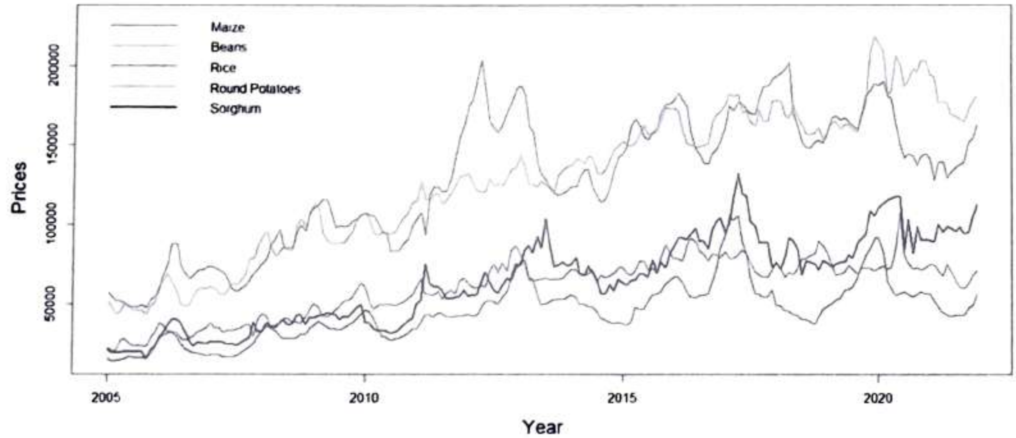

Figure 1: Plot of the Staple Food Prices in Tanzania

4.2 Descriptive Statistics of the Staple food in Tanzania.

Descriptive statistics show that rice and beans have recorded higher prices among the five staples food prices than the rest. Tanzania has a zonal-based production of these crops; rice is produced high in the lake zones and southern part of Tanzania. While beans in the northern and southern regions. Climate change may affect production; thus, if the demand is high, it may lead to shortages and eventually shooting prices. In addition, the global issues in the areas in which oil production is taking place may also increase the staple food prices due to the increase of crude oil per barrel. Table 1 below shows the summary statistics of the staple food prices.

Table 1 : Summary Statistics

Variables	N	mean	sd	median	min	max	skewess	kurtosis	Q0.25	Q0.75
Maize	204	45663.45	18598.3	44213.7	14298	106166.4	0.645635	0.694675	33415.71	56000.05
Potatoes	204	60733.59	18373.92	66710.85	20950	107448.8	-0.46889	-0.77442	46690.09	73232.55
Rice	204	128696.2	42166.44	136008.3	47654	204025.4	-0.27714	-1.07227	95458.35	163326.7
Beans	204	128399.3	45028.58	132272.4	43557	219078.9	-0.25725	-0.9483	94324	164788.3
Sorghum	204	64789.2	27960.09	66341.97	15801	132964.9	0.078813	-1.04384	39032.93	88872.58

4.3 Evaluation of the performance of the testing data set

The proposed hybrid SARIMA-FB Prophet model has performed better than the single ARIMA/ SARIMA model predicting staple food prices. Moreover, the single FB Prophet model has shown better performance than the proposed hybrid model in predicting beans, round potatoes, and sorghum. In predicting the staple food prices, the hybrid SARIMA-FB Prophet model ranks the numbers 3, 1, 1, 2, and 2 for beans, maize, rice, round potatoes, and sorghum, respectively. On the other hand, the proposed hybrid model ranked poorly only in predicting the maize price. The study that combines linear SARIMA and the non-linear generalized autoregressive conditional heteroscedasticity (GARCH) showed the hybrid SARIMA-GARCH performs better than the SARIMA model in forecasting gold price in India. Furthermore, the sophisticated hybrid models that compare SARIMA, Ensemble Learning Machine (ELM) and the hybrid SARIMA-MLP models in forecasting tourist arrivals to Bali concluded that the hybrid SARIMA-ELM model outperforms the hybrid SARIMA-MLP model. However, the complexity of the time series data proves that no unique model works better for each case. Thus, some developed hybrid models might perform the worst in some time series data. Forecasting financial market dynamics has become a focus of economic research. A hybrid Elman's Recurrent Neural Networks(ERNN) model developed to predict stock prices suggest a Minimal Mean Square Error (MMSE) for the h-step forecasts. Thus, the hybrid developed models prove that the integration of more than one model may result in better accuracy. Table 2 below shows the proposed hybrid model summary against its counterparts.

Table 2 : Proposed model summary

Variable	Model	RMSE	MAE	MAPE	Ranking of RMSE
Beans	FB Prophet	16224.10	13824.22	7.44	1
	SARIMA	31418.28	25847.75	13.39	2
	SARIMA-FB Prophet	31525.54	26278.86	13.66	3
Maize	FB Prophet	17565.67	15935.69	29.63	3
	ARIMA	15524.80	10609.04	16.52	2
	ARIMA-FB Prophet	15384.90	10573.17	16.43	1
Rice	FB Prophet	37831.31	31965.75	22.02	3
	SARIMA	17600.33	15530.32	9.89	2
	SARIMA-FB Prophet	17489.06	15650.63	9.94	1
Round Potatoes	FB Prophet	12637.03	11343.99	15.94	1
	SARIMA	31418.28	25847.75	13.39	3
	SARIMA-FB Prophet	26840.80	25138.17	35.63	2
Sorghum	FB Prophet	15634.96	13695.09	15.49	1
	ARIMA	17600.33	15530.32	9.89	3
	ARIMA-FB Prophet	16188.37	12513.40	12.07	2

4.4 The comparison between the developed hybrid and other machine learning models.

SARIMA-FB Prophet still outperforms single and hybrid machine learning models in terms of RMSE. The ELM model is ranked first in the prediction of beans and round potatoes, while the MLP model is ranked first in the prediction of maize and sorghum. The hybrid SARIMA-MLP model comes out on top in rice price prediction. In predicting the prices of beans, maize, rice, and round potatoes, the suggested hybrid SARIMA-FB Prophet model is ranked 5, 2, 2, 3, and 3, respectively. The model continues to beat other predictive models, although single models can still outperform hybrid models depending on the scenario. However, predicting the prices of staple foods is still difficult since each case uses a different model. Table 3 below shows the summary statistics of the proposed hybrid model against the machine learning models.

Table 3 : Proposed model Vs Machine learning models

Variables	Model	RMSE	MAE	MAPE	Ranking of RMSE
Beans	SARIMA-FB Prophet	31525.54	26278.86	13.66	5
	MLP	27543.08	22194.45	11.54	2
	ELM	27072.25	21018.99	10.81	1
	SARIMA-MLP	30725.68	25345.83	13.14	3
	SARIMA-ELM	31028.07	25422.16	13.15	4
Maize	SARIMA-FB Prophet	15384.90	10573.17	16.43	2
	MLP	13550.12	10395.35	17.84	1
	ELM	16459.24	11749.06	17.58	5
	SARIMA-MLP	15480.27	10591.57	16.53	4
	SARIMA-ELM	15028.45	10419.65	16.60	3
Rice	SARIMA-FB Prophet	17489.06	15650.63	9.94	2
	MLP	22529.09	18395.65	11.14	5
	ELM	21423.02	19486.97	12.66	4
	SARIMA-MLP	17441.12	15635.43	9.99	1
	SARIMA-ELM	18357.03	15692.01	9.69	3
Potatoes	SARIMA-FB Prophet	26840.80	25138.17	35.63	3
	MLP	22380.47	20732.71	29.36	2
	ELM	22826.42	21485.08	30.44	1
	SARIMA-MLP	27247.41	25531.13	36.18	5
	SARIMA-ELM	27057.56	25473.02	36.09	4
Sorghum	SARIMA-FB Prophet	16188.37	12513.40	12.07	3
	MLP	12656.46	9600.99	9.44	1
	ELM	16903.55	12946.46	12.38	5
	SARIMA-MLP	16390.57	12420.54	11.87	4
	SARIMA-ELM	14333.61	10821.17	10.52	2

4.5 The best model forecast of the staple food prices in Tanzania

The forecast of the staples food prices is based on the models' overall comparison performance, as indicated in tables 2 and 3. The next 2 years forecasts from the four staple food prices are shown in Fig. 2, 3, 4, 5, and 6. The forecast from the best models from December, 2021 to December, 2024 shows an increase in prices for staple food prices in Tanzania. The forecast alerts the government and other policymakers to take the initiative to stabilize the prices by combating climate changes and reducing the tax on the equipment and other related production facilities. The ongoing war in the oil-producing countries has already impacted the agricultural sector in different ways, such as increased cost of running production machinery and transportation of the harvest from farm to the consumers. Fig. 2,3,4,5, and 6 below shows the next 2 years' forecasts from the best models for beans, maize, rice, round potatoes, and sorghum.

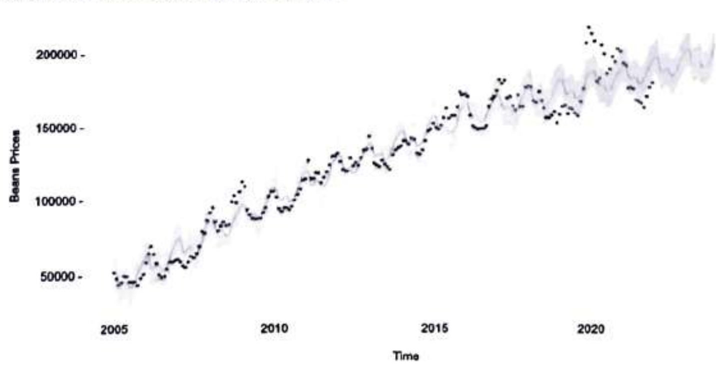

Figure 2 : 2 Years beans prices forecast using FB Prophet model

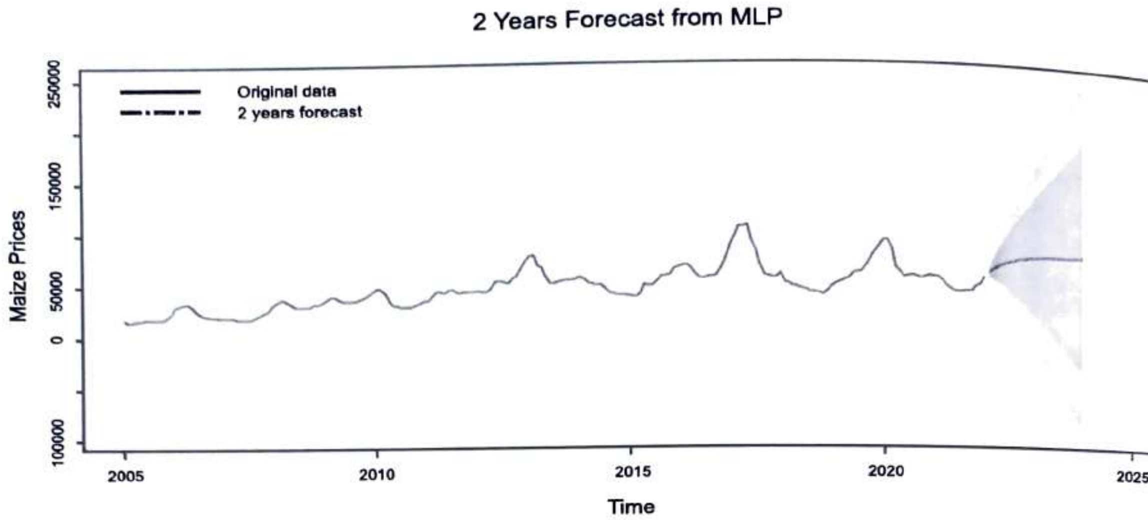

Figure 3 : 2 Year's maize prices forecast using MLP model

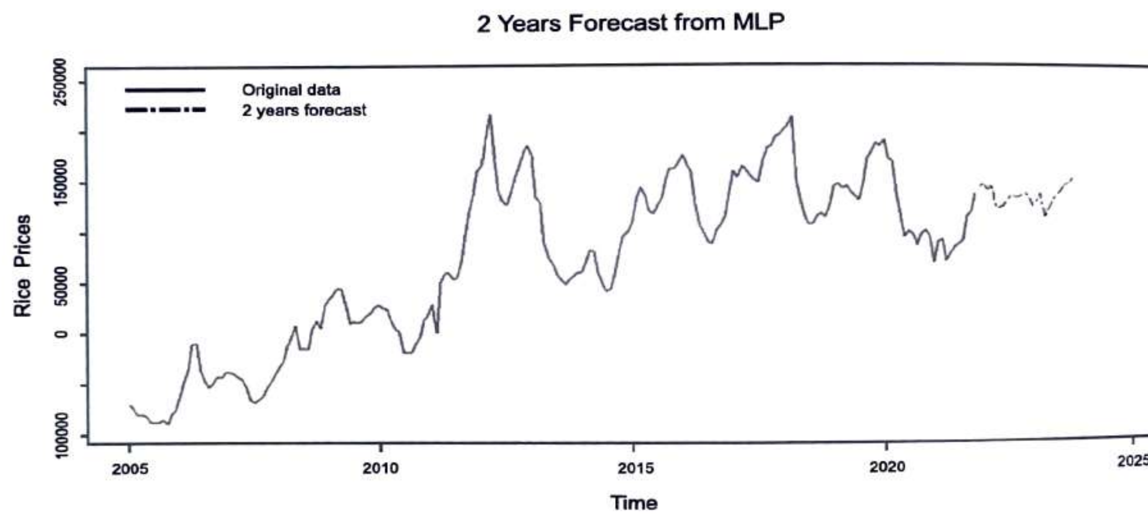

Figure 4 : 2 Year's rice prices forecast using SARIMA-MLP model

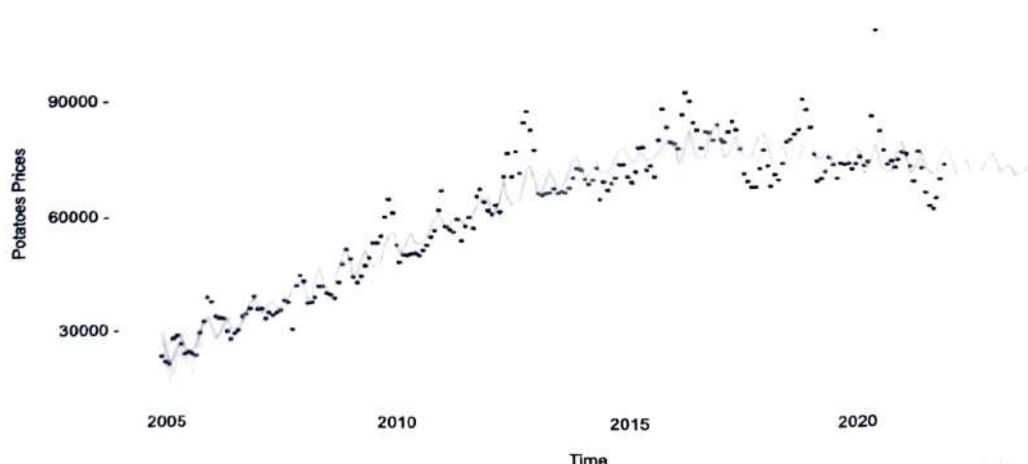

Figure 5 : 2 Years round potatoes prices forecast using FB Prophet model

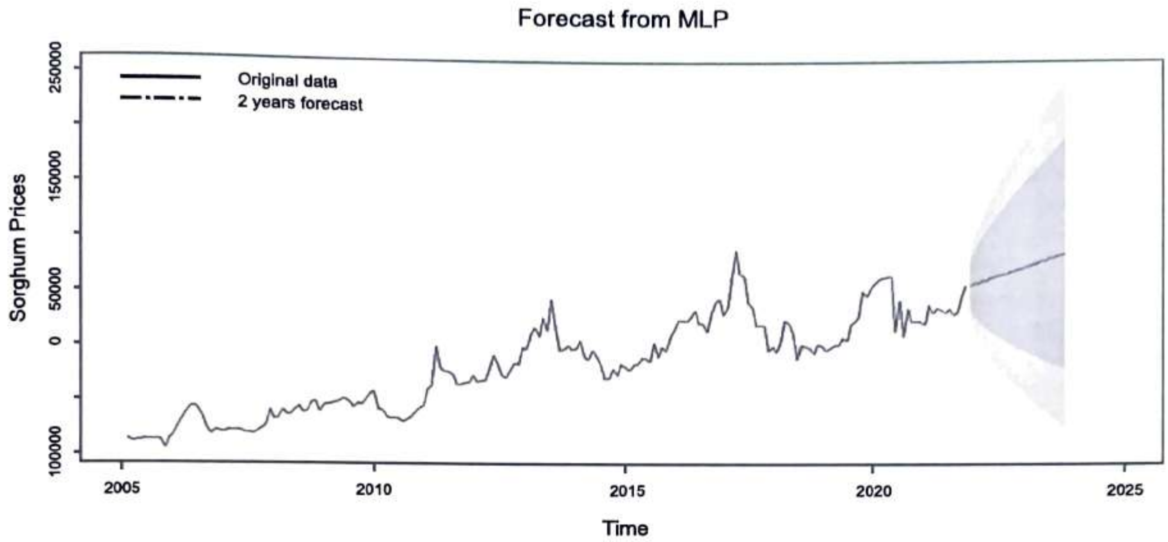

Figure 6: 2 Years sorghum prices forecast using MLP model

CONCLUSION

The hybrid SARIMA-FB Prophet model has shown a more significant performance than the single SARIMA model based on RMSE. The proposed model has shown a convincing performance except in predicting the maize price. Thus, the researchers may use the model to predict the other prices. Compared with a single model, the SARIMA-FB Prophet model can effectively improve a single model's prediction and forecasting accuracy and has more robust applicability. The comparison between the hybrid SARIMA-FB Prophet and machine learning models shows that the model still performs better, although the single ML model performed well in some prediction cases. The next 2 years' forecast of staple food prices using the best selected model shows an increase in prices. Therefore, the private and government planners may formulate more scientific replenishment planning to assure the country's food security and price stability. In this paper, only the hybrid SARIMA-FB Prophet model is developed other hybrid models may be design in the future for further improvement of the prediction and forecasting accuracy.

REFERENCES

Abadan, S., & Shabri, A. (2014). Hybrid empirical mode decomposition-ARIMA for forecasting price of rice. Applied Mathematical Sciences, 8(63), 3133-3143.

Babu, C. N., & Reddy, B. E. (2014). A moving-average filter based hybrid ARIMA-ANN model for forecasting time series data. Applied Soft Computing, 23, 27-38.

Babu, C. N., & Sure, P. (2016). Partitioning and interpolation based hybrid ARIMA-ANN model for time series forecasting. S?dhan?, 41(7), 695-706.

Baffes, J., Kshirsagar, V., & Mitchell, D. (2015). Domestic and external drivers of maize prices in Tanzania. Available at SSRN 2565953.

Borkar, P., & Bodade, V. M. (2017a). Modelling and Forecasting of Tur Production in India Using ARIMA Model. Agriculture Update, 12(2), 252-257.

Borkar, P., & Bodade, V. M. (2017b). Study on Modeling and Forecasting of Tobacco Production in India. International Research Journal of Agricultural Economics and Statistics, 8(2), 281-286.

Boudrioua, M. S. (2019). Modeling and forecasting the Algiers Stock Exchange returns using the Box-Jenkins methodology.

Box, G. (1970). EP and Jenkins, G. M. Time Series Analysis, Holden Day.

Box, G., & Jenkins, G. (1970). Time Series Analysis: Forecasting and ControlV Holden. Day, San Francisco.

Chen, Y. A., Li, R., & Hagedorn, L. S. (2019). Undergraduate international student enrollment forecasting model: An application of time series analysis. Journal of International Students, 9(1), 242-261.

Etuk, E. H., Eleki, A. G., & Sibeate, P. (2016). A Box-Jenkins Model for Monthly Natural Gas

Production in Nigeria. Journal of Multidisciplinary Engineering Science Studies, 2(11), 6.

Hassan, S. A., Abdelaziz, H. H., Ibrahimc, A. H., & Musa, K. S. M. (n.d.). Milk Production-Forecasting In Khartoum State, Sudan.

Hodson, T. O. (2022). Root mean square error (RMSE) or mean absolute error (MAE): When to use them or not. Geoscientific Model Development Discussions, 1-10.

Hussain, J. N., Raheem, S. H., & Enad, M. M. (2021). Use Box-Jenkins strategy for making a time sample to forecasting to Oil prices. Psychology and Education Journal, 58(2), 9943-9951.

Kadigi, I. L., Richardson, J. W., Mutabazi, K. D., Philip, D., Bizimana, J.-C., Mourice, S. K., & Waized, B. (2020). Forecasting yields, prices and net returns for main cereal crops in Tanzania as probability distributions: A multivariate empirical (MVE) approach. Agricultural Systems, 180, 102693.

Kangalawe, R. Y., Mung'ong'o, C. G., Mwakaje, A. G., Kalumanga, E., & Yanda, P. Z. (2017). Climate change and variability impacts on agricultural production and livelihood systems in Western Tanzania. Climate and Development, 9(3), 202-216.

Kangile, R. J., Mgeni, C. P., Mpenda, Z. T., & Sieber, S. (2020). The determinants of farmers' choice of markets for staple food commodities in Dodoma and Morogoro, Tanzania. Agriculture, 10(5), 142.

Karnaboopathy, R., Venkatesan, D., Manikandan, M., Muthukumar, T., Raja, D., & Ramasamy, D. J. (2018). Forecasting Dengue Incidence in Tamil Nadu, India: An Time Series Analysis. Journal of Comprehensive Health, 6(2), 62-66.

Khashei, M., & Bijari, M. (2012). A new class of hybrid models for time series forecasting. Expert Systems with Applications, 39(4), 4344-4357.

Luo, Z., Guo, W., Liu, Q., & Zhang, Z. (2021). A hybrid model for financial time-series forecasting based on mixed methodologies. Expert Systems, 38(2), e12633.

Magrini, E., Balié, J., & Morales-Opazo, C. (2018). Price Signals and Supply Responses for Staple Food Crops in Sub-Saharan Africa. Applied Economic Perspectives and Policy, 40(2), 276-296.

Manikandan, M., Velavan, A., Singh, Z., Purty, A. J., Bazroy, J., & Kannan, S. (2016). Forecasting the trend in cases of Ebola virus disease in west African countries using auto regressive integrated moving average models. Int J Community Med Public Health, 3, 615-618.

Maziku, P. (2019). Variation in Staple Food Prices in Tanzania. African Journal of Applied Research, 5(1), 74-84.

Mbunda, R. (2016). The developmental state and food sovereignty in Tanzania. Agrarian South: Journal of Political Economy, 5(2-3), 265-291.

Mkonda, M. Y., & He, X. (2017). Yields of the major food crops: Implications to food security and policy in Tanzania's semi-arid agro-ecological zone. Sustainability, 9(8), 1490.

Nijbroek, R. P., & Andelman, S. J. (2016). Regional suitability for agricultural intensification: A spatial analysis of the Southern Agricultural Growth Corridor of Tanzania. International Journal of Agricultural Sustainability, 14(2), 231-247.

Panigrahi, S., & Behera, H. S. (2017). A hybrid ETS-ANN model for time series forecasting. Engineering Applications of Artificial Intelligence, 66, 49-59.

Parvez, R., Ali Meerza, S. I., & Hasan Khan Chowdhury, N. (2021). Forecasting student enrollment using time series models and recurrent neural networks.

Pereira, C. M., de Almeida, N. N., & Velloso, M. L. (2015). Fuzzy modeling to forecast an electric load time series. Procedia Computer Science, 55, 395-404.

Priyadarshini, E., Gayathri, G. R., Vidhya, M., Govindarajan, A., & Chakkravarthi, S. (2020). Hybrid forecasting model for non-linear time series. AIP Conference Proceedings, 2277(1), 240005.

Rahman, A. (2017). Modeling of Tea Production in Bangladesh Using Autoregressive Integrated Moving Average (ARIMA) Model. Journal of Applied & Computational Mathematics, 6, 349.

Rajabu, J. K., Mgeni, C. P., Mpenda, Z. T., & Sieber, S. (2020). The Determinants of Farmers' Choice of Markets for Staple Food Commodities in Dodoma and Morogoro, Tanzania. Agriculture, 10(5), 142.

Ruiz-Aguilar, J. J., Turias, I. J., & Jiménez-Come, M. J. (2014). Hybrid approaches based on SARIMA and artificial neural networks for inspection time series forecasting. Transportation Research Part E: Logistics and Transportation Review, 67, 1-13.

Shafaei, M., Adamowski, J., Fakheri-Fard, M., Dinpashoh, Y., & Adamowski, K. (2016). A wavelet-SARIMA-ANN hybrid model for precipitation forecasting. Journal of Water and Land Development.

Shetty, D. K. (2015). Hybrid SARIMA-GARCH Model for Forecasting Indian Gold Price. 03(08), 7.

Shetty, D. K., & Ismail, B. (2021). Forecasting stock prices using hybrid non-stationary time series model with ERNN. Communications in Statistics-Simulation and Computation, 1-13.

Taylor, S. J., & Letham, B. (2018). Forecasting at scale. The American Statistician, 72(1), 37-45.

Todri, A. (2021). Empirical analysis on the USD/all exchange rate volatility in Albanian market: Preliminary results.

Waciko, K. J., & Ismail, B. (2020). SARIMA-ELM hybrid model versus SARIMA-MLP hybrid model. International Journal of Statistics and Applied Mathematics, 5(2), 01-08.

Wadi, S. A. L., Almasarweh, M., Alsaraireh, A. A., & Aqaba, J. (2018). Predicting closed price time series data using ARIMA Model. Modern Applied Science, 12(11).

Xu, S., Chan, H. K., & Zhang, T. (2019). Forecasting the demand of the aviation industry using hybrid time series SARIMA-SVR approach. Transportation Research Part E: Logistics and Transportation Review, 122, 169-180.

Yonar, H., Yonar, A., Tekindal, M. A., & Tekindal, M. (2020). Modeling and Forecasting for the number of cases of the COVID-19 pandemic with the Curve Estimation Models, the Box-Jenkins and Exponential Smoothing Methods. EJMO, 4(2), 160-165.

Zhang, G. P. (2003). Time series forecasting using a hybrid ARIMA and neural network model. Neurocomputing, 50, 159-175.