Original Research Article

Modelling and Forecasting High Frequency EAC Currency Exchange Rates Data Using Hybrid ARIMA-ANN Time Series Model

Abstract

Aims/ objectives: The study sought to compare the modeling and forecasting performance of the autoregressive integrated moving average (ARIMA), artificial neural network (ANN) and the hybrid ARIMA-ANN modeling strategies for high frequency data.
Methodology: The study made use of ARIMA, ANN, and hybrid ARIMA-ANN models to forecasts the the East Africa Community countries' daily currency exchange rates data which were obtained from the Central Bank of Kenya website and covered the period from January 2017 to December 2023. The ADF test was used to test for the stationarity of the data. The Ljung Box test and ACF plots were used to establish and compare the goodness-of-fit of the resultant models while RMSE and MAPE values were used to compare the forecasting performance.
Results: The study established that the hybrid ARIMA-ANN methodolgy provided better-fitting models for the currency exchange rates data compared to ARIMA and ANN modeling strategies since all the Lyung Box test statistics had p values greater than 5%. Comparatively, the hybrid methodology registered lower MAPE and RMSE values hence had better prediction accuracy compared to ARIMA and ANN methods.
Conclusion: The Hybid methodology improves the modeling and forecasting accuracy over the ARIMA and ANN models for high frequency time series data due to its ability to captures both the linear and nonlinear patterns in the time series data.

Keywords: ARIMA; ANN; Hybrid ARIMA-ANN; Forecasting; Modeling; Currency Exchange Rate
2010 Mathematics Subject Classification: 91B84; 62P20; 62M10

1 Introduction

1.1 Modeling and Forecasting high Frequency Currency Exchange Rates

Time series forecasting is an important area in which past observations of the same variable are collected and analyzed to develop a model describing the underlying relationship which is then applied to predict future values of the time series. As observed by Zhang (2003), the autoregressive integrated moving average (ARIMA) model is the most widely used time series model attributable to its statistical properties as well as its ease of implementation within the Box–Jenkins methodology. On their part, Zhang et al. (2001), studied the application of Artificial neural networks (ANNs) in time series forecasting. The methodology offers a major advantage of being flexible in nonlinear modeling. They further asserted that ANNs, do not require the specification of a particular model form, but rather, the model is adaptively formed based on the features presented from the data.

Zhang (2003) asserted the importance of using models in forecasting time series data since they will combine the advantages of the component model hence accomodate both the inear and non-linear components. He recommended the use of the hybrid ARIMA-ANN modeling strategy. Consequently, different studies have compared the performance of ARIMA, ANN, and the Hybrid ARIMA-ANN models (Siamba et al., 2023, Wang et al., 2013, Sanjeev and Nitin, 2022), and most of them have established the superiority of the hybrid models. This is attributable to the fact that it is often not easy to determine whether a time series under study is generated from a linear or non-linear process (Zhang, 2003).

Majority of the studies that have examined the performance of the modeling strategies have primarily used low frequency time series data measured either annually or monthly and little endeavors have been made to assess the performance of the ARIMA, ANN nd the hybrid ARIMA-ANN for high frequency time series frequency. More and more high frequency time series data such as daily and hourly data are being encountered and have resulted to longer and complex time series data, hence the need to develop an understanding on the performance of the time series forecasting methdologies for high frequency data. The study, therefore sought to establish the performance of ARIMA, ANN and the Hybrid ARIMA-ANN for high frequency time series data.

1.2 Objectives of the Study

i Compare the modeling perfomance of ARIMA, ANN and the Hybrid ARIMA-ANN for the EAC countries' currency exchange rates against the Kenya Shilling using high frequency time series data

ii Compare the forecasting perfomance of ARIMA, ANN and the Hybrid ARIMA-ANN for the EAC countries' currency exchange rates against the Kenya Shilling using high frequency time series data

2 Literature Review

2.1 Time Series Modeling and Forecasting of Currency Exchange Rates

Abounoori and Zabol (2020), observed that exchange rate forecasting has a substantial impact on the formulation of exchange rate policies that in turn influence the goods market hence the need for firms and individuals to pay attention to the prediction accuracy of the exchange rates. Various methods have been utilized in the forecasting of exchange rate in the literature with the ARIMA being the most extensively used method.

AHMAD, et al. (2023) applied ARIMA and GARCH in modeling Nigeria's Naira – US Dollar monthly exchange rates based on the monthly average official exchange rate (Naira/USD). Their study found out that the ARIMA (0, 2,2) and GARCH (1,1) models were the optimal models for the data. Tran (2016) applied ARIMA to monthly VND/USD exchange rates and established a model for the exchange rates of the Vietnam Dong vs. the United States dollar. The study identified the ARIMA(2,7,34) as the best-fitting model and concluded that ARIMA modeling was suitable for forecasting currency exchange rates.

Qimian Zhu (2023) utilized monthly data to develop an ARIMA(p,d,q)model for the USD/Euro exchange rates. He established an ARIMA(0,1,1) model for the data which was in contrast to a model identified by Dunis and Huang [2], who identified the ARMA (4, 4) model as the best for the USD/Euro exchange rate. On their part, AsadUllah et al. (2022) examined the accuracy of combined models with the individual models in terms of forecasting Euro against US dollar. They compared the forecasting performance of ARIMA, Naïve, exponential smoothing (ES) model, and nonlinear autoregressive distributive lags (NARDL) models and established that NARDL performed better than all the other models. They also considered a combination of different models and showed that the combination of NARDL and Naïve model had a higher forecasting accuracy compared to the individual models. This underscores the importance of considering hybrid models in modeling time series data.

Kamruzzaman and Ruhul employed the Artificial Neural Networks based on three learning algorithms: Standard Backpropagation (SBP), Scaled Conjugate Gradient (SCG) and Backpropagation with Bayesian Regularization (BPR) to predict six different currencies against the Australian dollar. The time series data frequency was weekly and found out that the SCG neural network model achieved very close prediction since it had the lowest NMSE and MAE metrics.

2.2 Conclusions and Knowledge Gap

Even though both the theoretical and empirical findings suggest that combining ARIMA and ANN methods as an effective way to improve time series data forecasting, there is little or no concrete focus on the extension od such methodology to forecasting the East African Community countries' currency exchange rates data. Moreover, most studies have applied ARIMA method using monthly and weekly data. Ignoring the use of high frequency time series data will limit the theoretical and methodological developments in the field of time series forecasting. This study compares he performance of ANN and ARIMA models in modeling and forecasting high frequency currency exchange time series data and then establish the modeling and forecasting performance of the hybrid ARIMA-ANN model under those frequencies.

3 Materials and Methods

3.1 Data Description

The data used was collected from the Central Bank of Kenya website, and each of the forex trading days was from 1st January 2017 to 31st December 2023. The data contains 1750 records of the currency exchange rates between the Kenya Shilling and the Tanzanian Shilling, Uganda Shilling, Rwanda Francs, and Burundi Francs.

3.2 Test for Stationarity

The augmented Dicky Fuller (ADF) test was applied for the hypothesis testing of stationarity. Dickey-Fuller test assumes a AR(1) type time series model represented mathematically as,

$$Y_t = \alpha y_{(t-1)} +_t \tag{3.1}$$

where Y_t is the variable of interest, t is the time index, α is a coefficient, and $_t$ is the error term. A unit root is present if $\alpha = 1$

3.3 Auto-Regressive Integrated Moving Average (ARIMA(p,d,q))

In the ARIMA(p,d,q) model, the parameters p, d, and q are non-negative integers that refer to the order of the autoregressive, integrated, and moving average parts of the model respectively. The ARIMA model will be distinguished by the autoregressive (AR) component, which estimates linear correlation with prior observations; an integrated (I) element that captures differencing and produces stationarity; and a moving average (MA) component. The ARIMA model parameters are identified based on the analysis of ACF and PACF plots. The obtained ARIMA model parameters allow for the analysis of the linear temporal trends of the data (Nyamao, 2014).

Definition 3.1. According to Liu et al (2016), if the differencing order is zero, we have an ARMA (p, q) which assumes that the time series is stationary and has the structural form:

$$\begin{aligned} X_t &= \alpha_1 X_{t-1} + \alpha_2 X_{t-2} +, \ldots, +\alpha_p X_{t-p} + \epsilon_t + \beta_1 \epsilon_{t-1} +, \ldots, +\beta_q \epsilon_{t-q} \\ &= \sum_{i=1}^{p} \alpha_i X_{t-i} + \sum_{j=0}^{q} \beta_j \epsilon_{t-j}, \quad \beta_0 = 1 \\ \varphi(B) X_t &= \theta(B)\epsilon_t \end{aligned} \tag{3.2}$$

Where (B) is backshift operator, $\varphi(B) = 1 - \alpha_1 B - \beta_2 B^2 - \cdots - \alpha_p B^p$ and $\theta(B) = 1 + \beta_1 B + \beta_2 B^2 + \cdots + \beta_q B^q$ are polynomials of orders p and q respectively. $\varphi(B)$ should be invertible, which means that the $det|\varphi(B)| \neq 0$

Definition 3.2. Suppose d is the non-seasonal differencing operator, such that $X_t^d = (1-B)^d X_t = X_t - X_{t-d}$, $X_t \sim ARMA(p, q)$ can be extended into an ARIMA(p,d,q) such that $X_t \sim ARIMA(p, d, q)$. Incorporating this into equation (3.2), yields the reduced model;

$$\begin{aligned} \varphi(B) X_t^d &= \theta(B)\epsilon_t, \\ \varphi(B) X_t^d &= \theta(B)\epsilon_t \end{aligned} \tag{3.3}$$

where $X_t \sim WN(0, \sigma^2)$ and $d \in \mathbb{Z}^+$ is the integration parameter. If $d = 0$, then $ARIMA(p, d, q) \equiv ARMA(p, q)$.

Therefore, Non-stationarity in time series of data is well handled by ARIMA (p, d, q) modeling.

If the time series data contains non-seasonal and seasonal components, a Seasonal ARIMA model denoted as $ARIMA(p, d, q)(P, D, Q)_S$ where p is the non-seasonal AR order, d is the non-seasonal differencing operator, q is the non-seasonal MA order, P is the seasonal AR order, D is the seasonal differencing, Q is the seasonal MA order and S is the period of repeating seasonal pattern is preferred for the modeling (Nyamao, 2013).

If $BX_t = X_(t-1)$, without differencing, a stationary SARIMA model is written as:

$$\phi(B)\Phi(B^S) X_t = \theta(B)\Theta(B^S)\epsilon_t, t = 0, 1, 2, \ldots \tag{3.4}$$

If the data is not stationary at both non-seasonal and seasonal levels, then the SARIMA model could be written as:

$$\phi(B)(1-B)^d \Phi(B^S)(1-B^S)^D X_t = \theta(B)\Theta(B^S)\epsilon_t \tag{3.5}$$

Where;
$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p$ is an autoregressive (AR) polynomial function of order p
$\theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q$ is a moving average (MA) polynomial of order q
$\Phi(B) = 1 - \Phi_1 B^S - \Phi_2 B^{2S} - \cdots - \Phi_P B^{PS}$

and
$\Theta(B) = 1 + \Theta_1 B^S + \Theta_2 B^{2S} + \ldots. + \Theta_Q B^{QS}$
are seasonal polynomial functions of order P and Q, respectively, that satisfy the stationarity and invertibility conditions

3.4 Artificial Neural Network

In the phase of modeling Artificial Neural Network (ANN), we seek to apprehend non-linear patterns and dependencies in high frequency time series data obtained from Central Bank of Kenya's daily exchange rates. The use of neural network architecture will be relatively simple since it consists only of an input layer, a hidden layer with the activation function ReLU and an output one containing linear units. The neural network is made to effectively acquire complicated relationships between input features, like historic exchange rates and relevant economic indicators from the target variable like SRR. The model is trained to historical data with the weights and biases adjusted through backpropagation using mean squared error loss function. The trained neural network can capture highly intricate, non-linear patterns in the data as it complements linear modeling approach of ARIMA model used within hybrid ARIMA–ANN forecasting method. Hyperparameters fine-tuning and model validation on a testing set improve ANN performance which in its turn improves the forecasting results. We will consider a simple feedforward neural network with one hidden layer.

$$h_j = \delta(\sum_{(i=1)}^{n} w_{ij} + b_i i) \tag{3.6a}$$

$$x_k = \delta \sum_{(j=1)}^{m} v_{jk} h_j + c_k \tag{3.6b}$$

Where;
x_i is the input feature
b_j is the bias of the hidden layer
w_{ij} is the weight between input and hidden layer
h_j is the output of the hidden later after applying the activation function δ
v_{kj} is the weight between the hidden and output layers
c_k is the bias of the output layer
y_k is the final output after applying the activation function

Remark 3.1. The logistic function is often made use of as the activation function (δ) in hidden layers and is expressed as:

$$g(x) = \frac{1}{1 + exp(-x)} \tag{3.7}$$

Equation (3.7) gives ANN models the ability to execute non-linear functional mappings [1].

3.5 Hybrid ARIMA-ANN Model

Sanjeev and Bhardwaj (2022) observe that Neither ARIMA nor ANN models are universally suitable for all kinds of time series when the data under consideration comprises both linear and nonlinear components, hence the need for Modeling techniques that can model simultaneously both the linear and nonlinear components. The study considered the hybrid methodology devised by Zhang (2003) that applies ARIMA and ANN and assessed its performance under various time series data frequencies. The ARIMA model handles linear temporal dependencies in stationary time series data, while the ANN model is good at recognizing non-linear patterns and long-term memory effect. The Hybrid ARIMA-ANN methodology is summarized in Figure 1:

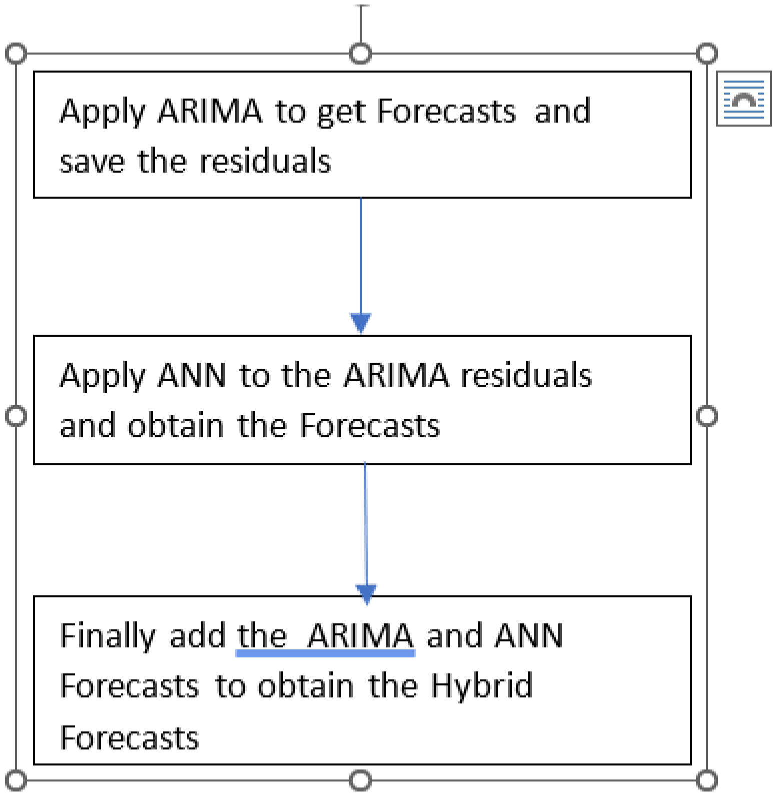


Figure 1: Hybrid ARIMA-ANN Methodology

Remark 3.2. forecasts from ARIMA and ANN models for a given time t are combined such that:

$$\hat{X}_{(Hybrid,t)} = \alpha\hat{X}_{(ARIMA,t)} + (1-\alpha)\hat{X}_{(ANN,t)} \tag{3.8}$$

Where;
$\hat{X}_{(Hybrid,t)}$ is the Hybrid model forecast at time t.
$\hat{X}_{(ARIMA,t)}$ is the ARIMA model forecast at time t.
$\hat{X}_{(ANN,t)}$ is the ANN model forecast at time t.
δ is the weight that determines the contribution of the ARIMA model within the interval $0 \leq \delta \leq 1$

3.6 Modeling and Forecasting Performance Metrics

To test the lack of fit in the ARIMA, ANN, and ARIMA-ANN hybrid time series models, the Ljung Box test. The Q* test statistic for the Ljung-Box test is given as:

$$Q^* = n(n+2)\sum_{(k=1)}^{m}\frac{\hat{r}_k^2}{(n-k)} \tag{3.9}$$

Where $\hat{r}_k$ is the estimated autocorrelation of the series at lag k and m is the number of lags being tested. The hypotheses tested were:

H_0:The model does not exhibit lack of fit.
H_1:The model exhibits lack of fit.
Box et al. (2015) observe that the null hypothesis (indicating that the model has a significant lack of fit) is rejected if

$$Q^* > \chi^2_{(1-\alpha,h)} \tag{3.10}$$

where $\chi^2_{(1-\alpha,h)}$ is the chi-square distribution table value with h degrees of freedom and α is the significance level

To assess the forecasting performance of the models, the Root Mean Squared Error (RMSE) and Mean Absolute Percentage Error (MAPE) were used, These are defined as:

$$MAPE = \frac{1}{n}\sum_{(i=1)}^{n}\frac{(|x_i - \hat{x}_i|)}{x_i} \times 100 \tag{3.11a}$$

$$RMSE = \sum_{(i=1)}^{n}\frac{||x_i - \hat{x}_i||^2}{n} \times 100 \tag{3.11b}$$

MAPE measures the percentage of the average absolute difference between predicted values and observed ones. Thus, it is a scale-free error metric since it takes the absolute value into account, it does not give information about the direction of the prediction error. Like MAPE, the RMSE value should be as small as possible for a good prediction performance.

4 Results and Discussion

4.1 Summary Statistics

The summary statistics are shown in Table 1

Table 1: Descriptive Statistics

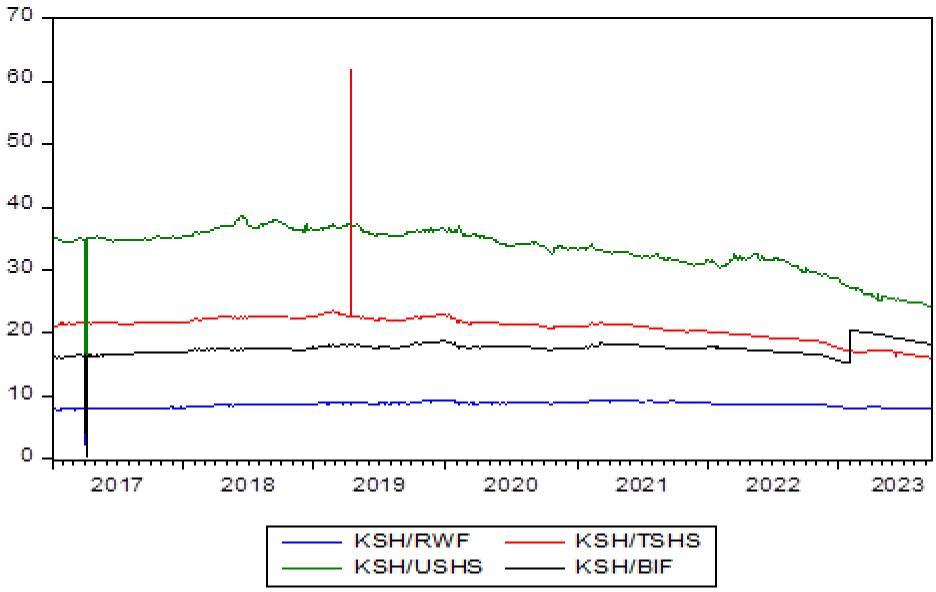
	KSH/RWF	KSH/TSHS	KSH/USHS	KSH/BIF
Mean	8.696767	20.97076	33.37707	17.66443
Median	8.777100	21.54320	34.53945	17.68915
Maximum	9.467400	61.93310	38.61960	20.62730
Minimum	0.121600	0.045900	0.028500	0.060300
Std. Dev.	0.462502	2.085097	3.464182	0.955076
Skewness	-3.869363	3.052612	-1.449287	-3.163397
Kurtosis	68.90642	92.78354	7.908600	69.06479
Jarque-Bera	321091.8	590505.3	2369.507	321167.6
Probability	0.000000	0.000000	0.000000	0.000000

The summary statistics table provides key descriptive statistics for the currency exchange rates between Kenyan Shilling (KSH) and other East African countries' currencies. For KES/BIF (Burundian Franc), the mean exchange rate is approximately 17.66, with a median rate also around 17.69 indicating a symmetric distribution. The standard deviation is 0.955, suggesting close spread of the KSH/BIF rates around the mean. However, the skewness value of -3.1633 does indicate that the KSH/BIF data are negatively skewed. This was true for the KSH/RWF and KSH/USHS exchange rates. The KSH/TSHS exchange rates exhibited a positive skewness and had a 20.97076 with other values deviating from the mean to the extend of 2.085097.

4.2 Currency Exchange Rates Time series Plots per Country

The plot in Figure 2 illustrates the exchange rates over time for East African countries from January, 2017, to December, 2023. The plot reveals fluctuations and trends in exchange rates across the observation period. While some currencies exhibit relatively stable rates, others display more volatile patterns, indicating potential economic or geopolitical influences. The overall trend suggests fluctuations in exchange rates over time, with periods of both increase and decrease observed. This visualization underscores the importance of monitoring exchange rate movements.

Figure 2: Time Series Plots for the Currency Exchange Rates



4.3 Test for Stationarity

The Augmented Dickey-Fuller test was used to test for serial autocorrelation. The hypothesis tested was H_0 There is a unit root vsH_1 Time series is stationary. The results of the ADF test in Table 2 show that before differencing, data for KSH/RWF, KSH/TSHS, and KSH/USHS were established to be non-stationary since the p-values were greater than 5%. The KSH/BIF data had a p-value of $0.0019 < 5\%$ hence we rejected H_0 that the KSH/BIF data had unit roots and concluded that without differencing, the KSH/BIF data were stationary. After differencing once, the KSH/RWF, KSH/TSHS, and KSH/USHS exchange rates data had p-values of zero and hence were established to be stationary.

4.4 ARIMA, ANN, and Hybrid ARIMA-ANN Models For the Currency Exchange Rates Data

The auto. arima and nnetar() functions in R automatically select the best ARIMA and ANN models for the given time series data. The Ljung Box statistics together with the corresponding p-values are shown in Table 3.

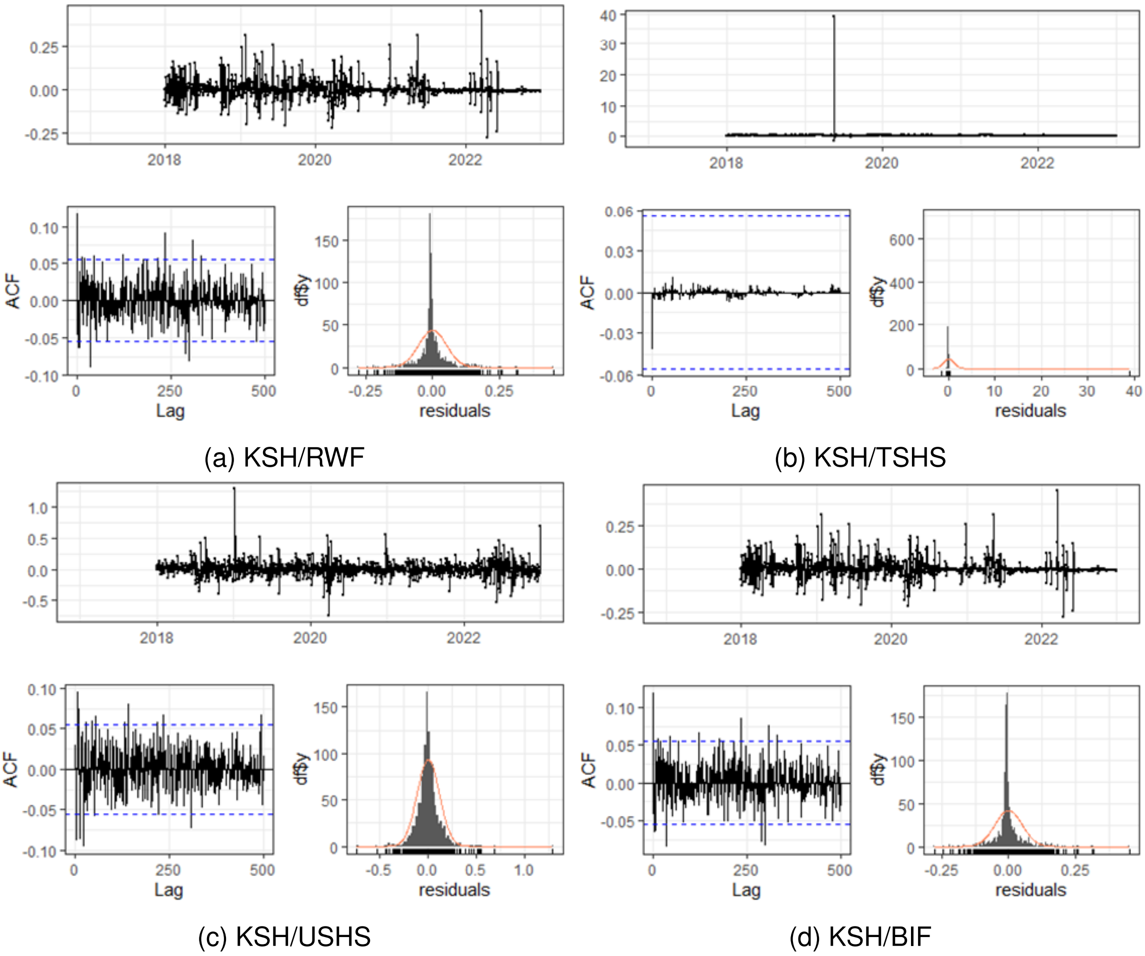
Table 2: ADF test for stationarity

Data Series	At Level		At 1st Difference	
	ADF Test Statistic	p-value	ADF Test Statistics	p-value
KSH/RWF	-1.70096	0.7508	-23.91471	.0000
KSH/TSHS	-2.974037	0.1399	-19.76881	.0000
KSH/USHS	-1.663701	0.7669	-22.70796	.0000
KSH/BIF	-4.443267	0.0019	-27.95786	.0000

Table 3: Comparison of the Modeling Performance of ARIMA, ANN, and Hybrid Method

Currency	ARIMA		ANN		ARIMA-ANN	
	Model	Q*(P-value)	Model	Q*(P-value)	Model	Q*(P-value)
KSH/RWF	(1,1,0)	0.0079	(10,1,6)	0.8182	(1,1,0)(5:1:2)	0.8748
KSH/TSHS	(0,1,1)	0.0001	(1,1,2)	0.9933	(0,1,1)(5:1:2)	1.000
KSH/USHS	(0,1,0)	0.0034	(1,1,2)	¡2.2E-16	(0,1,0)(5:1:2)	0.2033
KSH /BIF	(1,1,0)	0.0619	(11,1,6)	0.8349	(1,1,0)(5:1:2)	0.7476

Based on the Ljung Box test statistic, results in Table 3, show that the best-fitting model for the KSH/KRF exchange rates was the hybrid $ARIMA - ANN(1,1,0)(5:1:2)_{250}$ with the p-value for the Ljung Box statistic 0.8748 being greater than that of ANN(0.8182) and ARIMA (0.0079). The results show that the ANN approach also provided a good fit since its p-value was greater than 5%. The goodness-of-fit of the $ARIMA - ANN(1,1,0)(5:1:2)_{250}$ model for KSH/RWF is further illustrated in Figure 4 (a) and the ACF plot shows that the majority of the autocorrelations are within the confidence limits signifying that the model is a good fit for the data while the histogram plot shows that residuals were independently distributed. The study results also show that the best-fitting model for the KSH/TSHS currency exchange data was the hybrid $ARIMA - ANN(0,1,1)(5:1:2)_{250}$ with the p-value for the Ljung Box statistic being $1.000 > 0.05$. The p-value for the ANN model was $0.9993 > 0.05$ while that of ARIMA was $0.0001 < 0.05$. The results show that the ANN approach also provided a good fit since its p-value was greater than 5%. The goodness of fit of the model is further illustrated in Figure 4(b) and the ACF plot shows that all the autocorrelations lie within the confidence limits hence the model is a good fit for the data. The hybrid $ARIMA - ANN(0,1,0)(5:1:2)_{250}$ model with the p-value for the Ljung Box statistic being $0.2033 > 0.05$ was identified as the best fitting model for the KSH/USHS exchange rates data. Both ARIMA and ANN approches did not provide good fitting models. The residual plot for the hybrid $ARIMA - ANN(0,1,0)(5:1:2)_{250}$ model in Figure 4(c) shows that almost all the autocorrelations lie within the confidence limits hence the model is a good fit for the data. Finanlly, the study established that both the hybrid $ARIMA - ANN(1,1,0)(5:1:2)_{250}$ and ANN (11,1,6) provided good fits for the KSH/BIF data as evidenced by their respective p-values for the Ljung Box test statistic of 0.8349 and 0.7476 respectively that were all greater than 5%. However, based on the magnitude of the p-values, it can be inferred that the ANN model was a better fit than the hybrid model. The ACF plot in the residual plot shown in Figures 4 (d) show that almost all the autocorrelations lie within the confidence limits hence the models are a good fit for the data.



(a) KSH/RWF

(b) KSH/TSHS

(c) KSH/USHS

(d) KSH/BIF

Figure 3: Forecast Plots for the hybrid Models for the EAC Currency Exchange Rates

4.5 Forecasting Performance of ARIMA, ANN, and Hybrid ARIMA-ANN Models for the EAC countries' Currency Exchange Rates

To measure the forecasting performance of the ARIMA, ANN, and ARIMA-ANN models, Root Mean Squared Error (RMSE) and Mean Absolute Percentage Error were used and results are shown in Table 4.

The findings show that model for KSH/RWF resulted in RMSE values of 0.5218, 0.0423, and 0.0178 respectively for ARIMA, ANN and the hybrid ARIMA-ANN. This represents a decrease of 96.59% of the RMSE for the hybrid model compared to the ARIMA Model and a decrease of 57.91% for the hybrid model compared to the ANN model. The predictive accuracy of the models compared using the MAPE values show that the hybrid ARIMA-ANN model had the lowest MAPE value of 0.2709 compared to values of 18.42, and 0.2803 for the ARIMA and ANN models. In this case, the selected model for the ARIMA part is ARIMA (1,1,0) with a drift term while the ANN model using the residuals

Table 4: Forecasting Performance of ARIMA, ANN, and Hybrid ARIMA-ANN Models

Currency	ARIMA		ANN		ARIMA-ANN	
	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE
KSH/RWF	18.42	0.5218	0.2803	0.0423	0.2709	0.0178
KSH/TSHS	32.1576	1.1811	0.5815	1.2652	0.5262	0.1047
KSH/USHS	81.756	1.2803	0.4488	0.0197	0.4552	0.1162
KSH /BIF	18.420	0.5218	0.1460	0.0460	0.1397	0.0308

from the ARIMA(1,1,0) model was NNAR(5,1,2). Overall, $ARIMA - ANN(1,1,0)(5:1:2)_{250}$ was established to be a better model for forecasting the KSH/RWF currency exchange rates.

On KSH/TSHS exchange rates, the study result show that the hybrid ARIMA-ANN model had the least RMSE of 0.1047 compared to 1.1811 and 1.2652 for ARIMA and ANN respectively. The predictive accuracy of the models compared using the MAPE values show that the hybrid ARIMA-ANN model had the lowest MAPE value of 0.5262 compared to values of 32.1576, and 0.5815 for the ARIMA and ANN models. The selected model for the ARIMA part is ARIMA (0,1,1) with a drift term while the ANN model using the residuals from the ARIMA(0,1,1) model was NNAR(5,1,2) Overall, $ARIMA - ANN(0,1,1)(5:1:2)_{250}$ was established to be a better model for forecasting the KSH/TSHS currency exchange rates.

On KSH/USHS exchange rates, the study result show that the ANN model had the least MAPE and RMSE of 0.4488 and 0.0197 making it the best prediction model. The hybrid ARIMA-ANN model also had lower MAPE and RMSE values compared to the ARIMA model. The ANN model selected using nnetar() function was NNAR(1:1:2) and is a better model for forecasting the KSH/USHS currency exchange rates as it outperforms the hybrid ARIMA-ANN model.

On KSH/BIF exchange rates, the study result show that the hybrid ARIMA-ANN model had the least RMSE of 0.0308 compared to 0.5218 and 0.0460 for ARIMA and ANN respectively. The predictive accuracy of the models compared using the MAPE values show that the hybrid ARIMA-ANN model had the lowest MAPE value of 0.1397 compared to values of 18.420 and 0.1460 for the ARIMA and ANN models. The selected model for the ARIMA part is ARIMA (1,1,0) with a drift term while the ANN part was NNAR(5,1,2). Overall, $ARIMA - ANN(1,1,0)(5:1:2)_{250}$ was established to be a better model for forecasting the KSH/BIF currency exchange rates. Forecast Plots for the Hybrid Model for days of some months of 2024 are also shown in Figure 3. The results show a forecasted increasing trend for the KSH/BIF, KSH/RWF, and KSH/TSHS exchange rates while a slow downward trend is forecast for the KSH/USHS. This indicates that over the 2024 period, the Ugandan Shilling is expected to gain considerably over the Kenyan Shilling. Overall, the forecast results show that the Kenyan Shilling will be steady against the other EAC countries' currencies, and thus is major investment hub with the understanding of this risk and adjustments on the investor's part.

6 CONCLUSIONS

a On the modeling performance of the ARIMA, ANN and hybrid ARIMA-ANN, the study concludes that the hybrid ARIMA-ANN modeling strategy provided good fiiting models for the EAC countries' high frequency currency exchange rates data. The ARIME modeling strategy is not less efficient in modeling high frequency time series data.

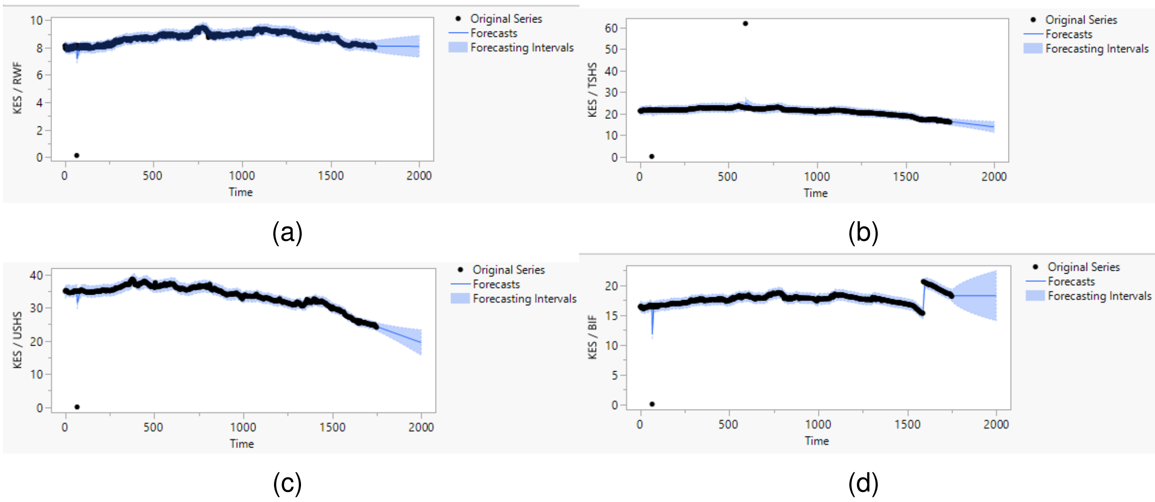


Figure 4: AForecast Plots for the Mixture Model for the EAC Currency Exchange Rates

b This study finds, based on the results described in Section 5.1, that the hybrid model ARIMA-ANN model significantly improves the ANN and ARIMA models in short-term and long-term forecasting accuracy, hence could be used in place of its constituent models in forecasting high frequency time series data.

.

Disclaimer (Artificial Intelligence)

Author(s) hereby declare that NO generative AI technologies such as Large Language Models (ChatGPT, COPILOT, etc) and text-to-image generators have been used during writing or editing of manuscripts

References

[1] Alshawarbeh, E.; Abdulrahman, A.T.; Hussam, E.(2023). Statistical Modeling of High Frequency Datasets Using the ARIMA-ANN Hybrid. *Mathematics*, 11, 45-94. https://doi.org/10.3390/math11224594

[2] Dunis C. L., and Huang X. (2002) Forecasting and trading currency volatility: An application of recurrent neural regression and model combination. *Journal of forecasting*, 21(5), 317-354, (2002).

[3] Crâşmăreanu, M. and Hreţcanu, C-E. (2008). Golden differential geometry. Chaos, Solitons & Fractals, 38(5), 1229–1238.

[4] Fischler, R-H. (2000). The Shape of the Great Pyramid Waterloo. Ontario: Wilfrid Laurier University Press.

[5] Fischler, R-H. (1998). A Mathematical History of the Golden Number. Dover Publications.

[6] Goldberg, S-I. and Yano, K. (1970). Polinomial structures on manifolds. Kodai Math. Sem. Rep., 22 (MR 42 #2380), 199–218.

[7] Goldberg, S-I. and Petridis, N-C. (1973). Differentiable solutions of algebraic equations on manifolds. Kodai Math Sem. Rep. 25, 111-128.

[8] Hreţcanu, C-E. and Crâşmăreanu, M. (2009). Applications of the Golden Ratio on Riemannian Manifolds. Turk J Math 33 (doi:10.3906/mat-0711-29), 179-191.

[9] Hreţcanu, C-E. and Crâşmăreanu, M. (2007). On some invariant submanifolds in Riemannian manifold with Golden Structure. Scientific Annals of Alexandru Ioan Cuza University - Mathematics, Iasi, Romania, s. I-a, Math. 53, 199-211.

[10] Hreţcanu, C-E. (2007). Submanifolds in Riemannian manifold with Golden Structure. In: Workshop on Finsler geometry and its applications, Hungary.

[11] Livo, M. (2002). The Golden Ratio: The Story of phi, the World's Most Astonishing Number. Broadway, MR 2003k:11025.

[12] Miron, R. and Anastasiei, M. (1994). The geometry of Lagrange spaces: theory and applications. Kluwer Academic Publishers FTPH No. 59, MR 95f: 53120.

[13] Perkins, K. (2006). The Cartan-Weyl conformal geometry of a pair of second order partial differential equations. University of Pittsburgh, *Unpublished Ph.D thesis*.

[14] Seiichi, I. and Akifumi, K. (2008). Golden duality in dynamic optimization. In: Proceedings of kosen workshop MTE2008-mathematics, technology and education-Ibataki national college of technology Hitachinaka, Ibaraki, Japan: February 15–18.

[15] Yano, K. and Kon, M. (1994). Structures on Manifolds. World Scientific, Singapore, Series in pure matematics, 3.

[16] Encyclopedia. (2011). Golden ratio. http://en.wikipedia.org/wiki/Golden ratio. *(Last accessed on 18 January 2011 at 18:02).*