Modeling export price of tea in Kenya: Comparison of artificial neural network and seasonal autoregressive integrated moving average

Mbiriri Ikonya[*], Peter Mwita, Anthony Wanjoya

Department of Statistics and Actuarial Science, Jomo Kenyatta University of Agriculture and Technology (JKUA T), Nairobi, Kenya

Email address:

ikonyaz@gmail.com (M. Ikonya)

To cite this article:

Mbiriri Ikonya, Peter Mwita, Anthony Wanjoya. Modeling Export Price of Tea in Kenya: Comparison of Artificial Neural Network and Seasonal Autoregressive Integrated Moving Average. *American Journal of Theoretical and Applied Statistics.* Vol. 3, No. 6, 2014, pp. 211-216. doi: 10.11648/j.ajtas.20140306.16

Abstract: Agriculture sector is a key driver of economic growth in Kenya. It remains the main source of livelihood for the majority of the Kenyan people. Tea, coffee, and horticulture are the main agricultural exports in Kenya. Export price of these commodities fluctuates mainly due to law of demand and supply. Other reasons include; quality of goods and inflation effect on the dollar or other hard currencies. Further, farmers and their cooperative societies are affected by the local foreign exchange. The government and other stake holders require prior information on price trends for ease of planning. *Thus it is important to forecast export prices of these commodities.* The purpose of this study is to compare the forecasting performance of artificial neural network (ANN) model and a SARIMA model in export price of tea in Kenya. Secondary data was obtained from Kenya National Bureau of Statistics (KNBS). A total of 185 monthly export prices were obtained. A three layer feed-forward artificial neural network was trained using *70% of the data*. The ANN model obtained was used to predict export prices for the remaining *30% of the data*. SARIMA model was also used to predict export prices for the same duration. Forecasting performance was evaluated using Root mean squared errors (RMSE), mean absolute error (MAE) and mean absolute percentage error (MAPE). ANN demonstrated a superior performance over SARIMA model. The authors' ANN has high performance compared to SARIMA and can accurately predict export price of tea.

Keywords: Artificial Neural Network (ANN), Seasonal Autoregressive Integrated Moving Average (SARIMA),
　　　　　　Kenya National Bureau of Statistics (KNBS)

1. Introduction

Agriculture sector is the key driver of economic growth in Kenya (Economic survey, 2014). It contributes approximately 25% of the Gross Domestic Product (GDP), employs 75% of the national labour force and accounts for about 65% of Kenya's total export. It also indirectly contributes approximately 27% to GDP through linkages with manufacturing, distribution etc. Major agricultural exports includes; tea (32%), horticulture (24%) and coffee (5.2%)-KNBS (2013). Small scale tea farmers are managed by Kenya Tea Development Agency (KTDA) and produce over 60% while the rest is produced by large scale producers (Tea performance report 2013, Tea Board of Kenya (TBK)).

Export price fluctuations are mainly due to law of demand and supply, crude oil prices, quality of goods and inflation effect on the dollar or other hard currencies. Fluctuations in export earnings can adversely affect the process of economic development. Low exports mean low foreign exchange and thus small purchasing capacity of a nation in the international market. Export fluctuations tend to cause balance of payment complexities leading to low confidence of people in the process of maintenance of the exchange rate (Deekay(2009)). This stimulates inflation. Due to export price fluctuations, farmers lack regular returns for their produce. Government is unable to predict tax revenue, plan for overall balance of trade and regulate foreign exchange. Financial institutions need to plan for farmer's loan advances.

Therefore price fluctuations are a matter of concern among consumers, farmers, financial institutions and policy makers. Its accurate forecast is extremely important for efficient monitoring and planning. In this study ANN and SARIMA

models were used and data was obtained from the KNBS.

2. Review of Previous Studies

Box and Jenkins (1976) showed that SARIMA model is suitable for stationary time series and has the advantage of accurate forecasting over short periods.

Kyungjoo Lee et al.(2007) showed that the Back Propagation Neural Network (BPNN) model was generally better than the SARIMA model in forecasting the Korean Stock Exchange Index(KOSPI) but for short term forecasting, the SARIMA model generally provided more accurate forecasts.

Chikobvu and Sigauke (2012) developed SARIMA and regression with SARIMA errors models to predict daily peak electricity demand in South Africa. Empirical results showed that the SARIMA model produces more accurate short-term forecasts. The regression-SARIMA modeling framework captured important drivers of electricity demand.

Hutchinson et al. (1994) examined stock option price data and shown that the ANN model is computationally less time consuming and more accurate non-parametric forecasting method, especially when pricing dynamics are unknown or when the pricing equation cannot be solved analytically.

Zhang et al. (1998) proposed the application of post-processing methods to improve the ANN's predictions.

Monica Adya and Fred Collopy (1998) concluded that; ANNs, when they are effectively implemented and validated, show potential for forecasting and prediction.

3. Methodology

3.1. Seasonal Autoregressive Integrated Moving Average

Let y_t : t = 0,1,2,… be a non-stationary time series containing seasonality i.e. a seasonal periodic component repeats itself after every s observations, y_t then depends on past values such as $y_{t-1s}, y_{t-2s}, ...$ as well as $y_{t-1}, y_{t-2}, ...$

$w_t = \nabla^d y_t$ where ∇ is a differencing operator and d is the order of non-seasonal differencing is an Autoregressive moving average (ARMA (p, q)) process.

$$A_p(z) w_t = B_q(z) \in_t \quad (1)$$

then, y_t is Autoregressive Integrated Moving Average (ARIMA (p, d q)) model.

If $w_t = \nabla^d \nabla^D y_t$, where D is the order of seasonal differencing, then the model is SARIMA written as ARIMA (p ,d, q)(P, D, Q) is given by

$$A_p(z) \, \mathcal{A}_P(z^s) w_t = B_q(z) \mathcal{B}_Q(z^s) \epsilon_t \quad (2)$$

Where z is the back shift operator, $A_p(z)$, $\mathcal{A}_P(z^s)$, $B_q(z)$ and $\mathcal{B}_Q(z^s)$ are polynomials of order p, P, q and Q respectively.

To obtain the forecasts from the SARIMA model, we adopted the Box and Jenkins' method. Basically, Box and Jenkins' method uses three-stage approach to select an appropriate model for the purpose of estimating and forecasting a time-series data namely; model identification, estimation of parameters and diagnostic check.

3.2. Artificial Neural Network Model

ANN is a parallel connection of a set of nodes called weights. ANN has an input layer, hidden layer and an output layer.
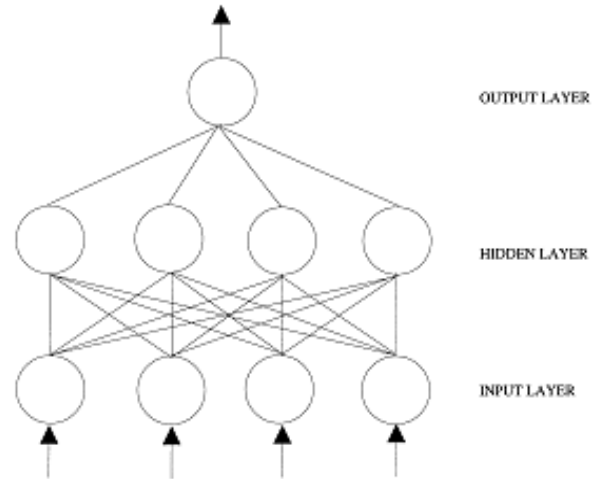

Figure 3.1. *Feed forward neural network.*

Figure 3.1 is an input-output map with d input nodes, one layer of H hidden nodes, one output node and an activation function $\psi(x)$. The inputs at hidden layer nodes are connected by weights W_{h0} for $h \in (1,...H)$ and $j \in (0,...d)$ where W_{ho} is the weight of the bias for the hth hidden node (Kyalo et al.). The hidden and output layers are connected by weights α_h for $h \in (0,...H)$.

Consider an input vector $x = x_1, x_2,...x_d \in \Re^2$.

The input $V_h(x)$ to the hth hidden node is;

$$V_h(x) = W_{ho} + \sum W_{hj} x_j \quad (3)$$

The output $\phi h(x)$ of the hth hidden node is the value;

$$\phi h(x) = \psi(V_h(x)) \quad (4)$$

The net input to the output node is the value;

$$\mathbb{Z}(x) = \alpha_0 + \sum \alpha_h \phi h(x) \quad (5)$$

Finally the output Z (x) of the net is the value;

$$Z(x) = \psi(\mathbb{Z}(x)) = \psi\left(\alpha_0 + \sum \alpha_h \phi h(x)\right)$$

3.2.1. Training a Neural Network

The Sum of squared error (SSE) is used to train faced forward networks. In this method the weights are adjusted in such a way that the SSE between the targets y and the goal of output Z is minimized. The SSE is defined as:

$$S^2(x_i, \theta) = \sum \left(y_i - Z(x_i, \theta)\right)^2$$

$$= \sum(y_i - \psi(\alpha_0 + \sum \alpha_n \psi(W_{ho} + \sum W_{hj} x_j)))^2 \quad (6)$$

which is an activation function.

3.2.2. Artificial Neural Network Forecasting

ANN model performs a nonlinear functional mapping from the past observations $(y_{t-1}, y_{t-2}, \dots y_{t-p})$, to the future value (y_t), i.e.

$$y_t = \phi(y_{t-1}, y_{t-2}, \dots y_{t-p}) + e_t \qquad (7)$$

where ϕ is a function determined by the network structure and connection weights.

To train the ANN a common Quasi Newton method known as BFGS was used in this study. This was suggested independently by Broyden, Fletcher, Goldfarb, and Shanno, in 1970.

4. Data and Empirical Results

This chapter describes the source of the data, properties of the export price series, how SARIMA and ANN models were used in forecasting. The chapter also describes how the forecast results from both models were compared. The data used in this research is export price per tonne of tea in Kenya. This is monthly data for the period August 1998 to December 2013. The data was obtained from the Kenya National Bureau of Statistics.

4.1. The Properties of the Export Price of Tea

Figure 4.1 is a time series plot for the export price of tea. On examining it, the following features were observed; an upward trend, an outlier at September 2001, random and cyclic components.
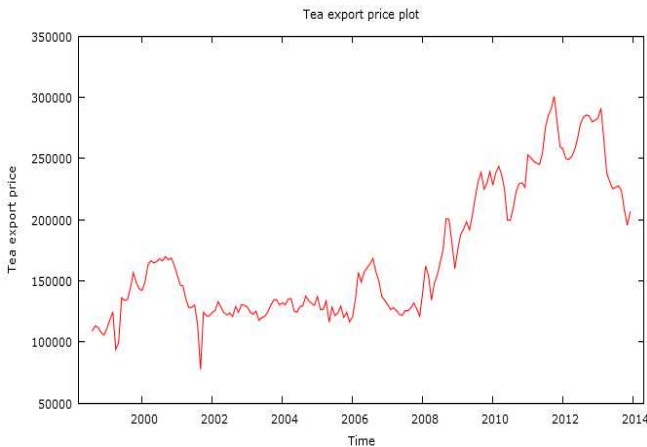

Figure 4.1. *time series plot.*

4.1.1. Summary Statistics for Export Price of Tea

Table 1. *Summary statistics*

Coefficient of Variation	0.32299
Skewness	0.79744
Excess Kurtosis	- 0.68543
Augmented Dickey-Fuller(p-value)	0.6408
Shapiro-Wilk (p-value)	< 0.0001
Box Pierce test (p-value)	< 0.0001
Breusch-Pagan (p-value)	< 0.0001

The summary statistics for export price of tea are shown in table 1. Skewness is a measure of symmetry and a positive value (0.79744) means that the data series is skewed to the right. Excess Kurtosis value (-0.68543) shows that the data has a flat distribution relative to a normal distribution. If the coefficient of variation is less than one such as ours (0.3229), then the distribution is considered to have a low variance. We reject the null hypothesis for normality using Shapiro-Wilk (p-value<0.0001). Also we reject the null hypothesis that the process is independent and identically distributed at 95% confidence interval using Box Pierce test statistics. Thus the export price of tea possess autocorrelation i.e. not independent and identically distributed. Augmented Dickey-Fuller test was used to test for unit root in the export price of tea series and p value > 0.05 obtained. The null hypothesis that a=1 was not rejected. Therefore the time series data is not stationary. Also the null hypothesis that heteroscedasticity is not present was rejected using Breusch-Pagan test whose p-value< 0.05 and thus concluded that the series is heteroscedastic.

4.2. SARIMA Modeling for Export Price of Tea

4.2.1. Model Identification and Parameter Estimation

Box and Jenkins' three-stage approach of selecting an appropriate model for the purpose of estimating and forecasting a time-series data was used. The model that was found to be reasonable based on the least values of the Akaike information criterion is SARIMA (0; 1; 1) (0; 1; 1)12. The fitted model can be written as;

$$(1 - z)(1 - z^{12})w_t = e_t + 0.19958e_{t-1} - 0.88744e_{t-12} - 0.17712e_{t-13} \qquad (8)$$
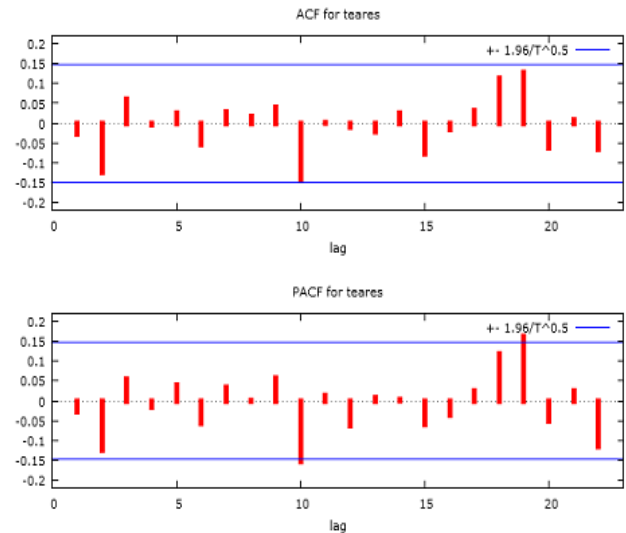
4.2.2. Diagnostic Check

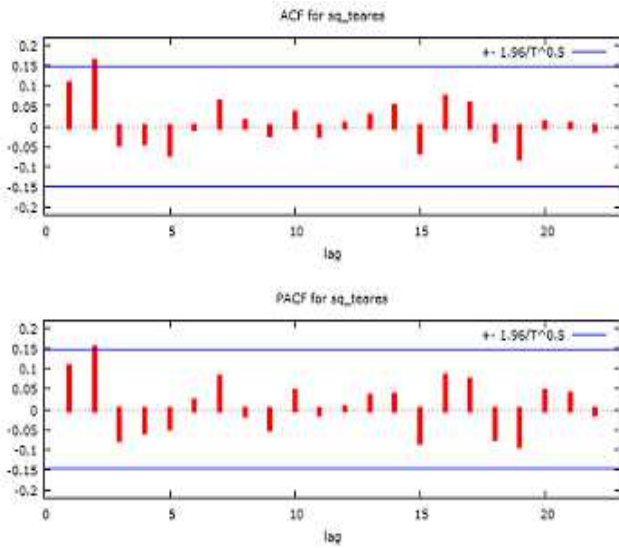**Figure 4.2.** *ACF and PACF of residuals*
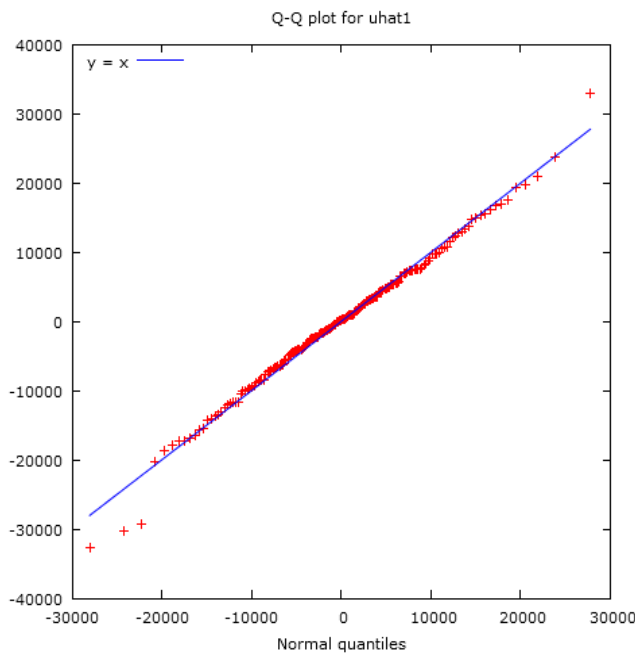
Figure 4.3. *ACF of squared residuals*

Figure 4.4. *Residual QQ plot for SARIMA model*

Quantile-Quantile (QQ) plot in Figure 4.4 approximately follows the QQ line visible on the plot. Also the p-value of Shapiro Wilk normality test (W = 0.991, p-value = 0.3509) is more than 0.05. This indicates that residuals are normally distributed at 5% level of significance. The Autocorrelation function (ACF) and Partial Autocorrelation function (PACF) of the residuals in Figure 4.2 are white noise. Most spikes are within 95% confidence interval. Results of Durbin Watson statistics d=2.03 is near two which indicates non-autocorrelations amongst residuals. The ACF and PACF of the squared residuals in figure 4.3 were examined. Most of their values are within the 95% confidence interval. Thus the residuals from this model are homoscedastic.
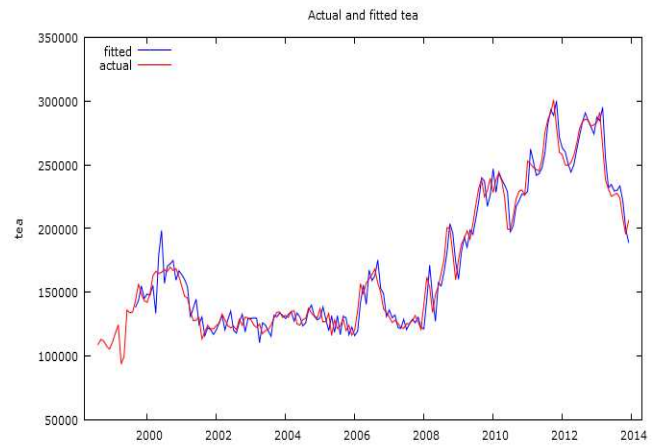
4.2.3. Forecast Graph

Figure 4.5. *Forecast graph*

4.3. ANN Modeling for Export Price of Tea

4.3.1. Training and Testing of Data

The log transformed price data was split such that 70% of the data was used as training data set (August 1998 to April 2009) while 30% was used as testing data set (May 2009 to December 2013). The training set was used to optimize the weights and the bias of the network, while testing set was used to determine the generalization ability of the network. The transformed data was scaled using a linear function to an interval of 0 and 1.The quasi-newton method known as BFGS (Broyden, Fletcher, Goldfarb, and Shanno) was used to train our model and validation of the model done using testing data set.

4.3.2. Selecting Best ANN Model for Tea Price

Errors were computed for various lag combinations. We selected the best ANN model based on the least errors. Some of the lag combinations are listed in table 2. ANN model selected had lag 1 and 3 and can be written as;

$$y_t = \phi(y_{t-1}, y_{t-3}) \qquad (9)$$

Table 2. *ANN model identification*

LAGS	MAE	RMSE	MAPE
1,5,9,13	8200.35	9772.11	3.506
1,3,5	3732.13	4029.54	1.569
1,5,9	6876.87	7977.23	2.706
2,5	1901.75	2142.00	0.795
2,3,5	5516.24	6935.87	2.276
1,2	2814.78	3801.48	1.091
1,5	2472.82	2766.23	1.033
1,7	2000.17	2238.81	0.869
1,3*	1032.95	1143.42	0.439
1,4	4315.12	5930.08	1.627

4.3.3. Residual Analysis for ANN Model

The Autocorrelation function (ACF) and Partial Autocorrelation function (PACF) of the residuals in Figure 4.6

are white noise since there are no significant spikes. Further results of Durbin Watson statistics d=1.83 is near two which indicates non-autocorrelations amongst residuals. The Shapiro wilk statistics yielded W = 0.9831, p-value =0.1498. The p value is more than 0.05. Thus we conclude that the residuals are normally distributed at 5% level of significance. This is also confirmed by residual Quantile-Quantile (QQ) plot in figure 4.7 which approximately follows the QQ line visible on the plot. The ACF and PACF of the squared residuals in figure 4.8 were examined. Most of their values are within the 95% confidence interval. Thus the residuals from this model are homoscedastic.
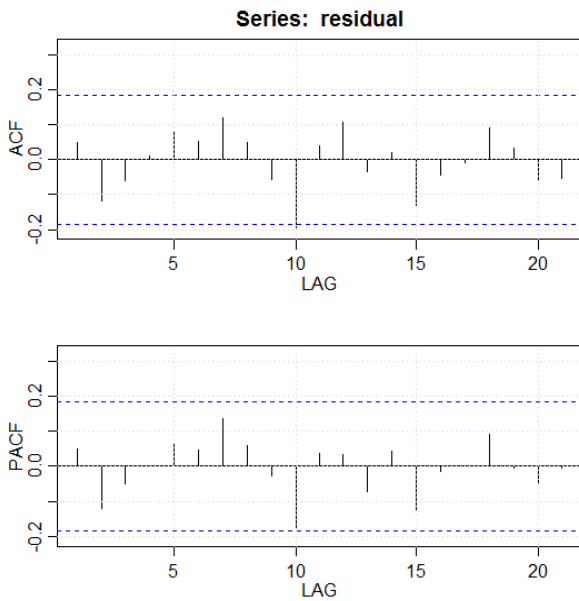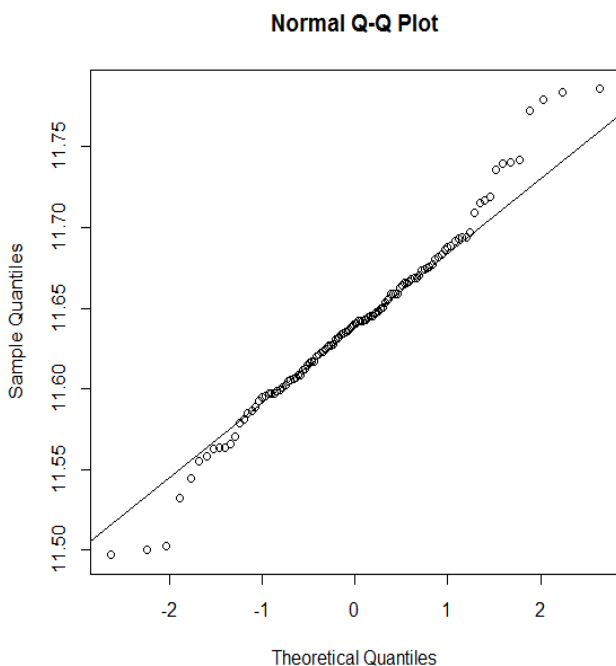

Figure 4.6. *Residual ACF and PACF of ANN model*

Figure 4.7. *Residual QQ plot of ANN model*
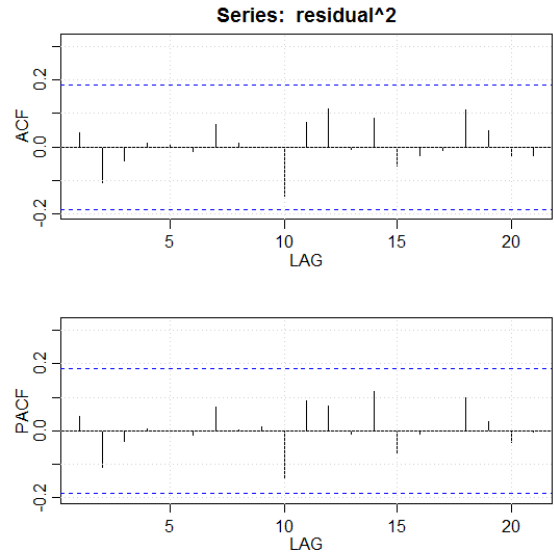

Figure 4.8. *ACF of squared residuals*
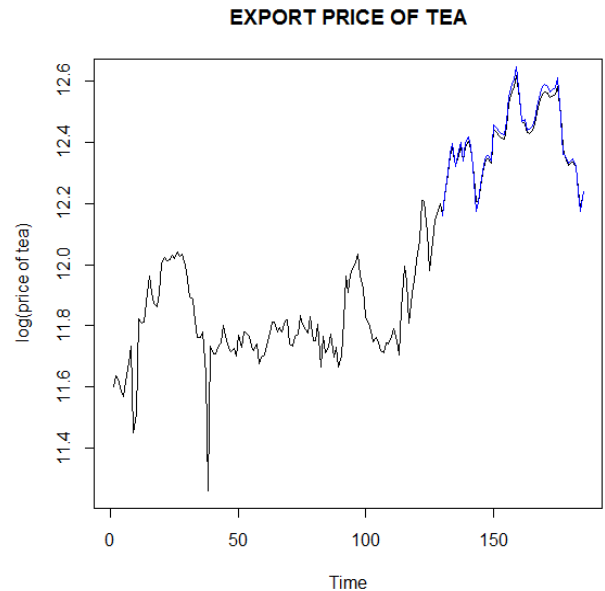
4.3.4. Forecast Graph

Figure 4.9. *Forecast graph*

4.4. Performance Statistics of SARIMA and ANN Models

Performance Statistics of SARIMA and ANN models are summarized in Table 3. Both models have a mean absolute percentage error. The mean absolute error, root mean squared error, mean absolute percentage error, are smaller for ANN model than for SARIMA model. It's clear that ANN model performs better than SARIMA model in forecasting export price of tea

Table 3. *Performance Statistics Table*

Model	MAE	RMSE	MAPE
SARIMA Model	8608.7	11019	3.6144
ANN Model	1032.9	1143.4	0.4399

5. Summary and Conclusion

The main objective of this study was to compare performance of SARIMA and ANN models in forecasting export price of tea. SARIMA model and ANN model can effectively forecast export price of tea in Kenya. However ANN is considered as the best model in this study.

We recommend that future research should target coffee and horticultural commodities. For ANN model we used lags as predictors. We recommend use of at least one other predictor such as international crude oil prices and foreign exchange rates.

References

[1]   Box, G. P. and Jenkins, G. M. (1976), "Time Series Analysis, Forecasting and Control". Holden-Day. San Francisco. CA.

[2]   Tanaka, H. (1987): "Fuzzy Data Analysis by Possibility Linear Models". *Fuzzy Sets and Systems*. 24(3), 363-375

[3]   Hornik, K., M. Stinchcombe, and H. White (1989): "Multi-Layer Feed forward Networks are Universal Approximators," *Neural Networks*, 2, 359—366

[4]   Refenes A. N., A. Zapranis, and G. Francis,(1994) "Stock Performance Modeling Using Neural Networks: A Comparative Study with Regression Models", *Elsevier B.V*, 5, 961-970.

[5]   Hutchinson J., W. Andrew, & T. Poggio (1994). "A non-parametric approach to pricing and hedging derivative securities via learning networks". *Journal of Finance*, 49, 851-889.

[6]   Zhang, G., and Hu, M.Y. (1998), Neural network forecasting of the British pound/US dollar exchange rate. *International Journal of Management Science*, 26, 495 – 506.

[7]   Monica Adya and Fred Collopy,(1998), "How effective are neural networks at forecasting and predicting? A review and evaluation", *Journal of Forecasting*, 17, 481-495.

[8]   Shekhar, S. (2004). "Recursive Methods for Forecasting Short-Term Traffic Flow using Seasonal ARIMA Time Series model". Master's thesis. Graduate Faculty of North Carolina State University. Raleigh. North Carolina. USA.

[9]   Kyungjoo Lee, Sehwan Yoo, John Jongdae Jin,(2007) "Neural network model vs. SARIMA model in forecasting Korean stock price index", *Issues in Information Systems*, 8,(2),372-378.

[10] Cao Q, Parry M (2009), Neural Network Earnings per Share Forecasting Models: A Comparison of Backward Propagation and the Genetic Algorithm. *Decision Support System. International journal,* 47, 32-41.

[11] Zhang, B. Eddy Patuwo, Michael Y. Hu ,(1998) "Forecasting with artificial neural networks: The state of the art", *International journal of forecasting*, 14 ,35–62.

[12] Kyalo Richard, Waititu Anthony, Wanjoya Anthony(2014). Artificial Neural Network Application in Modelling Revenue Returns from Mobile Payment Services in Kenya. *American Journal of Theoretical and Applied Statistics.* 3, 60-64.