

Forecasting the Unit Price of Coffee Export in Rwanda Using Arima Model

MWANGABWOBA Janvier¹, Dr. Joseph K. Mung'atu²

^{1,2}Jomo Kenyatta University of Agriculture and Technology, Kenya

Abstract: Unit price of coffee export which is the cost of coffee exported per kilogram is one of the indicators to see the economic stability of exported coffee in Rwanda. The value of unit price of coffee export in Rwanda on January 2010 - December 2017 is unstable. unit price which unstable will impede (slow) the economic development in Rwanda, therefore need to be undertaken against the value of the modeling unit price in the future with a method of ARIMA. The purpose of this study is to find an ARIMA model which appropriate to forecasting unit price of coffee export in Rwanda and to know the forecasting unit price in Rwanda on January 2018 - December 2019 using Eviews. The Box-Jenkins methods used in this study is a literature method i.e. authors collect, select and analyze readings related to the issues examined and methods documentation i.e. the author collected secondary data from January 2010 - December 2017 in Rwanda. Based on the research obtained, the model appropriate to forecasting unit price of coffee export in Rwanda is a model ARIMA (1,1,1). The results of the forecasting unit price of coffee export using Eviews software on January 2018– December 2019 is stable enough. The lowest unit price occurred on February 2018 while the highest occurred on December 2019.

Keywords: time series, unit price, stationarity, forecasting

1. Introduction

The socio-economic situation of the world reveals that man was and still need some items that cannot produce it or produces insufficient quantities to meet its needs. Those items are necessary to the survival of the individual, must be available and the most loyal to acquire is the exchange. Rwanda is a land locked eastern African nation in the Great Lakes region. It shares international borders with Uganda, Tanzania, Burundi and the Democratic Republic of Congo. Rwanda's economy primarily depends on agricultural productivity.

In all Rwandan products, coffee is an important part of export earnings. It is about 50% of total revenue. That is why we have limited the analysis on it. We can distinguish three periods in the development of Rwanda's exports sector since independence:

In 1962-1986 period merchandise exports averaged 8% of GDP, with 60% of income coming from coffee exports.

In 1986-1995 period Rwanda's exports sector collapsed due a large drop in global coffee prices, which fell by an estimated 70% between 1986 and 1992. This decline was amplified by the unsustainability of inward oriented policies and eventual economic instability following consecutive devaluations in the early 1990s and the political instability leading to the 1994 genocide.

The 1995-2011 period: While Rwanda's exports sector has fully recovered from the 1994 genocide – increasing from about 5% of GDP in 1994 to an average of about 12% since 2004. In this study we will analyze the change of unit price of coffee export from Rwanda in order to forecast the unit price in 24 months ahead.

2. Problem Statement

Government of Rwanda makes revenue projections in order to finalize the fiscal and monetary policy to do so it would be better to have a model that estimates the future revenue generated from coffee as one of the main export components of Rwanda. It is obvious that a successful time series forecasting depends on an appropriate model fitting. A lot of efforts have been done by researchers over many years for the development of efficient models to improve the forecasting accuracy. This study aims to formulate an ARIMA model that can be used to forecast the unit price of coffee export in Rwanda using mathematical and statistical knowledge.

3. Justification

Export in Rwanda is one of the macroeconomic framework, with the exportation of coffee one of the major export in Rwanda, predicting the future behavior of this agricultural product is uncertainty due to the ignorance of mathematical and statistical techniques while doing prediction. The main interest of this study is to help Rwanda planners and policy makers getting a fitted model for forecasting the future evolution of the export. The second interest of this study is to develop the knowledge of some readers who misunderstand the use of statistical notions in Economic Institutions. The last interest is to provide the tools for the future researchers in the same fields.

4. Objectives of the Study

4.1 General Objective

The main objective of the research of the study was to formulate an ARIMA model that can be used to forecast the unit price of exported Coffee in Rwanda.

Volume 7 Issue 11, November 2018

www.ijsr.net

Licensed Under Creative Commons Attribution CC BY

4.2 Specifics Objectives

The specific objectives of this research are:

- 1) To determine an ARIMA model that is fitted for forecasting the unit price of coffee exported from Rwanda.
- 2) The application of this model to predict the future unit price of coffee exported from Rwanda.
- 3) To test the Stationarity of time series data (monthly unit price of coffee export in Rwanda).

5. Research Hypothesis

It hypothesized that:

- 1) Statistical analysis can help to understand better the current situation of the unit price of exported coffee in Rwanda.
- 2) The forecasting model is sufficient to predict the future unit price of coffee exported from Rwanda using the data available from the national institute of statistics of Rwanda.

6. Data and Research Methodology

Box-Jenkins forecasting models are based on statistical concepts and principles and are able to model a wide spectrum of time series behavior. It has a large class of models to choose from and a systematic approach for identifying the correct model form. There are both statistical tests for verifying model validity and statistical measures of forecast uncertainty. In contrast, traditional forecasting models offer a limited number of models relative to the complex behavior of many time series with little in the way of guidelines and statistical tests for verifying the validity of the selected model.

6.1. Model building

The ARIMA methodology is carried out in three stages described by Box and Jenkins (1976), viz. identification, estimation and diagnostic checking. Parameters of tentatively selected ARIMA Model at the identification stage; parameters are estimated at the estimation stage and adequacy of tentatively selected model is tested at the at the diagnostic checking stage. If the model is found to be inadequate, the 3 stages are repeated until satisfactory ARIMA model is selected for the time series under consideration, to end up with a specific formula that replicates the patterns in the series as closely as possible and also produces accurate forecasts. Software packages EViews contain programs for fitting of ARIMA models.

6.1.1. Identification stage

A preliminary Box-Jenkins analysis with a plot of the initial data should be run as the starting point in determining an appropriate model. The input data must be adjusted to form a stationary series; one whose values vary more or less uniformly about a fixed level over time. Apparent trends can be adjusted by having the model apply a technique of "regular differencing," a process of computing the difference between every two successive values, computing a differenced series which has overall trend behavior removed.

The class of ARMA models is quite large, and in practice we must decide which of these models is most appropriate for the data at hand Y_1, Y_2, \dots, Y_n . The correlogram and partial correlogram are two simple diagrams which can help us to make this decision (i.e. to identify the model).

The sample correlogram is a plot against k of the estimated autocorrelations

$$r_k = \frac{\sum_{t=k+1}^n (y_t - \bar{y})(y_{t-k} - \bar{y})}{\sum_{t=1}^n (y_t - \bar{y})^2}$$

The correlogram for MA model and the partial correlogram for an AR model both cut off. As we know, the correlogram for AR model dies down (but does not cut off). It can be shown that the partial correlogram for an MA model dies down as well. Thus, if both diagrams die down, we can conclude that the appropriate model is ARMA.

Table 1: Characteristics of a stationary model

Model	Acf	Pacf
AR(p)	tails off toward zero (exponential decay)	Cuts off to zero after lag p
MA(q)	cuts off to zero after lag q	Tails off toward zero (exponential decay)
ARMA(p, q)	Tails off toward zero	Tails off toward zero

Another guiding principle in model identification is that of parsimony, the total number of parameters in the model should be as small as possible, this will almost certainly produce the best forecasts, and we can obtain more precise (stable) parameter estimates if the number of parameters is small.

6.1.2. Estimation stage

At the identification stage one or more models are tentatively chosen that seems to provide statistically adequate representations of the available data. Then we attempt to obtain precise estimates of parameters of the model by least squares as advocated by Box and Jenkins.

Estimation –stage results ask this question: Have we found a good model?

Characteristics of a good model:

1) **It is parsimonious:** fits the available data adequately without using any unnecessary coefficients.

2) **It is stationary:** Stationarity condition on coefficient

$$AR(p) = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \varepsilon_t$$

If p=0: we have a pure MA model or white noise ARMA (0, q) and always MA and white noise are stationary.

If p=1: AR(1) or ARMA(1,q) stationarity condition is required where the absolute value of ϕ_1 must be less than 1.

If p=2 : AR(2) or ARMA(2,q) stationarity condition are required where absolute value of ϕ_2 must be less than 1,

$$\phi_2 + \phi_1 < 1 \text{ and } \phi_2 - \phi_1 < 1.$$

If p>2, we check this condition $\phi_1 + \phi_2 + \dots + \phi_p < 1$.

3) It is invertible

Invertibility conditions on coefficients:

$$MA(q) = \varepsilon_t - \theta_1 B_1 - \theta_2 B_2 - \dots - \theta_q B_q$$

If $q=0$, we have a pure AR or a white noise. all pure AR or white noise are invertible`

If $q=1$, MA(1) or ARMA(p,1) ,invertibility require that the absolute value of $\theta_1 < 1$

If $q=2$, MA(2) or ARMA(p,2),condition for invertibility are stated as follow: the absolute value of $\theta_2 < 1$,

$$\theta_2 + \theta_1 < 1, \text{ and } \theta_2 - \theta_1 < 1$$

If $q>2$, we check this condition $\theta_1 + \theta_2 + \dots + \theta_q < 1$

4) It has statistically independent residuals

An important assumption is that the random shocks (ε_t) are independent in a process. We test the shocks for independence by constructing an acf using the residuals as input data. If residuals are statistically independent, this is important evidence that it cannot improve the model further by adding more AR or MA terms.

5) It fits the available data satisfactorily (available data sufficiently well at the estimation stage)

Of course no model can fit the data perfectly because there are random-shocks elements present in the data. The analyst must decide in each case if an ARIMA model fit available data well enough to be used for forecasting. Box-Jenkins suggest a minimum of 50 observations.

6) It produces sufficiently accurate forecasts (small forecast errors).

Though a good forecasting model will usually fit the past well, it is more important that it forecast the future satisfactorily. To evaluate a model by this criterion we must monitor its forecast performance.

6.1.3 Diagnostic checking stage

Once a model has been identified and estimated, it is usually taken to the true model and forecast can be obtain accordingly .to protect against disastrous forecasting errors, the least we can do is to check that the fitted model is a satisfactory one. The most commonly used method is to examine the correlogram of the residuals from the fitted model to see if the residuals are a white noise (as they should be if the model is correct). Box et al (1994). Once the appropriate ARIMA model has been fitted, we can examine the goodness of fit by means of plotting the ACF of residuals of the fitted model.

If the ACF and PACF of earlier lags are not in general within $\pm 2/\sqrt{N}$ band around 0 then there is probably left over serial dependence in the residuals or conditional heteroscedasticity.

6.2 Determining the Best Model

To determine the best model of several models of ARIMA can be used several criteria, among others: criteria for Mean Square Error (MSE), Akaike's Information Criterion (AIC)

and Schwartz's Bayesian Criterion (SBC). The best model was chosen that the value of the smallest message (Aswi & Sukarna, 2006).

6.3 Forecasting future values

According to Newborn and Granger (1974)"it is better to sort out the individual model to derive a preferred model that contains the useful features of the original model" Once a model has been created for a time series; EViews can use it to forecast future values beyond the end of the series.

7. Data Analysis

7.1 Introduction

The data analyzed are data on price per kilogram of coffee exported in January of 2010 to December 2017 in Rwanda. Data analyzed using Eviews software. The ARIMA models consist of three parts:

- a) The autoregressive part consists of a linear regression that establishes how past values of price per kilogram of coffee exported are related to future values,
- b) The "Integrated" part refers to how many times we have to take a difference to get a stationary series, and,
- c) The moving average part consists of how past forecast errors are related to future values of price per kilogram of coffee exported.

An ARIMA model was developed using the Box-Jenkins's methodology that will take into account past values and forecast errors to predict future coffee exportation levels. The Box-Jenkins's methodology aids in identifying a forecast model, estimating its parameters, checking the model's performance, and finally using it to forecast. All of these steps are illustrated below as I develop a simple model to forecast coffee exportation price.

7.2 Model Identification

The data series of unit price of coffee export in Rwanda are plotted in Figure below shows that the series raises through time, so its mean may not be stationary. But deciding if the mean is not stationary with visual inspection can be misleading.

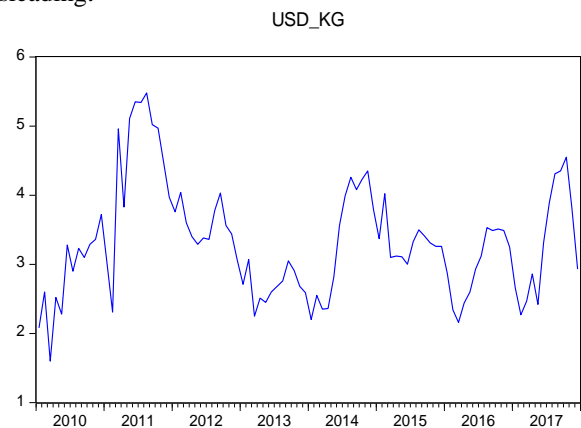


Figure 1: Plot of exported price of coffee in USD per KG against time

Where USD_KG: price value in usd per net weight in kilogram. We need the correlogram of our data series to ensure the stationarity

Date: 06/04/18 Time: 16:35
 Sample: 2010M01 2017M12
 Included observations: 96

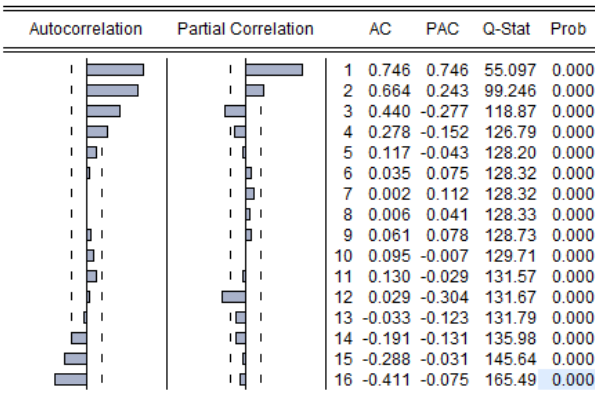


Figure 2: Correlogram of exported price of coffee in USD per KG against time

Q statistics is highly significant; the residuals are not white noise (there is a presence of serial correlation in the residuals). Need to re-specify the model, ACF dies out slowly, indicating that the mean of the data is non-stationary, PACF (the autocorrelation conditional on the in-between values of time series): two spikes. Through this affirmation we have to proceed by checking the stationarity of our data.

7.2.1. Test of stationarity for exported price of coffee in USD per KG

Unit root test will help us to know if our data are stationary or not using Augmented Dickey-Fuller and including intercept, Trend and Intercept or none of these (no intercept, no Trend). If the absolute values of calculated ADF are greater than the absolute values of 5% critical value, this means that there is stationarity. When the critical value is greater than ADF value; so we do not reject the null hypothesis (H0: non-stationarity).

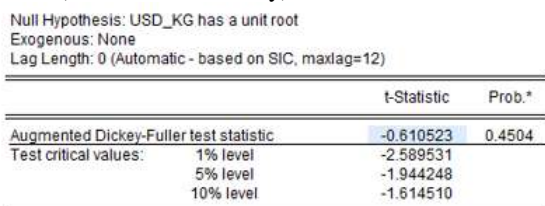


Figure 3: Unit Root Test Results of exported price of coffee in USD per KG

Based on the Figure 3 it shows that value at $\alpha = 5\%$ is -1.944 greater than the value of the statistic t of the ADF Test statistics "i.e. -0.6105 (notice the value used is the absolute value) this indicates that the data is nonstationary.

If we cannot reject the null hypothesis of non-stationarity, we cannot reject that our series is integrated of order d. We have to transform our series into first difference (the level of integration is one). The first difference of our series is performed in order to transform our series into stationarity condition, if it doesn't fulfill the requirement, the second difference will be performed.

7.2.2 Test of stationarity for the first difference of exported price of coffee in USD per KG

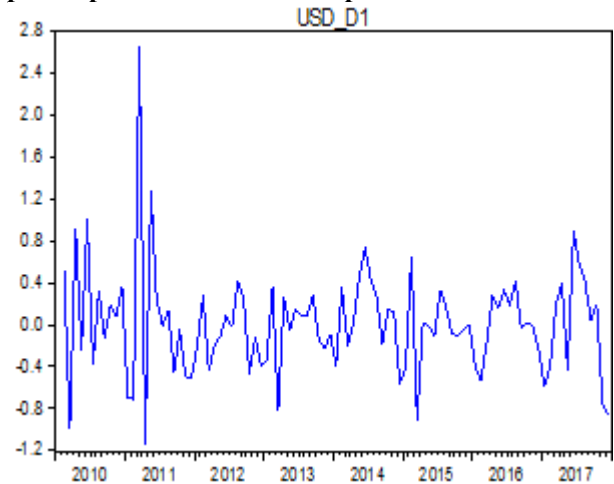


Figure 4: Plot of first difference of exported price of coffee in USD per KG

This plot above seems to be stationary, though the variance does not look constant. This will be confirmed with unit root test of the first difference of exported price of coffee in USD per KG.

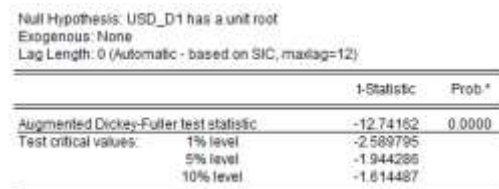


Figure 5: Unit Root Test Results of first difference of price of coffee exported (USD/KG).

Where: *USD_D1* is the first difference of price of coffee exported (USD/KG).

Based on the Figure 5 above, the ADF Test Statistic (-12.74) his absolute value is greater than the absolute value of 5% Critical Value (-1.94); this means that the fist difference of our data series is stationary. As we found that the first difference of our series is stationary, we have to proceed with the identification of tentative models, by using the correlogram of stationary data.

7.2.3. Identification of tentative models

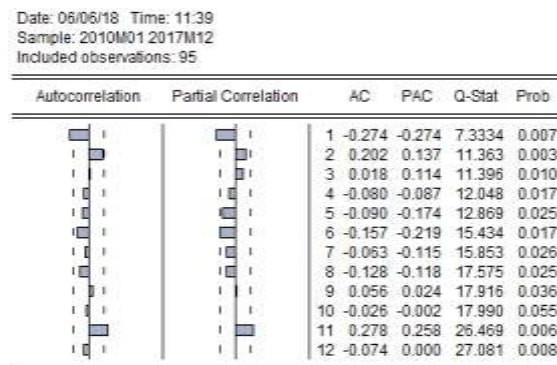


Figure 6: Correlogram of first difference of price per KG of coffee exported.

ACF has a large spike at first lag but oscillating and PACF has a large spike at first lag; this can be ARI(p,d),IMA(d,q) or an ARIMA(p,d,q). Base on the output of correlogram of first difference of our series, we may tentatively choose these models:

ARIMA (1,1,1), ARIMA (1,1,0), ARIMA (0,1,1), ARIMA (2,1,0) and ARIMA(0,1,2). We will have to apply diagnostic tests to find out if the chosen ARIMA model is reasonably accurate.

7.3. Model Estimation

At this stage it will do a test of the significance of the parameters. From Figure 7 below, the probability values obtained AR(1) in the table the Final Estimates of Parameters, namely in the amount of 0.0001. Because the probability value = 0.000 < $\alpha = 0.05$ then parameter AR(1) significant. The obtained probability value is also an MA(1) in the table the Final Estimates of Parameters i.e. amounting to 0.0084. Because the value of the probability = 0.0084 < $\alpha = 0.05$ then parameter MA(1) significant. Since the parameter of AR (1) and MA (1) significant then models ARIMA (1,1,1) can be inserted into a likely model.

Dependent Variable: USD_D1
 Method: Least Squares
 Date: 08/31/18 Time: 14:04
 Sample (adjusted): 2010M03 2017M12
 Included observations: 94 after adjustments
 Convergence achieved after 17 iterations
 MA Backcast: 2010M02

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.005814	0.046431	0.125209	0.9006
AR(1)	-0.706066	0.167362	-4.218801	0.0001
MA(1)	0.552830	0.205163	2.694593	0.0084

R-squared	0.110157	Mean dependent var	0.003579
Adjusted R-squared	0.090600	S.D. dependent var	0.519549
S.E. of regression	0.495455	Akaike info criterion	1.464712
Sum squared resid	22.33824	Schwarz criterion	1.525881
Log likelihood	-65.84148	Hannan-Quinn criter.	1.497499
F-statistic	5.632642	Durbin-Watson stat	2.072326
Prob(F-statistic)	0.004940		

Figure 7: The results of the estimation model of ARIMA (1,1,1)

The below are the results of overfitting some possible models of ARIMA.

Table 2: Overfitting models ARIMA using

Model	Significant	SSE	AIC	SBC
ARIMA(0,1,1)	Yes	0.506766	1.499292	1.553058
ARIMA(1,1,0)	Yes	0.501418	1.478295	1.532407
ARIMA(0,1,2)	Yes	0.507205	1.501026	1.554791
ARIMA(2,1,0)	Yes	0.502842	1.484188	1.538653
ARIMA(1,1,1)	Yes	0.495455	1.464712	1.525881

The next step is done best by doing a model election overfitting. Based on Table 5 shows that the best model is obtained that is ARIMA (1,1,1), due to the significant parameter values and the value of SSE, AIC and the SBC the smaller model from model AR(1) and MA(1).

The estimated equation is:

$$y_t = 0.0058 - 0.7061 y_{t-1} + \varepsilon_t + 0.5528 \varepsilon_{t-1} \quad (\text{see figure 7}).$$

7.4. Diagnostic Checking of Residuals

At this stage of the testing done to see if the selected model is already pretty well statistically. The trick is to test whether the residual estimation results already are white noise. When residual already white noise means the model is just right (Winarno,2011).

Date: 08/30/18 Time: 11:16
 Sample: 2010M03 2017M12
 Included observations: 94
 Q-statistic probabilities adjusted for 2 ARMA term(s)

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	0.151	0.151	2.2188	
		2	0.107	0.086	3.3475	
		3	0.015	-0.014	3.3687	0.066
		4	-0.045	-0.057	3.5754	0.167
		5	-0.039	-0.026	3.7315	0.292
		6	-0.014	0.006	3.7509	0.441
		7	-0.052	-0.045	4.0289	0.545
		8	-0.025	-0.014	4.0957	0.664
		9	0.057	0.071	4.4380	0.728
		10	-0.039	-0.055	4.5991	0.799
		11	-0.012	-0.017	4.6146	0.867
		12	-0.016	-0.008	4.6424	0.914

Figure 8: Correlogram of Residual ARIMA(1,1,1)

On the basis of the Figure 8 it appears that residual already are random. This is shown by the bar graph which are all located in Bartlett's line. From the independence of the residual test results, a model of ARIMA(1,1,1) are qualified white noise.

7.5. Forecasting

After a diagnostic checking the next step is to do forecasting by using models that have been chosen, namely ARIMA(1,1,1).

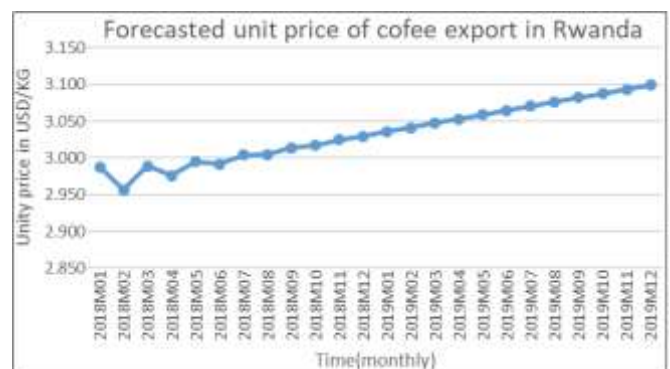


Figure 9: Forecasted unit price of coffee export in Rwanda (Jan2018-Dec2019)

Based on the model estimated the forecasted unit price of coffee export in Rwanda will increase on average of 0.23 percent monthly in the period of Jan2018 to Dec2019 and the average price will be 3.033 USD/KG, In Feb2018, Apr 2018 and Jun2018 the unit price will decrease by 1.04;0.43;0.12percent respectively. A low price will be in in Feb2018(2.957 USD/KG) while the highest will be in Dec2019(3.099 USD/KG).

8. Summary

Export of coffee in Rwanda have been and is still the major agricultural product to be exported, but the price of it remain uncertainty due to the terms of trade and the exchange rate of foreign currency. The aim of this study was to analyze the monthly unit price data obtained from National institute of statistics of Rwanda. The data was collected on Monthly basis from January 2010 to December 2017.

This research used the ARIMA model in modelling unit price of coffee export. Dickey-Fuller(ADF) statistical test was used to test the stationarity assumption. The SSE, AIC and the SBC test were used to select the good model among ARIMA models.

8.1 Conclusion

In business, people are frequently confronted with decision-making situation in which time is an important factor. It would be risky if someone is doing business without projecting it in the future.

Forecasting, which is one technique that business (country) leaders may use as an aide in monitoring present operations and in planning for future needs, plays a crucial role in business, industry, Government and Constitutional planning since any efficacy they do depends on the capacity to anticipate future events

The right to ARIMA Model forecasting unit price of coffee export in Rwanda is model ARIMA(1,1,1). The model ARIMA(1,1,1) is a better model when compared to the other models. This is indicated by the parameters in the model are already significant, the value of the SSE, MSE, AIC and the SBC that model which is smaller in compare to other models and already meets the test of independence.

The analysis done using the data collection from the secondary data of Rwanda Exports is very important since its forecasts the future evolution of the export and it can be applicable to the other institutions/activities economic and financial institutions.

8.2 Recommendation

ARIMA model are most powerful and popular in analyzing time series data. I would recommend many researchers to use it. These suggestions are addressed to planner of exportation in Rwanda in order to improve the evolution of export. It should take into consideration the results found in the paper in order to make a good decision about the strategies that can be made to manage the evolution of Rwandan coffee export. It should conduct some researchers each year in order to control the evolution of Rwandan coffee export. We would like to suggest to the government of Rwanda to keep the same pace of exportation and give more considerations to the export of Coffee.

Suggestion for further research: even though much effort has been put in this research study, there still remain many things to study in the domain of financial institutions using the statistical knowledge which will be studied by other researchers in the same field of the study.

References

- [1] DAMODAR, N. Gujarati, Basic econometrics 4th edition. Tata Mc Graw-Hill. Publishing Company limited, New Delhi, 2005.
- [2] John Wiley & Sons. Forecasting with Univariate Box-Jenkins Models. New York H.H, 1983.
- [3] John H. Cochrane. Time series for Macro economics and Finance. Chicago,1897.
- [4] Gebhard kirchgassner; Juigen Wolters. Introduction to Modern Time Series Analysis. Springer,2007.
- [5] Box, G. E. P., & Jenkins, G. M. Time series analysis: forecasting and control (2nded.). San Francisco: Holden-Day,1976
- [6] Chatfield, C. The Analysis of Time Series: An Introduction, Fifth Edition, Chapman &Hall, Boca Raton, FL.1996.
- [7] Fuller, W.A. Introduction to Statistical Time Series, Second Edition, John Wiley & Sons. New York,1996.
- [8] Gaynor, P.E. and R.C. Kirkpatrick. Introduction to Time-Series Modeling and Forecasting in Business and Economics. McGraw-Hill, New York,1994
- [9] Shumway R.H. and D.S. Stoffer. Time Series Analysis and Its Applications, Springer- Verlag, New York,2000.
- [10] Roman Kozhan. Financial Econometrics-with Eviews. ventus publishing APS,2010.
- [11] Barbara G. Tabachnick and Linda S.Fidel .Using Multivariate Statistics, 6th Edition, California State University-Northridge,2013.
- [12] Goodrich, R. L. "The Forecast Pro methodology", International Journal of Forecasting, 16(4), 533–535,2000.
- [13] Hendry, D. F. "The econometrics of macroeconomic forecasting", The Economic Journal, 107(444), 1330–1357,1997.
- [14] Hylleberg, S., R. Engle, C. Granger and B. Yoo. "Seasonal integration and cointegration", Journal of Econometrics, 44, 215–238,1990.
- [15] Hyndman, R. J. "It's time to move from 'what' to 'why'—comments on the M3-Competition", International Journal of Forecasting, 17(4), 567–570,2001.
- [16] Ord, J. K., A. B. Koehler and R. D. Snyder. "Estimation and prediction for a class of dynamic nonlinear statistical models", Journal of the American Statistical Association, 92,1621–1629,1997.
- [17] Reilly, D. "The AUTOBOX system", International Journal of Forecasting, 16(4), 531–533,2000.
- [18] Smith, J. and S. Yadav. "Forecasting costs incurred from unit differencing fractionally integrated processes", International Journal of Forecasting, 10(4), 507–514,1994.
- [19] Taylor, J. W. "Exponential smoothing with a damped multiplicative trend", International Journal of Forecasting, 19, 715–725,2003