M.C. Alibuhtto[1]

Modelling Fresh Coconuts Export using Time Series Approach

(1) South Eastern University of Sri Lanka, Faculty of Applied Sciences, Department of Mathematical Sciences, Sammanthurai, Sri Lanka. (mcabuhtto@seu.ac.lk)

Abstract: Coconut is one of major plantation crop in Sri Lanka. About 30 percent of the coconut production is processed into coconut products for exports. This paper analyses and modelling fresh coconut exports from Sri Lanka. For this study, secondary data on the exports of fresh coconuts were used from period of 1981 to 2010. Time series trend analysis, unit root test and Box-Jenkins approach were applied. The ARIMA (1, 1, 1) model was identified as the best fitted to forecast the fresh coconuts export from Sri Lanka. The error series of the fitted model was found to be a white noise process.

Keywords: ARIMA, Modelling, Unit root, White noise

Introduction

There are three major plantation crops such as coconut, tea and rubber in Sri Lanka. The extent under coconut is cultivated around 394836 ha (53.15%), tea is cultivated in 222000 ha (29.89%) while rubber is in 126000 ha (16.96%). The contribution of the total plantation sector to GDP is 2.5% while the contribution of coconut plantation is 1.1% (Central bank report, 2010).

Coconuts are cultivated in all parts of the country although its importance varies from region to region. The major coconut growing areas are Western, North Western and Southern provinces of Sri Lanka.

Sri Lanka had strong economic growth rates in the recent years compared to the South Asian region. Apparel, tea rubber and coconut export contribute more to the Sri Lankan economy, while export contributions in agricultural products show a dramatic increase in the export market. These exports contribute highly in foreign exchange.

Scope of this study was to find an appropriate forecasting model for exports of fresh coconuts and to support to formulate foreign earning by national coconut exports in the country.

Some of authors studied and analyzed coconut production and price of coconuts (Rangoda, Fernando, 2006) but so far none of authors were not interested to predict the time series model to export of fresh coconuts from Sri Lanka in one year advance.

Methodology

The secondary data on exports of fresh coconuts from 1981 to 2010 were considered for the analysis and it was extracted from bulletins of Coconut Development Authority (CDA).

Time Series plots and basic descriptive analysis were used to find the behavioral patterns of the data. Then trend analysis was applied to predict suitable model for coconut export. When the preliminary information extracted from the time series plot, the data were checked for the stationary. If the data series shows non-stationary pattern, the first order differencing method was used as the first attempt to convert it into a stationary series.

Time series approach is very suitable for describing this kind of stochastic process and also easy to establish the forecasting model to forecast export of fresh coconuts. Therefore Box-Jenkins (1976) ARIMA modeling approach was used in developing forecasting model.

Auto Regressive Integrated Moving Average (ARIMA) is the most general class of models for forecasting a time series. Different series appearing in the forecasting equations are called "Auto-Regressive" process. Appearance of lags of the forecast errors in the model is called "moving average" process. The ARIMA model is denoted by ARIMA (p,d,q), where "p" stands for the order of the auto regressive process, 'd' is the order of the data stationary and 'q' is the order of the moving average process. The general form of the ARIMA (p,d,q) can be written as described by Judge, et al. (1988).

$$\Delta^d y_t = \delta + \beta_1 \Delta^d y_{t-1} + \ldots + \beta_p \Delta^d y_{t-p} + e_t - \theta_1 e_{t-1} - \ldots\ldots - \theta_q e_{t-q} \qquad (1)$$

Where, Δd denotes differencing of order d, i.e., $\Delta y_t = y_t - y_{t-1}$, $\Delta^2 y_t = y_t - 2y_{t-1} + y_{t-2}$ and so forth, $y_{t-1}, \ldots\ldots, y_{t-p}$ are past observations are $\delta, \beta_1, \ldots\ldots, \beta_p$ parameters (constant and coefficient) to be estimated similar to regression coefficients of the Auto Regressive process of order "p" [AR(p)], where, et is forecast error, assumed to be independently distributed and $e_t, e_{t-1}, \ldots\ldots, e_{t-q}$ are past forecast errors, $\theta_1, \ldots\ldots, \theta_q$ are moving average (MA) coefficient that needs to be estimated.

The major problem in ARIMA modeling technique is to choose the most appropriate values for the p, d, and q. This problem can be partially resolved by looking at the Auto Correlation Function (ACF) and Partial Auto Correlation Functions (PACF) for the series (Rubinfeld, 1991). The degree of the homogeneity, (d) i.e the number of time series to be differenced to yield a stationary series was determined on the basis where the ACF approached zero.

The following accuracy measures were used to find an appropriate trend model to forecast exports of fresh coconuts.

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{(y_t - \hat{y}_t)}{y_t}\right| * 100 \qquad (2)$$

$$MAD = \frac{1}{n}\sum_{t=1}^{n}\left|y_t - \hat{y}_t\right| \qquad (3)$$

$$MSD = \frac{1}{n}\sum_{t=1}^{n}(y_t - \hat{y}_t)^2 \qquad (4)$$

Where:

y_t = *the actual value*

$\hat{y}_t$ = *the fitted value*

n = *number of observations*

Also, the AIC and BIC (SIC) can be used to find suitable forecasting ARIMA model for coconut exports.

$$AIC = \log\left(\frac{rss}{n}\right) + \left(\log(n) * \frac{k}{n}\right) \qquad (5)$$

$$BIC = \log\left(\frac{rss}{n}\right) + \left(2 * \frac{k}{n}\right) \qquad (6)$$

where;

k = number of coefficient estimated,
rss = residual sum of square,
n = number of observations

Results and Discussions

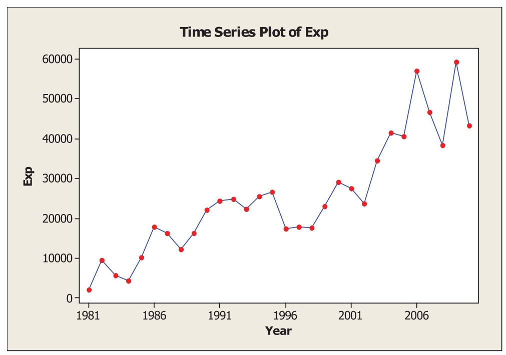


Figure 1: Time series plot of fresh coconuts export [1981-2010]

From the figure 1, it can easily be seen that coconuts exports have been increasing over time and variance is increasing with time. Thus it is obvious the series is not stationary.

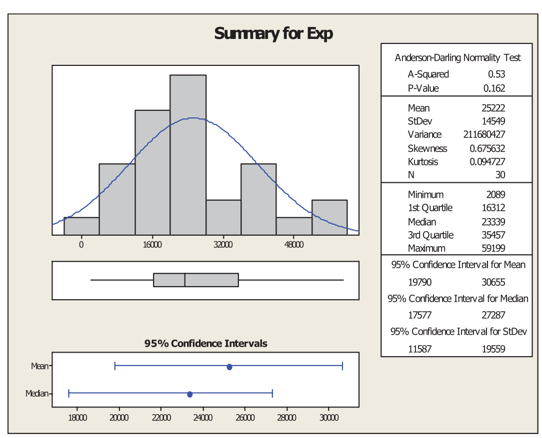


Figure 2: Summary statistics for export of fresh coconuts.

From the figure 2, the mean of coconut export is 25222 million nuts and its standard deviation is 14549 million nuts. According to the Anderson Darling (AD) statistic test, the export of coconuts were normally distributed at 5% significance level, (p=0.162). The average exports of coconuts lie between (19790, 30655 in million) at 5% significance level.

Due to the high standard deviation of original series, it was log transformed and fitted trend models. The following three models were fitted to estimate export of coconuts. The results were shown in table 1.

Table 1: Trend analysis for coconut export data

Model	MAPE	MAD	MSD	DW
$\log(y_t) = 8.79 + 0.07 * t$	3.05	0.29	0.14	1.09
$\log(y_t) = 8.39 + 0.45 * t - 0.002 * t^2$	2.89	0.27	0.12	1.28
$\log(y_t) = 8.80 * (1.01)^t$	3.13	0.29	0.15	1.36

Table 1 indicates that, quadratic trend model is better model to compare with linear and exponential models as it has less MAPE, MAD and MSD values. But this models residuals are positively correlated (DW=1.28). Plotting the residual of the estimated quadratic model is shown in figure 3 also it indicates that the residual are correlated and the model is not correctly specified.

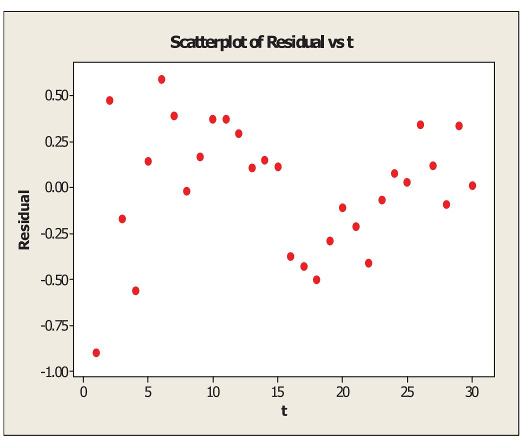


Figure 3: Residuals plot of quadratic model

The results of the ADF test for coconuts export shown in table 2 confirm that series is not stationary.

Table 2: Results of the unit root test for coconut export

ADF Test Statistic		t-Statistics	Prob.
		-0.168	**0.930**
Test Critical Values	1% Level	-3.738	
	5% Level	-2.992	

Table 2 indicates that the null hypothesis that the series in levels contain unit root could not be rejected for coconut export. [p=0.930]. Therefore, coconut export series is non-stationary. Then, the log transformed series was tested using ADF test and shown in table 3.

Table 3: Results of the unit root test for log transformed coconut export data.

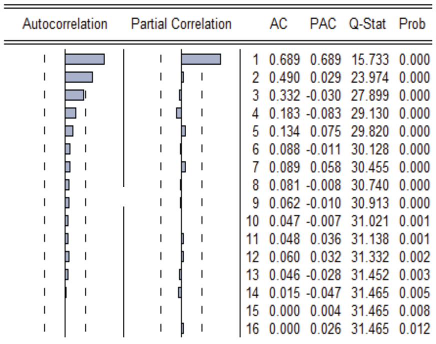
ADF Test Statistic		t-Statistics	Prob.
		1.153	**0.932**
Test Critical Values	1% Level	-2.647	
	5% Level	-1.610	

Table 3 indicates that the null hypothesis that the log series in levels contain unit root could not be rejected for coconut export. [p=0.932]. Therefore, log of coconut export series is non-stationary. Thus, the log of coconut export series needs to be differenced to obtain a stationary series. The process is continued until stationary series to be found. Thus the results of ADF test for the first differenced series is shown in Table 4.

Table 4: Results of the unit root test for first differenced of log transformed coconut export data.

ADF Test Statistic		t-Statistics	Prob.
		-8.001	**0.000**
Test Critical Values	1% Level	-2.650	
	5% Level	-1.953	

Table 4 indicates that the null hypothesis is rejected for the first differences of log transformed export coconut series. [p=0.000]. Therefore log of export coconut series is stationary at its first difference.



Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.689	0.689	15.733	0.000
		2	0.490	0.029	23.974	0.000
		3	0.332	-0.030	27.899	0.000
		4	0.183	-0.083	29.130	0.000
		5	0.134	0.075	29.820	0.000
		6	0.088	-0.011	30.128	0.000
		7	0.089	0.058	30.455	0.000
		8	0.081	-0.008	30.740	0.000
		9	0.062	-0.010	30.913	0.000
		10	0.047	-0.007	31.021	0.001
		11	0.048	0.036	31.138	0.001
		12	0.060	0.032	31.332	0.002
		13	0.046	-0.028	31.452	0.003
		14	0.015	-0.047	31.465	0.005
		15	0.000	0.004	31.465	0.008
		16	0.000	0.026	31.465	0.012

Figure 4: Sample ACF and PACF of the first order difference series log of coconut export.

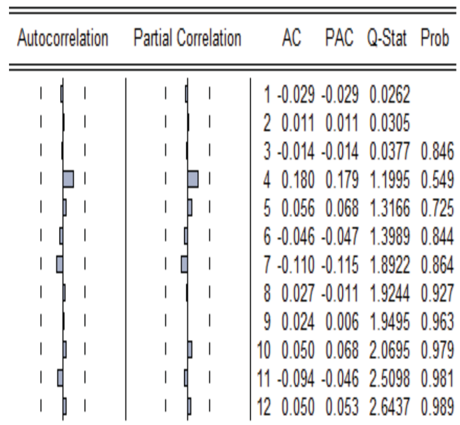
According to the figure 4, the sample ACF (SACF) has two significant autocorrelations at lag 1 and lag2. The sample PACF (SPACF) has one significant autocorrelation at lag 1. Thus it can be hypothesized in the ARMA model to be fitted AR order to be 1 and MA order to be less than or equal to 2. Based on the above, the following models were considered as possible models to represent the original series. They are: (i) ARIMA(1,1,1), (ii) ARIMA(1,1,0), (iii) ARIMA(1,1,2) and (iv) ARIMA(0,1,2)

Table 5: ARIMA models for D (LExpCoc)

Models	Parameter Estimates	P-Value	AIC , SIC	Log likeli hood	DW
ARIM A (1,1,1)	C=10.098 AR(1)= 0.545 MA(1)=0.780	0.000 0.000 0.000	0.542, 0.684	-4.861	2.06
ARIM A (1,1,0)	C=10.25 AR(1)= 0.669	0.000 0.000	0.772, 0.866	-9.19	1.83
ARIM A (1,1,2)	C=10.652 AR(1)=0.901 MA(1)= -0.143 MA(2)=-0.852	0.000 0.000 0.310 0.000	0.718, 0.807	-6.06	2.16
ARIM A (0,1,2)	C=10.028 MA(1)= 1.098 MA(2)= 0.283	0.000 0.000 0.004	0.711, 0.851	-7.669	1.22

Table 5 indicates that the coefficients of all AR (1), MA (1) [except ARIMA (1,1,2) model] and MA (2) items in the four models are significant at 5% significance level (p-value < 0.05). Hence, these four models are can be considered when selecting the best fitted model from the point of view of the parameter significance. Results in table 5 also indicate that of the four models the maximum log likelihood estimate and the lowest AIC and SIC values were obtained by ARIMA (1, 1, 1) model. Thus it can be concluded the best model out of the four is ARIMA (1,1,1).

The ACF plot of the residuals of ARIMA (1,1,1) model (Figure 5) shows that the residuals of the ACF are relatively small and not statistically significant. Therefore, it can be considered that the residuals of the fitted model are randomly distributed.



Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.029	-0.029	0.0262	
		2	0.011	0.011	0.0305	
		3	-0.014	-0.014	0.0377	0.846
		4	0.180	0.179	1.1995	0.549
		5	0.056	0.068	1.3166	0.725
		6	-0.046	-0.047	1.3989	0.844
		7	-0.110	-0.115	1.8922	0.864
		8	0.027	-0.011	1.9244	0.927
		9	0.024	0.006	1.9495	0.963
		10	0.050	0.068	2.0695	0.979
		11	-0.094	-0.046	2.5098	0.981
		12	0.050	0.053	2.6437	0.989

Figure 5: Plot of ACF of residuals of ARIMA (1, 1, 1)

In order to check the normality of the error series, the probability plot of the residuals was carried out (Figure 6).

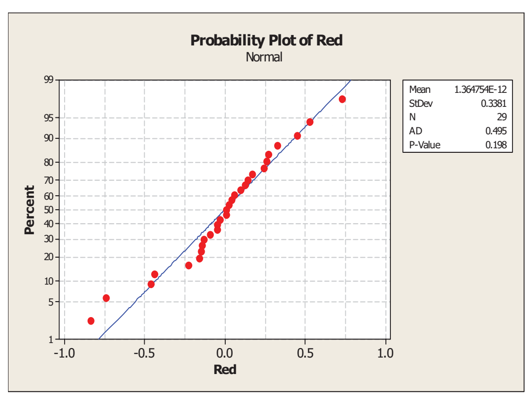


Figure 6: Normal probability plot for the residuals

Figure 6 shows that the respective Anderson-Darling statistic (AD=0.495, p=0.198) and thus it is confirmed the residual series is normally distributed.

Based on the above detailed analysis of residuals, it can be confirmed that the fitted ARIMA (1, 1, 1) model satisfies all the diagnostic tests. Hence, the ARIMA (1,1,1) model is considered as the best fitted model for the log transformed coconuts export data.

Thus, the model equation can be formed as:

$$Y_t = 10.098 + Y_{t-1} + 0.545\,(Y_{t-1} - Y_{t-2}) - 0.78 * e_{t-1} \quad (7)$$

The forecast accuracy measures of observed data were shown in figure 7.

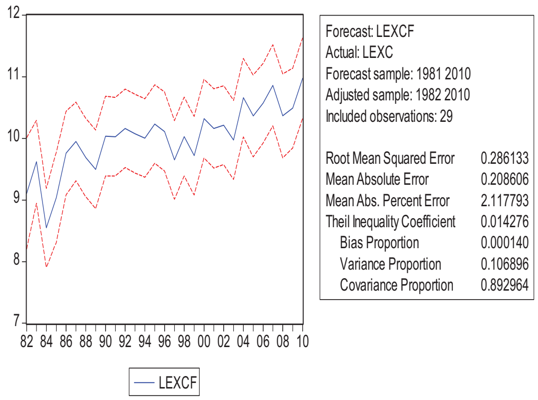


Figure 7: Forecast accuracy measures

Figure7 indicates that the RMSE and MAPE for ARIMA (1, 1, 1) model deviates from the observed data are 0.286 % and 2.118 % respectively, which is definitely be regarded as within the acceptable range. Forecasting values from 2011 to 2020 were estimated using fitted ARIMA (1,1,1) model for fresh coconuts export and it was tabulated in table 6.

Table 6: Forecasting export of coconuts from 2011 to 2020.

Year	Forecast values
2011	47449.6
2012	49763.9
2013	51499.7
2014	53061.8
2015	54571.8
2016	56066.1
2017	57555.7
2018	59044.0
2019	60531.7
2020	62019.4

Conclusions and Recommendations

This study aimed to modeling export of fresh coconuts from Sri Lanka using ARIMA model. The time series data is not stationary at level. By applying the ADF test for the series of the first order differences of log transformed series becomes stationary. Then various ARIMA models were estimated by using Box-Jenkins approach. The comparative performance of these ARIMA models have checked and verified by using the accuracy statistics (AIC and SIC). The comparison indicates that the ARIMA (1,1,1) model as the best model and performs much better than the rest of the estimated models. The residuals of ARIMA (1,1,1) models was found to be white noise.

References

Dickey and Fuller (1981), "Test for Autoregressive Time-Series with a Unit Root", Econometrica 49, pp. 1057-1072.

Rangoda, Abeywickrama, and Fernando, (2005), "An Analysis of different forecasting models for prices of coconut products in SriLanka" ,3rd Academic Session

Sivarajah P, (2011), "Sri Lanka's Agricultural Exports and International Market Integration: Case of Coconut Products Exports", 5th Annual Research Forum, Sri Lanka Agricultural Economics Association (SAEA).

Weiss E, (2000), "Forecasting commodity price using ARIMA", Technical Analysis of stocks & Commodities, vol. 18, no.1.