

Forecasting Model for Para Rubber's Export Sales

Tatiporn Pattranurakyothin and Kanchana Kumnungkit *

Department of Mathematics, Faculty of Science, King Mongkut's Institute of Technology
Ladkrabang, Bangkok, Thailand

Abstract

In this paper, monthly export values of para rubber are investigated using the Box-Jenkins method. To find the optimal predicting model, 12-year data, from January 2000 to December 2011, are used to analyze. Finally, the suitable mathematical model is seasonal ARIMA that use the analysis of time-series from the lowest level of the Mean Absolute Percent Error (MAPE). The best model is seasonal ARIMA $(1,1,1)(1,1,0)_{12}$.

Keywords: Mathematical model, Box-Jenkins Method, Exporting of Para rubber

1. Introduction

Thailand is a top country in the world exporting para rubber. Currently, Thailand economy is affected by the huge disaster in Japan and the slow-down world economy. Recently, the most severe flood in 70 years is also occurred in Thailand which enormously affects to an industrial and agricultural part of Thailand. Rubber trees are taken place in almost every part of Thailand such as north, east, north-east and especially south which can be easily affected by a natural disaster. Moreover, para rubber is an important material for many companies to produce rubber product. Therefore, para rubber plays a crucial part in provincial and national economy. From previous condition, circumstances of para-rubber value always vary. However, by considering domestic and national export values, Thailand's - exporting para rubber is still tended to be an important export to the world. Hence, it is interesting to analyze the moving characteristic and para-rubber export values and to find the best mathematical model to predict. The model can be used to stabilize or increase para-rubber export values which helps the investor to plan for some related investments.

There are many methods for forecasting, i.e. the Box-Jenkins method, smooth curve fitting method, exponential floating method or multiple regressions etc. The Box-Jenkins method has been used in this research because it's easy and suitable to use with time series data [3, 6, 7, 10] for forecasting. This paper attempts to identify time series forecast model of monthly export values of para rubber. The main objectives of this study to find the model 12 year data, from January 2000 to December 2011. The past observations of the same time series data are collected and analyzed to develop a model describing the underlying relationship. One of the most widely and important methods need time series model which is the autoregressive integrated moving average (ARIMA) model [1-2]. NoriZam Mohamed and his group research did serveral forecasting methods including multiple linear regression and nonlinear multivariable regression model [3-5]. Firstly, a time series need to be stationary [2, 6-7] and modeling process should take place on the

*Corresponding author: E-mail: kkkancha@kmitl.ac.th

next step. The notation, $\text{ARIMA}(p,d,q)(P,D,Q)_s$ is a form for analyzing of time series. The nonseasonal part of the model is (p,d,q) , the seasonal part of the model is (P,D,Q) and s is the seasonal length with abbreviated as seasonal ARIMA [7]. The specification of differencing orders (d, D) built for a seasonal ARIMA model. The method of differencing is one way to change nonstationary of seasonal to stationary [8].

In methodology section, we present the Box-Jenkins, ARIMA, seasonal ARIMA model, and then discuss the results of the model. Finally, our conclusions is shown based on the selected forecasting evaluation methods.

2. Methodology

There are many different forecasting approaches that are available to forecast time series data. Here, an approach is the Box-Jenkins autoregressive integrated moving average model (ARIMA) and seasonal autoregressive integrated moving average model (SARIMA).

2.1 Box- jenskins ARIMA Model:

The ARIMA linear models have dominated in many parts of time series forecasting for more than half a century. ARIMA is usually possible to find a process which provides an adequate description to our data. A nonseasonal time series, $\text{ARIMA}(p,d,q)$, can generally be modeled as a combination of past values and past errors. We can be written as follows [9]:

$$\phi_p(B)(1-B)^d Z_t = \theta_q(B)\varepsilon_t \quad (1)$$

with

$$\begin{aligned} \phi_p(B) &= 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p \\ \theta_q(B) &= 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q \end{aligned}$$

where Z_t denotes appropriately transformed in period t . $(1-B)^d$ denotes the nonseasonal differencing operator, B denotes the backward shift operator and ε_t denotes the purely random process.

2.2 Box- jenskins SARIMA Model:

In practice, many time series data contain a seasonal periodic component, which repeat every observation. To deal with seasonality, the ARIMA model is extended to a general multiplicative seasonal ARIMA $(p,d,q)(P,D,Q)_s$ model. A seasonal ARIMA $(p,d,q)(P,D,Q)_s$ model denotes SARIMA which is defined as follows [10]:

$$\varphi_p(B)\Phi_p(B^s)(1-B)^d(1-B^s)^D Z_t = \theta_q(B)\Theta_q(B^s)\varepsilon_t \quad (2)$$

with

$$\begin{aligned} \varphi_p(B) &= 1 - \varphi_1 B - \varphi_2 B^2 - \dots - \varphi_p B^p \\ \Phi_p(B^s) &= 1 - \Phi_1 B^s - \Phi_2 B^{2s} - \dots - \Phi_p B^{ps} \\ \theta_q(B) &= 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q \\ \Theta_q(B^s) &= 1 - \Theta_1 B^s - \Theta_2 B^{2s} - \dots - \Theta_q B^{qs} \end{aligned}$$

where Z_t is appropriately transformed in period t while $(1 - B)^d$ and $(1 - B^s)^D$ are the nonseasonal and seasonal different operators, respectively. B denotes the backward shift operator and ε_t denotes the purely random process. If the integer D is not zero, then the seasonal differencing is involved. The above model is called a SARIMA model or seasonal ARIMA model $(p, d, q)(P, D, Q)$. If d is non-zero, then there is a simple differencing to remove trend. The seasonal differencing, $(1 - B^s)^D$, may be used to remove seasonality. Basically, d and D values are usually zero or one but rarely being two.

2.3 Box-Jenkins ARIMA Modeling Step:

The modeling steps of Box-Jenkins ARIMA Model involves an iterative five-stage process as follows:

- i) Preparation of data including transformations and differencing,
- ii) Identification of the potential models by looking at the sample autocorrelations and the partial autocorrelations,
- iii) Estimation of the unknown parameters,
- iv) Checking the adequacy of fitted model by performing normal probability plot, and
- v) Forecast future outcomes based on the known data.

2.4 Forecasting Evaluation Method:

For the purpose of evaluating out of sample forecasting, different evaluation statistics such as the root mean square error (RMSE), the mean absolute error (MAE), the mean square error (MSE) and the mean absolute percentage error (MAPE) are considered [10]. In this study, MAPE is used. MAPE is the most widely used error measure in the forecasting literature. MAPE is an overall measure of forecast accuracy, computed from the absolute differences between a series of forecast and actual data points observed. This measure is commonly used in the forecasting literatures [10-11]. The formula of RMSE, MAE, and MAPE are defined as follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (z_i - \hat{z}_i)^2}{n}}, \quad (3)$$

$$MAE = \frac{\sum_{i=1}^n |z_i - \hat{z}_i|}{n}, \quad (4)$$

And

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{Z_i - \hat{Z}_i}{Z_i} \right| \times 100 \quad (5)$$

where Z_i denotes the actual value, \hat{Z}_i denotes the predicted value and n is the number of the predicted values.

3. Results

The observed data in this study obtained from the Bank of Thailand dueing January 2000 to December 2011 [12]. Following the rule of Box Jenkins ARIMA modeling steps, all data were arrange to stationality by the first order regular differencing and the difference of seasonality at lag 12 were performed. SPSS software is used to formulate seasonal ARIMA model. Plotting our the output after 1st differencing, it is clear to a stationary series ($D=1, s=12$). Seasonal ARIMA (1,1,1)(1,1,0)₁₂ is the best model for this data because of showing values in Table 1. All the parameters of this model are significant at 0.005 significant level and the equation is presented as follows :

$$\begin{aligned} Z_t &= y_t - y_{t-1} - y_{t-12} \\ &= 24.028 + 0.838(y_{t-1} - y_{t-2}) + 0.431(y_{t-12} + y_{t-24}) + 0.927\varepsilon_{t-1} + \varepsilon_t \end{aligned}$$

Table 1 reports the seasonal ARIMA (1,1,1)(1,1,0)₁₂ where RMSE, the root mean squared error, equals to 3074.043, MAPE, the mean absolute percentage, equals to 14.065%. The coefficient of determination is 8.57 and r (correlation coefficient value) equals to 0.93. The comparision between real data dueing January 2000 to September 2011 and the forecasting data of the exporting rubber price (million baht) are shown in Figure 1

Table 1. Details of R-squared, RMSE and MAPE of ARIMA model

Model	Number of Predictors	Model Fit statistics				Ljung-Box Q(18)			Number of Outliers
		Stationary R-squared	R-squared	RMSE	MAPE	Statistics	DF	Sig.	
rubber-Model_1	0	.166	.857	3074.043	14.065	32.842	15	.005	0

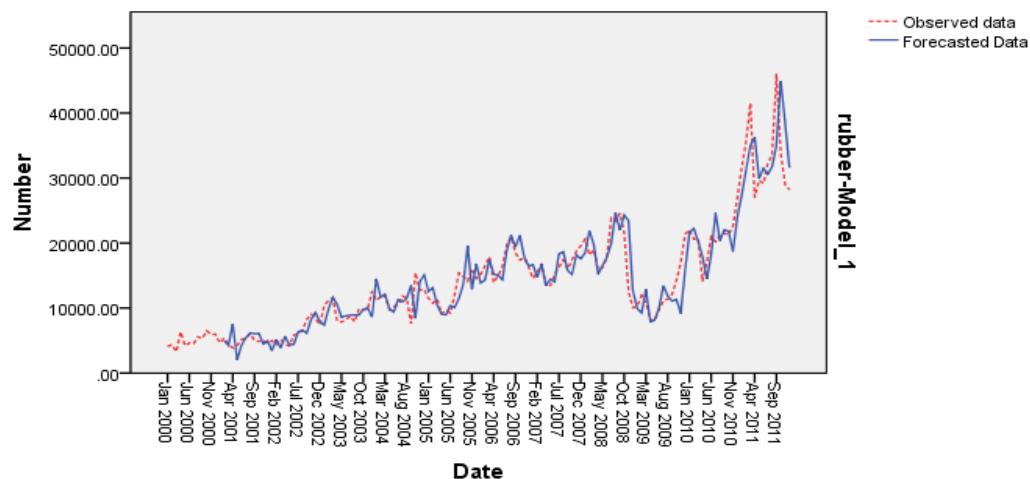


Figure 1 The comparision between real (observed) data dueing January 2000 to September 2011 and the forecasting data of the exporting rubber price (million baht)

4. Conclusions and Suggestion

The monthly export values of para rubber are investigated using the Box-Jenkins method to find the best model. Twelve year data, dueing January 2000 to September 2011 from Economic and Financial Statistics, the Bank of Thailand (BOT), Thailand [12] are used in this study. Finally, the suitable mathematical model is seasonal ARIMA that use the analysis of time-series from the lowest level of the MAPE. The suitable forcasting is seasonal ARIMA $(1,1,1)(1,1,0)_{12}$. The best model is obtained

$$\begin{aligned} Z_t &= y_t - y_{t-1} - y_{t-12} \\ &= 24.028 + 0.838(y_{t-1} - y_{t-2}) + 0.431(y_{t-12} + y_{t-24}) + 0.927\varepsilon_{t-1} + \varepsilon_t \end{aligned}$$

where coefficient of determination is 0.857 and 0.93 is the correlation coefficient.

Since 12 year monthly data are analyzed, the best model could fit or find out. For other forecasting and researchers, if the best model cannot find by monthly data, you have to manage to rearrange to a quarter, six months or annual time series data.

References

- [1] Brockwell, P.J. and Davis, R.A., **1996**. *Introduction to Time Serie and Forecasting*, Springer, New York.
- [2] Chatfield, C., **2004**. *The Analysis of Time Series: An Introduction*, 6th Ed, Chapman & Hall, New York.
- [3] Tsekouras, G.J., Dialynas, E.N., Hatzigergiou, N.D. and Kavatza, S., **2007**. A non-linear multivariable regression model for midterm energy of power systems. *Electric Power Systems Res.*, 77: 1560-1568.
- [4] Ghiassi, M., Zimbra, D.K. and Saidane, H., **2006**. Medium term system load forecasting with a dynamic artificial neural network model. *Electric Power System Res.*, 76: 302-316.
- [5] Mohamed, N., Ahmad, M. H., Suhartono and Ismail, Z., **2011**. Improving Short Term Load Forecasting Using Double Seasonal Arima Model., Indonesia. *World Applied Sciences Journal* 15(2): 223-231, 2011.
- [6] Zhang, G. P. and Qi, M., **2005**. Network Forecasting for Seasonal and Trend Time Series, *European Journal of Operation Research* 160, 501-514.
- [7] Mohamed, N., Ahmad, M. H., Ismail, Z. and Arshad, K. A., **2008**. A multilayer feedforward neural network model and Box-Jenkins model for seasonal load forecasting. Ultra Scientist of Physical Scientist, *International Journal of Physical Sciences*. 20(3): 767-772.
- [8] Makridakis, S., Wheelwright, S. C. and Hyndman, R. J., **1998**. *Forecasting: Methods and Application*. John Willey and Sons, New York.
- [9] Wei, W.W.S., **2006**. *Time Series Analysis, Univariate and Multivariate Methods*, 2nd Ed. NewYork: Pearson, Addison Wesley.
- [10] Box, G. E. P., Jenkins, G. M. and Reinsel, G. C., **2008**. *Time Series Analysis: Forecasting and Control*, 4th Ed. New Jersey: John Wiley & Sons.
- [11] Hobbs, B.F., Helman, U., Jipraikusarn, S., konda, S. and Maratukulan, D., **1998**. Artificial neural networks for short-term energy forecasting : Accuracy and economic value. *Neurocomputing*, 23 :71-84.
- [12] ศัชนีสินศักดิ์อกจำแนกตามกิจกรรมการผลิต (ส้านบท). **2010**. Bank of Thailand. At: <http://www.bot.or.th/Thai/Statistics/EconomicAndFinancial/EconomicIndices/Pages/Index.aspx>