Proceedings of the 2nd IMT-GT Regional Conference on Mathematics, Statistics and Applications
Universiti Sains Malaysia, Penang, June 13-15, 2006

ARIMA MODEL FOR FORECASTING OIL PALM PRICE

Rangsan Nochai[1], Titida Nochai[2]
[1]Department of Agribusiness Administration, Faculty of Agricultural Technology, King Mongkut's Institute of Technology Ladkrabang,
Ladkrabang, Bangkok, 10520
Thailand
knrangsa@kmitl.ac.th

[2]Department of Applied Statistics,
Assumption University
Huamark, Bangkok, 10240
Thailand
titida@s-t.au.ac.th

Abstract. This research is a study model of forecasting oil palm price of Thailand in three types as farm price, wholesale price and pure oil price for the period of five years, 2000 – 2004. The objective of the research is to find an appropriate ARIMA Model for forecasting in three types of oil palm price by considering the minimum of mean absolute percentage error (MAPE). The results of forecasting were as follows:

ARIMA Model for forecasting farm price of oil palm is ARIMA (2,1,0), ARIMA Model for forecasting wholesale price of oil palm is ARIMA (1,0,1) or ARMA(1,1), and ARIMA Model for forecasting pure oil price of oil palm is ARIMA (3,0,0) or AR(3) .

2000 Mathematics Subject Classification: 30C45, Secondary 30C80

Key words and phrases: ARIMA Model, Box-Jenkins Methodology, Oil-Palm Price, MAPE

1. Introduction

Nowadays, the increased oil prices worldwide are having a great impact on all economic activities. Like many countries, Thailand depends heavily on oil as a source of energy. Thai government by Prime Minister Thaksin Shinawatra called Thai people to help save energy by limiting their driving speed to 90 kilometers an hour, turn off air conditioners when not necessary, and include the use of alternative energy as the production of bio-diesel and ethanol. In the cabinet meeting on 8 March 2005, acknowledged a proposal by the Ministry of Energy has comprised six measures to cope with the oil crisis. The one of those measures is to involve the promotion of the production and use of bio-diesel. The Ministry of Agriculture and Cooperatives is helping develop oil palm cultivation for use to produce bio-diesel on a full-cycle basis. The special purpose vehicle (SPV), a new financial mechanism to support the agricultural business, offers to help manage and expand the bio-diesel market [**3**].

The Government has campaigned not only for energy saving, but also for the use of alternative sources of fuel. An action plan to help develop and promote bio-diesel won Cabinet approval in May 2005. In response to the plan, the Ministry of Agriculture and Cooperatives will have to complete its zoning for oil palm cultivation within the next six months. The eastern and southern regions will be developed as the bases for oil palm planting. The Cabinet also approved a budget of 1.3 billion baht as a revolving fund to produce oil palm. Out of this fund, 800 million baht will be used to promote the growing of oil palm, while the remaining 500 million baht will be spent on research and development and management [**3**].

The Ministry of Industry and the Ministry of Energy were instructed to take charge of producing bio-diesel and promoting the use of this alternative source. Construction of three bio-diesel factories, each with a production capacity of 100,000 liters a day, will begin this year. These factories will be able to sell bio-diesel from 2007 onwards [**3**].

From the government strategy, the government have the necessary to know the price in the future before to guarantee the minimum price. The department of agricultural economics is the government department which concerns forecasting the agricultural production and price in Thailand. This department has also the important mission to public the important agricultural data and purpose the executive for making the needed policy and planning of Thai government. The forecasting technique is used to be Regression Analysis or Moving Averages [**1**]. For Moving Averages can work better with stationary data. For a time series that contains a trend or seasonal or non-stationary data, the forecasting technique that should be considered is Auto Regressive Integrated Moving (ARIMA). ARIMA models have been already applied to forecast commodity prices [**4, 6**], such as oil [**5**]. And also there are many researchers interested to study those agricultural information which is useful to government and agriculturists, Areepong [**1**] purposed of methods to forecast the product and farm price of 5 vegetables and fruits: young corn, sweet corn, mangoes, coconuts, and pineapples which classified to 20 models by using forecasting techniques of regression analysis, classical decomposition model, exponential smoothing method, Box-Jenkins method. The minimum mean absolute percentage errors (MAPEs) of forecasting values were used in selecting an adequate model. This research used the secondary data during 1982-1998, and found that forecasting method by Box-Jenkins are almost suitable method forecasting model.

This paper focuses on a study model of forecasting oil palm price of Thailand in three types as farm price, wholesale price and pure oil price by using ARIMA models. That is, this paper provides ARIMA models to forecast the oil palm price.

2. Non-seasonal Box-Jenkins Models for a Stationary Series

The Box-Jenkins methodology refers to the set of procedures for identifying, fitting, and checking ARIMA models with time series data. Forecasts follow directly from the form of the fitted model [**4**].

(1) **A pth-order autoregressive model: AR(p),** which has the general form

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + \varepsilon_t \tag{2.1}$$

where

Y_t = Response (dependent) variable at time t
$Y_{t-1}, Y_{t-2}, \cdots, Y_{t-p}$ = Response variable at time lags $t-1$, $t-2, \ldots, t-p$, respectively.
$\phi_0, \phi_1, \phi_2, \cdots, \phi_p$ = Coefficients to be estimated
ε_t = Error term at time t

(2) **A qth-order moving average model: MA(q),** which has the general form

$$Y_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q} \tag{2.2}$$

where
Y_t = Response (dependent) variable at time t
μ = Constant mean of the process
$\theta_1, \theta_2, \cdots, \theta_q$ = Coefficients to be estimated
ε_t = Error term at time t
$\varepsilon_{t-1}, \varepsilon_{t-2}, \cdots, \varepsilon_{t-q}$ = Errors in previous time periods that are incorporated in the response Y_t

(3) **Autoregressive Moving Average Model: ARMA(p,q),** which has the general form

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + \varepsilon_t$$
$$-\theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q}$$

We can use the graph of the sample autocorrelation function (ACF) and the sample partial autocorrelation function (PACF) to determine the model which processes can be summarized as follows:

Model	ACF	PACF
AR(p)	Dies down	Cut off after lag q
MA(q)	Cut off after lag p	Dies down
ARMA (p,q)	Dies down	Dies down

TABLE 1. How to determine the model by using ACF and PACF patterns

3. The steps in the ARIMA model-building

STEP 1: Model Identification (Selecting an initial model)

1.1 Determine whether the series is stationary or not by considering the graph of ACF. If a graph of ACF of the time series values either *cuts off fairly quickly* or *dies down fairly quickly*, then the time series values should be considered *stationary.* If a graph of ACF *dies down extremely slowly*, then the time series values should be considered *non-stationary.*

If the series is not stationary, it can often be converted to a stationary series by differencing. That is, the original series is replaced by a series of differences. An ARMA model is then specified for the differenced series. Differencing is done until a plot of the data indicates the series varies about a fixed level, and the graph of ACF either cuts off fairly quickly or dies down fairly quickly.

The theory of time-series analysis has developed a specific language and a set of linear operators. According to equation (2.1), a highly useful operator in time-series theory is the lag or *backward linear operator* (B) defined by $BY_t = Y_{t-1}$.

Model for non-seasonal series are called **Autoregressive integrated moving average model,** denoted by ARIMA (p, d, q). Here p indicates the order of the autoregressive part, d indicates the amount of differencing, and q indicates the order of the moving average part. If the original series is stationary, $d = 0$ and the ARIMA models reduce to the ARMA models. The difference linear operator (Δ), defined by

$$\Delta Y_t = Y_t - Y_{t-1} = Y_t - BY_t = (1-B)Y_t$$

The stationary series W_t obtained as the dth difference (Δ^d) of Y_t ,

$$W_t = \Delta^d Y_t = (1-B)^d Y_t$$

ARIMA (p, d, q) has the general form:

$$\phi_p(B)\,(1-B)^d Y_t = \mu + \theta_q(B)\varepsilon_t$$

or $\phi_p(B)\,W_t = \mu + \theta_q(B)\varepsilon_t$

1.2 Once a stationary series has been obtained, then identify the form of the model to be used by using the theory in TABLE 1.

STEP 2 : Model Estimation

Estimate the parameters for a tentative model has been selected.

STEP 3 : Model Checking

In this step, model must be checked for adequacy by considering the properties of the residuals whether the residuals from an ARIMA model must has the normal distribution and should be random. An overall check of model adequacy is provided by the Ljung-Box Q statistic. The test statistic Q is

$$Q_m = n(n+2)\sum_{k=1}^{m}\frac{r_k^2(e)}{n-k} \sim \quad \chi^2_{m-r} \quad ;$$

where $r_k(e)$ = the residual autocorrelation at lag k

n= the number of residuals

m= the number of time lags includes in the test

If *the* p-value associated with the Q statistic is small (p-value $< \alpha$), the model is considered inadequate. The analyst should consider a new or modified model and continue the analysis until a satisfactory model has been determined.

Moreover we can check the properties of the residual with the following graph:

(1) We can check about the normality by considering the normal probability plot or the p-value from the One-Sample Kolmogorov – Smirnov Test.
(2) We can check about the randomness of the residuals by considering the graph of ACF and PACF of the residual. The individual residual autocorrelation should be small and generally within $\pm 2/\sqrt{n}$ of zero.

STEP 4: Forecasting with the Model
Forecasts for one period or several periods into the future with the parameters for a tentative model has been selected.

4. Building ARIMA Model and Forecasting

In this study, we used the data for oil palm price for the period 2000 – 2004.

By following of the steps in the model-building, we can obtain the following results:

FIGURE 1. Time Series Plot of the Farm Price of Oil Palm.

FIGURE 2. ACF of the Farm Price of Oil Palm.

FIGURE 3. PACF of the Farm Price of Oil Palm.

From FIGURE 1-3, show that the time series data are not stationary in mean value, we then corrected through appropriate differencing of the data. In this case, we applied ARIMA$(2,1,0)$ model. Model parameter were shown as following

Type	Coef.	SE Coef.	T	p
AR1	0.4621	0.1259	3.67	0.001
AR2	-0.3899	0.1259	-3.10	0.003

TABLE 2. Estimated model parameters of farm price model

We obtained the model in the form

$$\widehat{Y_t} = Y_{t-1} + 0.4621(Y_{t-1} - Y_{t-2}) - 0.3899(Y_{t-2} - Y_{t-3})$$

with the MAPE = 13.23%.

From FIGURE 4- 6, proof that the selected ARIMA$(2,1,0)$ is an appropriate model.

Similarly to the whole sales price model and pure oil palm price model, we have the following results:

FIGURE 4. ACF of the residuals of the Farm Price model.

FIGURE 5. PACF of the residual of Farm Price model.

FIGURE 6. Plots of residual of the model.

Type	Coef.	SE Coef.	T	p
AR1	0.8039	0.0887	9.06	0.000
MA1	-0.3466	0.1487	-2.33	0.023
constant	3.1060	0.3680	8.44	0.000
mean	15.842	1.877		

TABLE 3. Estimated model parameters of whole sale price model

We obtained the whole sale price model in the form

$$\widehat{Y_t} = 3.106 + 0.8039Y_{t-1} + 0.3466\varepsilon_{t-1}$$

with the MAPE = 9.01%.

Type	Coef.	SE Coef.	T	p
AR1	1.4313	0.1251	11.44	0.000
AR2	-0.8840	0.1969	-4249	0.000
AR3	0.3781	0.1270	2.98	0.004
constant	1.8778	0.2317	8.11	0.000
mean	25.173	3.106		

TABLE 4. Estimated model parameters of pure oil price model

From TABLE 4, we obtained the pure oil price model in the form

$$\widehat{Y_t} = 1.8778 + 1.4313Y_{t-1} - 0.8840Y_{t-2} + 0.3781Y_{t-3}$$

with the MAPE = 5.27%.

5. Conclusion

In this paper, we developed model for three types of oil palm price, were found to be ARIMA(2,1,0) for the farm price model, ARIMA(1,0,1) for whole sale price, and ARIMA(3,0,0) for pure oil price. Which we can see that the MAPE for each model very small.

Acknowledgements

The authors acknowledges the helpful comments of referees. This work was partially supported by King Mongkut's Institute of Technology Lardkrabang, and Assumption University, Bangkok, Thailand, we are very appreciative.

References

[1] Y. Areepong, *Forecasting of Strategy Agriculture Products: Vegetables and Fruits,* Master Thesis in Statistics, Chulalongkorn University, 1999.

[2] Javier Contier, ARIMA Models to Predict Next-Day Electricity Price ,*IEEE Transactions on Power Systems* Vol **18**; No. **3**, August, 2003.

[3] Foreign Office, The Government Public Relations Department, Thailand Review; http://thailand.pnd.go.th/ebook/review, 2006

[4] John E. Hanke and Dean W. Wichern, *Business Forecasting* 8th ed. Pearson Education Inc., USA, 2005.

[5] C. Morana, A semiparametric approach to short-term oil price forecasting, *Energy Economics*, Vol. **23,** No. **3**, pp. 325-338, May 2001.

[6] E. Weiss, Forecasting commodity prices using ARIMA, *Technical Analysis of Stocks & Commodities,* Vol. **18,** No.**1,** pp.18-19, 2000.