KMITL Sci. J. Vol. 5 No. 1 Feb. 2005

Modelling and Forecasting Daily Rubber Smoked Sheet Prices

Chatchaipun, P. and Atthirawong, W.*
Department of Applied Statistics, Faculty of Science,
King Mongkut's Institute of Technology Ladkrabang, Bangkok 10520, Thailand

ABSTRACT

This paper presents the comparison between transfer function method and neural networks for predicting rubber smoked sheet prices. The two procedures are based on time series data extracted from 2001 to 2003. The accuracy of two forecasting models is evaluated using (1) root mean squared error: RMSE and (2) mean absolute percentage error: MAPE. The results indicate that the traditional statistical approach has a potential to forecast rubber smoked sheet prices compared to neural networks model. In addition, the future directions are also discussed in the paper.

Keywords: Time series, Transfer function method, Neural networks, Box and Jenkins method, Rubber smoked sheet (RSS3) prices, Back propagation.

1. INTRODUTION

Natural rubber is regarded as an important agricultural commodity vital for the manufacturing of a wide variety of products and new materials. The increasing volumes of rubbers are being produced, consumed and exported. Thailand is one of the largest producing and exporting country in the world. It is therefore essential to have a potential tool for predicting rubber prices for producers and consumers. The reliable prices forecast is desirable for both buying and selling agents. Likewise, the prediction over time horizon will be beneficial for the Thai government and relevant agencies for making short- and long- term planning policies in this industry.

Throughout the literature, several techniques have been applied to perform time series forecasting such as Holt's linear models, Holt-Winter's trend and seasonally method, exponential smoothing techniques, as well as Box and Jenkins method [1, 2, 3, 4]. However, there are not many applications of statistical techniques to predict behaviour of rubber prices in Thailand. Supasiripinyo [5] proposed econometric model to study the fluctuation of rubber smoked sheet prices. Recently, Chatchaipun and Atthirawong [6] employed Box and Jenkins and transfer function models to predict rubber smoked sheet prices and compared the accuracy of both models. It was revealed that transfer function model has performed better than Box and Jenkins model.

Although those traditional methods perform well, several have inherent limitations due to many reasons. Such reasons due to the way in which those models are estimated [7]. Human interactions and evaluation are required during the process. It has been argued that neural networks could overcome these drawbacks [7, 8, 9]. Furthermore, the method itself does not oblige any assumptions about underlying population distributions. The focus of this paper is therefore to extend the earlier work [6]. It aims to present and report the comparative results of transfer function and neural networks models. The transfer function approach has been adopted here to compete with neural networks as it was shown to be the best-in-class among the traditional approaches in the previous study [6]. The procedures are based on time series analysis and applied to actual data of rubber smoked sheet (RSS3) prices at Hatyai central rubber market collected from 2001 to 2003 [10,11,12].

The outline of this paper is organised as follows. A statistical description of the proposed models is firstly introduced in section 2. Section 3 is devoted to the numerical results and discussions of the study. Finally, some conclusions and further work are discussed in the last section.

* Corresponding authors: Tel (662)-3264111 ext.6172 Fax(662)-3264305
E-mail address: walailak_ut@yahoo.com, me_porntip@yahoo.com

2. METHODS

In this section, the description of the two statistical methodologies to build a final model is briefly presented.

2.1 Transfer Function

Transfer function method is a dynamic regression model which allows the explanatory variable(s) to be included. The main objective of the model is to predict what happen to the forecast variable or output time series, called y_t, if the explanatory variable or input time series, called x_t changes.

Let x_t and y_t represent input and output data of transfer function model, respectively. These time series data consist of n =733 observations of

y_t = Rubber smoked sheet (RSS3) prices in days *t*

and x_t = Yen currencies in days *t*

The basis of the transfer function approach to modelling time series consists of four steps. The process of each step can be summarised as follows.

Step 1: Identification

In this step, the appropriate transfer function model will be identified. It is assumed that the input and output time series must be both stationary. If not, it is necessary to transform those data into stationary form.

1.1 Prewhitening of x_t and y_t

Once an appropriate model describing x_t can be identified, the relationship between x_t and y_t will be estimated.The prewhitened x_t and y_t values can be calculated by equations (1) and (2), respectively:

$$\alpha_t = \frac{\hat{\phi}_p^{(x)}(B)}{\hat{\theta}_q^{(x)}(B)} Z_t^{(x)} \qquad \text{(for } x_t\text{)} \qquad (1)$$

$$\beta_t = \frac{\hat{\phi}_p^{(x)}(B)}{\hat{\theta}_q^{(x)}(B)} Z_t \qquad \text{(for } y_t\text{)} \qquad (2)$$

1.2 Calculation of the sample cross-correlation function (SCC) and identification of a preliminary transfer function model

In order to identify a preliminary transfer function model describing the relationship between y_t and x_t, the sample cross-correlation function (SCC) between the α_t value and the β_t value must be computed from the following equation:

$$r_k(\beta_t,\alpha_t) = \frac{\sum_{t=1}^{n-k}(\alpha_t-\overline{\alpha})(\beta_{t+k}-\overline{\beta})}{\sqrt{\sum_{t=1}^{n}(\alpha_t-\overline{\alpha})^2\sum_{t=1}^{n}(\beta_t-\overline{\beta})^2}} ; k = 0,\pm1,\pm2,\ldots \qquad (3)$$

the general preliminary transfer function model is computed as:

$$Z_t = \mu + \frac{C\omega(B)}{\delta(B)} B^b Z_t^{(x)} + \eta_t \qquad (4)$$

where $z_t = y_t - y_{t-1}$

and $z_t^{(x)} = x_t - x_{t-1}$

In selecting the form of the model, the value of b, r and s must be determined from the correlogram of $r_k(\beta_t, \alpha_t)$. The valued of b is the number of periods before the input data (x_t) begins to influence output data (y_t). It is equal to the lag where the first spike in the SCC is encountered or the number of weights that are not significantly from żero. The value of r represents the number its own past value z_t. The value of s represents the number of past $z_t^{(x)}$ values influencing z_t [1].

1.3 Identification of a model describing η_t and of a final transfer function model

It is necessary to check whether the preliminary model is adequate by analyzing the residuals (η_t) with the values of the input x_t. The model describes η_t is determined by the following equation:

$$\eta_t = \frac{\hat{\phi}_{pn}(B)}{\hat{\theta}_{qn}(B)} \varepsilon_t \qquad (5)$$

where η_t is a disturbance term that follows an ARIMA model.
Hence, an appropriate final transfer function model is of the form:

$$Z_t = \mu + \frac{C\omega(B)}{\delta(B)} B^b Z_t^{(x)} + \frac{\phi_{pn}(B)}{\theta_{qn}(B)} \varepsilon_t \qquad (6)$$

Step 2: Parameters Estimation

In step 2, after all functions of the model have been structured, the parameters of these functions will be estimated. Good estimators of the parameters can be found using least squares method by assuming that those data are stationary.

Step 3: Diagnostic Checking

A diagnostic checking is employed to validate the model assumptions and to check whether the model is adequate. It is necessary to do diagnostic checking even if the selected model may perform to be the best among others. It checks whether the hypotheses made on the residuals are true or not. These residuals must be a white noise series: zero mean, constant variance, uncorrelated process and normal distribution. These requirements can be investigated by inspecting the autocorrelation function (ACF) and partial autocorrelations function (PACF) plots of the residuals and taking tests for randomness such as Ljung-Box statistics.

Step 4: Forecasting with transfer function

If the hypotheses on the residuals from step 3 are satisfied, the forecast prices of the final model are then computed and compared the results with the test data.

2.2 Neural Networks

A neural network or an artificial neural network is a computation model for inspired by the functioning of biological of neuron systems. It is structured in layers of basis processing unit called neurons or nodes. Basically, the neural network comprises of two layers i.e. input and output layers. Between them, hidden layers can be added to solve non-linear problem. The network is fully-linked together in adjacent layers. It means that neuron in any layer is connected to all nodes in the previous layer. Signals flow in a forward direction from left to right and from one layer to another. Figure 1 illustrates the architecture of neural network.

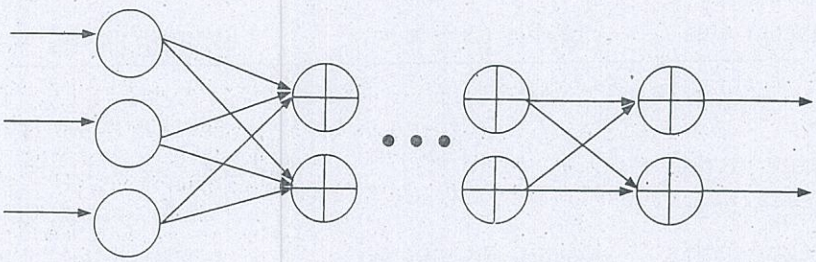

Figure 1: Architectural graph of a multilayer network with hidden layers

Each input neuron has activation value or a signal on it and fall in the closed interval range [0, 1]. Learning process from historical data is a very important task in the development of neural networks system. Basically, there are 2 types of learning process: (1) supervised learning process, and (2) unsupervised learning process. Back propagation, the most common supervised learning algorithm, is utilised for training multiple-layers networks in this study [7, 8]. The strength of the linkage is characterised by a weight (w_i).

The internal activation coming into neurons in the hidden or output layers are the sum of incoming activation level times its respective connection weight. The internal activation is then modified by the sigmoid transfer function and translates into outputs. The sigmoid transfer function is defined as:

$$f(x) = \frac{1}{1 + e^{-x}}$$

where e is the base of natural logarithms.

The range of the sigmoid transfer function is between zero and one [0, 1]. This function employed to helps to reduce the extreme input values. Thus, it produces some degree of robustness to the network [4].

The weights are adjusted to minimise the root mean squared error function (RMSE) between the model output and the desired output using the concept of the gradient steepest descent algorithm. The process repeats until no significant improvement in the network performance is obtained. In order to evaluate the effect of proposed approach to back propagation learning algorithm, the relevant parameters of this study are determined in Table 1.

Table 1: Neural networks parameters

Learning rate	0.3
Momentum	0.6
Training epochs	40,000
Minimum error	0.05

3. NUMERICAL RESULTS AND DISCUSSIONS

3.1 The observation data

The models explained in section 2 have been employed to forecast the rubber smoked sheet prices at Hatyai central rubber market. Secondary data were collected from Rubber Research Institute between January 3rd 2001 and December 29th 2003. The week of 5th to 9th January 2004 has been selected to be a set of test data. Moreover, time series data of yen currencies exchange rate has been included into both models as input data. These data were extracted at the same period from 2001 to

2003.The reason of using yen currencies because it seems to partly describe the pattern of rubber sheet prices behaviour [13].

3.2 Results and discussions

The final results of forecasting data obtained by transfer function model and neural networks are shown and compared with the test data in Table 2. Additionally, they are plotted in Figure 2. Root mean squared error (RMSE) and mean absolute percentage error (MAPE) are employed to evaluate the accuracy of the models. The results from Table 3 indicate that the error of transfer function model is slightly smaller. It is implied that the transfer function model is more suitable for predicting rubber smoked sheet prices. One possible reason in that the neural networks did not perform better than the transfer function is the design of neural networks is very complex. The procedure is mostly depended on trail and error process. Moreover, it requires the appropriate input variables, the necessary level of training of the system and the number of hidden nodes and hidden layers.

Table 2: Comparison of test data with transfer function and neural network models

Forecast horizon	Test data	Models	
		Transfer function	Neural networks
5 January 2004	44.56	43.22	44.76
6 January 2004	43.43	43.69	44.72
7 January 2004	43.20	43.48	44.56
8 January 2004	43.59	43.59	44.64
9 January 2004	44.07	43.67	44.54

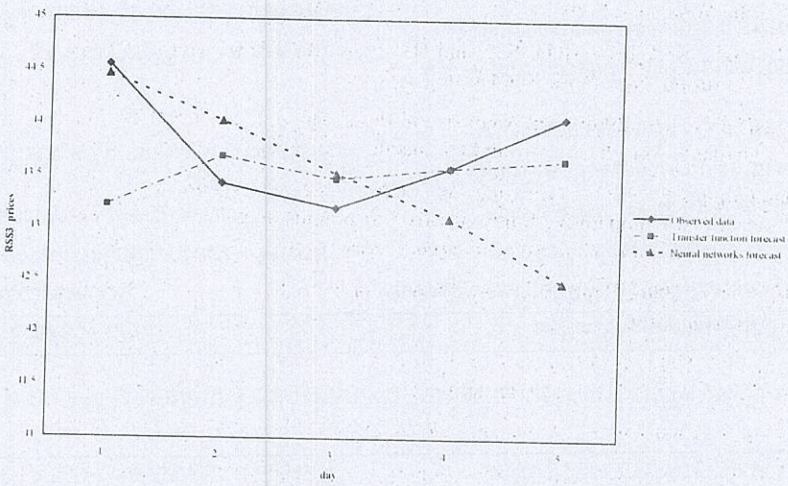

Figure 2: Plots of the forecast data in Table 1

Table3: Statistical accuracy measures

Model	Accuracy measures	
	RMSE	MAPE
Transfer function	0.648	1.062
Neural networks	0.7965	1.402

4. CONCLUSIONS AND FURTHER WORK

In this paper, we have presented efficient techniques to accurately predict time series data of rubber smoked sheet (RSS3) prices. The time series forecast based on transfer function method was compared to neural networks across the 5 periods ahead in the forecast horizon. The initial results obtained from the study revealed that the transfer function method is more suitable for constructing the model of predicting RSS3 prices.

There are several directions in which this research can be expanded:

- the neural networks developed for this study could be modified in terms of learning rule, different training techniques, different of hidden layers, neural network type and topology,
- to improve the prediction accuracy, one possible way is to incorporate more elements e.g. prices of natural rubber at various markets around the world and geographical factors as input data into both models,
- longer time series data might improve the prediction results.

REFERENCES

[1] Bowerman, B.L. and O'Connel, R.T. **1987**. *Time Series Forecasting.* Boston, Duxbury Press.

[2] Delurgio, S.A. **1998**. *Forecasting Principles and Applications.* Singapore, McGraw-Hill.

[3] Montgomery, D.C. and Johnson, L.A. **1977**. *Forecasting and Time Series Analysis.* New York, John Wiley and Sons.

[4] Makridakis, S., Wheelwright, S.C. and Hyndman, R.J.**1998**. *Forecasting: Method and Applications.* 3rd Edition. New York, John Wiley and Sons.

[5] Supasiripinyo, S. **2000**. An Analysis of Price Movement of Rubber Smoke Sheet Grade 3 of Thailand. Master thesis, Ramkhamhaeng University.

[6] Chatchaipun, P. and Atthirawong, W. **2003**. Time Series Forecasting of Rubber Smoked Sheet Prices: A Comparative Study of Box and Jenkins and Transfer Function Methods. *Proceedings of 1st Operations Research Conference,* Kasertsart University, Bangkok, pp. 7-17.

[7] Hill, T. et al. **1996**. Neural Network Models for Time Series Forecasts. *Management Science.* 42(7), pp. 1082-1092.

[8] Fausett, L.**1994**. *Fundamental of Neural Networks: Architectures Algorithms and Applications.* London, Prentice-Hall International.

[9] Haykin, S. **1994**. *Neural Networks: A Comprehensive Foundation.* New York, Macmillan College.

[10] www. http: www.rubberthai.co.th/ rubberprice.2001.html.

[11] www. http: www.rubberthai.co.th/ rubberprice.2002.html.

[12] www http: www.rubberthai.co.th/ rubberprice.2003.html.

[13] Phetgran, S.**1989**. *The Relationship of Natural Rubber Prices in Various Market Levels in Thailand and Effecting Factors.* Master thesis, Kasertsart University.