A STUDY OF THE HYBRID MODEL PERFORMANCE FOR TIME SERIES FORECASTING

Arayachon Dutsarak[1] and Jirapan Liangrokapart[2]
[1,2]Faculty of Engineering, Mahidol University, 25/25 Salaya, Phutthamonthon District, Nakhon Pathom 73170, Thailand

ABSTRACT

Automotive export is an important industry that contributes to Thailand economy. The current models to forecast the export quantity include Moving Average (MV), Holt-Winters (HW), Autoregressive Integrated Moving Average (ARIMA), and Artificial Neural Networks (ANNs). However, time series data often contain both linear and nonlinear patterns which the current forecasting models cannot provide much accuracy. In this study, a hybrid model is proposed to forecast automotive export quantity. The hybrid model combines the unique strength of ARIMA and ANN which is good for modeling linear and nonlinear behavior data. The comparison of ARIMA, ANN, Additive ARIMA-ANN and Multiplicative ARIMA-ANN is presented. Performance of the hybrid models is measured using the mean absolute deviation, the mean square error and mean absolute percentage error. The results indicate that the performance of the hybrid models is better in term of forecast accuracy than the other compared models.

KEYWORDS: Artificial Neural Networks, ARIMA, Hybrid model, Time Series Forecasting

1. Introduction

Thailand's automotive export industry has rapid expansion and high growth rate. The automotive industry including private cars, pickup trucks, and motorcycles which is an important industry and has created a number of entrepreneurs. The automotive industry has significant impact on the country's economic development, including manufacturing, employment, technology development and links with many other industries, including the steel industries, rubber, glass, and electronics. From Thailand Economic Report in 2016, the figures of automotive export value has fluctuated due to the fluctuated [1]. Forecasting is one of the factors that contribute to the determination and implementation of international trade policies. The forecasted figures will enable the government to plan for international trade.

A number of forecast models are used depending on the characteristics of the data. For export forecast, the popular forecast model is time series forecasting. Time series forecasting can be termed as the act of predicting the future by understanding the past. Many researchers used the models to forecast the export quantity in form of univariate individual models. For example, Co and Boosarawongse compared the performance of Artificial Neural Networks (ANNs) with exponential smoothing and Autoregressive Integrated Moving Average (ARIMA) model to forecast the volume of international trade of rice in Thailand [2]. Emang et al, studied export demand of moulding and chipboard volume (m^3) from Peninsular, Malaysia. Export demand for moulding and chipboard were estimated using univariate time series models including the Holt-Winters Seasonal, ARAR algorithms and the seasonal ARIMA models [3]. Wong *et al*, found similar results that the ARIMA models generated smaller forecasting errors over a longer time period [4]. Yavuz *et al,* using an ARIMA model to determine the beef market tendency and use this information as an aid to design public policy in Turkey [5]. However, the predictive capacity of types of univariate models is not always optimal.

Univariate models have limitations in predicting time series data that have a linear and non-linear combination in the same data set [6]. For example, the Moving Average (MV), Holt-Winters (HW) and Autoregressive Integrated Moving Average (ARIMA) model are linear predictive models. The predictive function is a linear combination of several past observations. This model captures linear data, but cannot deal with nonlinear relationships while the neural network model alone is able to handle nonlinear patterns equally well when compare with the linear model. As the time series data of Thailand's export figures are a combination of linear and non-linear, the individual model may not be as good as expected. Hence, the hybrid model using both ARIMA and ANNs was proposed by Zhang [6] combining the advantage of ARIMA and ANNs for modelling linear and nonlinear behaviour in data sets. However, the Zhang model was the additive hybrid model. Therefore, in this paper proposes a new multiplicative hybrid model for automotive export industry forecasting.

2. Methodology

2.1 Autoregressive Integrated Moving Average (ARIMA) Model.

The autoregressive integrated moving average (ARIMA) model was purposed by Box and Jenkins [7]. The ARIMA model has been one of the most popular approaches for forecasting. In ARIMA model, the future value of a variable is assumed to be a linear combination of several past observations and past errors, which can be denoted as ARIMA (p,d,q) or expressed as the following form:

$$\Delta_d y_t = \delta + \emptyset_1 \Delta_d y_{t-1} + \ldots + \emptyset_p \Delta_d y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \ldots - \theta_q \varepsilon_{t-q} \quad (1)$$

Where y_t is the value of observations and ε_t is the random error at time *t*, $\emptyset_p$ and θ_q are the coefficients, *p* and *q* are integers that are often referred to as autoregressive and moving average polynomials, respectively.

2.2 Artificial Neural Network (ANN) Model

Artificial neural networks (ANNs) approach has been suggested as an alternative technique to time series forecasting for non-linear component in the last few years. The most widely used ANNs in forecasting problems are multi-layer perceptrons (MLPs), which use a single hidden layer feed forward network (FNN) in diagrammatically depicted as below [6].

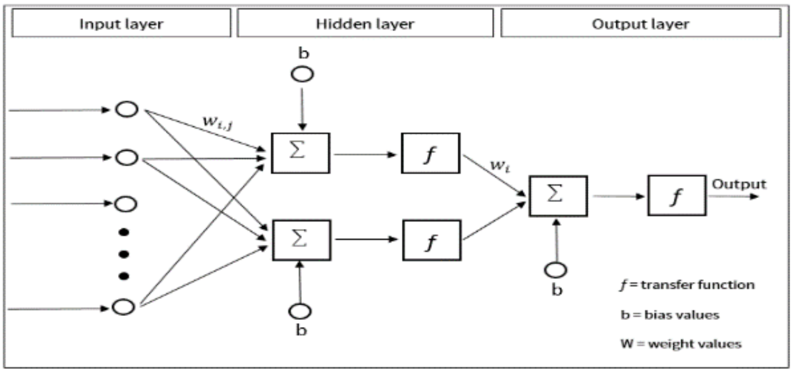


Figure 1 The feed forward architecture

The output of the model is computed using the following mathematical expression [8].

$$y_t = w_0 + \sum_{j=1}^{Q} w_j g\left(w_{0j} + \sum_{i=1}^{P} w_{i,j} y_{t-i}\right) \qquad (2)$$

where y_t is the output and y_{t-i} (*i = 1,2,..., P*) . The integers P, Q are the number of input and hidden nodes respectively, g is a sigmoid transfer function. w_j = (*j= 1,2,...,Q*) is a vector of weights from the hidden layer to output nodes, $w_{i,j}$ (i =1,2,..., Q; j =1,2,...,Q) are the weights from the input to hidden nodes. w_0 and w_{0j} are the bias terms. Usually, the sigmoid or hyperbolic tangent $\mathrm{f(x)} = \frac{1}{1+\mathrm{e}^{-\mathrm{x}}}$ is applied as the nonlinear activation function. Other activation functions, such as linear, logistic sigmoid function, etc. can also be used [8, 9].

2.3 Hybrid ARIMA-ANN Models

In general, the principle of the hybrid model based on effects of time series components namely, additive model = T+C+S+I and multiplicative model = T*C*S*I which Trend (T); Cyclical (C); Seasonal (S); and Irregular (I). Moreover, the complexity in the time series is composed of a linear component (L) and a nonlinear component (N). This study may assume two models in time series, an additive model (L+N) and a multiplicative model (L*N). The hybrid model exploits the unique feature and strength of ARIMA model as well as ANN model in determining different patterns. Thus, it could be advantageous to model linear and nonlinear patterns separately by using different models and then combine the forecasts to improve the overall modeling and forecasting performance.

2.3.1 Additive Hybrid ARIMA-ANN Model

The additive hybrid ARIMA-ANN model was proposed by Zhang [6]. It is based on the assumption that the given Time series is a sum of two components, linear and non-linear, given in:

$$y_t = L_t + N_t \qquad (3)$$

where L_t denotes the linear component and N_t denotes the nonlinear component. These two components have to be estimated from the data. In the first step, an ARIMA model is used to analyze the linear component, then the residuals from the linear model will contain only the nonlinear relationship. Let e_t denote the residual at time t from the linear model, then

$$e_t = y_t - \hat{L}_t \tag{4}$$

where $\hat{L}_t$ is the forecast value for time t from the estimated relationship in Eq (4). The second step, a neural network model is developed to model the residuals from the ARIMA model (see Figure 2).

$$e_t = f(e_{t-1}, e_{t-2}, \dots, e_{t-n}) + \varepsilon_t, \tag{5}$$

The hybrid model predictions are now obtained by summing the ARIMA and ANN predictions in Eq (6):

$$\hat{y}_t = \hat{L}_t + \hat{N}_t \tag{6}$$

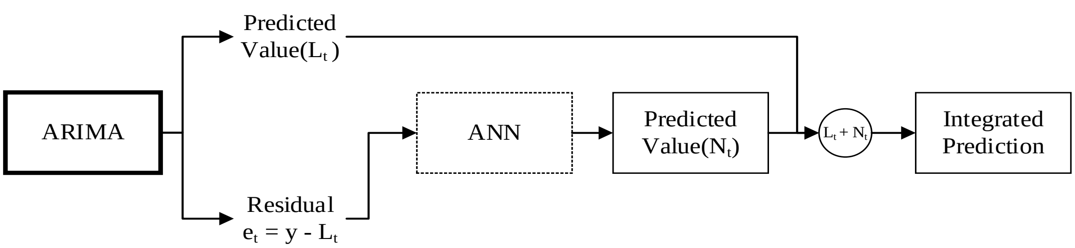


Figure 2 The Additive ARIMA-ANN Model

2.3.2 Multiplicative Hybrid ARIMA-ANN Model

We present the multiplicative model for forecasting time series data, in contrast to the additive model proposed by Zhang. The model assumes that a given Time series data is the product of a linear and a non-linear time series as:

$$y_t = L_t * N_t \tag{7}$$

The given time series data y_t is modeled using ARIMA as shown in Eq. (4), similar to the same step in Zhang model. Then, the predictions $\hat{L}_t$ obtained divide the original time series data to obtain the non-linear time series data series as:

$$e_t = y_t / \hat{L}_t \tag{8}$$

The series e_t is modeled and predicted using ANN. The obtained non-linear predictions $\hat{N}_t$ in Eq. (8) and linear predictions $\hat{L}_t$ are multiplied to obtain the final model forecasts as given by Eq. (9). The block diagram of this model is as shown in Figure 3.

$$\hat{y}_t = \hat{L}_t * \hat{N}_t \tag{9}$$

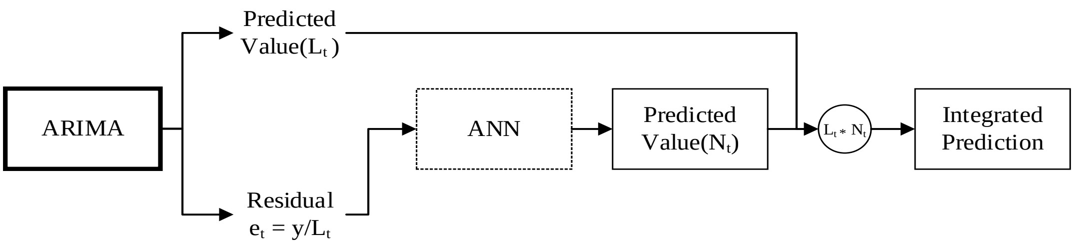
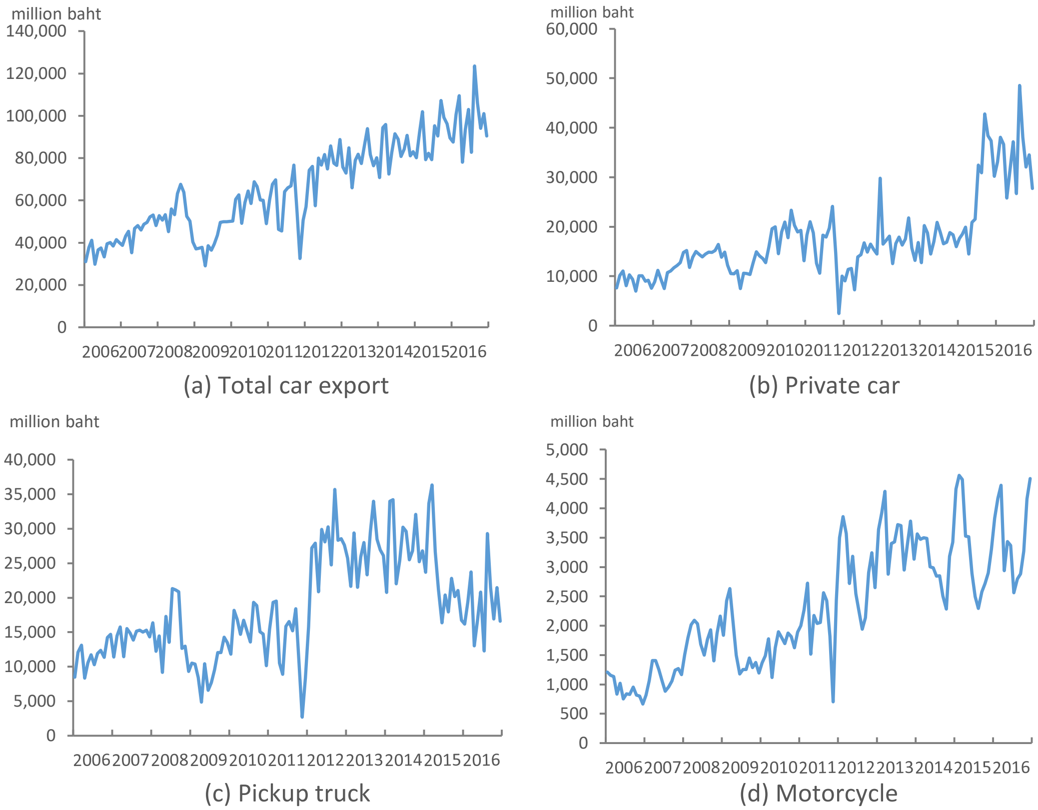


Figure 3 The Multiplicative ARIMA-ANN Model

3. Data Description and Forecast Evaluation Criteria

3.1 Data Description

The data of monthly automotive export value were used for forecasting the hybrid model performance. For data sets include the export value of total cars export, private cars, pickup trucks, and motorcycles. The data collection period is from January 2006 to December 2016 with a total of n =132 observations, as illustrated in Figure 4.



Source : Thailand Trading Report, Ministry of Commerce [10].

Figure 4 Thailand's automotive export monthly data from Jan 2006 to Dec 2016

To compare the performance of each model, the data set is divided into two samples, a training and a testing set. The training data set is used exclusively for model development and the testing data set is used to evaluate the established model. Many researchers in the literatures use appropriate ratios of splitting data not less than 80:20 for the training and the testing sets, respectively. Thus, the data were splited into 108 observations for training set and the remainder of 24 observations for testing set.

3.2 Forecast Evaluation Criteria

The performance of the each model was evaluated using Mean Absolute Deviation (MAD), Mean Square Error (MSE) and Mean Absolute Percentage Error (MAPE) which are widely used for evaluating result of time series forecasting [6, 11, 12]. The MAD, MSE, and MAPE are as follow:

$$\text{MAD} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \quad (10)$$

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \quad (11)$$

$$\text{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i-\hat{y}_i}{y_i}\right| * 100 \quad (12)$$

where y_i and $\hat{y}_i$ are the observed and predicted data and n is the number of observations.

4. Forecast procedure

4.1 Forecasted ARIMA Model

The ARIMA(p,d,q) were found and the best model under the Akaike Information Criterion (AIC) were selected from the training set [11,12,13]. The AIC formula is as follow:

$$\text{AIC}(p) = n\ln(\hat{\sigma}_e^2/n) + 2p \quad (13)$$

Here n is the number of effective observations, used to fit the model, p is the number of parameters in the model and $\hat{\sigma}_e^2$ is the sum of sample squared residuals. The optimal model order is chosen by the number of model parameters.

The forecasted ARIMA model started from determining “d” with unit root test method. Then, we are determines p and d which is the order of AR and MA term from 0 to 12 and 0 to 5, respectively. The result of the best ARIMA(q,d,q) provide to small AIC of each data set show in Table 1.

Table1 The best of structure ARIMA(p,d,q) model

Data set	ARIMA (p,d,q)	AIC
Total car export	0x1x2	2237.90
Private car	1x1x2	2040.25
Pickup truck	0x1x2	2092.73
Motorcycle	0x1x3	1602.48

4.2 Forecasted Neural Network Model

The ANN models are trained with the modifications of conventional backpropagation algorithm (Levenberg-Marquardt algorithm) over the training set and then tested over the testing set. The convergence criteria used for training is MSE less than or equal to 0.001 or a maximum of 1000 iterations. The experiment was repeated 10 times and searched for the best model under number input and hidden nodes of 1 to 15 for forecasting performance of neural network. The best model of neural network structures were delimited as 6x1x1, 11x1x1, 1x1x1, and 12x1x1, which were selected to model the total car export, private cars, pickup trucks, and motorcycles, respectively.

4.3 Forecasted Hybrid Models

In a similar fashion, the hybrid models forecasting of additive and multiplicative model use the same terms as the ARIMA and the ANN models. By selecting parameters of the best ARIMA model, than analysis residual from ARIMA with ANN model which follow method in headings 2.3.1 and 2.3.2. The criteria for selecting the best model mentioned in the residual section, the neural network structures of additive model were delimited as 2x3x1, 12x1x1, 6x2x1, and 13x1x1. For the multiplicative model were delimited as 2x9x1, 7x3x1, 9x2x1, and 13x3x1, which were selected to model the total car export, private cars, pickup trucks, and motorcycles, respectively.

5. Findings

The forecasting results for the automotive export data showed that the accuracy associated with the ARIMA model and the ANNs model are not good compared to the hybrid models (see Table 2).

Table 2 **The forecasting results of automotive export value.**

Category	**3 month forecasted**			**24 month forecasted**		
	MAD	MSE	MAPE %	MAD	MSE	MAPE %
Total car export						
ARIMA	9239.85	9.70E+07	9.83	9701.45	1.60E+08	10.22
ANN	10113.91	1.19E+08	10.74	9286.12	1.45E+08	9.56
Additive model	7614.68	7.32E+07	7.98	8939.07	1.28E+08	9.16
Multiplicative model	7762.28	6.55E+07	8.32	7938.15	8.96E+07	8.30
Private cars						
ARIMA	1083.45	1.64E+06	5.64	5180.05	5.52E+07	16.66
ANN	1815.21	3.61E+06	9.61	8657.07	1.13E+08	25.58
Additive model	1425.55	2.99E+06	7.56	5222.44	4.78E+07	16.95
Multiplicative model	551.74	4.66E+05	2.86	4668.62	4.36E+07	14.71
Pickup trucks						
ARIMA	6003.53	3.96E+07	19.02	4620.40	3.14E+07	23.14
ANN	6169.71	5.11E+07	18.61	4882.19	3.49E+07	24.33
Additive model	5285.04	3.13E+07	16.87	4158.08	2.65E+07	20.47
Multiplicative model	8285.18	7.01E+07	26.81	3370.93	2.28E+07	16.45
Motorcycles						
ARIMA	667.16	5.29E+05	15.07	464.82	3.22E+05	13.80
ANN	816.31	6.69E+05	18.33	430.30	2.66E+05	12.36
Additive model	477.40	2.47E+05	10.78	410.06	2.49E+05	12.44
Multiplicative model	528.56	3.07E+05	11.93	382.03	2.16E+05	11.57

For the of 3 month forecasted base on MAPE, an additive hybrid model was 7.98%, 16.87% and 10.78% for total car export, pickup trucks, and motorcycles, respectively outperformed than the ARIMA, ANN, and multiplicative model. Likewise, the additive model provided better predictions over the other models based on MAD and MSE. It should be noted that additive model do not provide better forecast for every data point especially the private car that provide accuracy less than multiplicative model. However, this model is generally more accurate than other models in this dataset. For the 24 month forecasted, a multiplicative hybrid model generated superior results as indicated by the improvements in MAPE by 8.30%, 14.71%, 16.45% and 11.57 for total car export, pickup trucks, private cars and motorcycles respectively. Also, it should be noted that the value of MAD and MSE associated with the multiplicative model was the lowest among the four models.

6. CONCLUSION

The forecast of automotive export value is important for managing Thailand automotive industry. The current forecasting models have some limitations in predicting time series data that have a linear and nonlinear combination in the same set data set. In this paper, a multiplicative hybrid model was proposed using the concept that the ARIMA model was used to analyze the linear part of the problem and then the residuals from the ARIMA model were modeled by using the ANNs model. The results from the multiplicative hybrid model indicated that for the long-term forecast approach gave more reliable prediction for automotive time series data than the additive hybrid model and the single ARIMA and ANNs models. For short-term forecast, the additive hybrid model has the highest accuracy

References

[1] World Bank. Thailand economic monitor: aging society and economy. Thailand: World Bank Group; 2016.

[2] Boosarawongse R. Forecasting Thailand's rice export: Statistical techniques vs. artificial neural networks. Computers & industrial engineering 2007;53(4):610-27.

[3] Emang D, Shitan M, Ghani AN, Noor KM. Forecasting with univariate time series models: a case of export demand for peninsular Malaysia's moulding and chipboard. Journal of Sustainable Development 2010;3(3):157.

[4] Wong HL, Tu YH, Wang CC. Application of fuzzy time series models for forecasting the amount of Taiwan export. Expert Systems with Applications 2010;37(2):1465-70.

[5] Yavuz F, Bilgic A, Terin M, Guler IO. Policy implications of trends in Turkey's meat sector with respect to 2023 vision. Meat science 2013;95(4):798-804.

[6] Zhang GP. Time series forecasting using a hybrid ARIMA and neural network model. Neurocomputing 2003;50:159-75.

[7] Box GE. P, and GM Jenkins. Time series analysis: forecasting and control. Revised edition. 1976.

[8] Zhang GP. A neural network ensemble method with jittered training data for time series forecasting. Information Sciences 2007;177(23):5329-46.

[9] Eswaran C, Logeswaran R. An enhanced hybrid method for time series prediction using linear and neural network models. Applied Intelligence 2012;37(4):511-9.

[10] Ministry of Commerce. The value of Thailand's automotive export [Internet]. 2006 [cited 2016 Jan 7]. Available from: http://www.ops3.moc.go.th/infor/menucomth/stru1_export/export_re/report.asp

[11] Faruk DÖ. A hybrid neural network and ARIMA model for water quality time series prediction. Engineering Applications of Artificial Intelligence 2010;23(4):586-94.

[12] Shi J, Guo J, Zheng S. Evaluation of hybrid forecasting approaches for wind speed and power generation time series. Renewable and Sustainable Energy Reviews 2012; 16(5):3471-80.

[13] Hipel KW, McLeod AI. Time series modelling of water resources and environmental systems. Elsevier; 1994.

Author's Profile

Arayachon Dutsarak is a student of master degree at the Faculty of Engineering, Mahidol University, Phuthamonthon sai 4, Salaya, Nakornprathom 73170 Thailand. Her contact number is 091- 8868810 and email arayachon.bom@gmail.com. Mr. Arayachon has got Bachelor of Engineering at Silpakon University. Her research areas of interest include Operations Improvement and Logistics and Supply Chain Management.

Jirapan Liangrokapart, Dr. is a lecturer at the Faculty of Engineering, Mahidol University, Phuthamonthon sai 4, Salaya, Nakornprathom 73170 Thailand. Her contact number is 02-889-2138 ext. 6221 and email: jirapan.lia@mahidol.ac.th. Dr.Jirapan has got an MBA degree from Thammasat University, Thailand and a Ph.D. in Industrial Engineering from Clemson University, USA. Her research areas of interest include Operations Improvement, Performance Measurement, Logistics and Supply Chain Management, and Transportation Management.