FORECASTING MODELS FOR FIRST YEAR PREMIUM OF LIFE INSURANCE

Somsri Banditvilai
Choojai Kuharattanachai
Department of Statistics
King Mongkut's Institute of Technology Ladkrabang
Bangkok 10520, Thailand
E-mail: somsri.ba@kmitl.ac.th

KEYWORDS
Forecasting, Ordinary Life Insurance, Holt-Winters method, Box-Jenkins method, ANN

ABSTRACT

The objective of this research is to study forecasting models for the first year premium of life insurance. The premium data are gathered from the Office of Insurance Commission (OIC) during January 2003 to November, 2021. The data are divided into 2 sets. The first set from January, 2003 to December 2020 is used for constructing and selection the forecasting models. The second one from January 2021 to November 2021 is used for computing the accuracy of the forecasting model. The forecasting models are chosen by considering the minimum Root Mean Square Error (RMSE). The Mean Absolute Percentage Error (MAPE) is used to measure the accuracy of the model. The results showed that the multiplicative model with initial values from 18 years Decomposition method give the appropriate model for the first year premium of life insurance and yields the MAPE = 17.29%

INTRODUCTION

The Office of Insurance Commission (OIC) of Thailand defines life insurance as it is a way for a group of people to share the dangers of death, dismemberment, disability and loss of income in old age. When any person has to face those dangers and received an average amount of help to alleviate suffering for themselves and their families. The life insurance company will act as the center for bringing such sums to pay to the victims.

First year premium is the amount of an insured person must pay to the life insurance company in the first year to purchase the coverage that they will receive from life insurance.

Life insurance is an essential tool in financial risk services for those who wish to prepare for unforeseen emergencies and retirement that will occur in the future. There are also many other benefits in terms of saving money with discipline and continuity. Most types of life insurance are insurance with long-term commitments. This will allow the insured to use this money for future needs. Both as scholarship money for children or as money saving for old age. It is a guarantee for one's life and family in helping the insured person feel more secure.

Life insurance therefore plays a direct role in the well-being of the people. In terms of financial planning to create stability for yourself not to be a burden on others and also helps to drive the economy. It is a source of long-term fund raising in the form of life insurance premium that can be invested for profits as government bonds causing money to circulate in the economy.

Forecasting is a powerful tool in helping life insurance companies make predictions the growth trend of life insurance premium in the future for the decision of the executives in driving the business to continue well.

At present, the most popular technique for forecasting is time series analysis. It is a statistical method that uses historical data collected in continuous chronological order to study the relationship patterns of the data by creating equations to guide future forecasts. There are several methods used in forecasting. This research will find a model for forecasting ordinary life insurance premium. This is because it is the most popular type of insurance in the market. which accounted for 79.76% (Thai Life Insurance Association 2019)

Artificial Neural Networks was used for forecasting model of insurance premium revenue. The actual and forecast value was almost the same. Therefore, the results confirmed to be strong and useful to deploy it for forecasting the insurance premium revenue. (Bahia 2013) Box-Jenkins method was employed for modeling the forecasting of life insurance premium and insurance penetration rate and the results were good and degree of accuracy were over 80%. (Namawejje and Geofrey 2020) Safyar (2010) studied the forecasting life insurance premium by ARIMA and Neural Networks and it was found that Neural Networks is the best model to predict life insurance premium. The Holt-Winters method is suitable for time series forecasts with both linear trends and seasonal influences. It is simple and easy and it is one of the most widely used time series forecasting methods. This has led to studies of different trend defaults in different studies. Changing the trend initiation affects the Holt-Winters forecasting performance. Different initiation configurations provide significantly different MAPE values (Booranawong and Booranawong 2018). Holt-Winters and Extended Additive Holt-Winters (EAHW) method, with 4 different initiation values were used for modeling crude palm oil yield and price

Communications of the ECMS, Volume 36, Issue 1,
Proceedings, ©ECMS Ibrahim A. Hameed, Agus Hasan,
Saleh Abdel-Afou Alaliyat (Editors) 2022
ISBN: 978-3-937436-77-7/978-3-937436-76-0(CD) ISSN 2522-2414

forecasting in Thailand and the results showed that different initiation values in both forecasting methods gave a significant different MAPE values. (Suppalakpanya et al. 2019) Therefore, this research employs the Holt-Winters Exponential Smoothing method with different initiation, Box-Jenkins, and Artificial Neural Networks in forecasting the first year premium of ordinary life insurance of Thailand.

DATA COLLECTION AND METHODOLOGY

The first year premium of ordinary life insurance are gathered from the Office of Insurance Commission (OIC) of Thailand. The monthly data are collected from January, 2003 to November 2021 and it is divided into 2 sets. The first set from January, 2003 to December 2020 is used for constructing the models and employed the minimum Root Mean Square Error (RMSE) for model selection. The second one from January 2021 to November 2021, it is used to compute the accuracy of forecasting models by using the Mean Absolute Percentage Error (MAPE). This research employs three forecasting methods which are the Holt-Winters Exponential Smoothing methods with different initiation, Box-Jenkins method, and Artificial Neural Networks to construct the forecasting models.

Holt-Winters method with different initial values

The Holt-Winters method of exponential smoothing involve linear trend and seasonal variation which are based on three smoothing equations: for level, for trend and for seasonal variation. The decision regarding which model to use depends on time series characteristics: the additive model is used when the seasonal component is constant. The multiplicative model is used when the size of the seasonal component is proportion to trend level (Chatfield 1996)

Additive Holt-Winters Model

If a time series has a linear trend with a fixed growth rate, β_1, and a fixed seasonal pattern, S_t with constant variation, then the time series may be described by the model.

$$Y_t = (\beta_0 + \beta_1 t) + S_t + \varepsilon_t$$

For this model, the level of the time series at time t-1 is $T_{t-1} = \beta_0 + \beta_1(t-1)$ and at time t is $T_t = \beta_0 + \beta_1 t$. Hence, the growth rate in the level from one time period to the next is β_1. Y_t is the observed data at time t and ε_t is an error at time t.

The estimate $\hat{T}_t$ for the level at time t, the estimate b_t for the growth rate at time t, and the estimate $\hat{S}_t$ for the seasonal factor at time t are given by the smoothing equations (Bowerman et al. 2005)

$$\hat{T}_t = \hat{T}_{t-1} + b_{t-1} + \alpha e_t$$

$$b_t = b_{t-1} + \alpha\gamma e_t$$

$$\hat{S}_t = \hat{S}_{t-L} + (1-\alpha)\delta e_t$$

The estimate e_t for the error of the time series in time period t where α, γ, δ are smoothing constants between 0 and 1. $\hat{T}_{t-1}$ and b_{t-1} are estimate in time period t-1 for the level and growth rate, $\hat{S}_{t-L}$ is the estimate in time period t-L for the seasonal factor, and L denotes the number of seasons per year.

Multiplicative Holt-Winters Model

If a time series has a linear trend with a fixed growth rate, β_1, and a fixed seasonal pattern, S_t with increasing variation, then the time series may be described by the model.

$$Y_t = (\beta_0 + \beta_1 t) \times S_t \times \varepsilon_t$$

The smoothing equations for the multiplicative Holt-Winters model are:

$$\hat{T}_t = \hat{T}_{t-1} + b_{t-1} + \frac{\alpha e_t}{S_{t-L}}$$

$$b_t = b_{t-1} + \frac{\alpha\gamma e_t}{S_{t-L}}$$

$$\hat{S}_t = \hat{S}_{t-L} + \frac{(1-\alpha)\delta e_t}{T_t}$$

The Holt-Winters method requires starting values in which this research separate starting values into 2 methods. The first method used the first year of data to calculate $\hat{T}_t, \hat{S}_t$ and b_t which separated into 5 different patterns which level component, and seasonal factor for additive and multiplicative model defined as follows:

$$\hat{T}_L = \frac{(Y_1 + Y_2 + \cdots + Y_L)}{L}$$

$$\hat{S}_i = Y_i - \hat{T}_L \quad ; i = 1, \ldots L$$

$$\hat{S}_i = \frac{Y_i}{\hat{T}_L} \quad ; i = 1, \ldots L$$

The growth rate of each pattern is defined in Table 1. The second one employed Decomposition method which spend the data from the first set 2-18 years (T=2,…,18) in calculating trend and seasonal variation by employed ratio to moving average method to decompose. Then, we get b_0, b_1 and $\hat{S}_t$ from linear trend and seasonal factor, then applied.

$$\hat{T}_L = b_0 + b_1 \times L$$

$$b_L = b_1$$

This research estimated the smoothing parameters α, γ and δ to minimize the root mean square error (RMSE) by using Solver module in Microsoft Excel.

Table 1: The growth rate for each patterns

Pattern	Growth rate
1	$b_L = \frac{1}{L}\sum_{i=1}^{L}\left(\frac{Y_{i+L}-Y_i}{L}\right)$ (Hyndman and Athannasopoulos 2018)
2	$b_L = Y_2 - Y_1$ (Kalekas 2020; Suppalakpanya et al. 2019)
3	$b_L = \frac{(Y_2-Y_1)+(Y_3-Y_2)+(Y_4-Y_3)}{3}$ (Kalekas 2020; Suppalakpanya et al. 2019)
4	$b_L = \frac{Y_L - Y_1}{L-1}$ (Kalekas 2020; Suppalakpanya et al. 2019)
5	$b_L = 0$ (Kalekas 2020; Suppalakpanya et al. 2019)

Box-Jenkins method

The Box-Jenkins method was developed in 1974 and it is widely used in modelling and forecasting the number of airport passengers. Box-Jenkins method uses Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) to identify the model under the stationary condition. Box-Jenkins method is a four-step process (Bowerman et al. 2005).

Step 1: Tentative Identification: historical data are used to identify an appropriate Box-Jenkins model.

Step 2: Estimation: historical data are used to estimate the parameters of the identified model.

Step 3: Diagnostic checking: various diagnostics are used to check the adequacy of the identified model. In some cases may need to suggest an improved model, which is then regarded as a new identified model.

Step 4: Forecasting: once a final model is obtained, it is used to forecast future time series values.

It is possible that several models may be identified, and the selection of an optimum model is necessary. Such selection of models is usually based on the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) defined respectively as follows:

$AIC = 2\ln L + 2k$ and $BIC = 2\ln L + \ln(n)k$

Where L represents the likelihood function, k is the number of estimated parameters from the model and n is the number of residuals that can be computed for the time series. The optimum model gives the minimum AIC and SBC.

Neural Networks

Artificial Neural Networks (ANN) are designed to mimic the characteristics of the biological neurons in the human brain and nervous system (Jiang et al. 2011). In the case of modelling first year premium of life insurance time series, the historical incidence are sent into the input neurons, and corresponding forecasting incidence is generated from the output neurons after the network is adequately trained. The network learns the information contained in the first year premium time series by adjusting the interconnections between layers. The structure and neural networks can only be viewed in terms of the input, output and transfer characteristics. The specific interconnections cannot be seen even after the training process. There is no easy way to interpret the specific meaning of the parameters and interconnections within networks trained using the first year premium time series data. There are two advantages of employing neural networks for forecasting time series data. First, they can fully extract the complex nonlinear relationships hidden in the time series. Second, they have no assumption of the underlining distribution for the collected data (Zhang et al. 1998).

Back Propagation Neural Networks are a type of feed forward artificial neural networks. In feed-forward neural networks, the data flow is in one direction and the answer is obtained solely based on the current set of inputs. Back Propagation Neural Networks consist of an input layer, a hidden layer, and an output layer. Each layer is formed by a number of nodes, and each node represents a neuron. The upper-layer and lower-layer nodes are connected by the weights.

The first year premium of life insurance was divided into three subsets. First year premium from January, 2003 to December, 2015 was employed as the training set used for training the network. The first year premium from January, 2021 to November, 2021 was employed as the validation set. The remaining set of the series was used as the test set.

This research employed Back Propagation Neural Networks training. The number of inputs of the neural networks was determined by trend and seasonal period of the time series. In this study, the number of input nodes was selected to be 13, 26, 39 and 52. The output layer of artificial neural networks contains only one neuron representing the forecast value of the first year premium for the next month. This research employed Weka 3.8.6 in running artificial neural networks. The different learning rate was examined from 0.005 to 0.1 with 0.005 increments. The different momentum was examined from 0.1 to 0.8 with 0.05 increments. The number of hidden neurons were varied from 2 to 26 at an increment of 1.

MODEL SELECTION CRITERION

The forecasting models were chosen by considering the smallest root mean square error (RMSE) and the mean absolute percentage error (MAPE) were used to measure the accuracy of the model.

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n} e_t^2}$$

$$MAPE = \frac{1}{n}\sum_{t=1}^{m}\left|\frac{e_t}{Y_t}\right| \times 100$$

RESULTS

The first set of data from January 2003 to December 2020 is employed to build the Holt-Winters models with different initial values, Box-Jenkins models, and artificial neural networks models.

The Holt-Winters method

The results from Holt-Winters models with different initial values are shown in Table 2 and 3.

Table 2: RMSE of Holt-Winters models with five different initital values

Pattern	Additive model	Multiplicative model
1	863,466.60	837,216.40
2	856,548.71	861,502.89
3	870,418.17	851,926.00
4	856,548.71	829,798.23
5	860,619.49	832,226.44

From Table 2, the Holt-Winters models with initial values from pattern 4 obtained the minimum RMSE for both additive and multiplicative models. Where multiplicative model yields the minimum RMSE (829,798.23).

Table 3: RMSE of Holt-Winters models with initial values from Decomposition method.

Number of years (T)	Additive model	Multiplicative model
2	836,940.92	844,950.91
3	828,661.13	834,436.60
4	824,061.29	831,282.64
5	820,765.35	824,791.16
6	819,956.77	821,765.97
7	818,993.45	813,264.40
8	816,004.33	809,525.52
9	814,967.62	809,131.77
10	814,439.14	803938.39
11	813,530.57	793,130.86
12	815,220.53	773,647.63
13	815,438.55	773,636.17
14	814,388.20	775,593.93
15	812,695.82	773,822.60
16	810,527.82	773,210.66
17	810,018.92	765,074.48
18	807,140.72	756,342.63

From Table 3, The RMSE of additive and multiplicative models are decreased while using more data in calculating trend and seasonal factor. The multiplicative Holt-Winters model with initial values from 18 years decomposition method yields minimum RMSE (756,342.63).

Box-Jenkins method

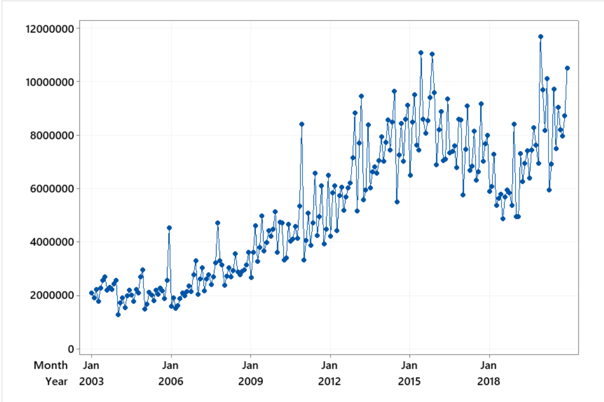

Figure 1: First year premium of life insurance from January 2003 to December 2020.

From Figure 1, the non-stationary is visible since the first year premium have non-linear trend and seasonal variation. One regular difference (d=1) and one seasonal difference (D=1) are taken to make time series stationary. Based on Autocorrelation Function (ACF) as shown in Figure 2 and Partial Autocorrelation Function(PACF) as shown in Figure 3 ACF are exponentially decayed, and PACF are significant spike at lag 3 then it suggests ARIMA(3,1,0). Additionally, PACF are significant spike at lag 12, and 48 suggest a seasonal SARIMA(4,1,0) However, ARIMA(3,1,0)xSARIMA$(4,1,0)_{12}$ model do not pass the diagnostic checking. Therefore, the improved model ARIMA(4,1,0)xSARIMA$(4,1,0)_{12}$ is suggested.

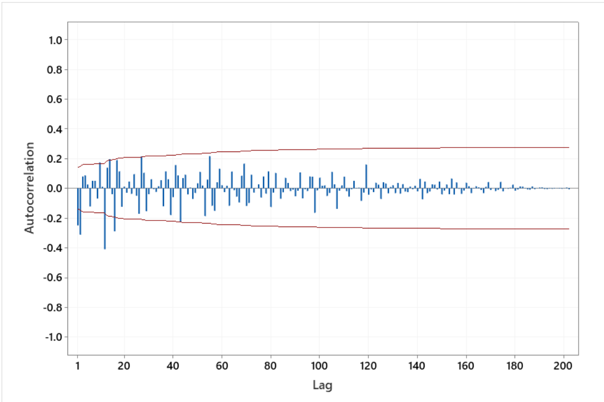

Figure 2: Autocorrelation Function of first year premium of life insurance with one regular difference and one seasonal difference.

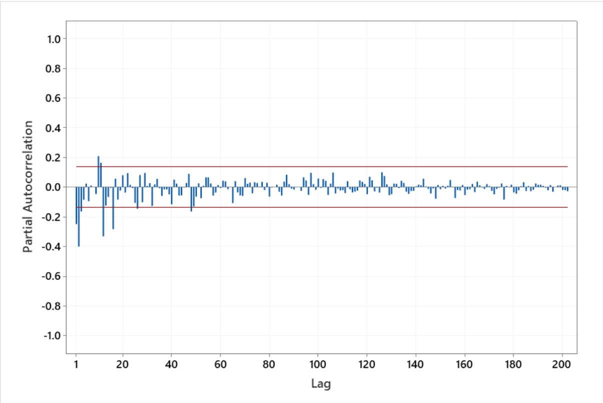

Figure 3: Partial Autocorrelation Function of first year premium of life insurance with one regular difference and one seasonal difference.

Table 4: Minitab output of Box-Jenkins model of ARIMA(4,1,0)xSARIMA(4,1,0)$_{12}$

Type	Coef	SE Coef	t-value	p-value
AR1	-0.5238	0.0720	-7.28	0.000
AR2	-0.5947	0.0789	-7.54	0.000
AR3	-0.2187	0.0793	-2.76	0.006
AR4	-0.1881	0.0720	-2.61	0.010
SAR12	-0.6719	0.0728	-9.23	0.000
SAR24	-03970	0.0865	-4.59	0.000
SAR36	-0.4860	0.0927	-5.24	0.000
SAR48	-0.4705	0.0858	-5.49	0.000
Modified Box-Pierce(Box-Ljung)Chi-Square statistic				
Lag	12	24	36	48
Chi-Square	7.39	22.15	36.71	45.62
DF	4	16	28	40
p-value	0.117	0.138	0.125	0.250

Table 4 shows all parameters $(\phi_1, \phi_2, \phi_3, \phi_4, \Phi_{12}, \Phi_{24}, \Phi_{36}, \Phi_{48})$ from ARIMA(4,1,0)xSARIMA(4,1,0)$_{12}$ model are statistically significant from zero (p-value is less than 0.05). From Box-Ljung test, residuals from the model are independent (p-value is greater than 0.05). Therefore, the model ARIMA(4,1,0)xSARIMA(4,1,0)$_{12}$ fits with first year premium of life insurance. In addition, ARIMA(0,1,3)xSARIMA(5,1,0)$_{12}$ and ARIMA(2,1,3)x SARIMA(5,1,0)$_{12}$ also pass the diagnostic checking. Table 3 shows all Box-Jenkins models with RMSE, AIC and SBC.

Table 5:Box-Jenkins models with RMSE, AIC and SBC

RMSE	AIC	SBC
ARIMA(0,1,3)xSARIMA(5,1,0)$_{12}$ model		
783,809.29	70.29	96.47
ARIMA(4,1,0)xSARIMA(4,1,0)$_{12}$ model		
794,947.80	70.34	96.53
ARIMA(2,1,3)xSARIMA(5,1,0)$_{12}$ model		
779,229.75	74.26	106.89

From Table 5, based on the minimum AIC and SBC values, the ARIMA(0,1,3)xSARIMA(5,1,0)$_{12}$ model are the optimum model for first year premium of life insurance.

Artificial Neural Networks

Increasing the input nodes, the RMSE of training set is decreasing significantly. However, the RMSE of testing set is increasing. To avoid the overfiting, the model which has closest training and testing RMSE is choosen. The optimal model for ANN is 13-4-1 with learning rate of 0.01, momentum of 0.6 and 500 iterations. The RMSE of the model are shown in Table 6.

Table 6: The RMSE of training set, test set and validation set.

Model	RMSE		
	Training set	Test set	Validation set
13-4-1	737,179.1	1,060,518.2	1,408,953.9

Table 7: The RMSE of three forecasting methods.

Forecasting model	RMSE
Multiplicative model with initial value pattern 4	829,798.23
Multiplicative model with initial values from 18 years Decomposition method	756,342.63
ARIMA(0,1,3)xSARIMA(5,1,0)$_{12}$	783,809.29
Artificial Neural Networks	1,060,518.23

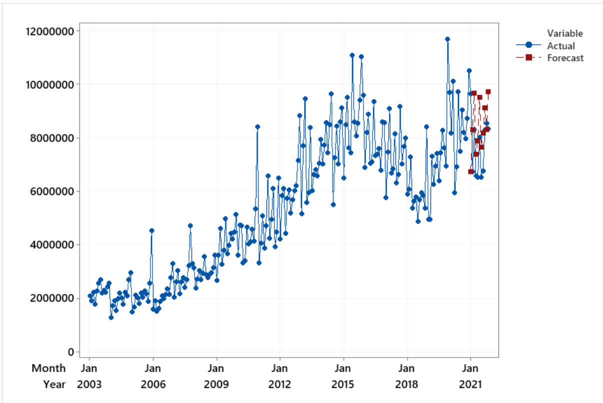

Figure 4. Actual and forecast by the multiplicative model with initial values from 18 years Decomposition method.

Model Selection and Performance Measure

From Table 7, the multiplicative model with in initial values from 18 years Decomposition method obtains the smallest RMSE. This research confirms that the different initial value can improve the efficient of Holt-Winters method significantly. Although the Holt-Winters method is suitable for time series that have linear trend and seasonal variation. This research shows that the different initial value can increase the efficiency of Holt-Winters method even though the time series have non-linear trend and seasonal variation as first year premium of life insurance.

CONCLUSION

This research presents three different forecasting methods which are Holt-Winters method, Box-Jenkins method and Artificial Neural Networks to model the first year premium of life insurance. The results showed that the multiplicative model with initial values from 18 years Decomposition method give the appropriate model for the first year premium of life insurance and yields the MAPE = 17.29% The Holt-Winters method with different initial value is simpler and require less data to estimate than Box-Jenkins method and Artificial Neural Networks. However, its performance can be outperformed both Box-Jenkins method and Artificial Neural Network. The increased complexity of Box-Jenkins method and Artificial Neural Networks do not guarantee the better results. As shown in this research, the simpler Holt-Winters method with different initial value can give higher accuracy in some cases. When the researcher models time series data with trend and seasonal factor, Holt-Winters method with different initial value is a good choice.

ACKNOWLEDGMENT

This work was supported by School of Science, King Mongkut's Institute of Technology Ladkrabang Research Fund. Grant no: 2564-02-05-002.

REFERENCES

Booranawong T. and Booranawong A. 2018. "Double exponential smoothing and Holt-Winters methods with optimal initial values and weighting factors for forecasting lime, Thai chili and lemongrass prices in Thailand". *Engineering and Applied Science Research*, 45(1), 32-38.

Bahia, Itedal Sabri Hashim 2013. "Using Artificial Neural Netrork Modeling in Forecasting Revenue: Case Study in National Insurance Company/Iraq". *International Journal of Intelligence Science*, Vol 3, 136-143.

Bowerman, B. L.; Richard T. O'; Connell and Anne B. Koehler. 2005. "Forecasting, Time Series, and Regression: An Applied Approach". 4th ed., Thomson Brooks/Cole. USA.

Chatfield C. 1996. "The Analysis of Time Series". 5th ed., Chapman & Hall, New York.

Hyndman R. J. and Athanasopoulos G. 2018. "Forecasting Principle and Practice". 2nd ed., OTexts: Melbourne, Australia.

Jiang L-H; Wang A-G; Tian N-Y; Zhang W-C; Fan Q-L. 2011. "BP Neural Network of Continuous Casting Technological Parameter and Secondary Dendrite Arm Spacing of Spring Steel". *Journal of Iron and Steel Research*, Vol 18, 25-29.

Kalekar, P.S. 2004. "Time Series Forecasting using Holt-Winters Exponential Smoothing"

Namawejje, H. and Geofrey, K. 2020. "Forecasting Life Insurance Premiums and Insurance Penetration Rate in Uganda". *International Journal of Scientific and Research Publications*, Vol 10, Issue 5, 708-718.

Safyar A. 2010. "Forecasting Life Insurance Premium by Arima Model s and Neural Network." *SANAAT-E-BIMEH.* Vol 25, no 2(98),121-138.

Suppalakpanya, K.; Nikhom, R.; Booranawong A.; Booranawong T. 2019. "An Evaluation of Holt-Winters Methods with Different Initial Trend Values for Forecasting Crude Palm Oil Production and Prices in Thailand". *Suranaree Journal of Science and Technology*, 26(1), 13-22.

Thai Life Assurance Association, 2019. "Overview of Thai Life Insurance Business 2018-2019 and Trends of 2020" Annual Report 2019, 78-110.

Zhang G.; Eddy Patuwo B.; Y. Hu M. 1998. "The State of the Art" *International Journal of Forecasting*, Vol 14, 35-62.