

# Factors Analysis on Stock Exchange of Thailand (SET) Index Movement

Phaisarn Sutheebanjard  
Graduate School of Information Technology  
Siam University  
Bangkok 10163, Thailand  
mr.phaisarn@gmail.com

Wichian Premchaiswadi  
Graduate School of Information Technology  
Siam University  
Bangkok 10163, Thailand  
wichian@siam.edu

**Abstract**— This paper proposes the prediction method for Stock Exchange of Thailand index (SET index). The proposed method was adapted from [1] by adding more important factors to the prediction function. This research takes into account both external and internal factors to forecast the SET index. The external factors are the historical movement of the world's major stock exchange market indices such as Dow Jones, Nikkei and Hang Seng index, while the internal factors are the historical movement of the SET index and Minimum Loan Rate (MLR). The experimental result on both short and long term data show that the proposed method has lower errors for both Mean Squared Error (MSE) and Mean Absolute Percentage Error (MAPE) measures. In addition, the proposed function can be put into the spreadsheet which makes it easier and faster for the investors to analyze the trend of stock exchange market.

**Keywords**—*SET index, Stock Exchange of Thailand, Thai stock market, Stock Exchange, Evolution Strategies, stock prediction function*

## I. INTRODUCTION

The analysis of stock market index movement is considered as a difficult and complex task. There are many factors that involve in predicting the direction of the Stock market index movement such as inflation rates, economic environments, political issues, and so on. In terms of stock exchange of Thailand, SET index mostly influences from the world's major stock market indices such as Hang Seng, Nikkei and Dow Jones.

This research extends the prediction function of [1] and applies the evolution strategies (1+1)-ES [2] in deriving the new prediction function which can forecast the SET index more accurately. The SET index prediction function uses the historical movement of the SET index itself as well as the world's major stock market indices which include Dow Jones (New York), Nikkei (Japan), Hang Seng (Hong Kong) including the Minimum Loan Rate (MLR) from Bank of Thailand.

Evolution Strategies (ES) was introduced by Rechenberg [4]. It is similar to genetic algorithm. The idea is to use the principles of organic evolution processes as the rules for optimum seeking procedures. The (1+1)-ES is a simple mutation-selection scheme called *two-membered* ES. It is based upon a "population" consisting of one parent individual (a real valued vector), and one descendant, created by means of

adding normally distributed random numbers with identical standard deviations. The resulting individual was evaluated and compared to its parent, and the better of both individuals survives to become parent of the next generation, while the other one is discarded.

This paper is organized as follow: Section II describes stock market prediction theories and methods. Section III shows the previous work on the Stock Exchange of Thailand index prediction. Section IV describes the solution technique using evolution strategies. Section V shows the experimental results. A conclusion is drawn in section VI.

## II. STOCK MARKET PREDICTION

### A. Stock Market Prediction Theories

In predicting stock market movement, two theories have had significant impact on market research, Efficient Market Hypothesis (EMH) and Random Walk Theory.

1) *Efficient Market Hypothesis (EMH)*: was an idea partly developed in the 1965 by Fama [7]. In EMH, the price of a security is a reflection of complete market information. Whenever a change in financial outlook occurs, the market will instantly adjust the security price to reflect the new information. This means that given the information, no prediction of future change in the price can be made. This is also a highly controversial and often disputed theory. Supporters of this model believe it is pointless to search for undervalued stocks or try to predict trends in the market through fundamental analysis or technical analysis. EMH contained three different levels of information sharing: the weak form, the semi-strong and the strong form. Within weak EMH, only historical information is embedded in the current price. The semi-strong form goes a step further by incorporating all historical and currently public information into the price. The strong form includes historical and current public information as well as private information such as insider information, in the share price. But in the 1991, Fama [8] stated that the strong form of the EMH was surely false.

2) *Random Walk Theory*: A random walk is one in which future steps or directions cannot be predicted on the basis of past actions. When the term is applied to the stock market, it means that the past movement or direction of the price of a

stock or overall market cannot be used to predict its future movement. The stock's future prices take a random and unpredictable path, it going up is the same as it going down.

### B. Stock Market Prediction Approaches

Efficient Market Hypothesis and Random Walk Theory have discouraged predicting stock's future prices, however there are still many methods that try to predict the stock's future prices. These methods can be grouped into two diametrically opposed approaches as Fundamental and Technical analysis.

1) *Fundamental Analysis*: are concerned with the company that underlies the stock itself. They evaluate a company's past performance as well as the credibility of its accounts. Many performance ratios are created that aid the fundamental analyst with assessing the validity of a stock, such as the P/E ratio. The goal of analyzing a company's fundamentals is to find a stock's intrinsic value, as opposed to the value at which it is being traded in the marketplace. If the intrinsic value is more than the current share price, your analysis is showing that the stock is worth more than its price and that it makes sense to buy the stock.

2) *Technical Analysis*: or chartists are not concerned with any of the company's fundamentals. They seek to determine the future price of a stock based solely on the (potential) trends of the past price (a form of time series analysis). Numerous patterns are employed such as the head and shoulders or cup and saucer. Alongside the patterns, statistical techniques are utilized such as the Time Series Analysis. Technical analysts test historical data to establish specific rules for buying and selling securities with the objective of maximizing profit and minimizing risk of loss. Technical trading analysis is based on two main premises. First, the market's behavior patterns do not change much over time, particularly the long-term trends. The patterns in market prices are assumed to recur in the future, and thus, these patterns can be used for predictive purposes.

### III. THE STOCK EXCHANGE OF THAILAND PREDICTION

There were some researchers proposed the stock price prediction for Stock Exchange of Thailand (SET) index.

In 2006, Meesad and Srikhacha [3] proposed an algorithm based on exponential smoothing and adaptive approach. Their work was proposed for the day-to-day stock price time series prediction. The system was designed to predict *Close* prices of each next bidding day. Their data were collected during bidding day period 4 January 2005 to 4 February 2005, 24 records. They used MSE to measure the error. MSE for SET index was 16.335 in test data. Their method is good for short term stock prediction with low error rate, but they do not mention about the long term stock prediction.

In 2005, Rimcharoen, et al. [1] proposed an adaptation of evolution strategies with evolving functional form and coefficients to predict the Stock Exchange of Thailand index. The potential parameters that drive the stock exchange index, namely Dow Jones index, Nikkei index, Hang Seng index, gold price and MLR were used in the algorithm by using the adaptive evolution strategies. It is not necessary to determine

the functional form of the prediction function a priori. The experimental data show that their method can be effectively used to forecast the Stock Exchange of Thailand index with error less than 3%. Function of Rimcharoen et al. is shown in (1).

$$\begin{aligned} SET_{(t)} = & 2.3645 + 5.5208\sin^3[0.3138_{(t-1)}] - \\ & 1.5430\text{hangseng}_{(t-1)} / -5.2054\text{mlr}_{(t-1)} + \\ & 2.8360\cos^2[0.6246_{(t-1)}] * 4.6811\sin[0.3651_{(t-1)}] - \\ & 1.5380\cos^3[0.7522_{(t-1)}] - 1.1618\cos^3[0.7724_{(t-1)}] + \\ & 3.3228\sin^3[1.5317_{(t-1)}] - 2.4620\cos[0.6676_{(t-1)}] * \\ & 2.3144\text{mlr}_{(t-1)} \end{aligned} \quad (1)$$

where (t) is today and (t-1) is yesterday.

They have eventually found that the SET index can be adequately explained by only two major factors, Hang Seng index and MLR as (2). The graph of (2) is show in figure 1.

$$SET_{(t)} = \frac{-1.5430\text{hangseng}_{(t-1)}}{-5.2054\text{mlr}_{(t-1)}} \quad (2)$$

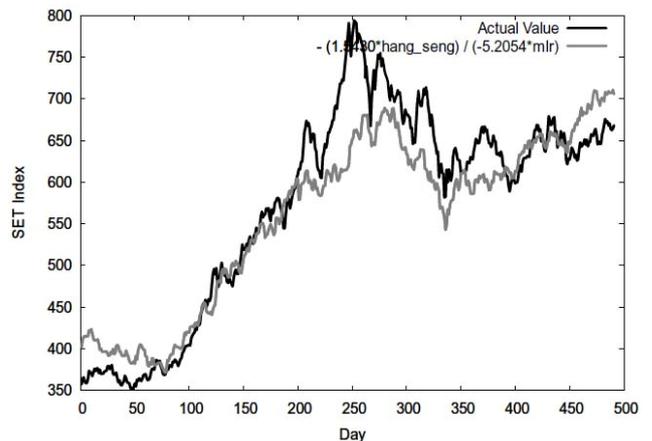


Figure 1. The graph of  $(-1.5430\text{hangseng})/(-5.2054\text{mlr})$  [1]

[1] Used MAPE to measure the error of the forecasted value. The research shows that their method is good for the long term SET index prediction with less than 3% error.

This research is intentionally conducted to find a better methodology for long term SET index prediction with less than 2% error.

This research takes into account both internal and external factors to forecast the SET index. The external factors are foreign major stock market indices, while the internal factors are SET index and MLR. From the assumption that both external and internal factors probably have great impact on the SET index, these factors include:

- SET index (Thailand)
- Dow Jones index (New York)
- Nikkei index (Japan)
- Hang Seng index (Hong Kong)
- Minimum Loan Rate (MLR)

In the prediction process, the parameters of the previous time period (yesterday) are used to predict today's stock exchange index as (3).

$$SET_{(t)} = F_{(t-1)} \quad (3)$$

where  $SET_{(t)}$  is the stock exchange index at day t, and  $F_{(t-1)}$  is a prediction function of day t-1.

#### IV. EVOLUTION STRATEGIES

Similar to Genetic Algorithms, Evolution Strategies (ES) are algorithms which imitate the principles of natural evolution as a method to solve parameter optimization problems. Evolutionary strategies (ES) was introduced by Ingo Rechenberg [4] for optimizing real value parameters. In ES, the representation used was a fixed-length real-valued vector. A vector of real values represented an individual. The standard deviation was used to control search strategy in ES. The main operator in ES was Gaussian mutation, in which a random value from a Gaussian distribution (normal distribution) was added to each element of an individual's vector to create a new offspring.

This research uses (1+1)-ES (*two-membered* ES) for the selection process. The (1+1)-ES consists of one parent individual (a real-valued vector), which produces one offspring by means of adding normal distribution random numbers. The better of both individual then serves as the ancestor of the following iteration/generation. An introduction to ES can be found in [4-6].

In forecasting task, the researcher applied the function of Rimcharoen et al. [1] by adding more important factors as shown in (4). The (1+1)-ES was used to find the coefficients of function.

$$SET_{(t)} = a_0 SET_{(t-1)} + a_1 \left( \frac{a_2 DJ_{(t-1)} + a_3 NK_{(t-1)} + a_4 HS_{(t-1)}}{a_5 MLR_{(t-1)}} \right) \quad (4)$$

where  $a_0$ - $a_5$  denote coefficients.

$SET$  is SET index (Thailand)  
 $DJ$  is Dow Jones index (New York)  
 $NK$  is Nikkei index (Japan)  
 $HS$  is Hang Seng index (Hong Kong)  
 $MLR$  is Minimum Loan Rate (MLR)

After the coefficient of prediction function was initialized, the evolution process began. The fitness function was used to evaluate an individual. This research used Mean Squared Error (MSE) in order to minimize error of fitting between the prediction function and actual value as shown in (5).

$$MSE = \frac{1}{n} \sum_{t=1}^n (g_{(t)} - f_{(t)})^2 \quad (5)$$

where  $g_{(t)}$  is an actual value.

$f_{(t)}$  is a forecasted value

$n$  is the number of data points.

In controlling the search strategy, an adjustment of standard deviation was considered and taken from the ratio of a better individual during the evolution process, refer to 1/5 success rule [4] as shown in (6).

$$\sigma' = \begin{cases} \sigma / 0.817 & \text{if } (p > 1/5) \\ \sigma \cdot 0.817 & \text{if } (p < 1/5) \\ \sigma & \text{if } (p = 1/5) \end{cases} \quad (6)$$

This research used Mean Squared Error (MSE) and Mean Absolute Percentage Error (MAPE) to measure the error of the forecasted value and actual value. MSE is one of many ways to quantify the amount by which an estimator differs from the actual value of the quantity being estimated. MSE measures the average of the square of the error. The error is the amount by which the estimator differs from the quantity to be estimated. MAPE expresses the error as a percentage. MAPE is commonly used in quantitative forecasting methods because it produces a measure of relative overall fit. The absolute values of all the percentage errors are summed up and the average is computed. The function of MAPE was shown in (7).

$$MAPE = \frac{\sum_{t=1}^n \left| \frac{g_{(t)} - f_{(t)}}{g_{(t)}} \right|}{n} \times 100 \quad (7)$$

where  $g_{(t)}$  is an actual value.

$f_{(t)}$  is a forecasted value

$n$  is the number of data points.

The mutations of a real value coefficient are accomplished by sampling a value from normal distribution and adding that to the coefficient. The implemented pseudo code is shown in figure 2.

1. Randomly assign standard deviation.
2. Create parent
  - 2.1 Randomly assign coefficients ( $a_1$ - $a_5$ ) of (4) with normal distribution.
  - 2.2 Evaluate the fitness by using MSE.
3. Create new offspring.
  - 3.1 Mutate and sum up ( $a_1$ - $a_5$ ).
 
$$a'_x = a_x + N(0, \sigma^2)$$

where  $x$  is 1-5
  - 3.2 Adjust standard deviation value by applying 1/5 success rule
  - 3.3 Evaluate the fitness by using MSE.
4. Select the better between parent and offspring to the next generation.
5. Repeat step 3 through 4 until 1,000 generations.

Figure 2. pseudo code.

TABLE I  
SHORT TERM PREDICTION EXPERIMENTAL RESULTS

No	Train		Test		$a_0$ ( $SET_{(t-1)}$ )	$a_1$	$a_2$ ( $DJ_{(t-1)}$ )	$a_3$ ( $NK_{(t-1)}$ )	$a_4$ ( $HS_{(t-1)}$ )	$a_5$ ( $MLR_{(t-1)}$ )
	MSE	MAPE(%)	MSE	MAPE(%)						
1	36.21	0.77	13.16	0.38	0.97	-0.005193	0.632565	-1.179130	0.455504	0.024606
2	47.75	0.89	13.21	0.38	0.94	-0.008849	0.628276	-1.184474	0.447989	0.029061
3	33.50	0.73	13.60	0.39	0.96	-0.006621	0.621288	-1.186140	0.450772	0.034985
4	27.40	0.65	13.92	0.39	0.96	0.160509	-0.364305	0.752382	-0.286521	0.828879
5	35.72	0.76	13.99	0.41	0.93	-0.007072	0.623421	-1.250866	0.468919	0.033811
6	40.45	0.81	14.06	0.38	1.06	1.530926	-0.177532	0.248756	-0.085295	1.354410
7	28.12	0.66	14.23	0.39	0.94	-0.975534	0.026066	-0.176908	0.100268	1.545628
8	27.20	0.66	14.39	0.38	0.94	-0.047632	0.099138	-1.089981	0.632183	0.549168
9	28.28	0.65	14.58	0.39	0.96	0.511800	-0.010432	0.194298	-0.129625	1.077547
10	27.33	0.65	14.60	0.39	0.95	0.507374	-0.013957	0.195559	-0.121431	1.089788

## V. EXPERIMENTAL RESULTS

The experimental data obtained from the Bank of Thailand consisted of historical data of SET index, Dow Jones index, Nikkei index, Hang Seng index, and Minimum Loan Rate.

The experimental results were compared with some of the well known methods for both short term and long term prediction. The MSE and MAPE of Simple Moving Average, Yesterday SET index, [1] and [3] were computed and compared with the proposed function as follows:

The simple moving average is a simple technique in time series forecasting. The weights of simple moving average used in this research were  $0.5_{(t-1)}$ ,  $0.3_{(t-2)}$  and  $0.2_{(t-3)}$  as shown in (8).

$$SET_{(t)} = 0.5SET_{(t-1)} + 0.3SET_{(t-2)} + 0.2SET_{(t-3)} \quad (8)$$

The error from simple technique like yesterday's SET index ( $SET_{(t-1)}$ ) was also analyzed as shown in (9).

$$SET_{(t)} = SET_{(t-1)} \quad (9)$$

### A. Short Term Prediction

The time series investigated in short term experiment started from January 4, 2005 to February 4, 2005. It contained 24 days of data. The data were divided into two groups: 9 days for training and 15 days for testing. The data period used in this experiment was the same as [3].

1) *Prediction by Evolution Strategies:* For each value of  $a_0$  from 0.01 – 1.06, the coefficient  $a_1$ - $a_5$  from (4) were assigned randomly to the mutation process which performed repeatedly 1,000 times. This process was repeated for 100 times, each time the new random value of  $a_1$ - $a_5$  were assigned to each value of  $a_0$  in order to compare the MSE and MAPE. The ten lowest MSE values are shown in Table I. The value of  $SET_{(t)}$  of function number one from table I can be calculated as shown in (10).

From Table I, all ten functions show that  $a_3$  (coefficient of Nikkie index) has the positive impact on SET index (From (4), when  $a_1$  multiply to  $a_2$ - $a_4$ , the sign of  $a_2$ - $a_4$  will change).  $a_2$

(coefficient of Dow Jones index) and  $a_4$  (coefficient of Hang Seng index) have the negative impact on SET index.

$$SET_{(t)} = 0.97SET_{(t-1)} - 0.005193 \times ((0.632565DJ_{(t-1)} - 1.179130NK_{(t-1)} + 0.455504HS_{(t-1)}) / 0.024606MLR_{(t-1)}) \quad (10)$$

2) *Comparison:* The error as measured by MSE and MAPE of the proposed function (10), the Simple Moving Average (8), Yesterday SET index (9) and Meesad and Srikhacha [3] were shown in table II. The result shows that the proposed function is the best prediction function to forecast the SET index in short term prediction with the lowest MSE and MAPE errors.

TABLE II SHORT TERM ERROR COMPARISON

Method	MSE	MAPE (%)
Proposed function (10)	13.16	0.38
Simple Moving Average (8)	35.57	0.68
Yesterday SET index (9)	19.38	0.45
Meesad and Srikhacha [3]	16.335	-

### B. Long Term Prediction

The time series investigated in the long term experiment started from January 2003 to December 2004. It contained 491 days of data. The data were divided into two groups: 420 days for training and 71 days for testing. The data period used in this experiment was the same as [1].

1) *Prediction by Evolution Strategies:* For each value of  $a_0$  from 0.01 – 1.06, the coefficient  $a_1$ - $a_5$  from (4) were assigned randomly to the mutation process which performed repeatedly 1,000 times. This process was repeated for 100 times, each time the new random value of  $a_1$ - $a_5$  were assigned to each value of  $a_0$  in order to compare the MSE and MAPE. The ten lowest MSE values are shown in Table III. The value of  $SET_{(t)}$  of function number one from table III can be calculated as shown in (11). The MSE and MAPE on the test data can be represented as graphs in figure 3 and figure 4 respectively.

TABLE III  
LONG TERM PREDICTION EXPERIMENTAL RESULTS

No	Train		Test		$a_0$ ( $SET_{(t-1)}$ )	$a_1$	$a_2$ ( $DJ_{(t-1)}$ )	$a_3$ ( $NK_{(t-1)}$ )	$a_4$ ( $HS_{(t-1)}$ )	$a_5$ ( $MLR_{(t-1)}$ )
	MSE	MAPE(%)	MSE	MAPE(%)						
1	116.65	1.50	55.49	0.88	0.86	0.021008	0.632529	0.317615	0.156434	0.509522
2	115.82	1.48	55.53	0.88	0.86	0.024501	0.611974	0.278751	0.164192	0.568226
3	117.17	1.49	55.56	0.88	0.86	0.021186	1.037871	0.717827	0.286161	0.951959
4	157.74	1.79	55.59	0.88	0.81	0.027236	1.008707	0.741167	0.333679	0.926067
5	160.09	1.79	55.73	0.88	0.81	-0.183472	-0.495648	-0.428838	-0.152022	3.221517
6	161.90	1.79	55.80	0.88	0.81	0.007578	1.486542	1.762930	0.504306	0.465326
7	162.51	1.74	55.85	0.88	0.81	0.041973	0.332829	0.835980	0.223377	0.968337
8	148.87	1.62	56.07	0.88	0.81	0.130244	0.135020	0.144157	0.111412	0.859299
9	217.14	2.07	56.10	0.88	0.76	-0.103523	-0.233277	-0.567247	-0.171876	1.323922
10	120.15	1.48	56.11	0.89	0.86	0.063879	0.551800	0.671816	0.151064	1.939403

$$SET_{(t)} = 0.86SET_{(t-1)} + 0.021008 \times ((0.632529DJ_{(t-1)} + 0.317615NK_{(t-1)} + 0.156434HS_{(t-1)}) / 0.509522MLR_{(t-1)}) \quad (11)$$

For the top five functions in table III,  $a_2$  has most impact on SET index, followed by  $a_3$  and  $a_4$ . For the last five functions,  $a_3$  has most impact on SET index, followed by  $a_2$  and  $a_4$ .

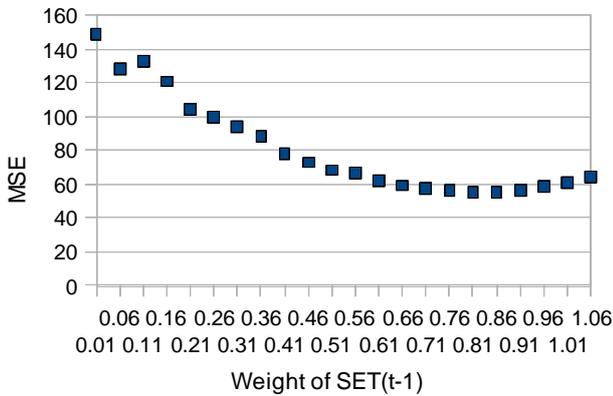


Figure 3. MSE of equation 4 at different  $a_0$  scale from 0.01-1.06

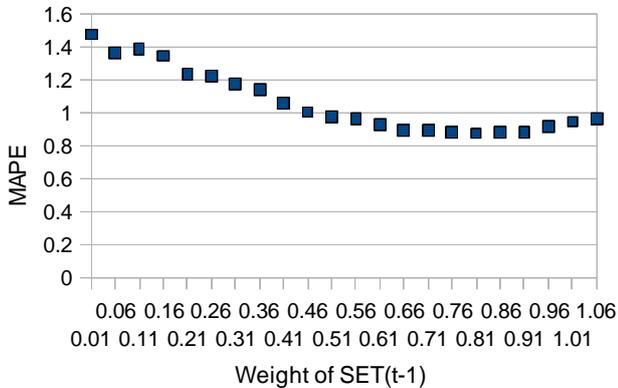


Figure 4. MAPE of equation 4 at different  $a_0$  scale from 0.01-1.06

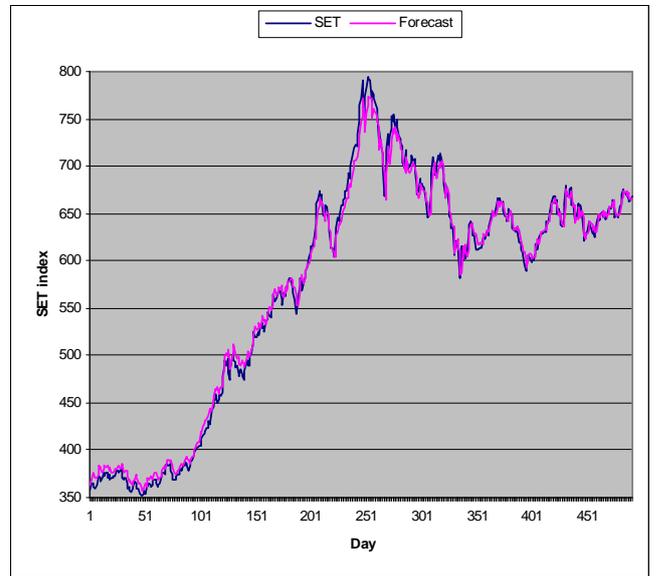


Figure 5. SET index comparison graph

Figure 5 illustrates the graph of actual SET index from the period of January 2003 to December 2004 compared to the SET index calculated by applying prediction function number one from table III. The SET index from prediction function is nearly the same as the actual SET index.

2) *Comparison:* The error as measured by MSE and MAPE of the proposed function (11), the Simple Moving Average (8), Yesterday SET index (9) and Rimcharoen et al. [1] were shown in table IV. The result shows that the proposed function is the best prediction function to forecast the SET index in long term prediction with the lowest MSE and MAPE errors.

TABLE IV LONG TERM ERROR COMPARISON

Method	MSE	MAPE (%)
Proposed function (11)	55.49	0.88
Simple Moving Average (8)	81.53	1.02
Yesterday SET index (9)	62.37	0.96
Rimcharoen et al. [1]	-	2.46

3) *Prediction on January 2005-March 2009*: Ten trained prediction functions from table III were used to forecast the SET index in January 2005 until March 2009 time period. The results show that the MAPE is less than 2% as shown in table V. Surprisingly, the simple functions with trained data of 420 days can forecast the SET index for the next 4 years more precisely. Therefore, it means that the factors in (4) have nearly the same effect to SET index in the time period of January 2003 until March 2009 (the time observed in this experiment).

TABLE V  
EXPERIMENTAL RESULT ON JANUARY 2005 – MARCH 2009

No	Test	
	MSE	MAPE(%)
1	220.59	1.65
2	221.03	1.66
3	205.45	1.58
4	292.25	1.90
5	295.50	1.90
6	276.12	1.82
7	221.33	1.62
8	220.41	1.68
9	293.75	1.88
10	195.66	1.51

## VI. CONCLUSION

This research proposed method to forecast Stock Exchange of Thailand index (SET Index) with lower errors than well-known approaches. The proposed method can be used for both short term and long term stock prediction with Mean Absolute Percentage Error (MAPE) less than 2%. The experimental results yield the meaningful prediction functions. They can suggest the positive and negative factors, which have an influence on SET index movement. Furthermore, the research results indicates that for long term prediction, the external factors as Dow Jones index<sub>(t-1)</sub> and Nikkei index<sub>(t-1)</sub> are the most important and have most impact on SET index. On the other hand, Hang Seng index<sub>(t-1)</sub> has least impact on SET index. For internal factors, SET index<sub>(t-1)</sub> is a positive factor while MLR<sub>(t-1)</sub> is an invert factor to SET index. Since this approach provides more accurate prediction on SET index movement, therefore it can be used by investors in accordance with other stock market analysis tools to forecast stock market trend and help in their decision making process. Further research can be done on finding better prediction function by taking into account more valid and effective factors to forecast the SET index with even smaller error.

## REFERENCES

[1] Sunisa Rimcharoen, Daricha Sutivong and Prabhas Chongstitvatana, "Prediction of the Stock Exchange of Thailand Using Adaptive Evolution Strategies," Tools with Artificial Intelligence, 2005. ICTAI 05. 17<sup>th</sup> 2005.

[2] Sunisa Rimcharoen and Prabhas Chongstitvatana, "An Adaptation of Evolutionary Strategies for Forecasting the Exchange Rate," The 8<sup>th</sup> Annual National Symposium on Computational Science and Engineering, 2004.

[3] Phayung Meesad and Tong Srihacha, "Universal Data Forecasting with an Adaptive Approach and Seasonal Technique," Computational Intelligence for Modelling, Control and Automation, 2006 and International Conference on Intelligent Agents, Web Technologies and Internet Commerce. 2006.

[4] Rechenberg, I. "Evolutionsstrategie: Optimierung technischer Systeme nach Prinzipien der biologischen Evolution." Stuttgart: Frommann-Holzboog Verlag. 1973.

[5] Back, T., Hoffmeister, F. and Schwefel, H.P. "A Survey of Evolution Strategies," Proceeding of the Fourth Conference on Genetic Algorithm, 1991.

[6] Beyer, H.G. and Schwefel, H.P. "Evolution Strategies-A Comprehensive Introduction," Natural Computing. Vol. 1, Issue 1, 2002

[7] Fama, E. "The Behavior of Stock-Market Prices," Journal of Business Vol. 38, No.1. 1965.

[8] Fama, E. "Efficient Capital Markets: II," Journal of Finance Vol. XLVI, No.5. 1991.