A Study of Stock Dynamism in Asian Emerging Markets after the 2008 Economic Crisis

Deng-Yiv Chiu*
Department of Information Management, Chung Hua University, Hsinchu, Taiwan, ROC
Email: chiuden@chu.edu.tw

Agus Andria
Department of Information Management, Chung Hua University, Hsinchu, Taiwan, ROC
Email: d09803024@chu.edu.tw

***Abstract*— The global financial crisis at the end of 2007 financially influenced various countries, including Indonesia. Because Indonesia achieved highest growth in the Southeast Asia region during the recession, global investors shift their investment in Indonesia market. Therefore, it is very important to explore the stock dynamism in Indonesia. We propose a hybrid approach of fuzzy theorem, support vector regression, genetic algorithm, and seasonal moving window to explore the Indonesian stock quarterly dynamism among the same quarter in continuous years using daily prices from 2006 to 2011. We find that the proposed method outperforms benchmark returns. We conclude that a hybrid approach is able to improve earning rate performances.**

***Index Terms*— emerging market, fuzzy c-means, genetic algorithm, moving window, support vector regression**

I. Introduction

The global financial crisis at the end of 2007 triggered by the collapse of major financial institutions in the U.S. began to take effect in various countries. In Indonesia, the economic crisis in the U.S. has forced investors of institutional U.S. Treasury to release their holdings in the Indonesian capital market to strengthen the liquidity of financial institutions. The value of the shares was dropped and the volume of sales of shares in the capital market in Indonesia was reduced. Jakarta composite index (JCI) was down in the fourth quarter of 2007 and it continued until the end of 2008. It happened to global market also.

Although economic growth slowed considerably during the recession, Indonesia achieved higher growth compared to the other G20 members with GDP 4.5% in 2009 [1]. Stable GDP growth is supported by domestic demand. Indonesia with 240 million of population is the fourth most populated country in the world which more than half of that population is under the age of 30. High growth GDP is followed by high investment in infrastructure. The Indonesian government has been eager to boost infrastructure financing up to 5% of GDP as well as attract private investment, which recorded US$ 941.5 million investment for infrastructure and US$ 711 million for mining sector in 2010. Considering global economy few years back, it appears that most of the foreign investors shift their investment from developed countries, like the U.S. to emerging countries in Asia, including Indonesia. Therefore, it is very important to explore the dynamism of the Indonesian stock market.

Previous research has examined the relationship between intermediaries, stock markets, and real activity in four East Asian countries, including Indonesia [2]. A recent study shows that the trading rules have the stronger predictive power in the emerging stock markets than in the more developed stock market [3]. In the past decade, various methods have been widely applied to explore the internal dynamism of the stock market. Genetic algorithm (GA) is an approach used to avoid local optimum. GA simulates the revolution in biology to keep better chromosome to reach the purpose of optimization. Some research applies the optimized search property of the GA algorithm to locate distribution centers for single product network such that the sum of facility location, pipeline inventory, and safety stock costs is minimized [4]. Ref. [5] combines the vector autoregression (VAR) and genetic algorithm (GA) with a neural network (NN) to model and forecast Asian Pacific stock markets. Their results show that their system is more robust and makes more accurate predictions than the benchmark NN.

Support vector machine (SVM) became a useful and popular method used by many researchers to avoid local optimum and achieve significant performance. Some research proposes a dynamic fuzzy model to explore the stock market dynamism. The fuzzy method combines various factors with an influential degree as the input variables, uses a GA algorithm to adjust the range of influential degree for each variable, and employs SVM to explore the stock market dynamism. The variables used in the experiment include technical indicators and macroeconomic variables [6].

Support vector regression (SVR) is extended from SVM. It adopts loss function and penalty parameter to avoid the effect of noise and outlier. SVR can convert nonlinear problems into high dimensional space and obtain good classification performance. Support vector regression is also used along with the fuzzy theorem in

* Corresponding author: Deng-Yiv Chiu
E-mail: chiuden@chu.edu.tw

doi:10.4304/jsw.8.9.2147-2154

some other researches to solve the two problems in prediction of financial time series: noise and non-stationarity [7]. Also, some researches use support vector regression and artificial intelligence approaches to study seasonality effect of stock dynamism [8,9].

The theory of fuzzy sets was first introduced by Loti Zadeh, primarily in the context of his interest in the analysis of complex systems. It introduces vagueness by eliminating the sharp boundary dividing members of the class from nonmembers. Fuzzy theory provides the forms for representing uncertainties. It is a tool for modeling uncertainties.

Some studies use fuzzy theorem and technical analysis to transfer technical indicators into fuzzy technical indicators. They set up fuzzy rules for membership function of each fuzzied technical indicator. When output is generated, the decisions are made under different rules. Therefore, when there is a change or certain trend in share price, an investment strategy can be planned based on fuzzy logic [10]. To make a stock investment decision, investors can refer to the If-Then rules generated from the fuzzy rule of fuzzy theorem and stock price fluctuation found from rough set theorem. Some studies combined Artificial Neural Network (ANN) and fuzzy regression to improve forecasting accuracy [11].

Some researches present the development of fuzzy portfolio selection model in investment. Using the formulated fuzzy portfolio model, a genetic algorithm (GA) is applied to find optimal values of risky securities [12]. The applied model estimates the investor's preference about risk-return trade-off. The obtained results from the modeling satisfy the efficiency of the presented fuzzy approach in portfolio selection. Some research present two fuzzy portfolio selection models where the objective is to minimize the downside risk constrained by a given expected return [13].

To examine dynamism of emerging stock market, in this paper we employ a hybrid approach of fuzzy theorem, support vector regression, genetic algorithm, and quarterly moving window to explore the Indonesian stock dynamism among same quarter in continuous years. In the experiment, we test the performance of the proposed approach by using daily prices of Indonesian stock market index from 2006 to 2011.

II. Related Method

A. Genetic Algorithm (GA)

GA is an efficient and better search method in the broad sense. With the simulation of biological evolution phenomenon, the parameter with better fitness function value is left. Also, with mechanisms of crossover and mutation, issue of partial minimization during a search is avoided and search time is shortened. The evolution process of genetic algorithms is shown as follows:

(1) Initialization: Each chromosome is created by randomly obtaining the diversity solutions.
(2) Selection: Select chromosome by evaluating the fitness value of each chromosome for searching near-optimization solution. The chromosomes with better fitness values are selected into the recombination pool using the roulette wheel or the tournament selection method.
(3) Crossover: Here, genes between two parent chromosomes are exchanged to obtain new offspring to attempt to get better solutions. Exchanging methods of genes crossover include one point crossover, two-point crossover, or homologous crossover between two chromosomes.
(4) Mutation: Using mutation to change the gene code from 0 to 1 or vice verse can differ from the population as a stochastic perturbation.
(5) Evolutionary cycle: Here, termination criteria are used to determine if the process should terminate or the process should go to step 2 repeatedly with the next generation.

B. Support Vector Regression (SVR)

Support vector regression (SVR) is often applied in the fields of pattern recognition and text classification. Theoretically, it is a learning system using linear-function hypothesis space in a high-dimensional feature space, and a kind of learning algorithm training from optimization theorem and minimized structure risk. Support vector regression consists of linear support vector regression and non-linear support vector regression [14].

(1) Linear support vector regression

SVR minimizes the error of training data to define a regression function. The equation is as below.

$$f(x,w)=\sum_{i=1}^{m}\langle x_i, w_i\rangle + b \quad (1)$$

where x is input vector, w is weight vector, m is count of training samples, b is bias.

In order to avoid noise and outlier, SVR employs the loss function and penalty parameter. Loss function is mainly used to find out the distance between a regression function and training data.

(2) Non-linear support vector regression

The non-liner separation problem can be solved using a mapping function Φ, which called the input spacetion, can map input space of training data into a higher-dimensional feature space. The inner product is replaced by the kernel function as below.

$k(x_i, x_j) := (\Phi(x_i)\Phi(x_j))$

Therefore, the function of optimization problem solved through non-linear support vector regression can be rewritten as:

$$\text{Maximize: } \mathrm{L_D} = -\frac{1}{2}\sum_{i,j=1}^{m}(\alpha_i-\alpha_i^*)(\alpha_j-\alpha_j^*)k\langle x_i, x_j\rangle - \varepsilon\sum_{i=1}^{m}(\alpha_i+\alpha_i^*)+\sum_{i=1}^{m}(\alpha_i-\alpha_i^*)\mathrm{y_i} \quad (2)$$

$$\text{st.} \ \sum_{i=1}^{m}(\alpha_i - \alpha_i^*) = 0, \ i = 1,2,...,m$$

$$0 \le \alpha_i, \alpha_i^* \le C, \ i = 1,2,...,m$$

The main kernel functions used are Linear function, Polynomial function, and Radial Basis function (RBF) shown as equations (3), (4), and (5).

Linear Kernel： $k(x_i, x_j) = x_i \cdot x_j^T$ (3)

Polynomial Kernel： $k(x_i, x_j) = (1 + x_i \cdot x_j)^d$ (4)

Radial Basis Function Kernel (RBF Kernel) ：

$$k(x_i, x_j) = \exp(-\gamma \left\| x_i \cdot x_j \right\|^2) \quad (5)$$

C. Fuzzy Clustering

Fuzzy c-means (FCM) algorithm combines the automatic clustering feature of k-means algorithm and the membership degree of fuzzy theorem to appropriately cluster objects and calculate the membership degree of each object belonging to each cluster. The main difference between FCM and k-means algorithm is that FCM employs fuzzy concept, thus, an object is not only clustered to a particular cluster, but may belong to each cluster with different membership degrees. These are used to express the correlation degree between objects and each cluster.

The main process of FCM algorithm is to calculate the cluster centroid through the cluster membership degree of each object to obtain the centroid which best represents each cluster. The membership degree is then adjusted through the distance between an object and new centroid of each cluster, in order to achieve a more suitable membership degree. The new membership degree among objects and the centroid is used to assess the performance of the clustering. The process is repeated until the clustering performance reaches the preset convergence threshold value.

III. The Proposed Fuzzy GA-SVR Approach

We use fuzzy c-means, fuzzy relation composition, and defuzzication methods to select number of trading days to calculate technical indicator values. This study adopts the fuzzy c-means algorithm to cluster values of all technical indicators appropriately to find more intensive clusters.

Then, we employ fuzzy relation composition and defuzzication methods to obtain the membership degree between technical indicators and transaction strategies. Finally, the results are compared with actual stock price fluctuation to assess the optimal number of trading days used to calculate technical indicator values. Then we utilize genetic algorithms to locate the approximate optimal combination of technical indicators. The corresponding values of those selected technical indicators are taken from the training data to form the input vectors of SVR, which is trained through the property of nonlinearity and high dimensionality.

In the design of the genetic algorithm in this study, each chromosome consists of 14 genes and each gene represents one technical indicator. When gene value is 1, its corresponding technical indicator value is used as a part of the input vector; when gene value is 0, its corresponding technical indicator value is not used as a part of the input vector.

To find the optimal combination of technical indicators, four key indicators, which are earning rate, transaction precision, recall, and precision, are considered in fitness function as below. Larger fitness function value is better.

Fitness function = earning rate × transaction precision + recall + precision

where

$$\text{earning rate} = \prod_{i=1}^{I_m}(P_{sell_i} - P_{buy_i}) / P_{buy_i}$$

I_m is transaction count of year m, P_{sell_i} is selling price of transaction i, P_{buy_i} is buying price of transaction *i.*

transaction precision = count of transactions predicted correctly / total transaction count

recall = count of transactions predicted correctly (TP) / (count of transactions predicted as rising correctly (TP) + count of transactions predicted as falling but actually rising (FN))

precision = count of transactions predicted correctly (TP) / (count of transactions predicted as rising correctly (TP) + count of transactions predicted as rising but actually falling (FP))

The kernel function of SVR model used in this study is Radial Basis Function (RBF) since RBF is a non-linear kernel function that can convert the data from original space to a higher-dimensional space to solve non-linear problems well. When the attribute original data is non-linear, this function has good effect. Parameter γ is set as 4, C value as 1. In our experiments, when the output value of SVR is +1, i.e. predictive trend is rising, the strategy is buying. If the stock has been bought, then it should be kept holding. When the output value of the SVR is -1, i.e. predictive trend is falling, the strategy is selling. If the stock is not held, then it should not be bought.

The proposed approach is explained as follows:

(1) Data collection: We collect Indonesian stock data from U.S.A. Yahoo! financial website for 6 years. The extracted period is from 2006/1/2 to 2011/12/30. The extracted attributes include opening price, highest price, lowest price, and closing price.

(2) Computation and normalization of technical indicator values for specific number of trading days: We adopt 14 kinds of technical indicators. We calculate technical indicator values for specific number of trading days. The range of number of trading days used to calculate technical indicator values is set between 3 and 60 days. Then we normalize the values as below:

normalized value = (original value – average value) /standard deviation.

(3) Fuzzy cluster with FCM algorithm: We employ FCM to cluster normalized technical indicator values for a specific cluster number. The cluster number is set between 3 and 6. Then we calculate the centroid of each cluster used to calculate the membership degree between a technical indicator value and a cluster.

(4) Calculating the membership degree between a cluster and a transaction: We compute the daily actual fluctuation percentage of technical indicator values clustered into same cluster with highest membership degree to obtain the membership degree between a cluster and a transaction.

(5) Calculating the membership degree between a technical indicator value and a transaction strategy: We employ fuzzy relation composition and defuzzication methods to compute the membership degree between a technical indicator value and a transaction strategy by employing the membership degree between a technical indicator value and a cluster and the membership degree between a cluster and a transaction strategy obtained in pervious steps.

(6) Calculating coincidence rate of a technical indicator: Here, we compute the coincidence rate for the specific number of trading days and the specific cluster number by comparing the membership degree between technical indicators and transactions to the corresponding actual rising of falling of stock price.

(7) Is maximum cluster number reached? : Here, we examine if the number of clusters employed in the loop reached the maximum limit. If so, go to step (8). Otherwise, the number of cluster is incremented by one. Go to step (3).

(8) Is maximum number of trading days used to calculate technical indicator values reached? : Here, we examine if the number of trading days used to calculate technical indicator values employed in this loop has reached the maximum limit. If so, go to step (9). Otherwise, number of trading days is incremented by one. Go to step (2).

(9) Selection of optimal number of trading days used to calculate technical indicator values: We select the number of trading days which leads to best coincidence rate as the optimal one. Then we form the training and testing data sets by using the selected number of trading days to calculate the corresponding technical indicator values.

(10) Initialization of chromosome for GA process: The first generation of GA process is initialized at random. A generation includes 20 chromosomes and each chromosome consists 14 genes. Each gene represents one technical indicator.

(11) Genotype converting: Genes are decoded to facilitate the combination of selected technical indicators.

(12) Training data selection: Corresponding values of selected technical indicators are extracted as training data to form the input of SVR.

(13) SVR training: The extracted training data is used to train SVR and produce values needed to evaluate fitness function. The kernel function employed is RBF.

(14) Evaluation of fitness function: To find the optimal combination of technical indicators, four key indicators, which are earning rate, transaction precision, recall, and precision, are considered in fitness function. Larger fitness function value means that the chromosome can make better financial earning.

(15) Termination criterion of genetic algorithm: the termination criterion is evolution of 50 generations. If the criterion is met, terminate the GA process and then go to step (17).

(16) Process of genetic algorithm: In genetic algorithm, a chromosome evolution includes three processes, selection, crossover, and mutation. In the selection process, from 20 chromosomes, one quarter of chromosomes with the highest fitness values are selected and duplicated using the roulette wheel selection method. In the crossover process, double-point crossover method is adopted. In mutation, the mutation rate is defined as 1%, and the process is redirected to step (11).

(17) Evaluation of testing data with trained SVR classifier: The trained SVR classifier is used to classify testing data to determine the proper transaction time point.

(18) Performance comparison: The performance of the proposed approach is compared with that of other methods to see how much the proposed method can outperform.

IV. Data and Empirical Results

Here we describe the empirical description, process, results and comparison.

A. Experiment Description

Here, we introduce the empirical data, technical indicators, and quarterly moving window employed in the experiments.

In order to explore the quarterly dynamism of Indonesian stock, we extract Indonesian Stock Exchange Composite Index (JCI) daily data from U.S.A. Yahoo! Financial website (http://finance.yahoo.com) for 6 years. The data period is between 2006/1/1 and 2011/12/30. The data count and index levels at the ends of years are as in Table I.

We refer to some researches and adopt 14 technical indicators as input variables [15]. They include Different (DIF), Moving average convergence and divergence (MACD), Relative strength (RS), Relative strength index (RSI), Relative strength volume (RSV), K line (K), D line

TABLE I.
EMPIRICAL DATA COUNT OF INDONESIA STOCK MARKET, INCLUDING INDEX LEVEL AT THE END OF A YEAR, 2006-2011

End of Year Date	Index Level (End of Year Index)	Number of Transaction Days
Dec 29 2006	1805.52	245
Dec 28 2007	2745.83	250
Dec 30 2008	1355.41	242
Dec 30 2009	2534.36	243
Dec 30 2010	3703.51	245
Dec 30 2011	3821.99	247

(D), J line (J), Psychological line (PSY), BIAS, Momentum (MTM), Williams overbought/oversold index (WMS), AR, and BR.

In moving window, data of a past period is treated as training data and data after that period as testing data to form a window. The period of training data moves subsequently to form another window. In quarterly moving window, the training data and testing data are from same quarter of various years. The data of previous year is treated as training data. The data of the same quarterly of later year is treated as testing data. The advantage of moving window model is that it has more training and testing data sets, so that the average of all data sets can be more representative. Moreover, the periods of training data and testing data of moving window are close, thus the data relativity is worth referencing. In this study, if the data in a certain quarter of a certain year is treated as training data, then the data in the same quarter of the following year is used as testing data of the same window. The data period is 6 years and 4 windows are formed each year from the first to the fifth year. Accordingly, there are 20 quarterly moving windows.

B. *Experiment Process*

Here, we introduce the selected optimal number of trading days used to calculate technical indicators values and the fitness values produced in the process of GA algorithm.

In order to find the optimal number of trading days used to calculate technical indicator values, we apply Fuzzy-c means algorithm, relation composition, and defuzzication methods. In each window, the best cluster number of each technical indicator is derived through Fuzzy c-means, so as to calculate the coincidence rate and determine number of trading days used to calculate technical indicator values. There are 20 quarterly moving windows.

As an example, Table II shows the coincidence rates of technical indicator RSV for the first quarterly moving window (training period: 1/2/2006~3/31/2006; testing period: 1/2/2007~3/30/2007). Through FCM method, the best coincidence rate for RSV in first window is 0.735. The optimal cluster number is 5 as well as the optimal number of trading days used to calculate RSV value is 6. The best coincidence rates, the optimal cluster number and the optimal number of trading days used to calculate

TABLE II.
THE COINCIDENCE RATE OF TECHNICAL INDICATOR RSV FOR THE FIRST QUARTERLY MOVING WINDOW (TRAINING PERIOD: 1/2/2006~3/31/2006; TESTING PERIOD: 1/2/2007~3/30/2007)

Number of trading days	Cluster Number			
	3	4	5	6
3	0.582	0.554	0.611	0.546
4	0.549	0.545	0.545	0.578
5	0.574	0.603	0.579	0.554
6	0.607	0.677	0.735	0.695
7	0.639	0.636	0.620	0.653
8	0.656	0.612	0.636	0.579
9	0.680	0.661	0.686	0.670
10	0.615	0.677	0.678	0.661
11	0.590	0.603	0.677	0.612
12	0.475	0.594	0.562	0.578
13	0.484	0.619	0.612	0.620
14	0.615	0.636	0.620	0.612
15	0.607	0.587	0.661	0.628
16	0.598	0.579	0.652	0.661
17	0.574	0.578	0.579	0.644
18	0.574	0.578	0.669	0.637
19	0.574	0.578	0.636	0.563
20	0.541	0.578	0.546	0.611
21	0.574	0.578	0.579	0.497
22	0.574	0.578	0.611	0.554
23	0.574	0.578	0.628	0.530
24	0.574	0.513	0.611	0.636
25	0.574	0.529	0.619	0.628
26	0.574	0.439	0.561	0.496
27	0.574	0.439	0.569	0.570
28	0.582	0.521	0.603	0.644
29	0.574	0.595	0.595	0.603
30	0.574	0.578	0.546	0.562
31	0.574	0.513	0.570	0.488
32	0.574	0.513	0.619	0.562
33	0.574	0.521	0.627	0.579
34	0.574	0.513	0.627	0.620
35	0.574	0.595	0.644	0.579
36	0.574	0.628	0.587	0.546
37	0.574	0.529	0.619	0.579
38	0.574	0.603	0.505	0.602
39	0.574	0.513	0.603	0.562
40	0.574	0.611	0.579	0.521
41	0.574	0.595	0.579	0.587
42	0.574	0.611	0.579	0.513
43	0.574	0.570	0.546	0.521
44	0.574	0.611	0.562	0.578
45	0.574	0.603	0.546	0.521
46	0.574	0.603	0.530	0.554
47	0.574	0.578	0.529	0.554
48	0.484	0.529	0.611	0.571
49	0.574	0.521	0.578	0.505
50	0.574	0.578	0.570	0.578
51	0.574	0.578	0.579	0.497
52	0.574	0.578	0.579	0.570
53	0.574	0.578	0.579	0.628
54	0.574	0.578	0.521	0.627
55	0.484	0.578	0.579	0.611
56	0.574	0.611	0.587	0.554
57	0.607	0.620	0.505	0.562
58	0.467	0.594	0.620	0.480
59	0.607	0.628	0.628	0.587
60	0.598	0.628	0.587	0.644

TABLE III.
THE BEST COINCIDENCE RATES, THE OPTIMAL CLUSTER NUMBER AND THE OPTIMAL NUMBER OF TRADING DAYS USED TO CALCULATE TECHNICAL INDICATOR VALUES IN 20 QUARTERLY MOVING WINDOWS 2006-2011

Moving Windows	AR			BIAS			BR			DIF			K		
	OTD	OCN	BCR	OTD	OCN	BCR	OTD	OCN	BCR	OTD	OCN	BCR	OTD	OCN	BCR
1	4	6	0.759	3	6	0.693	16	6	0.693	5	5	0.668	1	3	0.68
2	13	6	0.705	13	6	0.697	20	6	0.722	5	6	0.722	1	6	0.689
3	10	6	0.743	29	6	0.735	19	6	0.743	5	5	0.752	1	6	0.645
4	54	4	0.735	55	6	0.786	50	5	0.777	15	6	0.709	1	5	0.624
5	9	6	0.763	4	4	0.674	32	4	0.658	5	6	0.683	1	3	0.677
6	37	6	0.752	49	6	0.735	4	6	0.718	6	6	0.719	1	5	0.677
7	10	5	0.693	51	6	0.685	52	4	0.7	15	6	0.7	1	3	0.68
8	21	6	0.755	9	5	0.772	59	5	0.739	5	6	0.755	1	6	0.655
9	19	6	0.707	26	6	0.734	59	6	0.716	33	6	0.689	1	6	0.708
10	60	6	0.736	50	5	0.703	24	6	0.688	44	5	0.727	1	6	0.672
11	4	6	0.784	7	6	0.735	4	6	0.719	7	5	0.767	1	4	0.624
12	27	6	0.733	53	5	0.637	57	6	0.699	19	6	0.725	1	6	0.592
13	56	5	0.689	4	6	0.697	48	4	0.719	20	4	0.704	1	6	0.58
14	49	6	0.691	38	6	0.658	43	6	0.722	36	6	0.732	1	5	0.545
15	55	3	0.754	58	5	0.743	55	6	0.752	22	6	0.785	1	6	0.683
16	32	4	0.718	26	5	0.726	55	4	0.717	5	6	0.767	1	5	0.668
17	42	6	0.719	32	4	0.702	42	5	0.71	5	6	0.719	1	4	0.628
18	46	6	0.714	15	6	0.658	6	6	0.723	10	6	0.731	1	5	0.666
19	57	5	0.73	26	6	0.697	13	6	0.739	43	6	0.706	1	5	0.654
20	40	6	0.685	3	6	0.661	40	6	0.668	44	6	0.718	1	6	0.677
Average	32.25	5.5	0.728	27.55	5.55	0.706	34.9	5.45	0.716	17.45	5.7	0.724	1	5.05	0.651

Moving Windows	D			J			MACD			MTM			PSY		
	OTD	OCN	BCR	OTD	OCN	BCR	OTD	OCN	BCR	OTD	OC	BCR	OTD	OCN	BCR
1	1	5	0.636	1	3	0.566	44	5	0.64	60	6	0.69	18	5	0.71
2	1	6	0.689	1	5	0.672	12	4	0.69	22	6	0.70	44	5	0.70
3	1	6	0.661	1	4	0.645	33	6	0.71	37	5	0.71	58	6	0.69
4	1	6	0.631	1	6	0.735	11	4	0.69	31	3	0.75	50	6	0.78
5	1	6	0.618	1	6	0.626	3	6	0.63	5	6	0.67	51	5	0.73
6	1	6	0.644	1	6	0.628	4	6	0.69	3	5	0.73	12	6	0.71
7	1	4	0.606	1	6	0.677	7	6	0.68	56	6	0.69	59	6	0.72
8	1	5	0.68	1	5	0.739	17	5	0.69	56	5	0.73	40	6	0.72
9	1	6	0.645	1	4	0.655	21	6	0.67	35	6	0.72	31	6	0.74
10	1	5	0.64	1	4	0.686	37	6	0.74	31	6	0.73	11	6	0.67
11	1	3	0.627	1	3	0.643	18	6	0.68	5	6	0.72	22	6	0.73
12	1	6	0.556	1	6	0.575	33	6	0.67	24	5	0.67	40	6	0.70
13	1	5	0.638	1	5	0.605	10	5	0.68	22	5	0.73	40	5	0.73
14	1	4	0.634	1	5	0.61	15	6	0.70	32	6	0.72	36	6	0.71
15	1	4	0.667	1	6	0.667	8	6	0.70	39	6	0.75	38	6	0.80
16	1	4	0.644	1	4	0.678	4	6	0.71	32	6	0.76	31	6	0.74
17	1	4	0.578	1	6	0.545	17	4	0.66	47	6	0.71	47	4	0.77
18	1	6	0.65	1	5	0.601	31	6	0.71	52	5	0.69	27	6	0.71
19	1	4	0.68	1	6	0.739	21	6	0.76	60	6	0.75	58	6	0.72
20	1	4	0.636	1	4	0.726	21	6	0.69	41	6	0.69	46	6	0.72
Average	1	4.95	0.638	1	4.95	0.651	18.35	5.55	0.69	34.5	5.55	0.71	37.95	5.7	0.72

Moving Window	RS			RSI			RSV			WMS			Average		
	OTD	OCN	BCR	OTD	OCN	BC	OTD	OCN	BCR	OTD	OCN	BCR	OTD	OCN	BCR
1	13	6	0.693	13	6	0.69	6	5	0.73	6	6	0.71	17.09	5.21	0.68
2	45	6	0.705	8	6	0.70	46	6	0.72	27	6	0.706	23.18	5.71	0.70
3	25	5	0.727	12	6	0.73	31	5	0.71	39	6	0.727	27.09	5.57	0.71
4	52	6	0.803	51	6	0.81	3	6	0.64	32	6	0.666	36.73	5.36	0.72
5	14	6	0.683	14	6	0.72	60	6	0.68	35	4	0.65	21.09	5.29	0.68
6	47	5	0.727	7	6	0.74	45	6	0.65	44	6	0.644	23.45	5.79	0.70
7	10	6	0.701	33	4	0.70	7	6	0.70	20	6	0.7	29.09	5.29	0.69
8	52	6	0.756	16	5	0.73	18	5	0.75	18	6	0.756	28.27	5.43	0.73
9	19	6	0.716	20	6	0.72	21	6	0.68	23	6	0.663	27.91	5.86	0.70
10	23	5	0.759	59	6	0.73	55	5	0.69	56	6	0.696	40.91	5.50	0.71
11	19	6	0.72	52	5	0.72	5	6	0.70	5	5	0.736	13.45	5.21	0.71
12	12	6	0.69	59	6	0.67	45	6	0.61	11	6	0.601	34.55	5.86	0.65
13	56	5	0.714	40	5	0.71	57	6	0.67	25	6	0.679	34.36	5.14	0.68
14	37	5	0.707	23	6	0.72	43	5	0.73	13	6	0.731	33.18	5.57	0.69
15	38	6	0.768	11	6	0.76	49	6	0.71	21	5	0.709	35.82	5.50	0.73
16	48	6	0.76	31	6	0.77	56	6	0.67	48	6	0.677	33.45	5.29	0.72
17	56	5	0.735	45	6	0.69	35	3	0.67	31	6	0.694	36.27	4.93	0.68
18	57	6	0.788	57	5	0.74	21	6	0.65	39	5	0.666	32.82	5.64	0.69
19	55	6	0.747	57	6	0.75	14	6	0.71	16	6	0.697	38.18	5.71	0.72
20	40	5	0.677	40	5	0.69	13	6	0.69	5	5	0.685	30.27	5.50	0.69
Average	35.9	5.65	0.729	32.4	5.65	0.72	31.5	5.6	0.69	25.7	5.7	0.69	29.86	5.468	0.7

Note: OTD = Optimal number of trading days used to calculate technical indicator values, OCN = Optimal cluster number, BCR = Best Coincidence rate technical indicator values in 20 quarterly moving windows can be seen in Table III.

Number of trading days used to calculate technical indicator values are between 3 and 60 days. Take DIF in first moving window as an example. As shown in Table III, the number of trading days used to calculate DIF value is 5, that is, when calculating the DIF value of a day, the stock market data of the 5 days before the day should be used. In addition, Stochastic Oscillator (K, D and J line) has always adapted only one day before to calculate its value. Therefore, the number of trading days used to calculate technical indicator values is always 1 for K, D and J. We exclude those three indicators at calculating the average number of trading days as seen in Table III.

We utilize genetic algorithm to locate the approximate optimal combination of technical indicators. The corresponding values of these technical indicators are taken from the training data to form the input vectors of SVR, so as to train the SVR classifier. Finally, technical indicator combination with best performance is obtained by the method of evolution. The fitness function values of optimal selected chromosome in GA process are shown in Table IV. Take the first moving window in 2006-2007 as an example. Its fitness values are 3.357, 3.713, 3.945, and 3.457, respectively. The average of best fitness values for 20 moving windows is 3.817.

C. Experiment Result

The output of the proposed approach is used to determine the transaction time point. When the predicted trend for the day is up, we should keep holding if we hold the stock; we can buy in if we do not hold the stock. On the contrary, when the predicted trend for the day is down, we should sell out the stock in hand; we should keep watching without any actions if we do not hold the stock.

We apply the average yearly accumulated earning rate (AYAER) to evaluate the performance of our proposed approach. The AYAER can be calculated as below.

Average yearly accumated earning rate (AYAER)=

$$\left(\left(\prod_{m=s}^{f}\prod_{i=1}^{I_m} 1+(P_{sell_i}-P_{buy_i})/P_{buy_i}\right)-1\right)/\text{number\_of\_years} \times 100\%$$

where I_m is transaction count of year m, P_{sell_i} is selling price of transaction i, P_{buy_i} is buying price of transaction i, s is start year of transaction, and f is final year of transaction.

The AYAER of each methods is shown in Fig. 1. The AYAER of proposed approach (GA-Fuzzy-SVR) is 24.55%.

D. Experiment Comparison

First, we compare the proposed approach with the performance of experiment with same number of trading days to calculate technical indicator values. That is, there is no fuzzy theory adopted. It can be called GA-SVR method. The experimental result is illustrated by Fig. 1. The proposed approach (Fuzzy-GA-SVR) outperforms GA-SVR. The AYAER of proposed approach outperforms GA-SVR by 21.67%. Thus, FCM, fuzzy relation composition, and defuzzication methods employed work well in this study.

Then, we compare the performance of experiment with same technical indicators (Fuzzy-SVR method). That is, GA method is not employed to find the best combination of input variables. As can be seen in Fig. 1, the performance of of proposed approach outperforms Fuzzy-SVR method. The AYAER of proposed approach outperforms Fuzzy-SVR by 14.07%. Therefore, GA method should be applied to select suitable technical indicators.

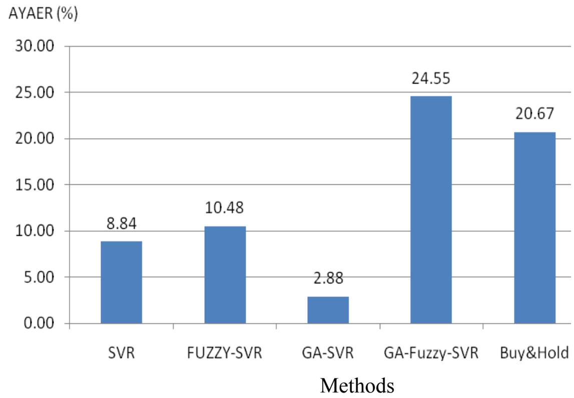

Figure 1. The average yearly accumulated earning rate of SVR, Fuzzy-SVR, GA-SVR, Fuzzy GA-SVR and Buy-and-Hold method

TABLE IV.
THE FITNESS FUNCTION VALUES OF OPTIMAL SELECTED CHROMOSOME IN GA PROCESS FOR EACH MOVING WINDOW, 2006-2011

Year	2006-2007				2007-2008				2008-2009			
Season	1	2	3	4	5	6	7	8	9	10	11	12
Fitness function values	3.357	3.713	3.945	3.457	3.620	3.635	4.047	3.728	3.306	3.635	4.301	3.502

Note: Fitness function = earning rate × transaction precision + recall + precision

Year	2009-2010				2010-2011				Average
Season	13	14	15	16	17	18	19	20	
Fitness function values	3.597	4.623	4.121	4.147	3.873	3.767	4.085	3.882	3.817

The proposed approach also outperforms other methods such as SVR as well as Buy-and-Hold method. As shown in Fig. 1, SVR and Buy-and-Hold method earn 8.84%, and 20.67%, respectively. The yearly earning rate in each year of buy-and-hold method is calculated by subtracting the stock price of the first day by that of the last day of the year. Thus, only the proposed approach can outperforms Buy-and-Hold method. Therefore, the hybrid approach should be applied to explore the dynamism of emerging markets.

V. CONCLUSIONS

As a member of G-20 major economies, Indonesia is one of the emerging market economies of the world. It is the largest economy in Southeast Asia and fifth-largest economy in Asia. We employ a hybrid approach of fuzzy theorem, support vector regression method, genetic algorithm, and quarterly moving window to explore the dynamism of Indonesia stock market. In the empirical results, the proposed approach outperforms SVR, GA-SVR, Fuzzy-SVR as well as Buy-and-Hold method. Therefore, the hybrid approach has referential values. However, weaknesses in one or more parts of the hybrid cause a decrease in performance.

The future directions of this research can be as follows:

(1) The proposed approach can be revised by using macroeconomic indicator, such as consumer price indices and interest rate.

(2) In addition to examining hybrid method, representative of developed and developing markets can be used in the next study.

REFERENCES

[1] Badan Pusat Statistik, "Economic Growth of Indonesia in 2009," *Berita Resmi Statistik*, vol. 12/02, p. 1, February 2010.

[2] J. H. Kim and P. L. Rousseau, "Credit buildups and the stock market in four East Asian economies," *Journal of Macroeconomics*, vol. 34, pp. 489-503, June 2012. doi: 10.1016/j.jmacro.2012.01.004

[3] H. Yu, G. V. Nartea, C. Gan, and L. J. Yao, "Predictive ability and profitability of simple technical trading rules: Recent evidence from Southeast Asian stock markets," *International Review of Economics & Finance*, in press. doi:10.1016/j.iref.2012.07.016

[4] K. Sourirajan, L. Ozsen, and R. Uzsoy, "A genetic algorithm for a single product network design model with lead time and safety stock considerations," *Eur. J. Oper. Res.*, vol. 197, pp. 599-608, September 2009. doi: 10.1016/j.ejor.2008.07.038

[5] S.I. Ao, "A hybrid neural network cybernetic system for quantifying cross-market dynamics and business forecasting," *Soft Comput.*, (2010). doi:10.1007/s00500-010-0580-4.

[6] D.Y. Chiu and P.J. Chen, "Dynamically exploring internal mechanism of stock market by fuzzy-based support vector machines with high dimension input space and genetic algorithm," *Expert Syst. Appl.*, vol. 36, pp. 1240-1248, March 2009. doi:10.1016/j.eswa.2007.11.022

[7] R. Khemchandani, Jayadeva, and S. Chandra, "Regularized least squares fuzzy support vector regression for financial time series forecasting," *Expert Syst. Appl.*,vol. 36, pp. 132-138, January 2009. doi:10.1016/j.eswa.2007.09.035

[8] D. Y. Chiu, M. H. Ying and P. C. Yu, "A study of Taiwan political elections and stock market dynamism," *Sixth International Conference on Genetic and Evolu.y Compu.* , pp.562-565, 2012.

[9] D. Y. Chiu and C. Y. Shiu, "Exploring seasonality effect of multinational stock dynamism with support vector regression and artificial intelligence approach," *Fourth International Conference on Innovative Computing, Information and Control* , December, 2009.

[10] H. Dourra and P. Siy, "Investment using technical analysis and fuzzy logic," *Fuzzy Sets and Systems*, vol. 127, pp. 221-240, April 2002. doi:10.1016/S0165-0114(01)00169-5

[11] M. Khashei, S.R. Hejazi, and M. Bijari, "A new hybrid artificial neural networks and fuzzy regression model for time series forecasting," *Fuzzy Sets and Systems,* vol. 159, pp. 769-786, April 2008. doi:10.1016/j.fss.2007.10.011

[12] R. H. Abiyev and M. Menekav, "Fuzzy portfolio selection using genetic algorithm," *Soft Comput*., vol. 11, pp. 1157–1163, August 2007. doi:10.1007/s00500-007-0157-z

[13] E. Vercher, J.D. Bermúdez, and J.V Segura, "Fuzzy portfolio optimization under downside risk measures," *Fuzzy Sets and Systems,* vol. 158, pp. 769-782, April 2007. doi: 10.1016/j.fss.2006.10.026

[14] C.L. Huang, and C.Y. Tsai, "A hybrid SOFM-SVR with a filter-based feature selection for stock market forecasting," *Expert Syst. Appl.*, vol. 36, pp. 1529-1539, March 2009. doi: 10.1016/j.eswa.2007.11.062

[15] K. J. Kim, "Financial forecasting using support vector machines," *Neural Comput. Appl.,* vol. 55, pp. 307-319, September 2003. doi:10.1016/S0925-2312(03)00372-2.

Deng-Yiv Chiu received the B.A. from Averett College, Virginia, USA in 1988, M.S. from University of Maryland, USA in 1990. He received the Ph.D. in Computer Science from Illinois Institute of Technology, USA in 1994. After working as an assistant professor at Dept. of Math and Computer Science, Chicago State University, USA, he is a full professor at Chung Hua University, HsinChu, Taiwan. His research interests include machine learning, information retrieval, and their applications to knowledge management and finance.

Agus Andria received B.S. in statistical computation from Institute of Statistics, Jakarta, Indonesia in 2001. Then, he received the M.Si. in statistics from Sepuluh Nopember Institute of Technology, Surabaya, Indonesia in 2008. His study interests include information retrieval, knowledge management, and machine learning.