Seasonality Effect of Stock Dynamism

Deng-Yiv Chiu, Cheng-Yi Shiu

Department of Information Management, Chung Hua University, chiuden@chu.edu.tw

Department of Finance, National Central University, cshiu@ncu.edu.tw

Abstract

We propose a hybrid approach of seasonal moving window, genetic algorithm, and support vector regression to explore seasonality effect for the stock indexes in two developed markets. First, we utilize genetic algorithm to locate the approximate optimal combination of technical indicators. Then the property of nonlinearity and high dimensionality of the support vector regression is employed to explore the stock price patterns. Finally, we adopt seasonal moving window to capture the seasonality effect of stock market returns. We find that the proposed method outperforms buy-and-hold returns.

Keywords: seasonality effect, moving window, genetic algorithm, support vector regression

1. Introduction

The seasonality, such as January effect or weekend effect, is well documented in literature. Therefore, it is very important to consider the seasonality effect. We try to establish an automatic GA-based nonlinear high-dimensional support vector regression method to explore the seasonality effect of stock dynamism.

In the past decade, various methods have been widely applied to explore the internal dynamism of stock market, such as linear and nonlinear mathematical models, multi-agent mechanism, and artificial neural network (ANN) of multiple layers used to simulate the potential stock market transaction mechanism [1, 2]. Because of the advantages of arbitrary function approximation and needlessness of statistics assumption, ANN is widely applied in the simulation of potential market transaction mechanism [3, 4, 5]. But there exist some problems in artificial neural network, such as local optimum and over-learning. In order to avoid those problems, some researches try to combine hybrid approaches of artificial intelligence methods and artificial neural network [6, 7, 8]. Genetic algorithm (GA) is an approach used to avoid local optimum. GA simulates the revolution of biology to keep better chromosome to reach the purpose of optimization. A noteworthy related work is study of the using genetic algorithms to find technical trading rules for S&P 500 index [9]. Some research applies the optimized search property of 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 [10]. Some research proposes a systematic approach to simulate the effect of foreign exchange management. GA is used to search optimal amount and time for forward and spot dealing per business day [11]. Ao [12] combines the vector auto regression (VAR) and genetic algorithm (GA) with neural network (NN) to model and forecast Asian Pacific stock markets. Their results show that their system is more robust and makes more accurate prediction than the benchmark neural network.

Then, support vector machine (SVM) became a useful and popular method used by many researchers to avoid local optimum and achieve significant performance. SVM has outstanding performances in handling high dimension entry space problems [13]. Such a feature leads to a better performance of SVM in simulating potential market transaction mechanism than other methods. Some researches proposed fuzzy linear SVM method for data mining [14]. Some researches developed a prediction model based on support vector machines [15]. Some researches adopt SVM to predict stock market dynamism with financial factors, such as macroeconomic variables and technical indicators [16]. Some researches utilize SVM and genetic algorithm (GA) to predict stock market dynamism with dynamically selecting stock market technical indicators in which genetic algorithm is used to reduce input feature dimension and select better model parameters to increase the forecast accuracy rate [17].

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. Some research mixes models of support vector regression and self-organizing feature map (SOFM) as well as filtering attribute screening method to predict the tertian closing price of Taiwan Stock Index Futures (FITX) [18]. First, input attributes with higher importance are selected through filtering attribute screening method. Then, training data is clustered through SOFM algorithm. Finally, the tertian closing price can be predicted through support vector regression. Support vector regression is also used along with fuzzy theorem in some other researches to solve the two problems in prediction of financial time series: noise and non-stationarity. Six data items, such as S&P500, Google, and Microsoft, from U.S. Yahoo are adopted in the prediction experiments [19]. The variations of foreign exchange are studied in some researches by combining with parametric techniques and nonparametric techniques. ARIMA (autoregression integrated moving average) and VAR (vector autoregression) are used in parametric techniques, while support vector regression and multilayer perceptron neural network are used in nonparametric techniques. The above-mentioned techniques are combined and tested on Euro, British Pound, Japanese Yen, and Australian Dollars [20].

Moreover, the variation of multinational stock market is a topic worth studying. Differences between returns and volatility in yen/dollar spot market of Tokyo, London and New York are discussed in some researches by adopting probability distribution techniques [21]. Some researches explored stock market dynamism in multi-nations with genetic algorithm, support vector regression, and optimal technical analysis [22].

Since the stock market may be affected by slow season and peak season, we try to explore this effect through seasonal moving window. We utilize genetic algorithm to locate the approximate optimal combination of technical indicators. The combination of technical indicators with best performance is found through the method of evolution. Then 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. Moreover, we divide a year into four seasons and use the seasonal moving window to capture the seasonality effect on stock market movement. Most important, we test the effectiveness of proposed model by using daily prices of two stock market indexes from 1996 to 2005. They represent the market performance in developed markets. We find that the proposed method outperforms buy-and-hold returns. Moreover, our results suggest that countries should use different technical indicators as their input variables. Also, the technical indicators influencing various markets should be different.

2. The proposed approach

We apply a hybrid approach of seasonal moving window, genetic algorithm, and support vector regression to explore the seasonality effect of stock dynamism. Since the affecting factors in various stock market may be different, we utilize genetic algorithm to locate the approximate optimal combination of technical indicators to make best performance. Then we explore the dynamism of stock market with the property of nonlinearity and high dimensionality of support vector regression method. Finally, we divide a year into 4 seasons and apply seasonal moving window to study the seasonality effect of stock dynamism.

Here, we introduce how to apply GA to local the approximate optimal combination of technical indicators and how to utilize SVR to make best performance. Finally, we describe the details of proposed integrated architecture.

2.1. GA-SVR method

In GA-SVR method, different combinations of technical indicators are selected through parameter optimization in genetic algorithm. Then, the corresponding values of these technical indicators are extracted from the training data to form the input vectors of SVR, so as to train SVR model. At last, technical indicator combination with best performance is obtained through the method of evolution to find the optimal SVR classifier.
2.1.1. 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 search is avoided and search time is shortened. Evolution process of genetic algorithm is shown as following:

(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 as shown in Figure 1.

(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 homologue crossover between two chromosomes as shown in Figure 2.

(4) Mutation: Using mutation to change gene code from 0 to 1 or vice versa can differ from population as a stochastic perturbation as shown in Figure 3.

(5) Evolutionary cycle: Here, termination criteria is used to determine if the process should terminate or the process should go to step 2 repeatedly with next generation as shown in Figure 4.

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 indicator, three key factors, which are earning rate, transaction precision, and F-measurement, are considered in fitness function as below. Larger fitness function value is better. The weights are 0.8 and 0.2 as shown in the equation. The weights are obtained with tried and error method.

Fitness function = (earning index × transaction precision) × 0.8 + (F1+F2) × 0.2

\[
\text{Fitness function} = (\text{earning index} \times \text{transaction precision}) \times 0.8 + (\text{F1}+\text{F2}) \times 0.2
\]

where

\[
\text{earning rate} = \prod_{i=1}^{t} \frac{1+(P_{sell} - P_{buy})}{P_{buy}}
\]

\[
\text{transaction precision} = \frac{\text{count of transactions predicted correctly}}{\text{total transaction count}}
\]

\[
F1 = \frac{\text{Recall of Rising Stock Market} \times \text{Precision of Rising Stock Market}}{\text{Recall of Rising Stock Market} + \text{Precision of Rising Stock Market}}
\]

\[
F2 = \frac{\text{Recall of Falling Stock Market} \times \text{Precision of Falling Stock Market}}{\text{Recall of Falling Stock Market} + \text{Precision of Falling Stock Market}}
\]
2.1.2. Support vector regression (SVR)

Support vector regression (SVR) is a regression technique extended from support vector machine [23, 24]. It is often applied in the fields of pattern recognition and text classification. Some researches used genetic algorithm optimization SVM to construct investment model [25]. 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. In recent years, it has been also used in the research of classification and prediction of finance. Support vector regression consists of linear support vector regression and non-linear support vector regression [18].

1) Linear support vector regression
   SVR minimizes the error of training data to define a regression function. The equation is as below.

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

   where \(w\) is weight vector, \(x\) is input vector, \(b\) is bias, \(m\) is count of training samples

   In order to avoid noise and outlier, SVR employs loss function and penalty parameter. Loss function is mainly used to find out the distance between regression function and training data. If the distance between predicted value and actual value is less than or equal to the set loss function value (\(\varepsilon\)), the loss function value is equal to 0; otherwise, the loss function value is not equal to 0, where the value of \(\varepsilon\) is defined by users themselves. Slack variable \((\xi_i, i = 1, ..., m)\) is used to solve the problem that applying loss function may exclude some training data. Moreover, each slack variable is multiplied by a penalty parameter to avoid the overlap problem of support vector regression. Therefore, the equation of SVR is shown as follows:

   Minimize \(\frac{1}{2}\|w\|^2 + C\sum_{i=1}^{m}(\xi_i + \xi_i^*)\)  

   Figure 1. Selection method in genetic algorithm.
   Figure 2. Two-point crossover method.
   Figure 3. Genetic mutation operation in GA process with the third gene to be mutated.
   Figure 4. Evolutionary cycle.
\[ w \cdot x + b - y \leq \epsilon + \xi, \]
\[ y - w \cdot x - b \leq \epsilon + \xi, \]
\[ \xi, \xi \geq 0, \quad i = 1, 2, \ldots, m \]

where \( \|w\| \) : Euclidean distance, is used to measure the flatness of \( f(x) \); \( \xi \) is the training error higher than margin; \( \xi \) is the training error lower than margin; \( y \) is the target value.

This equation can be converted into the optimization of Lagrange Multiplier. Its Lagrangian Function

\[
L_i = \frac{1}{2} \|w\|^2 - C \sum_{i=1}^{n} (\xi_i + \xi_i^*) - \sum_{i=1}^{n} (\eta_i + \eta_i^*) \xi_i - \sum_{i=1}^{n} \alpha_i (\epsilon + \xi_i - y_i + w \cdot x_i + b) \
- \sum_{i=1}^{n} \alpha_i^* (\epsilon + \xi_i^* - y_i - w \cdot x_i - b)
\]

is the coefficient of Lagrange.

(8)

Then, the statement below can be used to solve the original problem:

\[
\max_{\alpha, \eta} \min_{\epsilon, \xi} L_i \\
\text{s.t.} \quad \alpha, \eta \geq 0
\]

By differentiating with respect to \( w, b \), and \( \xi \) and simplifying, the equation below can be obtained.

\[
L_\epsilon = -\frac{1}{2} \sum_{i,j=1}^{n} (\alpha_i - \alpha_i^*) (\alpha_j - \alpha_j^*) \langle x_i, x_j \rangle - \epsilon \sum_i \alpha_i + \sum_j \alpha_j^* + \sum_i \sum_j \alpha_i (\alpha_j - \alpha_j^*)
\]

(9)

\[
\text{s.t.} \quad \sum_i (\alpha_i - \alpha_i^*) \text{ and } \alpha_i - \alpha_i^* \in [0, C]
\]

According to Karush Kuhn-Tucker (KKT) theory [26], \( b \) can be obtained through substituting \( w \), and the regression function produced by training data can be obtained accordingly (see equation below).

\[
f(x) = \sum_i (\alpha_i - \alpha_i^*) \langle x, x_i \rangle + b
\]

(10)

(2) Non-linear support vector regression

The non-liner separation question can be solved using a mapping function \( \Phi \), which called kernel function, can map input space of training data into a higher-dimensional feature space. The inner product is replaced by kernel function as below.

\[ k(x, x_i) := \langle \Phi(x), \Phi(x_i) \rangle \]

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

\[
\text{Maximize:} \quad L_i = -\frac{1}{2} \sum_{i,j=1}^{n} (\alpha_i - \alpha_i^*) (\alpha_j - \alpha_j^*) k \langle x_i, x_j \rangle - \epsilon \sum_i \alpha_i + \sum_j \alpha_i^* + \sum_i \sum_j (\alpha_i - \alpha_i^*) y_i
\]

\[
\text{s.t.} \quad \sum_i (\alpha_i - \alpha_i^*) = 0, \quad i = 1, 2, \ldots, m
\]

\[ 0 \leq \alpha, \alpha^* \leq C, \quad i = 1, 2, \ldots, m \]

The main kernel functions used are Linear function, Polynomial function, and Radial Basis function (RBF) shown as equations (12), (13), and (14). Users may need to adjust and set parameters for some
kernel functions to achieve better performance.

$$k(x, x') = x \cdot x'$$ \hspace{1cm} (12)

$$k(x, x') = (1 + x \cdot x')^d$$ \hspace{1cm} (13)

$$k(x, x') = \exp(-\gamma \|x - x'\|)$$ \hspace{1cm} (14)

The kernel function of SVR model used in this study is 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 original data and attribute are non-linear, this function has good effect. With trial and error method, parameter $\gamma$ 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.

2.2. The architecture of the proposed approach

The architecture is as shown in Figure 5. The detailed explanation is as follow.

1. Data collection: We collect stock data from U.S.A. Yahoo! financial website for 2 countries and 10 years. The extracted period is from 1996/1/1 to 2005/12/31. The extracted attributes include opening price, highest price, lowest price, and closing price.

2. Computation and normalization of technical indicator value: We adopt 14 technical indicators. We first calculate the technical indicator values and normalize them as below.

$$\text{normalized value} = \frac{\text{original value} - \text{average value}}{\text{standard deviation}}$$

3. Initialization of chromosome for GA process: The first generation of GA process is initialized at random. A generation includes 40 chromosomes and each chromosome consists 14 genes. Each gene represents one technical indicator.

4. Genotype converting: Genes are decoded to locate the combination of selected technical indicators.

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

6. 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.

7. Evaluation of fitness function: To find the optimal combination of technical indicators, three key factors, which are earning rate, transaction precision, and F-measurement, are considered in fitness function. Larger fitness function value means that the chromosome can make better financial earning.

8. Termination criterion of genetic algorithm: the termination criterion is evolution of 100 generations If the criterion is met, terminate the GA process and then go to step (10).

9. Process of genetic algorithm: In genetic algorithm, chromosome evolution includes three processes, selection, crossover, and mutation. In the selection process, from 40 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 (4).

10. Evaluation of testing data with trained SVR classifier: The trained SVR classifier is used to class testing data to determine the proper transaction time point.

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

3. Data and Empirical Results

In our study, we utilize GA-SVR method and seasonal moving window to explore the seasonality effect of stock dynamism, including 2 developed stock markets. Here we introduce the empirical description and the empirical results of the proposed approach.
3.1. Experiment description

3.1.1. Empirical data and technical indicator variable

In order to explore the dynamism of stocks, we extract stock data from U.S.A. Yahoo! Financial website (http://finance.yahoo.com) for 2 countries and 10 years, including 2 developed stock markets. The stock markets include U.S.A. S&P 500 and British FTSE 100. The data period is between 1996/1/1 and 2005/12/31. The detailed data counts are as in Table 1.

We refer to some researches and adopt 14 technical indicators as input variables [27, 28]. 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 (D), J line (J), Psychological line (PSY), BIAS, Momentum (MTM), Williams overbought/oversold index (WMS), AR, and BR.

3.1.2. Seasonal moving window model

In moving window model, 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 seasonal moving window, the training data and testing data are from same seasons of various years. The data of previous year is treated as training data. The data of the same season of later year is treated as testing data (see Figure 6). The advantage of moving
Table 1. Empirical data count of each stock market, including index level at the end of the a year

<table>
<thead>
<tr>
<th></th>
<th>U.S.A.</th>
<th>British</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Transaction Days</td>
<td>2,519</td>
<td>2,526</td>
</tr>
<tr>
<td>Index Level (End of Year Index)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dec 31 1996</td>
<td>740.74</td>
<td>4,118.50</td>
</tr>
<tr>
<td>Dec 31 1997</td>
<td>970.43</td>
<td>5,135.50</td>
</tr>
<tr>
<td>Dec 31 1998</td>
<td>1,229.23</td>
<td>5,882.60</td>
</tr>
<tr>
<td>Dec 31 1999</td>
<td>1,469.25</td>
<td>6,930.20</td>
</tr>
<tr>
<td>Dec 31 2000</td>
<td>1,320.28</td>
<td>6,222.50</td>
</tr>
<tr>
<td>Dec 31 2001</td>
<td>1,148.08</td>
<td>5,217.40</td>
</tr>
<tr>
<td>Dec 31 2002</td>
<td>879.82</td>
<td>3,940.40</td>
</tr>
<tr>
<td>Dec 31 2003</td>
<td>1,111.92</td>
<td>4,476.90</td>
</tr>
<tr>
<td>Dec 31 2004</td>
<td>1,211.92</td>
<td>4,814.30</td>
</tr>
<tr>
<td>Dec 31 2005</td>
<td>1,248.29</td>
<td>5,618.80</td>
</tr>
</tbody>
</table>

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 model are close, thus the data relativity is worth referencing. In this study, if the data in a certain season of a certain year is treated as training data, then the data in the same season of the following year is used as testing data of the same window. The data period is 10 years and 4 windows are formed each year from the first to the ninth year. Accordingly, there are 36 moving windows for each country.

3.2. Empirical results

3.2.1. GA algorithm process

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 model. 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 for each stock market are shown as Table 2. Take the seasonal windows of U.S.A. stock market in 1996 as an example. Its fitness values are 1.16, 1.20, 1.11, and 1.22, respectively. The average fitness values are 1.16 and 1.15 for each country respectively. The average values are all greater than 1.15. We can see that the fitness function values are satisfactory. It shows that the method of evolution works well here.

In our experiment, the GA converges well. As an example, the convergence figures for the final seasonal moving window of each country are shown as in Fig. 7. The fitness functions converge at most after 21 generations (see Figure 7).
<table>
<thead>
<tr>
<th>Year</th>
<th>Windows</th>
<th>Countries</th>
<th>U.S.A.</th>
<th>British</th>
</tr>
</thead>
<tbody>
<tr>
<td>1996</td>
<td>1</td>
<td>1.16</td>
<td>1.10</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1.20</td>
<td>1.11</td>
<td></td>
</tr>
<tr>
<td>1997</td>
<td>3</td>
<td>1.11</td>
<td>1.15</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>1.22</td>
<td>1.12</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>1.18</td>
<td>1.19</td>
<td></td>
</tr>
<tr>
<td>1997</td>
<td>6</td>
<td>1.21</td>
<td>1.25</td>
<td></td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>1.20</td>
<td>1.22</td>
<td></td>
</tr>
<tr>
<td>1998</td>
<td>8</td>
<td>1.23</td>
<td>1.23</td>
<td></td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>1.24</td>
<td>1.18</td>
<td></td>
</tr>
<tr>
<td>1998</td>
<td>10</td>
<td>1.16</td>
<td>1.11</td>
<td></td>
</tr>
<tr>
<td>1999</td>
<td>11</td>
<td>1.16</td>
<td>1.29</td>
<td></td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>1.29</td>
<td>1.15</td>
<td></td>
</tr>
<tr>
<td>1999</td>
<td>13</td>
<td>1.11</td>
<td>1.18</td>
<td></td>
</tr>
<tr>
<td></td>
<td>14</td>
<td>1.19</td>
<td>1.14</td>
<td></td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>1.08</td>
<td>1.04</td>
<td></td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>1.28</td>
<td>1.28</td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td>17</td>
<td>1.07</td>
<td>1.14</td>
<td></td>
</tr>
<tr>
<td></td>
<td>18</td>
<td>1.06</td>
<td>1.17</td>
<td></td>
</tr>
<tr>
<td>2001</td>
<td>19</td>
<td>1.14</td>
<td>1.10</td>
<td></td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>1.13</td>
<td>1.14</td>
<td></td>
</tr>
<tr>
<td>2001</td>
<td>21</td>
<td>1.17</td>
<td>1.07</td>
<td></td>
</tr>
<tr>
<td></td>
<td>22</td>
<td>1.26</td>
<td>1.13</td>
<td></td>
</tr>
<tr>
<td>2002</td>
<td>23</td>
<td>1.12</td>
<td>1.26</td>
<td></td>
</tr>
<tr>
<td></td>
<td>24</td>
<td>1.28</td>
<td>1.20</td>
<td></td>
</tr>
<tr>
<td>2002</td>
<td>25</td>
<td>1.00</td>
<td>1.08</td>
<td></td>
</tr>
<tr>
<td></td>
<td>26</td>
<td>1.03</td>
<td>0.84</td>
<td></td>
</tr>
<tr>
<td>2003</td>
<td>27</td>
<td>1.12</td>
<td>1.28</td>
<td></td>
</tr>
<tr>
<td></td>
<td>28</td>
<td>1.25</td>
<td>1.26</td>
<td></td>
</tr>
<tr>
<td></td>
<td>29</td>
<td>1.13</td>
<td>1.13</td>
<td></td>
</tr>
<tr>
<td>2003</td>
<td>30</td>
<td>1.23</td>
<td>1.20</td>
<td></td>
</tr>
<tr>
<td></td>
<td>31</td>
<td>1.12</td>
<td>1.11</td>
<td></td>
</tr>
<tr>
<td></td>
<td>32</td>
<td>1.19</td>
<td>1.17</td>
<td></td>
</tr>
<tr>
<td>2004</td>
<td>33</td>
<td>1.15</td>
<td>0.89</td>
<td></td>
</tr>
<tr>
<td></td>
<td>34</td>
<td>1.15</td>
<td>1.11</td>
<td></td>
</tr>
<tr>
<td>2005</td>
<td>35</td>
<td>0.92</td>
<td>1.12</td>
<td></td>
</tr>
<tr>
<td></td>
<td>36</td>
<td>1.12</td>
<td>1.09</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td>1.16</td>
<td>1.15</td>
<td></td>
</tr>
</tbody>
</table>
3.2.2. Empirical performance

The output of the 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 average yearly accumulated earning rate to evaluate the performance of our proposed approach. The average yearly accumulated earning rate can be calculated as below.

\[
\text{Average yearly accumulated earning rate (AYAER)} = \left( \prod_{i=s}^{f} \frac{1+(P_{i}-P_{i-1})/P_{i-1}}{1} \right) \times 100\%
\]

where \( I \) is transaction count of year \( m \), \( P_{i} \) is selling price of transaction \( i \), \( P_{s} \) is buying price of transaction \( i \), \( s \) is start year of transaction, and \( f \) is final year of transaction.

The average yearly accumulated earning rate is shown in Table 3. The average yearly accumulated earning rate of the U.S.A. is 12.63% and that of British is 12.61%.

The performances of the proposed method and buy-and-hold returns are compared (see Table 3). 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. The average yearly accumulated earning rate of the proposed approach of each country outperforms Buy-and-hold as follow: U.S.A. by 3.70% and British by 5.55%. The main reason may be that the technical indicators selected for each country have better influential degree, meanwhile, SVR exerts its dynamic exploration ability in high-dimensional space, and there exists seasonality effect in multinational stock dynamism.

3.2.3. Comparison with results of same technical indicators (SVR method)

Here, we compare the performance of experiment with same technical indicators (SVR method). That is, GA method is not employed to find the best combination of input variables. Every stock market uses the same technical indicators as input variables.

The experimental results are shown in Table 3. As seen, the performance of the experiments with different technical indicators selected through GA algorithm is better than that of the ones with the same technical indicators. Therefore, GA method should be applied to select suitable technical indicators. Countries should use different technical indicators as their input variables. That is, the technical indicators influencing various markets are different.

| Table 3. The performance comparison of the proposed GA-SVR, SVR and buy-and-hold method |
|-------------------------------|-------------------------------|
|                               | U.S.A. | British |
| GA-SVR                        | 12.63% | 12.61%  |
| SVR                           | 11.75% | 8.80%   |
| Buy-and-hold                  | 8.93%  | 7.06%   |
4. Conclusions

First, we utilize genetic algorithm to locate the approximate optimal combination of technical indicators. Then the property of nonlinearity and high dimensionality of the support vector regression is employed to explore the stock dynamism. Finally, we adopt seasonal moving window to study the seasonality effect of multinational stock market. In the experiment, we extract stock market data for 10 years and 2 countries. In the empirical results, we find that the performance of the proposed method is better than that of Buy-and-hold method. Therefore, the seasonality effect of stock market has referential value. Moreover, as seen, the technical indicators affecting various stock markets are different. Accordingly, each country should select its own factors carefully when exploring the dynamism of stock market. 

The future directions of this research can be as follows:
(1) In the future, fuzzy rule method can be used to combine with different fuzzied indicators to provide the decision models of stock market with more information.
(2) This paper explores the trend of stock market. We can modified the model to explore other derivatives, such as individual share or warrant.
(3) This paper focus on GA and support vector regression, however, other artificial intelligent techniques can be used to study the trend of stock market, such as artificial neural network, grey theorem, and case-base reasoning.

References