Expert Systems with Applications 38 (2011) 14178–14184

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

Enhanced stock price variation prediction via DOE and BPNN-based optimization

Ling-Feng Hsieh [a], Su-Chen Hsieh [b,c,*], Pei-Hao Tai [d]

[a] Department of Transportation Technology and Logistics Management, Chung Hua University, Hsinchu, Taiwan, ROC
[b] Program of Technology Management, Chung Hua University, Hsinchu, Taiwan, ROC
[c] Department of Finance, Chung Hua University, Hsinchu, Taiwan, ROC
[d] Delta Electronics Inc., Taoyuan, Taiwan, ROC

ARTICLE INFO

Keywords:
Stock price forecasting
Back-propagation neural network
Design of experiment
Taguchi method
Parameter optimization

ABSTRACT

Stock price variation predictions are at the core of many research issues, and neural networks (NNs) are widely applied and were proven to be more efficient than time series forecasting for stock price forecasting. However, this type of research always determines the parameter settings of the NNs rationally through a trial-and-error methodology. This paper integrates design of experiment (DOE), Taguchi method, and back-propagation NN (BPNN) to construct a robust engine to further optimize the prediction accuracy under a robust DOE-based predictor. Adopting data from Taiwan Stock Exchange (TWSE), the technical analytical indexes and β value of the listed stocks of TWSE were computed. The research results indicated that the proposed approach can effectively improve the forecasting rate of stock price variations.

1. Introduction

Predicting price activities in a stock market on the basis of either professional knowledge or stock analytical tools is a great concern of individual and institutional investors around the world. Price variations mean gains and losses for investors. Determining how to improve the forecasting accuracy is of great interest to investors and many researchers. One of the main concerns of predicting stock prices is determining what data to use to analyze variations in stock prices. Schadler and Cotten (2008) analyzed data of the American Association of Individual Investors (AAII) while Chavarnakul and Enke (2008) studied the importance of trading volumes. Bali, Demirtas, and Tehranian (2008) used the dividend payout ratio and aggregate earnings to forecast excess market returns. Harrington and Shrider (2007) and Romero-Meza, Bonilla, and Hinich (2007) analytically demonstrated the effects of events in true abnormal returns. Reilly, Wright, and Johnson (2007) focused on interest rate changes, and Peña and Rodríguez (2007) developed a model linking two stocks and bonds with the actual business cycle. Lee and Jo (1999) created a candlestick chart analysis expert system to predict the best stock market timing and future stock price movements, which can help investors obtain high returns on their stock investments. Technical indicators are often used to determine the abnormal return of stock markets (Ince & Trafalis, 2007; Terence Tai-Leung & Wing-Kam, 2008).

Another concern of many research papers is how to manipulate different tools to detect trends and find abnormal returns from a stock market. Conventional time-series models are applied to handle many forecasting problems, such as financial, economic, and weather forecasting. In stock markets, correct stock predictions can bring huge profits for investors. However, conventional time-series models produce forecasts based on some strict statistical assumptions about data characteristics, and so their forecasting accuracy rates are limited. A fusion model that combined the hidden Markov model (HMM), artificial neural networks (ANNs), and genetic algorithms (GAs) was presented in a paper by Hassan, Nath, and Kirley (2007). Huang and Tsai (2009) tried to improve the prediction accuracy and reduced the cost of training time by combining the stock price variation rate with the self-organizing feature map (SOFM) technique and a filter-based feature selection. Kim, Min, and Han (2006) used GAs that combined classifiers to predict the Korean stock price index (KOSPI). Liao, Ho, and Lin (2008) discovered stock market investment issues on the Taiwan Stock Exchange Capitalization Weighted Stock Index (TAIEX) using a two-stage data mining approach. Yu and Huarng (2008) applied neural networks (NNs) to fuzzy time-series forecasting.

It is widely accepted by many studies that non-linearity exists in financial markets, and that NNs can be effectively used to uncover such relationships (Enke & Thawornwong, 2005). Because of the superiority of NNs, Ko and Lin (2008) introduced an NN model to optimize investment weights of portfolios which showed the effectiveness of the rate on investment of the buy-and-hold trading strategy. Enke and Amornwattana (2008) developed a hybrid option trading system using a generalized regression NN to

* Corresponding author at: Program of Technology Management, Chung Hua University, Hsinchu, Taiwan, ROC. Tel.: +886 3 518 6056; fax: +886 3 518 6054.
E-mail address: schsieh@chu.edu.tw (S.C. Hsieh).

0957-4174/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2011.04.229

forecast the volatility and return of stock prices. Chang, Liu, Lin, Fan, and Ng (2009) predicted the buy/sell points of stock prices using a back-propagation network (BPN) to predict the rates of return for upwardly, steady, and downwardly trending stocks.

Zhu, Wang, Xu, and Li (2008) applied ANNs to prove that adopting the trading volume can improve the prediction performance of NNs under short-, medium-, and long-term forecasting horizons. Kim and Han (2000) combined a GA approach with ANNs to predict the stock price index. Huang, Yang, and Chuang (2008) utilized ANNs, support vector machines (SVMs), and ARIMA models which adopted 23 technical indices and then used a voting scheme to predict trends in the Korean and Taiwanese stock markets. Hyup Roh (2007) proposed hybrid models with NN and time-series models to forecast the volatility of the stock price index. However in previous studies, researchers set the parameters of ANNs intuitively or by trial-and-error processes to obtain the results. However, an alternative means of applying ANNs was proposed to improve the conventional Taguchi parameter design, and it is capable of effectively treating continuous parameter values. Since the conventional experimental design method and the Taguchi method are useful tools for determining optimal process parameter settings (Chen, Hsu, Hsieh, & Tai, 2010), this paper integrates design of experiment (DOE), the Taguchi method, and a back-propagation neural network (BPNN) to construct a robust engine and further optimize the prediction accuracy.

The remainder of this paper is organized as follows. Section 2 describes the optimization methodologies including NNs, DOE, and the Taguchi method. Section 3 proposes the detailed research model and formulae. Experimental results and discussion are presented in Section 4. Some concluding remarks are made in Section 5.

2. Optimization methodologies

Optimization methodologies including NNs, traditional experimental design, and Taguchi's parameter design method adopted to develop the proposed approach are briefly introduced below.

2.1. Neural networks (NNs)

The use of ANNs or NNs is well accepted in the arenas of telecommunications, signal processing, pattern recognition, prediction, process control, financial analysis, etc. (Widrow, Rumelhart, & Lehr, 1994). Much literature adopted a BPNN that has the advantage of a fast response and high learning accuracy (Chen, Fu, Tai, & Deng, 2009; Chen & Hsu, 2007; Dai & MacBeth, 1997; Maier & Dandy, 1998; Yao, Yan, Chen, & Zeng, 2005). The superiority of a network's functional approach depends on the network architecture, parameters, and problem complexity. If inappropriate network architecture and parameters are selected, then the analytical results may be undesirable. Conversely, analytical results will be more significant if a good network architecture and parameters are selected. The BPNN consists of the input layer, hidden layer, and output layer. Parameters for a BPNN include: the number of hidden layers, number of hidden neurons, learning rate, momentum, etc. All of these parameters have significant impacts on the performance of NNs. Fogel (1991) proposed a final information statistic (FIS) process based on the Akaike information criterion (AIC) to determine the number of hidden layers and neurons. A limitation of Fogel's research was that the process could only perform simple binary classifications. Murata and Yoshizawa (1994) and Onoda (1995) respectively proposed methods to improve the AIC. These two methods, respectively called the network information criterion (NIC) and NN information criterion (NNIC), use statistical probability and an error energy function to determine the number of hidden neurons.

In this research, the steepest descent method (or gradient descent learning algorithm) was used to find the weight change and minimize the error energy function. The activation function is a hyperbolic tangent function. According to previous studies (Cheng & Tseng, 1995; Haykin, 1999; Hush & Horne, 1993), there are a few conditions for network learning termination: (1) when the root mean square error (RMSE) between the expected value and network output value is reduced to a preset value; (2) when the preset number of learning cycles is reached; and (3) when cross-validation takes place between the training samples and test data. The first two methods are related to preset values. This research adopts the first and second approaches by gradually increasing the network training time to gradually decrease the RMSE until it is stable and acceptable. The RMSE is defined as follows:

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(d_i - y_i)^2}; \tag{1}$$

where N, d_i, and y_i are the number of training samples, the actual value for training sample i, and the predicted value of the NN for training sample i, respectively.

In network learning, input information and output results are used to adjust the weighting values of the network. The more-detailed the input training classification and the greater the amount of learning information which are provided, the better the output will conform to the expected result. Since the learning and verification data for the BPNN are limited by functional values, the data must be normalized by the following equation:

$$PN = \frac{P - P_{\min}}{P_{\max} - P_{\min}} \times (D_{\max} - D_{\min}) + D_{\min}; \tag{2}$$

where PN is the normalized data, P is the original data, $P_{\max}$ is the maximum value of the original data, $P_{\min}$ is the minimum value of the original data, $D_{\max}$ is the expected maximum value of the normalized data, and $D_{\min}$ is the expected minimum value of the normalized data. When applying an NN to a system, the input and output values of the NN fall in the range of $[-0.9, 0.9]$.

2.2. Design of experiment (DOE)

DOE is a critical methodology for applying statistical analytical tools to systematic experiments. DOE allows experimenters and engineers to develop a mathematical model to predict output variables or responses over interacting input variables. It can be used for a wide range of experiments with various purposes which includes nearly all engineering and science arenas and even marketing studies. In general, one can learn about the investigated procedures by using the DOE process, i.e., screening important factors, determining interacting factors, setting up a mathematical model for prediction, and optimizing the responses (Islam & Lye, 2009).

In engineering, one often-used approach is the best-guess (with engineering judgment) or the trial-and-error approach. Another strategy of experimentation prevalent in practice is the one-factor-at-a-time (OFAT) approach. Indeed, in the OFAT design, there is no way to determine interacting factors (Berger & Maurer, 2002), and the OFAT method can be considered a standard, systematic, and accepted method of scientific experimentation. Both of these methods were proven to be inefficient and disastrous. These methods of experimentation became outmoded in the early 1920s when Ronald A. Fisher created much more efficient methods of experimentation based on factorial designs. This type of experimental design includes the general factorial, two-level factorial, fractional factorial, and response surface methodology among

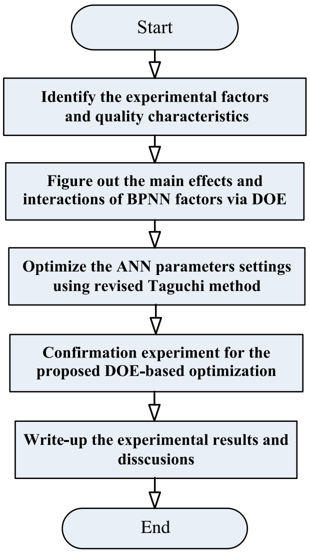

Fig. 1. Flowchart for artificial neural network stock return forecasting.

others. Thus, there are not only cost savings in using factorial designs; it is also the more-correct and -complete method of experimental design. However, these statistics-based experimental design methods are still not extremely widespread due to a lack of adequate training in basic statistical concepts and tools by product designers and process engineers (Montgomery, 2009).

Several aspects of the overall DOE method are shown and classified according to different aims such as the comparative objective, i.e., the primary goal of the experimental design is to determine whether or not that particular factor is significant; the screening objective, i.e., the purpose of the experiment series is to select or screen out the few important main effects from many less-important ones; response surface methodology (RSM) objective, i.e., the experiment is designed to allow the estimation of factor interactions and even quadratic effects; and optimal fitting of a regression model objective, i.e., if the experimental response is modeled as a mathematical function (either known or empirical) of a few continuous factors, then the model parameters have to be properly estimated using a multi-linear regression design. Once a suitable approximation for the true functional relationship between the independent variables and the surface response is found, the response variables can be optimized (Myers, Montgomery, & Anderson-cook, 2009).

Table 2
Information on the factors' assumed settings at different design of experiment levels.

Item	Control factor		Level 1	Level 2
A	Neuros	Number of neurons in the hidden layer	6	24
B	Lr	Learning rate	0.1	0.9
C	Mt	Momentum	0.5	0.9
D	Epochs	Number of epochs	10,000	60,000

2.3. Taguchi method

Taguchi's robust parameter design is a systematic method which normally selects an appropriate formulation of the signal/noise (S/N) ratio and calculates an S/N ratio for each treatment. There are three types of S/N ratios: nominal the best, the larger the better, and the smaller the better. Most engineers choose the highest S/N ratio treatment as the preliminary optimal initial process parameter setting. The Taguchi method was also used to design the parameters for NNs in previous research (Khaw, Lim, & Lim, 1995; Santos & Ludermir, 1999). Khaw et al. (1995) applied the Taguchi method to design parameters and verified that the method could rapidly and robustly design optimal parameters. Santos and Ludermir (1999) applied a factorial design to assist the design and implementation of an NN. The formulae of the three types of S/N ratios are given as follows:

$$\text{nominal the best} : S/N = 10 \times \log\left(\frac{\bar{y}^2}{\bar{S}^2}\right), \tag{3}$$

$$\text{the larger the better} : S/N = -10 \times \log\left(\frac{1}{n}\sum_{i=1}^{n}\frac{1}{y_i^2}\right) \cong -10 \times \log\left(\frac{1}{\bar{y}^2} + \frac{3\bar{S}^2}{\bar{y}^4}\right), \text{and} \tag{4}$$

$$\text{the smaller the better} : S/N = -10 \times \log\left(\frac{1}{n}\sum_{i=1}^{n} y_i^2\right) = -10 \times \log[\bar{y}^2 + \bar{S}^2]; \tag{5}$$

Table 1
Formulae of the proposed technical indicators.

Name of the indicator	Calculation formula (The definitions of variables see notes listed at the bottom)
Stochastic % D	$\frac{2}{3} \times K_{t-1} + \frac{1}{3}RSV$ $\left(RSV = \frac{C_t - L_n}{H_n - L_n} \times 100\right)$
Stochastic % K	$\frac{2}{3} \times D_{t-1} + \frac{1}{3}K_t$
Relative Strength Indicator (RSI)	$100 - \frac{100}{(1+RS)}$ $(RS = \frac{UP_{AVG(t)}}{DOWN_{AVG(t)}};$ $UP_{AVG(t)} = UP_{AVG(t-1)} \times (4/5) + UP_t \times (1/5)$ $DOWN_{AVG(t)} = DOWN_{AVG(t-1)} \times (4/5) + DOWN_t \times (1/5))$
Difference (DIF)	$EMA_{26} - EMA_{12}$ $(DI = (2 \times C + H_n + L_n)/4; SC = 2/(1+n)$ $EMA_t = EMA_{t-1} + SC(DI - EMA_{t-1})$
Moving average convergence and divergence (MACD)	$MACD_{t-1} + SC(DIF - MACD_{t-1})$
William's% R	$100 - \frac{H_n - C}{H_n - L_n} \times 100$
Ranking of relative return	$\frac{R_i}{M} \times 100$

Note:
1. RSV, raw stochastic value.
2. The first K_{t-1} and D_{t-1} are set to 50.
3. C_t, closing price of t; L_n, the lowest price during n days; H_n, the highest price during n days; n = 9 for Stochastic % D; n = 12 for the DIF; n = 5 for the $W\%R$.
4. $UP_{AVG(t-1)}$, the average value of the variation when the stock price goes up; UP_t, the increase in the stock price; $DOWN_t$, the decrease in the stock price; $DOWN_{AVG(t-1)}$, the average value of the variation when the stock price goes up; t = 5 in this case.
5. EMA, exponential moving average.
6. R_i, ranking of rate of daily return of a stock, the higher return the higher ranking; M, total number of stocks in the market.

where y_i is the response value of a specific treatment under i replications, n is the number of replications, $\bar{y}$ is the average of all y_i values, and $\bar{S}$ is the standard deviation of all y_i values.

A crucial procedure using the Taguchi method can be conducted as follows: using an analysis of variance (ANOVA) of the S/N ratio, determine factors having a significant effect on the S/N ratio; then identify levels of these factors to maximize the overall S/N ratio through the main-effect graphs of the S/N ratio. In case interactions are significant, information obtained from the plots of interactions is used to determine the optimal setting for the corresponding factors.

Table 3
Design-of-experiment results for testing R and hit ratio variables.

StdOrder	RunOrder	CenterPt	Blocks	Neuros	Lr	Mt	Epochs	Testing R	Hit ratio
11	1	1	1	6	0.9	0.5	60,000	0.020538	0.74
12	2	1	1	24	0.9	0.5	10,000	0.098521	0.72
6	3	1	1	24	0.1	0.9	10,000	0.077975	0.64
15	4	1	1	6	0.9	0.9	10,000	0.14354	0.78
16	5	1	1	24	0.9	0.9	60,000	0.070947	0.56
2	6	1	1	24	0.1	0.5	60,000	0.056034	0.72
17	7	0	1	15	0.5	0.7	35,000	−0.00198	0.38
1	8	1	1	6	0.1	0.5	10,000	0.259911	0.74
4	9	1	1	24	0.9	0.5	10,000	0.200272	0.74
10	10	1	1	24	0.1	0.5	60,000	0.036916	0.7
5	11	1	1	6	0.1	0.9	60,000	0.190672	0.74
3	12	1	1	6	0.9	0.5	60,000	0.127319	0.74
18	13	0	1	15	0.5	0.7	35,000	−0.03441	0.66
7	14	1	1	6	0.9	0.9	10,000	0.137458	0.6
13	15	1	1	6	0.1	0.9	60,000	0.148527	0.72
14	16	1	1	24	0.1	0.9	10,000	0.121676	0.7
9	17	1	1	6	0.1	0.5	10,000	0.229522	0.76
8	18	1	1	24	0.9	0.9	60,000	0.10492	0.46
19	19	0	1	15	0.5	0.7	35,000	0.239973	0.72

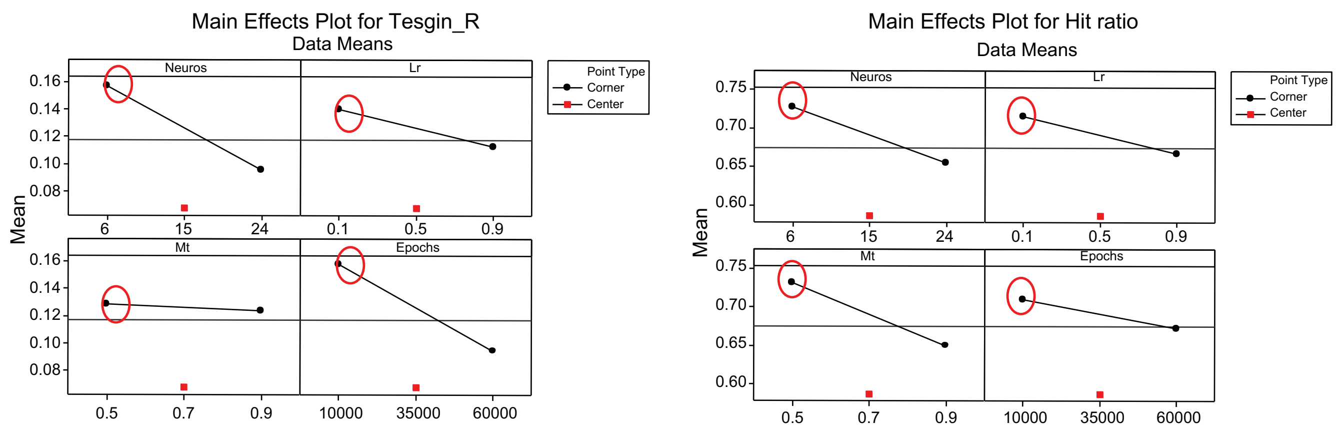

Fig. 2. Main effects plot of the BPNN design-of-experiment factors for discrete characteristics.

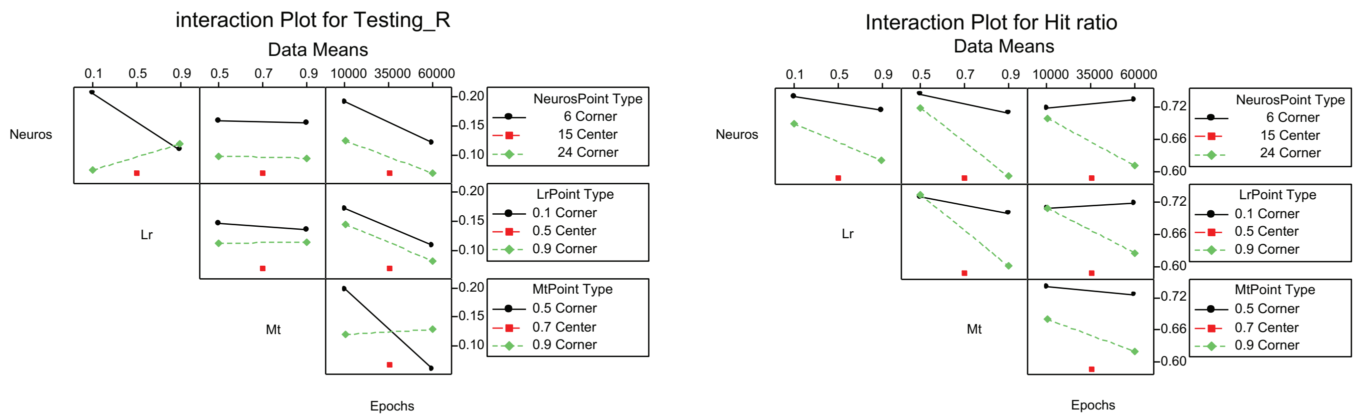

Fig. 3. Interactive plot of the BPNN design-of-experiment factors for discrete characteristics.

3. Model construction and formulae

3.1. Model construction

A flowchart for the proposed ANN stock returns forecasting model based on Chen et al. (2010) can be separated into four steps: step 1 of identifying the experimental factors and quality characteristics; step 2 of figuring out the main effects and interactions of BPNN factors via the DOE; step 3 of optimizing ANN parameter settings using the revised Taguchi method; and step 4 of running a confirmation experiment for the proposed DOE-based optimization as shown in Fig. 1.

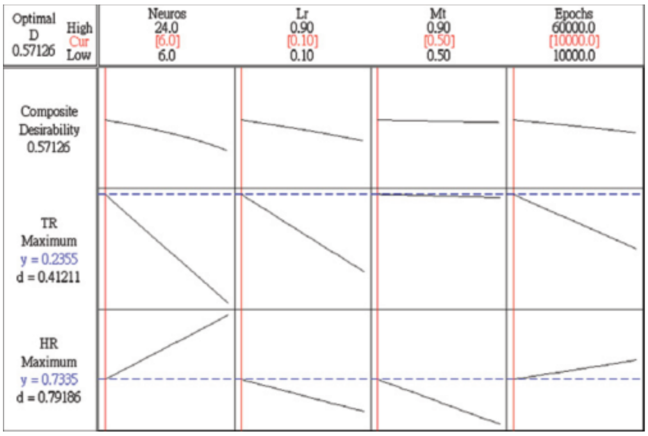

Fig. 4. Optimization plot of the BPNN design-of experiment factors for discrete characteristics.

Table 4
Information on the factors' assumed settings via the revised Taguchi method.

Item	Control factor	Level 1	Level 2	Level 3
Neuros	Number of neurons in the hidden layer	6	5	
Lr	Learning rate	0.1	0.08	0.06
Mt	Momentum	0.5	0.45	0.4
Epochs	Number of epochs	10,000	9000	8000

3.2. Formulae of research data

This research proposes an integrated approach to effectively improve the forecasting agreement effect compared to the one-step sign rate using a DOE-based BPNN predictor. The proposed study integrates a traditional experimental parameter design method and BPNN to analyze data from January 1 to 30 June, 2009 of the listed companies of the Taiwan Stock Exchange (TWSE). Because of the high turnover rate of stocks traded on the TWSE, technical analysis is a common and useful tool for analyzing short-term variations in the stock market. Kim and Han (2000) used technical indicators to predicate the daily change in direction of the KOSPI. To predict the stock price index by applying the practitioner's approach to the technical analysis, it is possible to more-effectively capture the information content in past prices (Loh, 2007). The forecasting formula of the BPNN is:

$$VR_t = f(TI_{i,t-1}; \quad i = 1, \ldots, 7). \tag{6}$$

VR_t shows the variation rate of the daily TAIEX index, that is $VR_t = \frac{(I_t - I_{t-1})}{I_t}$, where I_t is the market index of the TAIEX at time t. TI is the technical indicators of the listed companies adopted. The technical indicators used here are listed in Table 1. This paper uses technical indicators of seven listed companied as input components. Zhu et al. (2008) forecasted the market index using top-related companies. Different from research findings, the related companies depend on the β value of the listed companies. To find the listed companies for research, the β coefficient of a company was adopted. The formula for the β value of a stock, β_a is:

$$\beta_a = \frac{Cov(r_a, r_m)}{Var(r_m)}; \tag{7}$$

where r_a measures the rate of return of the stock, r_m measures the rate of return of the market, and $Cov(r_a, r_m)$ is the covariance between the rates of return. The β coefficient is a key parameter in finance; the β value of a stock shows the relation of its returns with that of the financial market as a whole. β is also a measure of the financial elasticity or correlated relative volatility of the market. The market (TAIEX) itself has an underlying β value of 1.0, and individual stocks (listed companies) are ranked according to how much they deviate from the market. An asset with a β value of 0 means that its price is not at all correlated with the market. A positive β value means that the asset generally follows the market. A negative β value shows that the asset inversely follows the market, i.e., the asset generally decreases in value when the market goes up and

Table 5
Experimental results for Y (the hit ratio) variable through the revised Taguchi method.

Neuros	Lr	Mt	Epochs	$Y1$	$Y2$	$\bar{Y}$	$\bar{S}$	S/N ratio
6	0.1	0.5	10,000	0.76	0.72	0.74	0.028284271	−2.624883811
6	0.1	0.45	9000	0.82	0.78	0.8	0.028284271	−1.946344131
6	0.1	0.4	8000	0.68	0.76	0.72	0.056568542	−2.893583337
6	0.08	0.5	10,000	0.74	0.78	0.76	0.028284271	−2.392751936
6	0.08	0.45	9000	0.78	0.72	0.75	0.042426407	−2.519626444
6	0.08	0.4	8000	0.72	0.82	0.77	0.070710678	−2.325161327
6	0.06	0.5	9000	0.6	0.7	0.65	0.070710678	−3.818903493
6	0.06	0.45	8000	0.8	0.74	0.77	0.042426407	−2.289967798
6	0.06	0.4	10,000	0.72	0.8	0.76	0.056568542	−2.419835863
5	0.1	0.5	8000	0.74	0.82	0.78	0.056568542	−2.192386908
5	0.1	0.45	10,000	0.76	0.72	0.74	0.028284271	−2.624883811
5	0.1	0.4	9000	0.78	0.78	0.78	0	−2.158107946
5	0.08	0.5	9000	0.74	0.76	0.75	0.014142136	−2.501091038
5	0.08	0.45	8000	0.8	0.78	0.79	0.014142136	−2.04954585
5	0.08	0.4	10,000	0.74	0.68	0.71	0.042426407	−2.998101132
5	0.06	0.5	8000	0.7	0.64	0.67	0.042426407	−3.504634231
5	0.06	0.45	10,000	0.76	0.74	0.75	0.014142136	−2.501091038
5	0.06	0.4	9000	0.8	0.76	0.78	0.028284271	−2.166674852

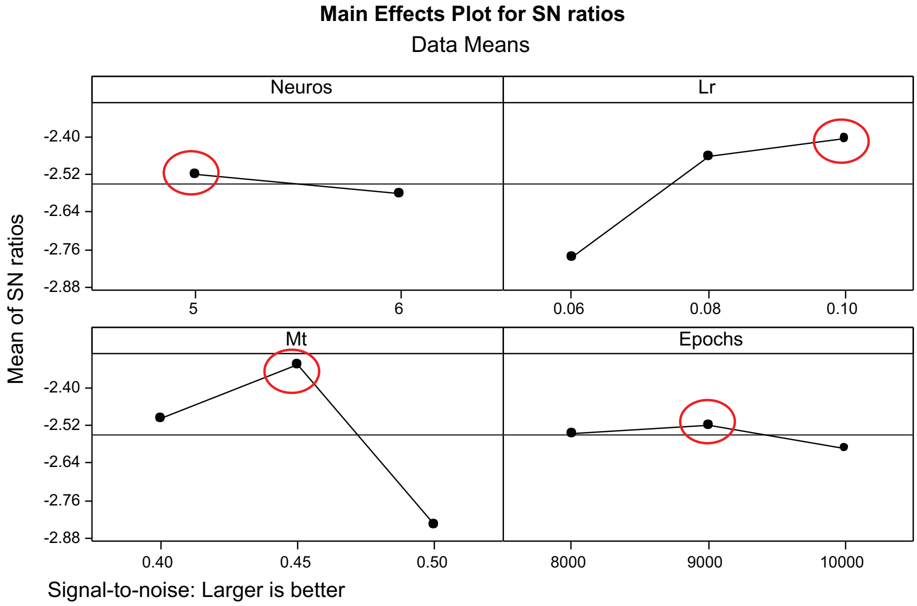

Fig. 5. Main-effects plots of BPNN factors via the revised Taguchi method.

Table 6
Confirmation experimental results for the final optimal parameter settings.

Neuros	Lr	Mt	Epochs	$Y1$	$Y2$	$Y3$	$Y4$	$Y5$	$\overline{Y}$
6	0.1	0.45	9000	0.84	0.78	0.80	0.82	0.74	0.80

vice versa. If a stock moves less than the market, the absolute value of the stock's β value is <1.0. This paper calculates the β values of the listed companies of the TWSE using five companies the β coefficients of which are very close to 1. To evaluate the efficiency of the forecasting rate, the paper adopts (Zhu et al., 2008) one-step sign prediction rate to calculate the hit ratio.

4. Experimental results and discussion

Step 1. Identify the experimental factors and quality characteristics

In estimating the quality characteristics, the proposed DOE-based research can be represented by two output variables: testing R (TR, Pearson correlation), and the hit ratio (HR, one-step sign with the same sign as the phase prediction rate), and both of them are the-larger-the-better quality characteristics, while the experimental factors include four input parameters: neurons of the hidden layer, the learning rate, momentum, and epochs.

In this research, we applied the DOE and revised Taguchi method to obtain the optimal parameter settings of the BPNN. Since the number of hidden layers did not have a significant effect on convergence, the number of hidden layer was set to 1; the transfer function (or active function) was set as the hyperbolic tangent for a data-normalized range of $[-0.9, 0.9]$. The controlling DOE factors are the number of hidden neurons (Neuros), learning rate (Lr), momentum (Mt), and the number of epochs (Epochs). Information on the factors' assumptive settings at different DOE levels is listed in Table 2, and the experimental results for DOE planning are given in Table 3.

Step 2. Figure out the main effects and interactions of BPNN factors via DOE

For the statistical graphic analysis of the aforementioned experimental data, the main effects (A–D) and their interactions are plotted in Figs. 2 and 3.

From the above main-effects plot and optimization plot (see Fig. 4), the optimal parameter settings of the proposed DOE-based BPNN predictor can be shown as Neuros of 6, Lr of 0.1, Mt of 0.5, and Epochs of 10,000.

Step 3. Optimize the ANN parameter settings using the revised Taguchi method

Apart from the DOE, the Taguchi method focused on an orthogonal array experiment regardless of the interactions. Under the condition of the four factors, one for two levels and three for three levels, and no interactions among the four factors, the total degrees of freedom were 7 (i.e., $1 \times (2-1) + 3 \times (3-1)$). An L_{18} $(2^1 \times 3^3)$ orthogonal array was suitable for arranging the factors and carrying out the experiment. The proposed Taguchi method is a novel concept not only in that it determines the equal-distance value but also that it determines a better trend for the factor levels in terms of the main-effects plots of the DOE in step 2 (i.e., the best parameter settings are arranged in level 1 and then set up better trend values in levels 2 and 3 with respect to the previous main-effects plots of the BPNN's DOE factors), as seen in Table 4. In this experiment, there were two replications, and the predicted performance of Y (the hit ratio) is an evaluation value for different combinations of factor levels. $\overline{Y}$ is the average of the two Y's in each replication. The optimal combination of factor levels is determined by the largest S/N ratio, and the experimental results of the revised Taguchi method are shown in Table 5. Subsequently, the main-effects plot of BPNN factors by the revised Taguchi method are drawn in Fig. 5. Finally, the optimal combination of factor levels is represented by the following: a BPNN architecture of 4–5–1, a hyperbolic tangent transfer function, the number of calculation generations of 9000, a learning rate of 0.1, and a momentum of 0.45.

Step 4. Confirmation experiment for the proposed DOE-based optimization

For the confirmation experiment of the final optimal parameter settings and the response of Y (the HR), this research conducted five replication tests with the number of epochs of 9000, a learning rate of 0.1, and a momentum of 0.45, and the results are presented in Table 6.

5. Conclusions

Investing in the stock market is an important financial practice for many people around the world. Determining how to improve the prediction rate of stock returns is a great concern of many investors and researchers, and ANNs were proven to be an effective forecasting tool for stock return forecasting by many researchers. Most researchers use trial-and-error to determine the parameters of ANNs, so it is difficult to obtain sound prediction rates in financial arenas.

This research integrated a conventional experimental design, Taguchi's parameter design method, and a back-propagation neural network (i.e., a design-of-experiment-based optimization) to improve the forecasting rate. For the short-term stock return forecasting compared to Zhu et al. (2008), the one-step sign prediction rate for short-term forecasting was revealed in three markets, DJIA, NASDAQ and STI, and they ranged 53.275–64.73%. The above-described experimental validation of the optimal parameter settings can effectively improve the forecasting rate to 84%. Thus, the proposed approach was proven feasible and effective for enhancing the accuracy of stock price variation predictions.

References

Bali, T. G., Demirtas, K. O., & Tehranian, H. (2008). Aggregate earnings, firm-level earnings, and expected stock returns. *Journal of Financial Quantitative Analysis, 43*(3), 657–684.

Berger, P. D., & Maurer, R. E. (2002). *Experimental design with applications in management. Engineering and the sciences*. Duxbury Press.

Chang, P.-C., Liu, C.-H., Lin, J.-L., Fan, C.-Y., & Ng, C. S. P. (2009). A neural network with a case based dynamic window for stock trading prediction. *Expert Systems with Applications, 36*(3, Part 2), 6889–6898.

Chavarnakul, T., & Enke, D. (2008). Intelligent technical analysis based equivolume charting for stock trading using neural networks. *Expert Systems with Applications, 34*(2), 1004–1017.

Chen, W. C., & Hsu, S. W. (2007). A neural-network approach for an automatic LED inspection system. *Expert Systems with Applications, 33*(3), 531–537.

Chen, W. C., Fu, G. L., Tai, P. H., & Deng, W. J. (2009). Process parameter optimization for MIMO plastic injection molding via soft computing. *Expert Systems with Applications, 36*(2), 1114–1122.

Chen, W. C., Hsu, Y. Y., Hsieh, L. F., & Tai, P. H. (2010). A systematic optimization approach for assembly sequence planning using Taguchi method, DOE, and BPNN. *Expert Systems with Applications, 37*, 716–726.

Cheng, C. S., & Tseng, C. A. (1995). Neural network in detecting the change of process mean value and variance. *Journal of the Chinese Institute of Industrial Engineers, 12*(3), 215–223.

Dai, H., & MacBeth, C. (1997). Effects of learning parameters on learning procedure and performance of a BPNN. *Neural Networks, 10*(8), 1505–1521.

Enke, D., & Amornwattana, S. (2008). A hybrid derivative trading system based on volatility and return forecasting. *Engineering Economist, 53*(3), 259–292.

Enke, D., & Thawornwong, S. (2005). The use of data mining and neural networks for forecasting stock market returns. *Expert Systems with Applications, 29*(4), 927–940.

Fogel, D. B. (1991). An information criterion for optimal neural network selection. *IEEE Transaction on Neural Network, 2*(5), 490–497.

Harrington, S. E., & Shrider, D. G. (2007). All events induce variance: Analyzing abnormal returns when effects vary across firms. *Journal of Financial & Quantitative Analysis, 42*(1), 229–256.

Hassan, M. R., Nath, B., & Kirley, M. (2007). A fusion model of HMM, ANN and GA for stock market forecasting. *Expert Systems with Applications, 33*(1), 171–180.

Haykin, S. (1999). *Neural networks: A comprehensive foundation*. Ontario, Canada: Prentice Hall.

Huang, C.-J., Yang, D.-X., & Chuang, Y.-T. (2008). Application of wrapper approach and composite classifier to the stock trend prediction. *Expert Systems with Applications, 34*(4), 2870–2878.

Huang, C.-L., & Tsai, C.-Y. (2009). A hybrid SOFM-SVR with a filter-based feature selection for stock market forecasting. *Expert Systems with Applications, 36*(2, Part 1), 1529–1539.

Hush, D. R., & Horne, B. G. (1993). Progress in supervised neural networks. *IEEE Signal Processing Magazine*, 8–39.

Hyup Roh, T. (2007). Forecasting the volatility of stock price index. *Expert Systems with Applications, 33*(4), 916–922.

Ince, H., & Trafalis, T. B. (2007). Kernel principal component analysis and support vector machines for stock price prediction. *IIE Transactions, 39*(6), 629–637.

Islam, M. F., & Lye, L. M. (2009). Combined use of dimensional analysis and modern experimental design methodologies in hydrodynamics experiments. *Ocean Engineering, 36*(3–4), 237–247.

Khaw, J. F. C., Lim, B. S., & Lim, L. E. N. (1995). Optimal design of neural network using the Taguchi method. *Neurocomputing, 7*, 225–245.

Kim, K.-J., & Han, I. (2000). Genetic algorithms approach to feature discretization in artificial neural networks for the prediction of stock price index. *Expert Systems with Applications, 19*(2), 125–132.

Kim, M.-J., Min, S.-H., & Han, I. (2006). An evolutionary approach to the combination of multiple classifiers to predict a stock price index. *Expert Systems with Applications, 31*(2), 241–247.

Ko, P.-C., & Lin, P.-C. (2008). Resource allocation neural network in portfolio selection. *Expert Systems with Applications, 35*(1–2), 330–337.

Lee, K. H., & Jo, G. S. (1999). Expert system for predicting stock market timing using a candlestick chart. *Expert Systems with Applications, 16*(4), 357–364.

Liao, S.-H., Ho, H.-h., & Lin, H.-w. (2008). Mining stock category association and cluster on Taiwan stock market. *Expert Systems with Applications, 35*(1–2), 19–29.

Loh, E. Y. L. (2007). An alternative test for weak form efficiency based on technical analysis. *Applied Financial Economics, 17*(12), 1003–1012.

Maier, H. R., & Dandy, G. C. (1998). Understanding the behaviour and optimising the performance of back-propagation neural networks: An empirical study. *Environmental Modelling & Software, 13*, 179–191.

Montgomery, D. C. (2009). *Design and analysis of experiments* (7th ed.). New York: Wiley.

Murata, N., & Yoshizawa, S. (1994). Network information criterion-determining the number of hidden units for an artificial neural network model. *IEEE Transaction on Neural Network, 5*, 865–872.

Myers, R. H., Montgomery, D. C., & Anderson-cook, C. M. (2009). *Response surface methodology: Process and product optimization using designed experiments* (3rd ed.). New York: Wiley.

Onoda, T. (1995). Neural network information criterion for optimal number of hidden units. *Proceedings of the IEEE International Conference on Neural Networks, 1*, 270–280.

Peña, J. I., & Rodríguez, R. (2007). On the economic link between asset prices and real activity. *Journal of Business Finance & Accounting, 34*(5/6), 889–916.

Reilly, F. K., Wright, D. J., & Johnson, R. R. (2007). Analysis of the interest rate sensitivity of common stocks. *Journal of Portfolio Management, 33*(3), 85–107.

Romero-Meza, R., Bonilla, C. A., & Hinich, M. J. (2007). Nonlinear event detection in the Chilean stock market. *Applied Economics Letters, 14*(13), 987–991.

Santos, M. S., Ludermir, B. (1999). Using factorial design to optimize neural networks. In *International joint conference on IEEE neural networks* (Vol. 2, pp. 857–861), Washington, DC.

Schadler, F. P., & Cotten, B. D. (2008). Are the AAII stock screens a useful tool for investors? *Financial Services Review, 17*(3), 185–201.

Terence Tai-Leung, C., & Wing-Kam, N. (2008). Technical analysis and the London stock exchange: Testing the MACD and RSI rules using the FT30. *Applied Economics Letters, 15*(14), 1111–1114.

Widrow, B., Rumelhart, D. E., & Lehr, M. A. (1994). Neural networks: Applications in industry, business and science. *Communications of the ACM, 37*, 93–105.

Yao, S., Yan, B., Chen, B., & Zeng, Y. (2005). An ANN-based element extraction method for automatic mesh generation. *Expert Systems with Applications, 29*, 193–206.

Yu, T. H.-K., & Huarng, K.-H. (2008). A bivariate fuzzy time series model to forecast the TAIEX. *Expert Systems with Applications, 34*(4), 2945–2952.

Zhu, X., Wang, H., Xu, L., & Li, H. (2008). Predicting stock index increments by neural networks: The role of trading volume under different horizons. *Expert Systems with Applications, 34*(4), 3043–3054.