Research on Stock Index Futures Trading Strategy and Risk Management Based on Genetic Algorithm

Wang Yufeng

School of Business Management, Qinghai Nationalities University, Xining, 810007, Qinghai, China

Keywords: Genetic algorithm, Stock index futures trading strategy, Risk management

Abstract: Neural network integration technology can effectively improve the learning ability and generalization ability of neural networks, and has become a research hotspot in the field of machine learning and neural computing. This paper makes use of different neural network algorithms to generate neural network ensemble individuals. Using the criterion of minimizing the sum of squared errors, we use the genetic algorithm to dynamically solve the non-negative weight coefficients of the integrated individuals to perform optimal combinational integration modeling research. And to establish a stock market prediction model. Through the Shanghai Stock Exchange Index opening price, closing price for example analysis.

1. Introduction

The calculation results show that this method is relatively simple and simple average integration model, and has the characteristics of high prediction accuracy, good stability and easy operation. People in today's society are increasingly aware of their investment and investment, and more and more investors have turned their eyes to the stock market. Although stocks can bring considerable returns to investors, investors must use reasonable scientific investment strategies to select stocks for investment in order to obtain a good return on investment. In terms of fundamentals, stock selection is to evaluate the intrinsic value of listed companies. The long-term memory of the stock market makes it possible to find out the long-term correlation of stock prices or yields through data analysis. At the same time, the stock market has a nonlinearity, and the application of intelligent algorithms can improve the accuracy and robustness of the analysis (Bekiros & Diks, 2008). Many of the existing researches on the stock market mainly use the optimization algorithm to predict the stock price and stock market situation. This paper proposes an improved stock selection strategy, which uses the improved K-means clustering algorithm based on genetic algorithm and the improved genetic algorithm to analyze the financial indicators of listed companies and obtain a reasonable stock selection model. First of all, in the preliminary processing of a large number of stock samples, a K-means algorithm based on genetic algorithm is adopted to perform cluster analysis on financial indicators so as to filter stock samples. Then using the improved genetic algorithm, adaptive crossover operator and mutation operator are used to find the best stock selection model parameters, so as to find the stocks with real investment value (Groote, 1994). A plate is a group of stocks that have common characteristics. The stocks sector can be divided according to the perspectives of the listed companies, such as industry, region, and featured topics. This article selected stocks of 50 listed companies in the power generation equipment sector as samples.

Clustering is the grouping of objects into multiple classes. Objects in the same class have great similarities. At present, according to the financial status of listed companies, the domestic stock market is generally divided into: blue chip stocks, general stocks and junk stocks. Therefore, this paper adopts a modified K-means clustering algorithm based on genetic algorithm to cluster the sample data into three categories according to financial index attributes, k=3. Excluding sample stocks in junk stocks, retaining samples in blue stocks and general stocks. In this way, stocks that are not conducive to investment are filtered out from a large amount of data, which reduces the amount of data that the stock chooses later and improves the accuracy and efficiency of selection. In securities investment theory, investment strategy is a general term for a series of investment

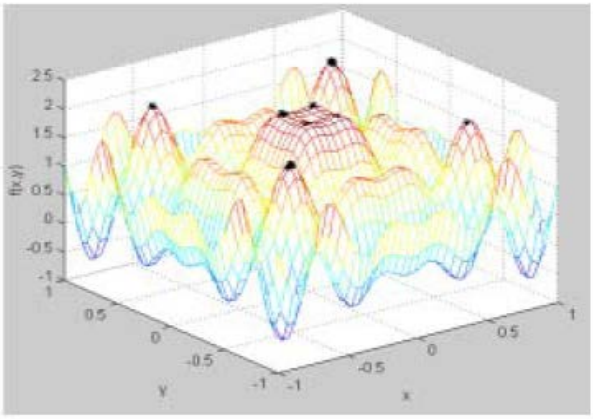
  DOI: 10.25236/ecomhs.2019.208

portfolios, divided into positive and negative investment strategies (Pesaran & Timmermann, 1994). Negative investment strategy refers to the average return obtained through fundamental analysis and portfolio investment. The higher the average return, the better the investment strategy. The average yield of stocks in portfolio investment refers to the total yield divided by the number of stocks, and the set investment is equal portfolio investment. In order to solve the problems of stock screening using genetic algorithms, it is necessary to strengthen the global search ability of genetic algorithms and eliminate too many non-optimal individuals in the process of genetic evolution to speed up the search and convergence of algorithms. In genetic algorithms, cross-operations achieve global search capabilities, and mutation operations enable local random search capabilities (Tapiero, 2005). Therefore, this paper proposes an improved adaptive crossover mutation probability to improve the global search ability of genetic algorithm in stock screening. The improvement is based on the following principles.

2. Stock market forecast

2.1. Sampling methods

A mutation operation is a change in the value of a gene at some locus in an individual string in a population. The mutation operation enables the genetic algorithm to have a local random search capability and maintain the diversity of the population to prevent premature convergence. The homogenous mutation operator replaces the original gene value with a random number from the value range of the corresponding gene position for each mutation point. The K-means algorithm is a gradient-based greedy algorithm. The traditional K-means algorithm has some weaknesses: affected by the randomly selected initial cluster centre, the algorithm easily converges to the local optimal solution, resulting in inaccurate or even wrong clustering results (Tse, 1995). The genetic algorithm's global optimization capability is used to improve the K-means clustering algorithm. The cluster centres in the population continue to iterate through selection, crossover, and mutation operations until they converge to the global optimal individuals, thereby improving the clustering quality. Since the stock trading, the prediction of the stock market index and stock price has become an important issue for the academic and securities sectors. People have been working hard to study the prediction of the behaviour of the stock market, hoping to find some rules from it, so as to minimize losses and maximize returns. In the traditional stock market forecast model research, multiple regression analysis, time series analysis, exponential smoothing, etc. are the most common methods. The stock market is a complex nonlinear dynamic system. Its prediction is a problem of estimating and extrapolating the value of a nonlinear function, and the market behaviour is influenced by a variety of factors. It has significant nonlinear and time-varying characteristics. Using traditional statistical prediction techniques, it is difficult to reveal its inherent laws. .



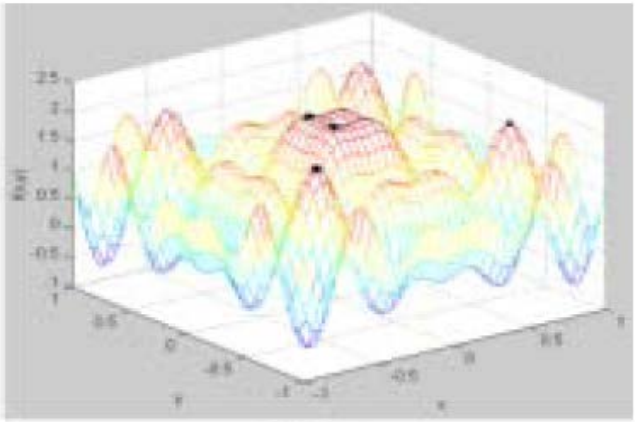
INGA

Figure 1. Genetic algorithm model

2.2. Neural network integration method based on genetic algorithm

However, due to the lack of strict theoretical guidance of the neural network method, its application effect completely depends on the user's experience. In practical applications, due to the lack of prior knowledge of the research questions, researchers often have to go through a lot of laborious and time-consuming experimentation in order to determine the appropriate network model and various parameter settings. The effect depends entirely on the user's experience, even if the same method is used to solve the same problem. Due to different operators, the results may be very different. It is often the case that neural networks have high accuracy in fitting training samples, but the prediction accuracy of new samples other than training samples is very poor, which greatly limits the application of neural networks in stock market prediction. Neural network integration uses a finite number of neural networks to learn the same problem. The output integrated in one input example is determined by the output of the neural networks that make up the integration. This method can significantly improve the generalization performance of neural network systems.

Even ordinary engineers and technicians lacking experience in neural computing can benefit from it and be considered as a very effective method of engineering neural computation. The genetic algorithm is a computational model that simulates Darwin's genetic selection and natural evolution of the biological evolution process. It is a global probabilistic search method based on biological mechanisms such as natural selection and genetic variation. In this paper, we use different neural network algorithms to generate integrated individuals. We use the least sum of squared errors as the criteria to integrate the non-negative weight coefficients of individuals with the dynamic solution of genetic algorithms to perform optimal combinational integration model research. In this way, a prediction model is established, and the example analysis is conducted through the opening and closing prices of the Shanghai Stock Exchange Index. The calculation results show that this method has the advantages of high prediction accuracy, good stability, and easy operation compared to the traditional simple average integration model.



SGA

Figure 2. Risk management system

3. Genetic algorithm optimization

3.1. Neural network prediction model

The GA algorithm is a global search algorithm. It integrates the BP neural network and the GA algorithm, and uses the GA algorithm to make up for the BP neural network connection weights and the randomness of the threshold selection. Not only can the BP neural network generalize the mapping ability, but also make the BP neural network have a fast convergence and strong learning ability. This paper combines genetic algorithm and BP neural network to form an improved genetic algorithm to optimize BP neural network prediction model. The optimal individual obtained by the genetic algorithm is decomposed into the connection weights and thresholds of the BP neural

network, and it is assigned as the initial weight and threshold of the BP neural network prediction model. After the BP neural network prediction model is trained, the chaotic time series forecast optimal solution output.

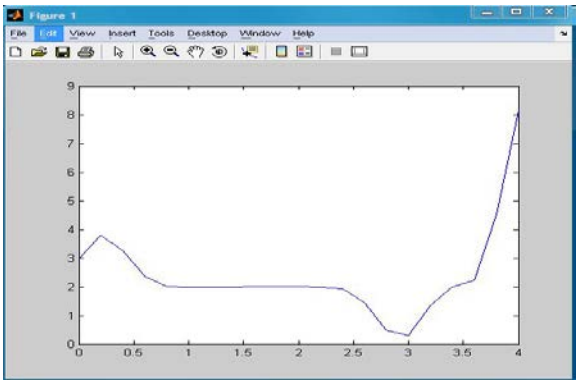


Figure 3. Mathematical function image

Aiming at the problem of BP neural network predicting the existence of local minimum defects and slow convergence speed, this paper proposes a genetic algorithm to optimize the time series prediction method of chaotic BP neural network. It was applied to the prediction of the Shanghai Composite Index and compared with the BP model. The results show that: This method reduces the probability of the BP neural network prediction model getting into local minimums and improves the convergence speed of the model. Compared with the BP prediction model, this method has better nonlinear fitting ability and higher prediction accuracy for the Shanghai Composite Index.

Test function:

$$f = 1 + x\sin(4\pi x) - y\sin(4\pi y + \pi) + \frac{\sin 6\sqrt{x^2 + y^2}}{6\sqrt{x^2 + y^2} + 0.000000000000001} \quad (1)$$

Probability measure:

$$P(A_k) = \frac{\oint_{s_1} ds}{\oint_{s} ds} = \rho \le 1 \quad (2)$$

The following new strategies for adaptively adjusting genetic algorithm control parameters are proposed:

$$P_c = \frac{1}{1 + \exp(-K_1 \bullet \Delta')} \quad (3)$$

$$P_m = \frac{1}{1 + \exp(-K_2 \bullet \Delta')} + 1.0 \quad (4)$$

This defines the variance of the t population as:

$$D_t^{(l)} = \sum [x_t^{i(i)} - \bar{x}_t^{i}]^2 / M, l \in \{1,2...L\} \quad (5)$$

The function we use is:

$$F_1 = 0.002 + \sum_{j=1}^{25} \frac{1}{j + \sum_{i=1}^{2} (x_i - a_{ij})^6} \quad (6)$$

The improved GAs presented in this paper are as follows:

$$x_i = a_i + (b_i - a_i) \bullet \sum_{j=1}^{L} (g_j^i \bullet 2^i) / (2^{L+1} - 1) \quad (7)$$

Solve unknown coefficients from initial conditions:

$$v = \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix}, \gamma = \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_n \end{bmatrix} = v^{-1} x(0) \quad (8)$$

So this equation of motion is given by:

$$x_1(t) = \frac{1}{2}\cos(\omega_1 t) + \frac{1}{2}\cos(\omega_2 t), x_2(t) = -\frac{1}{2}\cos(\omega_1 t) + \frac{1}{2}\cos(\omega_2 t) \quad (9)$$

$$\begin{bmatrix} -\frac{k_1 + k_2}{m} & \frac{k_2}{m} \\ \frac{k_2}{m} & -\frac{k_1 + k_2}{m} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \ddot{x}_1 \\ \ddot{x}_2 \end{bmatrix} \quad (10)$$

1) Introduction to rough set theory: the stock market is a “barometer“ for the development of the national economy. Stock price forecasting has always been a hot topic for investors and academic researchers. The stock market is a highly complex nonlinear system. Its changes are both internal and regular. At the same time, it is also influenced by many factors such as political events, macroeconomic conditions, and the psychological status of traders. In the past few decades, many researchers have used artificial intelligence methods to deal with such uncertainties. The stock price series has the characteristics of extreme dynamics and strong volatility, which makes it difficult for the traditional time series model to describe the changes of the stock market effectively. Therefore, the single neural network model cannot meet the requirements of stock prediction modeling.

2) Determination of fitness function: the fitness function is the only deterministic index for evaluating the individual fitness value. The form of the function directly determines the evolutionary behavior of the group. According to the definition of attribute reduction, the individual's fitness depends mainly on two aspects: the number of conditional attributes and the dependence of decision attributes on conditional attributes. It is difficult to obtain a minimal set of attribute reductions simply by means of the traditional rough set itself. Finding the minimum attribute reduction is an NP-hard problem. Genetic Algorithm is an adaptive search algorithm that simulates the genetic evolution process of organisms. Its main feature is the group search strategy and the exchange of information among individuals in the group, and its ability to optimize search in the global situation is outstanding.

3.2. Research tools and methods

When the individual fitness value is closer to the maximum fitness value of the population, the crossover rate and the mutation rate are smaller; when the two are completely equal, the values of the crossover rate and the mutation rate are zero. In the middle period of population evolution, this type of adjustment will be more appropriate for searching the solution space, but it will have an adverse effect in the early and late stages of genetic evolution. Because the better individuals in the early evolutionary population are almost in a state of no change, and the excellent individuals at this time are not necessarily the global optimal solutions, this will easily increase the probability that evolution will fall into a local optimum. In the later stages of evolution, the entire algorithm gradually converged, and the magnitude of crossover and mutation should be reduced as a whole.

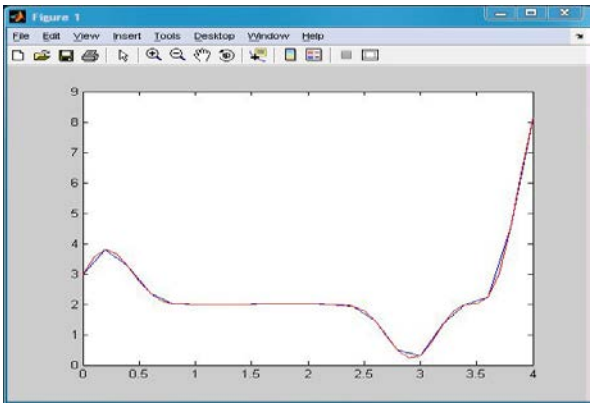


Figure 4. Image of the function's maximum value from a mathematical perspective

3.3. Stock forecasting model

The fundamental fluctuations of the stock market are regular, in other words, the movement of the stock price obeys a rhythm or an ordered sequence. According to the technical analysis theory of stock prediction, the future trend of the stock index is related to the current stock market conditions. Technical indicators are an important branch of technical analysis. Therefore, the main idea of this stock prediction model is to compose time series by a large number of auxiliary technical indicators, and to reveal the most likely trend of stock prices through machine learning. Using the technical indicators of several stock trading days after the attribute reduction as the final sample data, train NARX neural network for time series stock prediction. The stock price is a kind of data with strong randomness, which is not only influenced by the buyer and the seller, but also related to domestic and foreign policy factors. In order to improve the accuracy, the method of combining BP neural network and genetic algorithm is studied. Aiming at the problem of stochastic initial weights of BP neural network and easy to fall into the local optimization of genetic algorithm, a novel BP neural network stock price prediction model based on the optimization of adaptive genetic algorithm is proposed. This model optimizes the initial weights of BP neural network through adaptive genetic algorithm to predict the stock price. Experimental simulation results show that the model can obtain higher prediction accuracy and faster convergence rate. Since the establishment of the domestic stock market, in order to accurately predict the trend of the stock price, to avoid the risk of stock investment, and to obtain the maximum investment income, the stock price forecast has become the issue most concerned by investors.

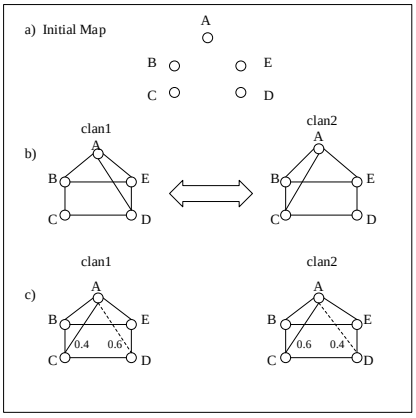


Figure 5. Multiple ant social rules

4. Empirical analysis

4.1. Genetic algorithm

Because the initial weights of the BP neural network are randomly generated, and the training speed is slow, there are problems such as local minimum values. Therefore, the combination of

genetic optimization and BP neural network can improve these defects to some extent. When genetic algorithm optimizes the initial weight of BP neural network, a fitness function needs to be set to determine the probability that the individual is selected. As the genetic algorithm continuously increases the fitness value during the search process of the optimal value, the fitness function can be set as the following expression. Determine the initial weights of the BP neural network and train the neural network until the stop request is satisfied. However, the traditional genetic algorithm, in the face of more complex optimization objective function, cannot get better optimization results, not only the optimization speed is slow, and cannot reach the standard optimization solution. In order to solve this problem, improve the ability of global search for optimal solutions and prevent from falling into local optimal solutions, a BP neural network model optimized by adaptive genetic algorithm was proposed to obtain the optimal BP neural network initial weights.

Table 1. Comparison of results of different TSP issues

TSP	MMAS	ACS	ASE	AS
Eil51	427.8	428.6	429.0	437.8
Chl130	6194.6	6260.9	6300.4	6410.1
D198	15972.8	16054.6	16207.2	16710.5
Eil51	427.8	428.6	429.0	437.8
Chl130	6194.6	6260.9	6300.4	6410.1
D198	15972.8	16054.6	16207.2	16710.5
Eil51	427.8	428.6	429.0	437.8
Chl130	6194.6	6260.9	6300.4	6410.1
D198	15972.8	16054.6	16207.2	16710.5
Eil51	427.8	428.6	429.0	437.8
Chl130	6194.6	6260.9	6300.4	6410.1
D198	15972.8	16054.6	16207.2	16710.5
Eil51	427.8	428.6	429.0	437.8
Chl130	6194.6	6260.9	6300.4	6410.1
D198	15972.8	16054.6	16207.2	16710.5

4.2. Adaptive genetic algorithm

Due to the random initial weights of BP neural network, the number of trainings is large, and the genetic algorithm easily falls into the local optimization situation during the global search. Therefore, a BP neural network model based on the optimization of adaptive genetic algorithm is given to optimize the search of BP neural network weights. From the above discussion, it can be seen that the BP neural network model based on the optimization of adaptive genetic algorithm mainly adjusts the crossover probability and the mutation probability to make the individual update the network weight continuously, thereby improving the prediction accuracy of the BP neural network.

Table 2. Solve some asymmetric TSP problems with BAS algorithm

TSP	m	f	opt.	best	worst	sol	σsol	ET	σET	GT	CPU
ry48p	400	20	14422	14507	15201	14734.1	224.0	24914	5050	30960	18.2
ftv55	400	20	1608	1612	1708	1645.7	26.7	34423	6322	42213	31.6
ftv64	400	20	1839	1842	1889	1864.2	15.2	34144	8276	42133	53.9
ft70	400	20	38673	39207	39500	39368.0	91.1	41468	6567	49520	86.8
Kro124p	400	20	36230	36956	38332	37539.6	393.6	60400	17392	72933	243.4
ftv170	400	20	2755	2808	2972	2888.7	46.7	95918	35866	119866	1295.8
ry48p	400	20	14422	14507	15201	14734.1	224.0	24914	5050	30960	18.2
ftv55	400	20	1608	1612	1708	1645.7	26.7	34423	6322	42213	31.6
ftv64	400	20	1839	1842	1889	1864.2	15.2	34144	8276	42133	53.9
ft70	400	20	38673	39207	39500	39368.0	91.1	41468	6567	49520	86.8
Kro124p	400	20	36230	36956	38332	37539.6	393.6	60400	17392	72933	243.4

4.3. Algorithm simulation and conclusion analysis

The BP neural network has certain flaws and there is a large error in the prediction of the stock price. With the BP neural network optimized by genetic algorithm, the prediction accuracy can be improved to a certain extent. However, there is a problem that the genetic algorithm has a local optimal value defect in searching for the optimal value. Aiming at these problems, a BP neural network stock prediction model based on adaptive genetic algorithm to optimize network weights is given. This not only accelerates the convergence of fitness function, but also improves the prediction accuracy of stock price relative to BP neural network model and genetic algorithm optimization BP neural network model.

Table 3. BP neural network algorithm flow chart optimized based on genetic algorithm

TSP	opt.	BAS	MMAS	dev.	ACS	AS	avg.ET	avg.GT
eil51	426	427.6	427.6	0.0	428.1	437.3	60126	72600
kro100	21282	21321.3	21320.3	<0.1	21420.0	22471.4	143346	172812
d198	15780	15979.7	15972.5	<0.1	16054.0	16702.1	294385	358266
ry48p	14422	14572.8	14553.2	0.1	14565.4	15396.4	53156	65133
ft70	38673	39260.7	39040.2	0.6	39099.0	39596.3	96176	115800
kro124p	36230	37583.1	36773.5	2.2	36857.0	38733.1	138992	168466
ftv170	2755	2880.2	2828.8	1.8	2826.5	3154.5	207203	258533
eil51	426	427.6	427.6	0.0	428.1	437.3	60126	72600
kro100	21282	21321.3	21320.3	<0.1	21420.0	22471.4	143346	172812
d198	15780	15979.7	15972.5	<0.1	16054.0	16702.1	294385	358266
eil51	426	427.6	427.6	0.0	428.1	437.3	60126	72600
kro100	21282	21321.3	21320.3	<0.1	21420.0	22471.4	143346	172812
d198	15780	15979.7	15972.5	<0.1	16054.0	16702.1	294385	358266

5. Conclusion

This article is an attempt to use the genetic algorithm in stock investment. Although it has a certain effect, the entire program is only an overall framework. In many details can be further improved by other methods: In addition to short-term predictions, we can also use genetic algorithms for long-term investment analysis. However, at present, China's stock market is not sound, and there are many long-term factors that affect stock prices, making it difficult to distinguish between factors. Therefore, it is difficult to determine the corresponding fitness function. Therefore, it is not yet ripe to use the genetic algorithm for long-term stock market analysis. The gene fitness function is a linear function of the four components. We believe that a nonlinear model of the function can be established, considering the combination of fuzzy techniques and neural networks, and dynamically determining the weight of each part to reflect the different periods. The importance of each part is not the same, making the model more authentic and objective. The final result given by the algorithm occasionally shows a loss of stock. After careful consideration, we discovered that because we chose the one with the highest fitness value, we represent a portfolio with the highest average return. So when the stock market is in a big ups and downs, the algorithm may choose to include a continuously rising stock.

References

[1] AïT-Sahalia, Y., & Lo, A. W. (1998). “Nonparametric risk management and implied risk aversion“. Journal of Econometrics, 94(1), 9-51.

[2] Bekiros, S. D., & Diks, C. G. H. (2008). “The relationship between crude oil spot and futures prices: cointegration, linear and nonlinear causality“.Energy Economics, 30(5), 2673-2685..,

[3] Christoffersen, P. F., & Diebold, F. X. (2006). “Financial asset returns, direction-of-change forecasting, and volatility dynamics“. Management Science, 52(8), 1273-1287.

[4] Groote, X. D. (1994). “The flexibility of production processes: a general framework“. Management Science, 40(7), 933-945.

[5] Kim, K. J. (2004). “Toward global optimization of case-based reasoning systems for financial forecasting“. Applied Intelligence, 21(3), 239-249.

[6] Laws, J., & Thompson, J. (2005). “Hedging effectiveness of stock index futures“. European Journal of Operational Research, 163(1), 177-191.

[7] Montagnoli, A., & Vries, F. P. D. (2010). “Carbon trading thickness and market efficiency“. Energy Economics, 32(6), 1331-1336.

[8] Pesaran, M. H., & Timmermann, A. (1994). “Forecasting stock returns an examination of stock market trading in the presence of transaction costs“. Journal of Forecasting, 13(4), 335-367.

[9] Roh, T. H. (2007). “Forecasting the volatility of stock price index“. Expert Systems with Applications, 33(4), 916-922.

[10] Tapiero, C. S. (2005). “Value at risk and inventory control“. European Journal of Operational Research, 163(3), 769-775.

[11] Tse, Y. K. (1995). “Lead-lag relationship between spot index and futures price of the nikkei stock average“. Journal of Forecasting, 14(7), 553-563.