

## DEVELOPING A MODULAR PORTFOLIO SELECTION MODEL FOR SHORT-TERM AND LONG-TERM MARKET TRENDS AND MASS PSYCHOLOGY

M. Jasemi<sup>1\*</sup> and A.M. Kimiagari<sup>2</sup>

<sup>1</sup>Department of Industrial Engineering  
Islamic Azad University, Masjed Soleyman Branch, Masjed Soleyman, Iran  
miladj@aut.ac.ir

<sup>2</sup>Department of Industrial Engineering  
Amirkabir University of Technology, Tehran, Iran  
kimiagar@aut.ac.ir

### ABSTRACT

In an effort to model stock markets, many researchers have developed portfolio selection models to maximise investor satisfaction. However, this field still needs more accurate and comprehensive models. Development of these models is difficult because of unpredictable economic, social, and political variables that affect stock market behaviour. In this paper, a new model with three modules for portfolio optimisation is presented. The first module derives the efficient frontier through a new approach; the second presents an intelligent mechanism for emitting trading signals; while the third module integrates the outputs of the first two modules. Some important features of the model in comparison with others are: 1) consideration of investors' emotions - the psychology of the market - that arises from the three above-mentioned factors; 2) significant loosening of simplifying assumptions about markets and stocks; and 3) greater sensitivity to new data.

### OPSOMMING

In 'n poging om aandelemarkte te modelleer het verskeie navorsers portefeulje-seleksie-modelle ontwikkel om beleggers se tevredenheid te maksimiseer. Desnieteenstaande word meer akkurate en omvattende modelle benodig. Die ontwikkeling van hierdie modelle word bemoeilik deur die onvoorspelbare ekonomiese, sosiale en politiese veranderlikes wat aandelemarkte se gedrag raak. In hierdie artikel word 'n nuwe model voorgedra wat bestaan uit drie modules vir portefeulje-optimalisering. Die eerste module bepaal die doelmatigheidsgrens op 'n nuwe metode; die tweede hou 'n intelligente meganisme voor om transaksies te lewer terwyl die derde module die uitsette van die eerste twee modules integreer. Sommige van die belangrike eienskappe van die model wat dit van ander onderskei is: 1) konsiderasie van die beleggers se emosies - die sielkunde van die mark - wat ontstaan vanweë die genoemde faktore; 2) betekenisvolle verslapping van die vereenvoudigende aannames oor market en aandele; en 3) verhoogde sensitiwiteit tot nuwe data.

---

\*Corresponding author

## 1. INTRODUCTION

Modern portfolio theory, based on the pioneering work of Markowitz [1,2] and Sharpe [3], is an effort to build on the traditional principles of portfolio selection. Portfolio theory has been organised to overcome the challenge of assigning one's wealth between different assets [4]. In mathematical programming, an asset is a random variable with a stochastic distribution for future returns, and a portfolio is a linear combination of these variables [5]; so any way of diversifying money between several assets can be called a portfolio [6]. Recognising the best portfolio of assets is one of the major challenges in the financial world [7], and is called 'portfolio selection'. In fact, portfolio selection is the process of building a portfolio that maximises investor satisfaction [6,8,9,10].

The Markowitz model, despite its theoretical importance, has never been extensively used to make large-scale portfolios [11]. Because of the computational difficulties that a large-scale quadratic programming problem with a dense covariance matrix has [12], there are some restricting assumptions, such as symmetric distribution of the returns. So the model has experienced much development in two directions: 'alternative portfolio selection models', and 'equilibrium models'. Some of the alternative models, with minor changes, are Mean-Semivariance, Mean-absolute deviation, Mean-Variance-Skewness, or Mean-Variance with logical constraints; and, if more fundamental changes are considered, we can include Robust Optimisation, Markov chain, Multi-objective decision-making, Possibility and Fuzzy theory, or Minimax modelling of portfolio optimisation. For the equilibrium models, the Capital Asset Pricing Models (CAPM) developed by Sharpe [13,14], Lintner [15] and Mossin [16], or the Arbitrage Pricing Theory formulated by Ross [17] and developed by Huberman [18] and Connor [19] can be named. But in spite of many improvements made by the above-mentioned contributions to portfolio theory, there is a long way to go to the ideal use of these models, and further contributions are still needed. In this paper, a portfolio selection model with three modules is presented. The first two modules are discussed in detail, but the third is only introduced. The improvements that this model makes to the theory are as follows:

- Loosening the restricting assumptions on the distribution of data
- Considering the stock market's psychology
- Making the model sensitive to both short- and long-term trends
- Decreasing reliance on raw past data, and moving towards processing them for particular purposes such as forecasting.

Sections 2 and 3 focus on the first and second modules respectively; section 4 describes and gives examples of the third module; and section 5 presents the conclusions of the study.

## 2. FIRST MODULE: ON DERIVING THE EFFICIENT FRONTIER

Deriving an efficient frontier (EF) on the basis of historical information is an essential initial step to removing inefficient portfolios; otherwise the complexity of the decision-making increases considerably [7]. The collection of portfolios that have maximum return at a specified level of risk, or have minimum risk at a specified level of return, is called the efficient frontier (EF) [1], and Ballesterro and Romero [20] recommend maximising investors' expected utility on EF. Our proposed model also considers EF as one of its bases - but with some changes to improve the results.

The model to derive EF that is generally well-known in the literature is as follows.

$$\begin{aligned} \text{Min} \quad & \text{Risk} (P(x_1, \dots, x_n)) \\ \sum_{i=1}^n x_i &= 1 \end{aligned} \tag{1}$$

$$\sum_{i=1}^n x_i \bar{r}_i = R_d \tag{2}$$

$$\sum_{i=1}^n y_i = a \quad (3)$$

$$l_i y_i \leq x_i \leq u_i y_i \quad i = 1, \dots, n \quad (4)$$

$$y_i = \begin{cases} 1 & x_i > 0 \\ 0 & x_i = 0 \end{cases} \quad (5)$$

$$x_i \geq 0 \quad i = 1, 2, \dots, n \quad (6)$$

where

*Risk* : Risk function

$x_i$  : Investment share of stock  $i$  in the portfolio

$P(x_1, \dots, x_n)$ : The portfolio whose shares of stocks are  $x_1, \dots, x_n$

$\bar{r}_i$  : Indicator of stock  $i$  past performance

$a$  : Pre-determined number of stocks in the portfolio

$R_d$  : Indicator of the portfolio past performance

$l_i$  : Lower admissible limit for investment in stock  $i$

$u_i$  : Upper admissible limit for investment in stock  $i$

In the model, constraints 1, 2 and 6 are necessary, but constraints 3, 4 and 5 (added to the Markowitz model by Perold [21]) are optional. The model also lets decision-makers choose any risk measure or any definition for  $\bar{r}_i$ .

The literature has introduced the arithmetic average of past returns of stock  $i$  as a definition for  $\bar{r}_i$ , but it does not encompass many situations, and may not be a basis for some investors' decisions. Investors may use any other statistic as a measure for the asset's past performance. For instance, a person who invests in lottery tickets considers the maximum potential return as an indicator of asset return, because the expected return of a lottery ticket is lower than many other investment opportunities with even smaller risks. The main reason for applying an arithmetic average in almost all the portfolio models is the mutual dependency of risk and return measures in this family of models. In fact, EF is the result of equilibrium between return and risk. For example, if you apply the risk measure of variance, for rational results the statistic of arithmetic average has to be used, as proposed by Markowitz.

The main contribution of our EF model is the risk measure of 'lower partial moment (LPM) of the first order' that replaces 'variance'. There is a considerable literature on the efficacy of the proposed measure. For example, Jasemi et al. [22] develop a new set of axioms for assessing risk measures to show the significance of the measure better. Unser [23] believes that LPM is of special importance for applications in financial decision-making; and symmetrical risk measures can clearly be dismissed in favor of shortfall measures [23]. Bawa [24,25], Harlow & Rao [26] and Unser [23] firmly recommend using it for the development of asset pricing models rather than earlier formulations. And it is to be noted that in the category of shortfall measures, LPM is more consistent with how individuals actually perceive risk [26].

In addition to the above literature in support of LPM, poor performance of the variance in some cases, compared with LPM, convinces us of the need to make the change. That is, if the computational difficulties associated with the variance-based models are overlooked, and if the applied data also satisfy the characteristic of being symmetric, the results still do not make sense. Let's consider the situation in which there are only two stocks with returns of  $r_1$  and  $r_2$  so that  $r_1 > r_2$ . If an investor wants to put together the best portfolio of  $r_1$  and  $r_2$ , which combination of them should be chosen? Naturally and obviously, a 100%

investment in  $r_1$  is recommended, and the associated EF would be only one point. To see the performance of variance-based EF models under the above conditions, a 301-day financial period with 300 positive returns of  $r_1$  and  $r_2$  is considered while  $r_2 = \frac{r_1}{2}$ .

According to the generated data and the arithmetic average as representative of past performance, the following model is achieved:

$$\begin{aligned} \text{Min var}(x_1 r_1 + x_2 r_2) &= x_1^2 \text{var}(r_1) + x_2^2 \text{var}(r_2) + 2x_1 x_2 \text{cov}(r_1, r_2) \\ &= 0.0107 x_1^2 + 0.0027 x_2^2 + 0.0107 x_1 x_2 \end{aligned}$$

$$x_1 + x_2 = 1$$

$$0.6053 x_1 + 0.3027 x_2 = R_d$$

$$x_1, x_2 \geq 0$$

After solving the model for 20 amounts of  $R_d$ , Table 1 with 20 points is achieved, in which each point represents a portfolio. The results of Table 1 are also shown in a chart in Fig.1. As can be seen, the model offers more than one point, while, surprisingly, a 100% investment in  $r_2$ , which in practice cannot be justified at all, is also recommended. It should be noted that the columns of  $x_1$  and  $x_2$  in Table 1 determine the share of investment in  $r_1$  and  $r_2$  respectively.

Point	$x_1$	$x_2$	Variance	Return
1	0.9957	0.0053	0.0106	0.6043
2	0.9434	0.0573	0.0098	0.5884
3	0.8905	0.1104	0.009	0.5725
4	0.8381	0.1628	0.0083	0.5566
5	0.7852	0.2158	0.0076	0.5406
...	...	...	...	...
16	0.2082	0.7908	0.003	0.3654
17	0.1556	0.8434	0.0029	0.3495
18	0.103	0.896	0.0028	0.3335
19	0.051	0.948	0.0027	0.3178
20	0.0011	0.9979	0.0027	0.3027

Table 1: The 20 portfolios that make up the EF

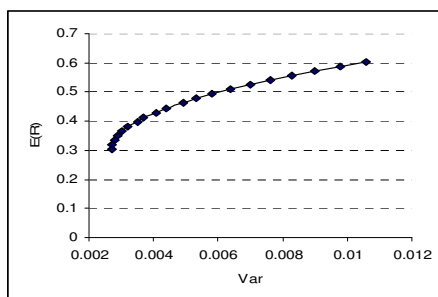


Figure 1: The EF associated with data of Table 1

After analysing the problem with additional stocks, again the same results were achieved. So it can be firmly stated that the model is not reliable, even if all of its theoretical problems are overlooked or resolved.

On the basis of what has been discussed, the new EF model that makes up the first module is as follows.

$$\text{Min } LPM \text{ 1st } (P(x_1, \dots, x_n)) = \int_{-100}^{\tau} (\tau - R_{P(x_1, \dots, x_n)}) f_{P(x_1, \dots, x_n)} dR_{P(x_1, \dots, x_n)} \tag{7}$$

$$\sum_{i=1}^n x_i = 1 \tag{7}$$

$$\sum_{i=1}^n x_i \bar{r}_i = R_d \tag{8}$$

$$x_i \geq 0 \quad i = 1, 2, \dots, n \tag{9}$$

where

$LPM \text{ 1st}$  : LPM of the first order

$f_{P(x_1, \dots, x_n)}$  : Probability density function for return of  $P(x_1, \dots, x_n)$

$R_{P(x_1, \dots, x_n)}$  : Random variable of return of  $P(x_1, \dots, x_n)$

$\bar{r}_i$  : Arithmetic average of the past returns that are greater than  $\tau$

To check the performance of the new EF model against the case discussed above, the new model is run with three amounts of 20%, 30% and 40% for  $\tau$ . Again, for each  $\tau$  20 amounts of  $R_d$  that completely cover the inputs range are assigned. Fig.2 shows the results, while - as usual - the x and y axes represent risk and return respectively.

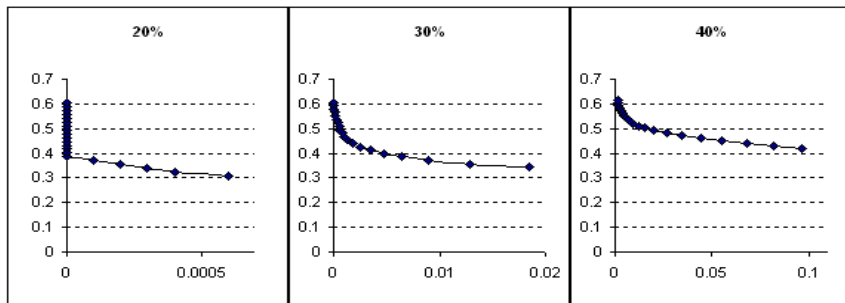


Figure 2: Results of solving the new model for three amounts of  $\tau$  and 20 amounts of  $R_d$

On the basis of the EF definition and Fig.2, for each  $\tau$  only the point on the furthest left, which recommends a 100% investment in  $r_1$ , can be accepted as EF; i.e., instead of an efficient frontier we have an efficient point, because the point offers maximum return and minimum risk, and is superior to all the others. So the Mean-LPM model generates acceptable results. It should be noted that the robustness of the new mechanism has also been confirmed by analysing cases with a greater number of assets (up to 10).

### 2.1. Comparing the new and old EF models in general cases

In this part, five sets of experiments with real data from the New York Stock Exchange (NYSE) are run, as shown in Table 2. The first five experiment sets cover respectively the periods of 1/2000-7/2009, 1/1990-12/1995, 1/1981-1/1986, 1/1971-12/1976, and 1/2006-8/2009; and they are referred to here respectively as 1<sup>st</sup>, 2<sup>nd</sup>, 3<sup>rd</sup>, 4<sup>th</sup> and 5<sup>th</sup>. In addition to the data periods, the experiments also differentiate on the basis of the stocks selected for each experiment.

Other than the above-mentioned differences, the experiments are identical, and make the following assumptions.

1. A typical EF is derived only for the next month by using the stock prices of the 13 latest months. For example, the EF for 8/1991 needs stocks prices from 1990/8 to 1991/8.
2. The input data are the closing prices of stocks at the beginning of each month.

In each experiment set, the EFs for each dataset are derived on the basis of new and old approaches. The performance of each portfolio is measured according to its return for the following month. Naturally the EFs - whether achieved by the old model or the new - are composed of several points, in which each represents an efficient portfolio; and each EF has to be totalised. In this paper, the three popular statistics of Max, Min and Mean are used. Table 3 summarises the final results of the experiments, while each grid determines the percentage of times that the model performs better than the other, according to the associated statistic. Obviously the first, second and third variance-based columns can only be compared with the first, second and third LPM-based columns respectively. After Table 3, in which the relative performances of Max, Min and Mean of the points on the EFs by both of the models are compared; Table 4 determines what would happen to the initial capital if an investor invests for the associated period in the Max, Min or Mean of variance-based or LPM-based portfolios. For example, Table 4 shows that, if for the 4<sup>th</sup> experiment set for each month only the average of efficient portfolios of the LPM-based model is used for investment during 1/1971 to 12/1976, the initial asset becomes 1.867 times bigger.

No	The stocks	1 <sup>st</sup>	2 <sup>nd</sup>	3 <sup>rd</sup>	4 <sup>th</sup>	5 <sup>th</sup>	No	The stocks	1 <sup>st</sup>	2 <sup>nd</sup>	3 <sup>rd</sup>	4 <sup>th</sup>	5 <sup>th</sup>
1	Alcoa	•	•	•	•	•	13	Johnson and Johns DC	•	•	•	•	
2	Analog Devices	•	•			•	14	Lowe's Companies	•	•			
3	BHP Billiton Ltd.	•	•			•	15	Millipore CP	•	•			
4	Boeing	•	•	•	•	•	16	Motors Liquidation	•	•	•	•	
5	Choice Hotel INT New	•	•			•	17	NIKE	•	•			
6	Dow Chemical	•		•		•	18	Pepsi	•	•	•		
7	General Electric	•	•	•	•	•	19	Ruby Tuesday	•				
8	Hawallan Elec Inds	•	•			•	20	Southern Company	•	•			
9	Hewlett Packard	•	•	•	•	•	21	Templeton Global Income Fund	•	•			
10	Hindustan Motors	•					22	Wells Fargo & Co New	•	•			
11	Honda Motor	•	•			•	23	Yazhou Coal MNG	•				
12	IBM	•	•	•	•								

**Table 2: List of the stocks and their periods in the experiments**

	Variance-based			LPM-based		
	Max	Min	Mean	Max	Min	Mean
1st	32%	54%	45%	68%	46%	55%
2nd	31%	46%	42%	69%	54%	58%
3rd	42%	50%	40%	58%	50%	60%
4th	42%	56%	49%	58%	44%	51%
5th	43%	61%	50%	57%	41%	50%

**Table 3: The final results of the experiments**

	Variance-based			LPM-based		
	Max	Min	Mean	Max	Min	Mean
1st	19.24	0.02	0.85	190.79	0.00	1.26
2nd	5.03	0.32	1.34	10.46	0.41	2.21
3rd	2.80	0.48	1.20	5.64	0.52	1.86
4th	4.56	0.72	1.85	8.78	0.40	1.87
5th	1.77	0.52	0.98	3.91	0.26	1.01

**Table 4: The amount of change in the initial capital during the associated period**

As can be seen in Tables 3 and 4, the new LPM-based model is decidedly better than the other according to the measures of Max and Mean, while its performance is slightly worse according to the measure of Min. That is nevertheless a satisfactory result.

### 3. SECOND MODULE: WHAT ABOUT THE NEAR FUTURE?

Unlike the past performance assessment of stocks done in the first module, evaluating stocks according to their probable future performance is not a routine task, mainly because of the many difficulties that come with forecasting. In the models based on past performance, it is assumed that the future state of stock markets will fully reflect past trends or behaviours - an assumption that cannot be easily accepted when stock markets experience continuous variations [27]. It is true that there are many works in the field of predicting stock return; but the lack has always been felt of an integrated portfolio selection model that may be capable of considering the future performance of stocks.

This module adds to our model the forecasting mechanism of technical analysis (TA) that contributes to the final results (of the portfolio optimisation model) that are not only on the basis of the stocks' past performance, but also their probable future state.

Technical analysis (TA, also known as charting technique) has been part of financial dealings for many years, and many people believe that it is the main tool for investment analysis [28]. Although there has been much academic opposition to TA, it has been proved that TA for stock prices is powerful, and has considerable popularity among economists and practitioners - because of the balance that TA maintains among human, political, and economic events [28] by considering market psychology. TA can provide a good mechanism to measure the irrational or emotional components that are present in all markets [29], while fundamental analysis only provides a gauge of the supply/demand situations, price/earnings ratios, economic statistics, and so forth; and there is no psychological component involved in such analysis [29]. In a nutshell, this module makes the model more sensitive to market psychology and to short-term data (such as daily movements). After

feeding the necessary data into this module, for each stock there will be a signal to buy, sell, or hold that stock - that is, the outputs of this phase are various signals.

### 3.1. The selected technique

Among the many technical analyses, 'Japanese candlestick' (the oldest [30], going back to 1730) is chosen for use in this module. Since its introduction to the Western world it has become almost universally available in software and online charting packages [31]. But although the supporting literature on candlestick charting and its antiquity are important factors, the main reasons for using Candlestick charting in our model are these:

1. Its use of open, high, low, and close prices, where other technical trading studies have used close price data only.
2. It is more resistant to the criticism of 'data snooping' than are other tests of technical trading rules, as it was developed for an entirely different purpose: forecasting rice markets. [30]

### 3.2. The structure

Generally the approach of artificial intelligence (AI) can be used to equip the model with TA, as shown in Fig.3. Performing TA in financial markets with AI is not new: it has been surveyed by many researchers, with promising results. For example, Lee and Jo [32] develop an expert system of candlestick charting analysis to forecast the best timing of the stock market. Fernandez et al. [33] study the applicability of a simple technical rule on the basis of neural networks. Yao and Tan [34] present some evidence for the applicability of neural network (NN) models in predicting currency exchange rates, where time series data and technical analyses - like the moving average to discover the principles of currency exchange rate movement - are fed to a NN. Leigh et al. [35] show the prospects for applying the modern approach of hybrid methods to assess buying opportunities in the stock market by TA and NN. Lam [36] studies the applicability of NNs - especially the back propagation algorithm - to integrating fundamental and technical analyses for the forecasting of financial performance. Chavarnakul and Enke [28] use an NN to perform an equivolume charting technique. It should be noted that, although in our study the concept of NN will be applied, any other kind of AI - including expert system, genetic algorithm, and fuzzy theory - can be also used.

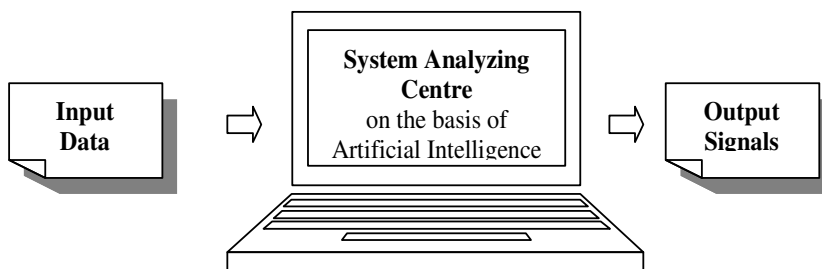


Figure 3: The future performance filtering module of the model

To discover patterns representing the input and output variables, a supervised feed-forward NN with back propagation learning will be applied. Unlike the input and output layer, there are no standard rules available for determining the appropriate number of hidden layers and neurons per layer. But since the model is a multiple regression one, the number of input and output nodes for the network configuration is determined according to dependent and independent variables.

The output of an NN is always the solution to the problem. In this study, based on what has been discussed, the network is intended to inform us about changes in market trends. So the output layer consists of only one node, which determines the stock price trend in the



near future using three states: ascending, descending, and neutral. For the input layer, the problem is not as easy, particularly because so many approaches can be taken to devise this layer. In this regard we focus on the raw input data of the Japanese candlestick, including Open, High, Low and Close prices of the stock over the last three days; while to cover the stock price trend, the close prices of the stock over the last seven days are also included.

Overall this approach produces 15 normalised indices of  $\frac{C_i}{C_1}$   $i = 2,3,4$ ; and  $\frac{O_i}{C_1}$ ,  $\frac{H_i}{C_1}$ ,

$\frac{L_i}{C_1}$  and  $\frac{C_i}{C_1}$   $i = 5,6,7$  while  $O_i$ ,  $H_i$ ,  $L_i$  and  $C_i$  denote respectively open, high, low and

close prices on  $i^{\text{th}}$  day while  $7^{\text{th}}$  day is today (last day),  $6^{\text{th}}$  day is yesterday, and so on. Based on the literature, feeding a particular NN with raw data is common because NNs are able to recognise high-level features such as serial correlations; and the NN is known as an appropriate classification and forecasting tool in business applications [37]. So there will be fifteen input nodes and one output node.

### 3.3. Empirical results

Here, the efficiency of this module is tested on the basis of the 5<sup>th</sup> experiment set. In total we have ten NNs that are trained by the daily data of the corresponding stocks during 2004 and 2005. These training data are used to determine the specifications of the models and the parameters of the forecasting technique, while the testing data of 2006-9 are used for out-of sample evaluation of the forecasting model.

The empirical results are divided into two parts. In the first, the correctness ratio of the emitted signals is evaluated; in the second, the amount of variation in the initial investment if the stocks are traded according to the emitted signals is surveyed. It should be noted that, to calculate the stocks returns, it is assumed that traders enter and exit at the close price of the signal day. Other researchers may make other assumptions; for example, Marshall et al. [30] assume that traders enter at the open price on the day following a signal.

#### 3.3.1. Correctness ratio

To determine the ratio of correct signals, the trend horizon and definition should be decided on. The Japanese candlestick is designed to get short-term price movements, and is most helpful over periods of about ten days [29,38]; so any period less than ten days seems wise. In this study, like Lee and Jo [32], a six-day stock market movement is selected as the unit of evaluation time. That is, each time a bull signal is followed by an actual upturn or a bear signal is followed by an actual downturn within six days, the signal is considered to be a success. In this regard Horton [39] uses a three-day period, and Marshall et al. [30] apply a ten-day period.

In this study, not all the positive and negative returns are considered as actual upturns or downturns. However, for the positive returns there is a lower bound of  $\frac{1}{5} mp$  and for the negative returns there is an upper bound of  $\frac{1}{5} mn$ ; where  $mp$  and  $mn$  are averages of positive and negative daily returns of the stock during the testing years respectively.

Table 5 shows the results of applying the trained networks to the 5<sup>th</sup> set of experiments for five different testing periods, in which column 1 shows the percentages of correct signals during a six-day period, and column 2 presents the total number of buy and sell signals emitted by the network.

Testing data	12/30/2005 to 8/31/2009		12/30/2005 to 12/29/2006		12/29/2006 to 12/31/2007		12/31/2007 to 12/31/2008		12/31/2008 to 8/31/2009	
	1	2	1	2	1	2	1	2	1	2
Alcoa	72.6	620	79.0	138	66.3	163	72.3	191	72.1	136
Analog Devices	74.7	525	73.3	135	77.0	139	72.5	153	76.5	102
BHP Billiton Ltd	76.2	638	71.8	170	79.4	160	77.3	194	72.8	125
Boeing	74.8	571	73.0	141	70.1	137	77.1	175	77.9	122
Choice Hotel INT New	72.8	523	77.6	116	69.2	117	68.6	169	79.6	113
Dow Chemical	69.3	564	77.3	128	69.9	136	65.7	178	65.9	135
General Electric	77.1	545	76.4	123	76.3	131	76.8	185	71.9	139
Hawallan Elec Inds	74.2	472	75.8	99	73.9	119	74.4	156	72.5	120
Hewlett Packard	77.0	379	79.6	93	75.8	95	73.3	131	78.8	80
Honda Motor	76.0	608	75.4	142	72.4	156	79.2	197	75.6	131

**Table 5: The percentages of correct forecastings for one- and six-day periods with total number of buy and sell signals**

Lee and Jo [32] believe that if the hit ratio defined by  $\frac{\text{Number of successes}}{\text{Total number of signals}}$  is above 51%,

the model can be regarded as useful and feasible. The total hit ratio of our new model is 74.2%; and even if the evaluation time unit is set at one day, a surprisingly high hit ratio of 43% is achieved. In fact, the model performs brilliantly well in some cases. For example, for Dow Chemical for the period of 12/30/2005 to 12/29/2006, the one-day period achieves a hit ratio of 55.5%; and Choice Hotel INT New for the period 12/31/2007 to 12/31/2008 achieves the excellent hit ratio of 79.6% using the six-day period.

### 3.3.2. Change in the initial investment

As discussed before, the output of this module is trading signals of 1, -1 and 0 whose correctness ratio was surveyed. In this part the returns of investing in stocks are investigated when the investor fully obeys the emitted signals to buy or sell the associated stocks. The following three strategies will be applied in converting the raw signals (of 1, -1 and 0) to trading commands (of 1 and -1).

1. Getting buy, sell and neutral signals and normalising them according to the following 2-day, 3-day, 4-day, and 5-day or 6-day trends.
2. Getting buy, sell and neutral signals without normalising them.
3. Getting only buy or sell signals without any neutral signal.

In the first strategy, at the first step the neutral trading signals of 0 are converted so that a typical zero is equal to the last non-zero signal. Then a kind of normalisation is performed on the logic that a buy or sell signal is credible for the next six days - i.e., if a buy signal does not become true for the next four days, there is a probability that it will become true on the fifth or sixth day. In this study the signals are normalised independently for 2, 3, 4, 5 and 6 days. For example, for a 3-day normalisation consider a buy signal on day  $t$ . A buy signal on day  $t$  means that the investor should buy the stock at the closing price of day  $t$ , because the next day the price will rise. But if what happens does not match what has been predicted, the buy signal will be applied on the next day as well, even if the predicted signal for day  $t + 1$  is a sell one. This is repeated for the following day too. For a six-day

normalisation, an unmet buy signal may overrule all the sell signals of the model on five days following the first. So this strategy is divided into five sub-strategies of S1.2, S1.3, S1.4, S1.5 and S1.6. Moreover, each of these sub-strategies is run with eight different parameters that affect the acceptance of a change in the stock price as a positive or negative trend.

In the second strategy, as in the first, the neutral trading signals are first converted to non-zero elements, and the resulting signals are used for trading without extra normalisation. And, finally, in the third strategy there is no neutral signal to be converted and no normalisation. Since in these two strategies there is no need for the recognition of trends, we shall have no parameters. So in total we end up with 42 strategies and sub-strategies.

In evaluating the efficacy of the new EF model presented in part 2, Markowitz's famous EF model was used as a benchmark; but in this part the investment strategy of Buy and Hold (B&H) will be applied as a benchmark. For example, for the BHP Billiton Ltd stock between 12/29/2006 and 12/31/2007, a return of 76% is achieved on the basis of B&H because the price changes from 39.75 to 70.04. It is worth noting that the best performance from among the 42 strategies will be compared with the benchmark of B&H. In other words, all three strategies are considered as a whole, and for each stock the best performance among them represents the module. Table 6 shows the results: the columns of 'ns' and 'bh' represent the amount of investment at the end of the associated period for the new and B&H strategies respectively, if the amount of investment at the beginning of the period is taken as 1. For example, for the Honda Motor stock, the new strategy performs much better than the B&H strategy for all periods; while for all stocks over the longest period, the new strategy is superior to B&H.

As mentioned earlier, this module is composed of 42 strategies and sub-strategies; or, to be more precise, it consists of S1.2, S1.3, S1.4, S1.5, S1.6, S2 and S3. So it seems wise to check which of these seven strategies and sub-strategies offer the best performance; and this is shown in Table 7. On the basis of Table 7, of the 50 cases, in 39 cases S1 is superior; in 10 cases it is S3; and in one case it is S2. And among the S.1s, S1.4 represents the best performance in 38% of the cases; S1.3 in 21% of the cases; S1.5 in 18%, S1.6 in 15%; and S1.2 in 8% of the cases.

	12/30/2005 to 8/31/2009		12/30/2005 to 12/29/2006		12/29/2006 to 12/31/2007		12/31/2007 to 12/31/2008		12/31/2008 to 8/31/2009	
	ns	bh	ns	bh	ns	bh	ns	bh	ns	bh
Alcoa	1.02	0.41	1.41	1.01	1.01	1.22	0.67	0.31	1.74	1.07
Analog Devices	1.74	0.79	1.11	0.92	1.23	0.96	1.15	0.60	1.55	1.49
BHP Billiton Ltd.	5.70	1.86	1.57	1.19	1.86	1.76	2.14	0.61	1.61	1.45
Boeing	1.23	0.71	1.24	1.26	0.99	0.98	0.95	0.49	1.57	1.16
Choice Hotel INT New	1.16	0.71	1.13	1.01	0.95	0.79	1.43	0.91	1.39	0.98
Dow Chemical	0.91	0.49	1.09	0.91	1.12	0.99	0.66	0.38	1.52	1.41
General Electric	1.20	0.40	1.03	1.06	1.13	1.00	1.27	0.44	1.12	0.86
Hawallan Elec Inds	0.95	0.67	1.09	1.05	1.04	0.84	1.07	0.97	0.93	0.79
Hewlett Packard	2.36	1.57	1.41	1.44	1.35	1.23	1.21	0.72	1.24	1.24
Honda Motor	2.52	1.08	1.48	1.36	0.99	0.84	1.45	0.64	1.69	1.47

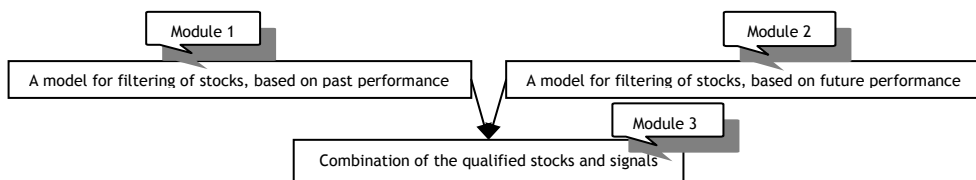
Table 6: The proportional amount of the investment at the end of the period for the new and B&H strategies

	12/30/2005 to 8/31/2009	12/30/2005 to 12/29/2006	12/29/2006 to 12/31/2007	12/31/2007 to 12/31/2008	12/31/2008 to 8/31/2009
<b>Alcoa</b>	S1.4	S1.5	S1.4	S1.5	S1.4
<b>Analog Devices</b>	S1.4	S1.5	S1.4	S1.3	S1.3
<b>BHP Billiton Ltd.</b>	S1.4	S1.3	S3	S1.4	S1.4
<b>Boeing</b>	S1.6	S1.6	S1.3	S1.4	S1.5
<b>Choice Hotel INT New</b>	S3	S1.3	S3	S1.6	S1.4
<b>Dow Chemical</b>	S1.3	S1.4	S1.5	S1.3	S1.2
<b>General Electric</b>	S3	S1.6	S3	S1.2	S3
<b>Hawallan Elec Inds</b>	S1.4	S3	S1.3	S1.4	S1.4
<b>Hewlett Packard</b>	S1.5	S3	S1.5	S1.4	S2
<b>Honda Motor</b>	S1.6	S1.2	S3	S1.6	S3

**Table 7: The strategies and sub-strategies with best performance among the others**

**4. THE MODEL**

In earlier parts of this paper, two independent modules of a particular portfolio selection model are discussed in detail, with promising results from their performances. The proposed modular model is shown in Fig.4. As can be seen in the figure, in addition to the first and second modules, there is a third one in which the outcomes of the first two modules are integrated. As a matter of fact it specifies a utility function of the investor for choosing the most attractive point



**Figure 4: Block diagram of the model**

In most cases, portfolio optimisation is limited to editing the current portfolio - that is, buying, selling, or holding stocks. This is called 'portfolio management'. The application of simple portfolio selection models to portfolio management has several inadequacies, such as a significant amount of repetitive work, or the low sensitivity of the results to new data. For instance, in the case of daily editing, the model should be solved every day and its results compared with those of the day before. If a particular stock is present in yesterday's portfolio but is absent from today's, the stock should be sold completely; while if the reverse is the case, the stock should be bought up to the amount specified by the new portfolio. And finally, if a stock is present in both of the portfolios, its investment share will be set on the basis of the new results. However, in the proposed new modular model, since the input data of TA processors cover intervals of up to seven days, the sensitivity of the model to new data will be much greater than that of the portfolio selection models, with their input data over several months. On the other hand, for exactly the same reason, data is fed into the model less often, and so the amount of repetitive work is considerably reduced.

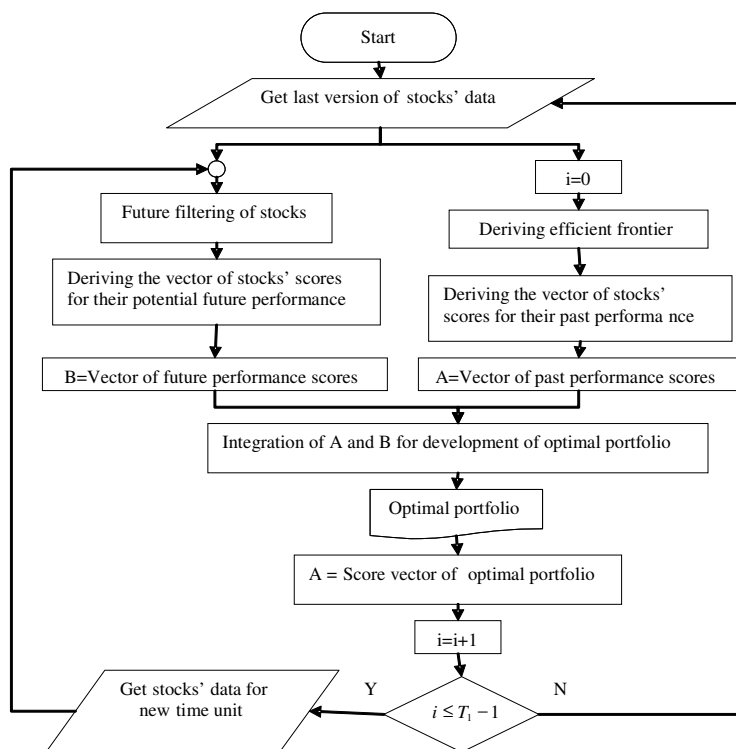


Figure 5: The portfolio management model flowchart

#### 4.1. A typical combination

In the concept of portfolio management, there should be an initial portfolio for editing. Here the first module delivers that initial portfolio. Naturally, if the stocks' data of  $T$  time units are fed to module 1, this module should be applied every  $T_1$  time units where  $T_1 \leq T$ . Fig.5 shows the flowchart of this typical portfolio management model.

According to Fig.5, at the first step the data should be fed into the model. For past performance filtering, the last  $T$  data for each stock - usually the closing price - are used and fed into the model. According to these data, the model yields the efficient frontier. For future performance filtering, the necessary data are fed in  $n$  (number of available stocks) times to get the signals for each stock. According to outputs and the pre-defined rule, an optimal portfolio results, and at this time the first run of the model is finished. The next run of the model will be after a specified time unit. For example, if a day is used as the time unit, the next run of the model will be the next day. After a time unit, new data are added to the data base and the new package of data are fed only into the second module, just as in the previous step, to get the new score vector of future performance. At this stage the old score vector of past performance is replaced with the new optimal portfolio score vector. In the next runs of the model, this vector will be replaced by the score vector of the new optimal portfolios. This procedure continues for  $T_1$  times, each time beginning like the first run, but with new data.

#### 5. CONCLUSION

After a short review of the field of portfolio selection and its major models, some of the shortcomings that seem to reduce their usefulness to stock traders in real markets are

introduced as the focal point of this study. To solve the problem, a modular model is presented. The first two modules are discussed in detail, including their development and validation; while for the third module the study is restricted to describing its concept.

The proposed model is to some extent sensitive to market psychology; it is free of the usual simplistic assumptions (for example, on the distribution of stock returns, the state of the market, and the varied direction of parameters); and its sensitivity to new market changes is good. On the other hand, the experiments show that the first and second modules perform well, and hold promise.

In the proposed model, the first module is a mechanism for deriving EF and covering long-term changes, while the second module emits trading signals and covers short-term changes in the market. The benchmark for the EF module was the Markowitz model, while for the trading signals module the benchmark was B&H.

To prove the efficiency of the model in a more practical manner, the third module should be devised more precisely, running a comprehensive set of experiments for a more solid assessment of its performance. Trying to fit the model into a given market to get better results than previous models would be a good area for research.

## 6. REFERENCES

- [1] Markowitz, H. 1952. Portfolio selection, *Journal of Finance*, 7, 77-91.
- [2] Markowitz, H. 1959. *Portfolio selection: Efficient diversification of investments*, Wiley, New York.
- [3] Sharpe, W.F. 1963. A simplified model for portfolio analysis, *Management Science*, 9, 277-293.
- [4] Deng, X.T., Li, Z.F. & Wang, S.Y. 2005. A minimax portfolio selection strategy with equilibrium, *European Journal of Operational Research*, 166, 278-292.
- [5] Liu, L. & Shenoy, P.P. 1995. Theory of coarse utility, *Journal of Risk and Uncertainty*, 11, 17-49.
- [6] Fernández, A. & Gómez, S. 2007. Portfolio selection using neural networks, *Computers & Operations Research*, 34, 1177-1191.
- [7] Ballesteros, E., Gunther, M., Pla-Santamaria, D. & Stummer, C. 2007. Portfolio selection under strict uncertainty: A multi-criteria methodology and its application to the Frankfurt and Vienna stock exchanges, *European Journal of Operational Research*, 181, 1476-1487.
- [8] Huang, X. 2007. Two new models for portfolio selection with stochastic returns taking fuzzy information, *European Journal of Operational Research*, 180, 396-405.
- [9] Elikyurt, U.C. & Ozekici, S. 2007. Multiperiod portfolio optimization models in stochastic markets using the mean-variance approach, *European Journal of Operational Research*, 179, 186-202.
- [10] Huang, X. 2008. Portfolio selection with a new definition of risk, *European Journal of Operational Research*, 186, 351-357.
- [11] Konno, H. & Yamazaki, H. 1991. Mean-absolute deviation portfolio optimization model and its applications to Tokyo stock market, *Management Science*, 37, 519-531.
- [12] Parra, M.A., Terol, A.B. & Uria, M.V.R. 2001. A fuzzy goal programming approach to portfolio selection, *European Journal of Operational Research*, 133, 287-297.
- [13] Sharpe, W.F. 1964. Capital asset prices: A theory of market equilibrium under conditions of risk, *The Journal of Finance*, 19, 425-442.
- [14] Sharpe, W.F. 1970. *Portfolio theory and capital markets*, McGraw-Hill, New York.
- [15] Lintner, J. 1965. The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets, *Review of Economics and Statistics*, 47, 13-37.
- [16] Mossin, J. 1966. Equilibrium in capital asset market, *Econometrica*, 34, 768-783.
- [17] Ross, S.A. 1976. The arbitrage theory of capital asset pricing, *Journal of Economic Theory*, 3, 341-360.

- [18] **Huberman, G.** 1982. Arbitrage pricing theory: A simple approach, *Journal of Economic Theory*, 28, 183-191.
- [19] **Connor, G.** 1982. A factor pricing theory of capital assets, Working Paper, Northwestern University, Evanston, IL.
- [20] **Ballesteros, E. & Romero, C.** 1996. Portfolio selection: A compromise programming solution, *Journal of Operational Research Society*, 47, 1377-1386.
- [21] **Perold, A.F.** 1984. Large-scale portfolio optimization, *Management Science*, 30, 1143-1160.
- [22] **Jasemi, M., Kimiagari, A.M. & Memariani, A.** 2010. Development of a more applied version of coherency called 'Sensible Coherency' for assessment of financial risk measures, *South African Journal of Industrial Engineering*, 21, 157-165.
- [23] **Unser, M.** 2000. Lower partial moments as measures of perceived risk: An experimental study, *Journal of Economic Psychology*, 21, 253-280.
- [24] **Bawa, V.S.** 1975. Optimal rules for ordering uncertain prospects, *Journal of Financial Economics*, 2, 95-121.
- [25] **Bawa, V.S.** 1978. Safety-first, stochastic dominance, and optimal portfolio choice, *Journal of Financial and Quantitative Analysis*, 13, 255-271.
- [26] **Harlow, W.V. & Rao, R.K.S.** 1989. Asset pricing in a generalized mean-lower partial moment framework: Theory and evidence, *The Journal of Financial and Quantitative Analysis*, 24, 285-311.
- [27] **Tanaka, H. & Guo, P.** 1999. Portfolio selection based on upper and lower exponential possibility distributions, *European Journal of Operational research*, 114, 115-126.
- [28] **Chavarnakul, T. & Enke, D.** 2008. Intelligent technical analysis based equivolume charting for stock trading using neural networks, *Expert Systems with Applications*, 34, 1004-1017.
- [29] **Nison, S.** 1991. *Japanese candlestick charting techniques: A contemporary guide to the ancient investment technique of the Far East*, New York Institute of Finance.
- [30] **Marshall, B.R., Young, M.R. & Rose, L.C.** 2006. Candlestick technical trading strategies: Can they create value for investors? *Journal of Banking & Finance*, 30, 2303-2323.
- [31] **Nison, S.** 2004. Candlestick trading principles, *Technical Analysis of Stocks and Commodities*, November, 22-27.
- [32] **Lee, K.H. & Jo, G.S.** 1999. Expert system for predicting stock market timing using a candlestick chart, *Expert Systems with Applications*, 16, 357-364.
- [33] **Fernandez-Rodríguez, F., Gonzalez-Martel, C. & Sosvilla-Rivero, S.** 2000. On the profitability of technical trading rules based on artificial neural networks: Evidence from the Madrid stock market, *Economics Letters*, 69, 89-94.
- [34] **Yao, J. & Tan, C.L.** 2000. A case study on using neural networks to perform technical forecasting of forex, *Neurocomputing*, 34, 79-98.
- [35] **Leigh, W., Purvis, R. & Ragusa, J.M.** 2002. Forecasting the NYSE composite index with technical analysis, pattern recognizer, neural network, and genetic algorithm: A case study in romantic decision support, *Decision Support Systems*, 32, 361- 377.
- [36] **Lam, M.** 2004. Neural network techniques for financial performance prediction: Integrating fundamental and technical analysis, *Decision Support Systems*, 37, 567-581.
- [37] **Law, R. & Au, N.** 1999. A NN model to forecast Japanese demand for travel to Hong Kong, *Tourism Management*, 20, 89-97.
- [38] **Morris, G.** 1995. *Candlestick charting explained: Timeless techniques for trading stocks and futures*, 2<sup>nd</sup> ed., New York: McGraw-Hill Trade.
- [39] **Horton, M.J.** 2009. Stars, crows, and doji: The use of candlesticks in stock selection, *The Quarterly Review of Economics and Finance*, 49, 283-294.

