

Prediction-Based Portfolio Optimization Model using Neural Networks

Fabio D. Freitas ^{a,*}, Alberto F. De Souza ^b, Ailson R. de Almeida

^a*Secretaria da Receita Federal do Brasil – RFB, Programa de Pós Graduação em Engenharia Elétrica – UFES, Pietrangelo de Biase, 56 sala 308, 29.010-190 – Vitória ES, Brazil.*

^b*Programa de Pós Graduação em Informática – UFES, Av. Fernando Ferrari, s/n, 29075-910 – Vitória, ES, Brazil.*

Abstract

This work presents a new prediction-based portfolio optimization model that can capture short-term investment opportunities. We used neural network predictors to predict stocks' returns and derived a risk measure, based on the prediction errors, that have the same statistical foundation of the mean-variance model. The efficient diversification effects holds thanks to the selection of predictors with low and complementary pairwise error profiles.

We employed a large set of experiments with real data from the Brazilian stock market to examine our portfolio optimization model, which included the evaluation of the Normality of the prediction errors. Our results showed that it is possible to obtain Normal prediction errors with non-Normal time series of stock returns, and that the prediction-based portfolio optimization model took advantage of short term opportunities, outperforming the mean-variance model and beating the market index.

Key words: Neural Networks; Time Series Prediction; Portfolio Optimization.

1. Introduction

Investment selection is a central problem in financial theory and practice and it is primarily concerned with the future performance of investments, mainly their *expected returns*. When investments are exposed to uncertainties, the investment selection framework must include a quantitative measure of the uncertainty of obtaining the expected return, i.e. a quantitative measure of *risk*.

* Corresponding author. Tel/Fax +552732228201.

Email addresses: freitas@computer.org (Fabio D. Freitas), alberto@lcad.inf.ufes.br (Alberto F. De Souza), ailson@ele.ufes.br (Ailson R. de Almeida).

The mean-variance model, proposed by Harry Markowitz [1], is a landmark in Modern Portfolio Theory (MPT). In this model, the risk of investment in a portfolio of stocks is minimized through optimal selection of stocks with low joint risk, which provides a mechanism of loss compensation known as *Efficient Diversification*. The portfolio optimization process consists of finding, in a large collection of stocks, the participation of each stock (i.e. the percentage in the portfolio value) that minimizes the portfolio's risk at a desired portfolio return, or, in the dual problem, maximizes the portfolio's return at a given risk. Varying the desired portfolio return (or risk) among all possible values for optimal portfolios outlines a *locus* in the risk-return space named *Efficient Frontier*. The portfolios that lie in Efficient Frontier are named *Efficient Portfolios*, and have the singular property that there are no other portfolios in the opportunity set that exhibit lower risk for the same return, or greater return for the same risk [2,3]. The model has the fundamental assumption that the time series of returns of each stock follows a Normal distribution and uses its mean as a prediction of the stock's future return, its variance as a measure of the stock's risk, and the covariance of each pair of time series as a measure of joint risk of each pair of stocks.

After the Markowitz's mean-variance model, many other models that use its fundamental assumption appeared [4, pp. 219–252]. In all these models, known today as classic models, the portfolio's expected return is given by the linear combination of the participations of the stocks in the portfolio and its expected returns (the mean returns). The portfolio risk measure varies, but it is often based on the linear combination of the participations and the moments about the mean of the time series of returns of its stocks. Despite the wide adoption of the classical models of portfolio selection, their fundamental assumption has been threatened by real world data in many ways. The distributions of the series of returns often departs from Normality, exhibiting kurtosis and skewness [5,6], which makes the variance (or standard deviation) of the returns an inappropriate measure of stocks' risk [3, pp. 156]. In addition, the use of mean returns as prediction of stocks' future returns impose a low pass filtering effect on the dynamic behavior of the stock markets, leading to imprecise estimates of short-term future returns, which is detrimental to the performance of the models on short-term investments.

The predictability of stock markets is still an open question in finance theory. The *Efficient Market Hypothesis* (EMH), the theoretical framework that guides the discussion around this question, has been under empirical testing and reviewing for the past several decades [7,8,9,10]. Market efficiency implies a *random walk* model for the prices of stocks, but pricing irregularities and predictable patterns like serial correlations, calendar effects, and even sports results effects do appear [10,11].

The forecasting of time series was traditionally tackled by the linear methods of *time series analysis* [12]. However, in the last two decades, many machine learning methods for time series prediction appeared, ranging from neural network models to support vector machines and fuzzy sets [13,14,15,16]. Among these methods, the non-linear mapping capabilities and the non-assisted estimation of the structural model's parameters of the neural networks are advantageous for its application in the prediction of stock's future returns, making then suitable for large scale applications [17,18]. The development of accurate neural predictors is still subject of many research efforts. Recently, Chen et al. [19] proposed a flexible neural tree ensemble to predict the NASDAQ-100 and S&P CNX NIFTY stock indexes, while Liu et al. [20] presented a reinforcement learning-based scheme for simultaneously determining the input dimension and time delay of a neural predictor that was employed to predict daily stock prices of General Motors Corp. These predictors were not used for building automated investment strategies, though. Nevertheless, in recent years, attention has been devoted to use neural networks for implementing automated investment strategies called *trading systems*. Neural-based trading systems typically employ neural network predictors to predict stock's future prices, returns or other related measures, and use these to generate signals to buy, sell or hold assets.

Lee et al. [21] proposed a divide and conquer approach that uses multiple specialized Q-Learning neural-based agents to build a trading system. The investment problem was divided into timing and pricing prob-

lems, and four separate agents were used for generating buy and sell signals for the respective timing and pricing problems. The system was used for trading on the Korean KOSPI index and presented superior results than other methods examined by the authors.

Thawornwong and Enke [22] used a data mining technique to select economic and financial variables with predictive power and used the selected variables as input of a neural predictor employed for predicting the direction (sign) of future return movements. The predictions were used to implement a trading strategy for deciding either to invest on the S&P 500 index portfolio or on the T-Bill (risk-free) during one month. The results showed that the monthly returns obtained with this adaptive strategy were always higher than those of the (non-adaptive) compared methods.

Moody and Saffell [23] developed a trading system based on the direct reinforcement method that does not have to learn a value function. In their method, the trading strategies are learned directly from data, thus without the need of forecasting intermediate values. They used their system to trade on currency, and for S&P 500/T-Bill asset allocation, achieving better results than compared systems.

Pantazopoulos et al. [24] presented a neurofuzzy volatility predictor that used partitioned training subsets with data vectors that preceded an increase, sustain or decrease in volatility, to derive a trading strategy for options on S&P 500 index during 10,000 days. A simulated portfolio based on the S&P 500 index and initiated with \$1,000 obtained a final balance of \$3,862, while a portfolio based on the author's trading system obtained a final balance of \$625,120, with the same initial value and in the same trading period, although with a higher volatility.

Hellstrom [25] developed a rank measure based on the relative returns of a large number of securities. Ranks built with this ranking measure were predicted with a linear model, and the sign of daily threshold-selected 1-day predictions achieved an accuracy of 63% (or a 63% *Hit Rate* – H_R). The author used the predicted ranks to build a trading strategy for the Swedish stock market that significantly outperformed the market index. In a subsequent work [26], the author optimized a set of *technical indicators* through a sliding time window modeling that generates different parameters for the optimized trading rules for each time window. Solutions covering too few examples were rejected and the generalization of the optimized trading rules was enhanced. His results achieved H_R between 59% and 64%, while the benchmark methods performed below 53%.

The trading systems approach has produced remarkable results, but these applications are typically designed for maximizing the investment's return by trading a single asset. Therefore, the higher returns are expected in periods of high volatility of the asset, which means high risk associated with the asset. But, although rational investors are profit-seekers, they are also averse to losses [27,28, pp. 595]. This makes the adoption of a risk controlling mechanism in the trading strategy an imperative issue.

The efficient diversification framework, allied with predictors of future returns and proper risk measures, is a suitable alternative for addressing the single asset and risk controlling issues associated with trading systems. As stated previously, neural networks are advantageous for large scale applications such as portfolio selection. In an earlier work [29], we investigated the Normality of the errors of weekly stock returns predictions produced by a new autoregressive neural network predictor we have proposed (the *autoregressive moving reference neural network* predictor – AR-MRNN). We found more evidence of Normality on these errors of prediction than on the series of returns, and explored this in a subsequent work [30], where we proposed a portfolio selection model that uses predicted returns as expected returns and the variance of the errors of prediction as risk measure.

In this work, we present an extensive evaluation of our prediction-based portfolio optimization model, comparing its short-term investment performance with that of the mean-variance model. This evaluation involved a large set of experiments with real data from the Brazilian stock market. Our experimental results showed that the prediction-based portfolio optimization model outperforms the mean-variance model and

the Brazilian IBOVESPA market index, achieving higher returns with lower risks, while using the same stocks in the same periods of time.

The remainder of the paper is organized as follows. After this introduction, Section 2 presents the AR-MRNN predictor [29] and Section 3 presents our prediction-based portfolio optimization model. Our experimental methods and results are presented in Section 4. We conclude with a brief discussion in Section 5 and our final conclusions in Section 6.

2. Autoregressive moving reference neural network (AR-MRNN) predictors

An investment project is typically planned and executed over a time frame named *investment horizon*. Investments are compared using a relative performance measure named *return on investment*, or simply *return*, that quantifies the wealth variation over the investment horizon. The one-period stock return in time t is defined as the difference between the price of the stock at time t and the price at time $t - 1$, divided by the price at time $t - 1$, as shown in Eq.1.

$$r_t = \frac{P_t - P_{t-1}}{P_{t-1}} \quad t \geq 1 \quad (1)$$

where r_t is the one-period stock return at time t , and P_t and P_{t-1} are the stock prices at times t and $t - 1$, respectively.

The series of N past returns of a stock, r' , is defined as:

$$r' = (r_1, r_2, \dots, r_N) \quad (2)$$

The one-period prediction of the future return of a stock can be defined as the process of using r' for obtaining an estimate of r_{t+1} . The classical neural network time series predictor is the *autoregressive neural network* (ARNN) predictor [13], $\mathcal{S}_{\mathcal{R}}$, with p inputs – the present value and the $p - 1$ past values of the series, whose output is an estimate of the value of the next time period, as shown in Eq.3:

$$(r_{t-(p-1)}, \dots, r_{t-1}, r_t) \rightarrow \mathcal{S}_{\mathcal{R}} \rightarrow \hat{r}_{t+1} \quad (3)$$

The number of inputs, p , is the *regression order*, and some methods for obtaining it are shown in [12]. After being trained, the neural network predictor implements a non-linear multiple regression model of the time series of returns [31, pp. 635-660].

We proposed a new autoregressive neural network model, named *autoregressive moving reference neural network* (AR-MRNN) predictor, to try and mimic a way one typically inspects a time series graph to guess its future value [29]. In this task, one tends to concentrate the visual attention in the last points of the graph, creating an imaginary frame that delimits this region of the graph and offers an image with, hopefully, sufficient visual information for extrapolating the next value of the series. In this prediction scheme, one uses some point inside this region as a reference from which the future value of the series is estimated. We mimic this in the AR-MRNN predictor by subtracting one of the past return values – the reference – from the returns that would be presented to the neural network inputs, and by presenting these differences as input instead.

The AR-MRNN(p, k) predictor has regression order p , the return at time $t - (p - 1) - k$ as reference, and inputs and outputs shown in Eq.4:

$$(r_{t-(p-1)} - z, \dots, r_t - z) \rightarrow \mathcal{S}_{\mathcal{R}} \rightarrow \widehat{r_{t+1} - z} \quad (4)$$

where z is the reference value given by:

$$z = r_{t-(p-1)-k} \quad (5)$$

After training, \hat{r}_{t+1} is obtained from the prediction $\widehat{r_{t+1} - z}$ using:

$$\hat{r}_{t+1} = \widehat{r_{t+1} - z} + z \quad (6)$$

The values encoded into the AR-MRNN weights are usually smaller than those encoded in the weights of a standard ARNN, which reduces the possibility of saturation at the neurons outputs. This increases the dynamic range of the network and its ability to represent the series of returns. Also, preprocessing demands like normalization and detrending are alleviated, and smaller weight values are produced. As a result, the network can be regularized, which enhances its generalization capabilities [31].

3. Prediction-based portfolio optimization model

Since the proposition of the mean-variance portfolio optimization model by Harry Markowitz, computational feasibility, model simplifications and the development of risk measures have received considerable research attention [32,33,34,35,36]. The prediction of future returns in the context of portfolio selection have received little attention, however, and most models use the same prediction method employed in the mean-variance model, i.e., the mean of past returns. Nevertheless, the mean returns are expected to be verified only in the long term; therefore, they are inadequate predictions of future returns in short-term investment strategies, such as active portfolio management. The use of better prediction methods for obtaining estimates of future returns, and associated risk measures, can be used to produce predictive portfolio selection models that are suitable for short-term investment strategies.

This section describes a prediction-based portfolio optimization model that, instead of using the mean returns as the mean-variance model, uses predicted returns as expected returns, and instead of using the variance of the returns, uses the variance of the errors of prediction as risk measure.

3.1. Expected return and risk of a stock

Let the return of a stock and its predicted return be related by:

$$r_t = \hat{r}_t + \varepsilon_t \quad (7)$$

where, r_t is the stock return at time t , \hat{r}_t is the predicted return for time t , obtained at time $t - 1$, and ε_t is the prediction error at time t . Rearranging terms, we can have the prediction error defined as:

$$\varepsilon_t = r_t - \hat{r}_t \quad (8)$$

The time series of n errors of prediction is then given by:

$$\varepsilon' = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n) \quad (9)$$

For a non-biased predictor, the series of errors of prediction must be statistically independent and identically distributed (iid), with mean and variance given by:

$$\mu_\varepsilon = \bar{\varepsilon} = 0 \quad (10)$$

$$\hat{v} = \sigma_\varepsilon^2 = \frac{1}{n-1} \sum_{t=1}^n \varepsilon_t^2 \quad (11)$$

The variance of the errors of prediction (Eq.11) reflects the uncertainty about the realization of the predicted return and is used in the model as measure of the individual risk of each stock (the higher the variance, the higher the risk).

3.2. Expected return and risk of a portfolio

A portfolio is a collection of M stocks and M weights, or participations. Each participation, X_i , $i = 1, \dots, M$, $0 \leq X_i \leq 1$, and $\sum X_i = 1$, represents the fraction of the portfolio value invested in the stock i . The predicted return of the portfolio, or portfolio expected return, \hat{r}_p , is the linear combination of the participation and predicted return of each stock in the portfolio:

$$\hat{r}_p = \sum_{i=1}^M X_i \hat{r}_i \quad (12)$$

Assuming that the time series of errors of prediction follows a Normal distribution, we model the portfolio risk as the variance of the joint Normal distribution of the linear combination of the participations and prediction errors of the stocks of the portfolio:

$$\hat{V} = \hat{\sigma}_p^2 = \sum_{i=1}^M \sum_{j=1}^M X_i X_j \gamma_{\epsilon_{ij}} \quad (13)$$

where \hat{V} is the total portfolio risk, that is equal to the variance of the linear combination of the participations and prediction errors of each stock in the portfolio, $\hat{\sigma}_p^2$; M is the number of stocks in the portfolio; X_i and X_j are the participations of stocks i and j in the portfolio, respectively; and $\gamma_{\epsilon_{ij}}$ is the interactive prediction risk of stocks i and j , which is the covariance of the errors of prediction of the stocks i and j (see also Eq.10 and Eq.11), and is given by:

$$\hat{\gamma}_{ij} = \gamma_{\epsilon_{ij}} = \frac{1}{n-1} \sum_{t=1}^n \epsilon_{it} \epsilon_{jt} \quad (14)$$

Eq.13 can be rewritten as:

$$\hat{V} = \hat{\sigma}_p^2 = \sum_{i=1}^M X_i^2 \sigma_{\epsilon_i}^2 + \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M X_i X_j \gamma_{\epsilon_{ij}} \quad (15)$$

where the first sum represents the contribution of the risk associated with each stock to the portfolio risk (sum of the square of the participation X_i times the variance of the errors of prediction $\sigma_{\epsilon_i}^2$ of each stock i), and the second group of sums represents the contribution of the interactive prediction risk of each pair of stocks i and j (sum of the participations X_i and X_j times the covariance $\gamma_{\epsilon_{ij}}$ of the stocks i and j).

3.3. Portfolio optimization model

With the fundamental measures defined, the prediction-based portfolio optimization model can be formulated as:

Minimize

$$\hat{V} = \sum_{i=1}^M X_i^2 \hat{v}_i + \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M X_i X_j \hat{\gamma}_{ij} \quad (16)$$

Subject to

$$\sum_{i=1}^M X_i \hat{r}_i = R_d \quad (17)$$

$$\sum_{i=1}^M X_i = 1 \quad (18)$$

$$X_i \geq 0, i = 1, \dots, M \quad (19)$$

Eq.16 is the minimized objective function, the prediction-based portfolio risk; Eq.17 is the desired return constraint that guarantees the desired portfolio return R_d ; Eq.18 guarantees total resource allocation and Eq.19 restrict the model for purchase trades only.

As in the mean-variance model, a set of minimum risk portfolios can be obtained by solving the minimization problem above for a range of desired portfolio return values. This set is named *Efficient Frontier*, these portfolios are named *Efficient Portfolios*, and the investment strategy based on this framework is named *Efficient Diversification*. Each efficient portfolio has the particular property that, according to the model, there is no other portfolio in the opportunity set that has a lower risk for a given return, or, in the dual problem, a higher return for a given risk.

3.4. Prediction-based portfolio optimization model versus mean-variance model

The prediction-based portfolio optimization model differs from the mean-variance model because: (i) in the prediction-based portfolio optimization model, the expected return of each stock is its predicted return, instead of the mean of its time series of returns, as is the case in the mean-variance model; (ii) in the prediction-based portfolio optimization model, the individual risk of each stock and the interactive risk between each pair of stocks are obtained from the variance and covariance of the time series of the errors of prediction (Eqs.11 and 14), instead of from the variances and covariances of the time series of returns. (iii) although both models are based on the Normal framework, in the prediction-based portfolio optimization model the Normal variable of interest is the error of prediction of the return of the stocks, while, in the mean-variance model, the Normal variable of interest is the return of the stocks.

Correctly predicting the time series of stock returns is recognized as a difficult task [37]. Stock return predictors do not have good performances, typically exhibiting high error levels. However, predictors can be combined in such a way that allows the exploitation of the complementarities of their errors, leading to good combined predictions, as is the case of the portfolio return of our model. The prediction-based portfolio optimization model is based on the assumptions that the mean of the errors of prediction is zero (Eq.10) and that the errors of prediction are Normal. These assumptions are supported by the experimental results shown in Section 4.2.

4. Experiments

This section presents the experiments we have used to evaluate our neural predictors, the Normality of the errors of prediction, and the performance of the prediction-based portfolio optimization model on artificial data, and on real data under several risk levels.

4.1. Data

From the 82 stocks that participated in the IBOVESPA index between December 2004 and September 2007, we selected a subset of 52 stocks with long enough time series for training the neural networks and cal-

culating the necessary parameters of the portfolio optimization models¹. For each one of these 52 stocks, we computed the 413 weekly returns of the period between 27-Oct-1999 and 19-Sep-2007 using closing prices sampled at Wednesdays to avoid the beginning-of-the-week and end-of-the-week effects [10,4, pp. 404]. In all cases of missing data, we used the last daily closing price available.

4.2. Prediction of returns

The topological and training parameters of the neural network predictors evaluated were determined empirically in previous works [29,30] and are as follows.

We used AR-MRNN(4,1) predictors (Section 2) implemented with a fully connected feedforward neural network with 2 hidden layers, sigmoidal activation function, and a 4:16:4:1 topology (4 input neurons, 16 neurons in the first hidden layer, 4 neurons in the second hidden layer, and 1 output neuron)².

To train and test (use for predictions) the neural networks involved in the experiments, we used a sliding window of 168 of the 413 weekly returns available. This allowed 246 predictions, i.e. the 413 returns (from our full data set) minus 168 (sliding window, including the 4 returns necessary as the first 4 inputs of the AR-MRNN(4,1)), plus 1 (the prediction of the initial window). Therefore, we had 12,792 training sessions (246 train and test cycles \times 52 stocks = 12,792). Each training session was conducted during 200,000 epochs using the back-propagation algorithm [31] with learning rate of 0.009 and inertia of 0.95.

The sliding window contains the training set (163 input-output pairs) and the testing set (1 input-output pair). To reduce overfitting, we employed a technique, described below, that required dividing the training set into two segments: a training segment with the first 156 input-output pairs, and a validating segment with the last 7 input-output pairs. The training segment was used for updating the neural networks' weights, while the validating segment was used to try and select, at each 1,000 training epochs (training blocks), the weights that provided the smaller root mean square prediction error (RMSE – see Eq.21 in Section 4.2.1).

The traditional use of training and validating segments for controlling overfitting [31] is challenged by non-stationary time series prediction, exactly because its structural regime changes over time, and more recent data are needed for both the training segment (for encoding more recent information into the network's weights) and the validating segment (for controlling overfitting). There are no standard methods to address this trade-off, but some alternatives do apply, like shifting the validation segment to other locations in time [24]. The overfitting controlling procedure we have used is inspired in the *Non-linear Cross Validation* (NCV) proposed in [39].

In our procedure, we first obtain a set of network weights at the end of a training block. Next, before starting the next training block, we perturb the weights by inserting new information through additional training with the validating segment during a small number of epochs (10 epochs), under the assumption that this perturbation will help finding a better set of weights for predicting the next time series value. Then, we use the network to predict the returns in the validation segment and examine its RMSE prediction error with the current set of weights. If it is the best performance so far, we save the current set of weights. We then proceed with the following training block. This procedure is repeated for the 200,000 epochs, at the end of which, we have the best set of weights that is used to predict the test set of the current sliding window. The sizes of the training and validating segments were obtained using the heuristic presented in [31, pp. 217].

¹ Since the end of the 1990's, the Brazilian stock market has been experiencing a growing and maturing process, with a steady increase in the number of initial public offerings (IPOs). Many of these new stocks entered the IBOVESPA index in the beginning of 2000's, and did not have long enough time series to be included in our analysis.

² Smaller networks typically underfitted the training data. This network size provides good generalization [38], but demands a smoother training, with small learning rate and large number of epochs to avoid saturation of the neurons' outputs.

The training and testing procedure described above was repeated for all 246 predictions by advancing the sliding window of 168 weeks, one week at a time. We performed the resulting 12,792 training sessions in the 64 ATHLON XP 1800 nodes Enterprise cluster of the High Performance Computing Lab of the *Departamento de Informática of Universidade Federal do Espírito Santo* (<http://www.lcad.inf.ufes.br>).

4.2.1. Evaluation metrics

We used the *Mean Error*, *Root Mean Square Error*, *Mean Absolute Percentage Error*, and *Hit Rates* metrics to evaluate our predictor's performance.

The Mean Error (ME) is the average difference between the realized and predicted returns, defined as:

$$ME = \frac{1}{n} \sum_{t=1}^n r_t - \hat{r}_t \quad (20)$$

where n is the length of the time series, and r_t and \hat{r}_t are the realized and predicted returns at time t , respectively. The ME is used to evaluate the assumption of Eq.10 and the Normality of the errors of prediction.

The Root Mean Square Error (RMSE) is a standard metric for comparing the differences between two time series, and it is defined as:

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (r_t - \hat{r}_t)^2} \quad (21)$$

The RMSE error may be interpreted as the standard deviation of the errors of prediction (see Eq.8) with respect to a zero mean, and it shows the distance of these errors from the ideal situation of zero mean error (see Eq.10). The RMSE has low *outlier* protection, good sensitivity for small changes in data, and does not display data asymmetry [40].

The Mean Absolute Percentage Error (MAPE) is defined as:

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{r_t - \hat{r}_t}{r_t} \right| \quad (22)$$

The MAPE error is a unit-free measure, has good sensitivity for small changes in data, does not display data asymmetry and has very low *outlier* protection [40]. It is important to note that, when r_t is very close to zero and \hat{r}_t is not, the MAPE error can average to a very large value.

The Hit Rates H_R , H_{R+} and H_{R-} measure the percentage of predictions whose signals of \hat{r} and r coincide: H_R is the percentage of predictions were both have the same signal and are different then zero, H_{R+} the percentage were both are positive, and H_{R-} were both are negative. These metrics are suitable for evaluating predictors as trading signals generators [37], and can be written as:

$$H_R = \frac{Count_{t=1}^n(r_t \hat{r}_t > 0)}{Count_{t=1}^n(r_t \hat{r}_t \neq 0)} \quad (23)$$

$$H_{R+} = \frac{Count_{t=1}^n(r_t > 0 \text{ AND } \hat{r}_t > 0)}{Count_{t=1}^n(\hat{r}_t > 0)} \quad (24)$$

$$H_{R-} = \frac{Count_{t=1}^n(r_t < 0 \text{ AND } \hat{r}_t < 0)}{Count_{t=1}^n(\hat{r}_t < 0)} \quad (25)$$

where, $Count(\cdot)$ is the counting of occurrences of its argument.

Table 1

Summary of the results obtained with the 52 predictors.

Error	mean	σ^2	σ
ME	-0.000578	0.000011	0.003241
RMSE	0.072500	0.000254	0.015950
MAPE	9.398462	163.102558	12.771161
H_R	0.512926	0.001255	0.035419
H_{R+}	0.610100	0.003013	0.054890
H_{R-}	0.383322	0.003304	0.057476

4.2.2. Prediction performance

Table 1 presents the results of the experiments we employed to assess the prediction performance of the 52 AR-MRNN(4,1) predictors (one for each stock). It shows the mean, variance (σ^2), and standard deviation (σ) of the results obtained with the 246 predictions of each predictor, measured according to each metric of the previous section.

As Table 1 shows, the average of the ME error was close to zero with a low standard deviation, thus validating the assumption of Eq.10. The RMSE, MAPE and H_R achieved typical levels for this application [20,37]. The H_{R+} of 61% showed that predictors achieved a performance 11% above pure chance (50%) in predicting the positive returns of the market, while the prediction of the negative returns obtained a H_{R-} near 38%. Apart from H_{R+} and ME error, the performance of the predictors was modest; however, our prediction-based portfolio optimization model was able to capture these two good collective results to achieve good portfolio performance, as we show in Section 4.4.

4.2.3. Normality of the errors of prediction and returns

We examined the Normality of the errors of prediction and the Normality of the returns of the precise time series used for portfolio optimization. For that, we measured the percentage of these series that had Normality accepted (i.e. not rejected) according the *Chi-square goodness-of-fit* test [41] at significance level of 1%. We called this measure *Normality Index*. Not all 52 stocks participated in all portfolios optimizations of our trading experiments (Section 4.4), but only those that were in the IBOVESPA index at the time of each portfolio optimization (the number of stocks used in each week varied between 40 and 49 of the 52). We did that to allow better performance comparisons of the portfolio optimization models under analysis with the IBOVESPA index (see Section 4.4).

We performed trading experiments during the last 142 weeks (from 5-Jan-2005 to 19-Sep-2007) of the 413 weeks available in our data set. To compute the parameters of the portfolio optimization models (the variances and covariances of the prediction errors for the prediction-based portfolio optimization model, and the mean, variances and covariances of returns for the mean-variance model) we used a sliding window of 104 weeks preceding each of the 142 weeks mentioned above.

The Normality Indexes computed at each one of the 142 weeks are shown in Fig.1. As Fig.1 shows, the Normality Indexes of the series of errors of prediction was slightly higher and varied less than the Normality Indexes of the series of returns.

Table 2 summarizes the 142 Normality Indexes of each series, showing the mean, variance (σ^2), and standard deviation (σ) of the Normality Indexes. We performed the *t-test* [41] of the means of the Normality Indexes at significance level of 5%. The result indicated that the mean of the Normality Index of the series of errors of prediction is larger than that of the series of returns (the *t-statistic* value of 3.41 was larger than

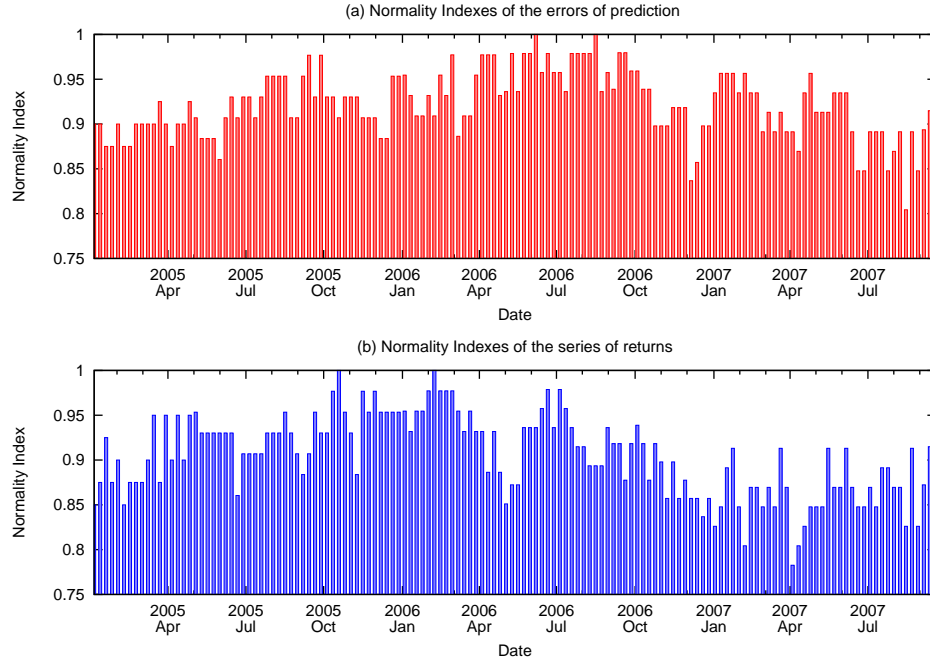


Fig. 1. Normality Indexes for (a) the series of errors of prediction, and (b) the series of returns, at each trading week. The series of errors of prediction are slightly more Normal than the series of returns.

Table 2

Summary of the Normality Indexes for the series of errors of prediction and for the series of returns.

	mean	σ^2	σ
Normality index for the series of errors of prediction	0.921621	0.001336	0.036561
Normality index for the series of returns	0.905141	0.001980	0.044503

the critical value $t_{0.05}$ of 1.645).

4.3. Trading with artificial data

We present a trading experiment with artificial data to show how the prediction-based portfolio optimization model takes advantage of predictive opportunities not visible to the mean-variance model. For that, we designed a trading experiment using the two simple artificial stocks of Fig.2(a): a stock R1 with constant weekly returns of 0.005, and a stock S1 with zero-mean sinusoidal weekly returns between -0.01 and 0.01 . The experiment consisted of implementing the trading strategy of the next section for investing in portfolios that maximize the expected return. To evaluate the portfolios' performance we used the metrics of Section 4.3.2 during 142 weeks.

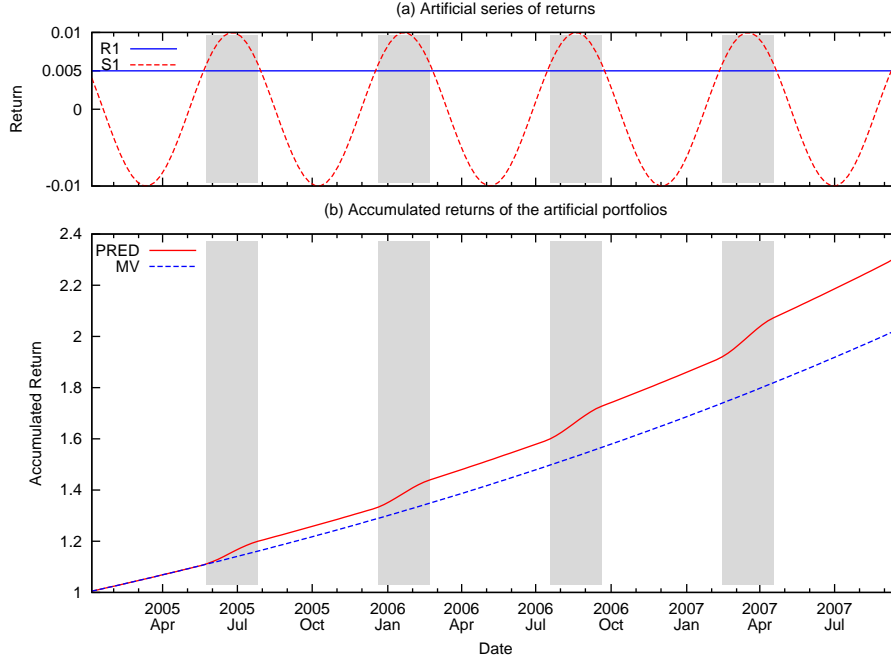


Fig. 2. (a) The artificial series of returns of R1 and S1; (b) Accumulated Returns obtained with the artificial stocks R1 and S2. The prediction-based portfolio optimization model (PRED) switched positions between R1 and S1, using this later when it predicted a superior return (shaded regions), while mean-variance model (MV) used only stock R1, because the mean return of stock S1 is zero.

4.3.1. Trading strategy

The trading strategy employed [3, pp. 139] consisted in: (i) to obtain an Efficient Frontier from the available data at time t , (ii) to select the desired efficient portfolio, (iii) to invest all the available wealth in the stocks according to the weights of the portfolio, and (iv) to sell the entire portfolio at time $t + 1$ and to reinvest the whole wealth obtained in a new portfolio obtained in the same way.

As usual, we implemented the above trading strategy for evaluating our models using the following underlying assumptions: (i) that the stocks are perfectly divisible; (ii) that it is possible to buy and sell any selected portfolio; (iii) that there is no friction (transactions costs, taxation, commissions, etc.); and (iv) that it is possible to buy and sell stocks at closing prices at any time t .

4.3.2. Portfolio evaluation metrics

The portfolio performance evaluation was based on the return and risk measures defined in Section 2 and Section 3, and on the *Accumulated Return*, *Portfolio Change Measure*, and *Turnover Index* metrics described below.

The Accumulated Return is defined as:

$$\text{Accumulated Return} = \prod_{i=0}^t (1 + r_i) \quad (26)$$

where r_i is the arithmetic return at time i , as defined in Eq.1. This is a standard performance measure for comparing investments and relates the wealth at time t , W_t , with the initial wealth, W_0 , as $W_t = W_0 \times \text{Accumulated Return}_t$. All the trading experiments in this work used an initial wealth $W_0 = 1$.

Table 3
Summary of the artificial trading during the 142 weeks

<i>Prediction-based portfolios selected at maximum \hat{r}_p</i>				<i>Mean-variance portfolios selected at maximum \bar{r}_p</i>			
Accumulated Return	2.3210			Accumulated Return	2.0303		
	<i>mean</i>	σ^2	σ		<i>mean</i>	σ^2	σ
weekly returns	0.00594865	$2.82215e-06$	0.00167993	weekly returns	0.005	$3.79465e-16$	$1.94799e-08$
<i>TI</i>	0.12676	0.240729	0.490642	<i>TI</i>	0	0	0
<i>PCM</i>	$5.75046e-05$			<i>PCM</i>	0		
	<i>mean</i>	<i>min</i>	<i>max</i>		<i>mean</i>	<i>min</i>	<i>max</i>
number of stocks	1	1	1	number of stocks	1	1	1

The Portfolio Change Measure (*PCM*) [42,43] is the covariance in time between the returns of the stocks and the variation of the participations on the same stocks, and is defined as:

$$PCM = \frac{1}{T} \sum_{t=1}^T \sum_{i=1}^M r_{it} (X_{it} - X_{it-1}) \quad (27)$$

where T is the length of the time period under analysis (in our case, number of weeks), M is the number of stocks in the portfolio, r_{it} is the return of the stock i at time t , X_{it} and X_{it-1} are the participations of stock i in the portfolios of times t and $t-1$, respectively. A stock contributes positively to the portfolio's *PCM* when it has a positive return in time t (see Eq.1) and the participation of this stock increases between times $t-1$ and t , or when it has a negative return in time t and the participation of this stock decreases between times $t-1$ and t .

We propose the Turnover Index (*TI*) that we use to measure the percentage of the invested wealth that is exposed to frictions at time t . The *TI* is defined as:

$$TI_t = \sum_{i=1}^{M^{t-1,t}} |X_{it} - X_{it-1}| \quad (28)$$

where $M^{t-1,t}$ is the number of the stocks belonging to both the portfolios of times t and $t-1$, X_{it} and X_{it-1} are the participations of stock i in the portfolios of times t and $t-1$, respectively, and $|\cdot|$ is the absolute value of its argument. The TI_t is bounded in the $[0,2]$ interval, and its bounds reflect holding the same portfolio between times t and $t-1$ (a 0% exposition to frictions), and selling the entire portfolio at time $t-1$ and buying a completely different one at time t (a 200% exposition to frictions), respectively.

4.3.3. Experimental results

Fig.2(b) shows the performance of the prediction-based portfolio optimization model and the mean-variance model, presenting their Accumulated Returns during the trading period defined in Section 4.2.3, i.e. the last 142 weeks (from 5-Jan-2005 to 19-Sep-2007) of the 413 weeks available in our data set. The neural predictors learned to predict both series of Fig.2(a) with zero error, allowing the prediction-based portfolio optimization model (the PRED curve in Fig.2(b)) to accumulate higher returns by alternating participations between R1 and S1, using that with the higher predicted return; while the mean-variance model (MV in Fig.2(b)) used only the R1 stock, because this stock always had a mean return superior than S1.

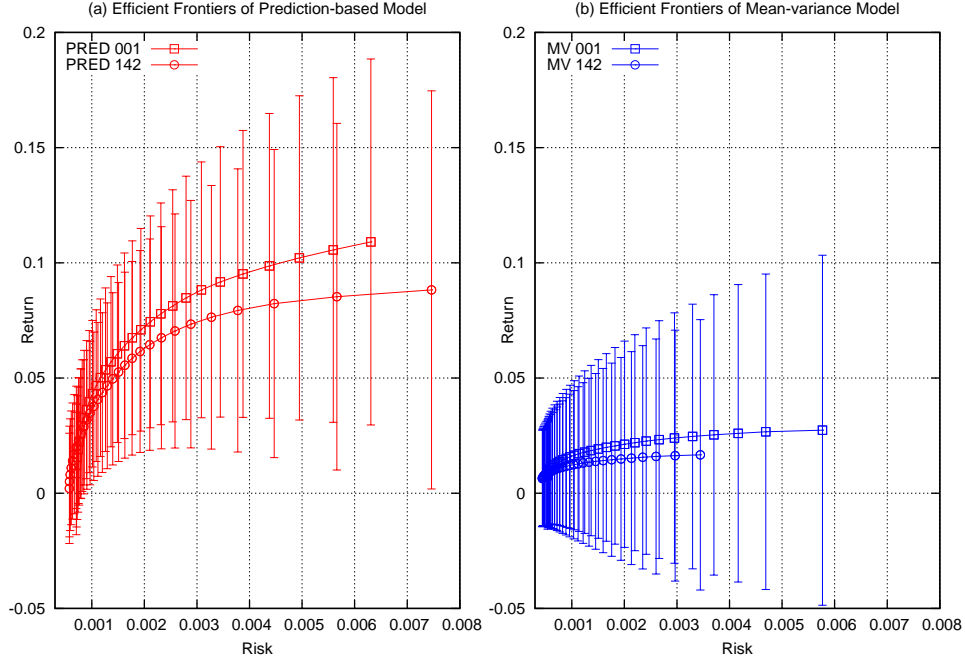


Fig. 3. Two Efficient Frontiers of the (a) prediction-based portfolio optimization model (PRED), and (b) mean-variance model (MV). The prediction-based portfolio optimization model obtained higher returns for the same risk levels, and have fewer negative $\hat{r}_p - \hat{\sigma}_p$ values than the mean-variance model's $\hat{r}_p - \hat{\sigma}_p$ values (vertical bars).

Table 3 summarizes the results obtained with all metrics employed in the evaluation. As Table 3 shows, the prediction-based portfolio optimization model presented a final Accumulated Return of 2.3210, a value 14.31% higher than the 2.0303 value that was achieved by the mean-variance model. The prediction-based portfolio optimization model accumulated returns at weekly rates of 0.005 or 0.01 (shaded area of Fig.2), with a mean weekly return of 0.00594865, while the mean-variance model accumulated returns only at the 0.005 rate. The prediction-based portfolio optimization model presented a small positive *PCM* due to the periods when the model correctly switched between R1 to S1, and a mean *TI* of 0.12676 also due to the 9 switches between R1 and S1 observed during in the 142 weeks. The mean-variance model presented zero *PCM* and zero mean *TI* because it used only the R1 stocks in all weeks. Both models obtained portfolios with 1 stock in all the times.

4.4. Trading with real data

This section presents the trading experiments we performed with real data from the Brazilian stock market. With these experiments, we evaluated the prediction-based portfolio optimization model comparing its performance with that of the mean-variance model and the IBOVESPA market index. We used the real data described in Section 4.1, and the same trading strategy, evaluation metrics and periods of time of Section 4.3.

4.4.1. Efficient frontiers

In all the following experiments, we computed Efficient Frontiers with 30 portfolios for both the prediction-based portfolio optimization model and the mean-variance model at each of the 142 weeks of the trading period. A total of 8,520 $((30 + 30) \times 142 = 8,520)$ quadratic optimization problems were solved in the experiments.

Fig.3(a) shows two of the 142 Efficient Frontiers of the prediction-based portfolio optimization model (PRED), obtained at the first and last trading week ($t = 1$ and $t = 142$), respectively; and Fig.3(b) shows the same for the mean-variance model (MV). The efficient portfolios are shown in the squares and circles of Fig.3, while the 1σ vertical bars, obtained by calculating the square root of the portfolio risk (variance), are plotted for each efficient portfolio. These 1σ bars outline the regions, or confidence intervals, where the portfolio returns are expected to lie with a probability of 68.27%, according to the Normal distribution [41].

As the curves and respective confidence intervals of Fig.3 shows, the prediction-based portfolio optimization model produced higher expected returns than the mean-variance model for the same levels of risk. That is, the efficient portfolios of the prediction-based portfolio optimization model with risk superior to 0.001 have probability of negative returns $p(\hat{r}_p - \hat{\sigma}_p < 0) < 15.87\%$ ($\frac{100-68.27}{2} = 15.87\%$), while all the efficient portfolios of the mean-variance model have probability of negative returns $p(\bar{r}_p - \sigma_p < 0)$ significantly higher than that. To our knowledge, it is the first time in the literature that efficient frontiers are analyzed with the help of the confidence intervals of the portfolios' expected returns.

4.4.2. Trading with real data – low risk

This experiment evaluates the portfolio models selecting portfolios of low level of risk. We selected portfolios with weekly expected returns \bar{r}_p and \hat{r}_p equal to 0.0135, since they typically lie at the beginning of the Efficient Frontiers of the models with our data set (see Fig.3).

Fig.4 shows the Accumulated Returns obtained in the trading period with the prediction-based portfolio optimization model (PRED), the mean-variance model (MV), and the IBOVESPA index (IBOV), according with the trading strategy of Section 4.3. As Fig.4 shows, the prediction-based portfolio optimization model outperformed the mean-variance model at the end of the 142 weeks, and tracked the market index with a slight superiority than the mean-variance model, which oscillated around the IBOVESPA index. It is important to note that the mean-variance model performed below IBOVESPA in the period.

Fig.5 shows an evaluation of the Accumulated Returns of the portfolio models in the context of the Normal framework. The figure shows the realized and expected Accumulated Returns curves for both models. The realized Accumulated Returns curve is obtained as usual (see Eq.26), while the expected Accumulated Returns curve is obtained as follows. At each time t , the expected return of the selected portfolio at time $t - 1$ is accumulated to the portfolio's Accumulated Return at time $t - 1$ and is plotted as the portfolio's expected Accumulated Return for time t . Then, we plot a 3σ bar (3 times the square root of variance, or risk, of the selected portfolio at time $t - 1$) centered at the expected Accumulated Return for time t . This 3σ bar outlines a region where the portfolio's Accumulated Return for time t is expected to lie with a probability of 99.73%, according to the Normal distribution [41]. As Fig.5 shows, most of the realized Accumulated Returns of both models are inside the 3σ bars.

Table 4 summarizes the results of the IBOVESPA market index, and Table 5 summarizes the results of the portfolio models. As the tables show, the prediction-based portfolio optimization model achieved an Accumulated Return of 2.184123, which is 29% higher than the 1.691034 achieved by the mean-variance model, and is very close to the 2.188907 achieved by IBOVESPA (see Table 4). The prediction-based portfolio optimization model obtained an *ex-post* risk (variance) of 0.000744, which is 57.40% below the 0.001747 of the mean-variance model, and 25.40% below the 0.000998 of the IBOVESPA (see Table 4). Thus, the prediction-based portfolio optimization model performed at market return rates with a significant lower risk.

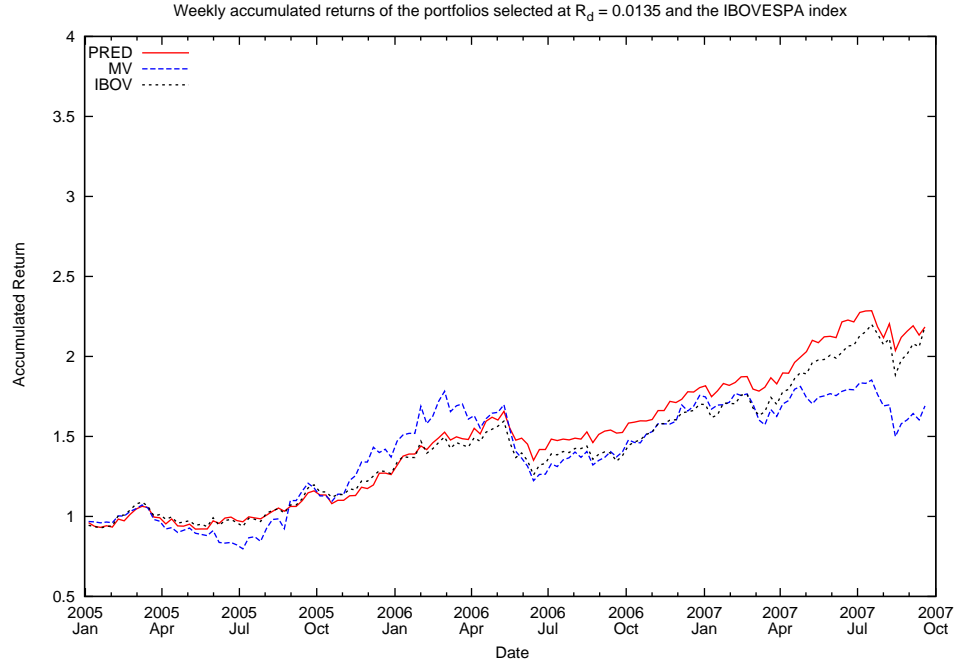


Fig. 4. Trading with real data – low risk: Accumulated Returns of the prediction-based portfolio optimization model (PRED) and the mean-variance model (MV), with $\hat{r}_p = 0.0135$ and $\bar{r}_p = 0.0135$, respectively, and of the IBOVESPA (IBOV) market index. The prediction-based portfolio optimization model outperformed the mean-variance model at the end of the 142 weeks, and tracked the IBOVESPA index better.

Table 4
Summary of the IBOVESPA market index.

Accumulated Return	2.188907		
	<i>mean</i>	σ^2	σ
weekly returns	0.006031	0.000998	0.031591
	<i>mean</i>	<i>min</i>	<i>max</i>
number of stocks	57	53	63

We used a paired *t-test* with significance level of 5% to compare the means of the 11 quarterly accumulated returns³ of the models and found a *p-value* of 0.384303, indicating that there is no difference between the mean of the prediction-based portfolio optimization model and the mean of the mean-variance model.

Both models presented a very small negative *PCM*, and the mean *TI* of 42% for the prediction-based portfolio optimization model was slightly better than the 45% for the mean-variance model.

The number of stocks in the prediction-based portfolios varied between 12 and 34, with an average of 21 stocks, while the number of stocks in the mean-variance portfolios varied between 2 and 21, with an average

³ Each quarterly accumulated return was calculated with Eq.26 during 13 ($\frac{52}{4} = 13$) weeks. We repeated the last weekly Accumulated Return of both models to obtain $142 + 1 = 143$ trading periods and exactly 11 quarterly Accumulated Returns ($\frac{143}{13} = 11$).

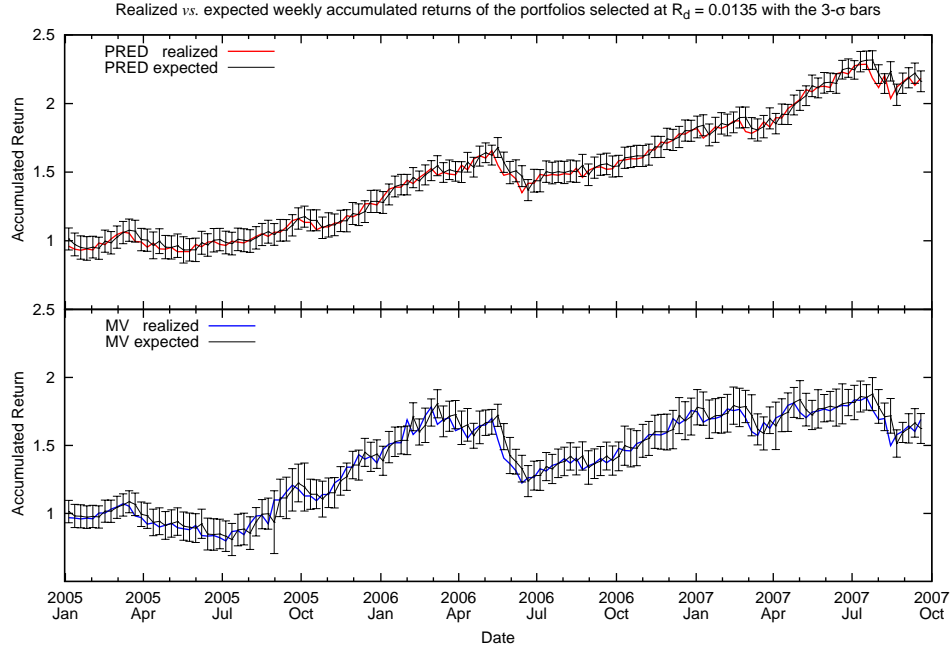


Fig. 5. Trading with real data – low risk: evaluation of the Accumulated Returns under the Normal framework's expected performance for the prediction-based portfolio optimization model (PRED) and the mean-variance model (MV).

Table 5
Trading with real data – low risk: summary of the trading during the 142 weeks

<i>Prediction-based portfolios selected at $\hat{r}_p = 0.0135$</i>				<i>Mean-variance portfolios selected at $\bar{r}_p = 0.0135$</i>			
Accumulated Return	2.184123			Accumulated Return	1.691034		
	<i>mean</i>	σ^2	σ		<i>mean</i>	σ^2	σ
weekly returns	0.005887	0.000744	0.027284	weekly returns	0.004564	0.001747	0.041804
<i>TI</i>	0.421122	0.057968	0.240766	<i>TI</i>	0.454311	0.104235	0.322854
<i>PCM</i>	-0.000590			<i>PCM</i>	-0.000636		
	<i>mean</i>	<i>min</i>	<i>max</i>		<i>mean</i>	<i>min</i>	<i>max</i>
number of stocks	21	12	34	number of stocks	7	2	21

of 7 stocks. Thus, the prediction-based portfolio optimization model achieved a better diversification level, and approached the Staltman's reference level of 30 stocks [44].

4.4.3. Trading with real data – medium risk

Investors frequently face the problem of diversifying into a portfolio of stocks and a risk-free asset (e.g. the Treasury Bill). In this case, it is shown that the optimal values of risk and return lie at a tangent of the

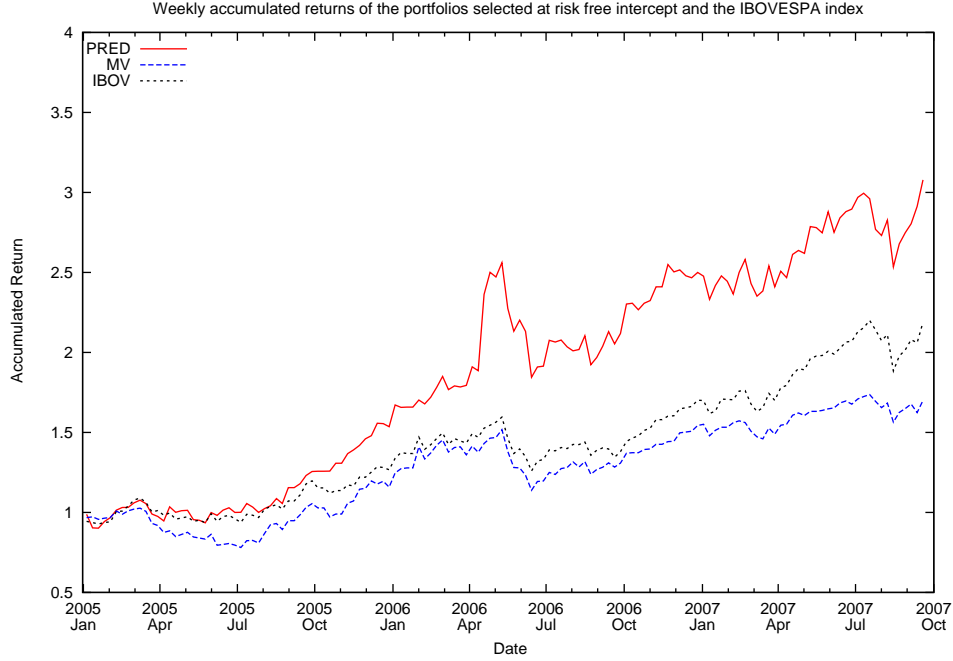


Fig. 6. Trading with real data – medium risk: Accumulated Returns for the prediction-based portfolio optimization model (PRED) and the mean-variance model (MV), selecting portfolios at the risk-free point of tangency, and for the IBOVESPA (IBOV) market index. The prediction-based portfolio optimization model outperformed the mean-variance model and the IBOVESPA index in almost all weeks.

Efficient Frontier (plotted with risks as standard deviations instead of variances) which intercepts the return axis at the risk-free return (R_{rf}) value. These optimal values correspond to the various combinations of the participation of the risk-free asset and the participation of the efficient portfolio that touches the tangent – all these combinations are also efficient [3,4]. The location of this point of tangency in the Efficient Frontier depends on the frontier's shape and the value of R_{rf} . In this experiment, we selected portfolios at this point of tangency, for a weekly risk-free return $R_{rf} = 0.001136$, which corresponds to a 6% annualized interest rate. These portfolios were located near the middle of the Efficient Frontiers of the models (see Fig.3).

Fig.6 shows the Accumulated Returns obtained in the trading period with the prediction-based portfolio optimization model (PRED), the mean-variance model (MV), and the IBOVESPA index (IBOV), according the same trading strategy of the previous experiment. As Fig.6 shows, the prediction-based portfolio optimization model outperformed the mean-variance model and the IBOVESPA index during almost all the weeks. It is important to note that the mean-variance model performed below IBOVESPA in the period.

Fig.7 shows the evaluation of the Accumulated Returns of the portfolio models in the context of the Normal framework as in the previous experiment. As Fig.7 shows, several Accumulated Returns of the prediction-based portfolio optimization model are outside the 3σ bars, while the major part of the Accumulated Returns of the mean-variance model are inside the 3σ bars. In this experiment, we are selecting portfolios with higher predicted returns (and risks) than in the previous experiment of Section 4.4.2. Hence, the prediction-based portfolio optimization model produced several predicted returns above the returns realized in the market, but in the right direction (see H_{R+} in Table 1), thus separating its expected and realized curves in several points.

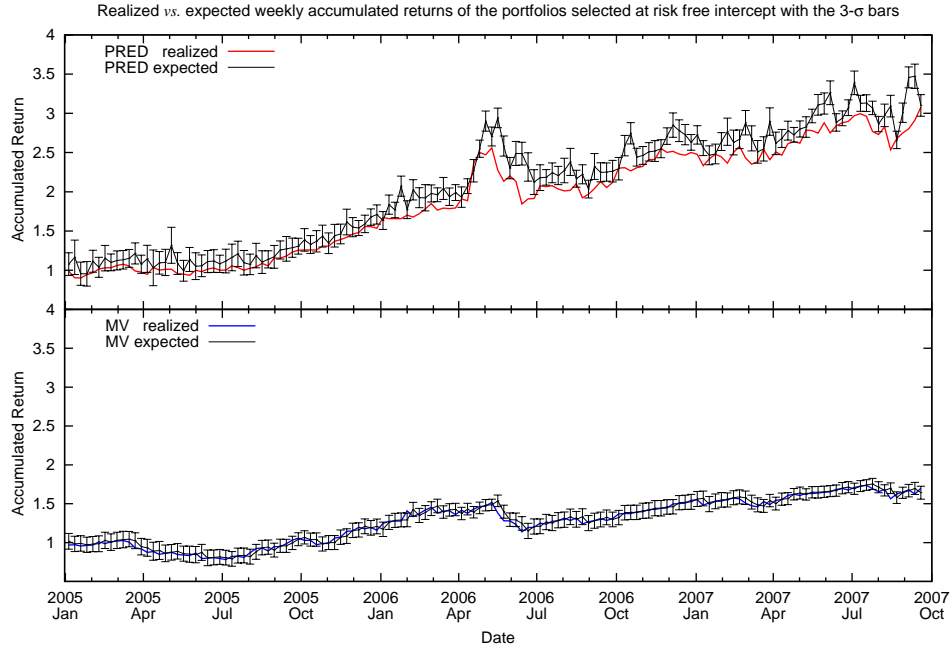


Fig. 7. Trading with real data – medium risk: evaluation of the Accumulated Returns under the Normal framework's expected performance for the prediction-based portfolio optimization model (PRED) and the mean-variance model (MV).

Table 6
Trading with real data – medium risk: summary of the trading during the 142 weeks

<i>Prediction-based portfolios selected at risk-free tangency point</i>				<i>Mean-variance portfolios selected at risk-free tangency point</i>			
Accumulated Return	3.078600			Accumulated Return	1.699988		
	<i>mean</i>	σ^2	σ		<i>mean</i>	σ^2	σ
weekly returns	0.008924	0.002003	0.044756	weekly returns	0.004292	0.001101	0.033184
<i>TI</i>	1.767750	0.071966	0.268266	<i>TI</i>	0.263251	0.015723	0.125393
<i>PCM</i>	0.0013431			<i>PCM</i>	-0.000645		
	<i>mean</i>	<i>min</i>	<i>max</i>		<i>mean</i>	<i>min</i>	<i>max</i>
number of stocks	6	2	14	number of stocks	11	6	24

Table 5 summarizes the results of the portfolio models. As table shows, the prediction-based portfolio optimization model achieved an Accumulated Return of 3.078600, which is 81.1% above the 1.699988 achieved by the mean-variance model, and 40.6% above the 2.188907 achieved by IBOVESPA (see Table 4). The prediction-based portfolio optimization model obtained an *ex-post* risk (variance) of 0.002003, which is 81.9% above the 0.001101 of the mean-variance model, and 100.7% above the 0.000998 of the IBOVESPA

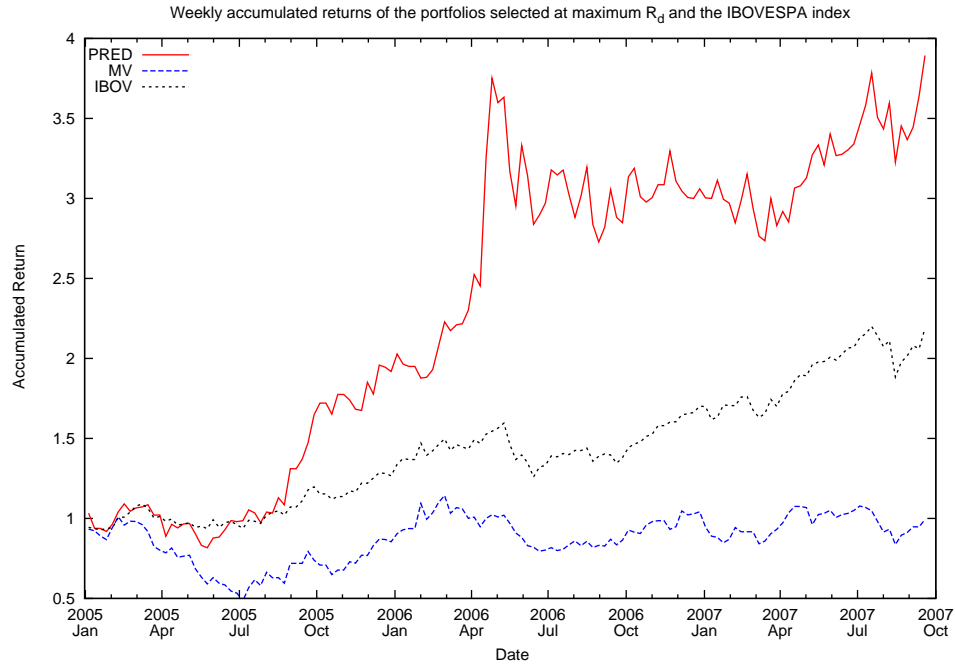


Fig. 8. Trading with real data – high risk: Accumulated Returns for the prediction-based portfolio optimization model (PRED) and the mean-variance model (MV), selecting portfolios with maximum expected return, and for the IBOVESPA (IBOV) market index. The prediction-based portfolio optimization model strongly outperformed the mean-variance model and the IBOVESPA index in almost all weeks.

(see Table 4).

We used a paired *t-test* with significance level of 5% to compare the means of the 11 quarterly accumulated returns of the models and found a *p-value* of 0.021143, indicating that the mean of the prediction-based portfolio optimization model is greater than the mean of the mean-variance model.

The prediction-based portfolio optimization model achieved a *PCM* of 0.0013431 while the mean-variance model achieved a *PCM* of -0.000645 . This shows that the prediction-based portfolio optimization model better explored the predictive opportunities than the mean-variance model.

The prediction-based portfolio optimization model achieved a mean *TI* of 176% while the mean-variance model achieved 26%. This shows that the prediction-based portfolio optimization model used a more aggressive strategy than mean-variance model, changing its participations more intensively.

The number of stocks in the prediction-based portfolios varied between 2 and 14, with an average of 6 stocks, while the number of stocks in the mean-variance portfolios varied between 6 and 24, with an average of 11 stocks. The prediction-based portfolio optimization model showed modest diversification because of its more aggressive strategy.

4.4.4. Trading with real data – high risk

This experiment evaluates the performance of the portfolio models employed for maximizing returns, as in traditional trading systems. In this application, the models are expected to select the stock with the higher expected return – the predicted return for the prediction-based portfolio optimization model, and the mean return for the mean-variance model; thus selecting portfolios at the far right end of the Efficient Frontiers.

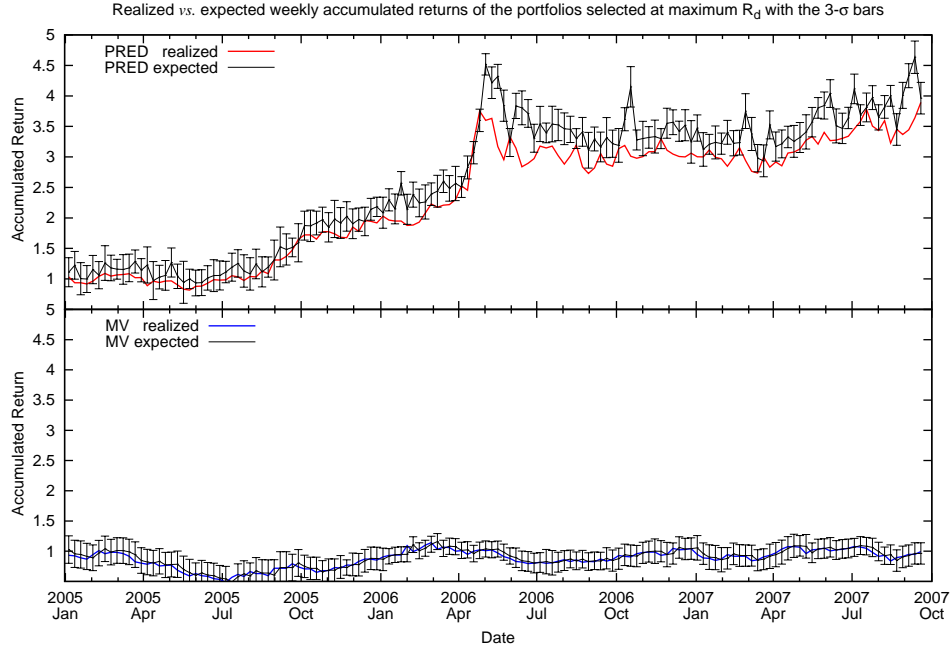


Fig. 9. Trading with real data – high risk: evaluation of the Accumulated Returns under the Normal framework’s expected performance for the prediction-based portfolio optimization model (PRED) and the mean-variance model (MV).

Fig.8 shows the Accumulated Returns obtained in the trading period with the prediction-based portfolio optimization model (PRED), the mean-variance model (MV), and the IBOVESPA index (IBOV), according the same trading strategy of the previous experiments. As Fig.8 shows, the prediction-based portfolio optimization model strongly outperformed the mean-variance model and the IBOVESPA index during almost all weeks. It is important to note that the mean-variance model performed far below IBOVESPA in the period.

Fig.9 shows the evaluation of the Accumulated Returns of the portfolio models in the context of the Normal framework as in the previous experiments. As Fig.9 shows, several Accumulated Returns of the prediction-based portfolio optimization model are outside the 3σ bars, while the major part of the Accumulated Returns of the mean-variance model are inside the 3σ bars – a behavior similar to that of the previous experiment of Section 4.4.3.

Table 7 summarizes the results of the portfolio models. As table shows, the prediction-based portfolio optimization model achieved an Accumulated Return of 3.893728, which is 291.9% above the 0.993455 achieved by the mean-variance model, and 77.9% above the 2.188907 achieved by IBOVESPA (see Table 4). The prediction-based portfolio optimization model obtained an *ex-post* risk (variance) of 0.003835, which is 21.2% above the 0.003164 of the mean-variance model, and 284.3% above the 0.000998 of the IBOVESPA (see Table 4).

We used a paired *t-test* with significance level of 5% to compare the means of the 11 quarterly accumulated returns of the models and found a *p-value* of 0.025636, indicating that the mean of the prediction-based portfolio optimization model is greater than the mean of the mean-variance model.

The prediction-based portfolio optimization model achieved a *PCM* of 0.006012 while the mean-variance model achieved a *PCM* of 0.002843. This shows that the prediction-based portfolio optimization model explored the predictive opportunities over two times more than the mean-variance model.

Table 7
Trading with real data – high risk: summary of the trading during the 142 weeks

<i>Prediction-based portfolios selected at maximum \hat{r}_p</i>				<i>Mean-variance portfolios selected at maximum \bar{r}_p</i>			
Accumulated Return	3.893728			Accumulated Return	0.993455		
	<i>mean</i>	σ^2	σ		<i>mean</i>	σ^2	σ
weekly returns	0.011444	0.003835	0.061933	weekly returns	0.001495	0.003164	0.056255
<i>TI</i>	1.900710	0.190071	0.435971	<i>TI</i>	0.453901	0.706788	0.840707
<i>PCM</i>	0.006012			<i>PCM</i>	0.002843		
	<i>mean</i>	<i>min</i>	<i>max</i>		<i>mean</i>	<i>min</i>	<i>max</i>
number of stocks	1	1	1	number of stocks	1	1	1

The prediction-based portfolio optimization model achieved a mean *TI* of 190% while the mean-variance model achieved 45%. This shows that the prediction-based portfolio optimization model used a much more aggressive strategy than the mean-variance model, selecting a different stock for its portfolios in almost all the times, since both models used only one stock in all of its portfolios.

5. Discussion

Portfolio selection is still an open question in financial theory. Investor behavior, expected utility and probabilistic frameworks, dynamic models, risk models and distributions of stock returns, transaction costs, predictive and robust models, only to cite a few, still remain in the research agenda [45,46,47,48,49,50].

The interest in exploiting prediction in portfolio selection is not new, but was traditionally focused on the difficult task of determining the distributions of the series of returns [51].

The performance of a neural network predictor in predicting future stock returns depends on many aspects of the network and the data, including the network's topology, training methods and noise. The central idea of this work is that, albeit predictors may not achieve good individual performance, when used in the investment in many stocks, their individual performances can be combined and produce superior results. Prediction-based portfolios can be more suitable for short-term investment than mean-variance portfolios. The experimental results showed in Section 4 are evidence of the validity of this hypothesis.

There are other works in the literature that explore prediction and diversification as strategy for short term investment. Lazo et al. [52] proposed a hybrid genetic-neural and statistical system for optimizing portfolios. A genetic algorithm was used for optimizing the mean-variance model with a set of stocks, and a subset of them with significant participations was selected for investment. A GARCH model was then built to forecast the volatility (variance) of these selected stocks, and a neural network was used for predicting its weekly returns using its past returns and this predicted volatility. A second genetic algorithm used the predicted returns, predicted variances and the covariances of returns for selecting portfolios with maximum return or minimum risk.

Although the system proposed by Lazo et al. uses predicted returns and efficient diversification, its risk measures are derived from the series of returns only (in the estimation of GARCH parameters and in the covariance of the stocks), thus giving equal treatment to predictors with possibly different performances. The prediction-based portfolio optimization model addresses this problem by deriving its risk measures

from the predictors' individual and collective performances using the variances and covariances of their prediction errors.

Hung et al. [53] extended the ASLD trading system [54] adding a risk control mechanism based on neural networks. During training, when future prices are "known", the extended ASLD system (EASLD) examines a set of N assets and generates instructions to buy, sell or hold assets based on a sequence of their "known" future prices. Optimal portfolio participations are then calculated for each asset that has a buy signal using the mean-variance or other classical portfolio selection models. N neural networks, one for each asset, are then trained to generate these participations, or hold and sell signals. During test, when future prices are unknown, the neural network's outputs are then used as signals to buy, hold and sell participations; i.e., the network is expected to learn the full operation of the system (trading instructions and portfolio optimization).

The EASLD trading system differs from our prediction-based portfolio optimization model in two fundamental ways: (i) EASLD trading system uses neural networks to predict trading signals, including participations, based on past prices and past trading signals, while the prediction-based portfolio optimization model uses neural networks to predict returns and employ these predicted returns and associated risks for optimizing portfolios. (ii) EASLD trading system uses diversification for addressing the issues of single asset and lack of risk control associated with trading systems. This is done by training neural networks on optimal portfolios, built in according to classical diversification frameworks. However, there is no guarantee that the resulting predicted portfolios will be optimal, as they are computed directly by neural networks. The prediction-based portfolio optimization model, on the other hand, assures optimality of its predicted portfolios by always selecting efficient portfolios at portfolio transitions.

6. Conclusion and Future Works

In this paper we predict stock returns using a new method named autoregressive moving reference neural network (AR-MRNN), and use AR-MRNN in the implementation of a new portfolio optimization model named prediction-based portfolio optimization model.

For each prediction, the AR-MRNN uses regression variables that are differences between the values of the time series and a determined past value used as reference. Some of the distributional aspects of the prediction errors produced by the AR-MRNN predictors were examined, and their Normality were verified on average for 92% of the stocks used. This showed that it is possible to have Normal prediction errors with the non-Normal series of returns and supported the development of our new prediction-based portfolio optimization model that uses the Normal framework.

Simulations with our model achieved returns 291% above the mean-variance model, with similar levels of risk. Also, the predictive portfolios showed a better market index tracking capability, achieving returns 77% above the IBOVESPA market index.

Our future works include: to research better neural predictors and training methods for minimizing the prediction errors, to insert market frictions in the re-balancing strategies, and to study other risk measures.

7. Acknowledgements

We would like to thank *Delegacia da Receita Federal do Brasil em Vitória-ES, Conselho Nacional de Desenvolvimento Científico e Tecnológico—CNPq-Brasil* (grants 308207/2004-1, 471898/2004-0, 620165/2006-5, and 309831/2007-5) and *Financiadora de Estudos e Projetos—FINEP-Brasil* (grants CT-INFRA-PRO-UFES/2005, CT-INFRA-PRO-UFES/2006) for their support to this research work.

References

- [1] H. M. Markowitz, Portfolio selection, *Journal of Finance* VII (1) (1952) 77–91.
- [2] H. M. Markowitz, *Portfolio Selection: Efficient Diversification of Investments*, 2nd Edition, John Wiley & Sons, New York, 1991.
- [3] W. F. Sharpe, G. J. Alexander, J. V. Bailey, *Investments*, 6th Edition, Prentice Hall, Upper Saddle River, New Jersey, 1999.
- [4] E. J. Elton, M. J. Gruber, S. J. Brown, W. N. Goetzmann, *Modern Portfolio Theory and Investment Analysis*, 7th Edition, John Wiley & Sons, Inc., 2007.
- [5] E. F. Fama, Portfolio analysis in a stable paretian market, *Management Science* 11 (3 Series A) (1965) 404–419.
- [6] S. J. Kon, Models of stock returns—a comparison, *The Journal of Finance* 39 (1) (1984) 147–165.
- [7] E. F. Fama, Efficient capital markets: A review of theory and empirical work, *The Journal of Finance* 25 (2) (1970) 383–417.
- [8] E. F. Fama, Efficient capital markets II, *Journal of Finance* 26 (5) (1991) 1575–1617.
- [9] E. F. Fama, Market efficiency, long-term returns, and behavioral finance, *Journal of Financial Economics* 49 (1998) 283–306.
- [10] B. G. Malkiel, The efficient market hypothesis and its critics, *The Journal of Economic Perspectives* 17 (1) (2003) 59–82.
- [11] A. Edmans, D. García, Ø. Norli, Sports sentiment and stock returns, *Journal of Finance* 62 (4) (2007) 1967–1998.
- [12] G. E. P. Box, G. M. Jenkins, G. C. Reinsel, *Time Series Analysis: Forecasting and Control*, 3rd Edition, Prentice Hall, 1994.
- [13] H. White, Economic prediction using neural networks: The case of IBM daily stock returns, in: *Proceedings of the IEEE International Conference on Neural Networks*, 1988, pp. 451–458.
- [14] F. Cao, L.J.; Tay, Support vector machine with adaptive parameters in financial time series forecasting, *Neural Networks, IEEE Transactions on* 14 (6) (Nov. 2003) 1506–1518.
- [15] K. Huarng, T. H. Yu, Ratio-based lengths of intervals to improve fuzzy time series forecasting, *IEEE Transactions on Systems, Man and Cybernetics – Part B: Cybernetics* 36 (2) (2006) 328–340.
- [16] K. Huarng, T. H. Yu, Y. W. Hsu, A multivariate heuristic model for fuzzy time-series forecasting, *IEEE Transactions on Systems, Man and Cybernetics – Part B: Cybernetics* 37 (4) (2007) 836–846.
- [17] R. Sharda, R. B. Patil, A connectionist approach to time series prediction: An empirical test, *Journal of Intelligent Manufacturing* 3 (5) (1992) 317–323.
- [18] R. Hansen, J.V.; Nelson, Neural networks and traditional time series methods: a synergistic combination in state economic forecasts, *Neural Networks, IEEE Transactions on* 8 (4) (Jul 1997) 863–873.
- [19] Y. Chen, B. Yang, A. Abraham, Flexible neural trees ensemble for stock index modeling, *Neurocomputing* 70 (4-6) (2007) 697–703.
- [20] F. Liu, G. S. Ng, C. Quek, RLDD: A novel reinforcement learning-based dimension and delay estimator for neural networks in time series prediction, *Neurocomputing* 70 (7–9) (2007) 1331–1341.
- [21] J. W. Lee, J. Park, J. O, J. Lee, E. Hong, A multiagent approach to Q-Learning for daily stock trading, *IEEE Transactions on Systems, Man and Cybernetics – Part A: Systems and Humans* 37 (6) (Nov. 2007) C1–853.
- [22] S. Thawornwong, D. Enke, The adaptive selection of financial and economic variables for use with artificial neural networks, *Neurocomputing* 56 (2004) 205–232.
- [23] J. Moody, M. Saffell, Learning to trade via direct reinforcement, *IEEE Transactions on Neural Networks* 12 (4) (2001) 875–889.
- [24] K. Pantazopoulos, L. Tsoukalas, N. Bourbakis, M. Brun, E. Houstis, Financial prediction and trading strategies using neurofuzzy approaches, *IEEE Transactions on Systems, Man and Cybernetics – Part B: Cybernetics* 28 (4) (Aug 1998) 520–531.
- [25] T. Hellström, Predicting a rank measure for stock returns, *Theory of Stochastic Processes* 22 (6) (2000) 64–83.
- [26] T. Hellström, Optimization of trading rules with a penalty term for increased risk-adjusted performance, *Advanced Modeling and Optimization* 2 (4) (2000) 135–149.
- [27] H. M. Markowitz, The utility of wealth, *The Journal of Political Economy* 60 (2) (1952) 151–158.
- [28] D. E. Fischer, R. J. Jordan, *Security Analysis and Portfolio Management*, 6th Edition, Prentice Hall International, 1995.
- [29] F. D. Freitas, A. F. De Souza, A. R. Almeida, Autoregressive neural network predictors in the Brazilian stock markets, in: *VII Simpósio Brasileiro de Automação Inteligente (SBAI)/II IEEE Latin American Robotics Symposium (IEEE-LARS)*, São Luis, Brasil, 2005, pp. 1–8.
- [30] F. D. Freitas, A. F. De Souza, A. R. Almeida, A prediction-based portfolio optimization model, in: *5th International Symposium On Robotics and Automation - ISRA 2006*, Hidalgo, Mexico, 2006, pp. 520–525.
- [31] S. Haykin, *Neural Networks: A Comprehensive Foundation*, 2nd Edition, Prentice Hall, Inc., 1999.
- [32] H. M. Markowitz, The optimization of a quadratic function subject to linear constraints, *Naval Research Logistics Quarterly* (3) (1956) 111–133.
- [33] F. Hamza, J. Janssen, Linear approach for solving large-scale portfolio optimization problems in a lognormal market, in: *IAA/AFIR Colloquium*, Nürnberg, Germany, 1996, pp. 1019–1039.
- [34] W. F. Sharpe, A simplified model for portfolio analysis, *Management Science* 2 (9) (1963) 277–293.

- [35] H. Konno, H. Yamazaki, Mean-absolute deviation portfolio optimization model and its applications to Tokyo stock market, *Management Science* 37 (5) (1991) 519–531.
- [36] F. A. Sortino, R. van der Meer, Downside risk—capturing what’s at stake in investment situations, *Journal of Portfolio Management* (1991) 27–31.
- [37] T. Hellström, Data snooping in the stock market, *Theory of Stochastic Processes* 5 (21) (1999) 33–50.
- [38] S. Lawrence, C. L. Giles, A. C. Tsoi, Lessons in neural network training: Overfitting may be harder than expected, in: *Proceedings of the Fourteenth National Conference on Artificial Intelligence (AAAI-97)*, 1997, pp. 540–545.
- [39] J. Moody, Prediction risk and architecture selection for neural networks, in: V. Cherkassky, J. H. Friedman, H. Wechsler (Eds.), *From Statistics to Neural Networks: Theory and Pattern Recognition Applications*, Springer, NATO ASI Series F, 1994.
- [40] J. S. Armstrong, F. Collopy, Error measures for generalizing about forecasting methods: Empirical comparisons, *International Journal of Forecasting* 8 (1) (1992) 69–80.
- [41] M. R. Spiegel, L. J. Stephens, *Schaum’s Outline of Theory and Problems of Statistics*, 3rd Edition, McGraw-Hill, 1998.
- [42] M. Grinblatt, S. Titman, Performance measurement without benchmarks: An examination of mutual fund returns, *The Journal of Business* 66 (1) (1993) 47–68.
- [43] R. R. Grauer, N. H. Hakansson, Applying portfolio change and conditional performance measures: The case of industry rotation via the dynamic investment model, *Review of Quantitative Finance and Accounting* 17 (3) (2001) 237–265.
- [44] M. Statman, How many stocks make a diversified portfolio?, *The Journal of Financial and Quantitative Analysis* 22 (3) (1987) 353–363.
- [45] L. Huang, H. Liu, Rational inattention and portfolio selection, *Journal of Finance* LXII (4) (2001) 1999–2217.
- [46] M. W. Brandt, Dynamic portfolio selection by augmenting the asset space, *Journal of Finance* LXI (5) (2006) 2187–2217.
- [47] D. Maspero, Portfolio selection for financial planners, *newfin Working Paper* 3/04 (2004).
- [48] M. Kritzman, S. Myrgren, S. Page, Optimal execution for portfolio transitions, *Journal of Portfolio Management* (2007) 33–39.
- [49] Y. Ait-Sahalia, M. W. Brandt, Variable selection for portfolio choice, *Journal of Finance* LVI (4) (2001) 1297–1351.
- [50] F. J. Fabozzi, P. N. Kolm, Robust portfolio optimization: Recent trends and future directions., *Journal of Portfolio Management* (2007) 40–48.
- [51] J. Fried, Forecasting and probability distributions for models of portfolio selection, *Journal of Finance* XXV (3) (1970) 539–554.
- [52] J. G. L. Lazo, M. A. C. Pacheco, M. M. B. R. Vellasco, Portfolio selection and management using a hybrid intelligent and statistical system, in: S. Chen (Ed.), *Genetic Algorithms and Genetic Programming in Computational Finance*, Kluwer Academic Publishers, Boston, 2002, pp. 221–238.
- [53] K. Hung, Y. Cheung, L. Xu, An extended ASLD trading system to enhance portfolio management, *IEEE Transactions on Neural Networks* 14 (2) (2003) 413–425.
- [54] L. Xu, Y. Cheung, Adaptive supervised learning decision networks for trading and portfolio management, *Journal of Computational Intelligence in Finance* 5 (6) (1997) 11–16.