A Prediction-Based Portfolio Optimization Model

Fábio Daros de Freitas, *Member, IEEE,* Alberto Ferreira De Souza, *Member, IEEE,* and Ailson Rosetti de Almeida

***Abstract*— This work presents a prediction-based portfolio optimization model that uses Normal prediction errors as risk measure. A new autoregressive neural network predictor is used to predict future stock returns and its prediction errors are used as risk measure. In this predictor, the differences between the values of the series of stock returns and a specified past value are the regression variables. A large set of experiments with real data from the Brazilian stock market was employed to evaluate our portfolio optimization model, which included the examination of the normality of the errors of prediction. Our results showed that it is possible to obtain Normal prediction errors with non-Normal series of stock returns, and that our prediction-based portfolio optimization model outperforms the Markowitz portfolio selection model showing better return for the same risk.**

***Keywords*— Neural Networks, Time Series Prediction, Portfolio Optimization.**

I. Introduction

Statistical Modeling can be used for providing uncertainty measures of realization of investments return. These uncertainty measures, named *risk*, make Statistical Modeling suitable for comparing investment alternatives. This feature is central to the Modern Portfolio Theory (MPT), first introduced by Harry M. Markowitz through his celebrated model of portfolio[1] selection [1]. In this model, the total risk of an investment in various stocks is minimized by the optimal selection of stocks with low joint risk, which provides a mechanism of loss compensation known as *Efficient Diversification*. The portfolio selection process, then, consists of finding, in a large collection of stocks, the participation (i.e. individual proportion) of each stock that minimizes the portfolio's risk at a given portfolio return, or maximizes the portfolio's return at a given risk. The model assumes that the historical series of returns of each stock follows a Normal distribution, uses the mean of the series as a prediction of the stock future returns, the variance as a measure of the risk of the stock, and the covariance of the stocks' returns as a measure of joint risk (the standard deviation, semi-variance and absolute deviation may also be used as a measure of risk in the model).

After the Markowitz model, many other models that use its basic assumptions (that the series of returns are Normal and that the moments of these series can be used as measure of future return and risk) appeared [2]. In all these models, known today as classic models, the portfolio's expected return is given by the linear combination of the expected returns of its stocks (the mean returns) and the participations of each stock in the portfolio. The portfolio risk, in turn, varies, but it is often related with the moments about the mean of the joint Normal distribution of the series of returns of its stocks.

Despite the wide adoption of the classical methods of portfolio selection, it is important to mention that the distributions of the series of returns often exhibit kurtosis and skewness [3], challenging the assumption of normality of the series of returns. In addition, the realization of mean returns tends to be verified only in the long term. This has stimulated the development of predictive models based on *Time Series Analysis* and other non-linear methods, like artificial neural networks, as a way of supporting the needs of investment on shorter horizons.

This paper presents a prediction-based portfolio optimization model that uses Normal prediction errors as risk measure. We have used a new autoregressive neural network (ARNN) predictor to predict future stock returns, and the variance of its prediction errors as risk measure. In this predictor, the differences between the values of the series of returns and a determined past value are the regression variables, instead of simply the returns, as in the traditional prediction methods. Our experimental results clearly suggest that the prediction errors are Normal (an earlier version of this verification of normality appears in [4]). Our experiments have also shown that our portfolio selection model outperforms the Markowitz model, presenting better returns with the same risk while using the same stocks in the same periods of time.

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

The one-period stock return in time t, r_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.

$$r_t = \frac{P_t - P_{t-1}}{P_{t-1}} \tag{1}$$

where, 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:

$$r' = (r_{t-N+1}, \ldots, r_{t-1}, r_t) \tag{2}$$

The time-series-based prediction of a future return of the stock can be defined as the process of using r' for obtaining an estimate of r_{t+l}, where $l \geq 1$. The value of l directly affects the choice of prediction method. For $l = 1$, or one-period

F. D. Freitas, Secretaria da Receita Federal, Programa de Pós Graduação em Engenharia Elétrica – UFES, e-mail freitas@computer.org
A. F. De Souza, Programa de Pós Graduação em Informática – UFES, e-mail albertodesouza@gmail.com
A. R. Almeida, Programa de Pós Graduação em Engenharia Elétrica – UFES, e-mail ailson@ele.ufes.br
Universidade Federal do Espírito Santo (UFES), Av. Fernando Ferrari, s/n, 29075-910, Vitória, ES, Brazil

[1]A portfolio is a collection of stocks, bonds, or other forms of investment. In this paper a portfolio is always a collection of stocks only.

prediction, the standard choice are the autoregressive methods [5], while for $l > 1$, or multiple-period prediction, techniques such as the method of Temporal Differences [6] are used; one can also make l recursive one-period predictions in this case. This work investigates only the one-period prediction problem.

The classical neural network time series predictor is an autoregressive based predictor [7], $\mathcal{S}_{\mathcal{R}}$, whose inputs are the present value and the $p-1$ past values of the series, and the output is an estimate of the value for the next time period, as shown in Eq.3 for the series of returns.

$$(r_{t-p+1}, \ldots, r_{t-1}, r_t) \to \mathcal{S}_{\mathcal{R}} \to \hat{r}_{t+1} \tag{3}$$

The number of inputs, $p\,(p > 0)$, is the *regression order*, and some techniques for obtaining it are shown in [5]. Thus, after trained with its input-output pairs, the neural network predictor implements a fully operational non-linear multiple regression model for the series of returns.

In this work, we propose a new ARNN method, inspired in the way one observes a time series graph to guess its future value. In this task, one tends to concentrate the visual attention in the last points of the graph, creating an imaginary frame that delimits a region of the graph and offers an image with sufficient visual information for extrapolating the next value of the series. In this prediction scheme, one uses the value of some point inside the region as a reference value. We mimic that in our method by subtracting a reference value from the returns presented to the neural network. This reference value is equal to one of the elements of the series inside the observation frame $t-p-k+1, ..., t\,(p, k > 0)$. We named this method *autoregressive moving reference neural network predictor - ARMRNN(p,k)*, where p is the regression order and k is the reference lag. Then, the AR-MRNN inputs and outputs become:

$$(r_{t-p+1} - z, \ldots, r_t - z) \to \mathcal{S}_{\mathcal{R}} \to \widehat{r_{t+1} - z} \tag{4}$$

were

$$z = r_{t-p-k+1} \tag{5}$$

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

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

With the autoregressive moving reference neural network method, the values encoded into the neural network weights are usually smaller than those of the series of returns, which minimizes the possibility of saturation at the neuron's output, enhancing the dynamic range of the network and its ability of representing the series of returns. This alleviates the needs of pre-processing, like normalization and detrending, and produces smaller weights values that regularize the network and enhance its generalization capabilities [8].

III. A NEW MODEL OF PORTFOLIO SELECTION WITH PREDICTED RETURNS

In this section we briefly present the Markowitz's model of portfolio selection and our new model, which is qualitatively compared with the Markowitz's model.

A. The Markowitz's portfolio selection model

The Markowitz's model is based on the return and risk measures of a linear combination of stocks, which are obtained from the individual return and risk measures of each stock in this combination. The original and until today widely used measure of expected return of one stock is the arithmetic mean of the series of returns of this stock, defined in Eq.7.

$$\bar{r} = \frac{1}{N} \sum_{t=1}^{N} r_t \tag{7}$$

Where $\bar{r}$ is the expected return of the stock in time $t+1$, N is the number of past times and r_t is the realized return in time t, defined in Eq.1.

The risk measure proposed by Markowitz, the statistical variance of the series of returns (Eq.8), reflects the uncertainties of realization of the expected return (Eq.7).

$$\nu = \sigma^2 = \frac{1}{N-1} \sum_{t=1}^{N} (r_t - \bar{r})^2 \tag{8}$$

The portfolio risk is defined as the variance of a linear combination of stocks [2] and is presented in Eq.9.

$$V = {\sigma_p}^2 = \sum_{i=1}^{M} {X_i}^2 \sigma_i^2 + \sum_{i=1}^{M} \sum_{j=1, j \neq i}^{M} X_i X_j \gamma_{ij} \tag{9}$$

In Eq.9, the first sum represents the individual risk components for each one of the M stocks, and the second group of sums represents the combined risks of each pair of stocks; V is the total portfolio risk, which is equal to the portfolio variance, ${\sigma_p}^2$; X_i is the participation of stock i in the portfolio; σ_i^2 is the individual risk of stock i; and term γ_{ij} is the covariance of stocks i and j, which is defined as:

$$\gamma_{ij} = \frac{1}{N-1} \sum_{t=1}^{N} ({r_i}_t - \bar{r}_i)({r_j}_t - \bar{r}_j) \tag{10}$$

where ${r_i}_t$ and ${r_j}_t$ are the returns of stocks i and j in the time t, $\bar{r}_i$ and $\bar{r}_j$ are the expected returns of stocks i and j, and N is the number of observed past times.

This covariance is the measure of the *interactive or joint risk* between two stocks, and its effect on the portfolio risk V is crucial for obtaining a portfolio with risk lower than the sum of the individual risks of its stocks: the foundation of the *Markowitz's Portfolio Optimization Model* is this interactive risk [1].

Markowitz's optimization model minimizes the portfolio risk V for a desired portfolio return R by solving the quadratic programming problem formulated as:

$Minimize$

$$V = \sum_{i=1}^{M} {X_i}^2 \nu_i + \sum_{i=1}^{M} \sum_{j=1, j \neq i}^{M} X_i X_j \gamma_{ij} \tag{11}$$

$Subject\ to$

$$\sum_{i=1}^{M} X_i \bar{r}_i \geq R \tag{12}$$

[2]This development is detailed in [9].

$$\sum_{i=1}^{M} X_i = 1 \tag{13}$$

$$X_i \geq 0, i = 1, \ldots, M \tag{14}$$

Eq.11 is the minimized objective function, the portfolio risk; Eq.12 is the desired return constraint, that guarantees a minimum portfolio return R; Eq.13 guarantees total resource allocation and Eq.14 restrict the model for purchase trades only.

The set of all portfolios with minimum risk for various desired returns is obtained via parameterizing of the model for R. These portfolios are named *Efficient Portfolios* and this investment strategy is named *Efficient Diversification*. The bidimensional locus of the risk-return space where all possible efficient portfolios lies is denoted *Efficient Frontier*. Each set of stocks has its own efficient frontier, which depends only on the individual expected returns and risks of each stock and its time series correlations (i.e. its covariances matrix).

B. Portfolio Selection with Predicted Returns

Since the original Markowitz's mean-variance proposition, computational feasibility, model simplifications and development of risk measures have received considerable research attention. These efforts strongly contributed to establish the portfolio theory and practice, producing remarkable results [10], [11], [12].

However, the mean estimates for the future returns are expected to be verified only in the long term; therefore, they are not suitable for using in active portfolio management and other short-term strategies. The using of better prediction methods for obtaining the estimates of future short-term returns, with suitable risk measures, may derive new predictive portfolio models for these short-term applications.

This section formulates a new portfolio optimization model, where the variance of the errors of predictions, provided by an AR-MRNN predictor, is the individual risk measure.

Let the real stock return in time t, r_t, be given by:

$$r_t = \hat{r}_t + \epsilon_t \tag{15}$$

where, $\hat{r}_t$ is the predicted return for time t, obtained at time $t-1$, and ϵ_t is the prediction error for time t. For a non-biased predictor, the series $\varepsilon' = (\epsilon_{t-n+1}, \epsilon_{t-1}, \ldots, \epsilon_t)$ of the n errors of prediction, defined from Eq.15 by:

$$\epsilon_t = r_t - \hat{r}_t \tag{16}$$

must be statistically independent and identically distributed (iid), with mean and variance given by:

$$\mu_\epsilon = \bar{\epsilon} = 0 \tag{17}$$

$$\hat{\nu} = \sigma_\epsilon^2 = \frac{1}{n-1} \sum_{t=1}^{n} {\epsilon_t}^2 \tag{18}$$

This variance of the errors of prediction (Eq.18) is the risk of the predicted value, used in the model as the individual risk of the stocks.

1) Return and risk: In this new portfolio optimization model, the predicted portfolio return, $\hat{r}_p$, is the sum of the predicted returns of its stocks weighted by the participation of each stock in the portfolio (Eq.19):

$$\hat{r}_p = \sum_{i=1}^{M} X_i \hat{r}_i \tag{19}$$

The portfolio risk is the variance of the joint Normal distribution of the combination of the errors of prediction, expressed by:

$$\hat{V} = \hat{\sigma}_p^2 = \sum_{i=1}^{M} {X_i}^2 \sigma_{\epsilon\, i}^2 + \sum_{i=1}^{M} \sum_{j=1, j\neq i}^{M} X_i X_j \gamma_{\epsilon ij} \tag{20}$$

In Eq.20, the first sum represents the individual prediction risk components for each one of the M stocks, and the second group of sums represents the combined prediction risks of each pair of stocks; $\hat{V}$ is the total prediction-based portfolio risk; X_i is the participation of stock i in the portfolio; $\sigma_{\epsilon\, i}^2$ is the prediction risk of stock i; and term $\gamma_{\epsilon ij}$ is the interactive prediction risk of stocks i and j, which is obtained from the covariance of the errors of prediction of the stocks i and j:

$$\hat{\gamma}_{ij} = \gamma_{\epsilon ij} = \frac{1}{n-1} \sum_{t=1}^{n} \epsilon_{it} \epsilon_{j\, t} \tag{21}$$

2) Portfolio optimization model: With the fundamental measures defined, the portfolio optimization model is formulated as:

$Minimize$

$$\hat{V} = \sum_{i=1}^{M} {X_i}^2 \hat{\nu}_i + \sum_{i=1}^{M} \sum_{j=1, j\neq i}^{M} X_i X_j \hat{\gamma}_{ij} \tag{22}$$

$Subject\ to$

$$\sum_{i=1}^{M} X_i \hat{r}_i \geq R \tag{23}$$

$$\sum_{i=1}^{M} X_i = 1 \tag{24}$$

$$X_i \geq 0, i = 1, \ldots, M \tag{25}$$

Eq.22 is the minimized objective function, the prediction-based portfolio risk; Eq.23 is the desired return constraint, that guarantees a minimum portfolio return R; Eq.24 guarantees total resource allocation and Eq.25 restrict the model for purchase trades only.

Despite of employing the joint Normal modeling, this model strongly differs from the Markowitz model because: (i) the distribution model of the stock is obtained with the distribution of the prediction errors centered in the predicted return, instead of with the distribution of the series of returns, (ii) the expected returns are the predicted returns instead of the mean of the series of returns, and (iii) the individual and interactive risks are obtained with the variances and covariances of the prediction errors (Eq.18 and Eq.21), instead of with the variances and covariances of the series of returns. This model is based on the assumptions of Eq.17 and on the normality of the errors of prediction, which are experimentally supported by the results that will be shown in Section V-A.

IV. Experimental methods

A set of experiments was designed with the goal of (i) verifying the normality of the errors of prediction, and (ii) evaluating our portfolio selection model.

A. Data

We selected 46 stocks from the Ibovespa index portfolio of the Brazilian Bovespa stock exchange of the first quarter of 2005. For each stock, we computed the weekly returns sampled at Wednesdays in the period between 16-feb-2000 and 09-mar-2005, obtaining 46 series of 265 returns as the data set for our experiments. We have used a moving window of 200 weeks for training and 5 weeks for testing, and performed 60 predictions, one for each stock ($46X60 = 2.760$ predictions), as described below.

B. Neural network predictors

We used fourth-order autoregressive one-lag moving reference predictors, AR-MRNN(4,1), with $p = 4$ and $k = 1$. They were implemented with a fully-connected feedforward neural network with 2 hidden layers, sigmoidal activation function, and 4:30:15:1 topology (4 input neurons, 30 neurons in the first hidden layer, 15 neurons in the second hidden layer, and 1 output neuron). We have trained one AR-MRNN(4,1) network for each one of the 46 stocks, employed 60 times ($46X60 = 2.760$ training sessions); we used the 64 ATHLON XP 1800 nodes Enterprise cluster of the Departamento de Informática of Universidade Federal do Espírito Santo (http://www.inf.ufes.br/~lcad) for that. The training was conducted during 200.000 epochs using the backpropagation algorithm with learning rate of 0.009 and inertia of 0.95 [8]. These topological and training parameters were determined empirically[3] [4]

To reduce overfitting, we employed a simplified form of cross validation [8], which consisted of dividing the training data set for each one of the 60 predictions into two segments: a training segment and a validating segment. The training segment for the first of the 60 predictions, the 200 weekly returns from 16-feb-2000 to 10-dec-2003, was used for updating the neural networks weights during the 200.000 training epochs, while the validating segment, the 5 weekly returns from 17-dec-2003 to 14-jan-2004, was used for choosing the best weights observed during the training. In order to select the best weights, at each 1.000 training epochs, the neural network of each stock was used for predicting the returns within the validating segment. The weights responsible for the smaller root mean squared prediction error (RMSE) observed with the validating segment were used for the prediction of the return of 21-jan-2004. This procedure was repeated for each of the remaining 59 predictions, advancing the window of 205 weeks, one week at a time, in all dates mentioned. The sizes of the segments were obtained with the heuristic proposed in [8].

[3]The series of stock returns are a high-frequency data, and smaller networks with fewer degrees of freedom tipically underfitted the training data.

[4]Such large networks (2 hidden layers, 46 neurons and 631 synapses) provided good training and generalization errors [13], but demanded a smoother training, with small learning rate and relatively large number of epochs, to avoid the saturation of the neurons' outputs.

α	H_0 not rejected	H_0 rejected
Chi-square test for the 46 series of prediction errors		
0.01	44 (96%)	2 (4%)
0.05	39 (85%)	7 (15%)
0.10	36 (78%)	10 (22%)
Chi-square Test for the 46 series of returns		
0.01	14 (30%)	32 (70%)
0.05	6 (13%)	40 (87%)
0.10	5 (11%)	41 (89%)

TABLE I
NORMALITY OF THE PREDICTION ERRORS FOR THE AR-MRNN PREDICTORS AND FOR THE SERIES OF RETURNS

V. Experiments and results

A. Normality of the errors of prediction

The normality of the errors of prediction of the AR-MRNN(4,1) predictors and of the series of returns was tested and the results are shown in Table I. The Chi-square tests of the 46 series of prediction errors and returns were conducted for the standard significance levels (α) of 0.01, 0.05 and 0.10. The normality of the prediction errors was accepted (i.e. not rejected) for 44 (96%), 39 (85%) and 36 (78%) of the AR-MRNN(4,1) predictors, respectively. However, for the same standard significance levels, the normality of the series of returns was accepted for 14 (30%), 6 (13%) and 5 (11%) of the series, respectively, only. These results suggest that it is possible to obtain Normal prediction errors from non-Normal series of returns. This fact is especially important for the developing of new predictive portfolio selection models that take the benefit of the theoretical developments of the Normal framework.

B. Performance of the portfolio selection with predicted returns

We have used a window of 40 predictions, starting with the first of the 60 predictions, for estimating the variances and covariances of the errors of prediction required for the predictive model of portfolio selection (Eq.22). The mean returns and variances required for the Markowitz model were computed from a window of 245 of the 265 weeks, starting from the first week. The portfolios formed with both predictive and Markowitz models were then evaluated for the remained 20 weeks of the dataset. The evaluation was conducted by selecting portfolios from the Efficient Frontier of each model, obtained via the optimization procedure described for each model in Section III.

We have simulated the investments in the period mentioned above and computed the accumulated returns. The investment strategy used was to re-balance the portfolios at the end of each period of 4 consecutive weeks; so, we formed a total of 5 portfolios for each model (20 weeks/4).

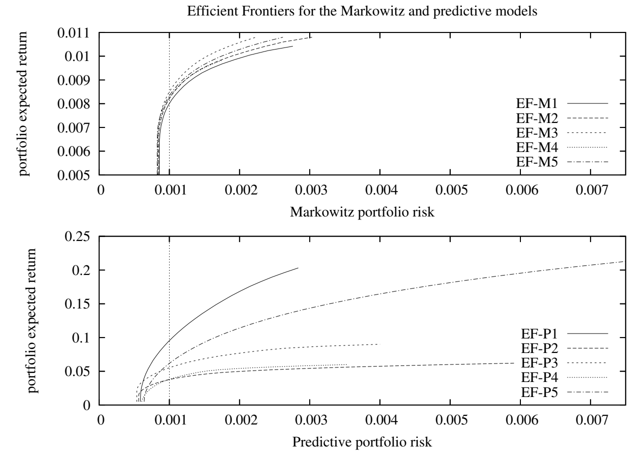

Fig. 1. **Efficient frontiers for the Markowitz (top) and predictive (bottom) models. The predictive model shows a greater diversity of solutions than the Markowitz model.**

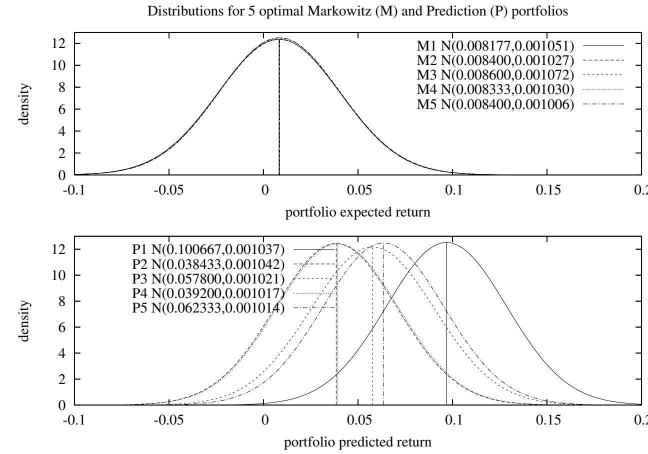

Fig. 2. **Normal distributions of each one of the 5 selected portfolios for the Markowitz (top) and predictive (bottom) models. The predictive portfolios present different values of expected returns, exhibiting the predictive short-term properties of the predictive model.**

Figure 1 shows the Efficient Frontiers for the Markowitz and predictive models, obtained via parameterizing of the models of Section III for R. Starting with a minimal value of 0.005, all other values of R where selected for obtaining Efficient Frontiers with 30 optimal portfolios for each model, in the space of each problem's viable solutions ($30X5X2 = 300$ quadratic problems). Thus, we used R values between 0.005 and 0.0108 for the Markowitz portfolios, and between 0.005 and 0.2128 for the predictive portfolios.

The top of Fig.1 shows the 5 Efficient Frontiers for the Markowitz model – EF-M1, EF-M2, EF-M3, EF-M4 and EF-M5, obtained at the initial time and in the sequence of the 4 re-balancing times during the simulations. Their optimal risk values stayed between 0.000827 and 0.003, for returns between 0.005 and 0.0108, respectively. The bottom of Fig.1 shows the 5 Efficient Frontiers for the predictive model – EF-P1, EF-P2, EF-P3, EF-P4 and EF-P5, obtained at same times and with optimal risks between 0.000537 and 0.0075, for returns between 0.005 and 0.2128, respectively. We can verify, in Fig.1, that the predictive portfolios showed risk and return values significantly better than those of the Markowitz portfolios, in Efficient Frontier terms. This superior performance can be explained by the greater diversity of solutions found by the predictive model (see the positions and shapes of the Efficient Frontiers of Fig.1), which better capture the short-term aspects than the Markowitz model.

After obtaining all Efficient Frontiers, we conducted investment simulation in the period above mentioned and calculated the accumulated returns for both models. The portfolios were selected from their Efficient Frontiers at risk level of 0.001 (variance of the joint Normal distribution of the portfolios – see the vertical dotted lines in Fig.1) and simulated during the 5 sequences of 4 consecutive weeks.

Figure 2 shows the Normal distributions of each one of the 5 selected portfolios for the Markowitz and predictive models, obtained with their respective expected returns and risks. The portfolios M1, M2, M3, M4 and M5 were obtained with Markowitz model and their distributions are shown at the top of Fig.2. They present very similar expected returns due to the very small variation of the weighted mean (by each X_i) of

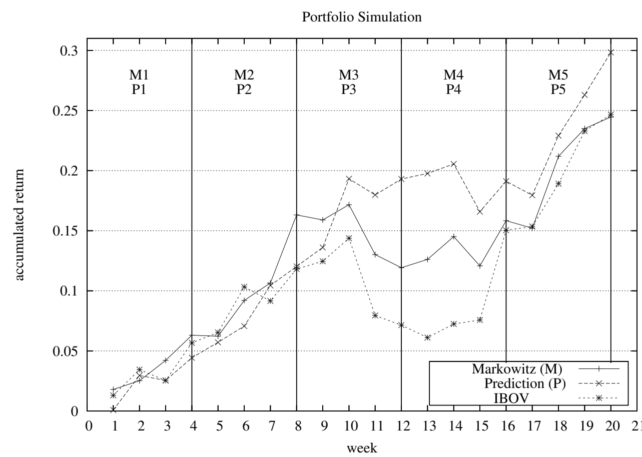

Fig. 3. **Accumulated returns for the Markowitz and predictive portfolios compared with the Ibovespa (IBOV) market index. The predictive portfolios (P) survived the IBOV drop after the 10th week better than the Markowitz (M) portfolios.**

the stocks returns in the 245-week periods. The portfolios P1, P2, P3, P4 and P5 were obtained with the prediction model and their distributions are shown at the bottom of Fig.2. As the graph shows, they present very different values of expected returns, exhibiting its predictive short-term properties.

Figure 3 shows a comparison of the accumulated returns of the 5 Markowitz portfolios, the 5 predictive portfolios and the Ibovespa (IBOV) market index. As mentioned above, the portfolios were re-balanced 4 times, at weeks 4, 8, 12 and 16 of the simulation period, using the data available at these moments. The simulation period for each one of the 5 pairs of portfolios is showed between the vertical lines in Fig.3. As Fig.3 shows, initially, Ibovespa and both portfolios presented equivalent performances. In week 8, the Markowitz portfolio showed a significant increase in performance, but it was surpassed by the predictive portfolio in week 10, which survived to an important market drop in following weeks (see Ibovespa graph). The remaining predictive portfolios sustained superior performance until the end of the period under analysis, when the predictive portfolios achieved accumulated return of 29.83%, while the Markowitz portfolios obtained 24.45%, close to the Ibovespa performance of 24.66%

VI. Discussion

Portfolio selection is still an open question in financial theory. Expected utility and probabilistic frameworks, risk models and distributions of stock returns still remains in the research agenda [2], [14]. The interest in exploiting predictive aspects in portfolio selection is not new [15], but prediction has been traditionally focused on the difficult task of determining the distributions of the series of returns only.

The prediction of future returns of a stock is a well-known difficult task [4], [16]. In recent years, artificial intelligence and machine learning techniques have been successfully applied in the trading of individual and multiple assets for predicting more predictable measures, like trading signals [17], [18] and rank measures [19]. These methods are more suitable for large scale applications than determining the distributions of the series of returns or fitting statistical models.

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 have a not so good individual performance, when combined for the investment in many stocks (portfolio), their individual performances can be combined by the efficient diversification mechanism with risk measures based on their expected performances. This expected performance is known as *prediction risk* [20], and in our study, after verifying the normality of the non-biased errors of prediction for the AR-MRNN predictors, we used their variance (of the errors of prediction) as this risk measure and took advantage of the Normal framework for portfolio optimization.

Through exploiting the predictive aspects of the series of returns, these predictive portfolios, then, can be more suitable for short-term investment and active portfolio management than the mean-variance portfolios. The experimental results shown in Section V are evidence of the validity of this hypothesis.

VII. Conclusion and Future Works

In this paper we have used a new prediction method, named autoregressive moving reference neural network (AR-MRNN) to predict stock returns. In this method, the regression variables are the differences between the values of the series and a determined past value. Some of the distributional aspects of the prediction errors produced by the AR-MRNN predictors were examined, and its normality was verified in 45 of the 46 series studied. This fact suggests the possibility of producing Normal prediction errors with the non-Normal series of returns. These properties of the prediction errors supported the development of a new prediction-based portfolio selection model that uses the Normal framework.

Simulations with this new model have shown that it can produce better results with lower risks than the Markowitz's portfolio selection model – it achieved accumulated returns 5.17% higher than the Markowitz model in investment period (20 weeks) studied. Also, the predictive portfolios have shown a better market index tracking capability, achieving returns 6.38% higher than the Ibovespa market index, while the Markowitz portfolios achieved only 1.21%.

Our future works includes researching better neural prediction systems and training methods for minimizing the prediction errors, and re-balancing strategies that dynamically capture more promising re-balancing opportunities.

Acknowledgements

We are very grateful to Enfoque Informações Financeiras Ltda (http://www.enfoque.com.br) for the providing of the stock market data used in this work.

References

[1] H. M. Markowitz, "Portfolio selection," *Journal of Finance*, vol. VII, no. 1, pp. 77–91, Mar. 1952.
[2] E. J. Elton, M. J. Gruber, S. J. Brown, and W. N. Goetzmann, *Modern Portfolio Theory and Investment Analysis*, 6th ed. John Wiley & Sons, Inc., 2003.
[3] S. J. Kon, "Models of stock returns – a comparison," *Journal of Finance*, vol. XXXIX, no. 1, pp. 147–165, Mar. 1984.
[4] F. D. Freitas, A. F. De Souza, and 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, Sept. 2005, pp. 1–8.
[5] G. E. P. Box, G. M. Jenkins, and G. C. Reinsel, *Time Series Analysis: Forecasting and Control*, 3rd ed. Prentice Hall, 1994.
[6] R. S. Sutton, "Learning to predict by the method of temporal differences," *Machine Learning*, vol. 3, pp. 9–44, 1988.
[7] 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.
[8] S. Haykin, *Neural Networks: A Comprehensive Foundation*, 2nd ed. Prentice Hall, Inc., 1999.
[9] H. M. Markowitz, *Portfolio Selection: Efficient Diversification of Investments*, 2nd ed. New York: John Willey & Sons, 1991.
[10] F. Hamza and J. Janssen, "Linear approach for solving large-scale portfolio optimization problems in a lognormal market," in *IAA/AFIR Colloquium*, Nürnberg, Germany, Oct. 1996, pp. 1019–1039.
[11] W. F. Sharpe, "A simplified model for portfolio analysis," *Management Science*, no. 9, pp. 277–293, 1963.
[12] F. A. Sortino and R. van der Meer, "Downside risk–capturing what's at stake in investment situations," *Journal of Portfolio Management*, pp. 27–31, 1991.
[13] S. Lawrence, C. L. Giles, and 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.
[14] D. Maspero, "Portfolio selection for financial planners," Newfin, Universitá Bocconi, Newfin Working Paper 3/04, 2004.
[15] J. Fried, "Forecasting and probability distributions for models of portfolio selection," *Journal of Finance*, vol. XXV, no. 3, pp. 539–554, June 1970.
[16] T. Hellström, "Data snooping in the stock market," *Theory of Stochastic Processes*, vol. 5, no. 21, pp. 33–50, Oct. 1999.
[17] J. Moody and M. Saffell, "Learning to trade via direct reinforcement," *IEEE Transactions on Neural Networks*, vol. 12, no. 4, pp. 875–889, July 2001.
[18] K. Hung, Y. Cheung, and L. Xu, "An extended ASLD trading system to enhance portfolio management," *IEEE Transactions on Neural Networks*, vol. 14, no. 2, pp. 413–425, Mar. 2003.
[19] T. Hellström, "Predicting a rank measure for portfolio selection," *Theory of Stochastic Processes*, vol. 22, no. 6, pp. 64–83, 2000.
[20] J. Moody, *Prediction Risk And Architecture Selection For Neural Networks*, ser. NATO ASI Series F. Springer-Verlag, 1994.