Expert Systems with Applications xxx (2009) xxx–xxx

Contents lists available at ScienceDirect

Expert Systems with Applications

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

Predicting the Brazilian stock market through neural networks and adaptive exponential smoothing methods

E.L. de Faria [a], Marcelo P. Albuquerque [a], J.L. Gonzalez [b,*], J.T.P. Cavalcante [a], Marcio P. Albuquerque [a]

[a] Centro Brasileiro de Pesquisas Físicas, Rua Xavier Sigaud 150, URCA, Rio de Janeiro 22290-180, Brazil
[b] Pontifícia Universidade Católica, Rua Marques de São Vicente 225, Gávea, Rio de Janeiro 22453-900, Brazil

ARTICLE INFO

Keywords:
Stock index forecasting
Neural networks
Adaptive exponential smoothing

ABSTRACT

The study of financial markets has been addressed in many works during the last years. Different methods have been used in order to capture the non-linear behavior which is characteristic of these complex systems. The development of profitable strategies has been associated with the predictive character of the market movement, and special attention has been devoted to forecast the trends of financial markets. This work performs a predictive study of the principal index of the Brazilian stock market through artificial neural networks and the adaptive exponential smoothing method, respectively. The objective is to compare the forecasting performance of both methods on this market index, and in particular, to evaluate the accuracy of both methods to predict the sign of the market returns. Also the influence on the results of some parameters associated to both methods is studied. Our results show that both methods produce similar results regarding the prediction of the index returns. On the contrary, the neural networks outperform the adaptive exponential smoothing method in the forecasting of the market movement, with relative hit rates similar to the ones found in other developed markets.

1. Introduction

In the last years the financial markets around the world have been modified by the rapid development of advance systems. The acquisition of high-frequency data in real time has developed new fields like *econophysics*, renewing also the interest in the forecasting of financial and stock market indexes. Forecasting stock market assumes that financial information (or other) can modify the actual behavior of stocks, in contraposition to the theory of efficient market (Fama, 1970), which stipulates that all the relevant information is efficiently incorporated in the price before anyone can use it. Nevertheless, there are some evidences that the financial returns can be predictable through the use of public information on financial indexes, as also by using the prices trajectory (Balvers, Cosimano, & MacDonals, 1990; Lo & MacKinlay, 1988).

Artificial neural networks (ARN) and statistical tools are different methods that can be used to predict financial indexes. Neural networks incorporate a large number of parameters which allows to learn the intrinsic non-linear relationship presented in time-series, enhancing their forecasting possibilities (Haykin, 2001; Specht, 1990). ARN have been successfully applied to predict important financial and market indexes, like for example, *Standart and Pool 500* (SP&500), *Nikei 225 Index*, and others (Chen, 1994; Enke & Thawornwong, 2005; Huarng & Yu, 2006; Huang, Lai, Nakamori, Wang, & Yu, 2007; Refenes, Zapranis, & Francis, 1994; Yu & Huarng, 2008). On the other hand, statistical methods have also proved to be efficient in the study of time-series, and in particular, the exponential smoothing methods have become very popular between researchers due to their robustness (Gardner, 2006). The importance of these methods has been stressed in recent works (Gardner, 2006; Taylor, 2006), and for example, the *RiskMetrics* document recommends the use of exponential methods to estimate the conditional volatility of financial markets (RiskMetrics, 1996). The adaptive exponential smoothing (AES) model has the advantage that their parameters vary as the time-series modifies its behavior allowing to forecast abrupt variations. As a consequence, the AES method has been used to forecast market volatility (Taylor, 2004) of important stock indexes like the Nikei 225 index of the Japan market (Leung, Daouk, & Chen, 2000).

Often the ARN forecasting results are better than the ones with classical statistical methods and some studies have demonstrated it (Desai & Bharati, 1998; Enke & Thawornwong, 2005; Refenes et al., 1994). In spite of that, in some special situations, when the details of the system are well specified, simple statistical methods, like multiple linear regression for example, can outperform ARN (Bansal, Kauffman, & Weitz, 1993; Warner & Misra, 1996). It should be noted that both techniques cited above, ARN and AES respectively, are adaptive in the sense that they can modify their answers as the market time-series modifies its behavior. Other point which

* Corresponding author. Tel.: +55 21 3527 1267.
E-mail address: jl@cbpf.br (J.L. Gonzalez).

0957-4174/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2009.04.032

have been intensely debated in forecasting studies is related to the prediction of the direction of market indexes (Enke & Thawornwong, 2005; Leung et al., 2000). This has been motivated by strong evidences in which the predictions with potential profit are those that try to identify the correct sign of the market returns (O'Connor, Remus, & Griggs, 1997; Wu & Zhang, 1997). A recent study compared the forecasting possibilities of ARN and AES methods regarding their capacity to predict the sign of returns in some market indexes (Leung et al., 2000) showing different results. Moreover, there are not works that have compared both methods in stock markets more volatiles like the emerging ones, which have shown a great interdependence with developed markets, and in which, some recently studies have focalized on the level estimation of the index by using ARN (Ding, Song, & Zen, 2008; Ma & Liu, 2008; Yu & Huarng, 2008).

Taking into account all these considerations, in this article we applied the two afore mentioned forecasting models (ARN and AES respectively) to predict the principal index of the Brazilian stock market. The principal objectives of this study are: verifying to which extension both methods can predict the principal index (*Ibovespa*) of this emergent market; to compare the relative number of times that both models can corrected predict the sign of the index returns; and finally to compare the results obtained with both methods. Another point which is also addressed here is the influence of the technical configuration used in both models on the accuracy of the results obtained. To this end, the article is organized as follows: after the introduction, the Section 2 introduces both techniques used in this work; in Section 3 the results obtained are shown and discussed. Finally the conclusions are drawn.

2. Neural network architecture and adaptive exponential smoothing methods

Artificial neural networks with a multi-layer feed-forward architecture, which has produced good results in others financial prediction studies (Vellido, Lisboa, & Vaughan, 1999), were used in this work. The training algorithm was the *backpropagation* in which the weights of the neurons are modified after the error in the training process is backpropagated along the network (Rumelhart, Hinton, & Williams, 1986). The network architecture included one input layer with the number of neurons equal to the number of days in the input window, one hidden layer with several neurons, while the output layer had only one neuron corresponding to the prediction result. It should be noted that there is not a guideline to choose the number of hidden layers, the number of neurons, and other technical details (Remus & O'Connor, 245–256), thus ARN with only one hidden layer and having several neurons were used in this simulation. All the simulations were performed by using the software MATLAB.

The interval to train the ARN (called training-set) was taken as the daily closing values of the Ibovespa index from september-1998 to april-2007. The results were tested by predicting the values of the same index since april-2007 to march-2008 (test-set). At all, 2132 points were used during the training process and 236 to test the trained ARN, which corresponds to an acceptable 90:10 relation between the training and the test-set respectively. The training of the ARN was performed through a windowing technique. An input window was built which was squeezed along all the training set. Different sizes for these windows were tested, namely 3, 5, 10, 15, 20, 25, 30, 40, 50 and 60 days. For example, to train the ARN with an input window of 10 days, the first ten days from the training-set were used to train the network with the target being the eleven day of the same set. This process was repeated always shifting the windows one day ahead, until the end of the training-set. A similar approach to this one has been used in similar works (Buscema & Sacco, 2000; Refenes & Francis, 1993). To predict the results a similar technique was used, keeping always the same window that the one used during the training process. If, for example, an input window of ten days was used during the training, a similar window (ten days) was also used to forecast all the days of the test-set.

The adaptive exponential smoothing method was developed initially by Trigg and Leach (1967) and it predicts the future values of a time-series according to:

$$F_{t+1} = \alpha_t x_t + (1 - \alpha_t)F_t \quad (1)$$

In the Eq. (1) above, F_{t+1} is the predicted value at time $t+1$, while x_t and F_t represent the real and the predicted values respectively at time t before the prediction. In the same equation α is a parameter which determinates the weights of x_t and F_t and it is update at each time step. This method represents an alternative to the traditional exponential smoothing in the sense that the parameter α is update along the prediction. The updating of α requires another parameter β, which is empirically determined by the model's performance in the experiment. At each time step the parameter α is calculated following:

$$\alpha_{t+1} = \left|\frac{E_t}{M_t}\right| \quad (2)$$

Being $E_t = \beta e_t + (1-\beta)E_{t-1}$ and $M_t = \beta|e_t| + (1-\beta)M_{t-1}$. In the last expressions e_t is the error of the prediction at time t:

$$e_t = x_t - F_t \quad (3)$$

As it can be noted from Eq. (3), as the prediction performance decrease, e_t is increased and consequently α_{t+1} is also increased (see Eq. 2). That increases the relative weight of the last real value of the time-series in the next prediction by using Eq. (1).

In order to evaluate the forecasting accuracy obtained with both methods two different metrics were used. The first is the root-mean squared error (RMSE), RMSE $= \sqrt{\frac{1}{n}\sum_{k=1}^{n}(Y_k - P_k)^2}$, where Y_k represents the forecast of the real value P_k, and n is the number of predicted events. The second forecasting accuracy measure was defined as the correct tendencies number hit by the model (N_{TEND}), and it represents the number of times that the predictions followed the real tendencies of the market (Buscema & Sacco, 2000).

3. Results and discussion

Using ARN with one hidden layer, different experiments were performed following the windowing method described above in Section 2. In this work, the numbers of neurons in the hidden layer were taken as 5, 10, 15, 20 and 25 neurons respectively for each one of the different input windows. Fig. 1 shows the neural network results with the lower RMSE and the best N_{TEND} predicted. The figure evidences that different number of neurons in the hidden layer did not produce significant variations concerning the two metrics used in this study. Also the size of the windows did not modify the results, since the RMSE and N_{TEND} were almost constant with the increasing of the number of days used in the input window. Thus, the ARN parameters used in our work and specifically the size of the input windows did not modify our forecasting results.

The adaptive estimation model was implemented in this work through the MatLab software and following Eqs. (1)–(3). The first point which was verified in this study was the parameter β which determinates α in Eqs. (1)–(3). There is not a clear indication about the optimal value of this parameter, and for example, in reference 16 their values were between 0.75 and 0.95, while other studies have suggested values between 0.1 and 0.2. The first point verified

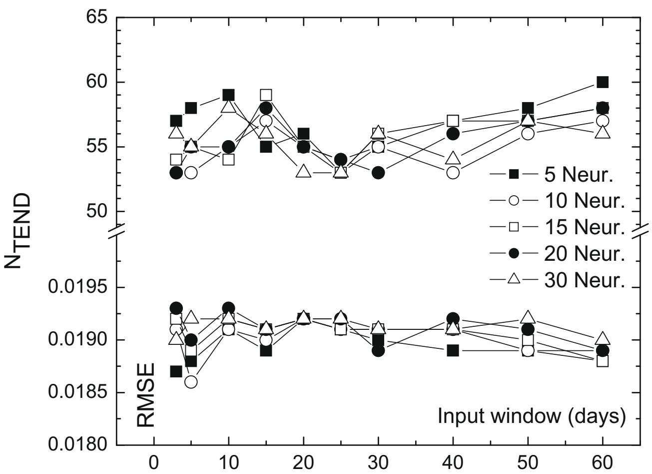

Fig. 1. Lower RMSE and maximum N_{TEND} obtained from all the ARN architecture studied.

was the influence of β in the prediction of the *Ibovespa* return, and to this end, different simulations were performed varying the values of β in the interval [0,1] with step of 0,05. The best results, according to RMSE and N_{TEND} respectively, were obtained for the same $\beta = 0.1$ where the RMSE and N_{TEND} were 0.021 and 0.51 respectively. It can be concluded that $\beta = 0.1$ produces the best results regarding the prediction of the *Ibovespa* returns. In a second approach, the same study was repeated but taking the absolute value of the index (absolute value of *Ibovespa*) as the variable to be predicted. In this case, the results were better as compared to the first case, namely, RMSE was about 0.0188 (for $\beta = 0.8$) and $N_{TEND} = 0.57(\beta = 0.1)$.

In order to compare both methods, the Table 1 shows the best ARN and exponential smoothing results according to RMSE and N_{TEND}. The notation $A \times B \times C$ for neural networks in the first column identifies A as the size of the input window, B is the number of neurons in the hidden layer and C is the output which is always one neuron. Also in the table, β represents the constant used in the updating of the smoothing constant for the AES method (see Eqs. (1)–(3)), and the asterisk in N_{TEND} represents a statistical one-sided test of $H_0 : p = 0.5$ (no predictive effectiveness) against $H_a : p > 0.5$ at a 95% level of confidence.

As shown in Table 1, for studies using neural networks the best results for RMSE (about 0.0186) and N_{TEND} (0.60) were obtained with different input windows in concordance with the conclusions obtained before from Fig. 1. It should be noted that in spite of the wide temporal interval used in the study (since 1998 to 2008) as higher as 60% of correct sign of returns were predicted through ARN in our test-set which was one year. Regarding the AES method, the predicting results obtained on the sign of the *Ibovespa* returns were not significant. The best results, N_{TEND} about 57%, were obtained by using the method to predict the absolute value of the index. This fact could be a consequence of a primary tendency presented by the *Ibovespa* index in the time period in which the model was tested. From Table 1 it is evident that predictions related to the Brazilian stock market with ARN are better than the AES ones when the direction of the market index movement is the principal goal of the forecaster. Finally, there is not a clear relation between the accuracy of the results obtained in this work and the ARN configurations used in the sense that ARN which produced the lowest RMSE values were not the same with the highest N_{TEND}.

Table 1
Best results obtained in our predicting experiments with artificial neural network and adaptive exponential smoothing (using return and absolute index), respectively.

Methods	Experiment architecture	RMSE	N_{TEND}	N_{TEND}
ARN	$5 \times 10 \times 1$	0.0186	123	0.52
ARN	$3 \times 5 \times 1$	0.0187	121	0.51
ARN	$5 \times 5 \times 1$	0.0188	116	0.49
ARN	$60 \times 5 \times 1$	0.0188	135	0.57*
ARN	$60 \times 15 \times 1$	0.0188	137	0.58*
ARN	$60 \times 15 \times 1$	0.0190	141	0.60*
ARN	$15 \times 15 \times 1$	0.0196	139	0.59*
AES (return)	$\beta = 0.1$	0.0215	121	0.51
AES (return)	$\beta = 0.15$	0.0219	113	0.48
AES (index)	$\beta = 0.8$	0.0188	125	0.53
AES (index)	$\beta = 0.1$	0.0241	135	0.57*

The highest N_{TEND} values obtained with ARN are slightly lower than the 0.63 and 0.68 values obtained for the S&P 500 in references 16 and 11. On the other hand, our values are around the same for the Nikkei 225 and significantly higher than the one found for the FTSE 100 index (Leung et al., 2000). It should be noted that all the above cited markets are well developed, and also that the latter studies focalized in the relationships between the market index and macroeconomics indicators by using neural networks. On the contrary, the AES method predicts the index returns with a lower N_{TEND} value which is almost equivalent to the one obtained on the S&P 500 index, and lower than the ones predicted in the Nikkei 225 ($N_{TEND} = 0.63$) (Leung et al., 2000). The prediction of the absolute index (*Ibovespa*) improved N_{TEND} which reached values as higher as the ones cited above. In summary, although both methods predicted the values of the *Ibovespa* returns with the same accuracy, our work evidenced that the neural networks are more efficient than the adaptive exponential technique to predict the direction of the Brazilian stock index. Technical parameters used in ARN techniques did not modify the forecasting results. Finally, the N_{TEND} results from ARN were similar to others obtained in developed markets and this opens the possibility to develop decision support strategies onto the Brazilian market.

4. Conclusions

In this work the principal index of the Brazilian stock market was studied through artificial neural networks and also by using the adaptive exponential smoothing method. It is shown that the neural networks have a superior performance to predict the correct sign of the index return. The relative number of times the correct movement of the market was predicted by neural networks was about 0.60 which is similar to other reports performed onto more developed markets by using ARN. The windowing technique used to provide the training data and technical details of ARN did not modify the forecasting results of the experience. The AES method did not contribute to predict the correct sign of the return, in spite that both methods, ARN and AES, produced almost the same RMSE in the prediction of the return values. Since the profitable strategies are related with the predictable character of the market movement our study shows the possibility to develop support decision systems for the Brazilian market based into the predictive possibilities of the neural networks.

References

Balvers, R. J., Cosimano, T. F., & MacDonals, B. (1990). Predicting stock returns in an efficient market. *Journal of Finance, 55*, 1109–1128.

Bansal, A., Kauffman, R., & Weitz, R. (1993). Comparing the modeling performance of regression and neural networks as data quality varies: A business value approach. *Journal of Management Information Systems, 10*, 11–32.

Buscema, M., & Sacco, P. (2000). Feedforward networks in financial predictions: The future that modifies the present. *Expert Systems, 17*, 149–170.

Chen, C. H. (1994). Neural networks for financial market prediction. *Proceedings of the IEEE International Conference on Neural Networks, 2*, 1199–1202.

Desai, V. S., & Bharati, R. (1998). A comparison of linear regression and neural network methods for predicting excess returns on large stocks. *Annals of Operations Research, 78*, 127–163.

Ding, Y. S., Song, X. P., & Zen, Y. M. (2008). Forecasting financial condition of Chinese listed companies based on support vector machine. *Expert Systems with Applications, 34*, 3081–3089.

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

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *Journal of Finance, 25*, 383–417.

Gardner, E. S. (2006). Exponential smoothing: The state of the art II. *International Journal of Forecasting, 22*, 637–666.

Haykin, S. (2001). *Neural networks, principles and practice*.

Huang, W., Lai, K. K., Nakamori, Y., Wang, S. Y., & Yu, L. (2007). Neural networks in finance and economics forecasting. *International Journal of Information Technology and Decision Making, 6*, 113–140.

Huarng, K., & Yu, T. H. (2006). The application of neural networks to forecast fuzzy time series. *Physica A, 363*, 481–491.

Leung, M. T., Daouk, H., & Chen, A. (2000). Forecasting stock indices: A comparison of classification and level estimation models. *International Journal of Forecasting, 16*, 173–190.

Lo, A. W., & MacKinlay, A. C. (1988). Stock market prices do not follow random walk: Evidence from a simple specification test. *Review of Financial Studies, 1*, 41–66.

Ma, J. H., & Liu, X. L. (2008). Multivariate nonlinear analysis and prediction of Shanghai stock market. *Discrete Dynamics in Nature and Society*. No. 526734.

O'Connor, M., Remus, W., & Griggs, K. (1997). Going up-going down: How good are people at forecasting trends and changes in trends? *Journal of Forecasting, 16*, 165–176.

Refenes, A. N., & Francis, G. (1993). Currency exchange rate prediction and neural network design strategies. *Neural Computing and Applications Journal, 1*, 46–58.

Refenes, A. N., Zapranis, A., & Francis, G. (1994). Stock performance modeling using neural networks: A comparative study with regression models. *Neural Networks, 7*(2), 375–388.

Remus, W., & O'Connor, M. (245–256). Neural networks for time-series forecasting. In J. S. Armstrong (Ed.), *Principles of forecasting: A handbook for researchers and practitioners* (pp. 245–256). Norwell, MA: Kluwer Academic Publishers.

RiskMetrics (1996). *RiskMetrics™ Technical Document*. J.P. Morgan/Reuters.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning internal representation by error propagation. In D. E. Rumelhart, J. L. McClelan, PDP Research Group (Eds.), *Parallel distributed processing: Explorations in the microstructure of cogbition* (pp. 318–362, Vol. I). Cambridge, MA: MIT Press.

Specht, D. (1990). Probabilistic neural networks. *Neural Networks, 3*, 109–118.

Taylor, J. W. (2004). Volatility forecasting with smooth transition exponential smoothing. *International Journal of Forecasting, 20*, 273–286.

Taylor, J. (2006). Discussion. *International Journal of Forecasting, 22*, 671.

Trigg, D. W., & Leach, A. G. (1967). Exponential smoothing with an adaptive response rate. *Operation Research Quartely, 18*, 53.

Vellido, A., Lisboa, P. J. G., & Vaughan, J. (1999). Neural networks in business: A survey of application (1992–1998). *Expert Systems with Applications, 17*, 51–70.

Warner, B., & Misra, M. (1996). Understanding neural networks as statistical tools. *American Statistician, 50*, 284–293.

Wu, Y., & Zhang, H. (1997). Forward premium as unbiased predictors of future currency depreciation: A non-parametric analysis. *Journal of International Money and Finance, 16*, 609–623.

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

Please cite this article in press as: de Faria, E. L., et al. Predicting the Brazilian stock market through neural networks and adaptive exponential smoothing methods. *Expert Systems with Applications* (2009), doi:10.1016/j.eswa.2009.04.032