European Journal of Operational Research 202 (2010) 285–293

Contents lists available at ScienceDirect

European Journal of Operational Research

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

Interfaces with Other Disciplines

Fuzzy adaptive decision-making for boundedly rational traders in speculative stock markets

Stelios D. Bekiros *,[1]

European University Institute, Via Delle Fontanelle 10, 50014 San Domenico di Fiesole (FI), Italy

ARTICLE INFO

Article history:
Received 10 December 2008
Accepted 16 April 2009
Available online 3 May 2009

Keywords:
Fuzzy sets
Bounded rationality
Technical trading
Bubbles
Direction-of-change forecasting

ABSTRACT

The development of new models that would enhance predictability for time series with dynamic time-varying, nonlinear features is a major challenge for speculators. Boundedly rational investors called "chartists" use advanced heuristics and rules-of-thumb to make profit by trading, or even hedge against potential market risks. This paper introduces a hybrid neurofuzzy system for decision-making and trading under uncertainty. The efficiency of a technical trading strategy based on the neurofuzzy model is investigated, in order to predict the direction of the market for 10 of the most prominent stock indices of U.S.A, Europe and Southeast Asia. It is demonstrated via an extensive empirical analysis that the neurofuzzy model allows technical analysts to earn significantly higher returns by providing valid information for a potential turning point on the next trading day. The total profit of the proposed neurofuzzy model, including transaction costs, is consistently superior to a recurrent neural network and a Buy & Hold strategy for all indices, particularly for the highly speculative, emerging Southeast Asian markets. Optimal prediction is based on the dynamic update and adaptive calibration of the heuristic fuzzy learning rules, which reflect the psychological and behavioral patterns of the traders.

1. Introduction

Ever since the introduction of the Efficient Markets Hypothesis, fully rational agents were considered the driving forces of markets, which in turn operated in a way to aggregate and process the beliefs and demands of traders reflecting all available information (Fama, 1970; Fama, 1991). But the empirical evidence from financial markets was not in full accordance with the Efficient Markets Hypothesis. The alternative heterogeneous agents' behavioral model is based on relaxing strict rational agent assumptions and introducing market frictions. Simon (1957) claimed that, boundedly rational agents using simple rules-of-thumb, provides a more accurate and realistic description of human behavior than perfect rationality with optimal decision rules. La Porta et al. (1997); Cenci et al. (1996) and Shiller (2002) argue that stock prices predictability reflects the psychological factors and fashions or fads of irrational investors in a speculative market. Similar results are reported in more recent studies of Madura and Richie (2004) and Sturm (2003). Overall, the study of bounded rationality and the rapid growth of the new field of behavioural economics over the past two decades led to the award of the Nobel Prize to V. Smith and D. Kahneman in 2002. The irrational market behavior has also been emphasized by Shleifer and Summers (1990) and Black (1986) in their exposition of noise traders who act on the basis of imperfect information and consequently cause prices to deviate from their equilibrium values. In general, there are two types of agents in heterogeneous agent models: "fundamentalists", who base their expectations upon dividends, earnings, growth or even macroeconomic factors, and "chartists" (noise traders and technical analysts) who instead base their trading strategies upon historical patterns and heuristics and try to extrapolate trends in future asset prices (Brock and Hommes, 1998; Hommes, 2005). The present study focuses on the latter. Specifically, the predictive return sign ability of trading rules that rely on a simple switching strategy is investigated: positive predicted returns are executed as long positions and negative returns as short positions. A similar strategy has been employed, with considerable success, by a number of other researchers (Gençay, 1998b; Gençay, 1998a; Fernández-Rodriguez et al., 2000) etc. In general terms they find that the returns from the switching strategy are higher than those from the passive one for annual returns, even when transaction costs are high. They also find that the asset return predictability is increased during volatile periods. The buy and sell signals are produced from technical trading strategies that incorporate various linear or non-linear econometric models.

* Tel.: +39 055 4685 698.
E-mail address: Stelios.Bekiros@eui.eu
[1] The author acknowledges financial support from the Max Weber Programme at the European University Institute (EUI) where he was a Fellow (2008–2009).

0377-2217/$ - see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.ejor.2009.04.015

2. Nonlinear modeling and forecasting with neural networks and fuzzy inference systems

The major challenge for "chartists" is the development of new models, or the modification of existing methods, that would enhance forecasting ability particularly for time series with dynamic time variant patterns. Conventional time series analysis, based on stationary stochastic processes does not always perform satisfactorily on economic and financial time series (Harvey, 1989). The reason is that economic data are not generally described by simple linear structural models, white noise or even random walks. The most commonly used techniques for financial forecasting are Regression methods and Autoregressive Moving Average (ARMA) models (Box and Jenkins, 1970). These methods have been used extensively in the past, but they often fail to give an accurate forecast for some series because of their nonlinear structures and some other inherent limitations. Even though ARCH/GARCH models (Bollerslev, 1986) deal with non-constant variance, still some series cannot be explained or predicted satisfactorily, due to inherent chaotic or noise patterns, fat tails, or other nonlinear components.

Extensive research in the area of nonlinear modeling has shown that neural networks enhance financial forecasting, mainly because they perform advanced mathematical and statistical processes such as nonlinear interpolation and function approximation. Neural networks are parallel computational models comprising input and output vectors as well as processing units (neurons) interconnected by adaptive connection strengths (weights), trained to store the "knowledge" of the network. Adya and Collopy (1998) demonstrated the advanced predictive ability of neural networks for time series forecasting. White (1989) and Kuan and White (1994), suggested that the relationship between neural networks and conventional statistical approaches for time series forecasting is complementary. Additionally, the function approximation properties of neural networks have been thoroughly investigated by many authors. The results in Cybenko (1989); Funahashi (1989); Hornik (1991); Hornik et al. (1989); Gallant and White (1992) and Hecht-Nielsen (1989) demonstrated that feedforward networks with sufficiently many hidden units and properly adjusted parameters can approximate any function to any desired degree of accuracy. Poddig (1993) applied a feedforward neural network to predict the exchange rates between American Dollar and Deutsche Mark, and compared results to regression analysis. Other examples using neural networks in stock and currency markets include Gençay (1998); Green and Pearson (1994); Rawani (1993); Weigend (1991); Yao et al. (1996) and Zhang (1994). However, conventional time series analysis techniques as well as neural networks incorporate in terms of input variables, only quantitative factors, such as stock returns, indices and other financial or economic magnitudes. A number of qualitative factors, e.g., macroeconomical or political effects as well as traders psychology may seriously influence the market trend, thus it is important to capture this inherent knowledge.

Fuzzy logic has been implemented initially in the area of control systems and decision theory and recently in economic applications with highly promising results. It provides a means of decision-making and learning under uncertainty. Specifically, in a fuzzy system numeric variables (inputs and outputs) are translated into fuzzy linguistic terms representing beliefs, e.g. "low" and "high". Each term is described by a membership function, which estimates the "degree" to which a variable belongs to a fuzzy set. Finally, fuzzy inference rules represented in IF–THEN statements are specified to associate the fuzzy input to the output fuzzy set. The specification of the rules could comprise an efficient mechanism of incorporating expectations and beliefs. In general, Fuzzy systems are widely applied in fields like classification, decision support, process simulation and control systems, exactly because are effective means of modeling human expert knowledge, experience, intuition, etc., (Slowinski, 1993; Slowinski and Stefanowski, 1994; Sugeno, 1988; Kosko, 1992; Klir and Yuan, 1995; Jamshidi et al., 1997; Mamdami, 1974; Mamdami, 1977). Financial and marketing applications have also been reported (Hashemi et al., 1998; Altrock, 1997). One important advantage of fuzzy inference systems is their linguistic interpretability. When implementing fuzzy systems, the focus is paid on modeling fuzziness and linguistic vagueness using membership functions. The fuzzy system approach has been applied to different forecasting problems whereby the operator's expert knowledge is used for prediction (Kaneko and Takaomi, 1996; Al-Shammari and Shaout, 1998). Although the fuzzy logic-based forecasting shows promising results, the process to construct a fuzzy logic system is subjective and depends on some ad-hoc assumptions. The learning rules derived in this way may not always yield the best forecast, and the choice of membership functions depends on trial and error. Neural networks' learning ability can be utilized to adjust and fine-tune the fuzzy membership functions. The combination of both techniques results in a hybrid neurofuzzy model which incorporates the learning ability of the neural network and the functionality of the fuzzy expert system. In a neurofuzzy system the basic concept is the derivation of various parameters of a fuzzy inference system by means of adaptive training methods obtained from neural networks (Buckley and Hayashi, 1994; Nishina and Hagiwara, 1997). Recent applications of neurofuzzy models for the prediction of financial prices and volatility can be found in the works of Pantazopoulos et al. (1998); Jalili-Kharaajoo (2004) and Cheng et al. (2007).

The present study advances the literature that has utilized separately neural networks or fuzzy logic systems in financial forecasting applications, by presenting a hybrid neurofuzzy approach that leads to superior predictions upon the *direction-of-change* of the market. The purpose of this paper is to illustrate this concretely through an investigation of the relative direction-of-change predictability of the proposed neurofuzzy trading model compared to other well-established nonlinear models. Finally, this study also provides a significant advancement of an earlier one by Bekiros and Georgoutsos (2007).

The remainder of this paper is organized as follows: Section 3 describes how the neurofuzzy model for "heuristic trading" is constructed. In Section 4 the other forecasting models used in this study are described. Finally, the empirical results are shown in Sections 5 and 6 provides concluding remarks.

3. Decision-making under uncertainty: a hybrid neurofuzzy inference model

The neurofuzzy architecture consists of the input, the rule layer and the output layer. In the input fuzzy layer all the input variables are translated into fuzzy linguistic terms. Each term is described by fuzzy membership functions. The type of membership functions is configured in this layer, whereas the parameters of these functions are processed and optimized via neural network training. Fuzzy learning, represented in IF–THEN statements, is specified to associate input and output variables of a system, which in this case is a heterogeneous financial market, while modeling psychology patterns and intuition of the agent. Consequently, the IF–THEN rules' set-up provides a very realistic model of the decision-making process under which rule-of-thumb traders operate. The rules modeled in the fuzzy rule layer consist of two parts, the "IF" part and "THEN" part. The "IF" part utilizes an "AND" association. This operator proposed by Zimmerman and Thole (1978) represents the minimum value among all the validity values of the "IF" part. The output fuzzy layer incorporates the fuzzy membership functions for outputs. Finally, in the defuzzification layer, the output is converted from fuzzy variables back into crisp values. A specific architecture where the fuzzy inference layer uses linear dependences of each rule on the system's input variables, whereby no defuzzification process is required, was introduced by Sugeno (1985). The more general 1st order Sugeno fuzzy model has rules of the form:

$$\text{IF } x_1 \text{ is } A \text{ AND } x_2 \text{ is } B \text{ THEN } z = \psi + \phi \cdot x_1 + \theta \cdot x_2 \tag{1}$$

where A and B are fuzzy sets while ϕ, θ, and ψ are all constants. Because of the linear dependence of each rule on the system's input variables the Sugeno system is suited for modeling nonlinear systems by interpolating multiple linear models.

In order to forecast the upward and downward trends of the financial market variables a two-input, two rule 1st order Sugeno model is used (Fig. 1), where the parameters ϕ, θ and ψ of the nth rule contribute via a first order polynomial $z_n = \psi_n + \phi_n \cdot x_1 + \theta_n \cdot x_2$. This model comprises two parameter sets, namely the membership function parameters and the polynomial parameters (ϕ, θ, ψ). In the proposed architecture two membership functions are used for each input corresponding to two regimes, namely "low" and "high". The hybrid training process uses a Levenberg–Marquardt neural backpropagation algorithm (Hagan and Menhaj, 1994) to optimize the membership parameters and a least squares-type algorithm to solve for the polynomial parameters. The polynomial parameters are updated first using a least squares-type algorithm and the membership parameters are then updated by backpropagating the errors. Finally, in order to solve for the neurofuzzy parameters the squared error objective function $E = \frac{1}{2}(y - y^t)^2$ is used where y^t the target output and y the system output for N size sample. The proposed model operates in five respective steps represented by $S_{l,i}$ where $l = 1, \ldots, 5$ the index of each step, i the ith node of step $S_{l,i}$ and j the number of inputs. In the first step the grades μ of the membership functions of each input j are generated as $S_{1,i} : \mu_{F_i}(x_j)$, while in the second step the rule weight coefficients are produced $S_{2,i} : \pi_i = \prod_{j=1}^{m} \mu_{F_i}(x_j)$. The third step normalizes the rule weight coefficients $S_{3,i} : \overline{\pi}_i = \frac{\pi_i}{\pi_1 + \pi_2}$. Next in the fourth step the rule outputs are calculated as follows:

$$S_{4,i} : y_i = \overline{\pi}_i \cdot z_i = \overline{\pi}_i \cdot (\phi_i \cdot x_1 + \theta_i \cdot x_2 + \psi_i) \tag{2}$$

Finally, in the fifth step all the inputs from the previous step are aggregated producing the output of the system as a piecewise linear interpolating function, dynamically calibrated by the input-dependent normalized weights:

$$\begin{aligned} S_{5,i} : y &= \sum_i y_i \\ &= \bar{\pi}_1 \cdot (\phi_1 \cdot x_1 + \theta_1 \cdot x_2 + \psi_1) + \bar{\pi}_2 \cdot (\phi_2 \cdot x_1 + \theta_2 \cdot x_2 + \psi_2) \end{aligned} \tag{3}$$

The last equation can be reformulated in the following matrix format:

$$\begin{aligned} y &= [\bar{\pi}_1 \cdot x_1 \ \bar{\pi}_1 \cdot x_2 \ \bar{\pi}_1 \ \bar{\pi}_2 \ \cdot x_1 \ \bar{\pi}_2 \cdot x_2 \ \bar{\pi}_2] \cdot [\phi_1 \ \theta_1 \ \psi_1 \ \phi_2 \ \theta_2 \ \psi_2]^T \\ &= \mathbf{X} \cdot \mathbf{\Pi} \end{aligned} \tag{4}$$

The solution for the weight vector $\mathbf{\Pi}$ to the above equation, if the $\mathbf{X}$ matrix was invertible and considering that the firing strengths are known, could be $\mathbf{\Pi} = \mathbf{X}^{-1} \cdot \mathbf{Y}$. Since this is not usually applicable, other regression methods are used such as lower triangular or more robust orthogonal decompositions. In this study Singular Value Decomposition method (SVD) (Golub and Reinsch, 1971; Golub and Van Loan, 1989; Horn and Johnson, 1991) is used. The SVD method has the advantage of using principal components to remove unimportant information related to white or heteroscedastic noise and thereby lessens the chance of overfitting. The $\mathbf{X}$ matrix is decomposed into a diagonal matrix $\mathbf{D}$ that contains the singular values, a matrix $\mathbf{U}$ of principal components, and an orthogonal normal matrix of right singular values $\mathbf{V}$. The weight matrix is finally solved for using:

$$\mathbf{\Pi} = \mathbf{V} \cdot \mathbf{D}^{-1} \cdot \mathbf{U}^{\mathbf{T}} \cdot \mathbf{Y} \tag{5}$$

For the fuzzification of the input variables, symmetric triangular membership functions are used, in order to optimize the neurofuzzy training performance (Ishibuchi et al., 1995). The triangular function contains two parameters, the a_i "peak" and the b_i "support" parameter, as follows:

$$\mu_{F_i}(x_j) = \begin{cases} 1 - \frac{|x_j - a_i|}{b_i/2}, & \text{if } |x_j - a_i| \leqslant \frac{b_i}{2} \\ 0, & \text{else} \end{cases} \tag{6}$$

The update rule of the gradient descent algorithm for the "peak" parameter is given as $a_{i,t+1} = a_{i,t} - \frac{\eta_a}{p} \cdot \frac{\partial E}{\partial a_i}$, where p the training sample size and η_α the learning rate (e.g. determines the change of the a_i values and eventually the convergence of the square error function). A similar rule applies for the "support" parameter. After chain partial derivation, the error derivatives are analyzed as $\frac{\partial E}{\partial a_i} = \frac{\partial E}{\partial y} \cdot \frac{\partial y}{\partial y_i} \cdot \frac{\partial y_i}{\partial \pi_i} \cdot \frac{\partial \pi_i}{\partial \mu_{F_i}} \cdot \frac{\partial \mu_{F_i}}{\partial a_i}$. The derived partial derivatives are

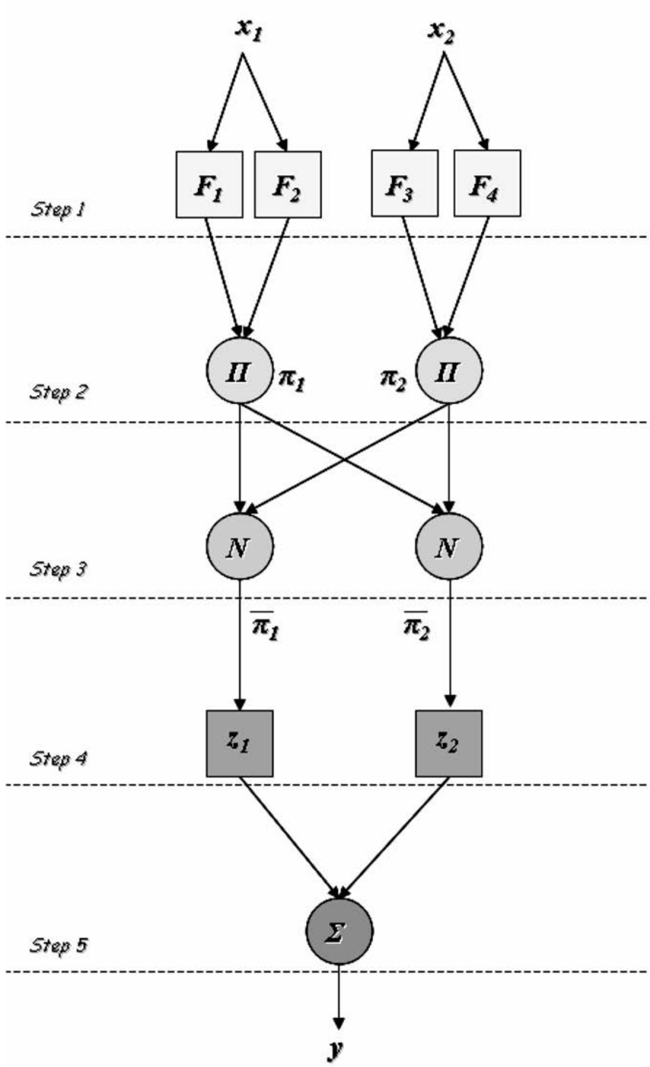

Fig. 1. Fuzzy adaptive decision-making model.

$\frac{\partial E}{\partial y} = y - y^t = e$, $\frac{\partial y}{\partial y_i} = 1$, $\frac{\partial y_i}{\partial \pi_i} = \frac{(z_i - y)}{\sum_{i=1}^n \pi_i}$ and $\frac{\partial \pi_i}{\partial \mu_{F_i}(x_j)} = \frac{\pi_i}{\mu_{F_i}(x_j)}$. The "peak" and "support" parameter partial derivatives are $\frac{\partial \mu_{F_i}}{\partial a_i} = \frac{2 \cdot sign(x_j - a_i)}{b_i}$ if $|x_j - a_i| \leqslant \frac{b_i}{2}$ or zero otherwise, and $\frac{\partial \mu_{F_i}}{\partial b_i} = \frac{1 - \mu_{F_i}(x_j)}{b_i}$. Substituting into the chain rule equation the result is $\frac{\partial E}{\partial a_i} = \frac{2e \cdot \pi_i \cdot (z_i - y) \cdot sign(x_j - a_i)}{b_i \cdot \mu_{F_i}(x_j) \cdot \sum_{i=1}^n \pi_i}$ and $\frac{\partial E}{\partial b_i} = \frac{e \cdot \pi_i \cdot (z_i - y) \cdot \left[1 - \mu_{F_i}(x_j)\right]}{b_i \cdot \mu_{F_i}(x_j) \cdot \sum_{i=1}^n \pi_i}$, where sign(*arg*) takes the value of 1 if the argument is positive and 0 otherwise. Finally, the update rule for the "peak" parameter is provided in the following recursive equation:

$$a_{i,t+1} = a_{i,t} - \frac{\eta_a}{p} \cdot \left[\frac{2e \cdot \pi_i \cdot (z_i - y) \cdot sign(x_j - a_i)}{b_i \cdot \mu_{F_i}(x_j) \cdot \sum_{i=1}^n \pi_i}\right] \quad (7)$$

whereas for the "support" parameter:

$$b_{i,t+1} = b_{i,t} - \frac{\eta_b}{p} \cdot \left[\frac{e \cdot \pi_i \cdot (z_i - y) \cdot [1 - \mu_{F_i}(x_j)]}{b}_i \cdot \mu_{F_i}(x_j) \cdot \sum_{i=1}^n \pi_i\right] \quad (8)$$

The presented model simulates in a way the adaptive decision-making process of the trader, which comprises two passes. In the "forward" pass the polynomial parameters are calculated using the SVD method, while the membership parameters remain fixed. Next, the outputs are produced using the previously calculated polynomial parameters and in the "reverse" pass the errors are backpropagated within the layers to determine the membership parameter updates.

4. A "memory-dependent neural network model

The neurofuzzy model is compared against a neural network model and a Buy & Hold strategy in order to examine its relative predictability and profitability performance. In general, a single hidden layer feedforward network with sufficiently hidden units and properly adjusted parameters can theoretically approximate any function to any desired degree of accuracy. Despite the importance of selecting the optimum number of hidden neurons, there is no explicit formula for that matter. The geometric pyramid rule proposed by Masters (1993) considers neurons for a three-layer network with n inputs and m outputs. Katz (1992) indicates that an optimal number of hidden neurons can be found between one-half to three times the number of inputs, whereas Ersoy (1990) proposes doubling the number of neurons until the network's RMSE performance deteriorates. The output of a neural network is produced via the application of a transfer function. The functionality is to modulate the output space as well as prevent outputs from reaching very large values which can "block" training. Levich and Thomas (1993) and Kao and Ma (1992) found that hyperbolic sigmoid and tan-sigmoid transfer functions are appropriate for financial markets data because they are nonlinear and continuously differentiable which are desirable properties for network learning. Learning typically occurs through training, where the training algorithm iteratively adjusts the connection weights. Common practice is to divide the sample into three distinct sets called the training, validation and testing (out-of-sample) sets; the training set is the largest and is used by the neural network to learn the patterns presented in the data, the validation set is used to evaluate the generalization ability in order to avoid overfitting and the training set should consist of the most recent observations that are processed for testing predictability. The validation error starts decreasing until the network begins to overfit the data and the error will then begins to rise. The weights are calculated at the minimum value of the validation error.

Specifically, if $\mathbf{X_t} = (x_{1,t}, \ldots, x_{p,t})$ is the input of a single layer feedforward network with q hidden units, the output is given by:

$$y_t = S\left[\beta_0 + \sum_{i=1}^q \beta_i G\left(\alpha_{i0} + \sum_{j=1}^p \alpha_{ij} x_{j,t}\right)\right] = f(\mathbf{X_t}, \mathbf{z}) \quad (9)$$

where $i = 1, \ldots, q$ and $j = 1, \ldots, p$. Consider $\mathbf{z} = (\beta_0, \ldots, \beta_q, \alpha_{11}, \ldots, \alpha_{ij}, \ldots, \alpha_{qp})^T$ as the weight vector and S, G transfer functions. The solution of the network considers estimation of the unknown vector $\mathbf{z}$ with a sample of data values. A recursive estimation methodology, which is called backpropagation is used to estimate the weight vector, as follows:

$$z_{t+1} = z_t + \eta \nabla f(\mathbf{X_t}, z_t) \cdot [y_t - f(\mathbf{X_t}, z_t)] \quad (10)$$

where $\nabla f(\mathbf{X_t}, z)$ is the gradient vector with respect to z and η the learning rate. The learning rate controls the size of the change of the weight vector on the t-th iteration. The z vector update is achieved via the minimization of the mean square error function. An alternative approach is the Bayesian updating (Foresee and Hagan, 1997), where the weights and biases of the network are assumed to be random variables with specified distributions. A major disadvantage of this method is that it generally takes longer to converge than backpropagation.

Whilst feedforward neural networks appear to have no memory since the output at any time instant depends entirely on the inputs and the weights at that instant, recurrent neural networks exhibit characteristics simulating short-term memory. In this study, Elman recurrent neural networks (Elman, 1990) have been utilized. In Elman networks with a single hidden layer the lagged outputs of the hidden neurons are fed back into the hidden neurons themselves. If $\mathbf{X_t} = (x_{1,t}, \ldots, x_{p,t})$ is the input with q hidden units and t the time index, the output of the network is given by:

$$y_t = G[\beta_0 + \sum_{i=1}^q \beta_i \cdot g_{i,t}] + \varepsilon_t \quad (11)$$

where $g_{i,t} = G\left(\alpha_{i0} + \sum_{j=1}^p \alpha_{ij} x_{j,t} + \sum_{h=1}^q \delta_{ih} g_{h,t-1}\right)$ with G the hyperbolic tangent sigmoid transfer function and $\mathbf{z} = (\beta_0, \ldots, \beta_q, \alpha_{11}, \ldots \alpha_{ij} \ldots, \alpha_{qp}, \delta_{11}, \ldots \delta_{ih} \ldots, \delta_{qq})^T$ the weight vector.

5. Empirical results

The performance of the models is examined using logarithmic returns of the most prominent indices of U.S.A, Europe and South-east Asia, covering developed and emerging markets with different capitalization and trading practices. Specifically for the United States, Standard & Poor's 500 and New York Stock Exchange index are considered while for Europe, FTSE100 (UK) and CAC40 (France). In case of Southeast Asia, KLCI Composite (Malaysia), Stock Exchange Weighted (Taiwan), HangSeng (Hong Kong), Jakarta Stock Exchange Composite (Indonesia), Straits Times (New) (Singapore) and SET 100 Basic Industries (Thailand). The sample spans between January 1, 1990 to March 2, 2001 (2915 observations). This sample contains diverse regimes and several "extreme" events including the Asian crisis and the *rise and fall of the tech-market bubble*, which makes the analysis for technical traders particularly interesting for trend forecasting. Furthermore, it provides an empirical benchmark also applicable to other turbulent periods such as the financial crisis of 2007–2009 which lead to global recession and was caused by the credit insolvency of investment institutions and high oil prices.

The daily returns are calculated as $r_t = \log(P_t) - \log(P_{t-1})$, where P_t denotes the daily index price. The predictive performance of the models is examined in the period May 5, 1997 to March 2, 2001 (1000 observations) with the use of a 1-day rolling window. To enhance robustness in the results, the out-of-sample period is segmented into three expanding sub-periods, namely PI: May 5, 1997 to April 2, 1999 (500 obs.), PII: May 5, 1997 to March 17, 2000 (750 obs.) and PIII: May 5, 1997 to March 2, 2001 (1000 obs.), the latter covering the entire backtesting period.

The inputs x_j of the neurofuzzy model correspond to the returns r_t of the previous p days while the output y is the forecasted one-day-ahead return $\hat{r}_t$. The inputs in the recurrent neural network correspond to the daily returns over the previous p days, following Gençay (1998) and Fernadez-Rodriguez et al. (2000). For the test period the models utilize a rolling window of all previous observations as a training sample and produce forecasts for each day within the corresponding period. The validation sample for each sub-period is the 30% of the training set, and is used to evaluate the generalization ability and avoid overfitting. The training set consists of the most recent observations that are processed in each sub-period. The training and validation samples utilize a moving window of all previous observations in order to produce forecasts for each day within each backtesting sub-period. The process is repeated in each of the expanding sub-periods.

The neurofuzzy model (symbolized as NF) corresponds to a specification with two lags of the returns ($j = 2$). The procedure for the selection of the lags involved the estimation of piecewise autoregressive models and the calculation of the Ljung-Box statistics for the first 10 lags of the series. Significant autocorrelations of up to the second lag of the return series were identified. Additionally, the Akaike and Schwarz Information Criteria (AIC, SIC) that were estimated for the first six lags provided the minimum value at the second lag. Sensitivity and RMSE analyses for different number of lags were conducted on all indices but the results were not found to be qualitatively different from those presented henceforth. In case of the recurrent neural network (RNN) the best forecasting ability was derived empirically by a topology which incorporated 10 neurons g in the hidden layer and an output layer with a single neuron y. This empirical result follows Katz (1992) and Ersoy (1990).

In order to account for the use of nonlinear models instead of linear, a test for the presence of nonlinear dependence in the series is conducted. To that end, the well-known BDS test statistic was used, which under the null of i.i.d. is given by (Brock et al., 1991):

$$W_{m,T}(\varepsilon) = T^{1/2}[C_{m,T}(\varepsilon) - C^{m}_{1,T}(\varepsilon)]/\sigma_{m,T}(\varepsilon) \qquad (12)$$

$C_{m,T}(\varepsilon)$ is the correlation integral from m dimensional vectors that are within a distance ε from each other, when the total sample is T and $\sigma_{m,T}(\varepsilon)$ is the standard deviation of $C_{m,T}(\varepsilon)$. Under the null hypothesis, $W_{m,T}(\varepsilon)$, has a limiting standard normal distribution. The BDS test has been applied on: (a) the original data, (b) the

Table 1
BDS test.

Index		Correlation dim.	m = 2		m = 3	
		Dim. distance	$\varepsilon = 1$	$\varepsilon = 1.5$	$\varepsilon = 1$	$\varepsilon = 1.5$
Malaysia	*KLCI Composite*	Raw data	17.91	16.03	21.99	19.65
		AFR	17.74	15.71	21.80	19.44
		NLSSR	12.82	13.79	16.55	18.03
Taiwan	*Stock Exchange Weighted*	Raw data	11.91	13.87	16.42	18.76
		AFR	12.06	13.75	16.51	18.65
		NLSSR	11.60	12.31	15.33	16.56
Hong Kong	*HangSeng*	Raw data	10.10	11.45	13.18	14.15
		AFR	10.19	11.73	13.35	14.51
		NLSSR	10.03	14.16	11.07	14.47
Indonesia	*Jakarta Stock Exchange Composite*	Raw data	23.10	21.64	27.31	24.32
		AFR	22.01	21.40	26.74	24.66
		NLSSR	16.33	13.55	17.89	15.07
Singapore	*Straits Times (New)*	Raw data	16.43	17.18	20.12	20.71
		AFR	16.52	16.94	20.86	21.22
		NLSSR	11.07	7.15	13.96	9.72
Thailand	*SET 100 Basic Industries*	Raw data	15.84	15.53	19.19	18.81
		AFR	15.55	15.39	19.29	18.91
		NLSSR	10.10	6.20	13.11	9.62
US	*SP500*	Raw data	6.06	6.60	9.60	10.18
		AFR	6.07	6.64	9.57	10.24
		NLSSR	5.20	4.35	8.16	6.67
US	*NYSE*	Raw data	6.11	6.81	9.49	9.98
		AFR	6.46	7.03	9.73	10.25
		NLSSR	6.17	5.34	8.32	6.90
UK	*FTSE100*	Raw data	5.29	6.36	8.10	9.36
		AFR	5.22	6.26	8.09	9.35
		NLSSR	5.47	6.17	8.52	7.85
France	*CAC40*	Raw data	4.23	5.98	5.22	7.22
		AFR	4.30	6.07	5.19	7.24
		NLSSR	7.53	7.31	8.19	8.83

Notes:

- Raw data = daily index returns, AFR = residuals from an autoregressive filter AR(2), NLSSR = natural logarithm of the squared standardized residuals from AR(2)–GARCH-M (1,1) model.
- m = dimension, ε = number of standard deviations of the data.
- Significance at the 1% level corresponds to the critical value 2.58.

residuals from an autoregressive filter AR(2) (based on the selected return lags), in order to ensure that the null is not rejected due to linear dependence, and (c) the natural logarithm of the squared standardized residuals from a AR(2)–GARCH-in-mean (1,1) model in order to ensure that rejection of the null is not due to conditional heteroscedasticity (De Lima, 1996).

In all cases the null of i.i.d. at the 1% marginal significance level could be rejected and the evidence seemed to suggest that a genuine non-linear dependence is present in the data (Table 1).

The trading rule works as follows; at the end of each trading day the models are being re-estimated over a rolling sample with a length equal to the training period. When the output of a model is greater than 0 this is used as a buy signal and a value less than 0 as a sell signal. The total return, when transaction costs are not considered, is estimated as:

$$R = \sum_{n+1}^{n+T+1} s \cdot r_t \tag{13}$$

where T indicates the out-of-sample horizon, r_t is the realized return and $\hat{s}_t$ is the recommended position which takes the value of (-1) for a short and (+1) for a long position (e.g. Gençay, 1998b; Jasic and Wood, 2004). In order to evaluate the forecast accuracy of the models, the percentage of correct predictions or correctly predicted signs was calculated as Sign Rate $= \frac{h}{T}$, where h is the number of correct predictions. One other comparative profitability measure is also considered: the Sharpe ratio (SR). The SR is the proportion of the mean return of the trading strategy over its standard deviation, $\mathrm{SR} = \frac{\mu_{R_T}}{\sigma_{R_T}}$. The higher the SR is the higher the return and the lower the volatility. Finally, as a measure of the predictability the Henriksson–Merton (HM) statistic (Henriksson and Merton, 1991) was employed. According to the test the number of correct forecasts has a hypergeometric distribution, asymptotically distributed as $N(0,1)$, under the null hypothesis of no market-timing ability.

The empirical results of the comparative implementation of all models are reported in Tables 2–4.

Considering total returns, a trading rule with the NF model dominates the RNN and the Buy & Hold (B&H) strategy consistently for all indices in all periods PI, PII, PII. Specifically, the total returns for the trading strategy based on the NF model ranges from the lowest 46.0% (NYSE in PI) to the highest 453.7% (SET 100 in PIII) and thus outperforms impressively the RNN and B&H strategy, the first ranging from −43.2% (SP500 in PII) to 112.4% (HangSeng in PIII) and the latter from −70.8% (SET100 in PIII) to 85.8% (CAC40 in PII). The same applies with the inclusion of transaction costs, which are estimated as 0.05% for each one-way trade, following Hsu and Kuan (2005) and Fama and Blume (1966). Again, the NF trading rule remains significantly profitable and by far better compared to that of the other models. The fact that the NF model outperforms RNN and the B&H strategy is also depicted in the proportion of correctly predicted signs, which is higher compared to the aforementioned models, although the impressive profitability of the NF model may be compromised some times with the marginal improvement of the sign rate (in some cases it is not much higher than 50%). Yet, it is due to the substantial improvement of the quantitative importance of the correctly forecasted signs. The HM test provides a further validation of the statistical significance of the sign rate in case of the NF rule with values such as 3.872 (Straits Times – Singapore in PI), 3.457 (CAC40 in PII) or even 4.525 (KLCI in PIII) at the one-sided 1% level. Instead, the RNN strategy achieves it's highest only twice in case of FTSE100 (PI and PIII) at 1% level, yet with many non-significant or even negative values. Additionally, the SR (annualized) measuring the profitability per unit of risk over the investigated market period, is much

Table 2
Out-of-sample performance of the trading models (sub-period PI: 5/5/1997–2/4/1999).

Index		Model	Total return	B&H return	RMSE	Sharpe ratio	Sign rate	HM test
Malaysia	*KLCI Composite*	**NF**	**156.0** (145.5)	−70.7	0.048	1.550	0.536	3.479*
		RNN	**44.4** (32.6)		0.038	0.443	0.470	−0.479
Taiwan	*Stock Exchange Weighted*	**NF**	**54.0** (45.6)	−13.4	0.017	1.044	0.484	1.698**
		RNN	−**40.3** (−52.2)		0.021	−0.775	0.456	−1.177
Hong Kong	*HangSeng*	**NF**	**94.8** (83.8)	−19.1	0.041	1.123	0.456	−1.804**
		RNN	**66.6** (54.3)		0.032	0.791	0.486	0.938
Indonesia	*Jakarta Stock Exchange Composite*	**NF**	**256.0** (247.0)	−50.4	0.032	3.067	0.524	3.108*
		RNN	**9.7** (−1.95)		0.034	0.111	0.478	−0.105
Singapore	*Straits Times (New)*	**NF**	**138.6** (127.7)	−18.9	0.023	2.055	0.540	3.872*
		RNN	−**10.7** (−22.7)		0.027	−0.158	0.490	0.346
Thailand	*SET 100 Basic Industries*	**NF**	**359.6** (348.6)	−47.8	0.037	3.131	0.522	2.900*
		RNN	−**24.8** (−37.7)		0.045	−0.206	0.474	−0.116
US	*SP500*	**NF**	**55.8** (47.6)	44.3	0.012	1.360	0.502	0.143
		RNN	−**7.5** (−19.4)		0.015	−0.190	0.462	−1.440***
US	*NYSE*	**NF**	**46.0** (35.3)	33.8	0.011	1.123	0.508	0.929
		RNN	−**9.4** (−21.7)		0.014	−0.269	0.448	−2.481*
UK	*FTSE100*	**NF**	**82.7** (73.4)	35.1	0.012	2.166	0.530	2.355*
		RNN	**27.6** (15.4)		0.015	0.712	0.518	2.178**
France	*CAC40*	**NF**	**80.1** (71.7)	45.9	0.016	1.629	0.542	3.150*
		RNN	**21.6** (9.8)		0.020	0.443	0.484	0.428

Notes:
- NF = neurofuzzy model. RNN = recurrent neural network. HT test = Henriksson and Merton (1981) test, asymptotically distributed as $N(0,1)$.
- In parenthesis total return after transaction costs (0.05% average fixed cost for each one-way trade).
- The sign rate measures the proportion of correctly predicted signs. The Sharpe ratio is defined as the ratio of the mean return of the strategy over its standard deviation (it has been annualized by multiplying it with the squared root of 250).

* Indicate significance at the one-sided 1% levels.
** Indicate significance at the one-sided 5% levels.
*** Indicate significance at the one-sided 10% levels.

Table 3
Out-of-sample performance of the trading models (sub-period PII: 5/5/1997–17/3/2000).

Index		Model	Total return	B&H return	RMSE	Sharpe ratio	Sign rate	HM test
Malaysia	*KLCI Composite*	**NF**	**206.7** (192.4)	15.8	0.040	1.581	0.532	3.845*
		RNN	**85.5** (68.0)		0.033	0.648	0.479	0.155
Taiwan	*Stock Exchange Weighted*	**NF**	**81.1** (68.5)	5.8	0.016	1.059	0.491	2.411*
		RNN	−**32.4** (−50.4)		0.021	−0.427	0.463	−0.663
Hong Kong	*HangSeng*	**NF**	**142.5** (125.1)	24.3	0.035	1.249	0.469	−1.039
		RNN	**47.9** (29.5)		0.029	0.411	0.476	−0.088
Indonesia	*Jakarta Stock Exchange Composite*	**NF**	**327.6** (314.0)	−9.9	0.029	2.783	0.508	2.010**
		RNN	−**11.5** (−29.3)		0.032	−0.095	0.477	−0.053
Singapore	*Straits Times (New)*	**NF**	**178** (160.9)	11.5	0.021	1.945	0.535	4.083*
		RNN	−**17.8** (−36.2)		0.025	−0.190	0.497	0.859
Thailand	*SET 100 Basic Industries*	**NF**	**439**(422.6)	−31.1	0.035	2.688	0.517	3.659*
		RNN	**2.3** (−16.8)		0.044	0.016	0.480	0.512
US	*SP500*	**NF**	**71.3** (58.6)	56.7	0.012	1.059	0.495	−0.632
		RNN	−**43.2** (−61.6)		0.012	−0.743	0.444	−3.125*
US	*NYSE*	**NF**	**56.3** (40.3)	38.3	0.011	0.964	0.501	0.759
		RNN	−**36.8** (−55.3)		0.014	−0.712	0.444	−3.226*
UK	*FTSE100*	**NF**	**86** (71.9)	38.6	0.012	1.518	0.519	1.627***
		RNN	**15.4** (−3.2)		0.015	0.269	0.499	1.154
France	*CAC40*	**NF**	**106.5** (92.0)	85.8	0.015	1.550	0.548	3.457*
		RNN	**3.6** (−14.4)		0.019	0.047	0.472	−0.784

Notation as in Table 2.

Table 4
Out-of-sample performance of the trading models (Total backtesting period PIII: 5/5/1997–2/3/2001)

Index		Model	Total return	B&H return	RMSE	Sharpe ratio	Sign rate	HM test
Malaysia	*KLCI Composite*	**NF**	**266.1** (248.7)	−44.3	0.036	1.708	0.534	4.525*
		RNN	**58.8** (35.0)		0.030	0.379	0.473	−0.002
Taiwan	*Stock Exchange Weighted*	**NF**	**96.4** (79.7)	−40.8	0.019	0.838	0.499	3.398*
		RNN	−**13.1** (−37.3)		0.023	−0.111	0.461	−1.166
Hong Kong	*HangSeng*	**NF**	**214.1** (190.4)	4.1	0.031	1.502	0.478	−0.313
		RNN	**112.4** (87.2)		0.027	0.791	0.487	1.006
Indonesia	*Jakarta Stock Exchange Composite*	**NF**	**366.5** (348.0)	−42.6	0.026	2.561	0.494	1.819**
		RNN	−**27** (−51.5)		0.029	−0.190	0.466	−0.763
Singapore	*Straits Times (New)*	**NF**	**186.8** (163.5)	2.1	0.019	1.613	0.522	3.557*
		RNN	−**11.5** (−36.5)		0.023	−0.095	0.502	1.447***
Thailand	*SET 100 Basic Industries*	**NF**	**453.7**(432.0)	−70.8	0.032	2.277	0.512	2.995*
		RNN	−**23.2** (−48.1)		0.040	−0.111	0.471	0.177
US	*SP500*	**NF**	**64.3** (47.5)	39.6	0.012	0.822	0.482	−1.072
		RNN	−**11.7** (−36.4)		0.016	−0.142	0.470	−0.996
US	*NYSE*	**NF**	**76.8** (55.2)	37.4	0.011	1.138	0.504	1.477***
		RNN	−**10.2** (−35.5)		0.013	−0.158	0.469	−1.350***
UK	*FTSE100*	**NF**	**77.7** (58.6)	27.4	0.012	1.059	0.512	1.597***
		RNN	**48.7** (23.8)		0.014	0.664	0.515	2.621*
France	*CAC40*	**NF**	**84.1** (75.8)	68.3	0.014	0.791	0.520	2.027**
		RNN	**21.9** (−2.2)		0.018	0.237	0.486	0.193

Notation as in Table 2.

higher than for the RNN in all indices and sub-periods examined. The fact that B&H strategy outperforms the RNN model in some cases is not in accordance with previous results derived by Fernández et al. (2000) as well as with the conclusions reached by Christoffersen and Diebold (2003). In that, it is noticeable in this study that for some indices and periods the nonlinear RNN model employing an "active" trading strategy compared to the "static" B&H, provides with worse even negative (loss) results, e.g. −43.2% for the SP500 in PII, or −11.5% for the SET100 in PIII. However, the B&H strategy never outperforms that of the NF model, with the latter producing remarkably higher profitability results when compared, in all periods and for all indices.

It is also worth noticing that the profitability of the NF model is significantly higher in the Asian compared to the US and European markets. These particular emerging markets as they are small and fragmented, illiquid, shallow and characterized by many extreme events provide the "ideal speculative environment" for noise traders and chartists who act on the basis of imperfect information and on behavioral and psychological patterns (Shleifer and Summers, 1990; Black, 1986). In this context, the NF model proves to be optimal in capturing the adaptive decision-making process of the specific class of market agents.

Overall, the predictive ability of the NF model is significantly higher compared to the other models. A plausible explanation is that a B&H strategy would be the best for the stock indices in the extreme case with no turning points in the testing period. However, when there are many turning points during a period and the more turning points occur, the better the NF model will be in prediction performance. The dynamic ad-hoc modification of the IF–THEN rules indicating a knowledge emergence mechanism, but most importantly the adaptive "calibration" of the membership function parameters to match the input regime, comprise the majors factors that lead to precise and prompt identification of market turning points and thus to better classification and prediction results.

6. Conclusions

This study introduces a neurofuzzy system for decision-making and trading under uncertainty. The development of new models that enhance predictability for time series with dynamic time-varying, nonlinear features is a major challenge for boundedly rational investors called "chartists", who eventually use advanced heuristics and rules-of-thumb to make profit by trading, or even hedge against potential market risks.

The present paper expands the literature that has utilized separately neural networks or fuzzy logic systems, nonlinear econometric models or Buy & Hold strategies in the evaluation of the return sign forecasting ability of trading rules by presenting a hybrid neurofuzzy model. The results suggest that with the inclusion of transaction costs, the performance of the proposed neurofuzzy model in terms of market direction-of-change for 10 US, European and Southeast Asian indices, is consistently superior to the recurrent neural network as well as a Buy & Hold strategy for all indices. The examined stock indices are Standard & Poor's 500, New York Stock Exchange, FTSE100 (UK), CAC40 (France), KLCI Composite (Malaysia), Stock Exchange Weighted (Taiwan), HangSeng (Hong Kong), Jakarta Stock Exchange Composite (Indonesia), Straits Times (New) (Singapore) and SET 100 Basic Industries (Thailand).

The neurofuzzy model produced a substantial improvement of the profitability per unit of risk over the investigated market period, as it provided valid information for a potential turning point on the next trading day. Specifically, these results seem to indicate that the neurofuzzy model has been optimally "trained" to correctly relate changes with the "sign" of the market one day ahead. The comparison between the trading models in terms of sign prediction demonstrated that the proposed model functioning as a dynamically adjusted piecewise linear interpolator compared to the static nonlinear neural predictor or the naive Buy &Hold strategy, leads to more precise identification of market turning points, while the dynamic change of the inference rules and its parameters allows for adaptive knowledge emergence and optimal regime recognition. In practice, it is because the fuzzy inference model efficiently simulates the adaptive decision-making process of the boundedly rational traders, that an investment strategy based on the proposed model allows them to earn significantly higher returns.

References

Adya, M., Collopy, F., 1998. How effective are neural networks at forecasting and prediction? A review and evaluation. Journal of Forecasting 17, 481–495.
Al-Shammari, M., Shaout, A., 1998. Fuzzy logic modeling for performance appraisal systems: a framework for empirical evaluation. Expert Systems with Applications 14, 238–323.
Altrock, C., 1997. Fuzzy Logic and Neurofuzzy Applications in Business and Finance. Prentice-Hall, Upper Saddle River, New Jersey.
Bekiros, S., Georgoutsos, D., 2007. Evaluating direction-of-change forecasting: neurofuzzy models vs. neural networks. Mathematical and Computer Modelling 46, 38–46.
Black, F., 1986. Noise. Journal of Finance 41, 529–543.
Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics 31, 307–327.
Box, G., Jenkins, G., 1970. Time Series Analysis: Forecasting and Control. Holden-day Inc., San Francisco, CA.
Brock, W.A., Hommes, C.H., 1998. Heterogeneous beliefs and routes to chaos in a simple asset pricing model. Journal of Economic Dynamics and Control, 22.
Brock, W., Hsieh, D., LeBaron, B., 1991. Nonlinear Dynamics, Chaos and Instability. The MIT Press.
Buckley, J.J., Hayashi, Y., 1994. Fuzzy neural networks: a survey. Fuzzy Sets and Systems 66, 1–13.
Cenci, M., Cerquetti, A., Peccati, L., 1996. Imitation and stability in a stock market. European Journal of Operational Research 91 (2), 301–305.
Cheng, P., Quek, C., Mah, M.L., 2007. Predicting the impact of anticipatory action on US stock market – an event study using ANFIS (a neural fuzzy model). Computational Intelligence 23 (2), 117–141.
Christoffersen, P., Diebold, F., 2003. Financial Asset Returns, Direction-of-change Forecasting, and Volatility Dynamics, Working Paper 10009, NBER.
Cybenko, G., 1989. Approximation by superposition of a sigmoidal function. Mathematics of Control, Signals and Systems 2, 303–314.
De Lima, P.J.F., 1996. Nuisance parameter free properties of correlation integral based statistics. Econometric Reviews 15, 237–259.
Elman, J.L., 1990. Finding structure in time. Cognitive Science 14, 179–211.
Ersoy, O., 1990. Tutorial at Hawaii International Conference on Systems Sciences.
Fama, E.F., 1970. Efficient capital markets: a review of empirical work. Journal of Finance 25, 383–417.
Fama, E.F., 1991. Efficient capital markets II. Journal of Finance 46 (5), 1575–1617.
Fama, E.F., Blume, M.E., 1966. Filter rules and stock-market trading. Journal of Business 39, 226–241.
Fernández-Rodriguez, F., Gonzalez-Martel, C., Sosvilla-Rivero, S., 2000. On the profitability of technical trading rules based on artificial neural networks: evidence from the Madrid stock market. Economic Letters 69, 89–94.
Foresee, F.D., Hagan, M.T., 1997. Gauss–Newton approximation to Bayesian learning. In: Proceedings of IEEE International Conference on Neural Networks, vol. 3, 1997, pp. 1930–1935.
Funahashi, K., 1989. On the approximate realization of continuous mappings by neural networks. Neural Networks 2, 183–192.
Gallant, A.R., White, H., 1992. On learning the derivatives of an unknown mapping with multiplayer feedforward networks. Neural Networks 5, 129–138.
Gençay, R., 1998a. The predictability of security returns with simple technical trading rules. Journal of Empirical Finance 5, 347–359.
Gençay, R., 1998b. Optimization of technical trading strategies and the profitability in security markets. Economics Letters 59, 249–254.
Golub, G.H., Reinsch, C., 1971. Singular Value Decomposition and Least Squares Solutions: Handbook for Automatic Computation. Linear Algebra, vol. II. Springer, Heidelberg.
Golub, G.H., Van Loan, Charles F., 1989. Matrix Computations, second ed. The Johns Hopkins University Press.
Green, H., Pearson, M., 1994. Neural nets for foreign exchange trading. In: Trading on the Edge: Neural, Genetic, and Fuzzy Systems for Chaotic Financial Markets. Wiley, New York.
Harvey, A.C., 1989. Forecasting Structural Time Series Models and the Kalman Filter. Cambridge University Press, Cambridge.
Hagan, M.T., Menhaj, M., 1994. Training feedforward networks with the Marquardt algorithm. IEEE Transactions on Neural Networks 5 (6), 989–993.
Hashemi, R., Le Blanc, L., Rucks, C., Rajaratnam, A., 1998. A hybrid intelligent system for predicting bank holding structures. European Journal of Operational Research 109, 390–402.
Hecht-Nielsen, R., 1989. Theory of the backpropagation neural networks. In: Proceedings of the International Joint Conference on Neural Networks I, Washington, DC, IEEE Press, New York, pp. 593–605.
Henriksson, R.D., Merton, R.C., 1981. On the market timing and investment performance II: statistical procedures for evaluating forecasting skills. Journal of Business 54, 513–533.
Hommes, C.H., 2005. Heterogeneous agent models in economics and finance. In: Tesfatsion, L., Judd, K.L. (Eds.), Handbook of Computational Economics, Agent-Based Computational Economics, vol. 2. Elsevier Science B.V.
Horn, R.A., Johnson, C.R., 1991. Topics in Matrix Analysis. Cambridge University Press.
Hornik, K., 1991. Approximation capabilities of multilayer feedforward networks. Neural Networks 4, 251–257.
Hornik, K., Stinchcombe, M., White, H., 1989. Multilayer feedforward networks are universal approximators. Neural Networks 2, 359–366.
Hsu, P.-H., Kuan, C.-M., 2005. Reexamining the profitability of technical analysis with data snooping checks. Journal of Financial Econometrics 3, 606–628.
Ishibuchi, H., Kwon, K., Tanaka, H., 1995. A learning algorithm of fuzzy neural networks with triangular fuzzy weights. Fuzzy Sets and Systems 71, 277–293.
Jalili-Kharaajoo, M., 2004. Stock trend prediction using neurofuzzy predictors based on brain emotional learning algorithm. Lecture Notes in Computer Science 3070, 308–313.
Jamshidi, M., Titli, A., Zadeh, L., Boverie, S. (Eds.), 1997. Applications of Fuzzy Logic. Prentice-Hall, New Jersey.
Jasic, T., Wood, D., 2004. The profitability of daily stock market indices trades based on neural network predictions: case study for the S&P 500, the DAX, the TOPIX, and the FTSE in the period 1965–1999. Applied Financial Economics 14, 285–297.

Kaneko, Takaomi, 1996. Building a financial diagnosis system based on fuzzy logic production system. Computers Industrial Engineering 31 (3/4), 743–746.

Kao, G.W., Ma, C.K., 1992. Memories, heteroscedasticity and prices limit in currency futures markets. Journal of Futures Markets 12, 672–692.

Katz, J.O., 1992. Developing neural network forecasters for trading. Technical Analysis of Stocks and Commodities, 58–70.

Klir, G.J., Yuan, B., 1995. Fuzzy Sets and Fuzzy Logic: Theory and Applications. Prentice-Hall, Upper Saddle River.

Kosko, B., 1992. Neural Networks and Fuzzy Systems. Prentice-Hall, Englewood Cliffs.

Kuan, C.-M., White, H., 1994. Artificial neural networks: an econometric perspective. Econometric Reviews 13, 1–91.

La Porta, R., Lakonishok, J., Shliefer, A., Vishny, R., 1997. Good news for value stocks: further evidence on market efficiency. Journal of Finance 52, 859–874.

Levich, R.M., Thomas L.R., 1993. The significance of technical trading rule profits in the foreign exchange market: a bootstrap approach. In: Strategic Currency Investing – Trading and Hedging in the Foreign Exchange Market, Probus, Chicago, pp. 336–365.

Mamdami, E., 1977. Application of fuzzy logic to approximate reasoning using linguistic systems. IEEE Transactions on Computers 26 (12), 1182–1191.

Mamdami, E., 1974. Applications of fuzzy algorithms for control of simple dynamic plant. Proceedings IEE 121, 1585–1588.

Masters, T., 1993. Advanced Algorithms for Neural Networks. John Wiley.

Nishina, T., Hagiwara, M., 1997. Fuzzy inference neural network. Neurocomputing 14, 223–239.

Pantazopoulos, K.N., Tsoukalas, L.H., Bourbakis, N.G., Brun, M.J., Houstis, E.N., 1998. Financial prediction and trading strategies using neurofuzzy approaches. IEEE Transactions on Systems, Man, and Cybernetics – Part B 28 (4), 520–531.

Poddig, A., 1993. Short term forecasting of the USD/DM exchange rate. In: Proceedings of First International Workshop on Neural Networks in Capital Markets, London.

Rawani, A., 1993. Forecasting and trading strategy for foreign exchange market. Information Decision Technology 1, 19.

Shiller, R.J., 2002. From Efficient Market Theory to Behavioral Finance. Cowles Foundation Discussion Paper No. 1385.

Shleifer, A., Summers, L.H., 1990. The noise trader approach to finance. Journal of Economic Perspectives 4 (2), 19–33.

Simon, H.A., 1957. Models of Man. Wiley, New York, NY.

Slowinski, R., 1993. Rough set learning of preferential attitude in multi-criteria decision making. In: Komorowski, J., Ras, Z.W. (Eds.), Methodologies for Intelligent System, Lecture Notes in Artificial Intelligence, vol. 68. Springer, Berlin, pp. 42–651.

Slowinski, R., Stefanowski, J., 1994. Rough classification with valued closeness relation. In: Diday, E. et al. (Eds.), New Approaches in Classification and Data Analysis. Springer, Berlin, pp. 482–488.

Sugeno, M., 1985. Industrial Applications of Fuzzy Control. Elsevier Science Publications Co.

Sugeno, M., 1988. Fuzzy Control. Nikkan Kougyou, Shinbunsha.

Weigend, A.S., 1991. Generalization by weight-elimination applied to currency exchange rate prediction. In: Proceedings of the IEEE International Joint Conference on Neural Networks, Singapore.

White, H., 1989. Learning in artificial neural networks: a statistical perspective. Neural Computing 1, 425–464.

Yao, J.T., Poh, H.-L., Jasic, T., 1996. Foreign exchange rates forecasting with neural networks. In: Proceedings of the International Conference on Neural Information Processing, Hong Kong, pp. 754–759.

Zhang, X.R., 1994. Non-linear predictive models for intra-day foreign exchange trading. International Journal of Intelligent Systems, Accounting and Finance Management 3 (4), 293–302.

Zimmerman, H.-J., Thole, U., 1978. On the suitability of minimum and product operators for the intersection of fuzzy sets. Fuzzy Sets and Systems 2, 173–186.