Available online at www.sciencedirect.com
ScienceDirect
Expert Systems with Applications 33 (2007) 816–823
Expert Systems with Applications
www.elsevier.com/locate/eswa

Computational intelligence approaches and linear models in case studies of forecasting exchange rates

André Alves Portela Santos [a], Newton Carneiro Affonso da Costa Jr. [a], Leandro dos Santos Coelho [b,*]

[a] Federal University of Santa Catarina, UFSC, Graduate Program in Economics, Box 476, 88040-900 Florianópolis, SC, Brazil
[b] Pontifical Catholic University of Parana, PUCPR/CCET/PPGEPS, Automation and System Laboratory, Imaculada Conceição, 1155, 80215-901 Curitiba, PR, Brazil

Abstract

Artificial neural networks and fuzzy systems, have gradually established themselves as a popular tool in approximating complicated nonlinear systems and time series forecasting. This paper investigates the hypothesis that the nonlinear mathematical models of multilayer perceptron and radial basis function neural networks and the Takagi–Sugeno (TS) fuzzy system are able to provide a more accurate out-of-sample forecast than the traditional auto regressive moving average (ARMA) and ARMA generalized auto regressive conditional heteroskedasticity (ARMA-GARCH) linear models. Using series of Brazilian exchange rate (R$/US$) returns with 15 min, 60 min and 120 min, daily and weekly basis, the one-step-ahead forecast performance is compared. Results indicate that forecast performance is strongly related to the series' frequency and the forecasting evaluation shows that nonlinear models perform better than their linear counterparts. In the trade strategy based on forecasts, nonlinear models achieve higher returns when compared to a buy-and-hold strategy and to the linear models.

Keywords: Neural networks; Fuzzy systems; Forecasting; Linear models; Exchange rates

1. Introduction

The literature related to financial time series has registered since the 1990's important advances with the incorporation of newly developed methods that attempt to determine patterns of relationships in financial market data. These approaches are, in general, computationally intensive and characterized by the capacity of modeling nonlinear dynamic systems, i.e., systems in which the variables of the environment possess complex patterns of interrelationships that alter throughout time.

Given the growing report of presence of nonlinear structures in financial time series, the use of deterministic linear models to describe and forecast financial prices movements has been criticized (Aguirre & Aguirre, 2000). One can argue that the traditional econometric approach based on models with simple specifications and constant parameters, such as Box-Jenkins ARIMA models, are unable to reply the dynamics inherent to economic and financial series (Timmermann & Granger, 2004). This aspect has stimulated researchers of diverse academic backgrounds to apply modern techniques of system identification to various problems in Finance, as asset pricing (Morelli et al., 2004), investment selection (Eakins & Stansell, 2003), game theory (Barr & Saraceno, 2005) and time series forecasting (Kaashoek & Van Dijk, 2002; Malhotra & Malhotra, 2002; Ocal, 2000; Pérez-Rodríguez, Torra, & Andrada-Félix, 2005; Tsaih, Hsu, & Lai, 1998). These approaches are suitable for dealing with intrinsic characteristics of financial time series, as the

* Corresponding author. Tel.: +55 41 3271 13 33; fax: +55 41 3271 13 45.
E-mail addresses: andreportela@gmail.com (A. Alves Portela Santos), newton@cse.ufsc.br (N. Carneiro Affonso da Costa Jr.), leandro.coelho@pucpr.br (L. dos Santos Coelho).

0957-4174/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2006.07.008

presence of multiple regimes in data generator process, ARCH effects and nonlinearities. However, the debate concerning the gains coming from the use of nonlinear models did not reach a consensus yet (Clements, Franses, & Swanson, 2004), stimulating further research in areas as nonlinear model selection, estimation and evaluation approaches.

Two relevant approaches, among many others, used for forecasting financial series are Takagi–Sugeno (TS) fuzzy systems and multilayer perceptron neural networks methods. Takagi–Sugeno fuzzy model, for instance, exhibits both high nonlinearity and simple structure (Sugeno & Kang, 1988; Takagi & Sugeno, 1985). The identification problem in T–S modeling consists of two major parts: structure identification and parameter identification. The structure identification is related to both the determination of the premise part and the consequent part of the production rules. It consists of determining the premise space partition and extracting the number of rules and determining the structure of the output elements (equations), respectively. Finally, the parameter-learning task consists of determining the system parameters, i.e., membership functions, so that a performance measure based on the output errors is minimized. Neural networks are originally inspired by biologic neural networks' functionality that may learn complex functional relations through a limited number of training data. Neural networks may serve as black-box models of nonlinear multivariable dynamic systems and may be trained using input-output data, observed from the system (Mcloone, Brown, Irwin, & Lightbody, 1998; Narendra & Parthasarathy, 1990). The usual neural network consists of multiple simple processing elements, called neurons, interconnections among them and the weights attributed to the interconnections. The relevant information of such methodology is stored in the weights.

In recent years, researchers (León, Liern, & Vercher, 2002; Shapiro, 2002; Tseng & Tzeng, 2002) have proposed a varied spectrum of methodologies for identification and nonlinear forecasting based upon fuzzy systems to deal with nonlinear systems. On the other hand, neural networks have received attention in recent years due to their abilities to perform learning, thus applied in a great number of situations in finance and business applications (Medeiros, Veiga, & Pedreira, 2001; Tino, Schittenkopt, & Dorffner, 2001; Wong & Selvi, 1998).

This paper investigates the hypothesis that the nonlinear mathematical models of multilayer perceptron and radial basis function neural networks and the Takagi–Sugeno (TS) fuzzy system are able to provide a more accurate out-of-sample forecast than the traditional auto regressive moving average (ARMA) and ARMA generalized auto regressive conditional heteroskedasticity (ARMA-GARCH) linear models.

The paper is organized as follows. The Section 2 describes the linear and nonlinear models. In Sections 3 and 4 we present the methodology of the study and the results, respectively. And finally, in Section 5, we bring concluding remarks and future works.

2. Fundamentals of forecast models

2.1. Linear models

Following the Box-Jenkins methodology, an auto-regressive moving average process ARMA (p, q) can be written as

$$y_t = \phi_0 + \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q} \tag{1}$$

where $\{e_t\}$ is a white noise process with zero mean and a constant variance of σ^2. The terms $\phi_0, \phi_1, \ldots, \phi_p$ and $\theta_1, \ldots \theta_q$ are parameters that must be estimated. Also considering the stylized fact of non-constant residual variance, we use ARMA models with an auto-regressive conditional heteroskedasticity process – ARMA(p,q)-GARCH(p.q) models (Mcloone et al., 1998). From the Eq. (1), we consider the generator process of $\{e_t\}$ as given by

$$\varepsilon_t = v_t \sqrt{h_t}, \quad v_t \sim \mathrm{IID}(0, 1) \tag{2}$$

$$h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i h_{t-1} \tag{3}$$

2.2. Nonlinear models

2.2.1. Multilayer perceptron neural network

The neural networks provide, usually, not-parametric quantitative knowledge and are suitable for identification, learning and adaptation of complex systems.

The multilayer perceptron is a generalization of a perceptron of a layer. The multilayer perceptron neural network (MLP-NN) contains three types of layers: the input layer, hidden layer (s) and the output layer. Any neuron of a layer is able to establish connection with another neuron of the following layer. The neurons of the input layer receive the signals from the external world and transmit them for the neurons of the next layer. While this, the neurons of the output layer send the information of the hidden layer (s) neurons for the external world, as presented in Fig. 1.

The output of a single layer feedforward neural network model of multilayer perceptron type with n_1 hidden units (and a linear component) can be defined as

$$y_{t+h}^h = w_0' Z_t + \sum_{i=1}^{n} w_{1i}' g(Z_t) + \varepsilon_{t+h} \tag{4}$$

where $g(x) = 1/(1 + \mathrm{e}^x)$ and $Z_t = (1, y_t, y_{t-1}, \ldots, y_{t-p-1})$.

The training process of the MLP-NN is supervised and usually realized with the back-propagation algorithm (BP). BP algorithm is based on the error correction learning rule, and can be seen as a generalization of the adaptive filtering algorithm or a special case of the least squares algorithm.

The BP algorithm is an iterative method based on gradient that seeks to minimize the addition of the quadratic error between the current output and the desired output.