EVOLVING FUNCTIONAL FUZZY MODEL FOR INTEREST RATE TERM STRUCTURE FORECASTING

Leandro Maciel, Fernando Gomide

School of Electrical and Computer Engineering - Department of Computer Engineering and Industrial Automation

University of Campinas

{maciel,gomide}@dca.fee.unicamp.br

Rosangela Ballini

Institute of Economics - Department of Economic Theory

University of Campinas

ballini@eco.unicamp.br

Abstract – Evolving fuzzy systems use data streams to continuously adapt the structure and functionality of fuzzy rule-based models. It gradually develops the model structure and its parameters from a stream of data, which is essential when dealing with complex amd nonstationary systems. In this paper, we suggest the use of functional evolving fuzzy modeling in the form of Takagi-Sugeno (eTS) model to forecast Brazilian government bond yields through the Nelson-Siegel function. In this case the eTS adaptively estimates the parameters Nelson-Siegel function to perform forecasts. This is a crucial procedure for bond portfolio management, derivatives and bonds pricing. The experiments reported here use daily data of the Brazilian National Treasury Bills of the period from January 2007 to December 2009 for one, three, six, nine and twelve months ahead forecasting horizons. The evolving model was compared with autoregressive and random walk models, in terms of root mean squared error. Results indicate that eTS is a promising approach to deal with government bond yields forecasting because it gives more accurate Nelson-Siegel parameters values than traditional approaches.

Keywords – Evolving Fuzzy Systems, Time Series Forecasting, Yield Curve, Interest Rate.

1. INTRODUCTION

Term structure of government bond yields modeling is a challenging task that has attracted investors, policy makers, researches and all market participants since it provides fixed-income instruments pricing and manage the risk of bonds and derivatives. It also allows monitoring observed and unobserved economic variables such as the risk premium, default risk, inflation and real activity, as well as forecasting future interest rates [1].

Yield curve forecasting is crucial for risk and bond portfolio management and for derivatives pricing. It is also used by treasuries that manage the emission and maintain of the stock of public debit, which continuously demands an assessment of current and future interest rates. Moreover, investors track the performance of their portfolios against the opportunity cost of investing in low-risk bonds and Central banks react to expected inflation and economic activity by adjusting the short rate, thereby affecting the whole curve [2, 3].

The literature related to finance, econometrics and macroeconomics have been focused term structure models, and few of them have analyzed the out-of-sample forecasting performance. Reference [4] investigates the relationship between forward and future spot rates as the first study devoted to yield curve predictability questions. More than one decade later, [5] applied affine models for US yields forecasts and observed poor results. Recently, [2] addressed a two-stage model based on Nelson-Siegel framework [6] to forecast the US term structure and better results were obtained when compared with competing models. Further, [7] showed that the inclusion of no-arbitrage conditions in latent models improves the out-of-sample fit.

In [8] the authors introduced a new model, namely Functional Signal Plus Noise with an Equilibrium Correction Model (FSN-ECM), which produces very good forecasts, in terms of mean squared forecast error, for one-month ahead horizon. The work reported in [9] compares the FSN-ECM model with the model developed by [2] to forecast 12-dimensional yields for Brazilian yields for one, three, six, and twelve months horizons. The results suggest that the FSN-ECM produces very good forecasts for short-term (one month), outperforming the benchmarks available.

Using a broad class of linear models, [1] showed that a simple parametric specification has the best predictive power for interest rate forecasting in USA and Brazilian markets. However, the random walk model is not outperformed and it was found that macroeconomic variables and no-arbitrage conditions have little effect to improve the out-of-sample fit, while a financial variable (stock index) increases the forecasting accuracy.

Although the models mentioned above have acceptable performance for term structure forecasting, they differ in accuracy for short and long horizons. Auto-regressive and random walk models are well-suited for short maturities while affine models and extensions perform better for medium and long horizons forecasting. Due to constraints, traditional statistical and econo-

metrics models have been outperformed by methodologies based on computational intelligence like artificial neural networks, evolutionary computing and fuzzy systems.

The purpose of this paper is to introduce the use of evolving fuzzy rule-based models to forecast the Brazilian yield curve by estimating the Nelson-Siegel function parameters. The concept of evolving fuzzy systems translates in the idea of gradual self-organization and parameter learning in fuzzy rule-based models [10]. Evolving fuzzy systems use data streams to continuously adapt the structure and functionality of fuzzy rule-based models. The evolving mechanism ensures greater generality of the structural changes because rules are able to describe a number of data samples. Evolving fuzzy rule-based models include mechanisms for rule modification to replace a less informative rule by a more informative one [10]. Overall, evolution guarantees gradual change of the rule base structure inheriting structural information. The idea of parameter adaptation of rules antecedent and consequent is similar in the framework of evolving connectionists systems [11], evolving Takagi-Sugeno (eTS) and extended Takagi-Sugeno (xTS) models, and their variations [10, 12, 13]. In particular, the eTS model is a functional fuzzy model in the Takagi-Sugeno (TS) form whose rule base and parameters continually evolve by adding new rules with higher summarization power and modifying existing rules and parameters to match current knowledge.

In the literature related to finance and economics there are a few and recent applications of evolving fuzzy rule-based models. For instance, [14] uses different evolving fuzzy structures to estimate the Value-at-Risk (VaR) and compare its accuracy against GARCH models using the São Paulo Exchange data. They showed that evolving fuzzy modeling outperforms traditional benchmarks for VaR estimation according to the number of failures. Similar results were found for sovereign bonds modeling [15]. Financial time series forecasting was addressed by [16] using an eTS model with memory for modeling and prediction of GBP/EUR closing price data and US Gross Domestic Product Data. The authors conclude that the predictive power of eTS with memory is higher, and its benefits can be appropriately exploited. Recently, [17] suggested an evolving fuzzy systems modeling approach for fixed income option pricing. Results, based on error measures and statistical tests, reveal that evolving fuzzy models outperform traditional methods based on Black-Scholes closed-form formula and alternative neural network approaches.

The main goal of this paper is to predict the Nelson-Siegel function parameters using an evolving Takagi-Sugeno model. To compare eTS model accuracy two benchmarks re considered , auto-regressive and random walk models, respectively. Actual daily data of the Brazilian Treasure Bonds is used to forecast the yield curve for different horizons. The out-of-sample forecasting performance is evaluated in terms of root mean squared error.

After this introduction, this paper proceeds as follows. Section 2 presents briefly the idea of evolving fuzzy rule-based modeling and the evolving Takagi-Sugeno method. Next, Section 3 describes the concepts of modeling and forecasting yield curves. Section 4 compares the performance of eTS model against auto-regressive and random walk models to forecasting the Brazilian yield curves for different horizons ahead. Finally, Section 5 concludes the paper, summarizing its contributions and suggesting issues for further investigation.

2. EVOLVING TAKAGI-SUGENO FUZZY SYSTEMS

In evolving systems, a key question is how modify the current model structure using the newest data sample which ensures greater generality of the structural changes. Evolving systems use incoming information to continuously develop their structure and functionality through online self-organization.

Fuzzy rule-based models whose rules are endowed with local models forming their consequents are commonly referred to as fuzzy functional models. The Takagi-Sugeno (TS) is a typical example of a fuzzy functional model. A particularly important case is when the rule consequents are linear functions. The evolving Takagi-Sugeno (eTS) model and its variations [12] assume rule-based models whose fuzzy rules are as follows:

$$\mathcal{R}_i : \textbf{IF } (x_1 \text{ is } \Gamma_{i1}) \textbf{ AND ... AND } (x_n \text{ is } \Gamma_{in}) \textbf{ THEN } y_i = a_{i0} + a_{i1}x_1 + \cdots + a_{in}x_n; \quad i = 1, \ldots, R$$

where $\mathcal{R}_i$ denotes the ith fuzzy rule, R is the number of fuzzy rules, $\boldsymbol{x}$ is the input vector, $\boldsymbol{x} = [x_1, x_2, \ldots, x_n]^T$, Γ_{ij} denotes the antecedent fuzzy sets, $j = 1, \ldots, n$, y_i is the output of the ith rule, a_{il} are its parameters, $l = 0, \ldots, n$.

The degree of firing of each rule is proportional to the level of contribution of the corresponding linear model to the overall output, for Gaussian antecedent fuzzy sets:

$$\Gamma_{ij}(x_j) = e^{-\alpha||x_j - x_{ij}^*||^2}; \quad i = 1, \ldots, R \text{ and } j = 1, \ldots, n \tag{1}$$

where $\alpha = 4/\varphi^2$ and φ is a positive constant, which defines the spread of the antecedent and the zone of influence of the ith model[1], and x_{ij}^* is the focal point of the ith rule antecedent.

The conjunction of respective fuzzy sets related to a rule defines its firing level as a Cartesian product:

$$\tau_i = \Gamma_{i1}(x_1) \times \Gamma_{i2}(x_2) \times \cdots \times \Gamma_{in}(x_n) = \bigcap_{j=1}^{n} \Gamma_{ij}(x_j) \tag{2}$$

The TS output is obtained by weighted averaging of individual rules' contributions:

$$y = \sum_{i=1}^{R} \lambda_i y_i = \sum_{i=1}^{R} \lambda_i x_e^T A_i \tag{3}$$

---

[1] According to [12], large values of φ leads to averaging, and too small values to over-fitting also, values of $\varphi \in [0.3; 0.5]$ can be recommended.

where $\lambda_i = \left(\tau_i / \sum_{j=1}^{R} \tau_j \right)$, which is the normalized firing level of the ith rule, $A_i = \begin{bmatrix} a_{i0} & a_{i1} & \ldots & a_{in} \end{bmatrix}^T$ is the vector of parameters of the ith linear model, and $x_e = \begin{bmatrix} 1 & x^T \end{bmatrix}^T$ is the expanded data vector.

The first subtask could be solved by subtractive clustering [18], which is an improved version of the so-called mountain clustering approach [19].

The clustering procedure starts with the first data point established as the focal point of the first cluster. Its coordinates are used to form the antecedent part of the fuzzy rule using for example Gaussian membership functions (1). Its potential is assumed equal to 1.

Starting from the next data point onwards the potential of the new data points is calculated recursively using Cauchy function of first order to measure the potential:

$$P_k\left(z_k\right) = \frac{1}{1 + \frac{1}{(k-1)}\sum_{t=1}^{k-1}\sum_{j=1}^{n+1}\left(d_{lj}^j\right)^2}, \quad k = 2, 3, \ldots \tag{4}$$

where $P_k(z_k)$ is the potential of the data point z_k calculated in k (where z is the augmented data vector $z^T = [x^T; y]$), $d_{lk}^j = z_l^j - z_k^j$ is the projection of the distance between two data points (z_l^j e z_k^j) on the axis z^j (x^j for $j = 1, 2, \ldots, n$ and on the axis y for $j = n + 1$).

When the new data is available, they influence the potential of the centers of the clusters, which are respective to the focal points of existing rules. The potential depends on the distance to all data points, including the new ones. If the potential of a new data is higher than the potential of the current cluster centers, then the new data becomes a new center and a new rule is created. If the potential of a new data is higher than the potential of the current centers, but it is close to an existing center, then the new data replaces the existing center. Using the potential instead of the distance to a certain rule center only for forming the rule-base result in rules that are more informative and more compact rule-base. This mechanism ensures an evolving rule-base by dynamically upgrading and modifying it while inheriting the bulk of the rules [12]. See [10] and [12] for more details.

Consequent parameters estimation in eTS models can be transformed into a least squares problem. Rewriting the equation (3) in an equivalent vector expression of y we have:

$$y = \psi^T \theta \tag{5}$$

where $\theta = \begin{bmatrix} A_1^T & A_2^T & \ldots & A_R^T \end{bmatrix}^T$ is a vector composed of the linear model parameters; $\psi = \begin{bmatrix} \lambda_1 x_e^T & \lambda_2 x_e^T & \ldots & \lambda_R x_e^T \end{bmatrix}^T$ is a vector of the inputs that are weighted by the normalized firing levels of the rules [12].

Let $\left(x_k^T, y_k\right)$, $k = 1, \ldots, N$, a set of input-output data with size N, we have the vector of linear model parameters θ that minimizes the objective function:

$$J = \sum_{k=1}^{N} \left(y_k - \phi_k^T \theta \right)^2 \tag{6}$$

where $\phi_k = \begin{bmatrix} \lambda_1\left(x_k\right) x_{ek}^T & \lambda_2\left(x_k\right) x_{ek}^T & \ldots & \lambda_R\left(x_k\right) x_{ek}^T \end{bmatrix}^T$, $x_{ek} = \begin{bmatrix} 1 & x_k^T \end{bmatrix}^T$. Equation (6) is estimated using the recursive least squares algorithm:

$$\hat{\theta}_k = \hat{\theta}_{k-1} + B_k \phi_k \left(y_k - \phi_k^T \hat{\theta}_{k-1} \right) \tag{7}$$

$$B_k = B_{k-1} - \frac{B_{k-1}\phi_k\phi_k^T B_{k-1}}{1 + \phi_k^T B_{k-1}\phi_k}, \quad k = 1, \ldots, N \tag{8}$$

The initial conditions are: $\hat{\theta}_0 = 0$ and $B_0 = \Omega I$, where Ω is a large positive number, I is the identity matrix, and B is a $(n+1) \times (n+1)$ dispersion matrix. [12] proposed a weighted average of the dispersion and parameters of the remaining R rules to estimate the respective dispersions and parameters of the new $(R+1)^{th}$ rule. When a new rule is added to the rule-base, the weighted average of the parameters of the other rules determines the new rule parameters since the weights are the normalized firing levels of the existing rules λ_i. Then, parameters of the old rules are inherited from the previous step as:

$$\hat{\theta}_k = \begin{bmatrix} \hat{A}_{1(k-1)}^T & \hat{A}_{2(k-1)}^T & \cdots & \hat{A}_{R(k-1)}^T & \hat{A}_{(R+1)k}^T \end{bmatrix}^T \tag{9}$$

where $\hat{A}_{(R+1)k} = \sum_{i=1}^{R} \lambda_i \hat{A}_{ik-1}$.

Dispersion matrices are reset as follows:

$$B_k = \begin{bmatrix} \gamma\Sigma_{11} & \cdots & \gamma\Sigma_{1R(n+1)} & 0 & \ldots & 0 \\ \cdots & \cdots & \cdots\cdots\cdots\cdots & \cdots & \cdots & \cdots \\ \gamma\Sigma_{R(n+1)1} & \cdots & \gamma\Sigma_{R(n+1)R(n+1)} & 0 & \ldots & 0 \\ 0 & 0 & 0 & \Omega & \ldots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & 0 & 0 & 0 & \ldots & \Omega \end{bmatrix} \tag{10}$$

where Σ_{ij} is co-variance matrix element ($i = j = [1, R \times (n+1)]$) and $\gamma = ((R^2 + 1)/R^2)$ is a coefficient. For more details about online recursive estimation of consequent parameters of eTS models see [12].

3. MODELING AND FORECASTING YIELD CURVES

In this section we recall the key ideas related to the yield curve modeling, including the relationships among the main theoretical constructs: the discount curve, the forward curve and the yield curve, as well as its forecasting.

3.1 Yield Curve Modeling

Let $P_t(\tau)$ denote the price of a τ-period discount bond, i.e., the present value at time t of $\$1$ receivable τ periods ahead, and let $y_t(\tau)$ denote its continuously compounded zero-coupon nominal yield to maturity. According to the yield curve we obtain the discount curve as:

$$P_t(\tau) = e^{-\tau y_t(\tau)} \tag{11}$$

and from the discount curve, the instantaneous (nominal) forward rate curve is obtained:

$$f_t(\tau) = \frac{-P_t'(\tau)}{P_t(\tau)} \tag{12}$$

We can write the relationship between the yield to maturity and the forward rate:

$$y_t(\tau) = \frac{1}{\tau} \int_0^\tau f_t(u)\,du \tag{13}$$

which implies that the zero-coupon yield is an equally-weighed average of forward rates. Given the yield curve or forward curve, we can price any coupon bond as the sum of the present values of future coupon and principal payments [2].

As yield curves, discount curves and forward curves are not observed in the markets, they must be estimated from observed bond prices. In this paper, we use Nelson-Siegel functional form [6], which is a convenient and parsimonious three-component exponential approximation. According to this model, the forward rate curve is given by:

$$f_t(\tau) = \beta_{1t} + \beta_{2t}e^{-\lambda_t\tau} + \beta_{3t}\lambda_t e^{-\lambda_t\tau} \tag{14}$$

The forward rate curve proposed by [6] is a constant plus a Laguerre function, which is a polynomial times and exponential decay term and is a popular mathematical approximation function. The corresponding yield curve is written as follows:

$$y_t(\tau) = \beta_{1t} + \beta_{2t}\left(\frac{1-e^{-\lambda_t\tau}}{\lambda_t\tau}\right) + \beta_{3t}\left(\frac{1-e^{-\lambda_t\tau}}{\lambda_t\tau} - e^{-\lambda_t\tau}\right) \tag{15}$$

The parameters β_{1t}, β_{2t} and β_{3t} are interpreted as long-term, short-term and medium-term components, respectively and, according to [2] they may also be interpreted in terms of level, slope and curvature. Expresion (15) means that the yield curve converges to β_{1t} as maturity increases, while it converges to $\beta_{1t} + \beta_{2t}$ as maturity decreases to present time. The parameters λ_t and β_{3t} control the possible presence of a hump in the yield curve. Specifically, λ_t determines the position (time) of the hump, while β_{3t} determines its magnitude and direction [20]. Note that β_{1t}, $\beta_{1t} + \beta_{2t}$ and λ_t should obviously be positive, while λ_t should have an upper bound, e.g. 30 years[2].

To estimate a zero-coupon yield curve, one would ideally use zero-coupon bonds traded in the market of interest rate and choose the parameters $\beta_i, \lambda, i = 1, 2, 3$ so as to minimize an error measure between the observed (i.e. obtained from the market) and fitted (i.e. calculated from the yield curve) yields or prices. Hence we have a non-linear function due to the λ parameter. In general, [2,3,6] fix the λ parameter to obtain a linear function for the corresponding β_i's. However, here we estimate the function considering all the parameters, using the non-linear least squares methodology (for details see [21]).

3.2 Forecasting Yield Curve Level, Slope and Curvature

In this paper, we consider the approach addressed by [2] for yield curve forecasting. The main idea is to model and forecast the Nelson-Siegel factors as univariate AR(1) process. The AR(1) models can be viewed as natural benchmarks determined a priori: the simplest great workhorse autoregressive models [2]. The yield forecasts based on underlying univariate AR(1) factor specifications are:

$$\hat{y}_{t+h/t}(\tau) = \hat{\beta}_{1,t+h/t} + \hat{\beta}_{2,t+h/t}\left(\frac{1-e^{-\hat{\lambda}_{t+h/t}\tau}}{\hat{\lambda}_{t+h/t}\tau}\right) + \hat{\beta}_{3,t+h/t}\left(\frac{1-e^{-\hat{\lambda}_{t+h/t}\tau}}{\hat{\lambda}_{t+h/t}\tau} - e^{-\hat{\lambda}_{t+h/t}\tau}\right) \tag{16}$$

where

$$\hat{\beta}_{i,t+h/t} = \hat{c}_i + \hat{\gamma}_i\hat{\beta}_{it} \quad i = 1, 2, 3, \tag{17}$$

$$\hat{\lambda}_{t+h/t} = \hat{d}_i + \hat{\alpha}_i\hat{\lambda}_t \tag{18}$$

---

[2]In this work, we suppose that the well-know Nelson-Siegel [6] curve is well-suited to our purpose, i.e., forecasting, which is a consensus in the related literature.

$\hat{c}_i$ and $\hat{\gamma}_i$ as well as $\hat{d}_i$ and $\hat{\alpha}_i$ are obtained by regressing $\hat{\beta}_{it}$ and $\hat{\lambda}_t$ on an intercept and $\hat{\beta}_{i,t-h}$ and $\hat{\lambda}_{t-h}$, respectively[3].

Differently from [2], we consider the λ parameter as time variant because, according to [20], the results are improved and it is not a good proxy for the Brazilian bonds market fix this parameter by observing the fixed-income market behavior. Thus, forecasting the yield curve is equivalent to forecasting $\{\hat{\beta}_{1t}, \hat{\beta}_{2t}, \hat{\beta}_{3t}, \hat{\lambda}_t\}$.

A random walk model was also chosen for comparison purposes because it provides good results for short-term forecasts. In this case, the forecast is always "no change", that is:

$$\hat{y}_{t+h/t}(\tau) = y_t(\tau) \tag{19}$$

In this paper we estimate the values of $\{\hat{\beta}_{1t}, \hat{\beta}_{2t}, \hat{\beta}_{3t}, \hat{\lambda}_t\}$ using evolving Takagi-Sugeno modeling. As in [2] we consider the parameters following an AR(1) process. The first lag of each variable $\{\beta_{it}, \lambda_t\}$ for $i = 1, 2, 3$ was considered as an input to the eTS model, when the output is the h-month-ahead forecast of the parameters, i.e., $\{\hat{\beta}_{i,t+h}, \hat{\lambda}_{t+h}\}$. Using the parameters forecast we evaluate the Nelson-Siegel function forecasts for different horizons, considering the proposed eTS model and the benchmarks, AR(1) and random walk models.

4. RESULTS AND DISCUSSION

In this section we present the fit of the well established Nelson-Siegel model and forecast its parameters using the autoregressive model as proposed by [2], and the random walk model to compare their performance against the evolving fuzzy rule-based method, i.e., eTS suggested here. We begin by describing the data.

4.1 Data Stream

The data stream was composed by daily series of the Brazilian Treasuries Bonds, more specifically, we consider the National Treasury Bills (LTN - *Letras do Tesouro Nacional*), which yields are determined (fixed rate) upon purchase and the form of payment is upon maturity. These are one of the most liquidity bond traded in the Brazilian bonds market. We cover the period from January 2007 to December 2009, composing a sample with 5,321 observations, taken from ANBIMA (Brazilian Financial and Capital Markets Association). The data was partitioned into two subsets. The first one, from January 2007 to December 2008 was used to estimate daily the Nelson-Siegel function parameters $\{\beta_{1t}, \beta_{2t}, \beta_{3t}, \lambda_t\}$, i.e., composing our in-sample base and resulting in a time series of cross sections. Finally, the out-of-sample data starts in January 2009 and finish in December 2009, used for forecasting purposes.

4.2 Fitting the Nelson-Siegel function

As discussed previously, we fit the yield curve using the Nelson-Siegel model as in equation (15). We estimated the parameters $\theta_t = \{\beta_{1t}, \beta_{2t}, \beta_{3t}, \lambda_t\}$ by nonlinear least squares for each day t. However, we do not fix λ_t at a prespecified value, which is more convenient since there is no rule in how to choose an appropriated value for this parameter.

The values of the parameters are chosen so as to minimize an error measure between the observed and fitted yields. Denoting by P_j the bond prices, obtained from the market, the goal is to minimize the error function E:

$$\min \sum_{j=1}^{N} \left(P_j - \hat{P}_j\right)^2 \tag{20}$$

where $\hat{P}_j, j = 1, \ldots, N$, is the fitted price of the jth bond in a N bond sample, according to the constraints $-\beta_i \leq 0, i = 1, 2, 3$ and $\tau - 30 \leq 0$.

Since the solution of an optimization problem depends largely on the starting point, we determine the start values for the parameters according to [20]:

$$\beta_1 = \frac{1}{M} \sum_{i=1}^{M} y_i, \quad \beta_2 = y_s - \beta_1, \quad \beta_3 = 0, \quad \lambda = 1 \tag{21}$$

where M refers to the bond with the longest maturity, and s is the rate related to the bond with the shortest horizon.

According to this methodology, we obtain a time series of estimates of $\{\hat{\beta}_{1t}, \hat{\beta}_{2t}, \hat{\beta}_{3t}, \hat{\lambda}_t\}$[4]. In Table 1 we present statistics for the estimated parameters.

Observing the autocorrelations of the four parameters, we can see that the $\hat{\beta}_{2,t}$ factor is the most persistent, followed by $\hat{\beta}_{1,t}$, $\hat{\beta}_{3,t}$ and $\hat{\lambda}_t$, respectively. Furthermore, the augmented Dickey-Fuller tests suggest that the level factor, $\hat{\beta}_{1,t}$, may have a unit root, and the other ones does not[5].

---

[3]In this way, the factors are directly regress at $t + h$ on factors at t, which is a standard method of coaxing least squares into optimizing the relevant loss function, h-month-ahead RMSE, as opposed to the usual l-month-ahead RMSE [2].

[4]In order to obtain these estimates we only have considered the period from January 2007 to December 2008.

[5]The critical values for rejection of hypothesis of a unit root are -3.4518 at the 1% level, -2.8704 at the 5% level, and -2.5714 at the 10% level.

Table 1: Descriptive Statistics for the Nelson-Siegel Function Estimated Parameters.

Parameter	Mean	Std. Dev.	Minimum	Maximum	$\hat{\rho}$ [1]	ADF[1]
$\hat{\beta}_{1,t}$	0.07065	0.04667	$2.2e10^{-12}$	0.12649	0.9117	-2.5665
$\hat{\beta}_{2,t}$	0.03624	0.04673	$2.4e10^{-14}$	0.12426	0.9255	-3.4094
$\hat{\beta}_{3,t}$	0.10777	0.08888	$1.9e10^{-07}$	0.45142	0.8655	-3.6683
$\hat{\lambda}_{t}$	3.53627	6.57317	0.0260054	29.9996	0.8616	-5.2345

[1]$\hat{\rho}$ represents the sample autocorrelation and ADF the augmented Dickey-Fuller unit root test statistics.

4.3 Out-of-Sample Forecasting Performance of the Nelson-Siegel Function Parameters

We forecast the Nelson-Siegel factors using univariate AR(1) and random walk processes and an evolving Takagi-Sugeno fuzzy model. For eTS model we consider as input the parameters $\{\beta_{1t}, \beta_{2t}, \beta_{3t}, \lambda_t\}$ in t and as output the parameters in $t + h$, that is the same idea as in autoregressive model AR(1)[6].

An accurate approximation to yield-curve dynamics should not fit only well in-sample, but it is more important forecast well out-of-sample, since is essential for decision making process in risk management. Therefore, we perform in this work models' evaluation based on out-of-sample forecasting, considering one, three, six, nine and twelve months ahead horizons. The data base from January 2007 to December 2008 was used to estimated the Nelson-Siegel parameters according to nonlinear least squares method and them obtain the factors time series. Then, the forecasts were compared with the factors estimated considering actual data for the period from January 2009 to December 2009.

The eTS model adopted the following values: $\Omega = 750$ and $\varphi = 0.6$. These are the only parameters of the eTS algorithm needs to be chosen by the user. We variate these parameters and verified that the results are not very sensitive to the Ω parameter but, the variation of φ affects the accuracy significantly. The number of rules from eTS model obtained by $\beta_{1,t}, \beta_{2,t}, \beta_{3,t}$ and λ_t parameters estimation were, respectively, 4, 3, 6 and 5 rules. The parameter $\beta_{3,t}$ requires the highest number of rules to describe its behavior in comparison with the other ones and, $\beta_{2,t}$ parameter resulted in the lowest number of rules.

To determine the accuracy of each model's estimates, we examine Root Mean Squared Error (RMSE):

$$RMSE = \sqrt{\left(\frac{1}{N}\right)\sum_{t=1}^{N}\left(\theta_t - \hat{\theta}_t\right)^2} \tag{22}$$

where θ represents the factors $\{\beta_{1t}, \beta_{2t}, \beta_{3t}, \lambda_t\}$ obtained by actual market data, $\hat{\theta}$ the factors for the models considered and N the sample size.

In Table 2 we compare the eTS model out-of-sample forecasting results from Nelson-Siegel parameters to those competitors, i.e., univariate AR(1) and random walk processes, for maturities of 1, 3, 6, 9 and 12 months.

Table 2: RMSE out-of-sample 1, 3, 6, 9 and 12-months-ahead forecasting results.

Parameters	Models	1month	3 months	6 months	9 months	12 months
$\hat{\beta}_{1,t}$	AR(1)	0.0025	0.0197	0.0099	0.0248	0.1142
	RW	0.0066	0.0202	0.0279	0.0533	0.1087
	eTS	0.0014	0.0041	0.0148	0.0241	0.1135
$\hat{\beta}_{2,t}$	AR(1)	0.0097	0.0014	0.0915	0.0247	0.0113
	RW	0.0101	0.0129	0.0642	0.0231	0.0540
	eTS	0.0078	0.0023	0.0686	0.0222	0.0439
$\hat{\beta}_{3,t}$	AR(1)	0.0195	0.0444	0.4072	0.0917	0.0232
	RW	0.0548	0.1715	0.4191	0.8003	0.3050
	eTS	0.0065	0.1630	0.3327	0.1159	0.1984
$\hat{\lambda}_{t}$	AR(1)	6.9789	7.8187	17.523	3.7070	0.5511
	RW	16.211	27.095	22.459	25.699	3.9291
	eTS	7.1899	7.2456	16.979	4.8536	1.7159

Random Walk for all months-ahead forecasting results, reported in Table 2, showed the highest error in terms of RMSE compared with eTS and autoregressive models. eTS model, in general, reveals better predictability performance than AR(1)

---

[6]The data base in eTS was normalized on the interval $[0, 1]$.

method for short, medium and medium term horizons. However, for long term horizons, AR(1) model performed slightly better than the eTS model in forecasting the curvature parameter, i.e., $\beta_{3,t}$.

In Figure 1 we selected fitted (model-based) yield curves for eTS and AR(1) models and compared with the curves obtained by actual market data. Clearly the estimated curves by eTS model are more capable of replicating the yields from the market. Furthermore, AR(1) works better with long term horizons, confirming the results described in [2]. Despite of that, eTS model showed good results also dealing with all horizons forecasting.
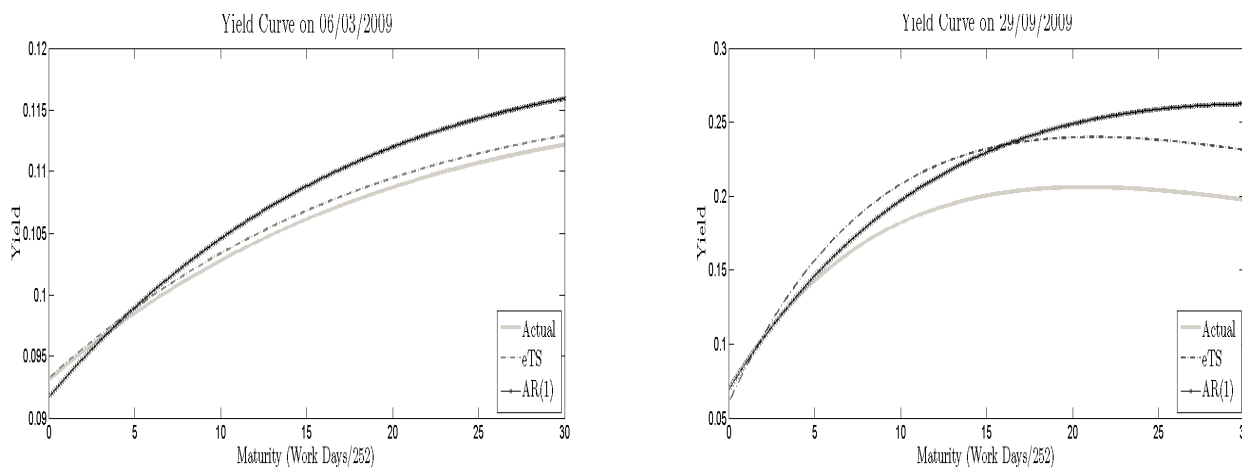
Figure 1: Fitted curves for selected dates (for short, 3 months, and long, 9 months, horizons), together with actual yields.

5. CONCLUSION

Takagi-Sugeno rule-based fuzzy models are a well established tool to deal with complex systems and the concept of evolving fuzzy systems use data streams to continuously adapt their structure and functionality. Evolving fuzzy rule-based models deal with rule modification to assemble a rule base that is more informative to the data samples, because evolution induces gradual change of the rule base structure inheriting structural information. In this paper we addressed the use of evolving Takagi-Sugeno (eTS) model to Brazilian yield curve forecasting. Term structure of interest rate forecasting is crucial for bond portfolio management, derivatives pricing and enable investors, policy makers and analysts to describe the market expectations about the future of interest rates as well as inflation and level of economic activity.

We introduced an approach which is a means to predict the parameters of the Nelson-Siegel [6] function with an eTS model. The eTS model is a Takagi-Sugeno rule-based model form whose rule base and parameters continually evolve by adding new rules with higher summarization power updating existing rules and parameters. For comparison purposes, we also addressed the methodology introduced in [2] which considers the Nelson-Siegel parameters time series as an autoregressive process AR(1), and a random walk model. The models were evaluated in terms of root mean squared error for one, three, six, nine and twelve months ahead forecasting horizons using daily series of the Brazilian National Treasury Bills for the period from January 2007 to December 2009.

According to the out-of-sample forecasts, the results obtained suggest that eTS model works well in Nelson-Siegel factors forecasting, which means more accurate estimates for short, medium and long term horizons. For all horizons, the random walk method displayed the highest errors, and the autoregressive model reveals better results for medium maturities when compared with eTS model. Further work shall include the application of statistical tests and residual analysis to evaluate the performance of the models.

Acknowledgment

REFERENCES

[1] M. S. Matsumura, A. R. B. Moreira and J. V. M. Valente. "Forecasting the Yield Curve with Linear Factor Models". *Brazilian Central Bank Working Paper*, vol. 1, no. 223, pp. 3–25, 2010.

[2] F. X. Diebold and C. Lin. "Forecasting the Term Structure of Government Bond Yields". *Journal of Econometrics*, vol. 130, no. 2, pp. 337–364, 2006.

[3] F. X. Diebold, C. Lin and V. Z. Yue. "Global Yield Curve Dynamics and Interactions: A Dynamic Nelson-Siegel Approach". *Journal of Econometrics*, vol. 146, no. 2, pp. 351–363, 2008.

[4] E. F. Fama and R. R. Bliss. "The Information in Long-Maturity Forward Rates". *American Economic Review*, vol. 77, no. 4, pp. 680–692, 1987.

[5] G. R. Duffee. "Term Premia and Interest Rates Forecasts in Affine Models". *Journal of Finance*, vol. 57, no. 1, pp. 405–443, 2002.

[6] C. Nelson and A. Siegel. "Parsimonious Modeling of Yield Curves". *Journal of Business*, vol. 60, no. 4, pp. 473–489, 1987.

[7] C. Almeida and J. Vicente. "The Role of No-Arbitrage on Forecasting: Lessons from a Parametric Term Structure Model". *Journal of Banking and Finance*, vol. 32, no. 12, pp. 2695–2705, 2008.

[8] C. G. Bowsher and R. Meeks. "The Dynamics of Economic Functions: Modeling and Forecasting the Yield Curve". *Journal of the American Statistical Association*, vol. 103, no. 484, pp. 1419–1437, 2008.

[9] D. O. Cajueiro, J. A. Divino and B. M. Tabak. "Forecasting the Yield Curve for Brazil". *Brazilian Central Bank Working Paper*, vol. 1, no. 197, pp. 3–16, 2009.

[10] P. Angelov. *Evolving Rule-Based Models: A Tool for Design of Flexible Adaptive Systems*. ISBN 3-7908-1457-1. Springer-Verlag, Heidelberg, New York, 2002.

[11] N. Kasabov and Q. Song. "DENFIS: Dynamic Evolving Neural-Fuzzy Inference System and Its Application for Time Series Prediction". *IEEE Trans. on Fuzzy Systems*, vol. 10, no. 2, pp. 144–154, 2002.

[12] P. Angelov and D. Filev. "An approach to on-line identification of evolving Takagi-Sugeno Fuzzy Models". *IEEE Trans. on Systems, Man and Cybernetics, part B*, vol. 34, no. 1, pp. 484–498, 2004.

[13] P. Angelov and D. Filev. "Simpl_eTS: A Simplified Method for Learning Evolving Takagi-Sugeno Fuzzy Models". *Proceedings of the 2005 International Conference on Fuzzy Systems, FUZZ-IEEE*, pp. 1068–1073, May 2005.

[14] R. Ballini, A. R. R. Mendonça and F. Gomide. "Evolving Fuzzy Modeling in Risk Analysis". *Intelligent Systems in Accounting, Finance and Management*, vol. 16, pp. 71–86, 2009.

[15] R. Ballini, A. R. R. Mendonça and F. Gomide. "Evolving Fuzzy Modeling of Sovereign Bonds". *Journal of Financial Decision Making*, vol. 16, pp. 3–15, 2009.

[16] S. McDonald and P. Angelov. "Evolving Takagi Sugeno Modeling with Memory for Slow Process". *International Journal of Knowledge-Based and Intelligence Engineering Systems*, vol. 14, no. 1, pp. 11–19, 2010.

[17] L. Maciel, F. Gomide and R. Ballini. "Evolving Fuzzy Systems for Pricing Fixed Income Options". In *IEEE Symposium Series on Computational Intelligence - SSCI*, pp. 54–61, Paris, France, May 2011.

[18] S. L. Chiu. "Fuzzy model identification based on cluster estimation". *Journal of Intelligent and Fuzzy Systems*, vol. 2, pp. 267–278, 1994.

[19] R. Yager and D. Filev. "Approximate Clustering via the Mountain Method". *IEEE Trans. on Systems, Man and Cybernetics*, vol. 24, no. 8, pp. 1279–1284, August 1994.

[20] P. Manousopoulos and M. Michalopoulos. "Comparison of non-linear optimization algorithms for yield curve estimation". *European Journal of Operational Research*, vol. 192, pp. 594–602, 2009.

[21] T. Strutz. *Data Fiting and Uncertainty: A Practical Introduction to Weighted Least Squares and Beyond*. Vieweg and Teubner Verlag, 2010.