2017 Evolving and Adaptive Intelligent Systems (EAIS)
Ljubljana, Slovenia, May 31 - June 2, 2017

Evolving participatory learning fuzzy modeling for financial interval time series forecasting

Leandro Maciel
School of Management and Accounting
Federal University of Rio de Janeiro
Rio de Janeiro, Brazil
Email: maciel@facc.ufrj.br

Rafael Vieira, Alisson Porto and Fernando Gomide
School of Electrical and Computer Engineering
University of Campinas
Campinas, São Paulo, Brazil
Emails: {giordano,alisport,gomide}@dca.fee.unicamp.br

Rosangela Ballini
Institute of Economics
University of Campinas
Campinas, São Paulo, Brazil
Email: ballini@unicamp.br

***Abstract*—Financial interval time series (ITS) describe the evolution of the maximum and minimum prices of an asset throughout time. These price ranges are related to the concept of volatility. Hence, their accurate forecasts play a key role in risk management, derivatives pricing and asset allocation, as well as supplements the information extracted by the time series of the closing price values. This paper addresses evolving fuzzy systems and financial ITS forecasting considering as the empirical application the main index of the Brazilian stock market, the IBOVESPA. An evolving participatory learning fuzzy model, named ePL-KRLS, is proposed. The model extends traditional ePL approach by considering Kernel functions to the identification of rule consequents parameters as well as a metaheuristic algorithm to automatically set model control parameters. One step ahead interval forecasts is compared against linear and nonlinear time series benchmark methods and with the state of the art evolving fuzzy models in terms of traditional accuracy metrics and quality measures designed for ITS. The results provide evidence for the predictability of of IBOVESPA ITS and significant forecast contribution of ePL-KRLS.**

I. Introduction

Prediction of asset prices in financial markets is the basis for asset allocation, portfolio management and derivatives pricing [1]. The temporal progress of assets, stock indices, and exchange rates is observed as single-valued financial time series [2]. This is useful in many cases, but it may be insufficient in situations where several values are observed at each time period. For instance, if only the opening (or closing) asset price is measured daily, the resulting time series will hide the intraday variability and important information is missed.

Besides that intraday time series could be forecasted, they reveal characteristics such as irregular temporal spacing, strong diurnal patterns and complex dependence, which result in obstacles for traditional time series models. Further, the accurate prediction of the whole sequence of intraday prices for one day ahead is almost impossible in practical situations. These limitations can be alleviated if the highest and the lowest values of prices are measured at each time period, what originates interval time series (ITS) [3].

In particular, financial ITS modeling and forecasting have received considerable attention in the recent literature. The literature has introduced several interval time series forecasting methods. Examples include [4], [5], [6], and [7]. The focus of this paper is on financial interval time series forecasting. The current literature advocates the use of ITS framework in economics and finance, since it provides appropriate mechanisms to analyze large data sets such as e.g. high-frequency data, and also supplements the information extracted by the time series of the closing values considering high and low prices in terms of a proxy measure of volatility. Despite the evidence of more accurate performance of nonlinear models, traditional econometric approaches adopt restrictive assumptions that are not observable in practice. This is because of the complexity of the noisy, nonlinear, non-stationary dynamical behavior of assets prices. Thus, this paper aims to address financial interval time series forecasting and evolving fuzzy systems in order to overcome the main limitations of traditional econometric approaches.

Hence, this paper suggests evolving fuzzy systems for financial ITS forecasting. An extension of the ePL+ method is proposed. In ePL+ consequent rules are comprised of local linear models. However, the use of local linear models to fit clusters whose data distribution are not necessarily linear may lead to a modeling error, thus reducing the overall model performance. As an alternative, an ePL Kernel Recursive Least Squares model, ePL-KRLS, is suggested. In ePL-KRLS nonparametric approaches for updating the model rule consequents are employed using Kernel functions, thus providing higher generalization power and the ability to approximate more complex dynamics. Further, differently from ePL+, control parameters, which were selected by the user, are now estimated recursively using a metaheuristic algorithm called GRASP (greedy randomized adaptive search procedure), providing a more autonomous user-free parameters algorithm.

ITS forecasts are constructed by the individual prediction of minimum and maximum stock prices. Experiments are conducted using the main stock index of the Brazilian financial market, the IBOVESPA, for the period from January 2000 to December 2015, focusing on one step ahead forecasts. ePL-KRLS is compared with traditional econometric benchmark approaches and with the state of the art evolving fuzzy models in terms of quality measures designed for the interval time series framework, as well as using additional forecast descriptive statistics such as efficiency and coverage rates.

This paper is organized as follows. A briefly literature review regarding financial interval time series forecasting and

978-1-5090-6444-1/17/$31.00 ©2017 IEEE

evolving fuzzy systems is presented in Section II. Section III describes the evolving participatory learning fuzzy model. Section IV comprises the computational experiments. Finally, Section V concludes and suggests issues for further investigation.

II. Related work

A. Financial interval time series forecasting

Analysis of large data sets, such as high-frequency data, based on the ITS framework is a new domain to study statistically detectable patterns. It has attracted a large number of researchers in economics and finance [8]. The daily high and low financial prices can be seen as reference values for investors in order to place buy or sell orders. Furthermore, these prices are related to the concept of volatility. [9] show that the difference between the highest and lowest log prices over a fixed sample interval, also known as the log range, is a highly efficient volatility measure. The literature that considers the high-low range prices as a proxy for volatility dates back to the 1980s with the work of [10]. [11] and [12] also stated that the range-based volatility estimator appears robust to microstructure noise such as bid-ask bounce, which overcomes the limitations of traditional volatility models based on closing prices that fall to use the information contents inside the reference period of the prices, resulting in inaccurate forecasts [13].

In general, the literature addresses ITS using extensions of classic data analytics and forecasting techniques [14], [15]. For instance, [14] use a vector error correction (VEC) model to forecast low, high and closing values of three foreign exchange rates: Dollar-Yen, Pound-Dollar, and Mark-Dollar. According to the authors, the data generating the process of the prices time series are different, and the prediction of low and high values improves closing prices forecasting. Similarly, [16] indicates that high and low values of the Dow Jones Industrial, S&P 500 and NASDAQ indices increase the explanatory capacity of a VEC model.

An interval linear model to forecast the stock market price movements is addressed in [17]. The intervals are constructed based on lower and higher equity prices, and the model is calculated from their midpoints (centers) using ordinary last squares. Further, [18] considered the same econometric model but estimated separately for the lower bounds and the upper bounds of the interval time series. In [19], both approaches are compared and the authors conclude that the former yields more accurate results.

The consideration of ITS to forecast Dow Jones Industrial index and Euro-Dollar exchange rate was conducted by [2] using univariate and multivariate techniques such as exponential smoothing, autoregressive integrated moving average (ARIMA), artificial neural networks, and pattern recognition methods. Forecasts are produced using independent (univariate) and joint (multivariate) predictions of the lower and upper bounds of intervals or, alternatively, their midpoints (center) and half-lengths (radius). Both univariate and multivariate approaches performed similarly. Approaches that independently forecasts the lower and upper bounds time series obtained the worst results.

Using intraday data, [20] compare ARIMA, naive, and interval linear regression models to forecast S&P 500, Dow Jones Industrial and Nasdaq indices. The authors claim that the interval-based approaches are superior when compared to single-valued time series models. Recently, [21] developed a support vector machine model to simultaneously forecast the minimum and maximum values of the S&P 500, FTSE 100, and Nikkei 225 interval indices. The results, when compared with classic econometric techniques and interval neural networks, show the superior potential of the interval-based model, especially in financial trading.

B. Evolving fuzzy systems

Evolving fuzzy systems are an advanced form of adaptive systems because they have the ability to simultaneously learn the model structure and functionality from flows of data. eFS has been useful to develop adaptive fuzzy rule-based models, neural fuzzy, and fuzzy tree models, to mention a few. Examples of the different types of evolving fuzzy rule-based and fuzzy neural modeling approaches include the pioneering evolving Takagi-Sugeno (eTS) modeling [22] approach and extensions (e.g. Simpl_eTS [23], eXtended eTS (xTS) [24]). An autonomous user-free control parameters modeling scheme called eTS+ is given in [25]. The eTS+ uses criteria such as age, utility, local density, and zone of influence to update the model structure. Later, ePL+ was developed in the realm of participatory learning clustering [26]. ePL+ extends the ePL approach [27] and uses the updating strategy of eTS+.

An alternative method for evolving TS modeling is given in [28] based on a recursive form of the fuzzy c-means (rFCM) algorithm. Clustering aims at learning the model structure. Later, the rFCM method was translated in a recursive Gustafson-Kessel (rGK) algorithm. Similarly as the original off-line GK, the purpose of rGK is to capture different cluster shapes [29]. The combination of rGK algorithm and evolving mechanisms such as adding, removing, splitting, merging clusters, and recursive least squares became a powerful evolving fuzzy modeling approach called eFuMo [30]. A comprehensive source of the evolving approaches can be found in [31].

III. ePL-KRLS approach

This section details the suggested ePL-KRLS modeling approach. First, the basic constructs of Takagi-Sugeno fuzzy rule-based models are described. Rules antecedents and consequents identification using participatory learning principles and Kernel recursive least squares are then discussed. Finally, the mechanism to adaptively update model parameters and the corresponding algorithm are presented.

A. Takagi-Sugeno fuzzy model

Takagi-Sugeno (TS) fuzzy model with affine consequents consists of a set of fuzzy functional rules in the form:

$$\mathcal{R}_i : \text{ IF } \mathbf{x} \text{ is } \mathcal{A}_i \text{ THEN } y_i = f_i(\mathbf{x}) = \theta_{i0} + \theta_{i1} x_1 + \ldots + \theta_{im} x_m, \tag{1}$$

where $\mathcal{R}_i$ is the i-th fuzzy rule, $i = 1, 2, \ldots, c$, c is the number of fuzzy rules, $\mathbf{x} = [x_1, x_2, \ldots, x_m]^T \in \Re^m$ is the input, $\mathcal{A}_i$ is the fuzzy set of the antecedent of the i-th fuzzy rule whose membership function is $\mathcal{A}_i(\mathbf{x}) : \Re \rightarrow [0, 1]$, $y_i \in \Re$ is the output of the i-th rule, and θ_{i0} and θ_{ij}, $j = 1, \ldots, m$, are the parameters of the consequent of the i-th rule.

Fuzzy inference using TS rules (1) proceeds as follows:

$$y = \sum_{i=1}^{c} \left(\frac{\mathcal{A}_i(\mathbf{x}) y_i}{\sum_{j=1}^{c} \mathcal{A}_j(\mathbf{x})} \right). \tag{2}$$

The expression (2) can be rewritten using normalized degrees of activation:

$$y = \sum_{i=1}^{c} \lambda_i y_i = \sum_{i=1}^{c} \lambda_i \mathbf{x}_e^T \theta_i, \tag{3}$$

where

$$\lambda_i = \frac{\mathcal{A}_i(\mathbf{x})}{\sum_{j=1}^{c} \mathcal{A}_j(\mathbf{x})}, \tag{4}$$

is the normalized activation level of the i-th rule, $\theta_i^T = [\theta_{i0}, \theta_{i1}, \ldots, \theta_{im}]$ is the vector of parameters, and $\mathbf{x}_e^T = [1 \ \mathbf{x}^T]$ is the expanded input vector.

The TS model uses parametrized fuzzy regions and associates each region with a local affine model. The nonlinear nature of the rule-based model emerges from the fuzzy weighted combination of the collection of the multiple local affine models. The contribution of a local model to the model output is proportional to its degree of activation.

TS modeling requires two tasks: i) learning the antecedent part of the model e.g. using a fuzzy clustering algorithm, participatory learning in this paper and ii) estimation of the parameters of the affine consequents.

B. Participatory learning

Participatory learning (PL) is a learning paradigm which assumes that the learning process depends on what the system has already learned from the data. The current knowledge is part of the learning process itself and influences the way in which new data are used for self-organization. An essential property of participatory learning is that the impact of new data in causing self-organization or model revision depends on its compatibility with the current knowledge, or equivalently in the context of this paper, on the compatibility of current input data with the existing cluster structure [27].

In participatory learning, a cluster structure is updated using a compatibility measure, $\rho_{ik} \in [0, 1]$ and an arousal index, $\alpha_{ik} \in [0, 1]$. While ρ_{ik} measures how much a data point is compatible with the current cluster structure, the arousal index α_{ik} acts as a critic to remind when current cluster structure should be revised in front of new information.

Due to its unsupervised, self-organizing nature, the PL clustering procedure may create or delete a new cluster, and merge or modify the existing ones at each learning step. If α_{ik} is greater than a threshold $\tau \in [0, 1]$, then a new cluster is created. The arousal index α_{ik} is updated as follows:

$$\alpha_{i,k+1} = \alpha_{ik} + \beta(1 - \rho_{i,k+1} - \alpha_{ik}), \tag{5}$$

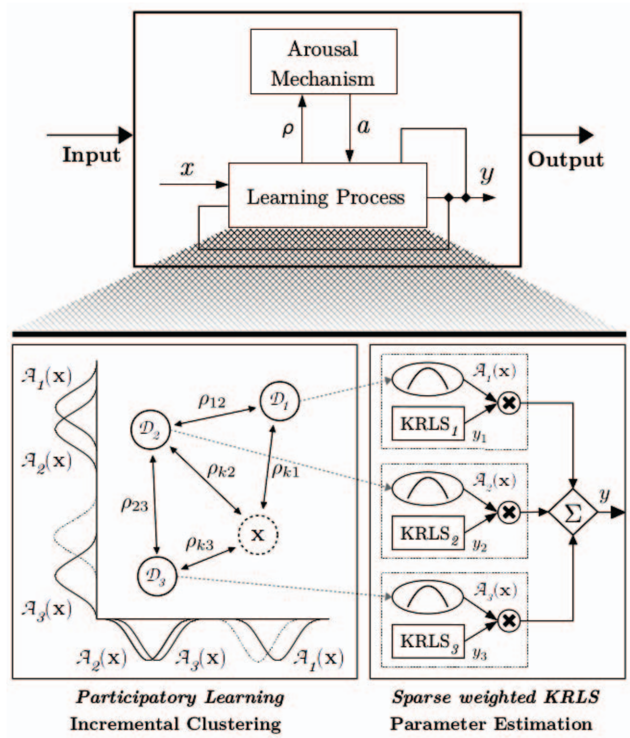

Fig. 1. Participatory learning for cluster structure adaptation in ePL-KRLS.

where $\beta \in [0, 1]$ controls the rate of change of arousal, the closer β is to one, the faster the system is to sense compatibility variations, $\alpha_{i0} = 0$, and

$$\rho_{ik} = 1 - \frac{||\mathbf{x}_k - c_i^k||}{n}, \tag{6}$$

is the compatibility index, with n as the size of the input space, and c_i^k is the i-th cluster center at step k.

Otherwise, the most compatible cluster, $\hat{c}_i^k | i = \text{argmax}_j \ \{\rho_j^k\}$, is updated as follows:

$$c_i^{k+1} = \hat{c}_i^k + G_i^k (x^k - \hat{c}_i^k). \tag{7}$$

If α_{ik} increases, the similarity measure has a reduced effect. The arousal index can be interpreted as the complement of the confidence we have in the truth of the current belief, the rule base structure. The arousal mechanism monitors the performance of the system by observing the compatibility of the current model with data. Figure 1 shows how participatory learning performs to adapt the cluster structure in ePL-KRLS.

In this paper, the quality of the clusters is monitored at each step considering the utility measure introduced in [25]. The utility measure is an indicator of the accumulated relative firing level of a rule:

$$\mathcal{U}_{ik} = \frac{\sum_{l=1}^{k} \lambda_i}{k - I^{i*}}, \tag{8}$$

where I^{i*} is the step (time tag) that indicates when fuzzy rule $i*$ was created. Once a rule is created, the utility indicates how much the rule has been used. This quality measure aims at avoiding unused clusters kept in the structure. Clusters corresponding to low-quality fuzzy rules can be deleted. In other words:

$$\text{IF} \;\; \mathcal{U}_{ik} < \epsilon \;\; \text{THEN} \;\; c \leftarrow c - 1, \tag{9}$$

where $\epsilon \in [0.03; 0.1]$ is a threshold that controls the utility of each cluster [25]. This principle guarantees high relevance cluster structure and corresponding fuzzy local models. Alternative quality measures such as age, support, zone of influence and local density may be adopted. Therefore, the rule-base structure is updated as follows: at each new data the compatibility and arousal are computed; if the arousal exceeds its threshold a new cluster is created; otherwise the most compatible is updated; if redundant clusters are verified they are merged; and finally, clusters with low utility are excluded.

C. Consequent parameters identification

After updating the rule-base structure of the proposed model, the next step aims to calculate the values of its rule consequents. This procedure is often performed in an online manner by using the famous recursive least squares (RLS) algorithm. Nevertheless, this paper considers using a more robust and flexible version of the traditional RLS, which is known as kernel recursive least square (KRLS) method [32].

The KRLS method is primarily based on a *kernel*, i.e., a function that performs an inner product between two samples mapped to a high-dimensional space (also called feature space). The basic idea behind using kernel is that problems which are tough to solve using the original dimensions of the input data can be easier fitted in the feature space. Furthermore, the use of a *kernel trick* technique can help to overcome the underlying problem of handling with vectors in such high dimensionality by replacing inner products in the feature space by kernel functions.

In this paper, the estimated output y for a given input data $\mathbf{x} = [x_1, x_2, \ldots, x_m]^T \in \Re^m$ is calculated through a weighted sum of the $y_i \in \Re$ local KRLS models, as follows:

$$y = \sum_{i=1}^{c} y_i \lambda_i \tag{10}$$

The value of each local KRLS model can be calculated from Eq. (10) in the following form:

$$y_i = \sum_{j=1}^{size(D_i)} \theta_{ij} \kappa(\mathbf{x}, d_{ij}) + y_i^{res} \tag{11}$$

where D_i is the i-th local dictionary[1], d_{ij} and θ_{ij} are the j-th stored data and the consequence parameter of the i-th local dictionary, y_i^{res} is the approximation error, or also called *residual* and $\kappa(\cdot, \cdot)$ is the Gaussian kernel function:

$$\kappa(x, y) = exp\left(\frac{||x - y||^2}{2\sigma^2}\right) \tag{12}$$

[1]Notice that the length of D_i may vary according to the problem requirements and the computational resources available.

Algorithm 1 ePL-KRLS algorithm

Input: Sample: $x^k \in [0,1]^m$ **and** lagged output: $y^{k-1} \in [0,1]$
Output: Predicted output: $y^k \in [0,1]$

for $x^k \in [0,1]^m, k = 1, 2, \ldots$ **do**
    if $k - 1$ **mod** $\frac{c}{a_i^k} = 0$ **then**
        Update $\{\alpha, \beta, \tau, \lambda\}$ using **Algorithm 2**
    end if
    for $i = 1, 2, \ldots, c$ **do**
        Compute the compatibility: $\rho_i^k = 1 - \frac{||x^k - \mathbf{v}_{ik}||}{p}$
        Compute the arousal: $a_i^{k+1} = \beta(1 - \rho_i^k - a_i^k)$
    end for
    if $a_i^{k+1} \geq \tau, \forall i = 1, 2, \ldots, c$ **then**
        Create a new cluster $c_i^k = c_i^k \cup x^k$ and set $c = c + 1$
    else
        Find the closest cluster: $\hat{c}_i^k | i = \text{argmax}_j \; \{\rho_j^k\}$
        Update i_{th} cluster position $c_i^{k+1} = \hat{c}_i^k + G_i^k(x^k - \hat{c}_i^k)$
    end if
    for $i \in \{1, 2, \ldots, c\}$ **and** $j \in \{1, 2, \ldots, c-1\}$ **do**
        if $\exists \; i, j \mid \frac{||c_i^k - c_j^k||}{p} > \lambda$ **then**
            Merge clusters i, j and make $c = c - 1$
        end if
    end for
    Compute utility of current clusters $U_i^k = \frac{\sum_{l=1}^k \lambda^l}{k - I^* i}$
    if $\exists \; i \mid U_i^k < \epsilon$ **then**
        Remove cluster i and make $c = c - 1$
    end if
    for $i = 1, 2, \ldots, c$ **do**
        Calculate residual $y_i^{res} = y_i - \sum_{j=1}^{size(D_i)} \theta_{ij} \kappa(x, d_{ij})$
        if $y_i^{res} < \nu$ **and** $d_i < d_{max}$ **then**
            $d_{i+1} = d_i + 1$
            $D_i = D_i \cup \{x_i\}$
        else
            Replaces $x_{oldest} \in D_i$ with x_i
        end if
    end for
    Update rules consequents using KRLS approach
    Calculate system output: $y = \sum_{i=1}^{c} y_i \lambda_i$
end for

having σ as kernel size. It is also worth to mention that one important property regarding kernel-based methods is the way the samples are mapped to the feature space. In KRLS, it occurs by calculating approximate linear combinations between the incoming data $\mathbf{x}$ and all the d_{ij} local dictionary samples. In this case, if the incoming data is enough representative, i.e., if $y_i^{res} < \nu$ (where ν represents a precision threshold for the linear combination), it is added to the dictionary. On the other hand, the local KRLS model replaces the less significant sample within the dictionary for the current data. For more details about consequent parameters updating using KRLS approach refer to [33]. The ePL-KRLS identification steps are described in Algorithm 1.

D. Adaptive tuning of regularization parameters

The procedure for obtaining near-optimal values for the regularization parameters of the proposed model (λ, τ, α and β) is performed by using a metaheuristic algorithm called GRASP (greedy randomized adaptive search procedure) [34].

The GRASP algorithm consists of two phases: a greedy adaptive randomized construction phase and a local search phase. At first, candidate elements in the problem search space are chosen at random when building up a constructive solution S. These candidates are assessed according to a greedy function φ in order to measure the contribution of each one in solving the problem. Once carried out this evaluation, the greedy function φ is used to find the locally optimal solution of the S neighborhood.

Algorithm 2 Adaptive parameter tuning using GRASP

Input: Parameter distributions: $[\alpha, \beta, \tau, \lambda] \subset [0, 1]^4$
Current best solution: $O_c^k \subset [0, 1]^4$
Output: Next best solution: $O_c^{k+1} \subset [0, 1]^4$

Let $S = \{O_c^k\}$
$\triangleright$ Greedy randomized construction
for $i = 1, 2, \ldots, r$ **do**
  Build the RCL: $R_c \in [\alpha, \beta, \tau, \lambda] \subset [0, 1]^{4 \times n}$
  Select from R_c an element v at random
  Let $S = S \cup v$
end for
$\triangleright$ Local search
while S is not locally optimal **do**
  Find $\hat{s} \in N$ such that $\varphi(\hat{s}) \leq s$
  $S = \hat{s}$
end while
Set the best solution: $O_c^{k+1} = S$

In the context of this paper, the candidate solutions are represented by a set of combinations among λ, τ, α and β, whose values can range from zero to one with a hundred data discretizations. The greedy function φ is computed as the normalized mean square error of ePL-KRLS using as input the acquired parameter combination. The evaluation of such candidates is performed in an online fashion, although not in all steps. It is because changes in the values of regularization parameters will not lead to immediate changes in the model performance. For this reason, one can specify a *time-interval analysis* t, i.e., a value that indicates the frequency which the new candidates will be evaluated through GRASP.

In this paper, t is defined at random, but considering a range between 0 and ten times the number of rules M at some iteration k. The procedures regarding the two phases of GRASP are described in the Algorithm 2.

IV. Computational experiments

The ePL-KRLS approach introduced in this paper gives a flexible modeling procedure and can be applied to a range of problems such as process modeling, time series forecasting, classification, system control, and novelty detection. This section evaluates the performance of ePL-KRLS for financial interval time series forecasting. The results of ePL-KRLS are compared with traditional econometric benchmarks such as Random Walk (RW), Exponential Smoothing, ARIMA, and Threshold Autoregressive (TAR) models and with state of the art evolving fuzzy, neuro and neuro-fuzzy modeling approaches, the eTS+ [25], ePL+ [26] and eFuMo [30].

A. Accuracy measures

In order to access models performance, traditional time series error measures such as root mean squared error (RMSE) and symmetric mean absolute percentage error (SMAPE) are considered. They are computed as follows:

$$\text{RMSE} = \sqrt{\frac{1}{T}\sum_{t=1}^{T}(y_t - \hat{y}_t)^2}, \tag{13}$$

$$\text{SMAPE} = \frac{1}{T}\sum_{t=1}^{T}\frac{|y_t - \hat{y}_t|}{(|y_t| + |\hat{y}_t|)/2}, \tag{14}$$

where $\hat{y}_t$ is the t-th forecasted value, y_t the t-th actual value, and T is the sample size.

To translate the performance metrics to intervals let us define interval time series. An ITS, $\{Y_t\}_{t=1}^{T}$, is a sequence of intervals observed in successive instants in time $t = 1, \ldots, T$. Each interval is represented by:

$$Y_t = [y_t^l, y_t^u] \in \Im, \tag{15}$$

where $\Im = \{[y_t^l, y_t^u] : y_t^l, y_t^u \in \Re, y_t^l \leq y_t^u\}$ is the set of closed intervals of the real line $\Re$, y_t^l the lower bound, and y_t^u the upper bound of the interval.

In the context of financial ITS, y_t^l and y_t^u correspond to the minimum and maximum prices over time t, respectively. Equivalently, the interval Y_t can be represented as

$$Y_t = \langle c_t, r_t \rangle, \tag{16}$$

where $c_t = (y_t^l + y_t^u)/2$ is the midpoint of the interval (center) and $r_t = (y_t^u - y_t^l)/2$ is the half-length of the interval (radius).

Note that traditional error measurements such as RMSE and SMAPE are not appropriate for ITS since the difference between the observed and the forecasted interval does not faithfully represent the concept of deviation besides considering interval arithmetics[2]. One way to overcome this limitation is measuring the error in each one of the four attributes of the ITS (lower and upper bounds, center and radius). This is done by considering the four attribute time series, $\{y_t^l\}$, $\{y_t^u\}$, $\{c_t\}$ and $\{r_t\}$, and then estimating the error for each series by comparing it to the respective forecasted attribute time series, $\{\hat{y}_t^l\}$, $\{\hat{y}_t^u\}$, $\{\hat{c}_t\}$ and $\{\hat{r}_t\}$ [2]. Hence, error metrics are computed for the four interval time series attributes.

One must note that the magnitude of the errors made in the four-time series can be quite different, mainly for the series of the radius, whose magnitude will usually be very different than the other threes (mainly concerning financial ITS data). Further, since the data has an interval structure it implies that both characteristics (lower and upper bounds or center and radius) that describe intervals have to be taken into consideration jointly using dissimilarity measures based on interval distance between observed interval and its forecast.

[2]For more details about interval arithmetics see [35].

To quantify the overall accuracy of the fitted and forecasted ITS, the mean distance error (MDE) of intervals is considered:

$$\mathrm{MDE} = \frac{\sum_{t=1}^{T} D(Y_t, \hat{Y}_t)}{T}, \tag{17}$$

where Y_t and $\hat{Y}_t$ are the observed and forecasted ITS and $D(\cdot)$ is an interval distance. As suggested in [2] and [8], the Euclidian distance for intervals is selected for lower and upper bounds representation:

$$D_{\mathrm{E}}(Y_t, \hat{Y}_t) = \sqrt{(y_t^l - \hat{y}_t^l)^2 + (y_t^u - \hat{y}_t^u)^2}. \tag{18}$$

Therefore, the MDE using Euclidian interval distance is computed for lower and upper bounds ($\mathrm{MDE}^{\mathrm{E}}$) interval representation. Another distance measure for interval time series forecasting is the normalized symmetric difference (NSD) of intervals, computed as [8]:

$$D_{\mathrm{NSD}}(Y_t, \hat{Y}_t) = \frac{w(Y_t \cup \hat{Y}_t) - w(Y_t \cap \hat{Y}_t)}{w(Y_t \cup \hat{Y}_t)}, \tag{19}$$

where $w(\cdot)$ indicates the width of the interval. The advantage is that NSD distance is a normalized distance measure and it is not affected by the data magnitude as the Euclidian distance.

The computation of descriptive statistics for ITS are also conducted as suggested by [8]. This paper calculates the coverage rate

$$\mathrm{R}^{\mathrm{C}} = \frac{1}{T}\sum_{t=1}^{T} \frac{w(Y_t \cap \hat{Y}_t)}{w(Y_t)}, \tag{20}$$

and the efficiency rate

$$\mathrm{R}^{\mathrm{E}} = \frac{1}{T}\sum_{t=1}^{T} \frac{w(Y_t \cap \hat{Y}_t)}{w(\hat{Y}_t)}. \tag{21}$$

These rates provide additional information on what part of the observed ITS is covered by its forecasts (coverage) and what part of the forecast covers the observed ITS (efficiency). If the observed intervals are fully enclosed in the predicted intervals then the coverage rate will be 100%, but the efficiency could be less than 100% and reveal the fact that the forecasted ITS is wider than the actual ITS. Hence, these statistics must be considered jointly. So, the indication of a good forecast is observed when the coverage and efficiency rates are reasonably high and the difference between them is small [8].

B. Results and discussion

This paper evaluates ePL-KRLS for financial interval time series forecasting. The empirical application concerns the Brazilian stock market. Data comprise daily minimum and maximum values of IBOVESPA for the period from January 2000 to December 2015[3]. Minimum and maximum values of IBOVESPA are used individually as intervals lower and upper bounds representations - Eq. (15). Thus, all forecasting models predict separately the minimum and maximum prices in order to produce the respective interval forecasts.

[3]Data were collected in Economatica.

Table I shows the descriptive statistics for the lower bound and upper bound of IBOVESPA ITS. As expected, time series of IBOVESPA intervals bounds are very similar, mainly in terms of mean and standard deviations. These series have heavy left-side tails as indicated by the negative skewness coefficients, and also lower values of kurtosis. The Jarque-Bera [36] statistics indicate that the series are non-normal with a 99% confidence interval.

TABLE I
DESCRIPTIVE STATISTICS OF LOWER AND UPPER BOUNDS OF IBOVESPA ITS FOR THE PERIOD FROM JANUARY 2000 TO DECEMBER 2015.

Statistic	Lower bound	Upper bound
Mean	40098.48	40984.33
Maximum	45401.00	73920.00
Minimum	8224.00	8513.00
Std. Dev.	19493.96	19836.37
Skewness	-0.1877	-0.1998
Kurtosis	1.5688	1.5649
Jarque-Bera	361.31	366.60
p-value	0.0000	0.0000

Data were split into two sets: in-sample and out-of-sample. The in-sample set includes the period from January 2000 to December 2011 and is used to estimate models parameters. The remaining data, from January 2012 to December 2015, comprises the out-of-sample set, for forecasting purposes. Note that this procedure is only required for the econometric models since the evolving algorithm can be applied from scratch. Models are evaluated for one-step ahead forecasts.

Models specifications, i.e. number of inputs and control parameters, were based on simulations using in-sample data in order to reach the best performance in terms of RMSE and SMAPE. ARIMA(2,1,3) and TAR (2) were selected for both lower and upper bounds of IBOVESPA ITS. The evolving fuzzy models take the following formulation:

$$\hat{y}_t = f(y_{t-1}, y_{t-2}, \ldots, y_{t-p}). \tag{22}$$

The number of lags, p, was also selected based on simulations in order to reach the lowest RMSE and SMAPE values. eTS+, ePL+ and eFuMo use $p = 4, 5, 3$, respectively. For the ePL-KRLS, experiments indicate $p = 5$. Also, the dictionary and kernel size of ePL-KRLS were defined as $N = 20$ and $\sigma = 0.9$, respectively.

Table II show the accuracy measures, RMSE and SMAPE, from all forecasting approaches for IBOVESPA ITS. All results concern the out-of-sample data set, i.e. the period from January 2012 to December 2015. Forecasts are one step ahead. Lower RMSE and SMAPE values indicate better accuracy. Best results are highlighted in bold. Considering both RMSE and SMAPE metrics, the evolving fuzzy methods, i.e. eTS+, ePL+, eFuMo and ePL-KRLS models, performed better for all lower and upper bounds time series in comparison with the econometric benchmarks. The suggested approach, ePL-KRLS, showed the lowest RMSE and SMAPE values in all cases (Table II). Amongst the evolving methods, results are similar. Nonetheless, in order to evaluate the better interval representation, it is necessary to consider the interval representatives forecasts jointly.

TABLE II
RMSE AND SMAPE VALUES BASED ON ONE STEP AHEAD FORECASTS OF LOWER AND UPPER BOUNDS IBOVESPA ITS REPRESENTATIONS FOR OUT-OF-SAMPLE DATA IN THE PERIOD FROM JANUARY 2012 TO DECEMBER 2015. BEST RESULTS ARE IN BOLD.

	RMSE		SMAPE	
Models	Lower b.	Upper b.	Lower b.	Upper b.
RW	655.78	669.84	0.00935	0.00931
ES	655.88	669.94	0.00935	0.00931
ARIMA	649.59	662.15	0.00922	0.00917
TAR	645.22	656.32	0.00916	0.00912
eTS+	379.46	399.12	0.00599	0.00647
ePL+	341.98	359.77	0.00531	0.00530
eFuMo	329.66	332.01	0.00478	0.00501
ePL-KRLS	**306.35**	**317.47**	**0.00398**	**0.00407**

In order to access the interval structure of the time series, IBOVESPA ITS forecasts were compared in terms of the mean distance error (MDE) of intervals using both the Euclidian distance ($\mathrm{MDE}^{\mathrm{E}}$) and the normalized symmetric difference (NSD) distance of intervals ($\mathrm{MDE}^{\mathrm{NSD}}$). Table III reports the results from $\mathrm{MDE}^{\mathrm{E}}$ and $\mathrm{MDE}^{\mathrm{NSD}}$ accuracy metrics of the actual and forecasted IBOVESPA ITS. Best results are in bold. Again, the lowest the MDE values, the highest the model's accuracy. The better results are again from the evolving fuzzy methods. Amongst the benchmark techniques, TAR method showed the lowest interval error values. ePL-KRLS method reports the lowest MDE values, using both Euclidian and normalized symmetric difference distance metrics (Table III).

TABLE III
MEAN DISTANCE ERROR (MDE) VALUES USING EUCLIDIAN DISTANCE, $\mathrm{MDE}^{\mathrm{E}}$, AND THE NORMALIZED SYMMETRIC DIFFERENCE (NSD), $\mathrm{MDE}^{\mathrm{NSD}}$, BASED ON ONE STEP AHEAD FORECASTS OF IBOVESPA ITS FOR OUT-OF-SAMPLE DATA IN THE PERIOD FROM JANUARY 2012 TO DECEMBER 2015. BEST RESULTS ARE IN BOLD.

Models	$\mathrm{MDE}^{\mathrm{E}}$	$\mathrm{MDE}^{\mathrm{NSD}}$
RW	765.81	0.5901
ES	765.91	0.5902
ARIMA	761.30	0.5899
TAR	757.83	0.5482
eTS+	392.01	0.4483
ePL+	358.19	0.4299
eFuMo	340.87	0.3701
ePL-KRLS	**321.50**	**0.3163**

Additionally, intervals descriptive statistics, coverage (R^{C}) and efficiency (R^{E}) rates, are reported in Table IV. They provide additional information about the adequacy of the forecasts. In this case, these statistics reveal the percentage of the actual ITS is covered by its forecast (coverage), and what part of the forecast covers the realized ITS (efficiency). To evaluate the models, these rates must be considered jointly and the higher their values the better is the forecast. Further, the closeness of the results of these two statistics can be taken as an indicator of the quality of the forecasts.

For both coverage and efficiency rates, the ePL-KRLS model showed the highest values, indicating more accurate predictions (Table IV). The other evolving approaches also provided similar results. Note that these two statistics values are very close, which corroborate the adequacy of the models. Again, the econometric forecasting methods performed worst.

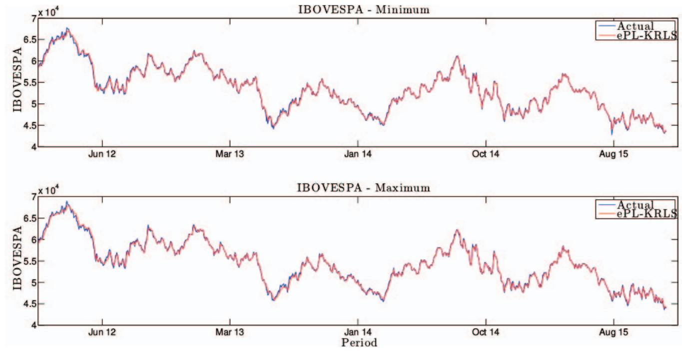

Fig. 2. IBOVESPA minimum and maximum actual and predicted values by ePL-KRLS for the period from January 2012 to December 2015.

TABLE IV
COVERAGE (R^{C}) AND EFFICIENCY (R^{E}) RATES VALUES BASED ON ONE STEP AHEAD FORECASTS OF IBOVESPA ITS FOR OUT-OF-SAMPLE DATA IN THE PERIOD FROM JANUARY 2012 TO DECEMBER 2015. BEST RESULTS ARE IN BOLD.

Models	R^{C}	R^{E}
RW	0.5747	0.5565
ES	0.5746	0.5564
ARIMA	0.5809	0.5654
TAR	0.6173	0.5783
eTS+	0.7431	0.7384
ePL+	0.7510	0.7499
eFuMo	0.7578	0.7411
ePL-KRLS	**0.7961**	**0.7859**

Summing up, the empirical results indicate the adequacy of ePL-KRLS for IBOVESPA interval time series forecasting. According to traditional and interval quality measures, in general, suggested approach showed better performance. Figure 2 illustrates the actual and forecasted values by the ePL-KRLS model of IBOVESPA lower (minimum) and upper (maximum) bounds for the out-of-sample data. One may note the high capability of the model to deal with nonlinear and time-varying dynamics. In line with [8], the results indicate that the contribution of nonlinear models to a good forecast performance is significant, also for the Brazilian equity market.

V. CONCLUSION

Interval time series (ITS) are time series where each period in time is described by an interval. In finance, ITS can be described as the evolution of the maximum and minimum prices of an asset throughout time. Their accurate forecasts play a key role in risk management, derivatives pricing, and asset allocation. This paper suggests an evolving participatory learning fuzzy rule-based model, named ePL-KRLS, for financial ITS. The model extends tradition evolving participatory learning (ePL) method by considering nonparametric approaches for updating the model rule consequents using

Kernel functions and an adaptive mechanism for models control parameters setting. As the empirical application, the main Brazilian stock market index, IBOVESPA, is considered, since there is no study in the current literature advocating the predictability of ITS in emergent economies, like Brazil. One step ahead forecasts is evaluated using data from January 2012 to December 2015 as the out-of-sample set. ePL-KRLS is compared against traditional linear and nonlinear time series benchmarks such as random walk, exponential smoothing, autoregressive integrated moving average, and threshold autoregressive methods, and with the state of the art evolving fuzzy techniques. Comparisons are conducted in terms of traditional error measures and also considering accuracy metrics designed for intervals. Results indicate the high predictability of IBOVESPA ITS by the ePL-KRLS method in comparison with the alternative approaches. Future works shall include the extension of the suggested method to model interval-valued data as a multivariate approach and its comparison against interval techniques.

ACKNOWLEDGMENT

The authors thank the Brazilian Ministry of Education (CAPES), the Brazilian National Council for Scientific and Technological Development (CNPq), and the Research of Foundation of the State of São Paulo (FAPESP) for their support.

REFERENCES

[1] D. Pettenuzzo, A. Timmermann, and R. Valkanov, "Forecasting stock returns under economic constraints," *Journal of Financial Economics*, vol. 144, no. 3, pp. 517–553, 2014.

[2] J. Arroyo, R. Espínola, and C. Maté, "Different approaches to forecast interval time series: A comparison in finance," *Computational Economics*, vol. 27, no. 2, pp. 169–191, 2011.

[3] R. F. Engle and J. Russel, "Analysis of high frequency data," in *Handbook of financial econometrics, Vol. 1: Tools and techniques*, Y. Aït-Sahalia and L. P. Hansen, Eds. Amsterdam: North Holland, 2009, pp. 383–346.

[4] W. Lu, X. Chen, W. Pedrycz, X. Liu, and J. Yang, "Using interval information granules to improve forecasting in fuzzy time series," *International Journal of Approximate Reasoning*, vol. 57, pp. 1–18, 2015.

[5] W. Froelich and J. L. Salmeron, "Evolutionary learning of fuzzy grey cognitive maps for the forecasting of multivariate, interval-valued time series," *International Journal of Approximate Reasoning*, vol. 55, no. 6, pp. 1319–1335, 2014.

[6] T. Xiong, Y. Bao, Z. Hu, and R. Chiong, "Forecasting interval time series using a fully complex-valued rbf neural network with dpso and pso algorithms," *Information Sciences*, vol. 305, pp. 77–92, 2015.

[7] L. Wang, X. Liu, and W. Pedrycz, "Effective intervals determined by information granules to improve forecasting in fuzzy time series," *Expert Systems with Applications*, vol. 40, no. 14, pp. 5673–5679, 2013.

[8] P. M. M. Rodrigues and N. Salish, "Modeling and forecasting interval time series with threshold models," *Advances in Data Analysis and Classification*, vol. 9, no. 1, pp. 41–57, 2015.

[9] S. Alizadeh, M. W. Brandt, and F. X. Diebold, "Range-based estimation of stochastic volatility," *The Journal of Finance*, vol. 57, no. 3, pp. 1047–1091, 2002.

[10] M. Parkinson, "The extreme value method for estimating the variance of the rate of return," *The Journal of Business*, vol. 53, no. 1, pp. 61–65, 1980.

[11] M. W. Brandt and F. X. Diebold, "A no-arbitrage approach to range-based estimation of return covariances and correlations," *The Journal of Business*, vol. 79, no. 1, pp. 61–74, 2006.

[12] J. H. Shu and J. E. Zhang, "Testing range estimators of historical volatility," *Journal of Future Markets*, vol. 26, no. 3, pp. 297–313, 2006.

[13] R. Y. Chou, "Forecasting financial volatilities with extreme values: The conditional autoregressive range (CARR) model," *Journal of Money, Credit and Banking*, vol. 37, no. 3, pp. 561–582, 2005.

[14] N. M. Fiess and R. MacDonald, "Towards the fundamentls of technical analysis: Analysing the information content of high, low and close prices," *Economic Modelling*, vol. 19, no. 3, pp. 353–374, 2002.

[15] E. A. Lima and F. A. T. d. Carvalho, "Center and range method for fitting a linear regression model to symbolic interval data," *Computational Statistics & Data Analysis*, vol. 52, no. 3, pp. 1500–1515, 2008.

[16] Y. W. Cheung, "An empirical model of dialy highs and lows," *International Journal of Finance & Economics*, vol. 12, no. 1, pp. 1–20, 2007.

[17] C. Hu and L. T. He, "An application of interval methods to stock marketing forecasting," *Reliable Computing*, vol. 13, no. 5, pp. 423–434, 2007.

[18] L. T. He and C. Hu, "Impacts of interval measurement on studies of economic variability: Evidence from stock market variability forecasting," *Journal of Risk Finance*, vol. 8, no. 5, pp. 489–507, 2008.

[19] ——, "Impacts of interval computing on stock market variability forecasting," *Computational Economics*, vol. 33, no. 3, pp. 263–276, 2009.

[20] W. Yang, A. Han, and S. Wang, "Forecasting financial volatility with interval-valued time series data," *Vulnerability, Uncertainty, and Risk*, pp. 1124–1233, 2014.

[21] T. Xiong, Y. Bao, and Z. Hu, "Multiple-output support vector regression with a firefly algorithm for interval-valued stock price index forecasting," *Knowledge-Based Systems*, vol. 55, pp. 87–100, 2014.

[22] P. Angelov and D. Filev, "An approach to online identification of Takagi-Sugeno fuzzy models," *IEEE Transactions on Systems Man and Cybernetics – Part B*, vol. 34, no. 1, pp. 484–498, 2004.

[23] ——, "Simpl_eTS: A simplified method for learning evolving Takagi-Sugeno fuzzy models," *IEEE International Conference on Fuzzy Systems, Reno, Nevada, USA*, pp. 1068–1073, 2005.

[24] P. Angelov and X. Zhou, "Evolving fuzzy systems from data streams in real-time," *International Symposium on Evolving Fuzzy Systems, Ambleside, Lake District, United Kingdom*, pp. 29–35, 2006.

[25] P. Angelov, "Evolving Takagi-Sugeno fuzzy systems from data streams (eTS+)," in *Evolving intelligent systems: Methodology and applications*, P. Angelov, D. Filev, and N. Kasabov, Eds. Hoboken, NJ, USA: Wiley & IEEE Press, 2010, pp. 21–50.

[26] L. Maciel, F. Gomide, and R. Ballini, "Enhanced evolving participatory learning fuzzy modeling: An application for asset returns volatility forecasting," *Evolving Systems*, vol. 5, no. 1, pp. 75–88, 2014.

[27] E. Lima, H. Hell, R. Ballini, and F. Gomide, *Evolving Fuzzy Modeling Using Participatory Learning*, P. Angelov, D. P. Filev, and N. Kasabov, Eds. John Wiley & Sons, Ltd, 2010.

[28] D. Dovžan and I. Škrjanc, "Recursive fuzzy c-means clustering for recursive fuzzy identification of time-varying processes," *ISA Transactions*, vol. 50, no. 2, pp. 159–169, 2011.

[29] ——, "Recursive clustering based on a Gustafson-Kessel algorithm," *Evolving Systems*, vol. 2, no. 1, pp. 15–24, 2011.

[30] D. Dovžan, V. Loga, and I. Škrjanc, "Solving the sales prediction with fuzzy evolving models," in *WCCI 2012 IEEE World Congress on Computational Intelligence*, June, Brisbane, Australia, 2012, pp. 10–15.

[31] E. Lughofer, "Evolving fuzzy systems – fundamentals, reliability, interpretability and useability," in *Handbook of Computational Intelligence*, P. Angelov, Ed. World Scientific, 2016, pp. 67–137.

[32] Y. Engel, S. Mannor, and R. Meir, "The kernel recursive least-squares algorithm," *IEEE Transactions on Signal Processing*, vol. 52, no. 8, pp. 2275–2285, 2004.

[33] S. Shafieezadeh-Abadeh and A. Kalhor, "Evolving takagi–sugeno model based on online gustafson-kessel algorithm and kernel recursive least square method," *Evolving Systems*, vol. 7, no. 1, pp. 1–14, 2016.

[34] P. Festa and M. G. Resende, "Grasp: An annotated bibliography," in *Essays and surveys in metaheuristics*. Springer, 2002, pp. 325–367.

[35] R. E. Moore, R. B. Kearfott, and M. J. Cloud, *Introduction to interval analysis*. Philadelphia: SIAM Press, 2009.

[36] A. Bera and C. Jarque, "Efficient tests for normality, homocedasticity and serial independence of regression residuals: Monte Carlo evidence," *Economics Letter*, vol. 7, pp. 313–318, 1981.