Stock Market Volatility Prediction Using Possibilistic Fuzzy Modeling

Leandro Maciel and Fernando Gomide
School of Electrical and Computer Engineering
University of Campinas
Campinas, São Paulo, Brazil
Emails: {maciel,gomide}@dca.fee.unicamp.br

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

***Abstract*—This paper addresses stock market assets return volatility forecasting and possibilistic fuzzy modeling. A recursive possibilistic fuzzy modeling (rPFM) approach is suggested to deal with the identification of systems affected by outliers and noisy data due to the use of memberships and typicalities to cluster data. Since financial markets are affected by news, expectations and investors psychology, the development of robust methodologies such as rPFM is essencial for market participants. The model is evaluated for volatility forecasting with jumps using intraday data of different stock markets. Results indicate that rPFM is a potential tool for volatility forecasting and outperforming some traditional recursive fuzzy and neural fuzzy models.**

I. Introduction

In finance and economics, volatility modeling and forecasting play an importante role in derivative securities pricing, portfolio and risk management, investment analysis, and hedge operations [1]. Several studies have focused on daily returns in forecasting daily volatility as well as the use of daily squared returns as a measure of "true volatility" [2]. Daily squared returns are obtained from closing prices and cannot capture price fluctuations during the day. To overcome this limitation, Andersen and Bollerslev [3] proposed realized volatility as a measure of true volatility. Realized volatility is computed as the sum of squared high-frequency returns within a day, that conveniently covers more information on daily transactions.

The growing availability of financial market data at intraday frequencies has led to the development of improved volatility measurements. The recent literature advocates using realized variation measures to improve gains in asset return volatility forecasting [4]. Furthermore, studies have shown the importance of explicitly allowing for jumps in the estimation of volatility models and in the pricing of options and other derivatives [5], [6].

Recently, Barndorff-Nielson and Shepard [7] introduced a fully nonparametric separation of the continuous sample path and jump components of realized volatility, combining the so-called bipower variation measures constructed from the summation of appropriately scaled cross-products of adjacent high-frequency absolute returns. According to the authors, the two components play different roles in forecasting and significant gains in performance are achieved when they are used separately.

Within this framework, Andersen et al. [8] used lagged realized volatility sample paths and jumps as separate regressors for volatility forecasting. They showed that, based on high-frequency prices from the foreign exchange, equity, and fixed-income markets, the inclusion of jumps works well empirically and that jumps dynamics are much less persistent (and predictable) than sample path dynamics. In addition, their model relies on a linear regression of lagged sample paths and jumps that are not observable in practice due to the complex and noisy nonlinear dynamical volatility behavior.

To overcome this limitation, recent studies have suggested the use of evolving fuzzy and neuro-fuzzy systems to model and forecast the realized volatility of financial asset returns, including lagged sample path and jump components, from high-frequency data. [9] and [10] suggested the use of a cloud-based evolving fuzzy model (eCloud) and a hybrid neural fuzzy network (eHFN) for volatility forecasting with jumps, respectively. The authors showed the high potential of evolving fuzzy modeling approaches in terms of accuracy, outperforming traditional econometric techniques.

This paper suggests a recursive possibilistic fuzzy modeling approach (rPCM), proposed by Maciel et al. [11], for volatility forecasting with jumps. The possibilistic idea acts as a mechanism to improve model robustness against noise and outlier data when model structure identification is performed by fuzzy clustering [12]. Since the financial markets are affected by news, expectations, and investors psychology, these events can lead in higher price variation not explained by economic factors, which may be seen as a noisy or outlier information to the financial system. In this case, the possibilistic framework is a potential tool to deal with assets volatility modeling. Computational experiments uses data of 1-minute returns from December 2009 through July 2013 for the main equity market indexes in the global markets: S&P 500 and Nasdaq (United States), FTSE (United Kingdom), DAX (Germany), IBEX (Spain) and Ibovespa (Brazil). Using goodness of fit, the performance of the recursive possibilistic fuzzy model is compared with linear regression suggested by [8] for realized volatility with jumps, multi layer feedforward neural network, and models representative of the current state of the art in evolving fuzzy modeling.

The paper is organized as follows. Section II briefly reviews the nonparametric approach for realized volatility with jumps. Section III describes the possibilistic fuzzy modeling approach. The empirical results and analysis are reported in Section IV. Finally, Section V concludes the paper and suggests issues for future research.

978-1-4673-8418-6/15/$31.00 ©2015 IEEE

II. REALIZED VOLATILITY WITH JUMPS

Consider a continuous-time jump diffusion process, expressed in stochastic differential equation form as follows

$$dp(t) = \mu(t)dt + \sigma(t)dW(t) + \kappa(t)dq(t), \quad 0 \leq t \leq T, \quad (1)$$

where $p(t)$ is logarithm of the asset price at time t; $\mu(t)$ is the continuous and locally bounded variation process; $\sigma(t)$ is a strictly positive stochastic volatility process with a right continuous sample path, well-defined left limits, which allows for occasional jumps in volatility; $W(t)$ is the standard Brownian motion; $q(t)$ is the counting process with time-varying intensity $L(t)$, that is, $P[dq(t) = 1] = L(t)dt$; and $\kappa(t)$ corresponds to the sizes of the discrete jumps in the price logarithm process. The quadratic variation for the cumulative return process, $r(t) = p(t) - p(0)$, is

$$r(t, 0) = \int_0^t \sigma^2(s)ds + \sum_{0<s\leq t} \kappa^2(s). \quad (2)$$

We define the daily realized volatility, or variation, as the sum of the corresponding $1/\Delta$ high-frequency intradaily squared returns:

$$RV_{t+1}(\Delta) = \sum_{j=1}^{1/\Delta} r^2_{t+j\cdot\Delta,\Delta}, \quad (3)$$

where $r_{t,\Delta} = p(t) - p(t - \Delta)$ is the discrete sample of the Δ-period return.

Accordingly to the theory of quadratic variation, the realized variation converges uniformly in probability to the increment of the quadratic variation process as the sampling frequency of the underlying returns increases [1], [8]:

$$RV_{t+1}(\Delta) \to \int_t^{t+1} \sigma^2(s)ds + \sum_{t<s\leq t+1} \kappa^2(s), \quad (4)$$

for $\Delta \to 0$.

Realized volatility inherits the dynamics of both, the continuous sample path, and the jump processes. One can directly model and forecast realized volatility without distinguishing jumps or applying nonjump contributions. However, literature suggests that superior forecasts are achieved measuring and modeling the two components of realized volatility separately [8]. This paper considers a nonparametric framework developed by [7], which identifies the two components of the quadratic volatility process based on the concept of the standardized realized bipower variation theory:

$$BV_{t+1}(\Delta) = \mu_1^{-2} \sum_{j=2}^{1/\Delta} |r_{t+j\cdot\Delta,\Delta}||r_{t+(j-1)\cdot\Delta,\Delta}|, \quad (5)$$

where $\mu_1 = \sqrt{(2/\pi)} = E(|Z|)$ is the mean of the absolute value of the standard normally distributed random variable, Z. According to [7], as $\Delta \to 0$,

$$BV_{t+1}(\Delta) \to \int_t^{t+1} \sigma^2(s)ds. \quad (6)$$

Therefore, from Eqs. (4) and (6), as shown in [7], the contribution to the quadratic variation process due to the jumps (discontinuities) is consistently estimated as $\Delta \to 0$:

$$RV_{t+1}(\Delta) - BV_{t+1}(\Delta) \to \sum_{t<s\leq t+1} \kappa^2(s). \quad (7)$$

To avoid negative values of the right side of Eq. (7) for a given finite sample, the empirical measurements at zero are truncated. Truncation ensures that all daily estimates are nonnegative [7]:

$$J_{t+1}(\Delta) = \max[RV_{t+1}(\Delta) - BV_{t+1}(\Delta), 0]. \quad (8)$$

This paper employs the Heterogeneous Autoregressive approach developed by [8] for the realized volatility (HAR-RV) formulation as benchmark. The HAR-RV approach is based on an extension of the heterogeneous ARCH, or HARCH, class of models studied in [13]. The conditional variance is parametrized as a linear function of the lagged squared returns over the identical return horizon, together with the squared returns over longer and/or shorter return horizons [8]. The HAR-RV model considers the multi period normalized realized variation

$$RV_{t,t+h} = \frac{1}{h} \sum_{i=1}^{h} RV_{t+i}. \quad (9)$$

Thus, the HAR-RV-J specification of [8] includes jump (J) components. It is expressed as

$$\begin{aligned} RV_{t+1} \quad = \quad & \beta_0 + \beta_D RV_t + \beta_W RV_{t-5,t} + \beta_M RV_{t-22,t} \\ & + \beta_J J_t + \varepsilon_{t+1}, \quad (10) \end{aligned}$$

where $h = 5$ and $h = 22$ refer to the normalized measures of weekly and monthly volatilities, respectively. ε is a white noise process, and $t = 1, 2, \ldots, T$.

III. RECURSIVE POSSIBILISTIC FUZZY MODELING

Construction of Takagi-Sugeno (TS) fuzzy rule-based models requires the identification of its structure and estimation of the parameters of the local affine models [14]. Structure identification can be done by using fuzzy clustering techniques, and rule consequent parameters can be estimated using a least squares algorithm [15].

The fuzzy c-means (FCM) [16] is one of the most widely used clustering algorithm. FCM may support both, off-line and on-line TS modeling. It is well known that the FCM algorithm is sensitive to outliers and noisy data. To overcome this problem, Krishnapuram and Keller [17] suggests a new clustering algorithm model called possibilistic c-means (PCM). The idea of PCM is to relax the unity membership degrees sum constraint of FCM. Furthermore, Pal et al. [12] introduced a possibilistic fuzzy c-means (PFCM) model. PFCM simultaneously produces memberships and typicalities through a hybridization of the possibilistic c-means (PCM) and the fuzzy c-means (FCM).

Recently, Maciel et al. [11] suggested a recursive possibilistic fuzzy c-means modeling (rPFM) approach for TS modeling. The rPFM uses a recursive form of the PFCM to update the model structure, and employs the recursive least squares algorithm to estimate the parameters of affine functions of the rule consequent. Computational experiments using benchmark data show that recursive possibilistic fuzzy modeling produces parsimonious models with comparable or better accuracy than alternative evolving fuzzy approaches. Thus, this paper addresses recursive possibilistic fuzzy modeling and volatility forecasting with jumps in order to account properly the financial markets movements reflected by noisy and outlier data. The following sections describe the main constructs of rPFM modeling approach.

A. Takagi-Sugeno fuzzy model

Takagi-Sugeno (TS) fuzzy model consists of a set of fuzzy functional rules of the following form:

$$\mathcal{R}_i: \text{ IF } \mathbf{x} \text{ is } \mathcal{A}_i \text{ THEN } y_i = \theta_{i0} + \theta_{i1}x_1 + \ldots + \theta_{im}x_m, \quad (11)$$

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 and its membership function $\phi_{\mathcal{A}_i}(\mathbf{x}) : \Re \rightarrow [0, 1]$, $y_i \in \Re$ 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 (11) may have a closed form as follows:

$$y = \sum_{i=1}^{c} \left(\frac{\phi_{\mathcal{A}_i}(\mathbf{x}) y_i}{\sum_{j=1}^{c} \phi_{\mathcal{A}_j}(\mathbf{x})} \right). \quad (12)$$

The expression (12) can be rewritten using normalized degree of activation:

$$y = \sum_{i=1}^{c} \lambda_i y_i = \sum_{i=1}^{c} \lambda_i \mathbf{x}_e^T \theta_i, \quad (13)$$

where

$$\lambda_i = \frac{\phi_{\mathcal{A}_i}(\mathbf{x})}{\sum_{j=1}^{c} \phi_{\mathcal{A}_j}(\mathbf{x})}, \quad (14)$$

is the normalized firing 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.

Usually, a TS model uses parametrized fuzzy regions and associates each region with an affine (local) model. The non-linear 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 the degree of firing of each rule.

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

B. Possibilistic fuzzy c-means clustering

This section gives a brief overview of the possibilistic fuzzy c-means clustering algorithm [12]. Let $\mathbf{x}_k^T = [x_{1k}, x_{2k}, \ldots, x_{mk}] \in \Re^m$ be the input data at step k. A set of n inputs is denoted by $X = \{\mathbf{x}_k | k = 1, 2, \ldots, r, \ldots, n\}$, $X \subset \Re^{m \times n}$. The aim of clustering is to partition the data set X into c subsets (clusters).

A possibilistic fuzzy partition of the set X is a family $\{\mathcal{A}_i | 1 \leq i \leq c\}$. Each $\mathcal{A}_i$ is viewed as possibilistic fuzzy subset characterized by a membership function and typicality (possibility distribution) specified by the fuzzy and typicality partition matrices $U = [u_{ik}] \in \Re^{c \times n}$ and $T = [t_{ik}] \in \Re^{c \times n}$, respectively. The entries of the i-th row of matrix U (T) are the values of membership (typicalities) degrees of the data matrix X in $\mathcal{A}_i$.

The possibilistic fuzzy c-means (PFCM) clustering algorithm produces c vectors that represent c groups in the data set. The PFCM algorithm derives from the solution of the following optimization problem:

$$\min_{U,T,V} \left\{ J = \sum_{k=1}^{n} \sum_{i=1}^{c} (a u_{ik}^{\eta_f} + b t_{ik}^{\eta_p}) ||\mathbf{x}_k - \mathbf{v}_i||_A^2 + \sum_{i=1}^{c} \gamma_i \sum_{k=1}^{n} (1 - t_{ik})^{\eta_p} \right\}, \quad (15)$$

subject to

$$\sum_{i=1}^{c} u_{ik} = 1 \ \forall\ k, \quad 0 \leq u_{ik} \leq 1, \quad 0 \leq t_{ik} \leq 1. \quad (16)$$

Here $a > 0$, $b > 0$, and $\eta_f, \eta_p > 1$, $\gamma_i > 0$ are user defined parameters, and $|| \cdot ||_A$ is a norm. The constants a and b define the relative importance of fuzzy membership and typicality values in the objective function, respectively. $V = [\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_c]^T \in \Re^{c \times m}$ is the matrix of cluster centers, η_f and η_p are parameters associated with membership degrees and typicalities, respectively.

If $D_{ikA} = ||\mathbf{x}_k - \mathbf{v}_i||_A > 0$ for all i, and X contains at least c distinct data points, then $(U, T, V) \in M_f \times M_p \times \Re^{c \times n}$ minimizes J, with $1 \leq i \leq c$ and $1 \leq k \leq n$, only if [12]:

$$u_{ik} = \left(\sum_{j=1}^{c} \left(\frac{D_{ikA}}{D_{jkA}} \right)^{2/(\eta_f - 1)} \right)^{-1}, \quad (17)$$

$$t_{ik} = \frac{1}{1 + \left(\frac{b}{\gamma_i} D_{ikA}^2 \right)^{1/(\eta_p - 1)}}, \quad (18)$$

$$\mathbf{v}_i = \frac{\sum_{k=1}^{n} (a u_{ik}^{\eta_f} + b t_{ik}^{\eta_p}) \mathbf{x}_k}{\sum_{k=1}^{n} (a u_{ik}^{\eta_f} + b t_{ik}^{\eta_p})}, \quad (19)$$

where

$$M_p = \left\{ T \in \Re^{c \times n} : 0 \leq t_{ik} \leq 1, \forall\ i, k; \forall\ k\ \exists\ i \ni t_{ik} > 0 \right\}, \quad (20)$$

$$M_f = \left\{ U \in M_p : \sum_{i=1}^{c} u_{ik} = 1\ \forall\ k; \sum_{k=1}^{n} u_{ik} > 0\ \forall\ i \right\}, \quad (21)$$

are the sets of possibilistic and fuzzy partition matrices, respectively.

Originally, [12] recommends to choose parameters γ_i as follows:

$$\gamma_i = K\frac{\sum_{k=1}^{n} u_{ik}^{\eta_f} D_{ikA}^2}{\sum_{k=1}^{n} u_{ik}^{\eta_f}}, \qquad 1 \leq i \leq c, \tag{22}$$

where $K > 0$ (usually $K = 1$), and u_{ik} are entries of a terminal FCM partition of X.

C. Recursive possibilistic fuzzy c-means

This section develops the recursive possibilistic fuzzy c-means (rPFM) clustering algorithm, and a mechanism to account for the cluster shapes similarly as in Gustafson-Kessel (GK) clustering as in [11].

Let $\mathbf{v}_{ir} = [v_{1,ir}, v_{2,ir}, \ldots, v_{m,ir}]^T \in \Re^m$ be the cluster center of the i-th rule at step r, using (19) to compute the current membership and typicality degrees. The cluster center after new data is input at $r+1$ is:

$$\begin{aligned}\mathbf{v}_{i,r+1} &= \frac{\sum_{k=1}^{r+1}(au_{ik}^{\eta_f} + bt_{ik}^{\eta_p})\mathbf{x}_k}{\sum_{k=1}^{r+1}(au_{ik}^{\eta_f} + bt_{ik}^{\eta_p})} \\ &= \frac{\sum_{k=1}^{r}(au_{ik}^{\eta_f} + bt_{ik}^{\eta_p})\mathbf{x}_k + (au_{i,r+1}^{\eta_f} + bt_{i,r+1}^{\eta_p})\mathbf{x}_{r+1}}{\sum_{k=1}^{r}(au_{ik}^{\eta_f} + bt_{ik}^{\eta_p}) + (au_{i,r+1}^{\eta_f} + bt_{i,r+1}^{\eta_p})}.\end{aligned} \tag{23}$$

Let us denote the relation between the previous cluster center and the new one by:

$$\mathbf{v}_{i,r+1} = \mathbf{v}_{ir} + \Delta\mathbf{v}_{i,r+1}, \tag{24}$$

$\Delta\mathbf{v}_{i,r+1}$ indicates the variation in the cluster center due to new input data $\mathbf{x}_{r+1}$.

According to [11], using (23), (24) can be rewritten as follows:

$$\mathbf{v}_{i,r+1} = \mathbf{v}_{ir} + \frac{(au_{i,r+1}^{\eta_f} + bt_{i,r+1}^{\eta_p})(\mathbf{x}_{r+1} - \mathbf{v}_{ir})}{\sum_{k=1}^{r}(au_{ik}^{\eta_f} + bt_{ik}^{\eta_p}) + (au_{i,r+1}^{\eta_f} + bt_{i,r+1}^{\eta_p})}. \tag{25}$$

Thus, the increment of the i-th cluster center $\Delta\mathbf{v}_{i,r+1}$ is:

$$\Delta\mathbf{v}_{i,r+1} = \frac{(au_{i,r+1}^{\eta_f} + bt_{i,r+1}^{\eta_p})(\mathbf{x}_{r+1} - \mathbf{v}_{ir})}{\sum_{k=1}^{r}(au_{ik}^{\eta_f} + bt_{ik}^{\eta_p}) + (au_{i,r+1}^{\eta_f} + bt_{i,r+1}^{\eta_p})}. \tag{26}$$

The denominator of cluster center increment (26) needs all the past r data to be computed. In principle, it can be computed recursively. However, in adaptive modeling and dynamic environments past data may become obsolete, and current data should be emphasized because they bring information about the current state. An exponential smoothing scheme to weight past membership and typicality degrees seems to be more appropriate.

Let the sum of the memberships and typicalities up to step r in the denominator of (26) be denoted by $s_{ik} \in \Re$:

$$s_{ir} = \sum_{k=1}^{r}(au_{ik}^{\eta_f} + bt_{ik}^{\eta_p}). \tag{27}$$

A way to upadte s_{ir} after new data is input at $r+1$ is as follows:

$$s_{i,r+1} = \gamma_v s_{ir} + au_{i,r+1}^{\eta_f} + bt_{i,r+1}^{\eta_p}, \tag{28}$$

The parameter γ_v, $0 \leq \gamma_v \leq 1$ is the forgetting factor of past membership degrees. Thus $\Delta\mathbf{v}_{i,r+1}$ can be rewritten as follows:

$$\Delta\mathbf{v}_{i,r+1} = \frac{(au_{i,r+1}^{\eta_f} + bt_{i,r+1}^{\eta_p})(\mathbf{x}_{r+1} - \mathbf{v}_{ir})}{s_{i,r+1}}. \tag{29}$$

Values of the membership degrees $u_{i,r+1}$ and typicalities $t_{i,r+1}$ of input data $\mathbf{x}_{r+1}$ at step $r+1$ are found using:

$$u_{i,r+1} = \left(\sum_{j=1}^{c}\left(\frac{D_{i,r+1,A}}{D_{j,r+1,A}}\right)^{2/(\eta_f-1)}\right)^{-1}, \tag{30}$$

$$t_{i,r+1} = \frac{1}{1 + \left(\frac{b}{\gamma_i} D_{i,r+1,A}^2\right)^{1/(\eta_p-1)}}. \tag{31}$$

D. Adaptive fuzzy dispersion matrix

Most of the fuzzy clustering algorithms assume clusters with spherical shapes. Actually, in real world applications clusters often have different shapes and orientations in the data space. A way to distinguish cluster shapes is to use information about the dispersion of the input data. We can supply rPFM modeling with a mechanism to handle cluster shapes and orientation similarly as in Gustafson-Kessel clustering [18] as follows. As in GK, the Euclidean norm is replaced by the Mahalanobis norm:

$$D_{ikA}^2 = (\mathbf{x}_k - \mathbf{v}_{ik})^T A_{ik}(\mathbf{x}_k - \mathbf{v}_{ik}), \qquad 1 \leq i \leq c, \tag{32}$$

where

$$A_{ik} = (\rho_i \det(F_{ik}))^{\frac{1}{m}} F_{ik}^{-1}, \tag{33}$$

is the cluster volume (usually $\rho_i = 1$ for all clusters), and $F_{ik} \in \Re^{m\times m}$ is the dispersion matrix.

The dispersion matrix for r input data is:

$$F_{ir} = \frac{\sum_{k=1}^{r}(au_{ik}^{\eta_f} + bt_{ik}^{\eta_p})(\mathbf{x}_{r+1} - \mathbf{v}_{ir})(\mathbf{x}_{r+1} - \mathbf{v}_{ir})^T}{\sum_{k=1}^{r}(au_{ik}^{\eta_f} + bt_{ik}^{\eta_p})}. \tag{34}$$

Therefore, the dispersion matrix after data $\mathbf{x}_{r+1}$ is input at step $r+1$ is:

$$\begin{aligned}F_{i,r+1} = &\frac{\sum_{k=1}^{r}(au_{ik}^{\eta_f} + bt_{ik}^{\eta_p})(\mathbf{x}_r - \mathbf{v}_{i,r+1})(\mathbf{x}_r - \mathbf{v}_{i,r+1})^T}{\sum_{k=1}^{r}(au_{ik}^{\eta_f} + bt_{ik}^{\eta_p}) + (au_{i,r+1}^{\eta_f} + bt_{i,r+1}^{\eta_p})} \\ &+ \frac{(au_{i,r+1}^{\eta_f} + bt_{i,r+1}^{\eta_p})(\mathbf{x}_{r+1} - \mathbf{v}_{i,r+1})(\mathbf{x}_{r+1} - \mathbf{v}_{i,r+1})^T}{\sum_{k=1}^{r}(au_{ik}^{\eta_f} + bt_{ik}^{\eta_p}) + (au_{i,r+1}^{\eta_f} + bt_{i,r+1}^{\eta_p})},\end{aligned} \tag{35}$$

where $\mathbf{v}_{i,r+1}$ is the cluster center of the i-th cluster computed after $\mathbf{x}_{r+1}$ is input at $r+1$.

From (34) and (35), and because of the same reasons as in the cluster center increment calculation, the dispersion matrix can be updated using exponential smoothing as follows:

$$F_{i,r+1} = \gamma_c \frac{s_{ir}}{s_{i,r+1}} F_{ir} + \frac{(au_{i,r+1}^{\eta_f} + bt_{i,r+1}^{\eta_p})}{s_{i,r+1}} (\mathbf{x}_{r+1} - \mathbf{v}_{i,r+1})^T, \quad (36)$$

where γ_c, $0 \leq \gamma_c \leq 1$, is the forgetting factor, and $F_{i0} \approx I \in \Re^{m \times m}$.

Actually, what is needed is the inverse of the dispersion matrix. The Woodbury inversion lemma gives [19], [11]:

$$F_{i,r+1}^{-1} = \frac{A - B}{C}, \quad (37)$$

where

$$A = \frac{1}{\gamma_c} \frac{s_{i,r+1}}{s_{ir}} F_{i,r}^{-1}, \quad (38)$$

$$B = \frac{1}{\gamma_c} \frac{s_{i,r+1}}{s_{ir}} F_{i,r}^{-1} \frac{(au_{i,r+1}^{\eta_f} + bt_{i,r+1}^{\eta_p})}{s_{i,r+1}} (\mathbf{x}_{r+1} - \mathbf{v}_{i,r+1}) (\mathbf{x}_{r+1} - \mathbf{v}_{i,r+1})^T \frac{1}{\gamma_c} \frac{s_{i,r+1}}{s_{ir}} F_{i,r}^{-1}, \quad (39)$$

$$C = 1 + (\mathbf{x}_{r+1} - \mathbf{v}_{i,r+1})^T \frac{1}{\gamma_c} \frac{s_{i,r+1}}{s_{ir}} F_{i,r}^{-1} \frac{(au_{i,r+1}^{\eta_f} + bt_{i,r+1}^{\eta_p})}{s_{i,r+1}} (\mathbf{x}_{r+1} - \mathbf{v}_{i,r+1}). \quad (40)$$

E. Consequent parameters estimation

The estimation of the parameters of the affine rule consequents is done using weighted recursive least squares algorithm (wRLS) as in [11].

IV. Computational results

A. Data sets

Source of data are those from the main equity market indexes in the global markets: S&P 500 and Nasdaq (United States), FTSE (United Kingdom), DAX (Germany), IBEX (Spain) and Ibovespa (Brazil). Data sets reflect distinct economies, from emergents to well developed countries. The purpose is to evaluate the forecasting performance of rPFM modeling in environments with different dynamics. The data sets are formed by 1-minute quotations from December 7, 2009, through July 31, 2013, for a total of 485,364 intraday observations.[1] High-frequency spot quotations are from Bloomberg. The stock market data produced a total of $1/\Delta = 507$ high-frequency return observations per day. Henceforth, reference to Δ is omitted, for short. This paper refers to the 1-minute realized volatility and jump measures as RV_t and J_t, respectively.

[1] The sample considered is justified by the increasing in high-frequency data availability in this period.

B. Methodology

We follow the nonparametric approach introduced in [1]. The realized volatility, given by Eq. (3), is computed using 1-minute equity indexes data. Based on bipower variation theory, the jump components of realized variation are obtained separately using Eqs. (6) and (8). We adopt the HAR-RV-J model [8] as a benchmark, namely

$$RV_{t+1} = \beta_0 + \beta_D RV_t + \beta_W RV_{t-5,t} + \beta_M RV_{t-22,t} + \beta_J J_t + \varepsilon_{t+1}.$$

For comparison purposes, volatility forecasts are done using alternative evolving fuzzy systems, such as eTS [20], xTS [21], eTS+ [22], eCloud [9], and eHFN [10]. Moreover, we consider also a multi layer feedforward neural network. The data set is divided into an in-sample, from December 7, 2009 through December 30, 2011, and an out-of-sample, assembled by all remaining data. The inputs for the rPFM, MLP and evolving fuzzy/neuro-fuzzy models are the same as the explanatory variables of the linear regression benchmark (RV_t, $RV_{t-5,t}$, $RV_{t-22,t}$, and J_t). The output of the models is the one-step ahead forecast of realized volatility RV_{t+1}.

Comparison of volatility forecasts uses the root mean squared error (RMSE), mean absolute deviation (MAD), and the mean average forecast error (MAFE):

$$\text{RMSE} = \frac{1}{T} \sum_{i=1}^{T} \left(RV_i - \hat{RV}_i \right)^2, \quad (41)$$

$$\text{MAD} = \frac{1}{T} \sum_{i=1}^{T} |RV_i - \hat{RV}_i|, \quad (42)$$

$$\text{MAFE} = \frac{1}{T} \sum_{i=1}^{T} \frac{|RV_i - \hat{RV}_i|}{RV_i}, \quad (43)$$

where T is the number of out-of-sample observations, RV_i is the actual realized volatility at period i, and $\hat{RV}_i$ is the forecast realized volatility at i.

C. Results and discussion

Volatility forecasts are compared in terms of goodness of fit using a linear regression (LR) [8], the recursive possibilistic fuzzy model (rPFM), a multi layer feedforward neural network (MLP), and the evolving fuzzy/neuro-fuzzy models eTS, xTS, eTS+, eCloud and eHFN.

The results summarized in Table I indicate the strong capability of the recursive possibilistic fuzzy model, the MLP and the evolving fuzzy/neuro-fuzzy models to model realized volatility forecasting. The LR gives the worst results, once the volatility is clearly nonlinear. One may see that for all error measures, the rPFM and the evolving models present similar results, especially rPFM and eHFN. Moreover, for all market indexes considered, the possibilistic approach performs significantly better than all remaining methodologies, except for IBEX index. The better results from rPFM may be due to its capability of dealing with noisy data and outliers. This becomes more evident if we look at the mean average forecast error (MAFE). The MLP performed slightly worse than the computational intelligence-based models due to its static, non adaptive nature. Considering the equity market indexes, the predictive accuracy of the models are similar despite their different volatility dynamics.

TABLE I. OUT-OF-SAMPLE ONE-STEP AHEAD REALIZED VOLATILITY FORECASTING PERFORMANCE

Model	S&P 500	Nasdaq	FTSE	DAX	IBEX	IBOV
RMSE						
LR	0.8726	1.0918	1.4450	0.9971	1.0117	1.1382
eTS	7.8E-06	4.4E-04	7.1E-01	2.0E-06	1.8E-06	4.0E-05
xTS	3.0E-06	2.0E-05	4.0E-04	1.1E-06	2.8E-06	1.2E-06
eTS+	2.3E-07	3.6E-07	9.7E-06	4.7E-07	4.1E-05	2.1E-06
MLP	4.7E-02	5.0E-01	6.3E-03	2.1E-02	1.0E-01	7.3E-04
eCloud	3.4E-07	2.7E-07	5.3E-08	2.7E-07	3.0E-06	5.5E-07
eHFN	1.8E-07	1.3E-07	4.5E-08	3.8E-07	**1.5E-06**	5.0E-07
rPFM	**4.1E-08**	**6.4E-08**	**2.0E-08**	**8.0E-08**	1.6E-06	**2.1E-07**
MAD						
LR	1.0092	0.7628	1.2422	1.0245	0.9887	1.1762
eTS	3.3E-02	2.8E-02	1.1E-03	1.1E-02	2.0E-03	1.1E-02
xTS	1.1E-02	2.2E-03	3.0E-03	2.7E-02	1.8E-04	3.5E-03
eTS+	4.2E-03	8.5E-04	4.6E-05	1.9E-03	2.2E-04	2.1E-03
MLP	2.1E-02	7.5E-01	3.8E-01	3.0E-02	5.7E-01	3.0E-02
eCloud	2.4E-04	2.2E-04	1.7E-04	2.6E-04	**5.6E-04**	7.1E-04
eHFN	1.7E-04	1.9E-04	1.2E-04	3.4E-04	6.7E-04	4.3E-04
rPFM	**1.1E-04**	**1.4E-04**	**8.3E-05**	**1.8E-04**	7.3E-04	**3.0E-04**
MAFE						
LR	7.826	8.092	11.715	9.012	7.072	8.664
eTS	2.981	3.165	5.505	2.800	2.362	1.992
xTS	2.873	2.902	5.452	2.783	2.374	1.872
eTS+	2.534	2.398	5.251	2.482	2.191	2.011
MLP	3.726	3.625	6.716	3.677	3.601	4.009
eCloud	2.450	2.316	5.029	2.784	2.156	1.882
eHFN	2.348	2.438	4.983	2.516	**1.933**	1.759
rPFM	**2.107**	**2.366**	**3.953**	**1.865**	3.114	**1.472**

V. CONCLUSION

This paper addresses recursive possibilistic fuzzy modeling (rPFM) and realized volatility forecasting with jumps. The approach combines recursive possibilistic fuzzy clustering to learn the model structure, and recursive least squares to estimate the model parameters. The possibilistic ideia improves model robustness to noisy and outlier data, which is appropriate for volatility dynamics, affected by news, expectations, and investors psychology. Computational experiments with main equity market indexes of global markets show the high capability of rPFM to model realized volatility forecasting and outperforms benchmark based on linear regression and some state of the art evolving fuzzy/neuro-fuzzy methods. Future work shall consider addition and removal of clusters, as well as the comparison in terms of computational cost and against GARCH-family models.

ACKNOWLEDGMENT

The authors thank the Brazilian Ministry of Education (CAPES), the Brazilian National Research Council (CNPq) grant 304596/2009-4, and the Research of Foundation of the State of São Paulo (FAPESP) for their support.

REFERENCES

[1] O. E. Barndorff-Nielson and N. Shephard, "Econometric analysis of realised volatility and its use in estimating stochastic volatility models," *Journal of the Royal Statistical Society*, vol. 64, no. 2, pp. 253–280, 2002.

[2] G. Chortareas, Y. Jiang, and J. C. Nankervis, "Forecasting exchange rate volatility using high-frequence data: Is the euro different?" *International Journal of Forecasting*, vol. 27, no. 4, pp. 1089–1107, 2011.

[3] T. G. Andersen and T. Bollerslev, "Answering the skeptics: Yes, standard volatility models do provide accurate forecasts," *International Economic Review*, vol. 39, no. 4, pp. 885–905, 1998.

[4] T. G. Andersen, T. Bollerslev, and N. Meddahi, "Correcting the errors: Volatility forecast evaluation using high-frequency data and realized volatilities," *Econometrica*, vol. 73, no. 1, pp. 279–296, 2005.

[5] D. S. Bates, "Post-'87 crash fears in the S&P 500 futures option market," *Journal of Econometrics*, vol. 94, no. 1–2, pp. 181–238, 2000.

[6] L. Khalaf, D. Saphores, and J. F. Bilodeau, "Simulation-based exact jump tests in models with conditional heteroskedasticity," *Journal of Economic Dynamics and Control*, vol. 28, no. 3, pp. 531–553, 2003.

[7] O. E. Barndorff-Nielson and N. Shephard, "Econometrics of testing for jumps in financial economics using bipower variation," *Journal of Empirical Finance*, vol. 4, no. 1, pp. 1–30, 2006.

[8] T. G. Andersen, T. Bollerslev, and F. X. Diebold, "Roughing it up: Including jump components in the measurement, modeling, and forecasting of return volatility," *The Review of Economics and Statistics*, vol. 89, pp. 701–720, 2007.

[9] L. Maciel, F. Gomide, R. Ballini, and R. Yager, "Simplified evolving rule-based fuzzy modeling of realized volatility forecasting with jumps," *IEEE Workshop on Computational Intelligence for Financial Engineering and Economics (CIFEr 2013), Cingapura*, pp. 76–83, 2013.

[10] R. Rosa, L. Maciel, F. Gomide, and R. Ballini, "Evolving hybrid neural fuzzy network for realized volatility forecasting with jumps," *IEEE Workshop on Computational Intelligence for Financial Engineering and Economics (CIFEr 2014), Londres*, vol. 1, pp. 1–8, 2014.

[11] L. Maciel, F. Gomide, and R. Ballini, "Recursive possibilistic fuzzy modeling," in *Evolving and Autonomous Learning Systems (EALS), IEEE Symposium Series on Computation Intelligence (SSCI)*, Orlando, FL, 2014, pp. 9–16.

[12] N. R. Pal, K. Pal, J. M. Keller, and J. C. Bezdek, "A possibilistic fuzzy c-means clustering algorithm," *IEEE Transactions on Fuzzy Systems*, vol. 13, no. 4, pp. 517–530, 2005.

[13] U. A. MÜLLER, M. M. Dacorogna, R. D. Davé, R. B. Olsen, O. V. Puctet, and J. Weizsäcker, "Volatilities of different time resolutions – Analyzing the dynamics of markets components," *Journal of Empirical Finance*, vol. 4, no. 2–3, pp. 213–239, 1997.

[14] T. Takagi and M. Sugeno, "Fuzzy identification of systems and its applications to modeling and control," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-15, no. 1, pp. 116–132, 1985.

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

[16] J. C. Bezdek, *Pattern recognition with fuzzy objective function algorithm*. New York: Plenum, 1981.

[17] R. Krishnapuram and J. Keller, "A possibilistic approach to clustering," *IEEE Transactions on Fuzzy Systems*, vol. 2, no. 1, pp. 98–110, 1993.

[18] E. E. Gustafson and W. C. Kessel, "Fuzzy clustering with a fuzzy covariance matrix," *IEEE International Conference on Decision and Control*, pp. 761–766, 1979.

[19] D. Dovžan and I. Škrjanc, "Recursive clustering based on a Gustafson-Kessel algorithm," *Evolving Systems*, vol. 2, no. 1, pp. 15–24, 2011.

[20] P. Angelov and R. Buswell, "Identification of evolving rule-based models," *IEEE Transactions on Fuzzy Systems*, vol. 10, pp. 667–677, 2002.

[21] P. Angelov and X. Zhou, "Evolving fuzzy-rule-based classifiers from data streams," *IEEE Transactions on Fuzzy Systems*, vol. 16, no. 6, pp. 1462–1475, 2008.

[22] 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.