

Enhanced evolving participatory learning fuzzy modeling: an application for asset returns volatility forecasting

Leandro Maciel · Fernando Gomide ·
Rosangela Ballini

Received: 10 February 2013 / Accepted: 26 September 2013 / Published online: 22 November 2013
© Springer-Verlag Berlin Heidelberg 2013

Abstract Evolving participatory learning (ePL) modeling joins the concepts of participatory learning and evolving fuzzy systems. It uses data streams to continuously adapt the structure and functionality of fuzzy models. This paper suggests an enhanced version of the ePL approach, called ePL+, which includes both an utility measure to shrink rule bases, and a variable cluster radius mechanism to improve the cluster structure. These features are useful in adaptive fuzzy rule-based modeling to recursively construct local fuzzy models with variable zone of influence. Moreover, ePL+ extends ePL to multi-input, multi-output fuzzy system modeling. Computational experiments considering financial returns volatility modeling and forecasting are conducted to compare the performance of the ePL+ approach with state of the art fuzzy modeling methods and with GARCH modeling. The experiments use actual data of S&P 500 and Ibovespa stock market indexes. The results suggest that the ePL+ approach is highly capable to model volatility dynamics, in a robust, flexible, parcimonious, and autonomous way.

Keywords Adaptive systems · Participatory learning · GARCH models · Forecasting · Volatility

L. Maciel · F. Gomide
School of Electrical and Computer Engineering, University of
Campinas, Campinas 13083-852, Brazil
e-mail: maciel@dca.fee.unicamp.br

F. Gomide
e-mail: gomide@dca.fee.unicamp.br

R. Ballini (✉)
Institute of Economics, University of Campinas,
Campinas 13083-857, Brazil
e-mail: ballini@eco.unicamp.br

1 Introduction

Real-world problem solving and decision making relies on information management, involving high-volume and complex non-stationary data streams. When handling streaming data, data processing systems must extract meaningful knowledge on line. In practice it is virtually impossible to store all data and batch methods may become unsuitable (Angelov and Zhou 2006). In the last decade, data-driven rule/knowledge extraction methods have emerged as a complementary approaches to develop intelligent systems to deal with data streams (Attar et al. 2010; Shaker and Hüllermeier 2012). They assemble a class of adaptive data processing and modeling systems (Yager and Filev 1994; Kasabov 1996; Ljung 1999). Evolving fuzzy systems (eFS) are an advanced form of adaptive systems because they have the ability to simultaneous learn the model structure and functionality from flows of data. eFS have been useful to develop adaptive fuzzy rule-based models, control systems, neural fuzzy, and fuzzy regression and classification trees.

Typically, eFS consider functional rule-based systems in which the structure (number of rules and antecedents/consequent components) evolves continuously using clusters created/excluded by recursive clustering algorithms (Lemos et al. 2011). After a pioneering approach for online learning of evolving Takagi-Sugeno (eTS) models (Angelov and Filev 2004) and its extensions, Simpl_eTS (Angelov and Filev 2005), and eXtended eTS (xTS) (Angelov and Zhou 2006); Angelov (2010) proposed the eTS+, an approach in which the antecedent parameters and rule-base structure are updated using criteria such as age, utility, local density and zone of influence, comprising a robust, flexible and autonomous technique (can start from scratch and does not use thresholds or user-specific parameters).

Dovzan and Skrjanc (2011a, b) suggest an alternative method for online identification of Takagi-Sugeno fuzzy functional models. To learn rule antecedents, they use a recursive fuzzy *c*-means (rFCM) algorithm. Distinctively from eTS (Angelov and Filev 2004), in which cluster centers can only be data samples with the largest potential, in rFCM the cluster centers depend on a weighted mean of the data (Dovzan and Skrjanc 2011a, b). They also suggest a recursive Gustafson–Kessel (rGK) method to capture the shapes of the clusters in the data set (Dovzan and Skrjanc 2011a, b). Combining the rGK algorithm and adaptation mechanisms such as adding, removing, splitting and merging, an evolving fuzzy model, called eFuMo (Dovzan et al. 2012), was constructed and shown to be effective for online identification of fuzzy rule-based models.

A distinct, but conceptually similar approach for adaptive TS modeling is the dynamic evolving neural-fuzzy inference system model (DENFIS) (Kasabov 2002). This approach uses a distance-based recursive clustering method to adapt the rule base structure, and a weighted recursive least squares with forgetting factor algorithm to update rules consequent parameters. A recursive clustering algorithm derived from a modification of the vector quantization technique, called evolving vector quantization, is another significant methodology to construct flexible fuzzy inference systems (FLEXFIS) (Lughofer 2008). More recent examples of eFS in the realm of neuro-fuzzy type-2 are the self-organizing fuzzy modeling and modified least-squares network (SOFMLS) (Rubio 2009), and the sequential adaptive fuzzy inference system (SAFIS) (Rong et al. 2006). SOFMLS employs an evolving nearest neighborhood clustering algorithm, whereas SAFIS uses a distance criterion in conjunction with an influence measure of the new rules created to construct and update the rule base.

The fuzzy self-organizing neural network (Leng et al. 2005) is another alternative evolvable system that adopts an error criterion to measure the generalization performance of the network. Similarly, Tung et al. (2011) suggest a self-adaptive fuzzy inference network (SaFIN) with categorical learning-induced partitioning as a clustering algorithm. The idea is to avoid the need for prior knowledge regarding the number of clusters in the input–output space. This model also shows the flexibility to incorporate new knowledge in the system. There exist other online techniques for TS fuzzy models identification considering similar mechanisms to construct and adapt their structure as the above-mentioned methods, such as the self-organizing fuzzy neural network (SOFNN) (Qiao and Wang 2008), the self-constructing fuzzy neural network (SCFNN) (Lin et al. 2001) and the generalized adaptive neuro-fuzzy inference system (GANFIS) (Azzem et al. 2003).

As reported in Lemos et al. (2011), a weakness of the recursive clustering algorithms adopted by most eFS approaches is the lack of robustness due to noise or outliers. In these situations, the algorithms may create new clusters instead of rejecting or smoothing noisy data or outliers. A method to improve robustness was suggested by Lima et al. (2010), namely, evolving participatory learning (ePL) fuzzy modeling. This approach joins the concept of participatory learning (PL) (Yager 1990) with the idea of evolving fuzzy modeling (Angelov and Filev 2004). Participatory learning induces an unsupervised clustering algorithm and is a natural candidate to develop rule base structures in dynamic environments. Similarly as in eTS modeling, structure identification and self-organization in ePL estimates the focal points of the rules, except that ePL uses PL fuzzy clustering instead of scattering, density, or information potential, as eTS does.

Alternatively Hartert et al. (2010) propose a validation phase in an adaptive fuzzy *K*-nearest neighbors (FKNN) method to delete non-representative groupings (outliers) and to merge similar clusters. In the eTS+ model, this problem is coped with the use of monitoring measures such as age and utility. The eFuMo approach uses the method of Hartert et al. (2010) to adapt the rule base and to handle outliers. A pruning criterion evaluates if a cluster was added due to an outlier. If the cluster does not gather enough support samples after a certain period of time, then it is removed.

In this paper, we introduce an enhanced version of the evolving participatory learning model, namely ePL+, which incorporates in ePL an utility measure and a variable zone of influence of clusters. The utility measure allows shrinking of fuzzy rule bases by removing rules that have low utility (the data pattern shifts away from the focal point of the rules), endowing rule bases with rules of higher summarization power. Unlike the original version of the ePL, ePL+ modeling adapts the cluster radius, an important parameter to determine the dispersion of membership functions, and hence, to govern the zone of influence and activation level of the fuzzy rules. Furthermore, ePL+ modeling gives multi-input, multi-output (MIMO) system models.

We evaluate ePL+ considering asset returns volatility modeling and forecasting. Volatility plays a central role in asset pricing, portfolio allocation and risk management. Despite the widespread use of GARCH models (Engle 1982) to address volatility forecasting, GARCH methodologies have been criticized for failing to capture dynamics of asset returns in highly unstable environments, as during financial crisis (Kung and Yu 2008; Tseng et al. 2008). To overcome this limitation, the idea is to use evolving fuzzy modeling to avoid the high computational effort needed to process data stored in databases, and to take advantage of

the inherent adaptability of evolving models, an essential requirement to handle nonlinear and non-stationary behavior such as in volatility dynamics.

Recent literature reveals several applications of evolving fuzzy rule-based models in finance and economics. Examples include: Value-at-Risk modeling and forecasting (Ballini et al. 2009a), sovereign bonds modeling (Ballini et al. 2009b), exchange rates forecasting (McDonald and Angelov 2010), fixed income option pricing (Maciel et al. 2012a, b, c), interest rate term structure forecasting (Maciel et al. 2012a, b, c), financial volatility forecasting (Luna and Ballini 2012a, b), and stochastic volatility prediction (Luna and Ballini 2012a, b).

Methods based on fuzzy set theory have been also developed to explain complex dynamics such as financial time series volatility (Chang et al. 2011; Helin and Koivisto 2011; Hung 2011a, b). Basically, these methods construct hybrid models combining fuzzy systems and GARCH models whose structure addresses both, time-varying volatility and volatility clustering. For example, Popov and Bykhanov (2005); Chang et al. (2011); Helin and Koivisto (2011); Hung (2011a, b) combine fuzzy and GARCH models to handle volatility modeling and forecasting. These modeling methods require high computational effort because they estimate the parameters of the model using all data available in the database. This may be troublesome in situations in which forecasts are needed whenever new data arrive. Therefore, current hybrid modeling methods may be prohibitive in dynamic environments which involve volatility forecasting, immunization strategies, portfolio allocation, and risk management.

Computational results reported in this paper include comparison of ePL+ modeling with state of the art evolving fuzzy modeling methods, and with GARCH modeling as benchmark. Simulations were done using actual data of S&P 500 (US) and Ibovespa (Brazil) stock market indexes from January 2000 to September 2011. Comparisons consider goodness of fit and statistical tests.

After this introduction, the paper proceeds as follows. Section 2 details ePL+ modeling, that is, a recursive scheme to learn the model structure and its parameters. Section 3 summarizes the computational results for assets returns volatility forecasting. Section 4 concludes the paper and suggests issues for further investigation.

2 Enhanced evolving participatory learning modeling (ePL+)

Evolving functional fuzzy participatory learning (ePL) modeling adopts the same philosophy as the classical evolving Takagi–Sugeno (eTS) methodology (Angelov and Filev 2004). After the initialization phase, data processing

is performed at each step to verify if a new cluster must be created, if an old cluster should be modified to account for the new data, or if redundant clusters must be eliminated, in an online-like mode. Cluster centers are the focal points of the rules, and to each cluster corresponds a fuzzy rule. Parameters of the consequent functions are computed using the local recursive least squares method.

The main difference between ePL and eTS concerns the procedure to update the rule base structure. Differently from eTS, ePL uses a fuzzy similarity measure to determine the proximity between new data and the existing rule base. The rule base structure is isomorphic to the cluster structure once each rule is associated with a cluster. Participatory learning assumes that model learning depends on what the system already knows about the model (Yager 1990). Therefore, in ePL the current model is part of the evolving process itself and influences the way in which new observations 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 rule base structure, or equivalently, on its compatibility with the current cluster structure.

In this paper, we suggest an enhanced version of the evolving participatory learning model, called ePL+. There are essentially three main additions in ePL+ that enhances ePL. First, in ePL, clusters radiuses, which represent the zone of influence of membership functions, are fixed and defined by expert knowledge. All membership functions of the fuzzy rules have the same dispersion, and hence they have the same shape in the input–output data space. In contrast, ePL+ adopts an adaptive mechanism to compute the zone of influence of each rule. Therefore, cluster radiuses are different and their values modify during learning. Because dispersion affects the activation degree of the fuzzy rules, the mechanism to determine their values has a major impact in the performance of the model. Second, the participatory learning clustering algorithm of ePL eliminates redundant clusters whenever the compatibility measure between cluster centers are greater than a user selected threshold, and no adaptive mechanism to evaluate the rule base quality is provided. On the contrary, ePL+ includes an utility measure to evaluate, recursively, the quality of the cluster structure, similar to eTS+ (Angelov 2010). The utility measure allows rule base shrinking by removing low utility rules, keeping only rules with high summarization power. Third, ePL modeling was developed for online identification of multi-input-single-output (MISO) fuzzy systems, whereas ePL+ models multi-input-multi-output (MIMO) systems as well. Therefore, ePL+ provides a more effective methodology than ePL because it is computationally efficient, requires less prior knowledge about the modeling environment, and enhances applicability in handling multivariable systems.

Next section describes the structure and parameter learning mechanisms of ePL+ modeling, emphasizing the points that differ from ePL. Details of the ePL modeling approach can be found in Lima (2008); Lima et al. (2010); Maciel et al. (2012a, b, c).

2.1 The structure of ePL+ models

ePL+ modeling assumes fuzzy rule-based systems whose rule consequents are functions of the rule antecedent variables. They form local models called local fuzzy functional models. From now on we consider, without loss of generality, Takagi–Sugeno (TS) type of fuzzy models with linear functions as rule consequents. More specifically, an ePL+ model is a set of functional fuzzy rules of the following form:

$$R^i: \text{IF } x_1 \text{ is } \Gamma_1^i \text{ AND } \dots \text{ AND } x_n \text{ is } \Gamma_n^i \text{ THEN}$$

$$y_1^i = \gamma_{0,1}^i + \sum_{j=1}^n \gamma_{j,1}^i x_j \text{ AND } \dots \text{ AND} \quad (1)$$

$$y_m^i = \gamma_{0,m}^i + \sum_{j=1}^n \gamma_{j,m}^i x_j$$

where R^i is the i th fuzzy rule ($i = 1, 2, \dots, R$), $x_j \in \mathfrak{R}$ is the j th input variable ($j = 1, 2, \dots, n$), Γ_j^i is the fuzzy set associated with the j th input variable of the i th fuzzy rule, $y_t^i \in \mathfrak{R}$ is the output of the i th rule ($t = 1, 2, \dots, m$)¹, and $\gamma_{0,t}^i$ and $\gamma_{j,t}^i$ are the parameters of the consequent of the i th rule.

TS models are developed for fuzzy domains identified by clusters, and to each cluster corresponds a linear local model. Overall, it forms a nonlinear model composed by a collection of loosely (fuzzily) coupled multiple local linear models. The contribution of a local linear model to the model output is proportional to the activation degree of each rule. We assume fuzzy sets of rule antecedents with Gaussian membership functions to ensure generalization of the description (Angelov and Filev 2004):

$$\mu^i(x_j) = e^{-\frac{(x_j - v_j^i)^2}{2(\sigma_j^i)^2}} \quad (2)$$

where $\mu^i(x_j)$ is the membership degree of the input x_j in Γ_j^i , v_j^i corresponds to the cluster center (focal point), and σ_j^i is the radius of Γ_j^i . The values of σ_j^i define the zone of influence of the i th rule.

ePL assumes that the values σ_j^i are the same and kept fixed for all fuzzy sets of all rules antecedents. Conversely,

ePL+ assumes that radiuses σ_j^i are different for each rule. σ_j^i is an important parameter because it characterizes the dispersion of the membership functions, the zone of influence and the activation degree of the rules. Therefore, in stream data processing and dynamic environments, adaptation of σ_j^i provides a more effective modeling framework and reduces the need of prior knowledge and user-specific parameter choices.

If the algebraic product t-norm is chosen to represent rule antecedents **AND**, then the activation degree of the i th rule is:

$$\kappa^i(x) = \prod_{j=1}^n \mu^i(x_j) = \mu^i(x_1)\mu^i(x_2) \dots \mu^i(x_n) \quad (3)$$

The output of the TS model at each step is the weighted average of the individual rule contributions:

$$y = \sum_{i=1}^R \theta^i y^i, \quad \theta^i = \frac{\kappa^i}{\sum_{l=1}^R \kappa^l} \quad (4)$$

where θ^i is the normalized activation degree of the i th rule, and R is the number of fuzzy rules.

As stated in Angelov and Filev (2004), identification of a TS model requires two sub-tasks: (1) learning the antecedent part of the model using e.g. a fuzzy clustering algorithm, and (2) learning the parameters of the linear rule consequent functions. In this paper, we focus on the evolving fuzzy participatory clustering algorithm for antecedent learning, and on the recursive least squares algorithm to estimate the consequent parameters.

2.2 Participatory learning clustering

The ePL+ modeling approach adopts the same mechanism to construct the rule base as ePL, i.e., the participatory learning clustering algorithm. Generally speaking, participatory learning assumes that model learning depends on what the system already knows about the model. Therefore, in ePL+ the current model is part of the evolving process itself and influences the way in which new data are used for self-organization. An essential property of PL is that the impact of new data in inducing self-organization or model revision depends on its compatibility with the current rule base structure or, equivalently, its compatibility with the current cluster structure (Lima et al. 2010).

In online mode, rather than being a fixed set, training data are collected continuously (Angelov and Zhou 2008). Let $v_k^i = [v_{1k}^i, v_{2k}^i, \dots, v_{nk}^i]^T$ be a vector that encodes the i th ($i = 1, \dots, R_k$) cluster center at step k . The aim of the participatory mechanism is to learn the value of v_k^i using a stream of data $x_k = [x_{1k}, x_{2k}, \dots, x_{nk}]^T$. In other words, each $x_k, k = 1, 2, \dots$, is used as a vehicle to learn about v_k^i . We say that the learning process is participatory if the

¹ One must note that ePL considers MISO models whereas ePL+ concerns MIMO models. This difference appears in consequent parameters learning, as will be described in the Subsect. 2.5.

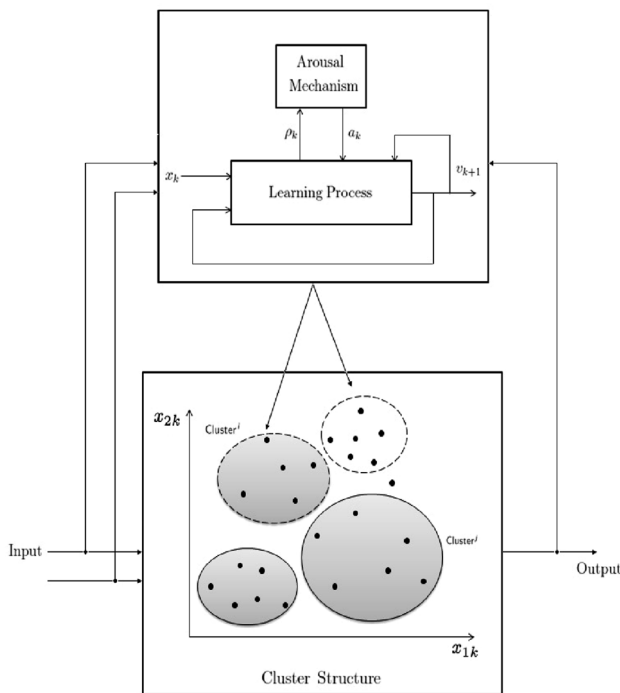


Fig. 1 Participatory learning clustering

contribution of each data x_k to the learning process depends upon its acceptance by the current estimate of v_k^i being valid. Implicit in this idea is that, to be useful and to contribute to the learning of v_k^i , data x_k must somehow be compatible with current estimates of v_k^i .

In ePL+, the main object of learning is the cluster structure. A cluster structure is formed by the collection of cluster centers (or prototypes) that assemble the input space partition. More formally, given an initial cluster structure $v_0^i, i = 1, \dots, R_0$, a collection of cluster centers $v_k^i, i = 1, \dots, R_k$, is constructed at each k using a compatibility measure $\rho_k^i \in [0, 1]$ and an arousal index $a_k^i \in [0, 1]$. While ρ_k^i measures the extent to which data x_k is compatible with the current cluster structure, the arousal index a_k^i acts as a reminder of when current cluster structure should be revised in face of new information contained in the data. Figure 1 shows the main constituents and functioning of PL clustering. If an initial cluster structure is not available beforehand, then the PL algorithm assumes the first data point as a cluster center.

Due to its unsupervised, self-organizing nature, the PL clustering procedure may create a new cluster or modify existing ones at each step k . If the arousal index is greater than a threshold value $\tau \in [0, 1]$, a new cluster is created. Otherwise, the i th cluster center, the one most compatible with x_k , is adjusted as follows:

$$v_{k+1}^i = v_k^i + G_k^i(x_k - v_k^i) \tag{5}$$

where

$$G_k^i = \alpha(\rho_k^i)^{1-a_k^i} \tag{6}$$

with $\alpha \in [0, 1]$ as the learning rate and

$$\rho_k^i = 1 - \frac{\|x_k - v_k^i\|}{n} \tag{7}$$

with $\|\cdot\|$ a norm, n the dimension of input space, and

$$i = \arg \max_j \{\rho_j^k\} \tag{8}$$

Note that the i th cluster center is a convex combination of the input data x^k and the closest cluster center.

The arousal index a_k^i in (6) is updated as follows:

$$a_{k+1}^i = a_k^i + \beta(1 - \rho_{k+1}^i - a_k^i) \tag{9}$$

The value of $\beta \in [0, 1]$ controls the rate of change of arousal: the closer β is to one, the faster the system is to sense compatibility variations. When $a_k^i = 0$, $G_k^i = \alpha\rho_k^i$, which is the PL procedure with no arousal. If the arousal index 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 the observations. Therefore learning is dynamic in the sense that (5) can be viewed as a belief revision strategy whose effective learning rate (6) depends on the compatibility among new data, the current cluster structure, and on model confidence as well.

Note that the learning rate is modulated by compatibility. Conventional learning models have no participatory considerations and the learning rate is usually set small to avoid undesirable oscillations due to spurious values of data far from cluster centers. While protecting against the influence of noisy data, low learning rate slow down learning. Participatory learning allows higher values of the learning rate, but the compatibility index lowers the effective learning rate when large deviations occur. On the other hand, high compatibility increases the effective rate and speeds up the learning process.

Whenever a cluster center is updated or a new cluster added, the PL clustering procedure verifies whether redundant clusters were created. Updating a cluster center using (5) can push a given center closer to another one and a redundant cluster may be formed. Thus, a mechanism to avoid redundancy is needed. One of such a mechanism is to verify if distinct rules produce similar outputs. In PL clustering, a cluster center is declared redundant whenever its similarity with another center is greater than or equal to a threshold value- $\lambda \in [0, 1]$. If this is the case, then we can either maintain the original cluster center or replace it by the average between the new data and the current cluster

center. Similarly as in (7), the compatibility index between cluster centers i and j is computed using:

$$\rho_k^{i,j} = 1 - \sum_{l=1}^n |v_{lk}^i - v_{lk}^j| \quad (10)$$

Therefore, if

$$\rho_k^{i,j} \geq \lambda \quad (11)$$

then the cluster i is declared redundant.

Participatory clustering requires the user to choose parameters α , β , λ and τ . The choice can be guided by the following consideration. If an input data is such that the arousal is greater than the threshold $\tau \in [0, 1]$, then this data becomes the focal point of a new cluster (Lima 2008). If the cluster with the highest compatibility s is updated, then $a_{k+1}^s < \tau$, and from (9) we get:

$$\begin{aligned} a_k^s + \beta(1 - \rho_{k+1}^s - a_k^s) &= (1 - \beta)a_k^s + \beta d_k^s < \tau \\ \Rightarrow d_k^s &< \frac{\tau - (1 - \beta)a_k^s}{\beta} < \frac{\tau}{\beta} \end{aligned} \quad (12)$$

where $d_k^s = d(v_k^s, x_k)$ is the distance between cluster center s and input x_k at k .

For any two distinct cluster $i, j = 1, \dots, c^k, i \neq j$, the compatibility measure $\rho_k^{i,j}$ is such that:

$$\rho_k^{i,j} = 1 - d_k^{i,j} < \lambda \Rightarrow d_k^{i,j} > 1 - \lambda \quad (13)$$

If $d_k^{i,j} < 1 - \lambda$, clusters i and j are considered redundant and become a single cluster. Here, $d_k^{i,j} = d(v_k^i, v_k^j)$ is the distance between clusters centers v_k^i and v_k^j .

To ensure that a new, non redundant cluster is added, (12) and (13) suggest to choose values of β , λ and τ such that:

$$0 < \frac{\tau}{\beta} \leq 1 - \lambda \leq 1$$

where

$$\tau \leq \beta \quad \text{and} \quad \lambda \leq 1 - \tau$$

Analysis of the dynamic behavior of the participatory learning considering the compatibility and arousal mechanisms simultaneously with learning rate is discussed in Lima et al. (2010). The learning rate α is a small value, typically $\alpha \in [10^{-1}, 10^{-5}]$.

Initial cluster structure can be chosen differently if data are available beforehand. In this case, we may consider the use of subtractive clustering (SC) algorithm (Chiu 1994) to obtain a corresponding initial rule base because SC does not require pre-specified number of clusters. This adds flexibility and increases PL clustering autonomy once the number of clusters and the clusters themselves are derived from data. The subtractive clustering proceeds as follows.

- (1) For all training data, select the one with the highest potential to be the first cluster center. The potential of a point is measured as the spatial proximity between all other data points computed by the Euclidean distance;
- (2) Reduce the potential of all other points by an amount proportional to the potential of the chosen point and inversely proportional to the distance to the current center;
- (3) Define two boundary (lower and upper) conditions as a function of the maximal potential. If a potential of data point is higher than the upper threshold, then create a new cluster center;
- (4) If the potential of a point lies between the boundaries and if this point is very close to some cluster center, then replace this cluster center by the current one.

See Chiu (1994) for further details.

2.3 Clusters quality measurement

Originally, the PL clustering algorithm removes redundant clusters, but gives no mechanism to recursively evaluate the summarization power of the rule base induced by the cluster structure. Differently, ePL+ uses an online monitoring scheme to evaluate the quality of a cluster structure based on the utility measure introduced in Angelov (2010). The utility measure is an indicator of the accumulated relative activation degree of a rule:

$$U_k^i = \frac{\sum_{l=1}^k \theta^l}{k - I^{i*}} \quad (14)$$

where I^{i*} denotes the step at which an input data became the focal point of the i th fuzzy rule, that is, it indicates when the i -th fuzzy rule was generated.

The utility measure gives a clue about the extent to which a fuzzy rule is useful. The use of the quality measure U_k^i aims at avoiding unused clusters in the cluster structure; that is, it eliminates fuzzy rules with steadily low activation degrees. For this purpose, the following principle for rule elimination is adopted (Angelov 2010):

$$\text{IF } U_k^i < \epsilon \text{ THEN delete } R^i, \quad R_k \leftarrow R_k - 1 \quad (15)$$

where $\epsilon \in [0.03; 0.1]$ is a threshold to control the utility of each cluster, and R_k is the number of rules in the rule base at step k .

This principle helps to underscore high quality rule bases because it emphasizes the rules which are relevant for the current structure of model. Angelov (2010) has suggested alternative quality measures such as age, support, zone of influence, and local density. This paper considers utility measure only because of it achieves good performance with low computational cost.

2.4 Adaptation of the zone of influence

A key parameter associated with Gaussian membership functions is its cluster radius σ_j^i because it defines the spread of the membership function and influences the activation degree of the fuzzy rules (Angelov 2010). Similarly as in Angelov and Zhou (2006), ePL+ modeling considers variable radiuses to capture the distribution of data at each step. They are updated as follows:

$$\sigma_{jk}^i = \pi \sigma_{jk}^i + (1 - \pi) q_{jk}^i \tag{16}$$

where π is a learning rate (suggested values are in range [0.3, 0.5]) (Angelov 2010), and $q_{jk}^i, j = 1, 2, \dots, n$, is called the local scatter:

$$(q_{kj}^i)^2 = \frac{1}{S_k^i - 1} \sum_{l=1}^{S_k^i - 1} (z_j^i - z_{lj})^2 \tag{17}$$

where $z = [x^T, y^T]^T$ is an input/output pair and $S_k^i = S_{k-1}^i + 1$ is the support of the cluster $i, i = \max_{i=1}^{R_k} \kappa^i(x)$. Essentially, S_k^i is the number of data samples associated with cluster i and gives an indication of the generalization power of the i -th fuzzy rule.

2.5 Parameter estimation

Estimation of the parameters of the linear rule consequent functions can be formulated as a least squared problem (Angelov and Filev 2004). Equation (4) can be put into vector form:

$$y = \Lambda^T \Phi \tag{18}$$

where $y = [y_1, y_2, \dots, y_m]^T$ is the m -dimensional output of the MIMO ePL+ model, $\Lambda = [\lambda_1 x_e^T, \lambda_2 x_e^T, \dots, \lambda_n x_e^T]^T$ is the fuzzily weighted extended input vector, $x_e = [1 \ x^T]^T$ is the expanded data vector, $\Phi = [\Psi_1^T, \Psi_2^T, \dots, \Psi_R^T]^T$ is the matrix of rule base parameters, and

$$\Psi_i = \begin{bmatrix} \gamma_{01}^i & \dots & \gamma_{0m}^i \\ \dots & \dots & \dots \\ \gamma_{n1}^i & \dots & \gamma_{nm}^i \end{bmatrix}^T \tag{19}$$

is the matrix of consequent function parameters of the i th linear local model.

Since the actual output value is available at each step, the parameters of the consequents can be updated using either the local or global recursive least squares algorithm

(Chiu 1994). In this paper we use the local recursive least squares whose purpose is to minimize local modeling errors:

$$\min E_L^i = \min \sum_{l=1}^k \theta^i(x_l) (y_l - x_{el}^T \Psi_l^i)^2 \tag{20}$$

Thus, the parameters of the i th local model is updated as follows (Angelov and Filev 2004; Angelov and Zhou 2008; Angelov 2010):

$$\Psi_{k+1}^i = \Psi_k^i + \Sigma_k^i x_{ek} \theta_k^i (y_k - x_{ek}^T \Psi_k^i), \Psi_1^i = 0 \tag{21}$$

$$\Sigma_{k+1}^i = \Sigma_k^i - \frac{\theta_k^i \Sigma_k^i x_{ek} x_{ek}^T \Sigma_k^i}{1 + \theta_k^i x_{ek}^T \Sigma_k^i x_{ek}}, \Sigma_1^i = \Omega I_{(n+1) \times (n+1)} \tag{22}$$

where I is the $(n + 1) \times (n + 1)$ identity matrix; Ω is a large number (usually $\Omega = 1,000$); and Σ is the dispersion matrix.

Whenever a new fuzzy rule is added, the corresponding dispersion matrix is set as $\Sigma_k^{R_k+1} = \Omega I$. Parameters of the new rule are estimated using the parameters of the existing R_k fuzzy rules as follows (Angelov 2010):

$$\Psi_k^{R_k+1} = \sum_{i=1}^{R_k} \theta^i \Psi_{k-1}^i \tag{23}$$

If no new rule is added, then their parameters are inherited from the previous step, and the dispersion matrices updated accordingly. Finally, after estimation of the consequent parameters values, the model output is produced using (4).

The use of the recursive least squares algorithm depends on the initial values of the parameters Ψ_0 , and on the initial values of the entries of the dispersion matrix Σ_0 . These initial values may be chosen if previous information about the system is available and exploring a database to get an initial rule base and Ψ_0 and Σ_0 . When no previous information is available we may proceed choosing large values for the entries of the matrix, as indicated above. In this paper, we assume the availability of a database to get the initial rule base and corresponding parameters.

2.6 ePL+ algorithm

The detailed steps of the ePL+ model are specified in pseudo-code form. All steps of the algorithm are non-iterative. The model can develop/evolve an existing model whenever the data pattern changes, and its recursive nature turns it computationally effective.

Enhanced Evolving Participatory Learning Fuzzy Modeling ePL+

Initialize rule base using SC or Fuzzy C-Means algorithm

Set the control parameters α , β , τ , λ , and $a_1^i = 0$

For $k = 1, 2, \dots$ **Do**

Read x_k

For $i = 1, 2, \dots, R_k$ **Do**

Compute the compatibility index: $\rho_k^i = 1 - \frac{\|x_k - v_k^i\|}{n}$

Compute the arousal: $a_{k+1}^i = a_k^i + \beta(1 - \rho_{k+1}^i - a_k^i)$

End For

If $a_{k+1}^i \geq \tau \forall i = 1, 2, \dots, R_k$ **Then**

Create a new cluster: $R_k = R_k + 1$

x_k is the center of the new cluster: $v_k^{R_k} = x_k$

Otherwise

Find the most compatible cluster: $s = \arg \max_{j=1}^{R_k} \rho_k^j$

Update the cluster center s : $v_{k+1}^s = v_k^s + \alpha(\rho_k^s)^{1-a_{k+1}^s}(x_k - v_k^s)$

End If

For $i = 1, 2, \dots, R_k - 1$ **and** $j = (i + 1), (i + 2), \dots, R_k$ **Do**

Compute the compatibility between the clusters v_k^i and v_k^j : $\rho_k^{i,j} = 1 - d(v_k^i, v_k^j)$

If $\exists i, j \mid \rho_k^{i,j} \geq \lambda$ **Then**

Remove the redundant cluster: $R_k = R_k - 1$

End If

End For

Update the utility measure: $U_k^i = \frac{\sum_{l=1}^k \theta^l}{k - I^{i*}}$

If $\exists i \mid U_k^i \leq \epsilon$ **Then**

Remove cluster i : $R_k = R_k - 1$

End If

Compute consequent parameters using recursive least square

Compute model output

End For

3 Computational experiments for volatility forecasting

The ePL+ model introduced in this paper is a flexible structure that can be employed to a range of problems such as regression, time series forecasting, classification, control, clustering, and novelty detection. In Maciel et al. (2012a, b, c), the ePL+ was used to forecast nonlinear benchmark problems such as Mackey–Glass time series and Box–Jenkins gas furnace data. The results have shown the efficiency of ePL+ to produce accurate time series forecasts.

This paper shows the effectiveness of ePL+ to model and forecast financial asset returns volatility. We compare the performance of ePL+ with state of the art evolving fuzzy system modeling approaches respectively, eTS, xTS,

eTS+ and ePL, and with the GARCH model, a benchmark in the area of volatility modeling.

3.1 Data

To illustrate the performance of the ePL+ model in forecasting stock market volatility, we took the daily values of the S&P 500 (US) and the Ibovespa (Brazil) stock market indexes over the period from January 3, 2000 through September 19, 2011. The daily stock return series were generated by considering the differences in the natural logarithm of the daily stock index and the previous day stock index as follows:

$$r_t = \ln(P_t) - \ln(P_{t-1}) \quad (24)$$

where r_t is the stock market return, and P_t is the stock index at t .

The data set was split into two subsets. The first is an in-sample set consisting of data from January 3, 2000 through September 21, 2007. The second is an out-of-sample set with data from September 24, 2007 through September 19, 2011. The second data set is used for forecasting. Data splitting is needed for the GARCH modeling only. The evolving fuzzy models learn recursively and do not require a pre-training phase.

S&P 500 and Ibovespa intra-day data were collected for the period from January 3, 2000 through September 19, 2011 to construct the realized volatility time series and use it as the 'true volatility'². Realized volatility RV_t at t is computed as the sum of squared high-frequency returns within a day. This mechanism conveniently avoids data processing complications while covering more information on daily transactions. Thus

$$RV_{t+1}(\Delta) = \sum_{j=1}^{1/\Delta} r_{t+j\Delta,\Delta}^2 \tag{25}$$

where $r_{t,\Delta}$ is the discretely sampled Δ -period returns.

The evolving fuzzy models inputs are previous lags of the realized volatility of each market index. The Bayesian information (BIC) and Akaike information (AIC) criteria were used to select lag values for the GARCH model (Akaike 1974; Schwarz 1978).

3.2 Performance assignment

Evaluation of the models was done considering the root mean square error (RMSE), and the non-dimensional error index (NDEI) which is the ratio between the RMSE and the standard deviation of the target data. They are computed as follows:

$$RMSE = \left(\frac{1}{N} \sum_{k=1}^N (y_k - \hat{y}_k)^2 \right)^{\frac{1}{2}} \tag{26}$$

$$NDEI = \frac{RMSE}{std(y_k)} \tag{27}$$

where y_k is the actual volatility, $std(\cdot)$ is the standard deviation function, \hat{y}_k is the model output, and N is the number of forecasts.

Model validation is an important task in system identification to access the goodness of fit of the model. As stated by Billings and Zhu (1994), nonlinear model validation can be approached using either correlation or model-comparison. Correlation based validation involves computing correlation functions of model residuals and system inputs,

and testing if these lie within confidence intervals (Billings and Zhu 1994). Model-comparison based validation concerns the use of statistical tests to compare models pairwise to test whether they are equally accurate or not in terms of their accuracy. Since ePL+ and the state of the art evolving models considered in this paper do not assume any probability distribution of data or model residuals, model-comparison based validation appear to be the one most appropriate.

In this paper, we employ the (Diebold and Mariano 1995) statistic test assuming equal predictive accuracy as the null hypothesis. This statistic test is widely used for forecasting model validation, especially in economics and finance. If N is the sample size and e_i^1, e_i^2 ($i = 1, 2, \dots, N$) are the forecast errors of two models, the mean square error (MSE) loss functions are computed as:

$$L_1^{MSE}(e_i^1) = (e_i^1)^2, L_2^{MSE}(e_i^2) = (e_i^2)^2 \tag{28}$$

The Diebold–Mariano test is based on the loss difference:

$$d_i^{MSE} = L_1^{MSE}(e_i^1) - L_2^{MSE}(e_i^2) \tag{29}$$

Thus, the null hypothesis is $H_0 : E(d_i^{MSE}) = 0$, meaning that both models are equally accurate. The alternative hypothesis $H_1 : E(d_i^{MSE}) \neq 0$ means that model 1 is more accurate than model 2. The Diebold–Mariano test statistic, DM, is found as follows:

$$DM = \frac{\bar{d}}{\sqrt{\hat{V}(\bar{d})}} \sim \mathcal{N}(0, 1) \tag{30}$$

where $\bar{d} = N^{-1} \sum_{i=1}^N d_i^{MSE}$ and $\hat{V}(\bar{d}) = N^{-1} \left[\hat{\varphi}_0 + 2 \sum_{k=1}^{N-1} \hat{\varphi}_k \right]$, $\varphi_k = cov(d_i, d_{i-k})$.

3.3 Control parameters

Control parameters were chosen based on simulation experiments, attempting to produce better performance in terms of RMSE and NDEI measures in-sample set. Table 1 shows the parameters values of all evolving fuzzy models for S&P 500 and Ibovespa indexes. The SC algorithm was employed to initialize the rule base of all evolving modeling approaches.

3.4 Results and discussion

Table 2 summarizes the basic statistics of the return series. The average daily returns are negative for the S&P 500 and positive for Ibovespa. The daily returns display evidence of skewness and kurtosis. The return series is skewed toward the left and characterized by a distribution with tails that are significantly thicker than for a normal distribution. The

² The database was provided by Bloomberg.

Table 1 Control parameters of the evolving fuzzy models for S&500 and Ibovespa data

Model and source	β	τ	α	λ	π	Ω	ϵ	σ
S&P 500								
eTS (Angelov and Filev 2004)	–	–	–	–	–	1,000	–	0.5
xTS (Angelov and Zhou 2006)	–	–	–	–	–	1,000	–	–
eTS+ (Angelov 2010)	–	–	–	–	0.5	1,000	0.1	–
ePL (Lima et al. 2010)	0.16	0.16	0.01	0.84	0.5	1,000	–	0.5
ePL+ current model	0.16	0.16	0.01	0.84	0.5	1,000	0.1	–
Ibovespa								
eTS (Angelov and Filev 2004)	–	–	–	–	–	1,000	–	0.5
xTS (Angelov and Zhou 2006)	–	–	–	–	–	1,000	–	–
eTS+ (Angelov 2010)	–	–	–	–	0.5	1,000	0.1	–
ePL (Lima et al. 2010)	0.02	0.02	0.02	0.11	0.5	1,000	–	0.4
ePL+ current model	0.02	0.02	0.02	0.11	0.5	1,000	0.1	–

Table 2 Descriptive statistics of S&P 500 and Ibovespa daily returns

	S&P 500	Ibovespa
Mean	–0.0073	0.0378
Max	10.9572	13.6766
Min	–9.4695	–12.0961
SD	1.3787	1.9493
Skewness	–0.1580	–0.1066
Excess Kurtosis	3.6701	7.4814
J–B ^a	234.9483*	277.1270*
$Q^2(10)^b$	789.7362*	683.9531*
ARCH Test (10) ^c	1109.1934*	1082.7409*

^a The statistics of Jarque–Bera normal distribution test

^b The Ljung–Box Q-test for the 10th order serial correlation of the squared returns

^c Engle’s ARCH test also examines for autocorrelation of the squared returns

* Significant at the 5 % level

Jarque–Bera test statistics further confirm that the daily returns are non-normally distributed. The Ibovespa index has a higher kurtosis than S&P 500, which explains the fact that emerging countries, generally, exhibit more leptokurtic behavior. Under the null hypothesis of no serial correlation in the squared returns, the Ljung–Box $Q^2(10)$ statistics infer linear dependence for both, S&P 500 and Ibovespa series. Moreover, Engle’s ARCH test for the squared returns reveals strong ARCH effects, an evidence supporting GARCH effects (that is, heteroscedasticity).

The stock indexes are shown in Fig. 2 and the corresponding returns in Fig. 3. In particular, in Fig. 3 volatility clustering becomes clearer, especially when the context of the recent US *subprime* crisis is considered, which translates the idea of higher non-stationarity during unstable events.

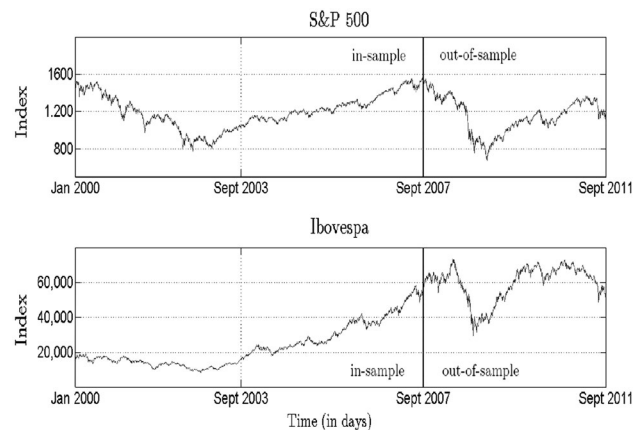


Fig. 2 Daily closing stock price indexes for the S&P 500 and Ibovespa

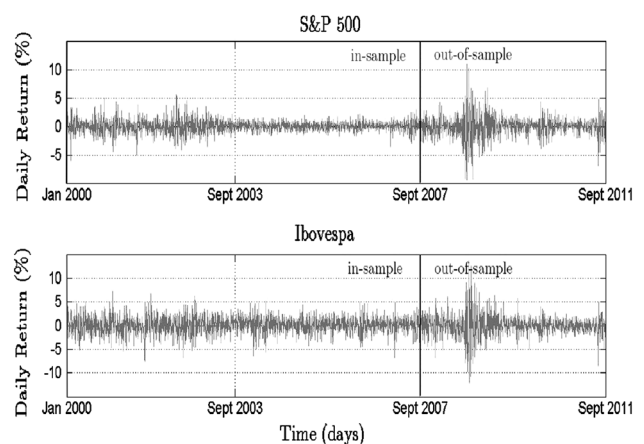


Fig. 3 S&P 500 and Ibovespa daily returns

For the S&P 500 index, the inputs of the evolving fuzzy models are the previous five lags of actual volatility, measured as the realized volatility, while for the Ibovespa

Table 3 Volatility modeling performance for one-step ahead forecast

Index	Model and source	RMSE	NDEI	Rules	Time (sec)	CP
S&P 500	GARCH (Engle 1982)	0.1685	0.1222	–	4.5698	–
	eTS (Angelov and Filev 2004)	0.1208	0.0876	5	3.6147	2
	xTS (Angelov and Zhou 2006)	0.1177	0.0854	4	3.2063	1
	eTS+ (Angelov 2010)	0.1098	0.0796	3	2.9881	3
	ePL (Lima et al. 2010)	0.1249	0.0906	5	3.1152	7
	ePL+ current model	0.0984	0.0714	3	3.0511	7
Ibovespa	GARCH (Engle 1982)	0.1954	0.1002	–	5.4117	–
	eTS (Angelov and Filev 2004)	0.1693	0.0869	7	4.8623	2
	xTS (Angelov and Zhou 2006)	0.1588	0.0815	6	4.0107	1
	eTS+ (Angelov 2010)	0.1473	0.0756	4	3.7719	3
	ePL (Lima et al. 2010)	0.1501	0.0770	6	4.9996	7
	ePL+ current model	0.1306	0.0670	4	4.1143	7

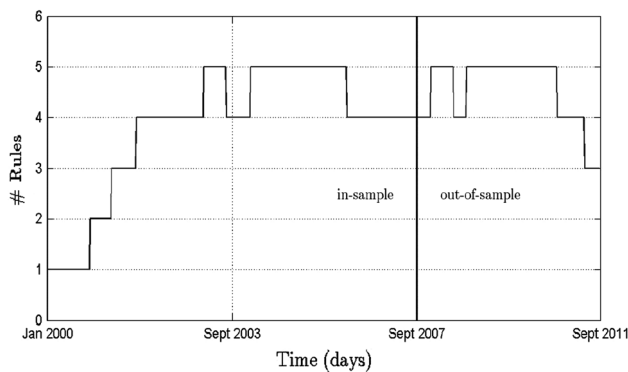


Fig. 4 Number of fuzzy rules of ePL+ for S&P 500

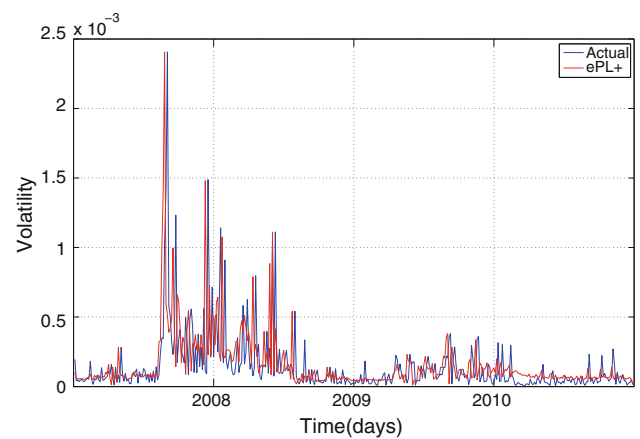


Fig. 6 S&P 500 actual and volatility forecast using ePL+

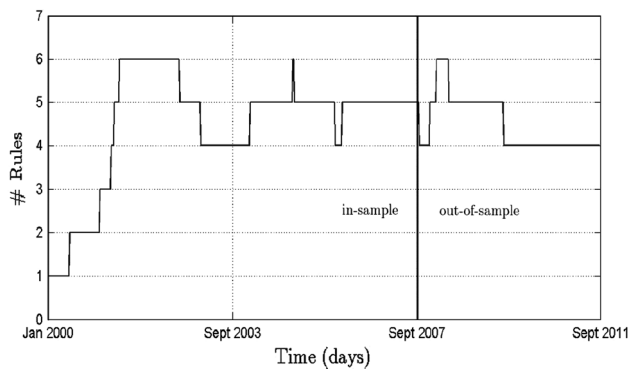


Fig. 5 Number of fuzzy rules of ePL+ for Ibovespa

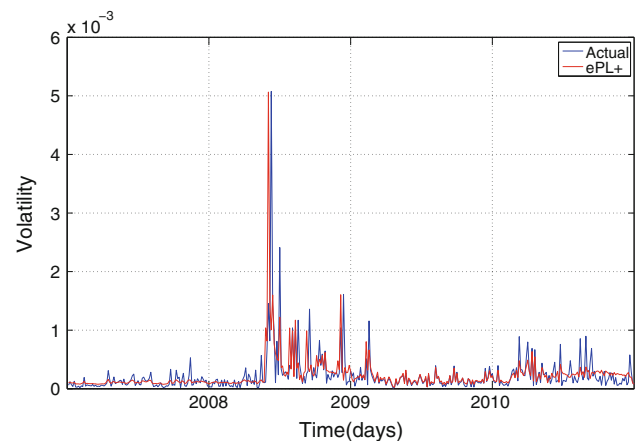


Fig. 7 Ibovespa actual and volatility forecasts using ePL+

index the past three lags of actual volatility were chosen as inputs. Models outputs are one-step ahead forecasts. A GARCH (1,1) model was selected as benchmark.

Table 3 summarizes the performance of the models when forecasting the S&P 500 and Ibovespa stock indexes volatility in terms of error measures (RMSE and NDEI), final number of rules (Rules), processing time, and number

of control parameters (CP). The ePL+ model performs better than the remaining models in terms of accuracy. All evolving fuzzy models achieve better performance than the GARCH model. The number of fuzzy rules and the

processing time are similar for the evolving models. From the point of view of accuracy and model complexity, ePL+ performs best since it requires less rules to achieve comparable or better modeling error values.

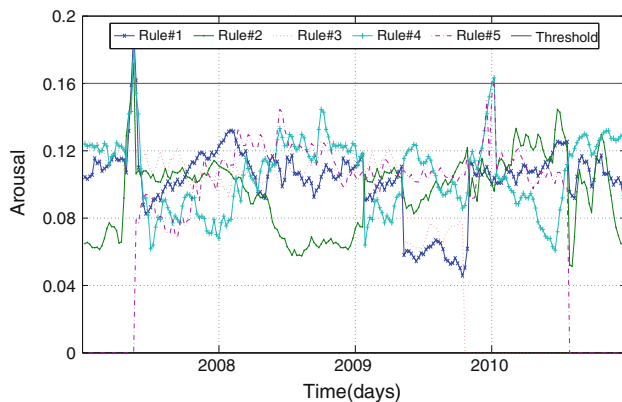


Fig. 8 Evolution of the arousal index of ePL+, out-of-sample learning for S&P 500 data

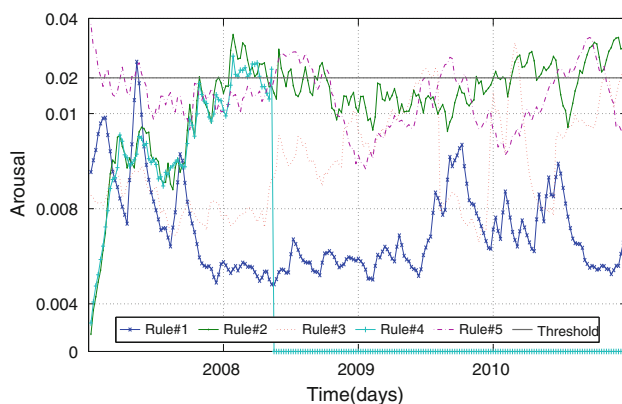


Fig. 9 Evolution of the arousal index of ePL+: out-of-sample learning for Ibovespa data

Figures 4 and 5 depict the continuous adaptation of the ePL+ model structure, i.e., the number of fuzzy rules in the rule base when modeling S&P 500 and Ibovespa stock market indexes, respectively. Notice that the number of rules increases for both indexes during 2007 and 2009, revealing the capability of ePL+ to capture instabilities due to crises. Interestingly, the 2007–2009 period corresponds to the US *subprime* mortgage crises.

To illustrate the capability of the ePL+ model to deal with volatility forecasting, Figs. 6, 7 show the “true volatility”, measured as the realized volatility, and the volatility forecasted using the ePL+ model for the S&P 500 and Ibovespa indexes, respectively. Periods of high volatility correspond to the volatility clustering behavior in the series of stock returns. The S&P 500 (Fig. 6) and Ibovespa (Fig. 7) indexes forecasts show that ePL+ successfully captures the instabilities suffered by the economies during the *subprime* crisis that started in the second semester of 2008.

Figures 8 and 9 show how the values of the arousal index of ePL+ change during learning with out-of-sample data sets for both, S&P 500 and Ibovespa indexes, respectively. For S&P 500, out-of-sample learning starts with five fuzzy rules and ends with three rules for arousal threshold $\tau = 0.16$ (Fig. 8). The Ibovespa index model starts with five rules and ends with four rules when arousal threshold $\tau = 0.02$.

The statistical performance evaluated using the (Diebold and Mariano 1995) using MSE loss function are summarized in Table 4. Diebold–Mariano tests were done between pairs of models, e.g., ePL+ versus GARCH. The null hypothesis, the equal forecast accuracy, is rejected in all cases, assuming 5% confidence level because $|DM| > 1.96$. Therefore, statistically evolving fuzzy forecast models are superior than GARCH. The eFS models can be viewed as equally accurate.

Table 4 Diebold–Mariano statistic for MSE loss function

Index	Model and source	eTS	xTS	eTS+	ePL	ePL+
S&P 500	GARCH (Engle 1982)	−2.98*	−2.54*	−2.01*	−2.22*	−2.95*
	eTS (Angelov and Filev 2004)	–	−0.44	−1.25	−0.15	−0.54
	xTS (Angelov and Zhou 2006)	–	–	−1.44	−1.13	−1.88
	eTS+ (Angelov 2010)	–	–	–	−0.62	−0.97
	ePL (Lima et al. 2010)	–	–	–	–	−1.12
Ibovespa	GARCH (Engle 1982)	−2.47*	−2.25*	−2.85*	−2.44*	−2.52*
	eTS (Angelov and Filev 2004)	–	−0.84	−0.55	−0.60	−0.34
	xTS (Angelov and Zhou 2006)	–	–	−1.01	−1.57	−1.13
	eTS+ (Angelov 2010)	–	–	–	−1.41	−0.59
	ePL (Lima et al. 2010)	–	–	–	–	−1.17

* Significant at the 5% level

4 Conclusion

This paper has introduced improvements of the evolving participatory learning fuzzy modeling (ePL), a method based on participatory learning and evolving systems paradigms to build adaptive fuzzy rule-based models. The enhanced version of ePL modeling, namely ePL+, uses an utility measure to monitor the quality of the rule base and variable zone of influence of the clusters that form fuzzy rules.

Computational results concerning the evaluation of ePL+ modeling to forecast volatility were reported. The ePL+ was compared with evolving fuzzy modeling methods representative of the current state of the art, and with GARCH modeling, an approach commonly used in the economics and finance literature. Volatility forecasting plays a central role in many financial decisions such as asset allocation and hedging, option pricing, and risk analysis. Since volatility mirrors behavior of non-stationary nonlinear environments, evolving modeling have been shown to be very suitable. Empirical evidence based on S&P 500 and Ibovespa index market data illustrates the potential of the ePL+ to forecast volatility. ePL+ perform better than ePL models in terms of both, model accuracy and model structure because it achieves low modeling error with smaller number of fuzzy rules than its counterparts.

Future work will address the use of the ePL+ model in financial decision making under volatility such as in option pricing, portfolio selection, and risk analysis.

Acknowledgements The authors thank the Brazilian Ministry of Education (CAPES), the Brazilian National Research Council (CNPq) Grants 304596/2009-4 and 306343/2011-8, and the Research Foundation of the State of São Paulo (FAPESP) Grant 13851-3/2011, for their support.

References

- Akaike H (1974) A new look at the statistical model identification. *IEEE Trans Automat Cont* 19(6):716–723
- Angelov P (2010) Evolving Takagi–Sugeno fuzzy systems from streaming data (eTS+). In: Angelov P, Filev D, Kasabov N (eds) *Evolving intelligent systems*, Wiley-IEEE-Press, NJ, chap 2, pp 21–50
- Angelov P, Filev D (2004) An approach to online identification of evolving Takagi–Sugeno fuzzy models. *IEEE Trans Syst Man Cybernetics Part B* 34(1):484–498
- Angelov P, Filev D (2005) Simpl_eTS: a simplified method for learning evolving Takagi–Sugeno fuzzy models. In: *Proceedings 14th IEEE International Conference Fuzzy System*, pp 1068–1073
- Angelov P, Zhou X (2006) Evolving fuzzy systems from data streams in real-time. In: *International Symposium on Evolving Fuzzy Systems*, IEEE Press, Ambelside, Lake District, pp 29–35
- Angelov P, Zhou X (2008) On-line learning fuzzy rule-based system classifiers from data streams. *IEEE Trans Fuzzy Syst* 16(6):1462–1475
- Attar V, Sinha P, Wankhade K (2010) A fast and light classifier for data streams. *Evolv Syst* 1(3):199–207
- Azzem MF, Hanmandlu M, Ahmad N (2003) Structure identification of generalized adaptive neuro-fuzzy inference systems. *IEEE Trans Fuzzy Syst* 11(5):668–681
- Ballini R, Mendonça ARR, Gomide F (2009a) Evolving fuzzy modeling in risk analysis. *Intell Syst Acc Fin Manag* 16(1–2):71–86
- Ballini R, Mendonça ARR, Gomide F (2009b) Evolving fuzzy modeling of sovereign bonds. *J Fin Decis Mak* 5(2):3–15
- Billings SA, Zhu QM (1994) Nonlinear model validation using correlation tests. *Intern J Con* 60(6):1107–1120
- Chang J, Wei L, Cheng C (2011) A hybrid ANFIS model based on AR and volatility for TAIEX forecasting. *Appl Soft Comput* 11(1):1388–1395
- Chiu SL (1994) Fuzzy model identification based on cluster estimation. *J Intell Fuzzy Syst* 2(3):267–278
- Diebold FX, Mariano RS (1995) Comparing predictive accuracy. *J Bus Econom Stat* 13(3):253–263
- Dovzan D, Logar V and Skrjanc I (2012) Solving the sales prediction with fuzzy evolving models. In: *WCCI 2012 IEEE World Congress on Computational Intelligence June, Brisbane*, pp 10–15
- Dovzan D, Skrjanc I (2011a) Recursive fuzzy c-means clustering for recursive fuzzy identification of time-varying processes. *ISA Trans* 50(2):159–169
- Dovzan D, Skrjanc I (2011b) Recursive clustering based on a Gustafson–Kessel algorithm. *Evolv Syst* 2(1):15–24
- Engle RF (1982) Autoregressive conditional heteroskedasticity with estimates of the variance of UK inflation. *Econometrica* 50(4):987–1007
- Hartert L, Mouchawed MS, Billaudel P (2010) A semi-supervised dynamic version of Fuzzy K-nearest neighbours to monitor evolving systems. *Evolv Syst* 1(1):3–15
- Helin T, Koivisto H (2011) The GARCH-fuzzy density method for density forecasting. *Appl Soft Comput* 11(6):4212–4225
- Hung J (2011a) Adaptive fuzzy-GARCH model applied to forecasting the volatility of stock markets using particle swarm optimization. *Inf Sci* 181(20):4673–4683
- Hung J (2011b) Applying a combined fuzzy systems and GARCH model to adaptively forecast stock market volatility. *Appl Soft Comput* 11(5):3938–3945
- Kasabov N (2002) DENFIS: dynamic evolving neural-fuzzy inference system and its application for time series prediction. *IEEE Trans Fuzzy Syst* 10(2):144–154
- Kasabov NK (1996) *Foundations of neural networks, fuzzy systems and knowledge engineering*. MIT Press, Cambridge
- Kung L, Yu S (2008) Prediction of index futures returns and the analysis of financial spillovers: a comparison between GARCH and the grey theorem. *Eur J Operat Res* 186(3):1184–1200
- Lemos A, Caminhas W, Gomide F (2011) Multivariable gaussian evolving fuzzy modeling system. *IEEE Trans Fuzzy Syst* 19(1):91–104
- Leng G, McGinnity TM, Prasad G (2005) An approach for on-line extraction of fuzzy rules using a self-organizing fuzzy neural network. *Fuzzy Sets Syst* 150(2):211–243
- Lima E (2008) *Modelo Takagi–Sugeno evolutivo participativo*. Master’s thesis, School of Electrical and Computer Engineering, University of Campinas
- Lima E, Hell M, Ballini R, Gomide F (2010) Evolving fuzzy modeling using participatory learning. In: Angelov P, Filev D, Kasabov N (eds) *Evolving intelligent systems*, Wiley-IEEE-Press, NJ, chap 4, pp 67–86
- Lin FJ, Lin CH, Shen PH (2001) Self-constructing fuzzy neural network speed controller for permanent-magnet synchronous motor drive. *IEEE Trans Fuzzy Syst* 9(5):751–759

- Ljung L (1999) System identification. Englewood Cliffs, NJ: Prentice Hall
- Lughofer ED (2008) FLEXFIS: a robust incremental learning approach for evolving Takagi–Sugeno fuzzy models. *IEEE Trans Fuzzy Syst* 16(6):1393–1410
- Luna I, Ballini R (2012a) Adaptive fuzzy system to forecast financial time series volatility. *J Intell Fuzzy Syst* 23(1):27–38
- Luna I, Ballini R (2012b) Online estimation of stochastic volatility for asset returns. In: *Proceeding of the IEEE Computational Intelligence for Financial Engineering & Economics (CIFER 2012)*
- Maciel L, Gomide F, Ballini R (2012a) MIMO evolving functional fuzzy models for interest rate forecasting. In: *Proceedings of the IEEE Computational Intelligence for Financial Engineering and Economics (CIFER 2012)*
- Maciel L, Gomide F, Ballini R (2012b) An enhanced approach for evolving participatory learning fuzzy modelling. In: *Proceedings of the IEEE Congress on Evolving and Adaptive Intelligent Systems (EAIS), Madrid, Spain*, pp 23–28
- Maciel L, Lemos A, Gomide F, Ballini R (2012c) Evolving fuzzy systems for pricing fixed income options. *Evolv Syst* 3(1):5–18
- McDonald S, Angelov P (2010) Evolving Takagi Sugeno modeling with memory for slow process. *Intern J Knowl Based Intell Eng Syst* 14(1):11–19
- Popov AA, Bykhanov KV (2005) Modeling volatility of time series using fuzzy GARCH models. In: *Annals of the 9th Russian–Korean International Symposium on Science and Technology*
- Qiao J, Wang H (2008) A self-organising fuzzy neural network and its applications to function approximation and forecast modeling. *Neurocomputing* 71(4–6):564–569
- Rong H, Sundararajan N, Huang G, Saratchandran P (2006) Sequential adaptive fuzzy inference system (SAFIS) for nonlinear system identification and prediction. *Fuzzy Sets Syst* 157(9):1260–1275
- Rubio JDJ (2009) SOFMLS: online self-organizing fuzzy modified least-squares network. *IEEE Trans Fuzzy Syst* 17(6):1296–1309
- Schwarz G (1978) Estimating the dimension of model. *The Annals of Statistics* 6(2):461–464
- Shaker A, Hüllermeier E (2012) IBLStreams: a system for instance-based classification and regression on data streams. *Evolv Syst* 3(4):235–249
- Tung SW, Quek C, Guan C (2011) SaFIN: A self-adaptive fuzzy inference network. *IEEE Trans Neural Netw* 22(2):1928–1940
- Tseng C, Chen S, Wang Y, Peng J (2008) Artificial neural network model of the hybrid EGARCH volatility of the taiwan stock index option prices. *Physica A: Stat Mech Appl* 387(13):3192–3200
- Yager R, Filev D (1994) Approximate clustering via the mountain method. *IEEE Trans Syst Man Cybernetics* 24(8):1279–1284
- Yager R (1990) A model of participatory learning. *IEEE Trans Syst Man Cybernetics* 20(5):1229–1234