



Revista Brasileira de Finanças

ISSN: 1679-0731

rbfin@fgv.br

Sociedade Brasileira de Finanças

Brasil

Maciel, Leandro

A Hybrid Fuzzy GJR-GARCH Modeling Approach for Stock Market Volatility Forecasting

Revista Brasileira de Finanças, vol. 10, núm. 3, 2012, pp. 337-367

Sociedade Brasileira de Finanças

Rio de Janeiro, Brasil

Available in: <http://www.redalyc.org/articulo.oa?id=305824788003>

- How to cite
- Complete issue
- More information about this article
- Journal's homepage in redalyc.org

redalyc.org

Scientific Information System

Network of Scientific Journals from Latin America, the Caribbean, Spain and Portugal

Non-profit academic project, developed under the open access initiative

# A Hybrid Fuzzy GJR-GARCH Modeling Approach for Stock Market Volatility Forecasting

(Modelo Híbrido GJR-GARCH Nebuloso para a Previsão da Volatilidade em Mercados de Ações)

Leandro Maciel\*

## Abstract

Forecasting stock market returns volatility is a challenging task that has attracted the attention of market practitioners, regulators and academics in recent years. This paper proposes a Fuzzy GJR-GARCH model to forecast the volatility of S&P 500 and Ibovespa indexes. The model comprises both the concept of fuzzy inference systems and GJR-GARCH modeling approach in order to consider the principles of time-varying volatility, leverage effects and volatility clustering, in which changes are cataloged by similarity. Moreover, a differential evolution (DE) algorithm is suggested to solve the problem of Fuzzy GJR-GARCH parameters estimation. The results indicate that the proposed method offers significant improvements in volatility forecasting performance in comparison with GARCH-type models and with a current Fuzzy-GARCH model reported in the literature. Furthermore, the DE-based algorithm aims to achieve an optimal solution with a rapid convergence rate.

**Keywords:** volatility; GARCH models; fuzzy systems; differential evolution.

**JEL codes:** C53; C61; G17.

## Resumo

A previsão da volatilidade dos retornos de ativos financeiros é uma abordagem desafiadora e tem atraído a atenção de participantes do mercado, reguladores e acadêmicos nos anos recentes. Este artigo propõe um modelo GJR-GARCH nebuloso para a previsão da volatilidade dos índices S&P 500 e Ibovespa. O modelo combina os conceitos de sistemas de inferência nebulosos e a abordagem GJR-GARCH para considerar os princípios de variância condicional, efeitos de alavancagem e agrupamentos de volatilidade, em que as flutuações são associadas por similaridade. Além disso, um algoritmo de evolução diferencial (ED) é sugerido para solucionar o problema de estimação dos parâmetros do modelo GJR-GARCH nebuloso. Os resultados indicaram que o método proposto permitiu uma melhor capacidade de previsão da volatilidade em comparação com modelos da família GARCH e com um modelo GARCH nebuloso recentemente proposto na literatura. Além disso, o algoritmo de ED considerado obteve soluções ótimas com uma rápida taxa de convergência.

---

Submitted 10 March 2012. Reformulated 22 May 2012. Accepted 1 August 2012. Published online 15 October 2012. The article was double blind refereed and evaluated by the editor. Supervising editor: Benjamin Tabak. I wish to thank the Coordination for the Improvement of Higher Education Personnel (Capes) for financial support. The suggestions and remarks from anonymous referees are also gratefully acknowledged.

\*UNICAMP, Campinas, SP, Brazil. E-mail: [maciel@dca.fee.unicamp.br](mailto:maciel@dca.fee.unicamp.br)

**Palavras-chave:** volatilidade; modelos GARCH; sistemas nebulosos; evolução diferencial.

## 1. Introduction

Accurately measuring and forecasting stock market volatility plays a crucial role for asset and derivative pricing, hedge strategies, portfolio allocation and risk management. Since the 1987 stock market crash, academics, practitioners and regulators have been investigated the development of financial time series models with changing variance over time in order to avoid huge investments losses due to their exposure to unexpected market movements (Allen & Morzuch, 2006, Carvalho *et al.*, 2005, Lin *et al.*, 2012). Indeed, volatility, as a measure of financial security prices fluctuation around its expected value, is one of the primary inputs in decision making processes under uncertainty, justifying its growing interest in the financial and economic literature (Kapetanios *et al.*, 2006, Lux & Kaizoji, 2007).

Stock returns volatility is often characterized by some stylized facts, such as volatility clusters, persistence, leptokurtic data behavior and time-varying volatility. A convenient framework for dealing with time-dependent volatility in financial markets concerns the autoregressive conditional heteroskedasticity (ARCH) model, proposed by Engle (1982), becoming a popular tool for volatility modeling. Providing a more flexible structure, Bollerslev (1986) introduced the Generalized ARCH (GARCH) model, which combines the ARCH and autoregressive moving average (ARMA) models. The GARCH model estimates jointly a conditional mean and conditional variance equation, and it is characterized by a fat tail and excess of kurtosis, regularly used in studying the daily returns of stock market data (Han & Park, 2008).<sup>1</sup>

Despite the success of GARCH model, it has been criticized for failing to capture the asymmetric volatility (Liu & Hung, 2010), since for stock prices, negative shocks to returns generally have large impacts on their volatility than positive shocks. To overcome these limitations, extensions of the GARCH model have been proposed, comprising a class of asymmetric GARCH models. The exponential GARCH (EGARCH), developed by Nelson (1991), and the GJR-GARCH, proposed by Glosten *et al.* (1993), are the main representatives of GARCH-family models assuming persistence. These models include the asymmetric responses of volatility to positive and negative shocks, but they do not simulate stock fluctuations with volatility clustering well. However, these effects can be modeled by modifications of linear models, while other require nonlinear approaches, less commonly used in practice due to their complexity (Hung, 2011b).

Methods based on artificial intelligence have been extensively applied as a flexible way to describe complex dynamics of various economic and financial problems (Haofei *et al.*, 2007, Racine, 2001). Hamid & Iqbal (2004) suggested the use of an artificial neural network (ANN) model to predict the volatility of S&P 500

<sup>1</sup>Several distinctive approaches have also been conducted for modeling and forecasting stock market volatility, including realized volatility models (Hansen & Lunde, 2006), intra-day volatility measure (Caretta *et al.*, 2011), and stochastic volatility (Fouque *et al.*, 2000).

index futures prices. Bildirici & Ersin (2009) enhanced GARCH-family models with ANN to forecast the volatility of daily returns in Istanbul Stock Exchange. On the other hand, Hajizadeh *et al.* (2012) proposed a scheme in which the estimates of volatility obtained by an EGARCH model are fed forward to an ANN model, considering the S&P 500 index prices.

A radial basis function neural network with Gaussian activation functions and robust clustering algorithms to model the conditional volatility of the Spanish electricity pool prices was suggested by Coelho & Santos (2011). The authors showed that their model performed better than traditional linear models to predict upward and downward movements in electricity future prices. Concerning the issue of derivative securities pricing, Wang (2009) integrated a GJR-GARCH model into an ANN option-pricing model and indicated that their approach provides higher predictability than other volatility methodologies. Providing similar results, Wang *et al.* (2012) and Tseng *et al.* (2008) also evaluated volatility forecasting performance in option pricing combining neural networks and GARCH-family models.

Tino *et al.* (2001) introduced a recurrent neural network model to simulate daily trading of straddles on financial indexes based on predictions of daily volatility. The authors showed that while GARCH models cannot generate any significantly positive profit, the use of recurrent networks can generate a statistically significant excess profit. Moreover, Tung & Quek (2011) jointed a self-organising neural network and option straddle-based approach to financial volatility trading. Compared with several benchmarks, the proposed methodology demonstrated that its ability to forecast the future volatility enhances investments profits. Despite the high ability to deal with the problem of volatility forecasting, ANN drawbacks include its “black box” nature, greater computational burden, proneness to overfitting, and the empirical nature of model development.

Due to these shortcomings, models based on fuzzy theory appear as an alternative methodology to evaluate high nonlinear systems (Zadeh, 2005, Savran, 2007). Popov & Bykhanov (2005) combined the concept of fuzzy rules and GARCH approach to model volatility of financial time series. The conditional volatility forecasting of foreign exchange rates returns was considered by Geng & Ma (2008), using a functional fuzzy inference system applied to the GARCH model. Hung (2009a) adopted the method of fuzzy logic systems to modify the threshold values for an asymmetric GARCH model. Based on simulations, the author showed that the forecasting performance is significantly improved if the leverage effect of clustering is considered along with the use of fuzzy systems and GARCH approaches.

Thavaneswaran *et al.* (2009) proposed a fuzzy weighted possibilistic model for option valuation based on the estimation and forecasting of financial volatility, considered as fuzzy numbers. They stated that fuzzy assumptions are more flexible and reveal promising results for option pricing as an intuitive way to look at the uncertainty in the models parameters. To capture the volatility conditional distribution on higher-order moments such as skewness, a GARCH-Fuzzy-Density method for volatility density forecasting was proposed by Helin & Koivisto (2011).

The model provided more accurate density forecasts for the higher-order moment varying processes than traditional GARCH models.

Combining the concepts of fuzzy systems and artificial neural networks, Chang *et al.* (2011) suggested the use of a hybrid adaptive network-based fuzzy inference system (ANFIS) to forecast the volatility of the Taiwan stock market. The authors indicated that the proposed model is superior to other methods with regard to error measures. Furthermore, Luna & Ballini (2012) introduced an adaptive fuzzy system to forecast financial time series volatility and compared their method with a GARCH model. The results indicate the higher performance of the adaptive fuzzy approach for volatility forecasting purposes.

A hybrid Fuzzy-GARCH model was suggested by Hung (2009b). The model comprises a functional fuzzy inference system with a GARCH model, optimized using a genetic algorithm framework. Similarly, Hung (2011a) and Hung (2011b) proposed a fuzzy system method to analyze clustering in GARCH models using genetic algorithms and particle swarm optimization to estimate the parameters, respectively. The author indicates that the model offers significant improvements in forecasting stock market volatility, overperforming some GARCH-family models.

Hence, this paper combines a GJR-GARCH model and fuzzy systems to develop a Fuzzy GJR-GARCH model. The proposed model is based on a collection of fuzzy rules in the form of IF-THEN statements, in which its structure comprises a GJR-GARCH model. The methodology considers both volatility asymmetry and volatility clustering as well. Moreover, the Fuzzy GJR-GARCH includes clustering estimation using subtractive clustering algorithm to provide a more autonomous model, based on data-knowledge, different from the approach proposed by Hung (2011a), which considers expert knowledge to determine the number of stock return clusters. Computational experiments illustrating the effectiveness of the Fuzzy GJR-GARCH model are provided by modeling and forecasting the volatility of S&P 500 (United States) and Ibovespa (Brazil) indexes from January 3, 2000 through September 30, 2011, in comparison with some GARCH-family models and a current Fuzzy-GARCH method proposed by Hung (2011a).

The choice regarding a GJR-GARCH specification is justified by the evidence of volatility asymmetry in stock market data (Martens *et al.*, 2009). Furthermore, other studies have found evidence in favor of the GJR-GARCH model (Brailsford & Faff, 1996, Taylor, 2004). For emerging stock market data, Ng & McAleer (2004) suggested that GARCH and GJR-GARCH models are superior to the Risk-Metrics (Morgan, 1996) model in forecasting stock market volatility; however, neither GARCH nor GJR-GARCH dominates the other. Despite extensive literature on volatility model evaluation, no consensus exists suggesting the most appropriate model for providing which model has the optimal performance in forecasting volatility (Liu & Hung, 2010).

The process of optimizing the Fuzzy GJR-GARCH model parameters is highly complex and nonlinear. A Differential Evolution (DE) based parameter estimation algorithm is suggested in this paper in order to provide the optimal solution for

the proposed model, since in this approach there is no requirement that the search space is differentiable or continuous, unlike some traditional optimization techniques. Nature-inspired metaheuristic, the DE algorithm, proposed by Storn & Price (1997), performs difference of the parameters vectors in order to search the optimal solution in a fitness function landscape. In a comparison to several other evolutionary computation algorithms, the main features that make the DE algorithm as an attractive optimization tool are: simplicity of implementation;<sup>2</sup> high accuracy and robustness to deal with many problems including separable, non-separable, modal and multimodal objective functions (Zhang & Sanderson, 2003); very few number of control parameters (three in classical DE); and finally, for expensive and large scale optimization problems, the space complexity of DE is quite slow as compared to most other evolutionary algorithms.

The Fuzzy GJR-GARCH model has novel features in comparison to the existing approaches in the literature. First, the proposed method combines a fuzzy scheme with a GRJ-GARCH model, providing a more realistic framework that captures both asymmetries and volatility clustering. Second, the Fuzzy GJR-GARCH model avoids the process of tuning the number of fuzzy rules by considering a subtractive clustering algorithm. Third, this paper proposes a simple and efficient differential evolution algorithm to optimize the model parameters. Finally, empirical results considers comparisons with GARCH-type models and also with a recent Fuzzy-GARCH approach proposed by Hung (2011a), estimated with particle swarm optimization as in its original version and with the differential evolution algorithm suggested.

After this introduction, the reminder of this paper is organized as follows. The proposed Fuzzy GJR-GARCH model and the current Fuzzy-GARCH approach related in the literature are described in Section 2. Section 3 discusses the DE-based optimization algorithm of the Fuzzy GJR-GARCH parameters estimation. Computational experiments evaluating the potential of the proposed method for stock market data volatility modeling and forecasting are reported in Section 4. Finally, Section 5 concludes and suggests issues for further investigation.

## 2. The Fuzzy GJR-GARCH model

### 2.1 GARCH-type models

The Generalized Autoregressive Conditional Heteroskedasticity model,  $GARCH(p, q)$ , considers the current conditional variance dependent on the  $p$  past conditional variances as well as the  $q$  past squared innovations. Let  $y_t = 100 \times (\ln P_t - \ln P_{t-1})$  denote the continuously compounded rate of stock returns from time  $t - 1$  to  $t$ , where  $P_t$  is the daily closing stock price at time  $t$ . The  $GARCH(p, q)$  model can be written as:

<sup>2</sup>Methods based on particle swarm optimization (PSO) are also straightforward to implement, however, Das *et al.* (2009) and Rahnamayan *et al.* (2008) pointed out a largely better performance of DE against PSO in some variety of problems.

$$y_t = \sigma_t \xi_t \quad (1)$$

$$\sigma_t^2 = \omega + \sum_{i=1}^q \alpha_i y_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2 \quad (2)$$

where  $\xi_t$  is a sequence of independent and identically distributed (i.i.d.) random variables with zero-mean and unit variance,  $\sigma_t^2$  is the conditional variance of  $\xi_t$ , and  $\omega$ ,  $\alpha_i$  and  $\beta_j$  are unknown coefficients to be estimated, assuming (Bollerslev, 1986):

$$\omega > 0, \quad \alpha_i \geq 0, \quad i = 1, 2, \dots, q, \quad q > 0,$$

$$\beta_j \geq 0, \quad j = 1, 2, \dots, p, \quad p > 0,$$

$$\sum_{i=1}^q \alpha_i + \sum_{j=1}^p \beta_j < 1 \quad (3)$$

The GARCH model reduces the number of parameters required by considering the information in the lag(s) of the conditional variance in addition to the lagged  $y_{t-i}^2$  term(s) as in ARCH-type models. GARCH models' simplicity and ability to capture persistence of volatility explain its empirical and theoretical appeal. However, it fails to capture the asymmetric influence of negative and positive shocks differently. On the other hand, the GJR-GARCH model provides a mechanism to account for leverage effects of price change on conditional variance. The GJR-GARCH( $p, q$ ) model is given by:

$$\sigma_t^2 = \omega + \sum_{i=1}^q \alpha_i y_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2 + \sum_{i=1}^q \gamma_i S_{t-i}^- y_{t-i}^2 \quad (4)$$

where  $S_{t-i}^- = 1$  if  $y_{t-i} < 0$ , and  $S_{t-i}^- = 0$  if  $y_{t-i} \geq 0$ , and  $\gamma_i$  are unknown parameters to be estimated. In this case, the following conditions must be satisfied (Glosten *et al.*, 1993):

$$\omega > 0, \quad \alpha_i \geq 0, \quad i = 1, 2, \dots, q, \quad q > 0,$$

$$\beta_j \geq 0, \quad j = 1, 2, \dots, p, \quad p > 0, \quad \alpha_i + \gamma_i \geq 0,$$

$$\sum_{i=1}^q \alpha_i + \sum_{j=1}^p \beta_j + \frac{1}{2} \sum_{i=1}^q \gamma_i < 1 \quad (5)$$

Despite the GJR-GARCH model includes the asymmetric responses of volatility to positive and negative shocks, it does not simulate stock fluctuations with volatility clustering well. This fact can lead to poor adequacy and forecast ability. Therefore, the proposal of a Fuzzy GJR-GARCH approach appears as a potential tool for volatility modeling and forecasting in a presence of volatility clustering and leverage effects.

## 2.2 The hybrid Fuzzy GJR-GARCH model

Fuzzy inference systems are universal approximations that can estimate nonlinear continuous functions uniformly with arbitrary accuracy (Hung, 2011b, Savran, 2007, Ji *et al.*, 2007). Besides the GJR-GARCH model considers the differential effects of the propagations of volatility caused by negative and positive shocks, a fuzzy approach provides the capability to simulate stock fluctuations with volatility clustering. The proposed Fuzzy GJR-GARCH model is describe by a collection of fuzzy rules in the form of IF-THEN statements in order to describe the stock market fluctuations via a GJR-GARCH model. Therefore, the  $k$ th rule of the Fuzzy GJR-GARCH( $p, q$ ) is written as:

$$\begin{aligned} \mathcal{R}^k : \quad & \text{IF } y_{t-1} \text{ is } \Gamma^k \text{ THEN} \\ & y_{kt} = \sigma_{kt}\xi_{kt}, \sigma_{kt}^2 = \omega_k + \sum_{i=1}^q \alpha_{ki}y_{t-i}^2 + \sum_{j=1}^p \beta_{kj}\sigma_{t-j}^2 \\ & + \sum_{i=1}^q \gamma_{ki}S_{t-i}^- y_{t-i}^2 \end{aligned} \quad (6)$$

where  $\Gamma^k$  is the  $k$ th fuzzy set (membership function) to describe the stock market return  $y$  (for  $k = 1, 2, \dots, R$ ),  $R$  is the number of fuzzy rules, and  $y_{t-1}$  is the previous value of the stock market's returns.

To describe the grade of membership of  $y_{t-1}$  in  $\Gamma^k$ , it is assumed a Gaussian membership function<sup>3</sup>:

$$\Gamma^k(y_{t-1}) = \exp\left(-\frac{1}{2}\left(\frac{y_{t-1} - c^k}{r^k}\right)^2\right) \quad (7)$$

where  $c^k$  is the center (focal point) of the  $k$ th local model, and  $r^k$  is a positive constant which defines the zone of influence of the respective cluster.

The collection of the  $R$  rules assembles a model as a combination of local models. The contribution of a local model to the overall output is proportional to the normalized degree of firing of each rule, expressed as:

<sup>3</sup>Wang and Mendel (1992) stated that a fuzzy system with Gaussian membership function has been shown to be universal approximation of any nonlinear function on a compact set.



$$\lambda^k(y_{t-1}) = \frac{\Gamma^k(y_{t-1})}{\sum_{l=1}^R \Gamma^l(y_{t-1})} \quad (8)$$

The Fuzzy GJR-GARCH model output is calculated by weighted averaging of individual rules' contributions:

$$\begin{aligned} y_t &= \sigma_t \xi_t, \text{ and } \sigma_t^2 \\ &= \sum_{k=1}^R \lambda^k(y_{t-1}) \left[ \omega_k + \sum_{i=1}^q \alpha_{ki} y_{t-i}^2 + \sum_{j=1}^p \beta_{kj} \sigma_{t-j}^2 + \sum_{i=1}^q \gamma_{ki} S_{t-i}^- y_{t-i}^2 \right] \\ &= \frac{\sum_{k=1}^R \Gamma^k(y_{t-1}) \Psi_k}{\sum_{k=1}^R \Gamma^k(y_{t-1})} \\ &= \frac{\sum_{k=1}^R \exp\left(-\frac{1}{2} \left(\frac{y_{t-1} - c^k}{r^k}\right)^2\right) \Psi_k}{\sum_{k=1}^R \exp\left(-\frac{1}{2} \left(\frac{y_{t-1} - c^k}{r^k}\right)^2\right)} \end{aligned} \quad (9)$$

where  $\Psi_k = \omega_k + \sum_{i=1}^q \alpha_{ki} y_{t-i}^2 + \sum_{j=1}^p \beta_{kj} \sigma_{t-j}^2 + \sum_{i=1}^q \gamma_{ki} S_{t-i}^- y_{t-i}^2$ .

The Fuzzy GJR-GARCH model in (9) is a nonlinear time-varying equation to model the behavior of complex dynamic systems as stock market volatility processes, including mechanisms to ensure the description of volatility stylized facts, such as volatility dependence, leverage effects and volatility clustering.

The parameters of the Fuzzy-GARCH model can be estimated by minimizing a mean squared error criterion, which is reasonable to obtain success (Haykin, 2001). Therefore, the optimization objective function is defined as:

$$E(\mathbf{x}) = \sum_{l=1}^N (y_l - \hat{y}_l)^2 \quad (10)$$

where  $\mathbf{x} = (\omega_1, \dots, \omega_R, \alpha_{11}, \dots, \alpha_{Rq}, \beta_{11}, \dots, \beta_{Rp}, \gamma_{11}, \dots, \gamma_{Rq})$  is the parameter set,  $y_l$  represents the stock market (actual) data sample for  $l = 1, 2, \dots, N$ ,  $N$  is the number of data points, and  $\hat{y}_l$  is the stock market data estimated by the Fuzzy GJR-GARCH model in Equation (9).

The objective function in (10) is a highly nonlinear function of  $\mathbf{x}$ . Simulations analysis showed that conventional gradient search methodologies produce poor estimates or even are not capable to find the global minimum of  $E(\mathbf{x})$  in Equation

(10). In this paper, to estimate the Fuzzy GJR-GARCH parameters, a differential evolution algorithm was proposed.

The number of rules and its respective focal points ( $c^k$ ) and spreads ( $r^k$ ) were set using the Subtractive Clustering (SC) algorithm, proposed by Chiu (1994). SC uses the data points as candidate prototype cluster centers. The capability of a point to be a cluster center is evaluated through its potential, a measure of the spatial proximity between a particular point  $y_i$  and all other data points:

$$\rho_i = \frac{1}{N} \sum_{j=1}^N e^{-r \|y_i - y_j\|^2}, \quad i = 1, 2, \dots, N \quad (11)$$

where  $\rho_i$  denote the potential of the  $i$ th data point,  $r$  is a positive constant which defines the zone of influence of the cluster, and  $N$  is the number of points.

The value of the potential is higher for a data point that is surrounded by a large number of close data points. Therefore, it is reasonable to establish such a point to be the center of a cluster. The potential of all other data points is reduced by an amount proportional to the potential of the chosen point and inversely proportional to the distance to this center. The next center is found also as the data point with the highest (after this subtraction) potential. The procedure is repeated until the potential of all data points is reduced below a certain threshold.<sup>4</sup>

### 2.3 Current Fuzzy-GARCH model approach

For comparison purposes, the suggested Fuzzy GJR-GARCH model was also compared with a current Fuzzy-GARCH approach related in the literature and proposed by Hung (2011a). Hence, the  $k$ th rule of the Fuzzy-GARCH( $p, q$ ) method is written as:

$$\begin{aligned} \mathcal{R}^k &: \text{ IF } y_{t-1} \text{ is } \Gamma^k \text{ THEN} \\ y_{kt} &= \sigma_{kt} \xi_{kt}, \quad \sigma_{kt}^2 = \omega_k + \sum_{i=1}^q \alpha_{ki} y_{t-i}^2 + \sum_{j=1}^p \beta_{kj} \sigma_{t-j}^2 \end{aligned} \quad (12)$$

In this case, rules consequents are comprised by a GARCH( $p, q$ ) model, instead of a GRJ-GARCH( $p, q$ ) process. Similarly as in Equation (9), the Fuzzy-GARCH model output is calculated by weighted averaging of individual rules' contributions:

$$y_t = \frac{\sum_{k=1}^R \exp\left(-\frac{1}{2} \left(\frac{y_{t-1} - c^k}{r^k}\right)^2\right) \left[\omega_k + \sum_{i=1}^q \alpha_{ki} y_{t-i}^2 + \sum_{j=1}^p \beta_{kj} \sigma_{t-j}^2\right]}{\sum_{k=1}^R \exp\left(-\frac{1}{2} \left(\frac{y_{t-1} - c^k}{r^k}\right)^2\right)} \quad (13)$$

<sup>4</sup>For more details about the subtractive clustering algorithm see Chiu (1994).

Hung (2011a) applied a particle swarm optimization (PSO) algorithm to estimate the parameter set  $\mathbf{x} = (\omega_1, \dots, \omega_R, \alpha_{11}, \dots, \alpha_{Rq}, \beta_{11}, \dots, \beta_{Rp})$  of the Fuzzy-GARCH model. Nevertheless, this paper adopts the same methodology as in Hung (2011a), i.e. using PSO, and also optimized Fuzzy-GARCH parameters with the differential evolution algorithm considered. Moreover, the number of rules and its focal points in Fuzzy-GARCH were performed with subtractive clustering as well.

### 3. Fuzzy GJR-GARCH parameters estimation with DE

The Fuzzy GJR-GARCH model estimation problem in Equation (10) is summarized as follows:

$$\min_{\mathbf{x} \in \Theta} E(\mathbf{x}) \quad (14)$$

where  $\Theta$  is the parameter space which defines viable solutions satisfying Equation (5) by the coefficients  $(\omega_1, \dots, \omega_R, \alpha_{11}, \dots, \alpha_{Rq}, \beta_{11}, \dots, \beta_{Rp}, \gamma_{11}, \dots, \gamma_{Rq})$ . The fitness (objective) function is given as follows:

$$\min_{\mathbf{x} \in \Theta} \sum_{l=1}^N \left( y_{lt} - \sqrt{\sum_{k=1}^R \lambda^k (y_{t-1}) \Psi_k} \xi_{lt} \right)^2 \quad (15)$$

where  $\Psi_k = \omega_k + \sum_{i=1}^q \alpha_{ki} y_{t-i}^2 + \sum_{j=1}^p \beta_{kj} \sigma_{t-j}^2 + \sum_{i=1}^q \gamma_{ki} S_{t-i}^- y_{t-i}^2$ .

Alternatively, the objective function associated with the Fuzzy-GARCH model is expressed as:

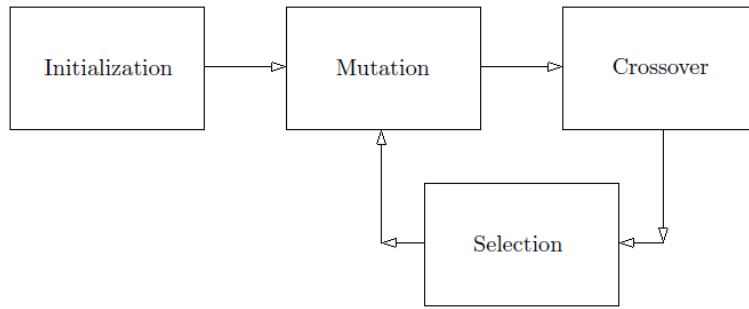
$$\min_{\mathbf{x} \in \Theta'} \sum_{l=1}^N \left( y_{lt} - \sqrt{\sum_{k=1}^R \lambda^k (y_{t-1}) \left[ \omega_k + \sum_{i=1}^q \alpha_{ki} y_{t-i}^2 + \sum_{j=1}^p \beta_{kj} \sigma_{t-j}^2 \right]} \xi_{lt} \right)^2 \quad (16)$$

where  $\Theta'$  is the parameter space of viable solutions satisfying Equation (3) by the coefficients  $(\omega_1, \dots, \omega_R, \alpha_{11}, \dots, \alpha_{Rq}, \beta_{11}, \dots, \beta_{Rp})$ .

In this paper, the optimization problems in (15) and (16) were solved by using a differential evolution algorithm for the Fuzzy GJR-GARCH and Fuzzy-GARCH models, respectively. Additionally, the PSO algorithm proposed by Hung (2011a) was considered to optimize the Fuzzy-GARCH approach as in its original version.

Differential Evolution is a vector population-based stochastic method for global optimization, introduced by Storn & Price (1997). The main idea behind the DE algorithm consists in a generation scheme of experimental parameters vectors that will disturb the population vector in order to evolve and find a desirable end. The main stages which comprises the DE algorithm are: initialization, mutation, crossover and selection. DE works through a simple cycle of these stages,

illustrated in Figure 1. Each stage is described as follows for the case of Fuzzy GARCH-type parameters optimization. Finally, the PSO algorithm (Hung, 2011a) is also reported.



**Figure 1**  
Main stages of the DE algorithm

### 3.1 Initialization

Consider an optimization problem in a  $D$ -dimensional real parameter space  $\mathbb{R}^D$ . The population  $\Phi$  is comprised by  $NP \times D$  dimensional real-valued parameter vectors, described as:

$$\Phi_{\mathbf{x},G} = [\mathbf{x}_{1,G}, \mathbf{x}_{2,G}, \dots, \mathbf{x}_{NP,G}]^T, \quad G = 0, 1, \dots, G_{max} \quad (17)$$

$$\mathbf{x}_{i,G} = [x_{1,i,G}, x_{2,i,G}, \dots, x_{D,i,G}], \quad i = 1, 2, \dots, NP \quad (18)$$

where  $G = 0, 1, \dots, G_{max}$  denotes the generation counter and  $NP$  the number of population vectors.

According to Storn & Price (1997), based on several benchmark optimization problems, a reasonable choice for  $NP$  is between  $5 \cdot D$  and  $10 \cdot D$ , but  $NP$  must be at least 4 to ensure that DE will have enough mutually different vectors with which to work. Moreover, the literature has been shown that DE control parameters setting is a specific empirical procedure (Das & Suganthan, 2011). In order to avoid non-convergence (low population size) and higher computational time complexity (high population size),  $NP = 10 \cdot D$  was considered in this paper.

Each population vector, called also as *chromosome*, forms a candidate solution to the problem. In the case of the Fuzzy GJR-GARCH model estimation problem the real parameter space has dimension  $D = (R + R \times q + R \times p + R \times q) = R(1 + 2q + p)$ . Otherwise, when the Fuzzy-GARCH model is take into account  $D = (R + R \times q + R \times p) = R(1 + q + p)$ . The initial population (at  $G = 0$ ), as in the classical DE algorithm, is set as:

$$x_{i,j,1} = \varepsilon_{i,j}, \quad j = 1, 2, \dots, D, \quad i = 1, 2, \dots, NP \quad (19)$$

where  $\varepsilon_{i,j}$  is an uniformly distributed random number lying between 0 and 1 and is instantiated independently for each component of the  $i$ th vector.

Das & Suganthan (2011) pointed out that evolutionary computation algorithms such as DE have the advantage of not depend on the choice of initial values, since any point in the parameter space has not null probability to be chosen. Moreover, the population is comprised only by viable solutions, i.e. satisfying Equations (3) and (5) for the Fuzzy-GARCH and Fuzzy GJR-GARCH models, respectively.

### 3.2 Mutation

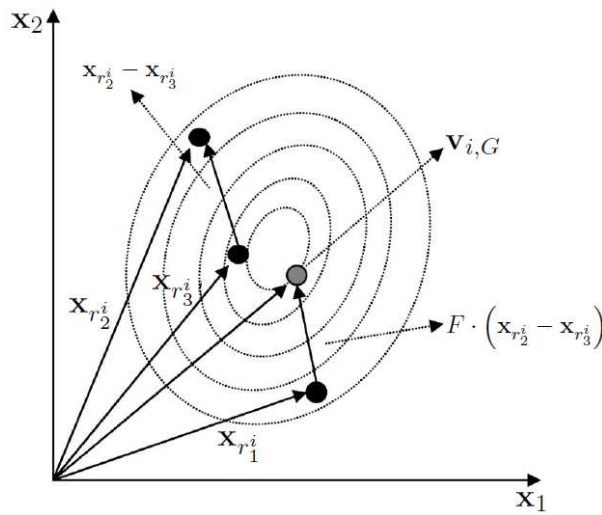
The mutation is a perturbation in the gene characteristic of a chromosome with a random element. In DE algorithms, a parent vector from the current generation is called *target* vector, a mutant vector obtained through the differential mutation operation is the *donor* vector, and an offspring formed by recombining the donor with the target vector is known as *trial* vector.

In the classical DE-mutation operator, three distinct parameter vectors  $\mathbf{x}_{r_1^i}$ ,  $\mathbf{x}_{r_2^i}$  and  $\mathbf{x}_{r_3^i}$  are sampled randomly from the current population, in order to create the donor vector for each  $i$ th target vector. The indices  $r_1^i$ ,  $r_2^i$  and  $r_3^i$ , different from the base vector index  $i$ , are mutually exclusive integers randomly chosen from the range  $[1, NP]$ . The mutation operator can be described as:

$$\mathbf{v}_{i,G} = \mathbf{x}_{r_1^i,G} + F \cdot (\mathbf{x}_{r_2^i,G} - \mathbf{x}_{r_3^i,G}) \quad (20)$$

where  $\mathbf{v}_{i,G}$  is the donor vector and  $F$  is the mutation scale factor, a scalar number usually in the interval  $[0.4, 1]$ .

The mutation process is shown in Figure 6 on a two-dimensional parameter space, including constant cost contours of an arbitrary objective function.



**Figure 2**  
DE mutation scheme in two-dimensional parametric space

In this case, the mutation uses a randomly vector  $\mathbf{x}_{r1}^i$  and only one weighted difference vector to perturb  $F \cdot (\mathbf{x}_{r2}^i - \mathbf{x}_{r3}^i)$  it. However, there are several variants of mutation operator which are different to Expression (20). The perturbations, for example, which setting the base vector to the current best vector or a linear combination of various vectors are widely applied. In this work, the classical mutation operator described was considered, since performs very well for a wide range of problems, and as stated by Price and Rönkkönen (2006), mutation operators variants need further investigation under which circumstances they perform well.

### 3.3 Crossover

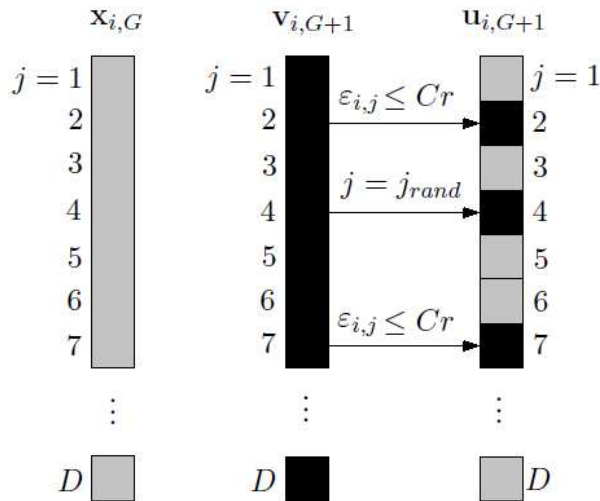
After generating the donor vector through mutation, the next step in the DE algorithm is to apply the crossover operator. This stage provides the potential diversity enhancement of the population, which exchanges parameters of the mutation vector  $\mathbf{v}_{i,G}$  and the target vector  $\mathbf{x}_{i,G}$  to generate the so-called trial vector  $\mathbf{u}_{i,G} = [u_{1,i,G}, u_{2,i,G}, \dots, u_{D,i,G}]$ . In this paper, it was considered the most common form of crossover operation, uniform (or binomial), defined as:

$$u_{j,i,G} = \begin{cases} v_{j,i,G} & \text{if } (\varepsilon_{i,j} \leq Cr \text{ or } j = j_{rand}) \\ x_{j,i,G} & \text{otherwise,} \end{cases} \quad \forall j = 1, 2, \dots, D \quad (21)$$

where  $\varepsilon_{i,j}$  is a uniformly distributed random number lying between 0 and 1, which is called a new for each  $j$ th component of the  $i$ th parameter vector,  $j_{rand} \in$

$[1, 2, \dots, D]$  is an index chosen randomly, which guarantees that  $\mathbf{u}_{i,G}$  gets at least one component from  $\mathbf{v}_{i,G}$ , and finally,  $Cr$  is called the *crossover rate*, a control parameter of DE crossover operator, which lies in the interval  $[0, 1]$  and defines the number of parameters in expectation are changed in a population member.

Figure 3 illustrates an example of a crossover process in differential evolution, where it is clear the enhance on the population diversity by exchanging some components of the target ( $\mathbf{x}$ ) and donor ( $\mathbf{v}$ ) vectors.



**Figure 3**  
DE uniform crossover scheme example

### 3.4 Selection

Finally, the *selection* step determines whether the target or the trial vector survives to the next generation  $G = G + 1$ , as the following condition:

$$\mathbf{x}_{i,G+1} = \begin{cases} \mathbf{u}_{i,G} & \text{if } E(\mathbf{u}_{i,G}) \leq E(\mathbf{x}_{i,G}) \\ \mathbf{x}_{i,G} & \text{if } E(\mathbf{u}_{i,G}) > E(\mathbf{x}_{i,G}) \end{cases} \quad (22)$$

where  $E(\mathbf{x})$  is the fitness function to be minimized.

The population size over generations is kept constant. Therefore, the new trial vector replaces the corresponding target vector in the next generation if provides an equal or lower value of the objective (fitness) function<sup>5</sup>. Otherwise, the target vector is retained in the population. It must be noted that, this selection criteria

<sup>5</sup>Objective function value and fitness concepts are interchangeable, since a lower objective function value corresponds to higher fitness, for minimization problems.

never deteriorates the fitness status, since the population either gets better or remains the same, in terms of objective function value. Nevertheless, this criteria enables DE-vectors to move over flat fitness landscapes with generations (Das & Suganthan, 2011).

### 3.5 Algorithm

Once described the classical DE algorithm steps, i.e. initialization, mutation, crossover and selection, the whole algorithm for the Fuzzy GJR-GARCH parameters identification in a pseudo-code is presented below.

---

**Algorithm.** DE algorithm for the Fuzzy GJR-GARCH identification

---

Initialize randomly a initial population  $\Phi_0 \in [0, 1]$

Initiate  $G = 1$

**while**  $G < G_{max}$  **do**

**for**  $i = 1$  to  $NP$  **do** for each individual sequentially

        Generate a donor vector  $\mathbf{v}_{i,G}$

        Generate a trial vector  $\mathbf{u}_{i,G}$

**if**  $E(\mathbf{u}_{i,G}) \leq E(\mathbf{x}_{i,G})$  **then**

$\mathbf{x}_{i,G+1} = \mathbf{u}_{i,G}$

**else**

$\mathbf{x}_{i,G+1} = \mathbf{x}_{i,G}$

**end if**

**end for**

$G = G + 1$

**end while**

---

In the algorithm suggested, the termination criteria is the number of iterations or generations,  $G_{max}$ . However, it can be defined as termination condition when the best fitness of the population does not change according to a threshold over successive generations, or alternatively, when some pre-specified objective function value was reached. The control parameters are:  $F$ ,  $Cr$  and  $G_{max}$ .

### 3.6 The Fuzzy-GARCH PSO-based algorithm

In order to optimize the Fuzzy-GARCH model proposed by Hung (2011a), this paper considered the PSO-based algorithm, as reported on its original formulation by the author, and also the differential evolution methodology suggested, previously described. This subsection presents the PSO algorithm.

Particle swarm optimization is a population-based stochastic algorithm that attempts to generate better solutions to find a desirable end (Hung, 2011a). Each swarm consists of many particles, which represents a possible solution to an optimization problem. Each particle is characterized by its position and velocity. In PSO, particles evolve in search space in three ways: i) by moving in the previous direction; ii) by moving toward the optimum (present location); or iii) by mov-



ing toward the best solution for the entire swarm (population). Over iterations an improvement in the performance is achieved.

The first step is to randomly generate a swarm of  $S$  particles in a form possible solution vectors for the Fuzzy-GARCH model. The position and velocity of the  $i$ th particle at the  $k$ th iteration are represented as  $\mathbf{p}_i^k = (p_{i,1}^k, p_{i,2}^k, \dots, p_{i,R(1+p+q)}^k)$  and  $\mathbf{v}_i^k = (v_{i,1}^k, v_{i,2}^k, \dots, v_{i,R(1+p+q)}^k)$ , respectively. Thus, the second step updates the position and velocity of each particle according to:

$$v_{i,j}^{k+1} = \chi \cdot v_{i,j}^k + \mu_1 \cdot \varepsilon_1 \cdot (l_{i,j} - p_{i,j}^k) + \mu_2 \cdot \varepsilon_2 \cdot (g_j - p_{i,j}^k) \quad (23)$$

$$p_{i,j}^{k+1} = p_{i,j}^k + v_{i,j}^{k+1} \quad (24)$$

where  $\chi$  is the linearly decreasing inertia weight,  $\mu_1$  and  $\mu_2$  are acceleration coefficients,  $\varepsilon_1$  and  $\varepsilon_2$  are uniformly distributed random variables lying between 0 and 1,  $l_{i,j}$  is the value of the objective function along the  $j$ th dimension for the best position for the  $i$ th particle, and  $g_j$  is the value of the fitness function along the  $h$ th dimension for the global best position found by all particles in the swarm (Hung, 2011a).

Parameters  $\mu_1$ ,  $\mu_2$  and  $\chi$  are related as follows (Clerc & Kennedy, 2002):

$$\chi = \frac{2}{|2 - (\mu_1 + \mu_2) - \sqrt{(\mu_1 + \mu_2)^2 - 4(\mu_1 + \mu_2)}|} \quad (25)$$

In the third step, the fitness function value  $E(\mathbf{x})$  of each particle, updated with Equations (23) and (24), is calculated. If the objective function value for the new particle is higher than that of the local optimum ( $l_{i,j}$ ) in its present location, then the local optimum is replaced by the new particle (Hung, 2011a). Moreover, if these previous condition holds and the fitness value for the new particle is higher than that of all the particles in the swarm, then this population's optimum value is also replaced by the new particle. This process repeats until the maximum number of iterations ( $G$ ) is achieved, as the termination criteria.<sup>6</sup> The PSO control parameters are:  $\mu_1$ ,  $\mu_2$ ,  $\chi$ ,  $S$  and  $G$ . Nevertheless, the number of rules and its focal points were set using subtractive clustering algorithm.

#### 4. Empirical Results and Analysis

To illustrate the performance of the proposed Fuzzy GJR-GARCH model for modeling and forecasting stock market volatility, this paper focus on daily prices of S&P 500 (US) and Ibovespa (Brazil) over the period from January 3, 2000 through September 30, 2011, in comparison with GARCH and GJR-GARCH models, and with the Fuzzy-GARCH model proposed by Hung (2011a). Moreover, the Fuzzy-GARCH approach was optimized using PSO as in its original version and using the DE algorithm suggested in this paper. The daily stock return series were generated

<sup>6</sup>For more details about the PSO-based Fuzzy-GARCH optimization algorithm see Hung (2011a).

by taking the natural logarithm difference of the daily stock index and the previous day's stock index and multiplied by 100, as described previously. The data sample was partitioned into two parts. The in-sample period consists on data from January 3, 2000 through December 29, 2005. On the other hand, the forecast out-of-sample period is from January 2, 2006 through September 30, 2011.

Table 1 shows the basic statistical characteristics of the return series. The average daily returns are negative for S&P 500 and positive for Ibovespa. The daily returns display evidence of skewness and kurtosis.

**Table 1**  
Descriptive statistics of daily returns of S&P 500 and Ibovespa.

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

<sup>a</sup> is the statistics of Jarque-Bera normal distribution test.

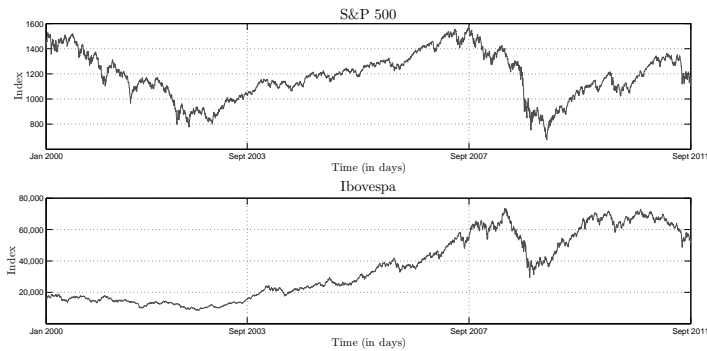
<sup>b</sup> is the Ljung-Box Q test for the 10th order serial correlation of the squared returns.

<sup>c</sup> Engle's ARCH test also examines for autocorrelation of the squared returns.

\* Significantly at 5%.

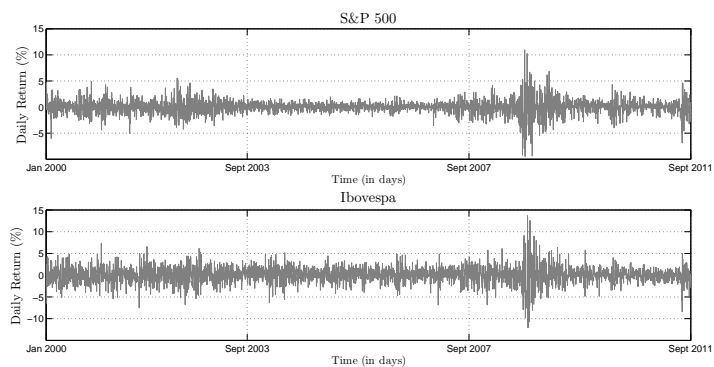
The returns series are skewed towards the left, characterized by a distribution with tails that are significantly thicker than for a normal distribution. J-B test statistics further confirms that the daily returns are non-normal distributed. As compared with Gaussian distribution, the kurtosis in S&P 500 and Ibovespa suggest that their daily returns have fat-tailed (Table 1). Ibovespa index has a higher kurtosis than S&P 500, which explains the fact that emerging countries show in general a more leptokurtic behavior. Under the null hypothesis of no serial correlation in the squared returns, the Ljung-Box  $Q^2(10)$  statistics infer a linear dependence for both series considered. Furthermore, the Engle's ARCH test for the squared returns reveals strong ARCH effects, which evidences in support of GARCH effects (i.e., heteroscedasticity). Accordingly, these preliminary analyses of the data encourage the adoption of a sophisticated model, which embody fat-tailed features, and of conditional models to allow for time-varying volatility.

Despite GARCH-type models are able to capture fat-tails and conditional volatility, they do not consider volatility clustering, as characterized by Fama (1965). The stock indexes are shown in Figure 4, and its correspondent returns are given in Figure 4. Particularly, in Figure 4 the context of volatility clustering became more clear.



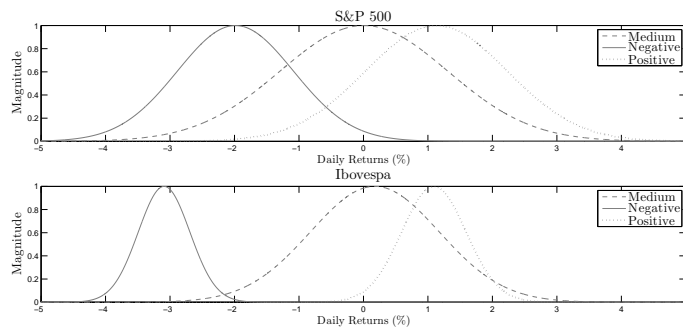
**Figure 4**  
Daily closing stock price indexes for S&P 500 and Ibovespa

In order to select best lag parameters for the Fuzzy GJR-GARCH and Fuzzy-GARCH models, and GARCH-type specifications, also considered in this work, the Bayesian information criterion (BIC) and Akaike's information criterion (AIC) were performed (Akaike, 1974, Schwarz, 1978). The models with various combinations of  $(p, q)$  parameters ranging from  $(1, 1)$  to  $(15, 15)$  were calibrated on return data. According to BIC and AIC criteria the best specification for all GARCH-type models was  $(1, 1)$ , i.e.  $p = 1$  and  $q = 1$ . The next step is to select the appropriate number of fuzzy rules for each data series, therefore, the subtractive clustering algorithm was applied. Figure 5 reports the membership functions for S&P 500 and Ibovespa indexes, which illustrates that these financial markets are asymmetric. In both economies evaluated, the clustering algorithm found three rules, that could be interpreted as *negative*, *medium* and *positive* returns.

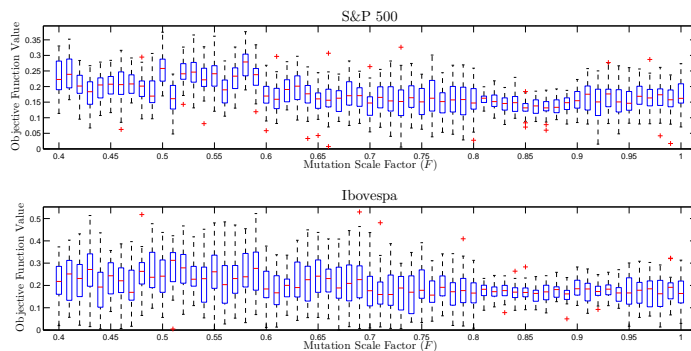


**Figure 5**  
S&P 500 and Ibovespa daily returns

Computational experiments were conducted to choose appropriate control parameters for the Fuzzy GARCH-type models, according to different values for these parameters and compared in terms of accuracy for a fixed number of generations. Simulations are comprised by 100 runs. Figures 7 and 8 display the boxplot analysis of the control parameters for S&P 500 and Ibovespa indexes, respectively, using the Fuzzy GJR-GARCH model. This approach is a convenient way of graphically depicting groups of numerical data through their five summaries: the sample minimum, sample maximum, median, upper quartile, and largest observation. A boxplot may also indicate which observations, if any, might be considered outliers. The spacings between the different parts of the box help indicate the degree of dispersion (spread) and skewness in the data. Therefore, control parameters identified by a low mean, in terms of objective function value, and low dispersion are chosen as the ones which provide better results to the problem of volatility forecasting.

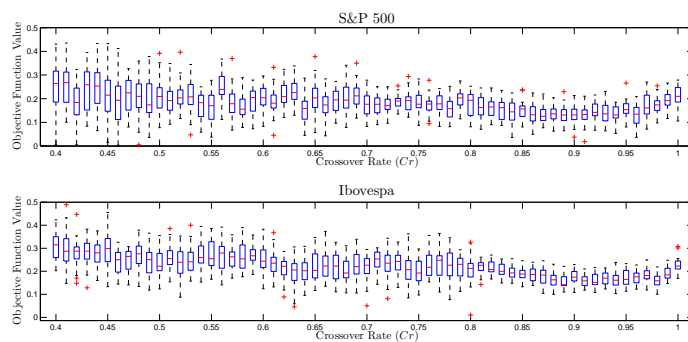


**Figure 6**  
The membership functions for the S&P 500 and Ibovespa indexes



**Figure 7**  
Boxplot of the DE algorithm for different values of mutation scaling parameter ( $F$ ) according to the objective function  $E(\mathbf{x})$  value for S&P 500 and Ibovespa

Considering the mutation scale factor ( $F$ ), shown in Figure 6, one must note that more stable results, in terms of lower median and spread, are attained on the intervals  $[0.8, 0.9]$  and  $[0.8, 0.95]$  for S&P500 and Ibovespa, respectively. Otherwise, higher spread and fitness value are observed, with high probability to provide suboptimal objective function values. On the other hand, the boxplot dispersions for the crossover rate ( $Cr$ ) are more stable for both economies evaluated (Figure 7). A range of parameters which performs better accuracy for S&P 500 lies in the interval  $[0.86, 0.89]$ , whereas for Ibovespa it corresponds the interval  $[0.89, 0.96]$ .



**Figure 8**  
Boxplot of the DE algorithm for different values of crossover rate ( $Cr$ ) according to the objective function  $E(\mathbf{x})$  value for S&P 500 and Ibovespa

Moreover, the Fuzzy-GARCH control parameters setting adopts the same procedure, i.e., by the boxplot analysis with 100 runs for both DE and PSO-based optimization techniques. Table 2 indicates the control parameters associated with each model, stock index and optimization procedure.

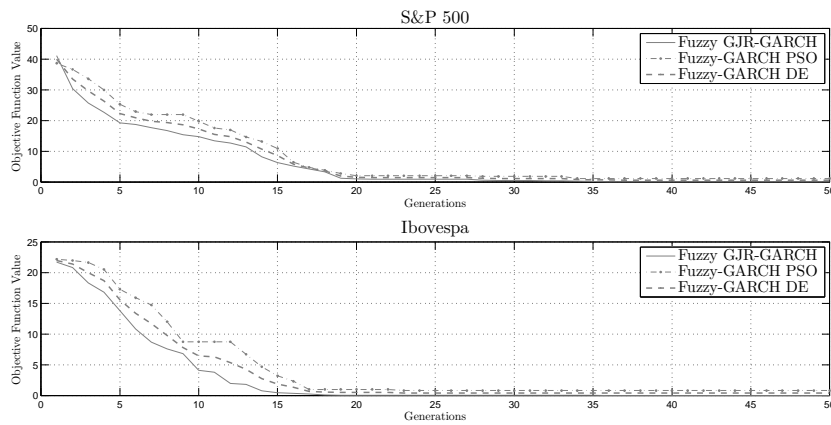
**Table 2**  
Fuzzy GARCH-family models control parameters

	Fuzzy-GARCH (PSO)		Fuzzy-GARCH (DE)		Fuzzy GJR-GARCH (DE)	
	S&P500	Ibovespa	S&P500	Ibovespa	S&P500	Ibovespa
$F$	—	—	0.80	0.79	0.85	0.84
$Cr$	—	—	0.89	0.98	0.91	0.94
$\mu_1$	2.06	2.04	—	—	—	—
$\mu_2$	2.03	2.11	—	—	—	—
$\chi$	0.74	0.68	—	—	—	—
$S$	45	50	—	—	—	—

Figure 9 shown the objective function value ( $E(\mathbf{x})$ ) behavior over the number of generations for S&P 500 and Ibovespa indexes to estimate the Fuzzy GJR-GARCH coefficients using the differential evolution algorithm proposed, and also the Fuzzy-GARCH model using particle swarm optimization and differential evolution. The fitness functions of the DE-based estimation method are exponential and rapidly converge at successful results in 25 generations, mainly when S&P 500 index is considered (Figure 9). Despite its simplicity, lower number of control parameters and no requirements about the parameter space stability, the DE also provides a fast convergence in a few number of iterations.<sup>7</sup> In terms of fit and convergence, the Fuzzy GJR-GARCH model provides better results in comparison with the Fuzzy-GARCH model optimized with PSO and DE algorithms.

<sup>7</sup>The parameters of GARCH and GJR-GARCH models were estimated using the traditional maximum likelihood method, which the log-likelihood function is computed from the product of all conditional densities of the prediction residuals.

However, these differences became more clear when the adjustment, in terms of error measures, is considered.



**Figure 9**  
Objective function  $E(\mathbf{x})$  evaluated using DE algorithm for S&P 500 and Ibovespa

The parameters estimates for in-sample data associated with the S&P 500 and Ibovespa indexes are shown in Table 3. Considering stable conditions, the parameters estimates all converge and by taking the  $t$ -statistic into account, the proposed Fuzzy GJR-GARCH model effectively estimate the daily return data.<sup>8</sup> Moreover, it must be noted that US and Brazilian markets are asymmetric and exhibit the leverage effect as confirmed by the estimated coefficients.

<sup>8</sup>A 95% confidence interval correspond as a  $t$ -statistic greater than 2, which is a measure of the number of standard deviations where the estimate for the parameter is non-zero (Hung, 2011b).

**Table 3**

Estimated parameters for the S&P 500 and Ibovespa stock indexes using the Fuzzy GJR-GARCH, Fuzzy-GARCH, optimized with PSO and DE, GARCH(1, 1) and GJR-GARCH(1, 1) models.

S&P 500				
Model	Parameter	Value	<i>t</i> -Statistic	
Fuzzy GJR-GARCH	$(\omega_1, \alpha_{11}, \beta_{11}, \gamma_{11})$	(0.004, 0.030, 0.937, 0.033)	(1.483, 3.928, 44.827, 4.092)	
	$(\omega_1, \alpha_{21}, \beta_{21}, \gamma_{21})$	(0.015, 0.028, 0.913, 0.029)	(2.546, 3.764, 41.039, 3.894)	
	$(\omega_1, \alpha_{31}, \beta_{31}, \gamma_{31})$	(0.080, 0.055, 0.798, 0.036)	(2.306, 2.192, 16.443, 9.117)	
Fuzzy-GARCH (PSO)	$(\omega_1, \alpha_{11}, \beta_{11}, \gamma_{11})$	(0.002, 0.027, 0.922, 0.028)	(1.446, 4.092, 42.112, 3.984)	
	$(\omega_1, \alpha_{21}, \beta_{21}, \gamma_{21})$	(0.015, 0.028, 0.913, 0.029)	(2.462, 3.532, 39.110, 3.913)	
	$(\omega_1, \alpha_{31}, \beta_{31}, \gamma_{31})$	(0.077, 0.061, 0.769, 0.033)	(2.376, 2.243, 17.283, 9.472)	
Fuzzy-GARCH (DE)	$(\omega_1, \alpha_{11}, \beta_{11}, \gamma_{11})$	(0.004, 0.031, 0.940, 0.035)	(1.500, 3.872, 44.956, 4.192)	
	$(\omega_1, \alpha_{21}, \beta_{21}, \gamma_{21})$	(0.011, 0.031, 0.942, 0.026)	(2.453, 3.664, 41.110, 3.980)	
	$(\omega_1, \alpha_{31}, \beta_{31}, \gamma_{31})$	(0.076, 0.060, 0.803, 0.034)	(2.299, 2.219, 15.973, 9.215)	
GARCH	$(\omega, \alpha, \beta)$	(0.059, 0.046, 0.865)	(6.532, 7.901, 40.103)	
GJR-GARCH	$(\omega, \alpha, \beta, \gamma)$	(0.022, 0.079, 0.850, 0.084)	(5.973, 8.183, 32.225, 10.927)	
Ibovespa				
Model	Parameter	Value	<i>t</i> -Statistic	
Fuzzy GJR-GARCH	$(\omega_1, \alpha_{11}, \beta_{11}, \gamma_{11})$	(0.002, 0.025, 0.844, 0.028)	(1.309, 2.827, 34.857, 4.938)	
	$(\omega_1, \alpha_{21}, \beta_{21}, \gamma_{21})$	(0.011, 0.021, 0.793, 0.034)	(3.029, 3.559, 36.294, 3.774)	
	$(\omega_1, \alpha_{31}, \beta_{31}, \gamma_{31})$	(0.077, 0.049, 0.766, 0.033)	(1.993, 2.574, 19.903, 7.358)	
Fuzzy-GARCH (PSO)	$(\omega_1, \alpha_{11}, \beta_{11}, \gamma_{11})$	(0.003, 0.029, 0.832, 0.021)	(1.422, 2.714, 32.991, 4.835)	
	$(\omega_1, \alpha_{21}, \beta_{21}, \gamma_{21})$	(0.013, 0.020, 0.769, 0.033)	(3.201, 3.445, 35.100, 3.866)	
	$(\omega_1, \alpha_{31}, \beta_{31}, \gamma_{31})$	(0.069, 0.047, 0.809, 0.030)	(2.082, 2.665, 21.022, 7.465)	
Fuzzy-GARCH (DE)	$(\omega_1, \alpha_{11}, \beta_{11}, \gamma_{11})$	(0.005, 0.030, 0.904, 0.031)	(1.509, 2.911, 36.445, 5.029)	
	$(\omega_1, \alpha_{21}, \beta_{21}, \gamma_{21})$	(0.010, 0.024, 0.765, 0.032)	(3.101, 3.448, 37.104, 3.833)	
	$(\omega_1, \alpha_{31}, \beta_{31}, \gamma_{31})$	(0.080, 0.047, 0.813, 0.030)	(2.092, 2.633, 19.223, 7.431)	
GARCH	$(\omega, \alpha, \beta)$	(0.036, 0.081, 0.794)	(4.726, 6.990, 36.772)	
GJR-GARCH	$(\omega, \alpha, \beta, \gamma)$	(0.034, 0.056, 0.881, 0.031)	(3.778, 5.994, 29.736, 9.693)	

Volatility forecasts comparison was conducted for one-step ahead horizon in terms of mean squared forecast error (MSFE), mean absolute forecast error (MAFE), and mean percentage forecast error (MPFE) defined as follows:

$$MSFE = \frac{1}{M} \sum_{i=1}^M (\sigma_i^2 - \hat{\sigma}_i^2)^2 \quad (26)$$

$$MAFE = \frac{1}{M} \sum_{i=1}^M |\sigma_i^2 - \hat{\sigma}_i^2| \quad (27)$$

$$MPFE = \frac{1}{M} \sum_{i=1}^M \frac{|\sigma_i^2 - \hat{\sigma}_i^2|}{\sigma_i^2} \quad (28)$$

where  $M$  is the number of out-of-sample observations,  $\sigma_i^2$  is the actual volatility at forecasting period  $i$ , measured as the squared daily return, and  $\hat{\sigma}_i^2$  is the forecast volatility at  $i$ .



Table 4 provides the performance of the evaluated models to predict the S&P 500 and Ibovespa stock indexes volatilities. The proposed Fuzzy GJR-GARCH model performs better than all remaining models since their structure provides a combination of rules for estimating forecast errors and also a mechanism to deal with leverage effects. The Fuzzy-GARCH model performs better than GARCH-family models, however, a better fit was attained when the model is optimized with the DE algorithm. The GJR-GARCH and GARCH models demonstrate similar results, but the GJR model presents slight lower errors than the GARCH model when the Ibovespa index is considered.

**Table 4**  
Models performance to volatility forecasting for one-step ahead.

Index	Model	MSFE	MAFE	MPFE
S&P 500	Fuzzy GJR-GARCH	0.1983	0.3652	0.3795
	Fuzzy-GARCH (PSO)	0.2573	0.5001	0.4822
	Fuzzy-GARCH (DE)	0.2209	0.4653	0.4407
	GARCH	0.5704	0.7076	0.8366
	GJR-GARCH	0.5839	0.7298	0.8420
Ibovespa	Fuzzy GJR-GARCH	0.6099	0.7912	0.2852
	Fuzzy-GARCH (PSO)	0.9726	0.9019	0.4494
	Fuzzy-GARCH (DE)	0.8801	0.8603	0.3855
	GARCH	1.4487	1.2403	0.6723
	GJR-GARCH	1.4230	1.1955	0.6511

Although all performance measures of forecasting accuracy that have been extensively employed in practice, they do not reveal whether the forecast of a model is statistically superior to another one. Therefore, it is imperative to use additional tests to help comparison among two or more competing models in terms of forecasting accuracy.

This paper adopts the parametric Morgan-Granger-Newbold (MGN) test, proposed by Diebold & Mariano (1995). This test is often employed when the assumption of contemporaneous correlation of errors is relaxed. The statistic for this test can be computed using:

$$\text{MGN} = \frac{\hat{\rho}_{sd}}{\left(\frac{1-\hat{\rho}_{sd}^2}{M-1}\right)^{\frac{1}{2}}} \quad (29)$$

where  $\hat{\rho}_{sd}$  is the estimated correlation coefficient between  $s = e_1 + e_2$ , and  $d = e_1 - e_2$ , with  $e_1$  and  $e_2$  the residuals of two models adjusted, e.g. Fuzzy GJR-GARCH model versus GARCH-type approaches. In this case, the statistic is distributed as Student distribution with  $M - 1$  degrees of freedom, and  $M$  is the number of out-of-sample observations. For this test, if the estimates are equally accurate (null hypothesis), then the correlation between  $s$  and  $d$  will be zero.

The results from MGN test, shown in Table 5, are in line with our results. MGN statistics reveal that the proposed Fuzzy GJR-GARCH model is superior predictor for forecasting S&P 500 and Ibovespa indexes than traditional GARCH-type models and the Fuzzy-GARCH model, optimized both by PSO and DE algorithms.

Considering the Fuzzy-GARCH model, one may infer that, for the different optimization algorithms, the results are competitive. Moreover, this model overperforms both GARCH-type specifications. MGN test also suggests that GARCH and GJR-GARCH are equally accurate.

**Table 5**  
MGN statistics for volatility forecast for S&P 500 and Ibovespa

S&P 500				
	Fuzzy-GARCH (PSO)	Fuzzy-GARCH (DE)	GARCH	GJR-GARCH
Fuzzy GJR-GARCH	3.0928 (0.0020)	2.7364 (0.0063)	3.7645 (0.0001)	3.6661 (0.0002)
Fuzzy-GARCH (PSO)	– –	1.8040 (0.0714)	2.9817 (0.0029)	2.8911 (0.0038)
Fuzzy-GARCH (DE)	– –	– –	2.7114 (0.0068)	2.7991 (0.0052)
GARCH	– –	– –	– –	0.3108 (0.7559)
Ibovespa				
	Fuzzy-GARCH (PSO)	Fuzzy-GARCH (DE)	GARCH	GJR-GARCH
Fuzzy GJR-GARCH	2.8115 (0.0050)	2.7365 (0.0063)	3.3937 (0.0007)	3.4928 (0.0004)
Fuzzy-GARCH (PSO)	– –	1.7265 (0.0845)	2.6998 (0.0070)	2.5874 (0.0098)
Fuzzy-GARCH (DE)	– –	– –	2.8009 (0.0052)	2.7991 (0.0052)
GARCH	– –	– –	– –	0.9530 (0.3407)

The relevant *p*-values are shown in beneath each test statistic in parenthesis.

Finally, one must compare the volatility forecasting methodologies in terms of computational complexity. Table 6 reports the processing time of the algorithms to estimate its respective parameters. All algorithms were carried out using Matlab® on a laptop equipped with 4 GB and Intel®i3CPU.

**Table 6**  
Computational processing time (in seconds) for S&P 500 / Ibovespa indexes.

Models	Fuzzy GJR-GARCH	Fuzzy-GARCH (PSO)	Fuzzy GARCH (DE)	GARCH	GJR-GARCH
Time	23.11/22.87	28.55/27.62	22.35/21.19	12.87/11.82	14.17/13.09

According to Table 6, GARCH and GJR-GARCH models are computational more efficient than all remaining models. However, the time required to estimate the Fuzzy GJR-GARCH model do not differ significantly from GARCH-type methods, which is also effective in dynamic environments. Comparing the optimizers, PSO and DE, for Fuzzy-GARCH estimation, it must be noted that DE provides a lower computational time than PSO algorithm.

The proposed hybrid model exhibits high capability to forecasting volatility of the real market returns evaluated by considering both stock market asymmetry and volatility clustering, also overperformed Fuzzy-GARCH, GARCH and GJR-GARCH methodologies in statistical terms. Moreover, a DE-based optimization algorithm improves the capability of the Fuzzy GJR-GARCH and Fuzzy-GARCH to obtain accurate parameters in a simple mechanism without restrictions about the parameter space and with a fast convergence.

## 5. Conclusion

Since the 1987 stock market crash, modeling and forecasting financial market volatility has received a great deal of attention from risk managers, regulators, academics and all markets participants in general. Volatility forecasting plays a central role in several financial applications such as asset allocation and hedging, option pricing and risk analysis. GARCH-family models are the most applied approach to model and forecast assets returns volatility, since accommodate some volatility stylized facts as persistence and time-varying volatility. However, despite the success of the GARCH model, it has been criticized for failing to capture asymmetric volatility. Therefore, the literature has been introduced a recent class of asymmetric GARCH specifications, as the GJR-GARCH model, for example. These models include the asymmetric responses of volatility to positive and negative shocks, but they do not simulate stock fluctuations with volatility clustering well.

Hence, this paper proposed a new hybrid model named Fuzzy GJR-GARCH for stock market return volatility modeling and forecasting. The model combines both fuzzy inference systems and the conditional variance GJR-GARCH model to deal with time-varying volatility in a presence of leverage effects and also volatility clustering. The proposed model is based on a collection of fuzzy rules in the form of IF-THEN statements, in which its structure comprises a GJR-GARCH model. This paper also suggested the use of a Differential Evolution (DE) algorithm to estimate the parameters of the proposed methodology, since it comprises a high nonlinear and complex optimization problem. Computational experiments illustrating the effectiveness of the Fuzzy GJR-GARCH model were provided by modeling and forecasting the volatility of S&P 500 (United States) and Ibovespa (Brazil) indexes from January 3, 2000 through September 30, 2011, in comparison with some GARCH-family models and with a current Fuzzy-GARCH method proposed by Hung (2011a), optimized with Particle Swarm Optimization (PSO), as in its original version, and with the DE algorithm proposed.

The results revealed the potential of the Fuzzy GJR-GARCH approach to the problem of volatility forecasting, providing more accurate results than GARCH-type models and the Fuzzy-GARCH method, estimated with PSO and DE algorithms, in terms of fitness accuracy and statistical tests. Moreover, the DE-based algorithm converged with stable condition in a simple structure, with low computation complexity and few control parameters, not requiring stable conditions to the

parameter space like traditional gradient search methods. Future research shall include applications of the Fuzzy GJR-GARCH approach in financial decision making problems related to volatility such as option pricing, portfolio selection and risk modeling (Value-at-Risk).

## References

- Akaike, Hirotugu. 1974. A New Look at the Statistical Model Identification. *IEEE Transactions on Automatic Control*, **19**, 716–723.
- Allen, Philip G., & Morzuch, Bernard J. 2006. Twenty- Ve Years of Progress, Problems, and Confronting Evidence in Econometric Forecasting. What About the Next 25 Years? *International Journal of Forecasting*, **22**, 475–492.
- Bildirici, Melike, & Ersin, Özgür O. 2009. Improving Forecasts of GARCH Family Models with the Artificial Neural Networks: An Application to the Daily Returns in Istanbul Stock Market Exchange. *Expert Systems with Applications*, **36**, 7355–7362.
- Bollerslev, Tim. 1986. Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, **31**, 307–327.
- Brailsford, Timothy J., & Faff, Robert W. 1996. An Evaluation of Volatility Forecasting Techniques. *Journal of Banking and Finance*, **20**, 419–438.
- Caretta, Paulo S., de Barba, Fernanda G., Vieira, Kelmara M., & Casarin, Fernando. 2011. Previsão Da Volatilidade Intradiária: Análise Das Distribuições Alternativas. *Revista Brasileira de Finanças*, **9**, 209–226.
- Carvalho, Marcelo C., Freire, Marco A. S., Medeiros, Marcelo C., & Souza, Leonardo R. 2005. Modeling and Forecasting the Volatility of Brazilian Asset Returns: A Realized Variance Approach. *Revista Brasileira de Finanças*, **4**, 321–343.
- Chang, Jing-Rong, Wei, Liang-Ying, & Cheng, Ching-Hsue. 2011. A Hybrid ANFIS Model Based on AR and Volatility for TAIEX Forecasting. *Applied Soft Computing*, **11**, 1388–1395.
- Chiu, Stephen L. 1994. Fuzzy Model Identification Based on Cluster Estimation. *Journal of Intelligent Fuzzy Systems*, **2**, 267–278.
- Clerc, Maurice, & Kennedy, James. 2002. The Particle Swarm Explosion, Stability, and Convergence in a Multidimensional Complex Space. *IEEE Transactions on Evolutionary Computation*, **6**, 58–73.
- Coelho, Leandro S., & Santos, André Alves P. 2011. A RBF Neural Network Model with GARCH Errors: Application to Electricity Price Forecasting. *Electric Power Systems Research*, **81**, 74–83.

- Das, Swagatam, & Suganthan, Ponnuthurai N. 2011. Differential Evolution: A Survey of the State-of-the-Art. *IEEE Transactions on Evolutionary Computation*, **15**, 4–31.
- Das, Swagatam, Abraham, Ajith, Chakraborty, Uday K., & Konar, Amit. 2009. Differential Evolution Using a Neighborhood Based Mutation Operator. *IEEE Transactions on Evolutionary Computation*, **13**, 526–553.
- Diebold, Francis X., & Mariano, Roberto S. 1995. Comparing Predictive Accuracy. *Journal of Business and Economic Statistics*, **13**, 253–265.
- Engle, Robert F. 1982. Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, **50**, 987–1007.
- Fama, Eugene F. 1965. The Behavior of Stock Market Price. *Journal of Business*, **38**, 34–105.
- Fouque, Jean-Pierre, Papanicolaou, George, & Sircar, Ronnie. 2000. Mean Reverting Stochastic Volatility. *International Journal of Theoretical and Applied Finance*, **3**, 101–142.
- Geng, Liyan, & Ma, Junhai. 2008. *TSK Fuzzy Inference System Based GARCH Model for Forecasting Exchange Rate Volatility*. Annals of the Fifty International Conference on Fuzzy Systems and Knowledge Discovery.
- Glosten, Lawrence R., Jagannathan, Ravi, & Runkle, David E. 1993. On the Relation Between the Expected Value and the Volatility of the Nominal Excess Return on Stocks. *Journal of Finance*, **48**, 1779–1801.
- Hajizadeh, Ehsan, Seifi, Abbas, Fazel Zarandi, Mohammad H., & Turksen, Burhan. 2012. A Hybrid Modeling Approach for Forecasting the Volatility of S&P 500 Index Return. *Expert Systems with Applications*, **39**, 431–436.
- Hamid, Shaikh A., & Iqbal, Zahid. 2004. Using Neural Networks for Forecasting Volatility of S&P 500. *Journal of Business Research*, **57**, 1116–1125.
- Han, Heejoon, & Park, Joon Y. 2008. Time Series Properties of ARCH Processes with Persistent Covariates. *Journal of Econometrics*, **146**, 275–292.
- Hansen, Peter R., & Lunde, Asger. 2006. Realized Variance and Market Microstructure Noise (with Discussion). *Journal of Business and Economic Statistics*, **24**, 127–161.
- Haofei, Zou, Guoping, Xia, Fangting, Yang, & Han, Yang. 2007. A Neural Network Model Based on the Multi-Stage Optimization Approach for Short-Term Food Pricing Forecasting in China. *Expert Systems with Applications*, **33**, 347–356.

- Helin, Tomi, & Koivisto, Hannu. 2011. The GARCH-FuzzyDensity Method for Density Forecasting. *Applied Soft Computing*, **11**, 4212–4225.
- Hung, Jui-Chung. 2009a. A Fuzzy Asymmetric GARCH Model Applied to Stock Markets. *Information Sciences*, **179**, 3930–3943.
- Hung, Jui-Chung. 2009b. A Fuzzy GARCH Model Applied to Stock Market Scenario Using a Genetic Algorithm. *Expert Systems with Applications*, **36**, 11710–11717.
- Hung, Jui-Chung. 2011a. Adaptive Fuzzy-GARCH Model Applied to Forecasting the Volatility of Stock Markets Using Particle Swarm Optimization. *Information Sciences*, **181**, 4673–4683.
- Hung, Jui-Chung. 2011b. Applying a Combined Fuzzy Systems and GARCH Model to Adaptively Forecast Stock Market Volatility. *Applied Soft Computing*, **11**, 3938–3945.
- Ji, Yanqing, Massanari, Robert M., Ager, Joel, Yen, John, Miller, Richard E., & Ying, Hao. 2007. A Fuzzy Logic-Based Computational Recognition-Primed Decision Model. *Information Sciences*, **177**, 4338–4353.
- Kapetanios, George, Labhard, Vincent, & Price, Simon. 2006. Forecasting Using Predictive Likelihood Model Averaging. *Economic Letters*, **91**, 373–379.
- Lin, Edward M. H., Chen, Cathy W. S., & Gerlach, Richard. 2012. Forecasting Volatility with Asymmetric Smooth Transition Dynamic Range Models. *International Journal of Forecasting*, **28**, 384–399.
- Liu, Hung-Chun, & Hung, Jui-Chung. 2010. Forecasting S&P-100 Stock Index Volatility: The Role of Volatility Asymmetry and Distributional Assumption in GARCH Models. *Expert Systems with Applications*, **37**, 4928–4934.
- Luna, Ivette, & Ballini, Rosangela. 2012. Adaptive Fuzzy System to Forecast Financial Time Series Volatility. *Journal of Intelligent & Fuzzy Systems*, **23**, 27–38.
- Lux, Thomas, & Kaizoji, Taisei. 2007. Forecasting Volatility and Volume in the Tokyo Stock Market: Long Memory, Fractality and Regime Switching. *Journal of Economic Dynamics and Control*, **31**, 1808–1843.
- Martens, Martin, Van Dijk, Dick, & de Potter, Michiel. 2009. Forecasting S&P 500 Volatility: Long Memory, Level Shifts, Leverage Effects, Day-of-the-Week Seasonality, and Macroeconomic Announcements. *International Journal of Forecasting*, **25**, 282–303.
- Nelson, Daniel B. 1991. Conditional Heteroskedasticity in Asset Returns: A New Approach. *Econometrica*, **59**, 347–370.

- Ng, Hock G., & McAleer, Michael. 2004. Recursive Modeling of Symmetric and Asymmetric Volatility in the Presence of Extreme Observations. *International Journal of Forecasting*, **20**, 115–129.
- Popov, Alexander A., & Bykhanov, Kolya V. 2005. *Modeling Volatility of Time Series Using Fuzzy GARCH Models*. Annals of the 9th Russian-Korean International Symposium on Science and Technology, 687-692, Novosibirsk, Russia.
- Racine, Jeffrey. 2001. On the Nonlinear Predictability of Stock Returns Using Financial and Economic Variables. *Journal of Business and Economic Statistics*, **19**, 380–382.
- Rahnamayan, Shahryar, Tizhoosh, Hamid R., & Salama, Magdy A. 2008. Opposition-Based Differential Evolution. *IEEE Transactions on Evolutionary Computation*, **12**, 64–79.
- Savran, Aydogan. 2007. An Adaptive Recurrent Fuzzy System for Nonlinear Identification. *Applied Soft Computing*, **7**, 593–600.
- Schwarz, Gideon. 1978. Estimating the Dimension of Model. *The Annals of Statistics*, **6**, 461–464.
- Storn, Rainer, & Price, Kenneth V. 1997. Differential Evolution – A Simple and Efficient Heuristic for Global Optimization over Continuous Spaces. *Journal of Global Optimization*, **11**, 341–359.
- Taylor, James W. 2004. Volatility Forecasting with Smooth Transition Exponential Smoothing. *International Journal of Forecasting*, **20**, 273–286.
- Thavaneswaran, Aerambamoorthy, Appadoo, Srimantoora S., & Paseka, Alexander. 2009. Weighted Possibilistic Moments of Fuzzy Numbers with Applications to GARCH Modeling and Option Pricing. *Mathematical and Computer Modeling*, **49**, 352–368.
- Tino, Peter, Schittenkopf, Christian, & Dorffner, Georg. 2001. Financial Volatility Trading Using Recurrent Neural Networks. *IEEE Transactions on Neural Networks*, **12**, 865–874.
- Tseng, Chih-Hsiung, Cheng, Sheng-Tzong, Wang, Yi-Hsien, & Peng, Jin-Tang. 2008. Artificial Neural Network Model of the Hybrid EGARCH Volatility of the Taiwan Stock Index Option Prices. *Physica A: Statistical Mechanisms and its Applications*, **387**, 3192–3200.
- Tung, Whye L., & Quek, Chai. 2011. Financial Volatility Trading Using a Self-Organising Neural-Fuzzy Semantic Network and Option Straddle-Based Approach. *Expert Systems with Applications*, **38**, 4668–4688.

- Wang, Ching-Ping, Lin, Shin-Hung, Huang, Hung-Hsi, & Wu, Pei-Chen. 2012. Using Neural Network for Forecasting TXO Price under Different Volatility Models. *Expert Systems with Applications*, **39**, 5025–5032.
- Wang, Yi-Hsien. 2009. Nonlinear Neural Network Forecasting Model for Stock Index Option Price: Hybrid GJR-GARCH Approach. *Expert Systems with Applications*, **36**, 564–570.
- Zadeh, Lotfi. 2005. Toward a Generalized Theory of Uncertainty (GTU) – An Outline. *Information Sciences*, **172**, 1–40.
- Zhang, Jingqiao, & Sanderson, Arthur C. 2003. JADE: Adaptive Differential Evolution with Optional External Archive. *IEEE Transactions on Evolutionary Computation*, **13**, 945–958.