

Research Article

Improving Forecasts of the EGARCH Model Using Artificial Neural Network and Fuzzy Inference System

Geleta T. Mohammed ¹, Jane A. Aduda,² and Ananda O. Kube^{2,3}

¹Mathematics Department, Pan African University (PAU), Nairobi 62000 00200, Kenya

²Mathematics Department, Jomo Kenyatta University, Nairobi, Kenya

³Mathematics Department, Kenyatta University, Nairobi, Kenya

Correspondence should be addressed to Geleta T. Mohammed; geletah2016@gmail.com

Jane A. Aduda and Ananda O. Kube contributed equally to this work.

Received 3 March 2020; Accepted 16 May 2020; Published 24 June 2020

Academic Editor: Pierpaolo D'Urso

Copyright © 2020 Geleta T. Mohammed et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper proposes an innovative *semiparametric nonlinear fuzzy-EGARCH-ANN* model to solve the problem of accurate modeling for forecasting stock market volatility. This model has been developed by a combination of the FIS, ANN, and EGARCH models. Because the proposed model is highly nonlinear and gradient-based parameter estimation methods might not give global optimal parameters for highly nonlinear models, the study has decided to use evolutionary algorithms instead. In particular, a differential evolution (DE) algorithm is suggested to solve the parameter estimation problem of the proposed model. After this, the semiparametric nonlinear fuzzy-EGARCH-ANN model has been developed mathematically from the three models mentioned before, and the study has simulated data by it. After the simulation, parameter estimation of the proposed model using a differential evolution algorithm on the simulated data is done. Finally, it is seen that the proposed model is good in capturing the volatility clustering and leverage effects of highly nonlinear and complicated financial time series data that were overlooked by the EGARCH model.

1. Introduction

Modeling and forecasting volatility of the stock market have gained great attention from researchers in academia as well as in financial markets due to its wide range of applications [1]. Investment decisions in financial markets powerfully depend on the forecast of expected returns and volatility of the assets. Volatility delivers a measure of fluctuation in a financial security price around its expected value [1]. Expected market return is correlated to predictable stock market volatility [1]. Because of the usefulness of volatility prediction, a lot of time series models have been established for financial data with time-varying variance. But all these models are still deficient in several aspects [2]. An important task in applied research is to decide which of the many possible volatility models one should employ in any given situation. The work by [3, 4] attempts to provide some guidance in this respect by evaluating the modeling ability of

various statistical procedures. Unfortunately, their results suggest that none of the popular employed models are able to adequately capture asymmetric effects and volatility clustering, at least for the particular stock market indices and time periods [2].

Even though GARCH family models play a successful role in modeling and forecasting volatility, they fail to capture the asymmetric volatility [5]. This is because the volatility of stock prices is significantly affected by negative shocks to returns rather than positive shocks. Though many financial time series observations have a nonlinear dependence structure, a linear correlation structure is usually assumed among the time series data [1]. However, conventional time series models produce forecasts based on some strict statistical assumptions about data distributions, and, therefore, they are not very appropriate to forecast financial data sets [6]. Also, due to their complexities, nonlinear models are in very limited use today. Considering

these difficulties, there is currently a demand for more flexible models.

In the evolution of time series models, researchers have made wide efforts to take advantage of artificial intelligence to process a broad amount of information and to forecast financial markets leading to an increase in the investment return. To increase the predictive power of the financial time series models, the traditional time series models are hybridized with the neural networks for volatility prediction of different benchmark financial data sets [7–9]. The attempt for the hybrid system is to outperform the forecast results and overcome the shortcomings by extracting input variables from statistical methods and include them in the ANN learning process. Regardless of the high ability to deal with the problem of volatility forecasting, ANN drawbacks include its “black box” nature, more computational burden, proneness to overfitting, and the empirical nature of model development. Again, to overcome the drawbacks of neural network models and to tackle the uncertainties that exist in precise forecasting of stock market volatility, several fuzzy time series models have been evolved [1].

The hybrid fuzzy time series models proposed in references [10, 11] have shown important enhancements in forecasting stock market volatility, outperforming the traditional time series models, neural networks, other hybrid models, etc. Adaptive neurofuzzy information system (ANFIS) is a different popular hybrid model used in volatility forecasting [12–15]. The ANFIS model gives the advantage of combining the rules in the rule base of fuzzy theory to describe the complex relationships between the variables and the learning capability of the neural network to amend the membership functions and the rule base [1]. Furthermore, the ANFIS model offers a superior approach of using the low-level learning, the computational power of the neural network, and the high-level human-like thinking and reasoning of the fuzzy system [1].

Although there have been numerous studies on volatility model evaluation for stock markets, no optimal model exists to provide an accurate forecast of volatility indices [5]. Thus, the motivation of this study is to increase the ability of the time series model in forecasting return volatility by integrating it with a suitable machine learning approach [1]. Therefore, the proposal of numerous hybrid time series models seems like a potential device for volatility modeling and forecasting. Hence, the study has developed a new model by combining the EGARCH model with the ANN and FIS to handle the issue of not simulating stock fluctuations with volatility clustering and asymmetric effects overlooked by the EGARCH model.

The paper is organized into seven sections. Not together from the introduction in Section 1, GARCH family models are presented in Section 2. In Section 3, detail of the EGARCH-ANN model is given. In Section 4, a new fuzzy-EGARCH-ANN model is presented to handle the stochastic nature of stock market volatility. A differential evolution approach is given in Section 5 to update the weights of the fuzzy-EGARCH-ANN model, and Section 6 provides the simulation study of the stock market volatility using the

newly proposed model. Concluding remarks are given in Section 7.

2. GARCH Family Models

Among a number of time series models, the GARCH models proposed by Bollerslev appear to be the most successful and popular form for modeling and forecasting the conditional variance of the return of volatility [16, 17]. The GARCH (p, q) model considers the current conditional variance dependent on the p past conditional variances as well as the q past squared innovations [18]. Let P_t be the daily closing stock price at time t and $Y_t = 100(\ln P_t - \ln P_{t-1})$ be the continuously compounded rate of stock returns from time $t - 1$ to t . Now, unpredictable component of Y_t is given by $y_t = Y_t - E(Y_t | Y_{t-1}, \dots, Y_{t-p}) = Y_t - E(Y_t) = \sigma_t \varepsilon_t$ [19]. The GARCH (p, q) model can be written as [5]

$$y_t = \sigma_t \varepsilon_t, \quad (1)$$

$$\sigma_t^2 = w + \sum_{i=1}^p \alpha_i y_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2,$$

where $w > 0, \alpha_i \geq 0, \beta_j \geq 0$, and (ε_t) is the white noise with $E(\varepsilon_t) = 0$ and $E(\varepsilon_t^2) = 1$. In this model, σ_t^2 is the conditional variance of y_t . In addition to the nonnegativity of the parameters here, there is parameter restriction

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

to ensure the positivity of the conditional variance.

GARCH models' simplicity and ability to capture persistence of volatility explain its empirical and theoretical appeal [5]. However, despite the success of the GARCH model, it fails to capture the asymmetric volatility. This limitation has been managed somehow by introducing more flexible volatility treatments by accommodating the asymmetric responses of volatility to positive and negative shocks. Two simple classes of models that capture the asymmetric behavior of returns are the GJR-GRACH model proposed by Glosten and the exponential GARCH (EGARCH) introduced by Nelson [5]. The GJR-GARCH model allows the conditional variance to respond differently to past negative and positive innovations. The GJR-GARCH (p, q) model is given by [1]

$$y_t = \sigma_t \varepsilon_t, \quad (3)$$

$$\sigma_t^2 = w + \sum_{i=1}^p \alpha_i y_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2 + \sum_{i=1}^p \gamma_i s_{t-i}^- y_{t-i}^2,$$

where $s_{t-i}^- = 1$ if $y_{t-i} < 0$ else $s_{t-i}^- = 0$.

At this point, the variable s_{t-i}^- distinguishes negative shocks from positive shocks, so that asymmetric effects on the data are captured by $w > 0, \alpha_i \geq 0, \beta_j \geq 0$, and $\gamma_i \geq 0$ parameters with the constrains, $\alpha_i + \gamma_i \geq 0$ and

$$\sum_{i=1}^p \alpha_i + \sum_{j=1}^q \beta_j + \sum_{i=1}^p \gamma_i < 1. \quad (4)$$

The exponential GARCH (EGARCH) model provides an alternative asymmetric model by considering the leverage effects of a price change on the conditional variance. That means a large price decline can have a bigger impact on volatility than a large price increase. The EGARCH (p, q) model can be represented as

$$y_t = \sigma_t \varepsilon_t,$$

$$\ln \sigma_t^2 = w + \sum_{j=1}^q \beta_j \ln \sigma_{t-j}^2$$

$$+ \sum_{i=1}^p \gamma_i \left(\frac{|y_{t-i}|}{\sigma_{t-i}} - E \left(\frac{|y_{t-i}|}{\sigma_{t-i}} \right) \right) + \sum_{i=1}^p \alpha_i \frac{y_{t-i}}{\sigma_{t-i}},$$
(5)

where (ε_t) is a sequence of i.i.d random variables, $E(\varepsilon_t) = 0$, and $E(\varepsilon_t^2) = 1$, with no parameter restrictions. Finally, the use of the log form allows the parameters to be negative without the conditional variance becoming negative [1].

Although the GJR-GARCH and EGARCH models include the asymmetric responses of volatility to positive and negative shocks, they do not simulate stock fluctuations with volatility clustering accurately. This fact can lead to poor adequacy and forecasting ability. Therefore, the proposal of several hybrid time series models appears as a potential tool for volatility modeling and forecasting [1]. In this paper, the basic idea is to propose a unified framework of a computing model that adopts the merits of both ANN and FIS with an efficient statistical time series model like EGARCH.

3. EGARCH-ANN Model

As mentioned in the previous section, ANFIS is also a hybrid time series model for volatility prediction. ANNs have been broadly applied to answer many hard problems in different areas, including pattern recognition, signal processing, and language learning, as mentioned in [20, 21]. Since the pioneer work of White [22], there have also been abundant applications of ANN in economics and finance [1]. Unfortunately, the ANN literature is not easy to enter, so it is hard for applied economists to understand why ANN works and how it can be applied appropriately [22, 23].

As in [24], the study prefers to consider a network with three layers called an input layer, a hidden layer, and an output layer. The input layer contains m input nodes, and the output layer contains v output nodes in such a way that each nodes of the input layer corresponds to a particular input variable, and each nodes of the output layer corresponds to a particular output variable, respectively. There are total of s hidden nodes in the hidden layer which are connected to all input and output nodes. The strengths of such connections are assigned by parameters which are unknown and are called the network connection weights. In particular, $W_h = (w_{h,1}, \dots, w_{h,m})'$ represents the vector of the connection weights between the h^{th} hidden node and all m input nodes, and $\xi_j = (\xi_{j,1}, \dots, \xi_{j,s})'$ represents the vector of the connection weights between the j^{th} output node and all s hidden nodes.

Formally, let us say that the input vector is $Z_t = (z_{t,1}, \dots, z_{t,m})$. The input nodes receives the information Z_t and send to all hidden nodes, weighted by the connection weights between the input and hidden nodes. The information Z_t is then transformed from each input nodes by the activation function G to each hidden node. That is, the h^{th} hidden node receives $Z_t W_h + w_{h,0}$ and transforms it to $G(Z_t W_h + w_{h,0})$ where $w_{h,0}$ is the bias term in the h^{th} hidden node. The information generated by all hidden nodes is further passed to the output nodes, again weighted by the corresponding connection weights, and transformed by the activation function F in each output node. Hence, the j^{th} output node receives $\sum_{h=1}^s \xi_{j,h} G(Z_t W_h + w_{h,0}) + \xi_{j,0}$ and transforms it into the network output:

$$o_{t,j} = F \left(\xi_{j,0} + \sum_{h=1}^s \xi_{j,h} G(Z_t W_h + w_{h,0}) \right), \quad \text{for } j = 1 : v,$$
(6)

where $\xi_{j,0}$ is the bias term in the j^{th} output node and the output $o_{t,j}$ is used to describe or predict the behavior of the j^{th} target Z_t . As far as model specifications are concerned, the building blocks of an ANN model are the activation functions F and G . Different choices of the activation functions result in different network models. But, this study has chosen the most common activation functions. That is, $F(x) = x$ and $G(z) = \tanh(Z)$. In addition to this, the study has considered only the output layer of one hidden unit.

From the given lessons learned from the existing literature and the above preliminary concepts [19, 24–27], the hybrid model from the EGARCH model and ANN is given by EGARCH-ANN (p, q, m, s):

$$y_t = \sigma_t \varepsilon_t,$$

$$\ln \sigma_t^2 = E + \sum_{j=1}^q \beta_j \ln \sigma_{t-j}^2 + \sum_{i=1}^p \gamma_i \left(\frac{|y_{t-i}|}{\sigma_{t-i}} - E \left(\frac{|y_{t-i}|}{\sigma_{t-i}} \right) \right)$$

$$+ \sum_{i=1}^p \alpha_i \frac{y_{t-i}}{\sigma_{t-i}} + \sum_{h=1}^s \xi_h G(Z_t W_h + w_{h,0}),$$
(7)

where

$$G(Z_t W_h + w_{h,0}) = \tanh(Z_t W_h + w_{h,0}),$$

$$W_h = (w_{h,1}, \dots, w_{h,m})',$$

$$z_{t-d} = \frac{y_{t-d} - E(y)}{\sqrt{E(y^2)}}, \quad d = 1 : m,$$

$$E = \xi_0 + w,$$
(8)

in which σ_t^2 is the asymmetric function of y_t and $G(Z_t W_h + w_{h,0})$, where $E(y)$ and $E(y^2)$ are the in-sample mean and the in-sample variance, respectively; both $w_{h,0}$ and the elements of W_h are going to be generated by a uniform random number generator, and they lie between -1 and 1 [2, 5]. After having the values of p, q, m , and s in hand, the

estimation method for EGARCH-ANN (p, q, m, s) can be MLE based on the preferred distribution of ε_t . This EGARCH-ANN hybrid model can handle the leverage which is overlooked by the EGARCH model by the help of ANN's universal approximation nature [2]. That is, because the ANN component can add the ability to handle the nonlinear relationship of the past return innovations and its future volatility to the EGARCH model [28]. In other words, this hybrid model is good in modeling and forecasting volatility than that of the EGARCH model because the ANN component helps it to handle the leverage effect stylized fact of volatility overlooked by the EGARCH model [1, 2].

In the same way, one among many means of combining GARCH-type models with FIS to build a hybrid model is the fuzzy-GARCH-type hybrid models [20]. That is, just by applying FIS on GARCH-type models. In other words, it is just building fuzzy-GARCH-type hybrid models by applying a fuzzy inference system on GARCH-type models [28]. For the same reason mentioned before, now the study wants to focus only on the fuzzy-EGARCH hybrid model among other fuzzy-GARCH-type hybrid models [1]. In this model, the fuzzy component gives the hybrid model the ability to simulate stock price variations with volatility clustering. And also, in this model, leverage effects of a price change on volatility are captured by the EGARCH model.

4. Proposed Fuzzy-EGARCH-ANN Model

From the given lessons learned from the existing literature and some other preliminary concepts of FIS [29–32], the study has developed a hybrid model from FIS, and the EGARCH-ANN model by the same methodology of R. Dash and D. K. Dash [1] has been applied to build the fuzzy-EGARCH model. In a similar fashion to the formation of the fuzzy-EGARCH hybrid model, the study has considered the fuzzy approach as it provides the capability to simulate stock fluctuations with volatility clustering, whereas the asymmetry and leverage effects of price change on the conditional variance is captured by the EGARCH-ANN model. This fuzzy-EGARCH-ANN model is described by a collection of fuzzy rules in the form of If-Then statements in order to describe the stock fluctuations with volatility clustering overlooked by the EGARCH-ANN model via fuzzy rules and the asymmetric responses of volatility to positive and negative shocks via an EGARCH-ANN model.

The k^{th} rule of the fuzzy system for the EGARCH-ANN (p, q, m, s) model is described as follows.

Rule ^{k} : if x_1^t is F_{k1} and \dots x_d^t is F_{kd} , then

$$y_{kt} = \sigma_{kt} \varepsilon_{kt},$$

$$\begin{aligned} \ln \sigma_{kt}^2 = & E_k + \sum_{j=1}^q \beta_{kj} \ln \sigma_{t-j}^2 + \sum_{i=1}^p \gamma_{ki} \left(\frac{|y_{t-i}|}{\sigma_{t-i}} - E \left(\frac{|y_{t-i}|}{\sigma_{t-i}} \right) \right) \\ & + \sum_{i=1}^p \alpha_{ki} \frac{y_{t-i}}{\sigma_{t-i}} + \sum_{h=1}^s \xi_{kh} G(Z_t W_h + w_{h,0}), \end{aligned} \quad (9)$$

where

$$\begin{aligned} G(Z_t W_h + w_{h,0}) &= \tanh(Z_t W_h + w_{h,0}), \\ W_h &= (w_{h,1}, \dots, w_{h,m})', \\ z_{t-d} &= \frac{y_{t-d} - E(y)}{\sqrt{E(y^2)}}, \quad d = 1: m, \\ x^t &= [y_{t-1}, y_{t-2}, \dots, y_{t-p}, \sigma_{t-1}^2, \sigma_{t-2}^2, \dots, \sigma_{t-q}^2]^T \\ &= [x_1^t, x_2^t, \dots, x_d^t]', \end{aligned} \quad (10)$$

in which x^t is the input vector with $d = (p + q)$ at instance t , F_{kl} is the fuzzy set to describe the stock market return and volatility for $l = 1, 2, 3, \dots, d$, R is the number of rules, and x_l^t is the premise variable. In addition, the study has assumed that the distribution of residual series follows either the Gaussian normal or Student's t -distribution. That is, if it is not following the Gaussian normal distribution, then it follows the Student's t -distribution directly.

In this fuzzy-EGARCH-ANN hybrid model, the premise variables include both the previous value of the stock market return and the previous volatility of the stock market return. The then part clearly expresses a local input and output relation by considering the EGARCH-ANN model parameters. Like in the previous EGARCH-ANN hybrid model, this fuzzy-EGARCH-ANN hybrid model σ_t^2 is not only an asymmetric function of y_t but also an asymmetric function of $G(Z_t W_h + w_{h,0})$. Wang and Mendel [33] stated that a fuzzy system with the Gaussian membership function has been shown to be a universal approximation of any nonlinear function on a compact set.

The final output of this fuzzy-EGARCH-ANN model is the weighted average output of each individual rule output and is obtained by using FIS fundamental steps [1] as follows.

Step 1 (fuzzification layer). Using the Gaussian membership function, the grade of membership of the input x_l^t in F_{kl} is found as follows:

$$F_{kl}(x_l^t) = \exp \left[-\frac{1}{2} \left(\frac{x_l^t - c_{kl}}{a_{kl}} \right)^2 \right], \quad (11)$$

where c_{kl} is the center and a_{kl} is the spread of k^{th} rule membership function corresponding to the l^{th} premise variable.

Step 2 (firing strength layer). The firing strength of the k^{th} rule by assuming the product T-norm of the antecedent fuzzy sets is found as

$$u_k(x^t) = \prod_{l=1}^d F_{kl}(x_l^t) = \prod_{l=1}^d \left[-\frac{1}{2} \left(\frac{x_l^t - c_{kl}}{a_{kl}} \right)^2 \right]. \quad (12)$$

Step 3 (normalization layer). Finding the normalized firing strengths, i.e., the ratio of the k^{th} rule's firing strength to the sum of all rule's firing strengths:

$$\bar{w}_k = \frac{u_k(x^t)}{\sum_{k=1}^R u_k(x^t)}. \quad (13)$$

Step 4 (consequent and defuzzification layer). Combining the normalized firing strengths and the consequent corresponding rule to produce the model output as the weighted average of each individual rule:

$$\ln \sigma_t^2 = \sum_{k=1}^R \bar{w}_k f_k, \quad (14)$$

where f_k is the output of the k^{th} rule. That is,

$$\begin{aligned} f_k = & E_k + \sum_{j=1}^q \beta_{kj} \ln \sigma_{t-j}^2 + \sum_{i=1}^p \gamma_{ki} \left(\frac{|y_{t-i}|}{\sigma_{t-i}} - E \left(\frac{|y_{t-i}|}{\sigma_{t-i}} \right) \right) \\ & + \sum_{i=1}^p \alpha_{ki} \frac{y_{t-i}}{\sigma_{t-i}} + \sum_{h=1}^s \xi_{kh} G(Z_t W_h + w_{h,0}). \end{aligned} \quad (15)$$

This implies that

$$\ln \sigma_t^2 = \sum_{k=1}^R \bar{w}_k f_k = \sum_{k=1}^R \left[\frac{u_k(x^t)}{\sum_{k=1}^R u_k(x^t)} f_k \right]. \quad (16)$$

The collection of the R rules assembles the fuzzy-EGARCH-ANN model as a combination of R local models. Every local model contributes a value to the overall output of the fuzzy-EGARCH-ANN model that is proportional to the normalized degree of the firing strength of each rule. As a result, the exponential of this weighted average value gives the predicted stock market return volatility using the newly proposed model.

Finally, Figure 1 explains the general framework of the proposed model. Figure 2 explains the architecture of the proposed model for $p=1$, $q=1$, $m=1$, $s=1$, and $R=3$. Moreover, f_i for $i=1:R$ are the EGARCH-ANN local models. In order to select the optimum value of p , q , m , and s for the fuzzy-EGARCH-ANN model, first the study obtains five different randomly produced sets of W (connection weight between the input and hidden layers including the bias term). This is what makes the model a *semiparametric* model. Then, for each of these five random W sets and for each for $p, q, m, s \in [0, 5]$, it estimates the parameters of the EGARCH-ANN (p, q, m, s) model using the maximum likelihood estimation method [2]. As a result, one can have now an optimal value for W, p, q, m , and s of EGARCH-ANN (p, q, m, s) [2]. The $[0, 5]$ grid more encompasses the usual range of models employed in the literature since for each model and index, the data always select a specification well within the inferior of the search grid, and the study believes that this grid search does not unduly constrain the parameter space [2].

The number of fuzzy rules (R) and its respective cluster center (c_{kl}) and spreads (a_{kl}) are going to be determined

from the data sample using the subtractive clustering (SC) algorithm, proposed by Chiu in 1994 [34]. 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 x_i and all other data points [35]:

$$\theta_i = \frac{1}{N} \sum_{j=1}^N e^{-a|x_i - x_j|}, \quad i = 1: N, \quad (17)$$

where θ_i denotes the potential of the i^{th} data point, a is a positive constant which defines the zone of influence of the cluster, and N is the number of data 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 [35].

After having the optimum values for p, q, m, s, R, c_{kl} , and a_{kl} in hand, the parameters of the fuzzy-EGARCH-ANN (p, q, m, s) model can be estimated by minimizing a mean-squared error criterion, which is reasonable to obtain success [25]. Therefore, the objective function of this optimization is defined as

$$E(X) = \frac{1}{N} \sum_{i=1}^N (\sigma_i^2 - \hat{\sigma}_i^2)^2, \quad (18)$$

where $X = (E_1, \dots, E_R, \alpha_{11}, \dots, \alpha_{Rq}, \beta_{11}, \dots, \beta_{Rp}, \gamma_{11}, \dots, \gamma_{Rq}, \xi_{11}, \dots, \xi_{Rs})$ is the parameter vector of the fuzzy-EGARCH-ANN hybrid model, σ_i^2 represents the realized volatility of the stock market (actual) data sample for $i = 1, 2, \dots, N$, N is the number of data points, and $\hat{\sigma}_i^2$ is the volatility of the stock market data estimated by the fuzzy-EGARCH-ANN model.

5. Fuzzy-EGARCH-ANN Model Parameter Estimation with DE

The design of the fuzzy-EGARCH-ANN model includes the determination of the unknown parameters, namely, the parameters of the consequent parts of the fuzzy rules by minimizing the prediction error. Now, the corresponding optimization problem can be written as

$$\text{Min}_{X \in \mathbb{R}^D} E(X), \quad (19)$$

where $D = R(1 + 2p + q + s)$. Such an optimization problem is usually called an unconstrained optimization problem. It is unconstrained because the parameter space of the EGARCH model is unconstrained. That is, the EGARCH model does not need restriction of its parameters unlike other GARCH-type models. This is one of the advantages of the EGARCH model give us, and that is why the study has selected it. Clearly, the objective function is a highly nonlinear function

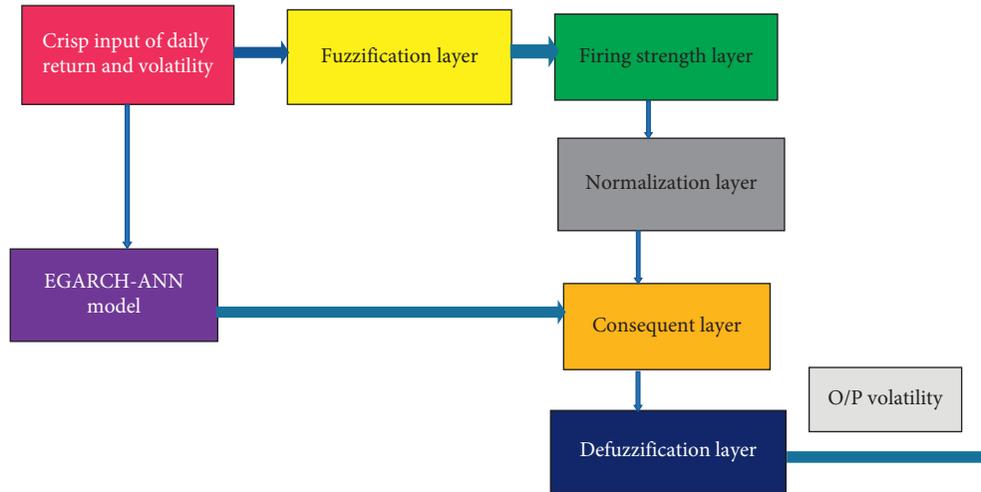


FIGURE 1: Framework of the fuzzy-EGARCH-ANN model.

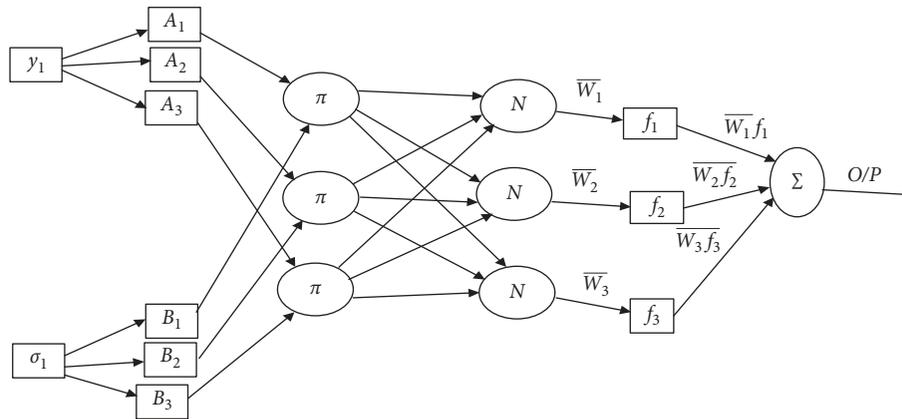


FIGURE 2: Architecture of the fuzzy-EGARCH-ANN model with $p = 1, q = 1, m = 1, s = 1,$ and $R = 3.$

of X . Simulation analysis showed that conventional gradient search methodologies produce poor estimates or even are not capable to find the global minimum of $E(X)$ [1]. Therefore, to estimate parameters the fuzzy-EGARCH-ANN model without suffering from local optimum, the study has proposed a differential evolution (DE) algorithm.

Differential evolution (DE) is a population-based stochastic function optimizer, which uses a rather greedy and less stochastic approach for problem solving in comparison to classical evolutionary algorithms, such as genetic algorithms, evolutionary programming, and PSO [36]. DE combines simple arithmetical operators with the classical operators of recombination, mutation, and selection to evolve from a randomly generated starting population to a final solution. DE also incorporates an efficient way of the self-adapting mutation using small populations. It is able to reproduce the same results consistently over many trials, unlike PSO which is more dependent on the randomized

initialization of individuals. DE has the advantage of not depending on the choice of initial values since any point in the parameter space has no null probability to be chosen [36].

The optimization process is conducted by means of three main operations: mutation, crossover, and selection. In each generation, individuals of the current population become target vectors. For each target vector, the mutation operation produces a mutant vector, by adding the weighted difference between two randomly chosen vectors to a third vector. The crossover operation generates a new vector, called a trial vector, by mixing the parameters of the mutant vector with those of the target vector. If the trial vector obtains a better fitness value than the target vector, then the trial vector replaces the target vector in the next generation.

A number of mutation strategies like rand1, rand2, best1, best2, and current to best are available in the literature [1]. In

TABLE 1: Procedure of JADE with archive and its description [38].

Procedure of JADE with archive
Start
Initialize:
$\mu_{CR} = 0.5; \mu_F = 0.5; \mathbf{A} = \emptyset$; and create a random initial population $\{X_{i,0}; i = 1: NP\}$
For $g = 1: G$; generation loop
$S_F = \emptyset; S_{CR} = \emptyset$;
For $i = 1: NP$; population size loop
Generate:
$CR_i = \text{randni}(\mu_{CR}, 0.1), F_i = \text{randci}(\mu_F, 0.1)$.
Randomly choose:
$X_{\text{best},g}^p$ as one of the 100 p % best vectors, $X_{r1,g} \neq X_{i,g}$ from current population \mathbf{P} ,
And $\tilde{X}_{r2,g} \neq X_{r1,g} \neq X_{i,g}$ from $\mathbf{P} \cup \mathbf{A}$.
Compute
$V_{i,g} = X_{i,g} + F_i * (X_{\text{best},g}^p - X_{i,g}) + F_i * (X_{r1,g} - \tilde{X}_{r2,g})$
Generate:
$j_{\text{rand}} = \text{randint}(1, D)$
For $j = 1: D$ dimension loop
If $j = j_{\text{rand}}$ or $\text{rand}(0, 1) < CR_i$ then $u_{j,i,g} = v_{j,i,g}$ else $u_{j,i,g} = x_{j,i,g}$.
End if
End for dimension loop
If $E(X_{i,g}) \leq E(U_{i,g})$ then $X_{i,g+1} = X_{i,g}$ else $X_{i,g+1} = U_{i,g}$;
$X_{i,g} \rightarrow \mathbf{A}; CR_i \rightarrow S_{CR}; F_i \rightarrow S_F$.
End if
End for population size loop
Randomly remove solutions from \mathbf{A} so that the cardinal of \mathbf{A} is less or equal to NP
$\mu_{CR} = (1 - c) * \mu_{CR} + c * \text{mean}_A(S_{CR})$
$\mu_F = (1 - c) * \mu_F + c * \text{mean}_L(S_F)$
End for generation loop
End

c is a positive constant between 0 and 1, $\text{mean}_A(\cdot)$ is the usual arithmetic mean, and $\text{mean}_L(\cdot)$ is the Lehmer mean given by

this application, we have taken the current to best mutation strategy. Like other evolutionary algorithms, the performance of DE algorithms depends on the settings of the control parameters: scaling factor F and crossover rate CR . As the optimal settings of these parameters are problem dependent, it is often necessary to tune the control parameters in order to achieve the desired results [37]. In this paper, instead of using a fixed crossover rate CR and scaling factor F , the JADE [38] approach has been used to update these iteratively. Initially, the crossover rate and the scaling factor associated with each individual are generated according to a normal (Cauchy) distribution with means $\mu_{CR}(\mu_F)$ [38]. At the end of each generation, the values of μ_{CR} and μ_F are updated according to the CR, F pair that resulted in the generation of the successful trial vector in that generation [38]. As the search progresses, μ_{CR} and μ_F should gradually approach the optimal values for the given problem [38]. Furthermore, step-by-step self-explanatory summary of the algorithm that the study has used is given in Table 1.

For your understanding, the software used throughout this study is MATLAB R2018a.

$$\text{mean}_L(S_F) = \frac{\sum_{F \in S_F} F^2}{\sum_{F \in S_F} F}. \quad (20)$$

6. Simulation Study

In this section, first of all, the study has simulated data by the fuzzy-EGARCH-ANN model with two different error distribution types, namely, the Gaussian and the Student's t -distribution. Then, the study has estimated back the parameters of the fuzzy-EGARCH-ANN model with the corresponding identified error distribution types. That is, the study has simulated data for a certain given parameters (including the connection weights of the ANN component) with a given return shocks distributions in the next two Sections 6.1 and 6.2.

6.1. Return Innovation with Gaussian Distribution. For $p = 1, q = 1, s = 2, m = 1$, sample size = 1200, Gaussian distribution type, and the parameters given in Table 2, the study has simulated data by the fuzzy-EGARCH-ANN model. As a result, the obtained basic statistical properties of these data are given in Table 3.

Furthermore, Figures 3–7 are the corresponding figures of the simulated stock price series, return series, true volatility series, residual series, and histogram of the simulated residual series, respectively.

Now, by using the first 1000 (that is, for $T=1000$) simulated data points, the study has obtained the following results. For p and q in the grid search specified in the previous section, all of EGARCH (p, q) models are not fit at all. The study has identified that the residual series follows the Gaussian normal distribution. Then, by using this distribution type, the study has obtained that the EGARCH-ANN (1, 2, 3, 1) model is the best-fitted model to the selected data. Furthermore, its fitness summary is presented in Table 4.

Now, to proceed to the parameter estimation of the fuzzy-EGARCH-ANN model, only the number of rules (R) and the realized volatility are remaining to be determined. For our case, as mentioned earlier, R is going to be determined by the subtractive clustering algorithm (SCA), and the realized volatility at time t is the square root of the square of residual series at time t . As a result, the study has obtained $R=3$, and the graph of the realized volatility is given in Figure 8.

Therefore, for $p = 1, q = 2, m = 3$, and $s = 1$, by JADE with archive, the estimated parameter vector is given in Table 5. In addition, the corresponding mean-squared forecast error (MSFE) value is $5.5127e - 04$.

TABLE 2: The parameters used for the fuzzy-EGARCH-ANN model to simulate data.

Parameters	Coefficients
E_1	0.140625
α_{11}	-0.084375
γ_{11}	0.28125
β_{11}	0.140625
ξ_{11}	-0.196875
ξ_{21}	-0.140625
E_2	-0.28125
α_{12}	-0.140625
γ_{12}	0.140625
β_{12}	0.140625
ξ_{12}	-0.196875
ξ_{22}	-0.084375
E_3	0.28125
α_{13}	-0.253125
γ_{13}	0.309375
β_{13}	0.140625
ξ_{13}	-0.225
ξ_{23}	-0.225

TABLE 3: Statistical properties of the residual data.

Statistical properties of residual data	
Min.	-0.362951
Max.	0.368273
Mean	-0.000000
Median	-0.002849
Variance	0.011601
Skewness	-0.012069
Kurtosis	3.014560
Arch test	p value = $6.1696e - 07$
Jarque-Bera test (for alpha = 0.05)	p value = 0.983

Mean < median usually flags left skewness. Negative skewness means left-skewed data. Kurtosis > 3 means larger peakedness than Gaussian. Arch test result with very small p value implies existence of strong heteroskedasticity. Jarque-Bera test result with very big p value implies that the residual series strongly follows the normal distribution.

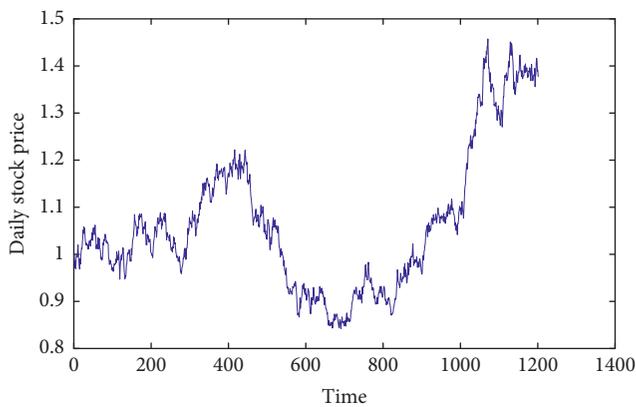


FIGURE 3: The simulated stock price.

Figure 9 presents a summary and analysis of the MSFE that has been obtained by *JADE with archive* during the parameter estimation of the proposed model.

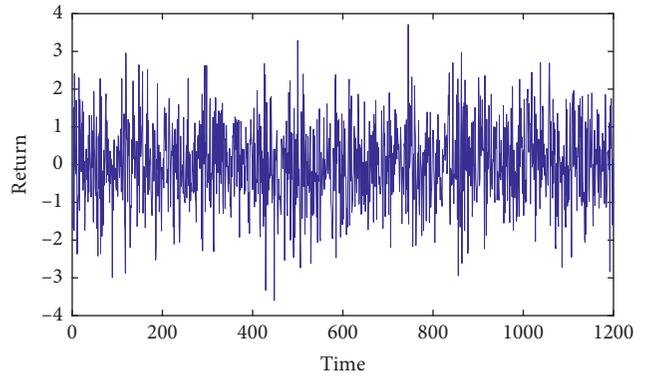


FIGURE 4: The simulated return series.

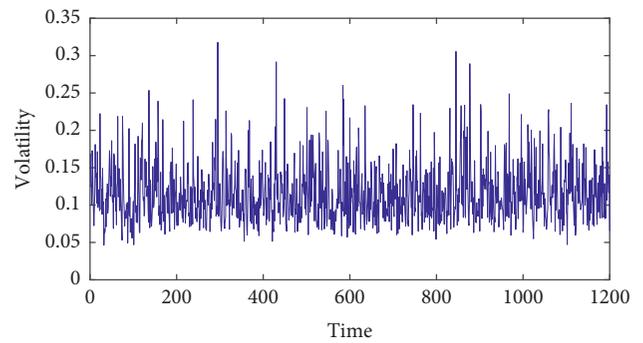


FIGURE 5: The simulated true volatility series.

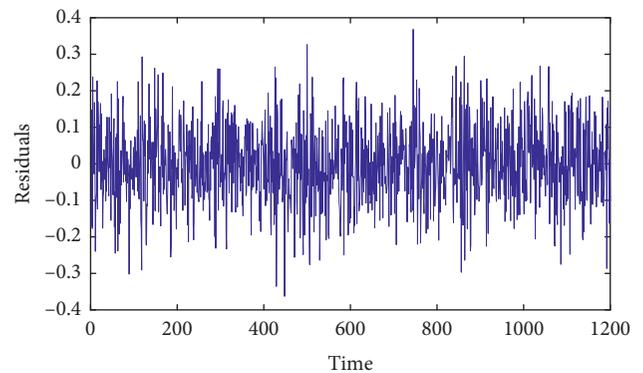


FIGURE 6: The simulated residual series.

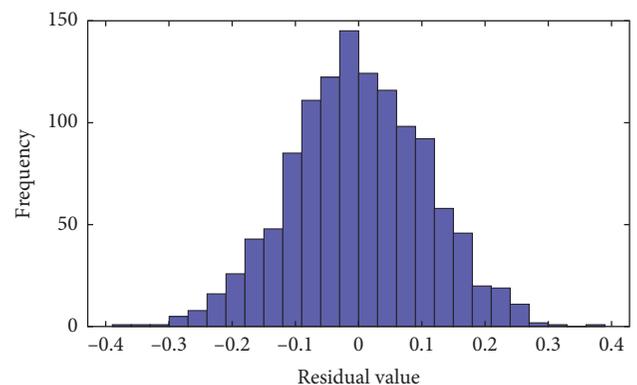


FIGURE 7: The histogram of simulated residual series.

TABLE 4: The EGARCH-ANN (1, 2, 3, 1) model fitted to the given data.

Parameters	Coefficients	Std. errors	<i>t</i> -stats
K	-3.7087	0.5434	-6.8248
ARCH1	-0.1665	0.0696	-2.3914
GARCH1	0.2668	0.0860	3.1021
GARCH2	0.6137	0.2389	2.5683
Leverage1	-0.4321	0.2121	-2.0373
LeveragebyANN1	-0.3534	0.1003	-3.5246

With addition information like log likelihood: 831, Akaike information criterion: -1649, Bayesian information criterion: -1620, and Hannan-Quinn information criterion: -1638.

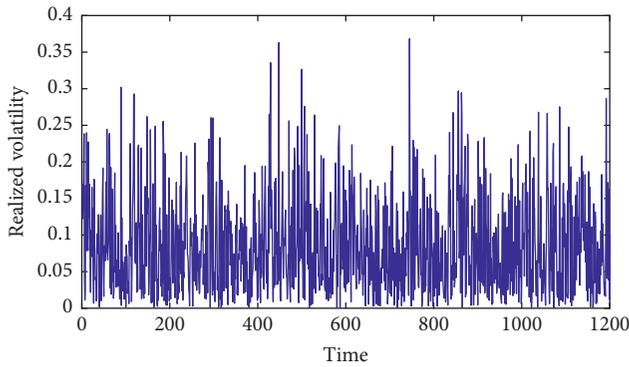


FIGURE 8: The realized volatility from the simulated residual series.

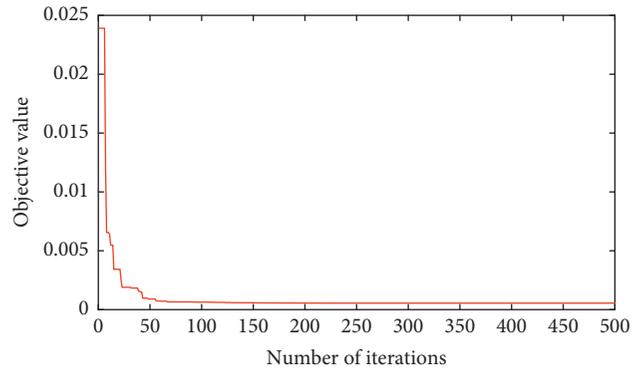


FIGURE 9: The MSFE obtained by JADE with archive.

TABLE 5: The estimated parameters of the fuzzy-EGARCH-ANN model on the first 1000 simulated data points.

Parameters	Coefficients
E_1	2.384303
α_{11}	-0.489804
γ_{11}	0.725490
β_{11}	1.2661273
β_{21}	0.456758
ξ_{11}	-0.164144
E_2	-2.026308
α_{12}	0.697262
γ_{12}	0.1068855
β_{12}	0.272422
β_{22}	-0.056937
ξ_{12}	-0.356486
E_3	-0.003542
α_{13}	1.106064
γ_{13}	2.169738
β_{13}	0.293315
β_{23}	-0.024236
ξ_{13}	-0.20663

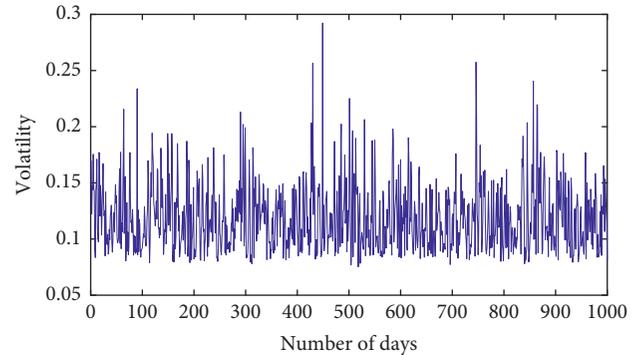


FIGURE 10: The estimated volatility for $T=1000$.

Finally, Figures 10 and 11 present the summary of estimated volatility by the proposed model for $T=1000$ and validation for $T=1200$, respectively. Here, the validation data set is the last 200 data points.

Note. The mean-squared forecast error (MSFE) for this validation is $4.7030e-04$, which is clearly less than that of the mean-squared forecast error (MSFE) obtained

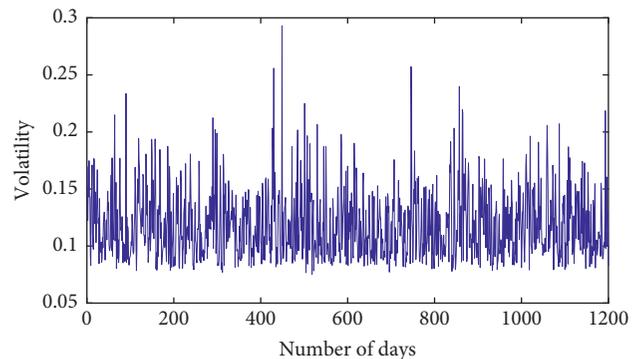


FIGURE 11: The validation of volatility for $T=1200$.

TABLE 6: 5 days ahead volatility prediction analysis of the fuzzy-EGARCH-ANN model.

The last 10 true value	$T=1190$	$T=1193$	$T=1195$	$T=1197$	$T=1199$
0.08284	0.09130				
0.13254	0.03657				
0.23432	0.01444				
0.10648	0.005619	0.10727			
0.14338	0.00215	0.00560			
0.15761		0.00002	0.14810		
0.13004		0.00002	0.01311		
0.11667		0.00002	0.00080	0.09362	
0.14062			0.00002	0.01987	
0.06576			0.00002	0.00512	0.04048
—				0.00156	0.00184
—				0.00055	0.00002
—					0.00002
—					0.00002

during the parameter estimation time. Therefore, the model has fitted the data very well. As a result, the study has obtained the forecast values for different periods as given in Table 6.

As one can see from the above forecasted values of volatility, the newly proposed model fuzzy-EGARCH-ANN has given us almost nearly the same values to the data simulated by itself. In the next section, the same thing with the same procedure was done except that Student's t -distribution is the distribution of the return innovation series. Moreover, in this section and the next section, the realized conditional variance series is the square root of the square of the residual series.

6.2. *Return Innovation with Student's t-Distribution.* For $p = 1, q = 1, s = 2, m = 4$, sample size = 1200, the Student's t -distribution type, degree of freedom $\nu = 10$, and the parameters given in Table 7, the study has simulated data by the fuzzy-EGARCH-ANN model. As a result, the obtained basic statistical properties of these data are given in Table 8.

Furthermore, Figures 12–16 are the corresponding figures of the simulated stock price series, return series, true volatility series, residual series, and histogram of the simulated residual series, respectively.

Now, by using the first 800 (that is, for $T = 800$) simulated data points, the study has obtained the following results. Since the study has been identified as the residual series which follows Student's t -distribution, then by using this distribution type, the EGARCH – ANN(2, 4, 0, 0) model is obtained as the best-fitted model to the selected data. Moreover, its fitness summary is presented in Table 9.

To proceed to the parameter estimation of the fuzzy-EGARCH-ANN model, only number of rules (R) and the realized volatility are remaining to be determined. For our

TABLE 7: The parameters used for the fuzzy-EGARCH-ANN model to simulate data.

Parameters	Coefficients
E_1	0.140625
α_{11}	-0.084375
γ_{11}	0.28125
β_{11}	0.140625
ξ_{11}	-0.196875
ξ_{21}	-0.140625
E_2	-0.28125
α_{12}	-0.140625
γ_{12}	0.140625
β_{12}	0.140625
ξ_{12}	-0.196875
ξ_{22}	-0.084375
E_3	0.28125
α_{13}	-0.253125
γ_{13}	0.309375
β_{13}	0.140625
ξ_{13}	-0.225
ξ_{23}	-0.225

TABLE 8: Statistical properties of residual data.

Statistical properties of residual data	
Min.	-0.453063
Max.	0.600486
Mean	0.000000
Median	-0.000187
Variance	0.014241
Skewness	0.169104
Kurtosis	4.270338
Arch test	p value = $8.0384e - 09$
Jarque–Bera test (for alpha = 0.05)	p value = $4.4409e - 16$

Mean < smedian usually flags left skewness. Negative skewness means left-skewed data. Kurtosis >3 means larger peakedness than Gaussian. Arch test result with very small p value implies existence of strong heteroskedasticity. Jarque–Bera test result with very small p value implies that the residual series strongly follows Student's t -distribution.

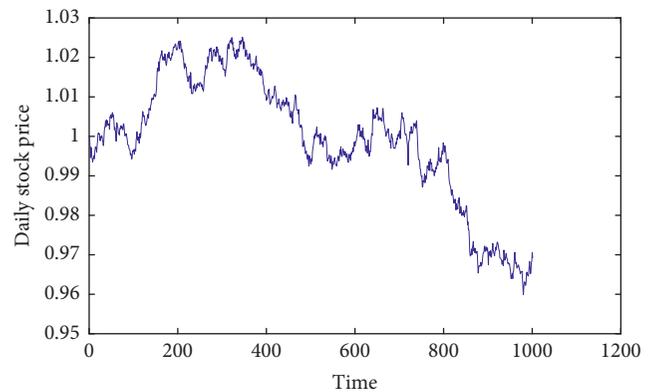


FIGURE 12: The simulated stock price.

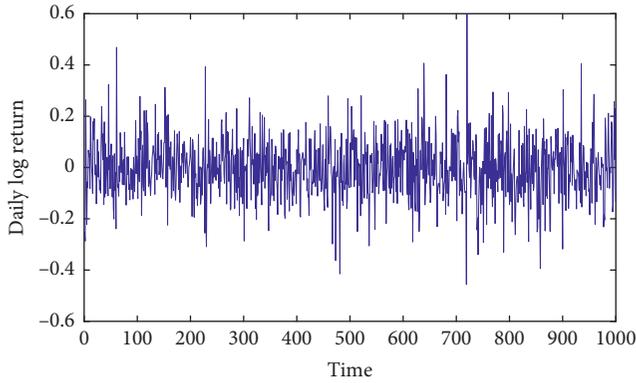


FIGURE 13: The simulated return series.

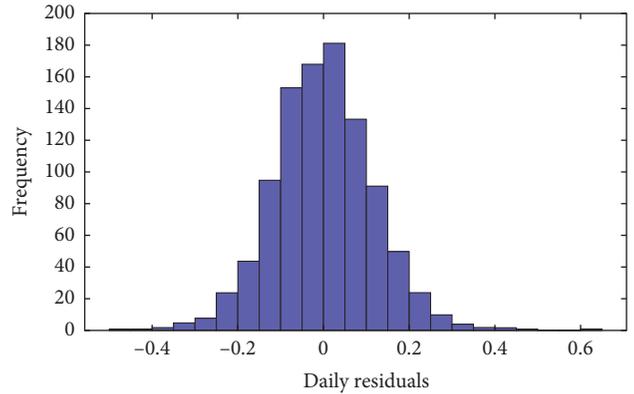


FIGURE 16: The histogram of simulated residual series.

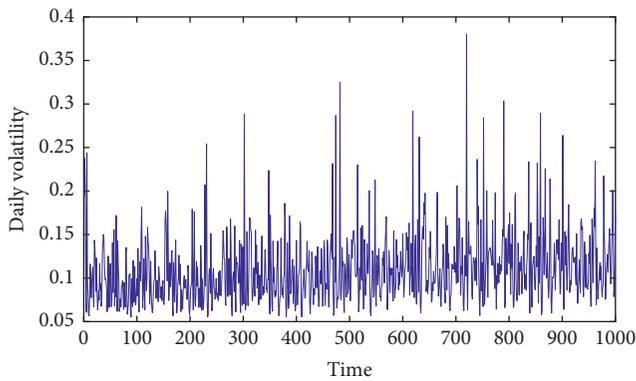


FIGURE 14: The simulated true volatility series.

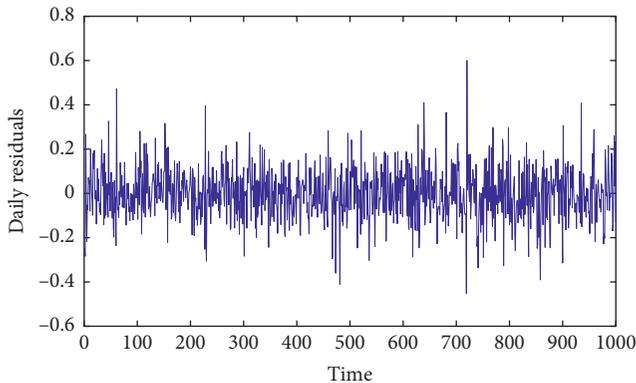


FIGURE 15: The simulated residual series.

TABLE 9: The EGARCH-ANN (2, 4, 0, 0) model fitted to the given data.

Parameters	Coefficients	Std. errors	<i>t</i> -stats
K	-2.5633	0.5817	-4.4067
ARCH1	0.1988	0.0744	2.6727
ARCH2	-0.2074	0.0777	-2.6704
GARCH1	1.1701	0.1324	8.8384
GARCH2	-1.1034	0.1666	-6.6238
GARCH3	0.8452	0.1675	5.0462
GARCH4	-0.4982	0.1543	-3.2291
Leverage1	-0.1534	0.0423	-3.6284
Leverage2	0.2297	0.0457	5.0208
Dof	15.002	7.1161	2.1082

With addition information like log likelihood: **608**, Akaike information criterion: -1196, Bayesian information criterion: -1149, and Hannan-Quinn information criterion: -1178.

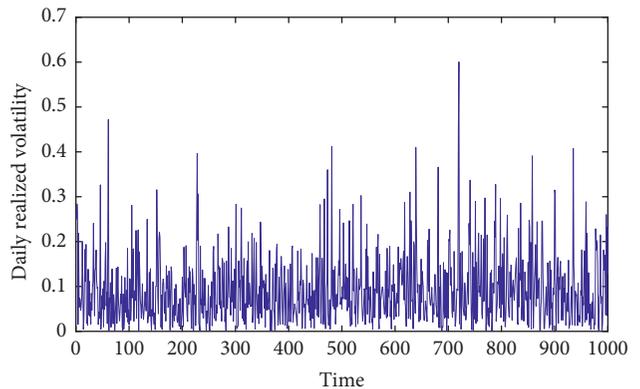


FIGURE 17: The realized volatility from the simulated residual series.

case, as mentioned earlier, R is going to be determined by the subtractive clustering algorithm (SCA), and the realized volatility at time t is the square root of the square of residual series at time t . $R = 3$ is obtained, and the graph of the realized volatility is given in Figure 17.

Therefore, for $p = 2$, $q = 4$, $m = 0$, and $s = 0$, by *JADE with archive*, the estimated parameter vector is given in Table 10. In addition, the corresponding mean-squared forecast errors (MSFE) value is $1.662787591217562e - 04$.

Figure 18 presents a summary and analysis of the MSFE obtained by *JADE with archive* during the parameter estimation of the proposed model.

Finally, Figures 19 and 20 present a summary of estimated volatility by the proposed model for $T = 800$ and validation for $T = 1000$, respectively. Here, the validation data set is the last 200 data points.

TABLE 10: The estimated parameters of the fuzzy-EGARCH-ANN model on the first 800 simulated data points.

Parameters	Coefficients
E_1	-0.135551 3333
α_{11}	-0.748310
α_{21}	0.400564
γ_{11}	0.902197
γ_{21}	-0.357340
β_{11}	0.503465
β_{21}	0.138148
β_{31}	0.117398
β_{41}	-0.042051
E_2	-0.743902
α_{12}	-0.486475
α_{22}	0.476170
γ_{12}	0.681418
γ_{22}	-0.476920
β_{12}	1.045201
β_{22}	-0.565138
β_{32}	0.220023
β_{42}	-0.071607
E_3	-0.625098
α_{13}	-0.289472
α_{23}	0.601564
γ_{13}	0.596101
γ_{23}	-0.473408
β_{13}	1.026555
β_{23}	-0.507309
β_{33}	0.218524
β_{43}	-0.044271

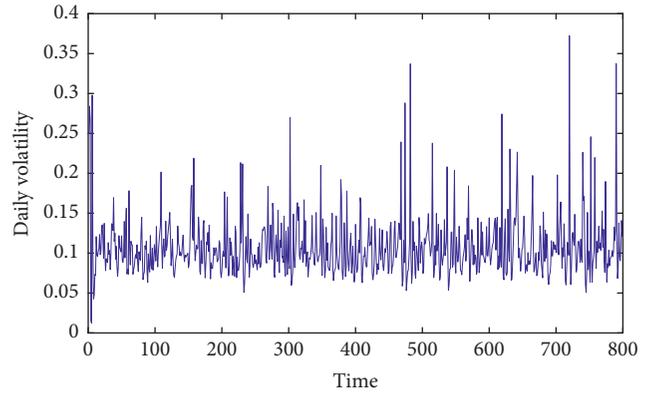


FIGURE 19: The estimated volatility for $T = 800$.

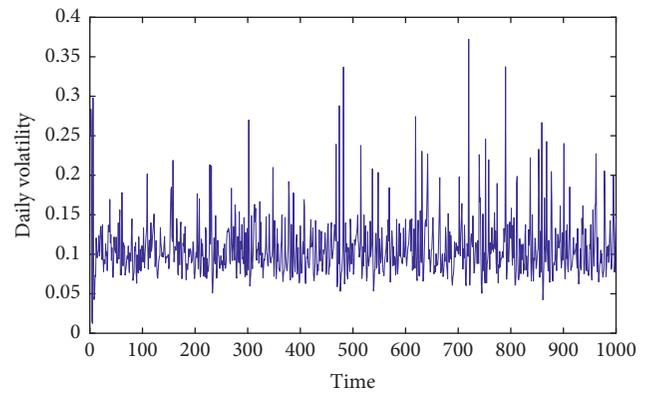


FIGURE 20: The validation of volatility for $T = 1000$.

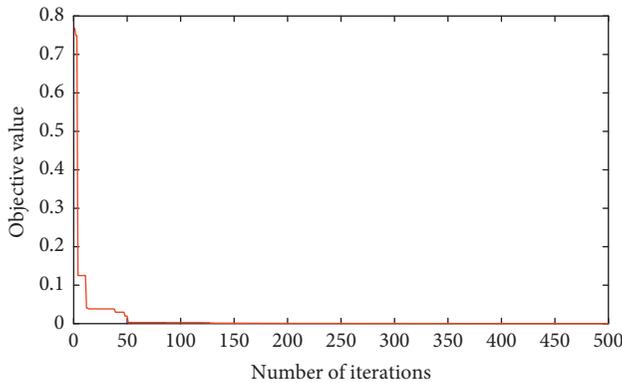


FIGURE 18: The MSFE obtained by JADE with archive.

Note. The mean-squared forecast error (MSFE) for this validation is $2.225849991650711e - 04$, which is very near to that of the mean-squared forecast error (MSFE) obtained during the parameter estimation time. Therefore, the model has fitted the data very well. As a result, the forecasted values have been obtained for different periods as given in Table 11.

Here also, as one can see from the above forecasted values of volatility, the newly proposed model fuzzy-EGARCH-ANN has given us almost nearly the same values to the data simulated by itself. In the next paper, we the authors of this research article are going to publish the

TABLE 11: 3 days ahead volatility prediction analysis of the fuzzy-EGARCH-ANN model.

The last 10 true value	$T = 990$	$T = 991$	$T = 995$	$T = 997$	$T = 999$
0.09341	0.09645				
0.09764	0.06920	0.12499			
0.13427	0.05264	0.05123			
0.158434		0.01873			
0.20099					
0.15670			0.19983		
0.07870			0.12573		
0.09213			0.08569	0.07734	
0.11422				0.05674	
0.17372				0.04385	0.12499
					0.05123
					0.018729

model calibration of the proposed model on some well-known financial time series data.

7. Conclusions

From the previous section, as one can see from the above simulation study results, the proposed semiparametric nonlinear fuzzy-EGARCH-ANN model has fitted the simulated data very well. Clearly, one can see from Table 6 that

the proposed model is good to forecast only one ahead forecast. This problem is the inherited problem from the EGARCH model. That is, the EGARCH model is good only for one ahead volatility forecasting. Therefore, the proposed model is very good in both capturing the leverage effect that was overlooked by the EGARCH model and volatility clustering that was overlooked by the EGARCH model. Hence, for both cases of return innovation distributions types, the proposed model is good only for one ahead period volatility forecasting.

Data Availability

The simulated financial time series data (Gaussian distribution-type simulated data and Student's t -distribution-type simulated data) used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

Methodology was done by Geleta Mohammed and Jane Aduda and Ananda Kube conceptualized and supervised the study. Finally, all the authors have read and approved the final manuscript.

Acknowledgments

Geleta T. Mohammed has paid the article processing charge from his monthly stipend.

References

- [1] R. Dash and P. K. Dash, "An evolutionary hybrid fuzzy computationally efficient EGARCH model for volatility prediction," *Applied Soft Computing*, vol. 45, pp. 40–60, 2016.
- [2] R. G. Donaldson and M. Kamstra, "An artificial neural network-GARCH model for international stock return volatility," *Journal of Empirical Finance*, vol. 4, no. 1, pp. 17–46, 1997.
- [3] A. R. Pagan and G. W. Schwert, "Alternative models for conditional stock volatility," *Journal of Econometrics*, vol. 45, no. 1-2, pp. 267–290, 1990.
- [4] A. R. Pagan and G. W. Schwert, "Measuring and testing the impact of news on volatility," *The Journal of Finance*, vol. 48, no. 5, pp. 1749–1778, 1993.
- [5] M. Bildirici and O. Ersin, "Modeling Markov switching ARMA-GARCH neural networks models and an application to forecasting stock returns," *The Scientific World Journal*, vol. 2014, Article ID 497941, 21 pages, 2014.
- [6] J. C. Hung, "Forecasting volatility using robust Kalman filter under Fuzzy-GARCH model," in *Proceedings of the International Conference on Business and Information*, pp. 166–177, Poznań, Poland, June 2013.
- [7] C.-H. Tseng, S.-T. Cheng, and Y.-H. Wang, "RETRACTED: new hybrid methodology for stock volatility prediction," *Expert Systems with Applications*, vol. 36, no. 2, pp. 1833–1839, 2009.
- [8] E. Hajizadeh, A. Seifi, M. H. Fazel Zarandi, and I. B. Turksen, "A hybrid modeling approach for forecasting the volatility of S&P 500 index return," *Expert Systems with Applications*, vol. 39, no. 1, pp. 431–436, 2012.
- [9] A. Tarsauliya, R. Kala, R. Tiwari, and A. Shukla, "Financial time series volatility forecast using evolutionary hybrid artificial neural network," *Advances in Network Security and Applications*, vol. 196, pp. 463–471, 2011.
- [10] J.-C. Hung, "Applying a combined fuzzy systems and GARCH model to adaptively forecast stock market volatility," *Applied Soft Computing*, vol. 11, no. 5, pp. 3938–3945, 2011.
- [11] L. Maciel, "A hybrid fuzzy GJR-GARCH modeling approach for stock market volatility forecasting," *Brazilian Review of Finance*, vol. 10, no. 3, pp. 337–367, 2012.
- [12] J.-R. Chang, L.-Y. Wei, and C.-H. Cheng, "A hybrid ANFIS model based on AR and volatility for TAIEX forecasting," *Applied Soft Computing*, vol. 11, no. 1, pp. 1388–1395, 2011.
- [13] L.-Y. Wei, "A GA-weighted ANFIS model based on multiple stock market volatility causality for TAIEX forecasting," *Applied Soft Computing*, vol. 13, no. 2, pp. 911–920, 2013.
- [14] L.-Y. Wei, C.-H. Cheng, and H.-H. Wu, "A hybrid ANFIS based on n-period moving average model to forecast TAIEX stock," *Applied Soft Computing*, vol. 19, pp. 86–92, 2014.
- [15] A. Kablan, "Adaptive neuro fuzzy inference for financial trading using intraday seasonality observation," *World Academy of Science, Engineering and Technology*, vol. 34, pp. 479–488, 2009.
- [16] R. F. Engle, "Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation," *Econometrica*, vol. 50, no. 4, pp. 987–1007, 1982.
- [17] T. Bollerslev, "Generalized autoregressive conditional heteroskedasticity," *Journal of Econometrics*, vol. 31, no. 3, pp. 307–327, 1986.
- [18] R. S. Tsay, *Analysis of Financial Time Series*, Vol. 543, John Wiley and Sons, Hoboken, NY, USA, 2005.
- [19] W. Kristjanpoller, A. Fadic, and M. C. Minutolo, "Volatility forecast using hybrid neural network models," *Expert Systems with Applications*, vol. 41, no. 5, pp. 2437–2442, 2014.
- [20] J. Arneric, T. Poklepovic, and Z. Aljinovic, "GARCH based artificial neural networks in forecasting conditional variance of stock returns," *Croatian Operational Research Review*, vol. 5, no. 2, pp. 329–343, 2014.
- [21] X. Lu, D. Que, and G. Cao, "Volatility forecast based on the hybrid artificial neural network and GARCH-type models," *Procedia Computer Science*, vol. 91, pp. 1044–1049, 2016.
- [22] H. White, "Economic prediction using neural networks: the case of ibm daily stock returns," in *Proceedings of the IEEE International Conference on Neural Networks*, San Diego, CA, USA, July 1988.
- [23] H. White, "Some asymptotic results for learning in single hidden-layer feedforward network models," *Journal of the American Statistical Association*, vol. 84, no. 408, pp. 1003–1013, 1989.
- [24] C.-M. Kuan and H. White, "Artificial neural networks: an econometric perspective," *Econometric Reviews*, vol. 13, no. 1, pp. 91–1, 1994.
- [25] S. S. Haykin and S. S. Haykin, *Kalman Filtering and Neural Networks*, Wiley Online Library, Hoboken, NY, USA, 2001.
- [26] K. Hornik, "Approximation capabilities of multilayer feed-forward networks," *Neural Networks*, vol. 4, no. 2, pp. 251–257, 1991.
- [27] K. Hornik, "Some new results on neural network approximation," *Neural Networks*, vol. 6, no. 8, pp. 1069–1072, 1993.

- [28] A. A. Popov and K. V. Bykhanov, "Modeling volatility of time series using Fuzzy GARCH models," in *Proceedings of the 9th Russian-Korean International Symposium on Science and Technology, 2005, KORUS 2005*, pp. 687–692, Novosibirsk, Russia, July 2005.
- [29] T. Takagi and M. Sugeno, "Fuzzy identification of systems and its applications to modeling and control," in *Readings in Fuzzy Sets for Intelligent Systems*, pp. 387–403, Elsevier, Amsterdam, Netherlands, 1993.
- [30] H. Ying, *Fuzzy Control and Modeling: Analytical Foundations and Applications*, Wiley-IEEE Press, Hoboken, NJ, USA, 2000.
- [31] L. A. Zadeh, "Fuzzy sets," *Information and Control*, vol. 8, no. 3, pp. 338–353, 1965.
- [32] L. A. Zadeh, "Soft computing and fuzzy logic," in *Fuzzy Sets, Fuzzy Logic, and Fuzzy Systems: Selected Papers by Lotfi a Zadeh*, pp. 796–804, World Scientific, Singapore, 1996.
- [33] L.-X. Wang and J. M. Mendel, "Fuzzy basis functions, universal approximation, and orthogonal least-squares learning," *IEEE Transactions on Neural Networks*, vol. 3, no. 5, pp. 807–814, 1992.
- [34] S. L. Chiu, "Fuzzy model identification based on cluster estimation," *Journal of Intelligent and Fuzzy Systems*, vol. 2, no. 3, pp. 267–278, 1994.
- [35] L. Maciel, F. Gomide, and R. Ballini, "Evolving Fuzzy-GARCH approach for financial volatility modeling and forecasting," *Computational Economics*, vol. 48, no. 3, pp. 379–398, 2016.
- [36] R. Storn and K. Price, "Differential evolution—a simple and efficient heuristic for global optimization over continuous spaces," *Journal of Global Optimization*, vol. 11, no. 4, pp. 341–359, 1997.
- [37] J. Zhang and A. C. Sanderson, "JADE: self-adaptive differential evolution with fast and reliable convergence performance," in *Proceedings of the 2007 IEEE congress on evolutionary computation*, pp. 2251–2258, Singapore, September 2007.
- [38] J. Zhang and A. C. Sanderson, "JADE: adaptive differential evolution with optional external archive," *IEEE Transactions on Evolutionary Computation*, vol. 13, no. 5, pp. 945–958, 2009.