A Differential Harmony Search based Hybrid Interval Type2 Fuzzy EGARCH Model for Stock Market Volatility Prediction

5.1 Introduction

This chapter presents an extension of the fuzzy computationally efficient EGARCH model discussed in chapter 4 by introducing interval type2 fuzzy sets in antecedent part. The new hybrid model integrates an interval type2 fuzzy logic system (IT2FLS) with a computationally efficient functional link artificial neural network (CEFLANN) and an Exponential Generalized Autoregressive Conditional Heteroskedasticity (EGARCH) model for accurate forecasting and modeling of financial data with changing variance over time. The proposed model denoted as IT2F-CE-EGARCH helps to enhance the ability of EGARCH model through a joint estimation of the important features of EGARCH like leverage effect, asymmetric shock by leverage effect with the secondary membership functions of interval type2 TSK FLS and the functional expansion and learning component of a CEFLANN. The secondary membership functions with upper and lower limit of IT2FLS provide a forecasting interval for handling more complicated uncertainties involved in volatility forecasting compared to type1 FLS. The model performance has been observed with two membership functions i.e. Gaussian with fixed mean, uncertain variance and Gaussian with fixed variance and uncertain mean. The proposed model has also been compared with few other fuzzy time series models and Fuzzy Computational Efficient EGARCH model discussed in chapter 4, based on four performance metrics: MSFE, RMSFE, MAFE and Rel MAE.

Again a differential harmony search (DHS) algorithm has been suggested for optimizing the parameters of all the fuzzy time series models. In chapter 2, the differential harmony search based parameter optimization has been discussed in details. So in this chapter only the new models used for volatility prediction has been highlighted. To test the model performance the two benchmark stock market indices i.e. BSE SENSEX and CNX NIFTY dataset discussed in chapter 4 are taken as experimental dataset.
5.2 Models used for Volatility Prediction

5.2.1 Interval Type2 Fuzzy Logic System

Fuzzy inference systems are universal approximators that can estimate nonlinear continuous functions uniformly with arbitrary sets of inputs. Fuzzy modeling methods are promising techniques for describing complex dynamics and asymmetries in systems. Commonly, two types of fuzzy logic system (FLS) are seen in the literature: type1 FLS and type2 FLS. As type1 fuzzy sets express the belongingness of a crisp value of a base variable in a fuzzy set by a crisp membership value, they cannot capture the uncertainties due to imprecision in identifying membership functions. To overcome this limitation, type2 FLS has been introduced to minimize the effects of the uncertainty in the rule base. Type2 Fuzzy logic systems are extensions of type1 FLS in which type2 fuzzy sets are used in antecedent or consequent parts. Type2 fuzzy sets are characterized by a 3 dimensional fuzzy membership function that provides additional degree of freedom. Therefore a type2 FLS has potential to outperform a type1 FLS. A type2 FLS includes a fuzzifier, rule bale, fuzzy inference engine and an o/p processing unit composed of a type reducer and a defuzzier. But general type2 FLSs is computationally intensive because of its intensive type-reduction. Hence instead of using T2-FLS with complex computation, IT2-FLS is preferred with much simplified computation in numerous fields. An interval type2 FLS is the practical extension of a general T2FLS with a uniform secondary membership function. In interval type-2 FLS two membership functions i.e. a lower and an upper membership function are used. Each of two membership function can be represented by a type-1 fuzzy set membership function. The interval between these two membership values represents the foot print of uncertainty (FOU) and this FOU is used to describe an interval type-2 fuzzy sets. FOU is the union of all primary membership functions and consists of a bounded region.

Most of the steps for type-1 and interval type-2 fuzzy sets are similar except the type reducer one which marks the difference between these two fuzzy logic systems. This step is necessary because membership grades of interval type-2 fuzzy sets are fuzzy set in (0, 1) and thus it has to be reduced to type-1 fuzzy set before defuzzifier is able to reduce the output fuzzy set further into crisp value. The detailed architecture and working principle of IT2FLS is specified in [117-124].
5.2.2 Interval Type2 Fuzzy Computationally Efficient EGARCH Model

Neuro Fuzzy Networks (NFNs) provide a better approach of using the low level learning, computational power of neural networks and the high-level human-like thinking and reasoning of fuzzy systems. Introducing a nonlinear function, especially a neural structure, to the consequent part of the fuzzy rules has yielded the neural networks designed on approximate reasoning architecture (NARA) and the coactive neuro fuzzy inference system (CANFIS) models. In the literature different NNs like MLP, WNN [46, 125] and FLANN [25] have been used in the consequent part of the CANFIS models. The proposed Interval Type2 Fuzzy computationally efficient EGARCH (IT2F-CE-EGARCH) model given in Fig. 5.1 represents a joint estimation of the membership function parameters of IT2FLS, the functional expansion and learning component of a CEFLANN along with the leverage and asymmetric shock due to leverage used in EGARCH model.

The proposed model uses a CEFLANN to the consequent part of the fuzzy rules and the IT2FLS in antecedent part. Computationally efficient FLANN (CEFLANN) is single-layer neural structure capable of forming arbitrarily complex decision regions by generating nonlinear decision boundaries with nonlinear functional expansion. The functional expansion of CEFLANN effectively increases the dimensionality of the input vector providing faster convergence rate and less computational loading. Thus, the consequent part provides a nonlinear combination of input variables. Again the parameters of EGARCH model i.e. the leverage effect by the market shock, the asymmetric shock by leverage effect and the log value of the volatility has used in functional Expansion of the CEFLANN. The model realizes a collection of fuzzy if-then rules in the following form. The $k^{th}$ rule is described by:

$$ R^k : \text{if} \quad x'_1 \text{ is } \tilde{F}_{k_1} \quad \text{and} \ldots \quad x'_d \text{ is } \tilde{F}_{k_d}, \quad \text{then} \quad y = CE(x) \quad (5.1) $$

where $x' = [r_{t-1}, r_{t-2}, \ldots, r_{t-q}, \sigma_{t-1}^2, \sigma_{t-2}^2, \ldots, \sigma_{t-p}^2]^T = [x_1, x_2, \ldots, x_l, x_d]^T$ is the input vector for $l=1, 2, \ldots, d$, with $d=(p + q)$ at instance $t$. $\sigma_t^2$ is the model output. $\tilde{F}_{kl}$ is the interval type2 fuzzy set used to describe the stock market return and volatility, $R$ is the number of rules and $x_l$ is the premise variable. Here, the premise variables include both the previous value of stock market return and the previous volatility of the stock market.
where as the consequent part expresses a local input output relation considering a CEFLANN with the EGARCH parameters as input. The number of fuzzy rules matches with the number of expanded units of CEFLANN.

The model output is obtained using the following steps:

Step1: Using the Gaussian membership function, find the grade of membership of each input \( x_i \) in \( \tilde{F}_{kl} \). As \( \tilde{F}_{kl} \) is the interval type2 fuzzy set so for each input two membership grades are obtained using two membership functions i.e. lower and upper. Gaussian membership function can be used in two ways:

Case 1: With Gaussian membership function having fixed mean and uncertain variance

\[
\tilde{F}_{kl}(x_i) = \exp \left( -\frac{1}{2} \left( \frac{x_i - c_{kl}}{a_{kl}} \right)^2 \right) \quad \text{where} \quad a_{kl} \in [a_{1kl}, a_{2kl}] \quad (5.2)
\]

The lower membership grade of input \( x_i \) using lower membership function is
\[ F_{kl}(x'_l) = \exp\left(-\frac{1}{2} \left( \frac{x'_l - c_{kl}}{a_{1kl}} \right)^2 \right) \] (5.3)

The upper membership grade of input \( x'_l \) using upper membership function is

\[ \overline{F}_{kl}(x'_l) = \exp\left(-\frac{1}{2} \left( \frac{x'_l - c_{kl}}{a_{2kl}} \right)^2 \right) \] (5.4)

where \( c_{kl} \) is the center and \( a_{1kl}, a_{2kl} \) are the two spreads of \( k^{th} \) rule membership functions corresponding to the \( l^{th} \) premise variable.

Case 2: With Gaussian membership function with fixed variance and uncertain mean

\[ \bar{F}_{kl}(x'_l) = \exp\left(-\frac{1}{2} \left( \frac{x'_l - c_{kl}}{a_{kl}} \right)^2 \right) \text{ where } c_{kl} \in [c_{1kl}, c_{2kl}] \] (5.5)

The lower membership grade of input \( x'_l \) using lower membership function is

\[
\underline{F}_{kl}(x'_l) = \begin{cases} 
\exp\left(-\frac{1}{2} \left( \frac{x'_l - c_{2kl}}{a_{kl}} \right)^2 \right), & x'_l \leq \frac{c_{1kl} + c_{2kl}}{2} \\
\exp\left(-\frac{1}{2} \left( \frac{x'_l - c_{1kl}}{a_{kl}} \right)^2 \right), & x'_l > \frac{c_{1kl} + c_{2kl}}{2}
\end{cases}
\] (5.6)

The upper membership grade of input \( x'_l \) using upper membership function is

\[
\overline{F}_{kl}(x'_l) = \begin{cases} 
\exp\left(-\frac{1}{2} \left( \frac{x'_l - c_{1kl}}{a_{kl}} \right)^2 \right), & x'_l < c_{1kl} \\
1, & c_{1kl} \leq x'_l \leq c_{2kl} \\
\exp\left(-\frac{1}{2} \left( \frac{x'_l - c_{2kl}}{a_{kl}} \right)^2 \right), & x'_l > c_{2kl}
\end{cases}
\] (5.7)

where \( c_{1kl}, c_{2kl} \) are the centers and \( a_{kl} \) is the spread of \( k^{th} \) rule membership functions corresponding to the \( l^{th} \) premise variable.
Step 2: Find the firing interval of $k^{th}$ rule, assuming the product T-norm of the antecedent fuzzy sets i.e.

$$[w_{Lk}, w_{Rk}] = \prod_{l=1}^{d} F_{kl}(x'_l)$$  \hspace{1cm} (5.8)$$

$$w_{Lk} = u_k(x') = \prod_{l=1}^{d} F_{kl}(x'_l)$$  \hspace{1cm} (5.9)$$

$$w_{Rk} = \overline{u_k}(x') = \prod_{l=1}^{d} F_{kl}(x'_l)$$  \hspace{1cm} (5.10)$$

Step 3: Find the normalized firing strengths i.e. the ratio of the $i^{th}$ rule’s firing strength to the sum of all rules’ firing strengths.

$$\overline{W}_{Lk} = \frac{W_{Lk}}{\sum_{k=1}^{R} W_{Lk}} = \frac{u_k(x')}{\sum_{k=1}^{R} u_k(x')}$$  \hspace{1cm} (5.11)$$

$$\overline{W}_{Rk} = \frac{W_{Rk}}{\sum_{k=1}^{R} W_{Rk}} = \frac{\overline{u_k}(x')}{\sum_{k=1}^{R} \overline{u_k}(x')}$$  \hspace{1cm} (5.12)$$

Step 4: Using the functional expansion block of the CEFLANN with order $n$, expand the input pattern as $CE(x) = [cx_1, cx_2 \ldots cx_d, cx_{d+1}, cx_{d+2}, \ldots cx_m]^T$ with $d = p + 2q$ and $m = d + n$ which is equal to the number of rules $R$.

$lv_i = \ln \sigma_{r-i}^2$ represents the logarithmic of volatility.

$L_i = \frac{r_{r-i}}{\sigma_{r-i}}$ represents the Leverage.

$le_i = \left(\frac{r_{r-i}}{\sigma_{r-i}} - \sqrt{\frac{2}{\Pi}}\right)$ represents the asymmetric shock due to leverage effect.

Each $cx_l$ (with $d+1 < l < m$) represents the o/p of the nonlinear hyperbolic tangent ($\tanh()$) function that takes weighted sum of logarithmic volatility, leverage and leverage effect i.e. the components of the input pattern $[cx_1, cx_2 \ldots cx_d]$ as its input.
for \( d + 1 \leq l \leq m \)

\[
    cx_i = \tanh \left( w_i + \sum_{j=1}^{p} \beta_j \ln \sigma_{r-j}^2 + \sum_{i=1}^{q} \gamma_i \left( \frac{r_{r-i}}{\sigma_{r-i}} - \sqrt{\frac{2}{\Pi}} \right) + \sum_{i=1}^{q} \alpha_i \frac{r_{r-i}}{\sigma_{r-i}} \right) \quad (5.13)
\]

Step 5: Perform the type reduction to combine the normalized firing strengths and the corresponding rule consequents to produce two weighted average of each individual rule as follows:

\[
y_L = \sum_{k=1}^{R} W_{Lk} \times cx_k = \frac{\sum_{k=1}^{R} u_k(x') \times cx_k}{\sum_{k=1}^{R} u_k(x')} \quad (5.14)
\]

\[
y_R = \sum_{k=1}^{R} W_{Rk} \times cx_k = \frac{\sum_{k=1}^{R} \bar{u}_k(x') \times cx_k}{\sum_{k=1}^{R} \bar{u}_k(x')} \quad (5.15)
\]

Step 6: Find the crisp output using defuzzification as follows:

\[
y = \lambda y_L + (1 - \lambda) y_R \text{ where } \lambda \text{ lies between 0 and 1} \quad (5.16)
\]

With \( \lambda = 0.5 \), \( y = \frac{y_L + y_R}{2} \)

Hence the model output is

\[
\ln \sigma_i^2 = \frac{\left( \sum_{k=1}^{R} u_k(x') \times cx_k + \sum_{k=1}^{R} \bar{u}_k(x') \times cx_k \right)}{\left( \sum_{k=1}^{R} u_k(x') + \sum_{k=1}^{R} \bar{u}_k(x') \right)} \quad (5.17)
\]

So the predicted volatility using the model is obtained by taking the exponential of defuzzified value.
5.2.3 Interval Type2 Fuzzy EGARCH Model

The Interval Type2 fuzzy EGARCH (p,q) model of Fig. 5.2 represents an IT2 fuzzy inference system based on EGARCH model. The fuzzy approach provides the capability to simulate stock fluctuations with volatility clustering where as the leverage effects of price change on conditional variance is captured by the EGARCH model. The model is described by a collection of type2 fuzzy rules in the form of if-then statements in order to describe the asymmetric responses of volatility to positive and negative shocks via an EGARCH model. The \( k^{th} \) rule of the IT2F-EGARCH (p, q) model is described by:

\[
R^k : \text{if} \quad x_1' \text{ is } \tilde{F}_{k1} \text{ and} \ldots x_d' \text{ is } \tilde{F}_{kd}, \text{ then} \\
r_i = \sigma_i \varepsilon_i \\
\ln \sigma_i^2 = w + \sum_{j=1}^{d} \beta_j \ln \sigma_{r-i}^2 + \sum_{i=1}^{p} a_i \left( \frac{r_{i-j}}{\sigma_{r-i}} - \sqrt{p} \right) + \sum_{i=1}^{q} \alpha_i \frac{r_{i-j}}{\sigma_{r-i}} \quad (5.18)
\]

where \( x' = [r_{1-1}, r_{1-2}, \ldots r_{1-q}, \sigma_{r-1}^2, \sigma_{r-2}^2, \ldots \sigma_{r-p}^2] \)^\(T\) is the input vector for \( l=1, 2, \ldots d \), with \( d=(p + q) \) at instance t. \( \sigma_i^2 \) is the model output. \( \tilde{F}_{kl} \) is the interval type2 fuzzy set used to describe the stock market return and volatility, \( R \) is the number of rules and \( x_{i} \) is the premise variable. Here, the premise variables include both the previous value of stock market return and the previous volatility of the stock market where as the consequent part expresses a local input output relation considering the EGARCH parameters. The model output is obtained using the following steps:

Step 1: Using the Gaussian membership function of equation (5.2) or (5.5), find the grade of membership of each input \( x_{i} \) in \( \tilde{F}_{kl} \).

Step 2: Find the firing interval \([w_{Lk}, w_{Rk}]\) of \( k^{th} \) rule, assuming the product T-norm of the antecedent fuzzy sets using equation (5.8) to (5.10).

Step 3: Find the normalized firing strengths \([w_{Lk}^{\prime}, w_{Rk}^{\prime}]\) using equation (5.11) and (5.12).

Step 4: Perform the type reduction to combine the normalized firing strengths and the corresponding rule consequents to produce weighted average of each individual rule as follows:
\[
y_L = \frac{\sum_{k=1}^{R} u_k(x^t)}{\sum_{k=1}^{R} u_k(x^t)} \left[ w_k + \sum_{j=1}^{p} \beta_{kj} \ln \sigma^2_{t-j} + \sum_{i=1}^{q} \gamma_{ki} \left( \frac{r_{t-i}}{\sigma_{t-i}} - \sqrt{\frac{2}{\Pi}} \right) + \sum_{i=1}^{q} \alpha_{ki} \frac{r_{t-i}}{\sigma_{t-i}} \right]
\]

\[
y_R = \frac{\sum_{k=1}^{R} u_k(x^t)}{\sum_{k=1}^{R} u_k(x^t)} \left[ w_k + \sum_{j=1}^{p} \beta_{kj} \ln \sigma^2_{t-j} + \sum_{i=1}^{q} \gamma_{ki} \left( \frac{r_{t-i}}{\sigma_{t-i}} - \sqrt{\frac{2}{\Pi}} \right) + \sum_{i=1}^{q} \alpha_{ki} \frac{r_{t-i}}{\sigma_{t-i}} \right]
\]

(5.19)

(5.20)

Step 5: Find the model output as a crisp value using defuzzification as follows:

\[
y = \lambda y_L + (1 - \lambda) y_R \text{ where } \lambda \text{ lies between 0 and 1} \quad (5.21)
\]

With \( \lambda = 0.5 \), \( y = \frac{y_L + y_R}{2} \)

Then the predicted volatility using the model is obtained by taking the exponential of the defuzzified value.

\[
\text{Defuzzify}
\]

\[
\sum_{i=1}^{R} \frac{r_{t-i}}{\sigma_{t-i}}
\]

Fig. 5.2 Architecture of IT2F-EGARCH Model with \( p=1, q=1 \) and \( R=3 \)
5.3 **Empirical Study**

In this section the model performance is evaluated over two different stock indices namely: BSE SENSEX and CNX NIFTY index of Indian stock market. The statistical analysis of both the data sets is clearly discussed in chapter 4. Hence in this section a comparative analysis of the results obtained from new hybrid fuzzy time series models and the fuzzy computationally efficient EGARCH model (proposed in chapter 4) are highlighted based on four error metrics and a significance test using Superior Predictive Ability Test. Finally the model validation is checked using the auto correlation and cross correlation analysis.

### 5.3.1 Data Set Description

In this study the sample data set of BSE SENSEX and CNX NIFTY index of Indian stock market comprising the daily closing prices have been taken to illustrate the performance of proposed Fuzzy time series models. The total number of samples for BSE is 333 from 2nd January 2012 to 30th April 2013 and for CNX are 255 from 2nd January 2012 to 4th January 2013. The daily stock return series has been generated by taking the natural logarithm difference of the daily stock index and the previous day’s stock index and multiplied by 100. The stock returns and daily volatility for BSE and CNX data sets are already shown in Figs. 4.4 and 4.5 in chapter 4. Both figures show that the return series are mean stationary and exhibit the typical volatility clustering phenomenon.

### 5.3.2 Performance Evaluation Criterion

In this study volatility forecasts comparison are conducted for one-step a head horizon in terms of mean squared forecast error (MSFE), root mean squared forecast error (RMSFE), mean absolute forecast error (MAFE) and relative mean absolute error (Rel MAE). The definitions of MSFE, RMSFE, and MAFE are already given in chapter 2. The Rel MAE is defined as follows:

\[
\text{Rel MAE} = \frac{1}{N} \sum_{i=1}^{N} \left| \frac{\sigma_i^2 - \hat{\sigma}_i^2}{\sigma_i^2} \right|
\]

(5.22)
where \( \sigma_i^2 \) is the realized volatility and \( \hat{\sigma}_i^2 \) is the predicted volatility. \( N \) is the number of data points. The realized volatility on day \( t \) is calculated using equation (2.34). Smaller error values are considered as the measure of better forecasting performance, but if results are not consistent among three criterions, then the MSFE is taken as the benchmark one.

### 5.3.3 Parameter Setup

Initially comparing the performance of GARCH (1, 1), GJR-GARCH (1, 1) and EGARCH (1, 1) models, EGARCH has shown the best performance and thus is selected for construction of hybrid models. Few other lag values for these GARCH family models have also been tested. As no significant better output is obtained with other lag values, so here the same lag order has been considered for the fuzzy time series models. Then the proposed IT2F-CE-EGARCH model with lag order \( p=1, q=1 \) has been compared with type1 Fuzzy EGARCH (1, 1), type1 fuzzy CE-EGARCH (1, 1) and IT2F-EGARCH (1, 1) model based on four performance metrics: MSFE, RMSFE, MAFE and Rel MAE. The size of parameter space of different fuzzy time series models has been specified in Table 5.1.

<table>
<thead>
<tr>
<th>MODEL</th>
<th>Size of Parameter Space</th>
</tr>
</thead>
<tbody>
<tr>
<td>IT2F-CE-EGARCH(1,1) (using Gaussian function with uncertain variance)</td>
<td>38</td>
</tr>
<tr>
<td>IT2F-CE-EGARCH(1,1) (using Gaussian function with uncertain mean)</td>
<td>38</td>
</tr>
<tr>
<td>IT2F-EGARCH(1,1) (using Gaussian function with uncertain variance)</td>
<td>30</td>
</tr>
<tr>
<td>IT2F-EGARCH(1,1) (using Gaussian function with uncertain mean)</td>
<td>30</td>
</tr>
<tr>
<td>Type1 Fuzzy Computationally Efficient EGARCH(1,1)</td>
<td>28</td>
</tr>
</tbody>
</table>

The following controlling parameters are set for optimization of the fuzzy systems through DHS algorithm. With the harmony memory size 20, the value of each harmony i.e. the unknown parameters of the model comprising the antecedent and
consequent parameter values are initialized randomly. Maximum number of iterations is set to 100. Through a number of simulations the upper bound of HMCR, PAR and BW value are set to 0.95, 0.5, and 0.55, respectively whereas the lower bound of HMCR, PAR and BW value are set to 0.8, 0.2, and 0.15, respectively. RMSE is used as the objective function for the DHS algorithm. The performance of both IT2F-CE-EGARCH and IT2F-EGARCH model are observed using both type of Gaussian function specified in equations (5.2) and (5.5).

5.3.4 Result Analysis

To measure the generalization ability of the model initially data sets are divided into training and testing sets. The training part has been used to train the model, while the testing part has been used to compute the predictions. For BSE data set the training set consists of 218 patterns and the rest is set for testing and for CNX dataset the training set consists of 166 patterns leaving the rest for testing. Originally the dataset is divided into a single train and test set. Then the performance metrics have been observed as an average of 20 independent runs of a particular model over the same training and testing sets. Tables 5.2 and 5.3 list the various forecast statistics of the different volatility models obtained during testing.

The convergence curve of DHS algorithm for training of IT2F-CE-EGARCH (1, 1) model for both the datasets are shown in Figs. 5.3 to 5.6. Further the statistical analysis in term of best, mean, standard deviation and average CPU time of DHS algorithm with different harmony memory size and number of iterations for training the proposed model is listed in Table 5.4. The one day a head volatility forecasting of BSE and CNX data sets using different models are shown in Figs. 5.7 to 5.16.

Again to avoid the data snooping problem, the upper and lower bound of SPA test over MSFE error are calculated for both the datasets. Tables 5.5 and 5.6 present the p-value of upper and lower bound of SPA test with 3 different sizes of bootstrap re-samples. In this study individually each model has been considered as the benchmark model to evaluate whether a particular model is significantly outperformed by other competing modes. For both the datasets, the p-value of SPA test over MSFE error with different bootstrap size is showing higher value indicating that the other competing
models are not giving better performance compared to the proposed model. In both the data sets the performance of the proposed IT2F-CE-EGARCH model using Gaussian function with uncertain variance and fixed mean performs better than the other models, since its structure provides a combination of rules and functional learning component for estimating forecast errors and also a mechanism to deal with leverage effects.

Table 5.2 Performance comparison of different volatility models over BSE Data set

<table>
<thead>
<tr>
<th>MODEL</th>
<th>MSFE</th>
<th>RMSFE</th>
<th>MAE</th>
<th>Rel MAE</th>
</tr>
</thead>
<tbody>
<tr>
<td>IT2F-CE-EGARCH(1,1) (using Gaussian function with uncertain variance)</td>
<td>0.0333</td>
<td>0.1826</td>
<td>0.1251</td>
<td>0.5234</td>
</tr>
<tr>
<td>IT2F-CE-EGARCH(1,1) (using Gaussian function with uncertain mean)</td>
<td>0.0349</td>
<td>0.1867</td>
<td>0.1255</td>
<td>0.4413</td>
</tr>
<tr>
<td>IT2F-EGARCH(1,1) (using Gaussian function with uncertain variance)</td>
<td>0.0401</td>
<td>0.2004</td>
<td>0.1433</td>
<td>0.6343</td>
</tr>
<tr>
<td>IT2F-EGARCH(1,1) (using Gaussian function with uncertain mean)</td>
<td>0.0408</td>
<td>0.2019</td>
<td>0.1438</td>
<td>0.7393</td>
</tr>
<tr>
<td>Type1 Fuzzy Computationally Efficient EGARCH(1,1)</td>
<td>0.0367</td>
<td>0.1915</td>
<td>0.1223</td>
<td>0.4354</td>
</tr>
</tbody>
</table>

Table 5.3 Performance comparison of different volatility models over CNX Nifty Data set

<table>
<thead>
<tr>
<th>MODEL</th>
<th>MSFE</th>
<th>RMSFE</th>
<th>MAE</th>
<th>Rel MAE</th>
</tr>
</thead>
<tbody>
<tr>
<td>IT2F-CE-EGARCH(1,1) (using Gaussian function with uncertain variance)</td>
<td>0.0414</td>
<td>0.2034</td>
<td>0.1249</td>
<td>0.5598</td>
</tr>
<tr>
<td>IT2F-CE-EGARCH(1,1) (using Gaussian function with uncertain mean)</td>
<td>0.0428</td>
<td>0.2069</td>
<td>0.1240</td>
<td>0.5423</td>
</tr>
<tr>
<td>IT2F-EGARCH(1,1) (using Gaussian function with uncertain variance)</td>
<td>0.0434</td>
<td>0.2084</td>
<td>0.1459</td>
<td>0.7146</td>
</tr>
<tr>
<td>IT2F-EGARCH(1,1) (using Gaussian function with uncertain mean)</td>
<td>0.0454</td>
<td>0.2130</td>
<td>0.1402</td>
<td>0.6272</td>
</tr>
<tr>
<td>Type1 Fuzzy Computationally Efficient EGARCH(1,1)</td>
<td>0.0435</td>
<td>0.2086</td>
<td>0.1238</td>
<td>0.5053</td>
</tr>
</tbody>
</table>
Table 5.4 Performance comparison of proposed IT2F-CE-EGARCH model with different harmony memory size and iterations

<table>
<thead>
<tr>
<th>Model</th>
<th>Data set</th>
<th>HMS</th>
<th>NI</th>
<th>RMSE Train</th>
<th>RMSE Test</th>
<th>CPU time in sec</th>
</tr>
</thead>
<tbody>
<tr>
<td>IT2F-CE-EGARCH</td>
<td>BSE</td>
<td>20</td>
<td>100</td>
<td>0.1152</td>
<td>0.1165</td>
<td>0.0011 0.1826 0.2015 0.0158 41.18</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.1093</td>
<td>0.1126</td>
<td>0.0033 0.1954 0.2133 0.0169 89.55</td>
</tr>
<tr>
<td></td>
<td></td>
<td>40</td>
<td>100</td>
<td>0.1194</td>
<td>0.1224</td>
<td>0.0036 0.2091 0.2152 0.0073 86.65</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.1214</td>
<td>0.1257</td>
<td>0.0041 0.2107 0.2156 0.0055 175.5</td>
</tr>
<tr>
<td>(using Gaussian function</td>
<td>CNX</td>
<td>20</td>
<td>100</td>
<td>0.1488</td>
<td>0.1557</td>
<td>0.0064 0.2034 0.2263 0.0237 45.08</td>
</tr>
<tr>
<td>with uncertain variance)</td>
<td></td>
<td></td>
<td></td>
<td>0.1451</td>
<td>0.1521</td>
<td>0.0062 0.2210 0.2490 0.0326 71.49</td>
</tr>
<tr>
<td></td>
<td></td>
<td>40</td>
<td>100</td>
<td>0.1602</td>
<td>0.1636</td>
<td>0.0022 0.2154 0.2563 0.0327 75.69</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.1449</td>
<td>0.1481</td>
<td>0.0031 0.2257 0.2797 0.0476 167.8</td>
</tr>
<tr>
<td></td>
<td>BSE</td>
<td>20</td>
<td>100</td>
<td>0.1246</td>
<td>0.1266</td>
<td>0.0027 0.1867 0.1946 0.0075 45.21</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.1147</td>
<td>0.1167</td>
<td>0.0016 0.1934 0.2138 0.0141 97.81</td>
</tr>
<tr>
<td></td>
<td></td>
<td>40</td>
<td>100</td>
<td>0.1208</td>
<td>0.1233</td>
<td>0.0017 0.1985 0.2165 0.0135 98.68</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.1195</td>
<td>0.1209</td>
<td>0.0013 0.2094 0.2137 0.0039 156.8</td>
</tr>
<tr>
<td>(using Gaussian function</td>
<td>CNX</td>
<td>20</td>
<td>100</td>
<td>0.1688</td>
<td>0.1761</td>
<td>0.0058 0.2069 0.2168 0.0093 48.74</td>
</tr>
<tr>
<td>with uncertain mean)</td>
<td></td>
<td></td>
<td></td>
<td>0.1423</td>
<td>0.1496</td>
<td>0.0085 0.2130 0.2253 0.0101 100.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>40</td>
<td>100</td>
<td>0.1736</td>
<td>0.1750</td>
<td>0.0018 0.2160 0.2360 0.0300 93.14</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.1525</td>
<td>0.1634</td>
<td>0.0083 0.2086 0.2661 0.0432 142.4</td>
</tr>
</tbody>
</table>
Fig. 5.3 Convergence curve of DHS for IT2F-CE-EGARCH (1, 1) model using Gaussian MF with uncertain variance over BSE dataset

Fig. 5.4 Convergence curve of DHS for IT2F-CE-EGARCH (1, 1) model using Gaussian MF with uncertain variance over CNX dataset

Fig. 5.5 Convergence curve of DHS for IT2F-CE-EGARCH (1, 1) model using Gaussian MF with uncertain mean over BSE dataset

Fig. 5.6 Convergence curve of DHS for IT2F-CE-EGARCH (1, 1) model using Gaussian MF with uncertain mean over CNX dataset
Fig. 5.7 Volatility of BSE data using IT2F-CE-EGARCH (1, 1) using Gaussian MF with uncertain mean

Fig. 5.8 Volatility of CNX data using IT2F-CE-EGARCH (1, 1) using Gaussian MF with uncertain mean

Fig. 5.9 Volatility of BSE data using IT2F-CE-EGARCH (1, 1) using Gaussian MF with uncertain variance

Fig. 5.10 Volatility of CNX data using IT2F-CE-EGARCH (1, 1) using Gaussian MF with uncertain variance
Fig. 5.11 Volatility of BSE data using type1 fuzzy Computationally Efficient EGARCH (1, 1)

Fig. 5.13 Volatility of BSE data using IT2F-EGARCH (1, 1) using Gaussian MF with uncertain mean

Fig. 5.15 Volatility of BSE data using IT2F-EGARCH (1, 1) using Gaussian MF with uncertain variance

Fig. 5.12 Volatility of CNX data using type1 fuzzy Computationally Efficient EGARCH (1, 1)

Fig. 5.14 Volatility of CNX data using IT2F-EGARCH (1, 1) using Gaussian MF with uncertain mean

Fig. 5.16 Volatility of CNX data using IT2F-EGARCH (1, 1) using Gaussian MF with uncertain variance
Further to evaluate the model quality a residual analysis including the whiteness and independence test has also been performed. In whiteness test the autocorrelation analysis of residuals of the output volatility is observed to verify whether the residual autocorrelation function is inside the confidence interval of the corresponding estimates or not. Similarly in Independence test Cross-correlation between the input and the residuals for each input-output pair is checked to verify the correlation between residuals and past input. Again for model validation only the positive lags of cross correlation analysis is taken in to consideration. For a good model residuals are
uncorrelated with other residuals and with past inputs. An analysis of Auto correlation and Cross correlation of the proposed model over two data sets has been shown in Figs. 5.17 to 5.20.

Fig. 5.17 Auto correlation and cross correlation analysis of residuals of BSE data obtained from IT2F-CE-EGARCH (1, 1) using Gaussian MF with uncertain variance

Fig. 5.18 Auto correlation and cross correlation analysis of residuals of BSE data obtained from IT2F-CE-EGARCH (1, 1) using Gaussian MF with uncertain mean
Fig. 5.19 Auto correlation and cross correlation analysis of residuals of CNX data obtained from IT2F-CE-EGARCH (1, 1) using Gaussian MF with uncertain variance

Fig. 5.20 Auto correlation and cross correlation analysis of residuals of CNX data obtained from IT2F-CE-EGARCH (1, 1) using Gaussian MF with uncertain mean

In each figure the top axes show the autocorrelation of residuals for the output. The horizontal scale is the number of lags, which is the time difference between the
signals at which the correlation is estimated. The blue solid lines on the plot represent the confidence interval of the corresponding estimates. Any fluctuations within the confidence interval are considered to be insignificant. In all figures the residual autocorrelation values are almost within the confidence interval, which clearly indicates that the residuals are uncorrelated. The bottom axes show the cross-correlation of the residuals with the input. As the model has been designed with lag 1 and the cross correlation of residuals with input daily return and daily volatility are within the confidence interval at positive lag, so it clearly indicates the residuals are uncorrelated with past daily return and daily volatility.

5.4 Summary

In this chapter, an integrated model combining the features of time series model, reasoning of IT2FLS and learning ability of ANN is explained in details for stock market volatility modeling and forecasting. From experimental study it is clearly established that the IT2F-CE-EGARCH model helps to deal with the uncertainties involved in volatility prediction in an efficient manner. The use of a Differential Harmony Search (DHS) algorithm is also suggested 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 IT2F-CE-EGARCH model are provided by modeling and forecasting the volatility of BSE SENSEX and CNX NIFTY indexes in comparison with type1 fuzzy CE-EGARCH model and IT2F-EGARCH model optimized with DHS algorithm. The computational results demonstrate that the IT2F-CE-EGARCH model using Gaussian function with uncertain variance and fixed mean offers significant improvements in forecasting performance compared to all other volatility models.

Future scope of this research can be extended for multistep ahead volatility prediction and can also be applied to some other indices of international stock market. The model has taken the lagged return and volatilities as its input. The performance accuracy can also be enhanced by combining few technical indicators with the model. Again future work will include exploring new methods for choosing appropriate lag value of the input and in choosing number of rules in the fuzzy time series models.