produce multiple numeric outputs and the expected value of consequent is calculated. In this way, the proposed method forms a model of the power system which is then used for load forecasting.

One major advantage of this method is that it doesn't require secondary time series like temperature, humidity, holiday index, etc. for Electric Load prediction. Most of the other research papers in this context have used the secondary time series. So our method could be very helpful in the cases where weather data and other parameters for load forecasting are not available. The IT2FS model is not dependent on secondary indices as well. Due to the superior ability of IT2FS to model knowledge uncertainty, the proposed method is expected to perform better than the previous methods. Another advantage of our method is that it is applicable not only on STLF but on MTLF and LTLF as well.

The performance metric used is the *Root Mean Square Error* (RMSE) and *Mean Absolute Probability Error* (MAPE). The average RMSE and MAPE due to the proposed method is lower than that of many of the available methods of Electric Load prediction.

The following figure shows the block diagram for prediction of test data using the trained system

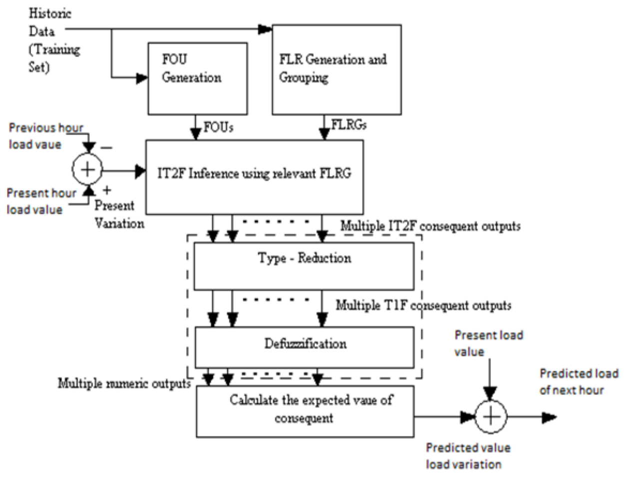

Fig 1. Load Forecasting using IT2FS block diagram

II. Elementary Concepts

The following definitions are essential to the understanding of the new Interval type-2 Fuzzy Set based prediction method proposed in this paper.

A. Definition 1

Let $X \subseteq \mathbb{R}$, be the universe of discourse. Let $f(t) \in X$ be a time dependent variable. Then the sequence $f(t_1)$, $f(t_2)$, . . . , $f(t_n)$ of samples of $f(t)$ at time instants t_1, t_2, . . . , t_n is known as a time series in X.

B. Definition 2

An *interval type-2 fuzzy set* (IT2FS)$\tilde{A}$, in the universe of discourse X, is characterised by a membership interval, known as the footprint of uncertainty (FOU($\tilde{A}$)), encompassing all the embedded primary membership functions $J_x$of $\tilde{A}$.

FOU ($\tilde{A}$) is bounded by an upper membership function (UMF) $\overline{\mu}_{\tilde{A}}(x)$ and a lower membership function (LMF) $\underline{\mu}_{\tilde{A}}(x)$. $\overline{\mu}_{\tilde{A}}(x)$ and $\underline{\mu}_{\tilde{A}}(x)$ at all x, respectively take up the minimum and maximum of the membership functions of the embedded T1FSs in the FOU.

$$FOU(\tilde{A}) = \bigcup_{x \in X} J_x \tag{1}$$

C. Definition 3

Let$X \subseteq \mathbb{R}$, the set of real numbers, be the universe of discourse. Let $\tilde{A}_i \subseteq X$ for $i = 1, 2, \dots, m$ be m IT2FS. Let $G_{t,k}: f(t) \to \tilde{A}_i$ be a mapping from the time varying variable $f(t) \in X$ to an IT2FS $\tilde{A}_i$.

Then for the time series $f(t_1)$, $f(t_2)$, . . . , $f(t_n)$, we obtain an *interval fuzzy time series* (IT2FTS) $\tilde{A}_{k_1}$, $\tilde{A}_{k_2}$, . . . , $\tilde{A}_{k_n}$, where $k_i \in [1, m] \subseteq \mathbb{I}$, the set of integers according to the mapping function $G_{t,k}$. Let the IT2FTS be denoted by $\tilde{F}(t)$.

Example: Consider the time series $[f(t_1), f(t_2), \dots, f(t_n)]$ = $[40, 37, 14, 10, 29, 19, 11, 25, 39, 38]$. Let $X = [10, 40]$ be the universe of discourse, partitioned into non-overlapping intervals$I_1 = [10, 20)$, $I_2 = [20, 30)$and $I_3 = [30, 40]$. Let $\tilde{A}_1$, $\tilde{A}_2$ and $\tilde{A}_3$ be 3 IT2FSs in X.

Then using the mapping rule: IF $f(t) \in I_i$, THEN $\tilde{F}(t) = \tilde{A}_i$, we get the IT2FTS $\tilde{F}(t)$ as follows.

$$[\tilde{F}(t_1), \tilde{F}(t_2), \dots, \tilde{F}(t_n)] = [\tilde{A}_3, \tilde{A}_3, \tilde{A}_1, \tilde{A}_1, \tilde{A}_2, \tilde{A}_1, \tilde{A}_1, \tilde{A}_2, \tilde{A}_3, \tilde{A}_3]. \tag{2}$$

Let $f_{3-mn}(t)$ and $f_{3-sd}(t)$ be two series obtained using the following formulae.

$$f_{3-mn}(t) = \frac{1}{3} \sum_{k=1}^{3} f(t - (k-1))$$

$$f_{3-sd}(t) = \sqrt{\frac{1}{3} \sum_{i=1}^{3} (f^2(t - i + 1)) - f_{3-mean}{}^2(t)} \tag{3}$$

Suppose, the FOUs of $\tilde{A}_1$, $\tilde{A}_2$ and $\tilde{A}_3$ are formed based on the rule:

If $f_{3-mn}(t) \in I_i$, THEN Gaussian($f_{3-mn}(t)$, $f_{3-sd}(t)) \in FOU(\tilde{A}_i)$ where Gaussian($f_{3-mn}(t)$, $f_{3-sd}(t)$) is the gaussian distribution with a mean of $f_{3-mn}(t)$ and a standard deviation of $f_{3-sd}(t)$.

The FOUs are obtained as follows.

$$FOU(\tilde{A}_i) = \bigcup_{\forall f_{3-mn}(t)\in I_i} \left\{ \frac{1}{\sqrt{2\pi} f_{3-sd}(t)} e^{\frac{1}{2}\left(\frac{-(x-f_{3-mn}(t))}{f_{3-sd}(t)}\right)^2} \right\} \quad (4)$$

UMF $\overline{\mu}_{\tilde{A}}(x)$ and LMF $\underline{\mu}_{\tilde{A}}(x)$ are given by the following formulae:

$$\overline{\mu}_{\tilde{A}_i}(x) = max_{\forall f_{3-mn}(t)\in I_i} \left\{ \frac{1}{\sqrt{2\pi} f_{3-sd}(t)} e^{\frac{1}{2}\left(\frac{-(x-f_{3-mn}(t))}{f_{3-sd}(t)}\right)^2} \right\}$$

$$\underline{\mu}_{\tilde{A}_i}(x) = min_{\forall f_{3-mn}(t)\in I_i} \left\{ \frac{1}{\sqrt{2\pi} f_{3-sd}(t)} e^{\frac{1}{2}\left(\frac{-(x-f_{3-mn}(t))}{f_{3-sd}(t)}\right)^2} \right\} \quad (5)$$

The case in point is illustrated in figures 2 and 3.

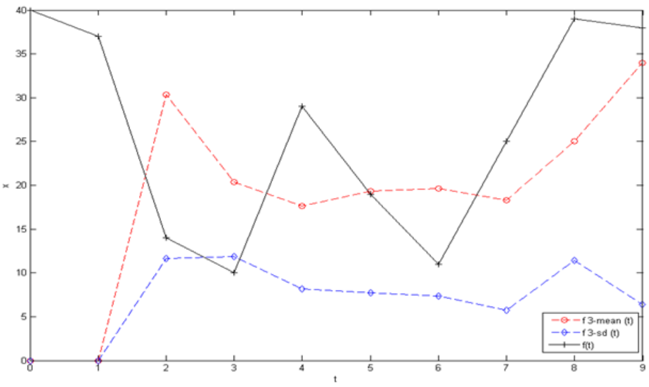

Fig. 2. Diagram showing the fuzzification of f(t) to obtain IT2FTS $\tilde{F}(t)$

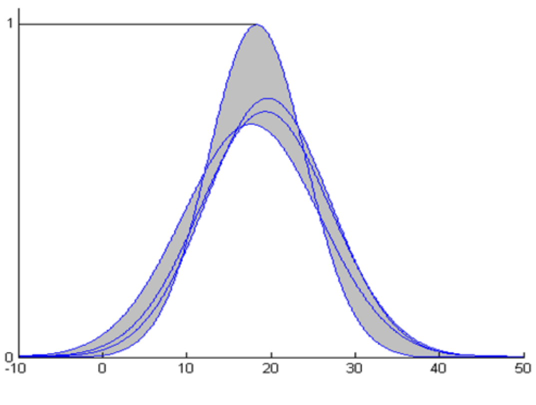

Fig. 3. FOU of $\tilde{A}_1$ obtained for $f_{3-mn}(2)$, $f_{3-mn}(3)$,$f_{3-mn}(4)$ and $f_{3-mn}(5) \in I_1$

It should be noted that, mean and standard deviation series for more than 3 hour scan also be used. In fact, 5-hour moving mean and standard deviation series is, used in this paper, for predicting electrical load.

D. Definition 4

Let $\tilde{A}_i$ and $\tilde{A}_j$ be two IT2FS in the universe of discourse X. Also, let $\tilde{A}_i \to \tilde{A}_j$ be an FLR in X. Then, the interval type 2 fuzzy implication to obtain the consequent IT2FS $\tilde{A}_j'$, for an antecedent data point $x_1 \in \tilde{A}_i$, is defined as follows.

$$\begin{aligned} \overline{\mu}_{\tilde{A}_j'}(x) &= min\{\overline{\mu}_{\tilde{A}_i}(x_1), \overline{\mu}_{\tilde{A}_j}(x)\} \\ \underline{\mu}_{\tilde{A}_j'}(x) &= min\{\underline{\mu}_{\tilde{A}_i}(x_1), \underline{\mu}_{\tilde{A}_j}(x)\} \end{aligned} \quad (6)$$

where $\overline{\mu}_{\tilde{A}_i}(x_1)$ and $\underline{\mu}_{\tilde{A}_i}(x_1)$ are the values of UMF and LMF of IT2FS $\tilde{A}_i$ at $x = x_1$; $\overline{\mu}_{\tilde{A}_i}(x)$, $\underline{\mu}_{\tilde{A}_i}(x)$ and $\overline{\mu}_{\tilde{A}_j}(x)$, $\underline{\mu}_{\tilde{A}_j}(x)$ are UMF and LMF of $\tilde{A}_i$ and $\tilde{A}_j$, respectively.

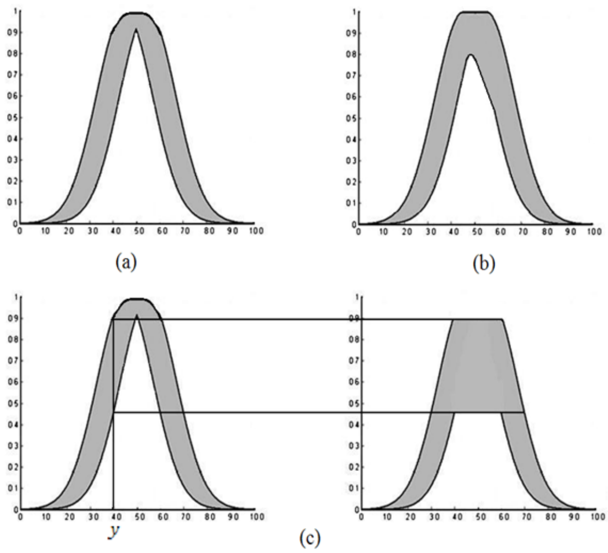

Fig. 4(a). FOU ($\tilde{A}_i$), (b). FOU ($\tilde{A}_j$), (c). Implication from $\tilde{A}_i$ to $\tilde{A}_j$ for $x_1 \in \tilde{A}_i$ to obtain $\tilde{A}_j'$

In figure 4, 4(a) and 4(b) show the FOUs of sets $\tilde{A}_i$ and $\tilde{A}_j$, respectively. For a given $x_1 \in \tilde{A}_i$ and the FLR $\tilde{A}_i \to \tilde{A}_j$, we can find the consequent set $\tilde{A}_j'$ by using the implication shown in 4(c).

E. Example

1. Let us choose a dataset load requirements of first 23 hours of 1st January 2004 to predict the 24th hour load demand. Here, $[f(1), f(2), f(3), f(4), f(5), f(6), f(7), f(8), f(9), f(10), f(11), f(12), f(13), f(14), f(15), f(16), f(17), f(18), f(19), f(20), f(21), f(22), f(23)]$ = [12930,12311,11805,11629,11674,11972,12433,12744,13370,14246,15042,15672,16064,16053,15960,16047,17033,18190,17964,17450,16708,15580,14186].
2. Since, the maximum and minimum values are $f(18)$ = 18190 and $f(4)$ = 11629, respectively, we define the universe of discourse $U = [11629, 18190]$ and divide into equal, non-overlapping intervals $I_1 = [11629, 13816)$, $I_2 = [13816, 16003)$ and$I_3 = [16003, 18190]$.
3. We define IT2FS $\tilde{A}_1$, $\tilde{A}_2$ and $\tilde{A}_3$. Then according to Equation (2), we obtain the IT2FTS

$$\tilde{F}(1), \tilde{F}(2), \tilde{F}(3), \tilde{F}(4), \tilde{F}(5), \tilde{F}(6), \tilde{F}(7), \tilde{F}(8), \tilde{F}(9), \tilde{F}(10), \tilde{F}(11), \tilde{F}(12), \tilde{F}(13), \tilde{F}(14), \tilde{F}(15), \tilde{F}(16), \tilde{F}(17), \tilde{F}(18), \tilde{F}(9), \tilde{F}(19), \tilde{F}(20), \tilde{F}(21), \tilde{F}(22), \tilde{F}(23)] = \begin{bmatrix} \tilde{A}_1, \tilde{A}_1, \tilde{A}_1, \tilde{A}_1, \tilde{A}_1, \tilde{A}_1, \tilde{A}_1, \tilde{A}_1, \tilde{A}_1, \tilde{A}_2, \tilde{A}_2, \tilde{A}_2, \tilde{A}_3, \tilde{A}_3, \\ \tilde{A}_2, \tilde{A}_3, \tilde{A}_3, \tilde{A}_3, \tilde{A}_3 \end{bmatrix}.$$

4. Using equations (3), we get

$$\begin{bmatrix} f_{3-mn}(3), f_{3-mn}(4), f_{3-mn}(5), f_{3-mn}(6), f_{3-mn}(7), \\ f_{3-mn}(8), f_{3-mn}(9), f_{3-mn}(10), f_{3-mn}(11), f_{3-mn}(12), \\ f_{3-mn}(13), f_{3-mn}(14), f_{3-mn}(15), f_{3-mn}(16), f_{3-mn}(17), \\ f_{3-mn}(18), f_{3-mn}(19), f_{3-mn}(20). f_{3-mn}(21). f_{3-mn}(22). \\ f_{3-mn}(23) \end{bmatrix}$$

$$= \begin{bmatrix} 12349,11915,11703,11758,12026,12383,12849, \\ 13453,14219,14987,15593,15930,16026,16020, \\ 16347,17090,17729,17868,17374,16579,15491 \end{bmatrix}$$

and

$$\begin{bmatrix} f_{3-sd}(3), f_{3-sd}(4), f_{3-sd}(5), f_{3-sd}(6), f_{3-sd}(7), \\ f_{3-sd}(8), f_{3-sd}(9), f_{3-sd}(10), f_{3-sd}(11), f_{3-sd}(12), \\ f_{3-sd}(13), f_{3-sd}(14), f_{3-sd}(15), f_{3-sd}(16), f_{3-sd}(17), \\ f_{3-sd}(18), f_{3-sd}(19), f_{3-sd}(20), f_{3-sd}(21), \\ f_{3-sd}(22), \\ f_{3-sd}(23) \end{bmatrix}$$

$$= \begin{bmatrix} 460,289,74,152,312,317,389,616,682,583,420, \\ 182,46,42,486,875,500,309,515,768,1031 \end{bmatrix}.$$

5. We find the consequent IT2FS for $f(23)$ =14186 using the FLRG $\tilde{A}_3 \rightarrow \tilde{A}_2, \tilde{A}_3$ using equation (12). The inference is shown in figure 5.

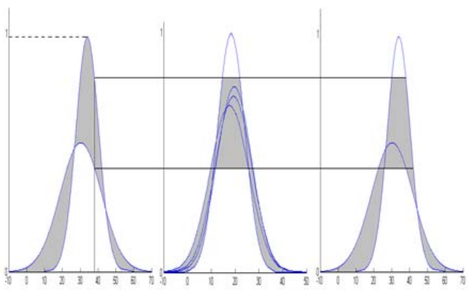

Fig. 5. Inference using $\tilde{A}_3 \rightarrow \tilde{A}_2, \tilde{A}_3$ for $f(23)$ = 14186

6a. The two consequent IT2FS $\tilde{A'}_1$ and $\tilde{A'}_3$ are type reduced to obtain T1FS A'_1 and A'_3. This is illustrated in figure 6.

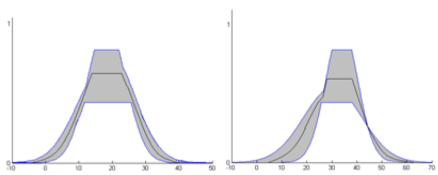

Fig. 6. Type reduction of IT2FS $\tilde{A'}_1$ and $\tilde{A'}_3$ to obtain T1FS A'_1 and A'_3

6b. The T1FS A'_1 and A'_3 obtained in Step 6a are defuzzified using the formula in equation (13) and (14) to obtain crisp consequent values, $C_{A'_1}$ and $C_{A'_3}$.

7. The expected value of the consequent (predicted value of $f(23)$) is obtained using equation (15).

III. PROPOSED ALGORITHM

IT2FS Based Time Series Forecasting Algorithm

1. Define the universe of discourse and divide into intervals
2. Define IT2FS for each of the intervals and fuzzify the time series
3. System Training
   3a. Construct FOUs for each of the IT2FSs
   3b. Construct FLRs using historic time series data and group into FLRGs
   3c. Compute Transition Probability Matrix (TPM)

Repeat steps 4, 5 and 6 for each data point to be predicted

4. Find the consequent IT2FS using the FLRs for the previous data point corresponding to the IT2FS having maximum transition probability.
5. Type Reduction and Defuzzification:
   a) Type reduce the consequent IT2FS using equation (13)
   b) Defuzzify the consequent T1FS using equation (14)
6. Predict the next value as the centroid and calculate errors RMSE and MAPE with respect to original data.

IV. NEW METHOD APPLIED TO ELECTRICAL LOAD FORECASTING

This new method can be applied for predicting any time series, but we have applied it on historic electric load data to forecast electric load.

A. Choosing a Time Series

1: Select the time series to be predicted and also the time-span relevant for training and prediction (testing).

2: Find the variation for the training data set. The quantity variation is defined as follows.

$$variation(t) = load(t) - load(t-1)$$

In the proposed method, the time series used for prediction of is the hourly variation of the electric load. Figure 7.shows the plots of the hourly load value and the hourly variation of an electric load over a period of time.

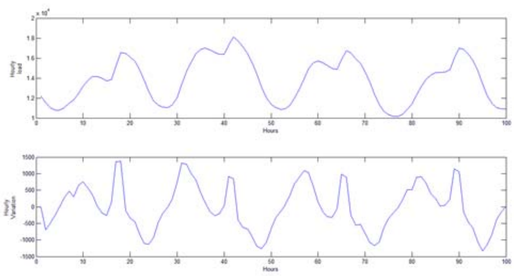

Fig. 7 hourly plot electric load value and its variation

In the strict sense, a stationary time series is such that the joint probability distribution of any n consecutive observations of the series remains constant. In the weak sense, a time series is said to be stationary if mean and standard deviation of the time series data do not vary much with time. An inspection of figure 3 shows us that the variation data maintains a steadier mean and standard deviation compared to hourly load data and can be considered to be a weakly stationary time series. Therefore, the use of daily variation in the proposed method is expected to improve performance. Another advantage, arising due to the bounded nature of the variation data, is that the universe of discourse becomes smaller compared to the hourly load data used. This also results in a decrease in computational complexity, since computational complexity is directly proportional to the size of the universe of discourse.

In this paper we have used electric load demand of 2004-2008 provided by ISO New England IRC (available on the internet on the Mathworks Inc. website) as our time series of interest. The hourly load data for each hour of each day from January through October is selected as the training series and the rest of the series from November through December is selected as the test data set.

B. Define the universe of discourse and Intervals

1. The universe of discourse is defined as

$$U = [D_{min} - m_1, D_{max} + m_2] \quad (7)$$

where D_{min} and D_{max} are the minimum and maximum value of training variation, respectively; while m_1 and m_2 are margins of safety which act as buffer zones in case the minimum and maximum values of test variation exceeds that of the training variation.

2. Divide the universe of discourse U into n intervals in the following way, where n_{1_i} and n_{2_i} are appropriate positive real values so that the universe of discourse U is divided into n intervals of equal length $I_1, I_2, \ldots, I_n$.

$$I_i = [D_{min} - m_1 + n_{1_i}, D_{max} + m_2 - n_{2_i}] \quad (8)$$

For example, we define the universe of discourse and the intervals for electric load as follows.

After inspecting the daily variation over the period of 1st January through 30th September of the year 2006, we obtain the minimum variation D_{min} = -2695 and the maximum variation D_{max} =2398.

We select the margins of safety m_1 = 5 and m_2 = 2. These margins are required to ensure that maximum and minimum variations, for the test period of November and December, remain within the universe of discourse U. Hence, the universe of discourse is

$$U = [-2700, 2400]$$

We then divide the universe of discourse into 5 intervals of length 1000 each as

$$I_1 = [-2700, -1700], I_2 = [-1700, -700], I_3 = [-700, 300], \ldots, I_5 = [1300, 2400].$$

C. FOU Construction

1: Find b-day mean and standard deviation of training variation series.

Calculate the b-day mean of training deviation as

$$mean(t) = \frac{1}{b}\sum_{k=1}^{b} variation(t-(k-1)) \quad (9)$$

Calculate the b-day standard deviation (sd) of training variation as follows.

$$sd(t) = \sqrt{\frac{1}{b}\sum_{i=1}^{b-1}(variation^2(t-i+1)) - mean^2(t)}$$

2: Define interval type 2 fuzzy sets and find the memberships.

a: Define n IT2FS $\tilde{A}_1, \tilde{A}_2, \ldots, \tilde{A}_n$ corresponding to the intervals $I_1, I_2, \ldots, I_n$.

b: Find the interval type 2 membership of each set using the following rules.

Rule 1: If $mean(t) \in I_i$, then $mean(t)$ and $sd(t)$ will contribute to the membership of set $\tilde{A}_i$, using rule 2.

Rule 2: The UMF and LMF of the set $\tilde{A}_i$ at point x, i.e., $\overline{\mu}_{\tilde{A}_i}(x)$ and $\underline{\mu}_{\tilde{A}_i}(x)$ are obtained as follows:

$$\overline{\mu}_{\tilde{A}_i}(x) = max_{\forall mean(t)\in I_i}\left\{\frac{1}{\sqrt{2\pi}sd(t)}e^{\frac{1}{2}\left(\frac{-(x-mean(t))}{sd(t)}\right)^2}\right\}$$

$$\underline{\mu}_{\tilde{A}_i}(x) = min_{\forall mean(t)\in I_i}\left\{\frac{1}{\sqrt{2\pi}sd(t)}e^{\frac{1}{2}\left(\frac{-(x-mean(t))}{sd(t)}\right)^2}\right\} \quad (11)$$

Rule 3: FOU must be scaled such that the highest value becomes 1.

For example, we calculate the mean and standard deviation of the training variation data for the electric load over 5 days; using equations (9) and (10).

We select b = 5 for the simple reason that there are 5 working days in a week. One may also opt for other values of b. The values of electric load demand for a single day may fluctuate considerably from the 5-day mean due to irregularities in the factors affecting load demand, which can be likened to noise in a physical system. Therefore, we use the 5-day mean as a measure of the true value of load over the period. Standard deviation, on the other hand, serves as a measure of the uncertainty associated with the true value. By using this approach, we have eliminated the need to use highest and lowest values of electric load.

D. Rule Construction

1: Form the FLRs and the FLRGs.

a: If $variation(t-1) \in I_i$ and $variation(t) \in I_j$, then the following FLR is formed.

$$A_i \rightarrow A_j$$

b: FLRGs are formed when there are multiple FLRs with common LHS, as given in definition 4.

In the present example, we form the following FLRs from the data shown in Table 3.

$$A_3 \rightarrow A_3, A_3 \rightarrow A_4, A_4 \rightarrow A_4, A_4 \rightarrow A_3$$

These FLRs are then grouped into FLRGs as follows.

$$A_3 \rightarrow A_3, A_4$$
$$A_4 \rightarrow A_4, A_3 \quad (12)$$

E. Probability of Occurrence

Calculate the probabilities of occurrence of each FLR for a given FLRG and store in the Transition Probability Matrix (TPM). Since, there may be multiple consequent outputs for a single antecedent depending on the number of IT2FS in the RHS of the FLRG in question, the expected value of the consequent is determined using the probabilities of occurrences of each FLR in the FLRG.

F. Draw IT2FS Inference

In the Inference stage, the consequent output interval type 2 fuzzy set, for a given input value (which is the previous day variation for stock index prediction), is obtained using the FLRs and the FOUs of the corresponding sets.

G. Type Reduction & Defuzzification

1: Reduce the type of the interval type 2 fuzzy output sets obtained for all the sets in the RHS, of the FLRG corresponding to the previous variation, by using centroid technique. The concerned formula is as follows:

$$\mu_{\tilde{A}_j}(y) = \frac{\overline{\mu}_{\tilde{A}_j}(y) + \underline{\mu}_{\tilde{A}_j}(y)}{2} \quad (13)$$

2: Defuzzify the type reduced sets using centroid defuzzification to obtain crisp values V_{A_j} for each set in the RHS of the FLRG. The centroid formula is as follows:

$$V_{A_j} = centroid_{\tilde{A}_j} = \frac{\int_{D_{min}-m_1}^{D_{max}+m_2} \mu_{\tilde{A}_j}(y) \times y \ dy}{\int_{D_{min}-m_1}^{D_{max}+m_2} \mu_{\tilde{A}_j}(y) \, dy} \quad (14)$$

To obtain a crisp output value, the consequent IT2FS must be type reduced and then defuzzified. Fig 6 illustrates type reduction and defuzzification of IT2FS.

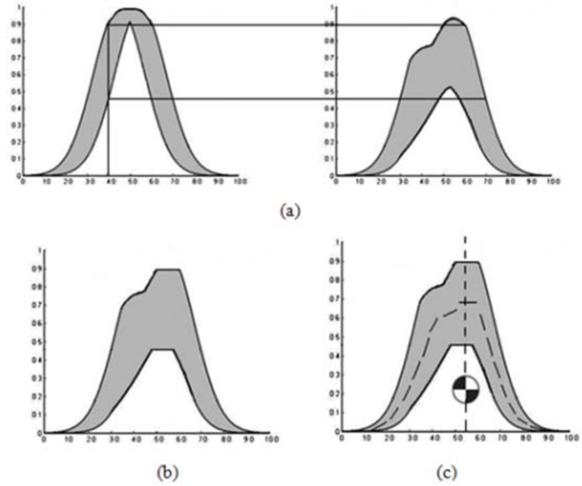

Fig.6. Fuzzy Implication from an interval type 2 fuzzy set A_i onto another set A_j; (b). The resultant output interval type 2 fuzzy set O_j; (c). The centroid and the crisp value 51 of O_j

H. Expected Value of the Consequent:

1: Find the expected value of the predicted variation by combining all the obtained crisp values according to their probability from the TPM.

$$variation(t+1) = V_{A_{jk}} such\ that\ a_{i,jk} \geq a_{i,jl} \forall k \neq l \quad (15)$$

As explained in the previous stages, there can be multiple consequents for a single antecedent. Hence, to obtain a single output value, we find the expected value of the consequent using the probabilities of occurrence found in step 3c. The consequent values obtained for each of the FLRs in the relevant FLRG is the one corresponding to the maximum probability of occurrence.

Finally, after following the steps 1 through 6, the predicted variation for the next day is found. This is then added with the present load demand value, to predict the next value.

$$load(t+1) = variation(t+1) + load(t) \quad (16)$$

Finally, it should be noted that the proposed interval type 2 fuzzy time series prediction method is a static one. The nature of the FOUs and the FLRs keep changing dynamically over time due to the high degree of non-linearity of the electric load. Therefore, the system must be retrained periodically to achieve satisfactory results. This has been illustrated in algorithm proposed in Section III. The plot of the predicted values vs. the actual value can be seen in Fig. 7.

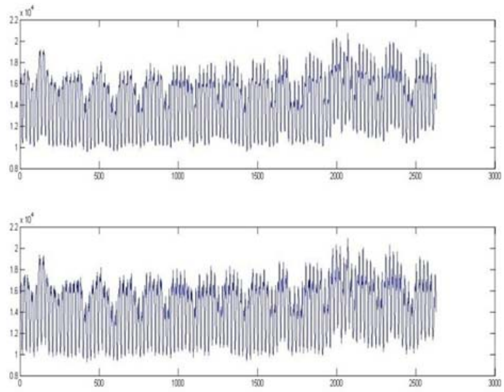

Fig. 7. Plot of Actual vs. Predicted values of electric load for the year 2006

V. EXPERIMENTAL RESULTS

Dataset and Experimental setup used: The proposed algorithm has been tested against the ISO New England electric load dataset. Hourly Electric Load variations of ISO New England from 01.01.2004 to 31.12.2009 were available on [1]. Experiments has been performed in MATLAB 7.8.0 environment on a HP Pavilion g6 laptop with 4GB RAM, Intel Core-i5 CPU (2.60 GHz) and 64bit Windows-8 Operating System. In this paper we have given two commonly used performance metrics Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE) for different experimental setting of the algorithm on different range of the prediction interval and training series.

$$RMSE=\sqrt{\frac{\sum(Predicted\ Value-Original\ Value)^2}{Number\ of\ points\ in\ test\ series}}$$

$$MAPE=\frac{\sum|Predicted\ Value-Original\ Value|}{Number\ of\ points\ in\ test\ series} \quad (17)$$

Algorithm Parameters: Here we have tuned two algorithm parameters namely the Interval Size based on the deviation series (I) and the Moving Average Window Size (W). We have tuned these two give the optimum results. An Interval size of 1000 and Moving Average Window Size of 5 gave us satisfactory results with respect to both of the performance measure said above.

Results: We have tested this algorithm on the aforesaid database multiple times by varying the training and testing series length and their ratio i.e. a certain portion of the total historic load data available (2004-2008) was used for the experiment. Now training series to testing series ratio is varied and the algorithm is run and tested on them. The results of the experiments are listed in different tables below.

TABLE I
HOURLY VARIATION OF LOAD IN ONE DAY (01.01.2004) AND COMPARISON BETWEEN ACTUAL AND PREDICTED DATA.

Hour	Predicted Load Value (in MW)	Actual Load Value (in MW)
1	16.56588	16.854
2	17.322	16.978
3	17.45024	17.877
4	18.33228	19.296
5	19.7459	19.313
6	18.91501	18.839
7	18.43758	18.149
8	17.74872	17.256
9	16.23276	16.232
10	15.20868	15.102
11	14.7139	14.312
12	13.92351	13.935
13	13.56868	13.807
14	14.25687	13.834
15	14.30717	14.137
16	14.60781	15.009
17	15.46591	16.459
18	16.90892	17.633
19	18.1007	18.482
20	18.95475	18.996
21	19.47055	19.241
22	18.87456	19.167
23	18.76285	18.859
24	16.78654	16.854

TABLE II
COMPARISON OF AVERAGE PREDICTED VALUE AND AVERAGE ACTUAL VALUE OVER A WEEK (01.01.2016-07.01.2016)

Day of the Week	Average Predicted Value	Average Actual Value
Sunday	16.9304	17.0143
Monday	17.0084	17.1495
Tuesday	16.0216	16.0494
Wednesday	15.7216	15.7441
Thursday	15.4284	15.5002
Friday	15.3088	15.4012
Saturday	15.2085	15.198

TABLE III
PREDICTION ERROR FOR DIFFERENT AMOUNT OF DATA POINT CONSIDERED AND SPAN OF THE TRAINING SERIES

Total time series data used for experiment (no. of data points)	Training Series : Testing Series division (--% --%)	RMSE	MAPE
200	50-50	0.442	0.176
--	60-40	0.429	0.1609
--	70-30	0.438	0.0671
--	80-20	0.565	0.242
--	90-10	0.562	0.219
400	50-50	0.336	0.1596
--	60-40	0.442	0.1974
--	70-30	0.462	0.1478
--	80-20	0.417	0.0219
--	90-10	0.445	0.167

Total time series data used for experiment (no. of data points)	Training Series : Testing Series division (--% --%)	RMSE	MAPE
500	50-50	0.449	0.3178
--	60-40	0.443	0.3344
--	70-30	0.427	0.2975
--	80-20	0.441	0.4225
--	90-10	0.445	0.431
1000	50-50	0.449	0.35
--	60-40	0.456	0.4051
--	70-30	0.444	0.4
--	80-20	0.443	0.45
--	90-10	0.451	0.43
2000	50-50	0.42	0.39
--	60-40s	0.42	0.3831
--	70-30	0.426	0.4147
--	80-20	0.428	0.3886
--	90-10	0.445	0.3923

TABLE IV
CALCULATION OF RMSE AND MAPE FOR THE CORRESPONDING YEARS OF THE DATASET

Year	Training Series Span: Test Series Span (in months)	RMSE	MAPE
2004	6-6	0.454	0.655
--	7-5	0.446	0.663
--	8-4	0.432	0.669
--	10-2	0.462	0.795
--	11-1	0.481	0.848
2005	6-6	0.425	0.591
--	7-5	**0.285**	**0.0258**
--	8-4	0.442	0.288
--	10-2	0.361	0.412
--	11-1	0.385	0.469
2006	6-6	0.45032	0.78
--	7-5	0.42056	0.60
--	8-4	0.29434	0.35
--	10-2	0.38255	0.42
2007	6-6	0.44262	0.86
--	7-5	0.38832	0.54
--	8-4	0.2825	0.40
--	10-2	0.4204	0.66
2006 & 2007	6-6	0.48523	1.21
	7-5		0.67
	8-4	0.44383	0.71
	10-2	0.44562	0.81

Comparison with other research articles: In [10], a particle swarm optimization based memetic algorithm for model selection in STLF using support vector regression, has been proposed. The paper uses hourly electricity load in New York City during 2003-2004. This paper shows that MAPE for methods like SVM, Hybrid Network, and Wavelet Neuro have MAPE 3.03, 2.29 and 2.02 and their CLPSO-MA-SVR gives MAPE of 1.43. Our method, though tested on different datasets, gives better results.

In [11] load forecasting results using SSA-SVR on ISO-NE 2006 dataset gave MAPE in between 1.0 to 2.1 whereas the same due to our method on ISO-NE 2006 never exceeds 1.0 Although in [9] weather forecasts and actual weather data has been used but we haven't considered any such secondary time series.

VI. CONCLUSION

A new Interval Type-II Fuzzy Set based Time Series Prediction has been presented. The method is used for the purpose of Electric Load Forecasting. This comes out to be an efficient method for Load Forecasting with very small amount of error in prediction. This method doesn't use any secondary time series like Weather data, week day – holiday index, etc. which are commonly used in other methods. In future we hope to apply this method for some other forecasting scenarios also.

REFERENCES

[1] Park, D.C.; El-Sharkawi, M.A.; Marks, R.J.; Atlas, L.E.; Damborg, M.J., "Electric load forecasting using an artificial neural network," *IEEE Transactions on Power Systems,* vol. 6, no. 2, pp. 442-449, 1991.

[2] Swaroop R, Hussein Ali Al Abdulqader., "Load Forecasting for Power System Planning using Fuzzy-Neural Networks," *Proceedings of the World Congress on Engineering and Computer Science*, 2012 Vol I.

[3] P. Lauret, E. Fock, R. N. Randrianarivony, and J. F. M. Ramsamy, "Bayesian neural network approach to short time load forecasting," *Energy Convers. Manage,* vol. 49, no. 5, pp. 1156–1166, 2008.

[4] Amral, N.; Ozveren, C. S.; King, D., "Short term load forecasting using Multiple Linear Regression," *IEEE 2007 42nd International Universities Power Engineering Conference,* Brighton, UK, 2007.

[5] J. Nowicka-Zagrajek, R. Weron "Modeling electricity loads in California: ARMA models with hyperbolic noise," *Elsevier Signal Processing*, vol. 82, no. 12, pp. 1903-1915, 2002.

[6] Cheng-Ming Lee; Chia-Nan Ko, "Short-term load forecasting using lifting scheme and ARIMA models," *Elsevier- Expert Systems with Applications*, vol. 38, no. 5, pp. 5902-5911, 2011.

[7] Nima Amjady, Farshid Keynia, "Mixed price and load forecasting of electricity markets by a new iterative prediction method," *Elsevier Electric Power Systems Research*, vol. 79, no. 9, pp. 2678-2687, January 2009.

[8] El Desouky, A.A., Elkateb, M.M., "Hybrid adaptive techniques for electric-load forecast using ANN and ARIMA," *IEEE Proceedings - Generation Transmission and Distribution*, vol. 147, no. 4, pp. 213-217, July 2000.

[9] H.S. Hippert, C.E. Perdreira, and R.C. Souza, "Neural networks for short-term load forecasting: a review and evaluation," *IEEE Trans. Power Syst.,* vol. 16, pp. 44-55, Feb. 2001.

[10] Zhongyi Hu, Yukun Bao, Tao Xiong, "Comprehensive learning particle swarm optimization based memetic algorithm for model selection in short-term load forecasting using support vector regression," *Elsevier Applied Soft Computing*, vol. 25, pp. 15-25, September 2014

[11] Ervin Ceperic, Vladimir Ceperic, Adrijan Baric, "A Strategy for Short-Term Load Forecasting by Support Vector Regression Machines," *IEEE Transactions on Power Systems,* vol. 28, no. 4, pp. 4356-4364, June 2013.