An Enhanced Fuzzy Linguistic Term Generation and Representation for Time Series Forecasting

Atakan Sahin, Tufan Kumbasar, Engin Yesil, M. Furkan Dodurka, Onur Karasakal, Sarven Siradag

Abstract—This paper introduces an enhancement to linguistic forecast representation using Triangular Fuzzy Numbers (TFNs) called Enhanced Linguistic Generation and Representation Approach (ElinGRA). Since there is always an error margin in the predictions, there is a need to define error bounds in the forecast. The interval of the proposed presentation is generated from a Fuzzy logic based Lower and Upper Bound Estimator (FLUBE) by getting the models of forecast errors. Thus, instead of a classical statistical approaches, the level of uncertainty associated with the point forecasts will be defined within the FLUBE bounds and these bound can be used for defining fuzzy linguistic terms for the forecasts. Here, ElinGRA is proposed to generate triangular fuzzy numbers (TFNs) for the predictions. In addition to opportunity to handle the forecast as linguistic terms which will increase the interpretability, ElinGRA improved forecast accuracy of constructed TFNs by adding an extra correction term. The results of the experiments, which are conducted on two data sets, show the benefit of using ElinGRA to represent the uncertainty and the quality of the forecast.

Keywords—*forecasting; fuzzy time series; fuzzy numbers; fuzzy estimator; Prediction Interval; fuzzy linguistic terms*

I. INTRODUCTION

In the recent years, computational intelligence methods have been widely employed as prediction and estimation approaches [1]. It has been stated in [2] that there are two main problems with state of art forecasting methods: (i) the models become unreliable in the presence of uncertainty and (ii) no indication of the accuracy of the single point forecasts is provided. The accuracy of the forecast is usually measured with performance indexes such as Mean Absolute Percentage Error (MAPE), Percentage of Error (POA), etc. [3]. Though, since there is always an error margin in the predictions, there is a need to define error bounds in the forecast with its Confidence Interval (CI), Prediction Interval (PI) or using other novel approaches.

The CIs handle with the accuracy of the prediction of the regression while the PIs consider the accuracy with the prediction to the targets values [4]. A PI is constructed from interval bound which covers the future unknown value with a

A. Sahin, T. Kumbasar, E. Yesil and O. Karasakal are with the Control and Automation Engineering Department, Istanbul Technical University, Istanbul, Turkey (e-mail: {sahinata, kumbasart, yesileng, karasakalo} @itu.edu.tr).

M. F. Dodurka and S. Siradag are with Getron Bilisim Hizmetleri A. S., Istanbul, Turkey (e-mail: {furkan.dodurka, sarven.siradag}@getron.com).

prescribed probability called a confidence level $((1 - \alpha)\%,$ $\alpha \in [0, 1])$ [5]. The availability of PIs allows the decision makers to quantify the level of uncertainty associated with the point forecasts. A relatively wide PI indicates the presence of high level of uncertainties in the underlying system operation. On the other hand, narrow PIs give the decision makers the opportunity to decide more confidently with less chance of confronting an unexpected condition in the future. This useful information can guide the decision makers to avoid the selection of risky actions under uncertain conditions. Thus, the construction of PIs has been a subject of much attention [6]. Thus, different methods haven been proposed for the construction of PIs such as delta technique [7, 8], Bayesian technique [9], bootstrap [10], mean-variance estimation [11], lower and upper bound estimation method [12]

Recently, a Fuzzy logic based Lower and Upper Bound Estimator (FLUBE) is proposed to estimate the uncertainty in the forecast [13]. Then, the FLUBE is used for the Linguistic Term Generation and Representation Approach (LinGRA) so that the forecasts are represented with the linguistic terms, which are defined with Triangular Fuzzy Numbers (TFNs). In the LinGRA, the support of the TFN is constructed on the output interval generated by the FLUBE while its center is directly assigned to the forecast value [13].

In this study, we will enhance the performance of the FLUBE based LinGRA by implementing an extra FLS which generates a Center Point Correction Term (CPCT) for the TFN representation. It will be shown that, in comparison with the LinGRA, the Enhanced LinGRA (ElinGRA) will increase the information about the accuracy and success of the single point forecast by providing a relative membership degree (μ). Thus, the linguistic forecast representation will give the opportunity to the decision maker to quantify the uncertainty of the point forecasts with the linguistic terms which might increase the interpretability for further assessments. The proposed approach consists of two main parts, the FLUBE design and the linguistic forecast generation with the ElinGRA via TFNs. Thus, we will start by presenting the internal structure and the design steps of the FLUBE. We will then present the ElinGRA in comparison with the LinGRA. To illustrate the superiority of the proposed approach, experimental results are presented on the benchmark data sets. Based on the presented results, it will be concluded that ElinGRA is an efficient and useful approach to evaluate the success of the single point forecast in comparison with the LinGRA.

Section II introduces the preliminaries. Section III presents the FLUBE based LinGRA and the ElinGRA. Section IV includes the experimental results and discussion, and Section V presents the conclusions and future works.

II. Preliminaries

A. SARIMA Model

AutoRegressive Integrated Moving Average (ARIMA) and Seasonal ARIMA (SARIMA) models are enhancements of ARMA class in order to include more dynamics, respectively, non-stationary in mean and seasonal behaviors. The multiplicative SARIMA $(p, D, q)(p_s, D_s, q_s)_s$ model is defined as follows:

$$\phi(L)\Phi(L^s)(1 - L)^D(1 - L^s)^{D_s}S_t = c + \theta(L)\Theta(L^s)\varepsilon_t \quad (1)$$

where D is the integrated order and s defines the periodicity, c is the constant of the model. The SARIMA model express with the conditional mean of both past observations $(S_{t-1}, \ldots, S_{t-p})$ of the target data which are parameterized as S, and past innovations $(\varepsilon_{t-1}, \ldots, \varepsilon_{t-q})$ where the number of past observations and innovations determined with the autoregressive (AR) parameter (p) and the moving average parameter (q). Addition to the observations $\Delta^D S_t$ denotes a Dth differenced time series, and ε_t is an uncorrelated innovation process with mean zero. Seasonality which is difference between ARIMA and SARIMA can be described with lag operator (L) which commonly using define as $L^i S_t = S_{t-i}$ on following parameter definition:

$$\begin{aligned}
\phi(L) &= \left(1 - \phi_1 L - \cdots - \phi_p L^p\right) \\
\theta(L) &= \left(1 - \theta_1 L - \cdots - \theta_q L^q\right) \\
\Phi(L) &= \left(1 - \Phi_1 L^s - \cdots - \Phi_p L^{sP}\right) \\
\Theta(L) &= \left(1 - \Theta_1 L^s - \cdots - \Theta_Q L^{sQ}\right)
\end{aligned} \quad (2)$$

where p and P degree of AutoRegressive (AR) and Seasonal AutoRegressive (SAR) operator parameterized as ϕ and Φ, q and Q degree of Moving Average (MA) and Seasonal Moving Average (SMA) operator parameterized as θ and Θ [14, 15]. The performance of the SARIMA forecast model can be evaluated with respect to its performance indexes such as MAPE and POA values [3, 14].

B. Performance Measures for PI Methods

PI Coverage Probability (PICP) and PI Normalized Averaged Width (PINAW) which are commonly used in literature to evaluate the PI performance. [16] The PICP is measured by counting the number of target values covered by the constructed PIs. The PICP shows in which probability target values will be covered by the lower and upper bounds and thus is defined as:

$$PICP = \frac{100}{n}\sum_{k=1}^{n} j_k \quad (3)$$

where n is the number of samples and j_k is [16]:

$$j_k = \begin{cases} 1, & \text{if } y_k \in [L_i \ U_i] \\ 0, & \text{if } y_i \notin [L_i \ U_i] \end{cases} \quad (4)$$

The second measure is the PINAW which provides a measure about the width of the PIs and is defined as [16]:

$$PINAW = \frac{100}{nR}\sum_{k=1}^{n}\left(U(X_k) - L(X_k)\right) \quad (5)$$

where R is the range of underlying targets. Same as in the PICP measure, using the extreme target values as lower and upper bounds of PIs will result with 100% PINAW. From a practically point of view, it is important to have narrow PIs (relatively small PINAW value) with a high coverage probability (relatively high PICP value) [17].

C. Data Sets

1) Data Set-1: The Australian monthly electrical consumption data set

The Australian monthly electricity consumption data set [18] will be used to illustrate the proposed approach. This data set involves the electricity consumption values from January 1956 to August 1995, thus the data set has a total of 476 samples. As it can be clearly seen in Fig. 1, the consumption of the electricity has always increased with respect the time which shows the trend property of the data. Moreover, the data has a seasonality characteristic of 12 months described with parameter *s* which can be clearly seen from subplots presented in Fig. 1. Thus, the data has trend and seasonality characteristics which are commonly encountered in time series analysis [14, 15].
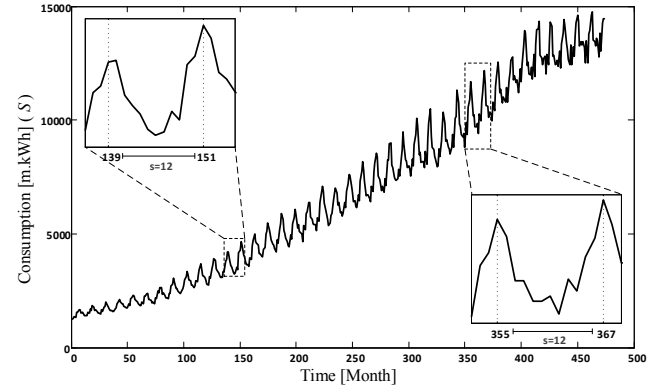

Fig. 1. The Australian monthly electricity consumption data set.

2) Data Set-2: The air passenger data set

The air passenger data set [18] will also be used to illustrate the proposed approach. This data set involves monthly counts of the international airline passengers, measured in thousands, for the period January 1949 through December 1960, thus the data set has a total of 144 samples. Just as electricity consumption data set, air passenger data set has same characteristic as it can be seen in Fig. 2.
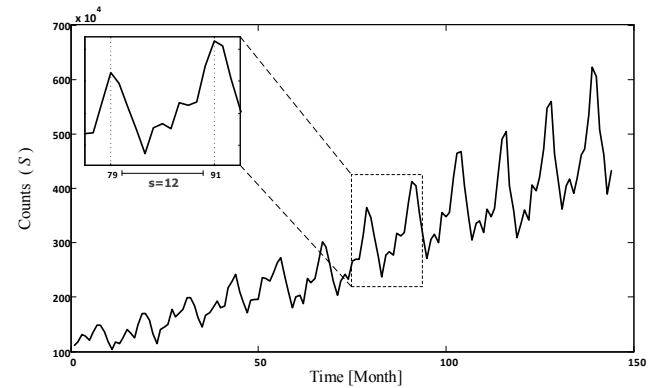

Fig. 2. The air passenger data set.

III. FORECAST REPRESENTATION OF TFN

In this section, we will firstly present the FLUBE which will define the uncertainty interval of the single point forecast value [13]. Then instead of conventional PI representation, the level of uncertainty associated with the point forecasts will be quantified by defining TFNs within the uncertainty interval provided by the FLUBE.

A. Design of the FLUBE

The FLUBE consists of two Fuzzy Logic Systems (FLSs) which will define the uncertainty bounds of the point forecast error. The FLUBE is constructed by choosing the input to be the target data (S) and the output as the forecast (F) error terms ($E = F - S$) [13]. The required data sets are chosen by the Selection Algorithm (SA). Then, for the design of the FLUBE two unique data sets which are $[S_{min}, E_{min}]$ for the training of the Lower FLS (LFLS) and $[S_{max}, E_{max}]$ for the training of the Upper FLS (UFLS) are collected. For the training of the FLUBE, the ANFIS toolbox/MATLAB will be used to generate the LFLS and UFLS. The fuzzy rule base structures of the LFLS and UFLS are as follows:

LFLS: $\quad R_\omega^L$: If S is A_ω^L, Then E^L is D_ω^L $\quad$ (6)

UFLS: $\quad R_\omega^U$: If S is A_ω^U, Then E^U is D_ω^U $\quad$ (7)

where A_ω^L and A_ω^U, are the antecedent membership functions (MFs), D_ω^L and D_ω^U are the consequent crisp sets and W ($\omega = 1, ... W$) is the total number of rules. E^L and E^U represent the lower and upper error forecast values [13].
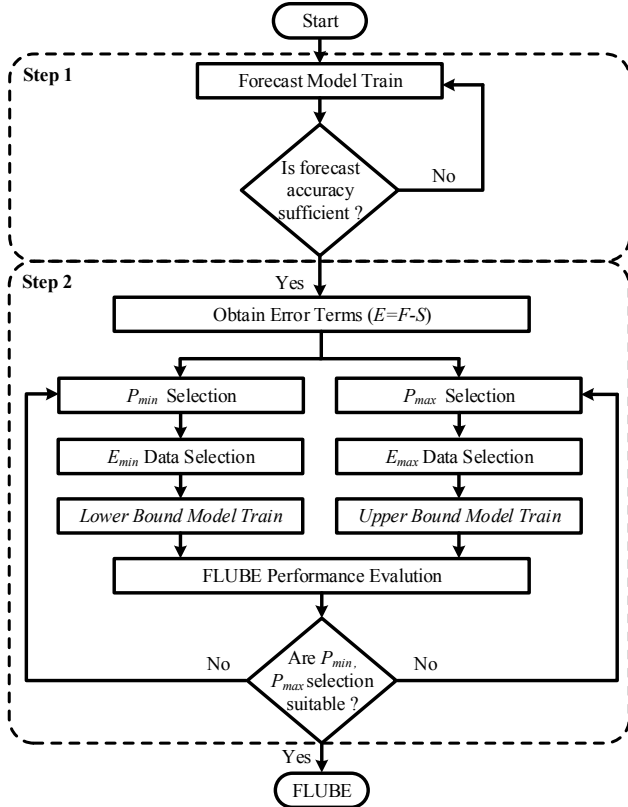

Fig. 3. The flow chart of the training procedure of the FLUBE.

In Fig. 3, the flow chart of the training procedure of the FLUBE is presented which consists of two main steps. The first step is to construct a conventional forecast model while in the second step the FLUBE is designed. We will explain the presented analysis on the Data Set-1 for illustrative purposes.

Step 1: Forecast phase

Here, the time series forecast model is constructed, by using the target data (S) to predict the forecast value (F). The forecast model can be constructed from different structures such as time series regression, conditional means/variance and multivariate models [14]. In general, the forecast model should provide a satisfactory performance to result with an acceptable FLUBE training. In this study, we will prefer a SARIMA forecast model since it results with an acceptable forecast performance for the handled electrical consumption data set. Note that the FLUBE can be also easily employed if other forecast models are preferred.

Step 2: FLUBE training

In this step, we will collect the required training data sets which are defined as follows:

$$E_k = F_k - S_k \quad (8)$$

where F_k is the forecast value at the kth sample. In Fig. 4, the training data set is illustrated for the electricity consumption data set. Then, the SA is employed to collect the required the data sets which will be used at the training of FLUBE. The data set for LFLS will be constructed with negative error terms and will be labeled as $[S_{min}, E_{min}]$ which are illustrate with red circles in Fig. 4. In a similar manner, the dataset for the UFLS will be labeled as $[S_{max}, E_{max}]$ which are illustrate with blue circles in Fig. 4. As it has been asserted, these two data sets will be chosen by the SA with a tuning parameter P. The design parameter (P) of the SA decomposes the minimum and maximum target data with equal and fixed subintervals. The SA picks the maximum/minimum error values for each decomposed subinterval. Thus, it can be concluded that the parameter P defines the maximum size of the training data set.
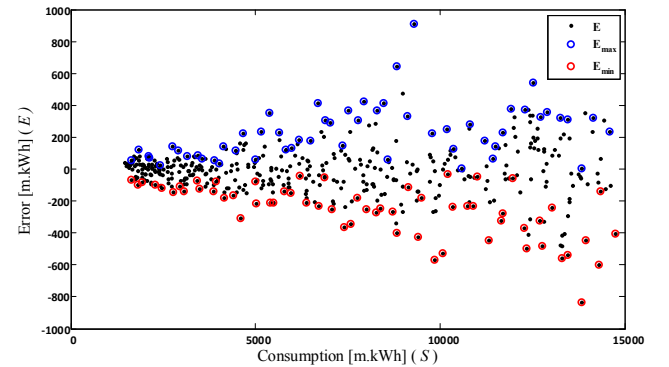

Fig. 4. Selected error terms by Selection Algorithm according to target data of the Data Set-1

For the collection of training data set of the UFLS $[S_{max} \; E_{max}]$, the SA will perform as follows:
- If $E_k > 0$, then the SA will select and insert the corresponding data into training data set of the UFLS $[S_{max}, E_{max}]$.
- If $E_k < 0$, then the SA will not update the training data set for this subinterval.

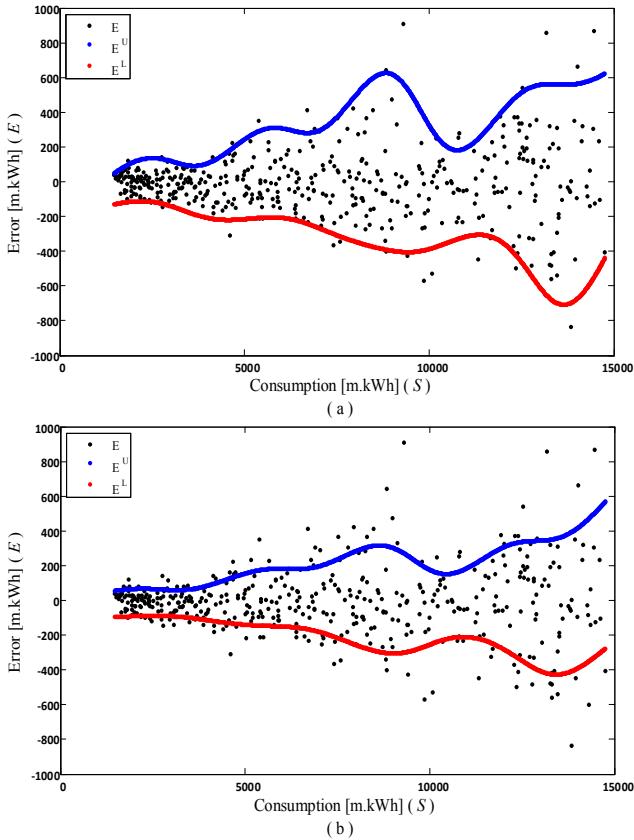In a similar manner, the $[S_{min}, E_{min}]$ data set will be collected.

Fig. 5. The selection parameter effect on the FLUBE bounds determination on the Data Set-1 for a) P=60, b) P=320.

It can be observed that selection of the SA parameter P affects size of the train data sets. The SA will collect more training data points for relatively big P values. Thus, the generated FLUBE might result with a general approximation while neglecting the characteristics of the error values. Therefore, P an important design parameter and needs to be tuned for each handled data set [13]. To illustrate this concept, for the Data Set-1, the performance of the FLUBE is illustrated in Fig. 5 for the P values $P = 60$ and $P = 320$.

(Note that, we have employed $P = P_{min} = P_{max}$ throughout the paper). It can be clearly seen that the relatively small value of P ($P = 60$) provides bounds which cover almost all error terms while the relatively big value of P ($P = 320$) results a smoother bound characteristic as van be seen at Fig. 5b. Hence, it can be concluded that the P parameter must be selected such that to provide a tradeoff between coverage performance and bound characteristic.

Remark: The mean value of the error distribution must be shifted from the error terms (E_{min}, E_{max}) to make the expected value of the error as zero. Thus, the shifted mean value, at the end, has to be added to the output of the FLUBE.

B. Linguistic Forecast Generation

In this subsection, the ElinGRA, which is the improved version of the online LinGRA [13] will be presented. In this context, we will present two novel approaches that try to explain how the target values (S) are similar to the forecast values (F) by using linguistic terms which cannot be accomplished by the conventional PI representations [13]. As it has been asserted, the FLUBE has been constructed where the target data (S_k) is the input. However, since the S_k value will not be available for evaluating FLUBE at the kth sample, we will use the single point forecast value F_k as the input of the FLUBE by assuming the expected value (mean) of the error terms is zero. This will give the opportunity to the FLUBE to generate the E_k^L and E_k^U [13]. Thus, we firstly give a brief overview of LinGRA [13] and then the ElinGRA.

1) LinGRA

In the LinGRA presented in [13] is illustrated in Fig. 6a. Here, we will use the generated uncertainty bounds of the FLUBE to construct TFN. In this context, we will prefer and employ TFNs which are represented with triplet (L, C, U) [19] which can be seen at Fig. 6b. Thus, at each sample k, the TFN parameters will be assigned in an online manner as follows:

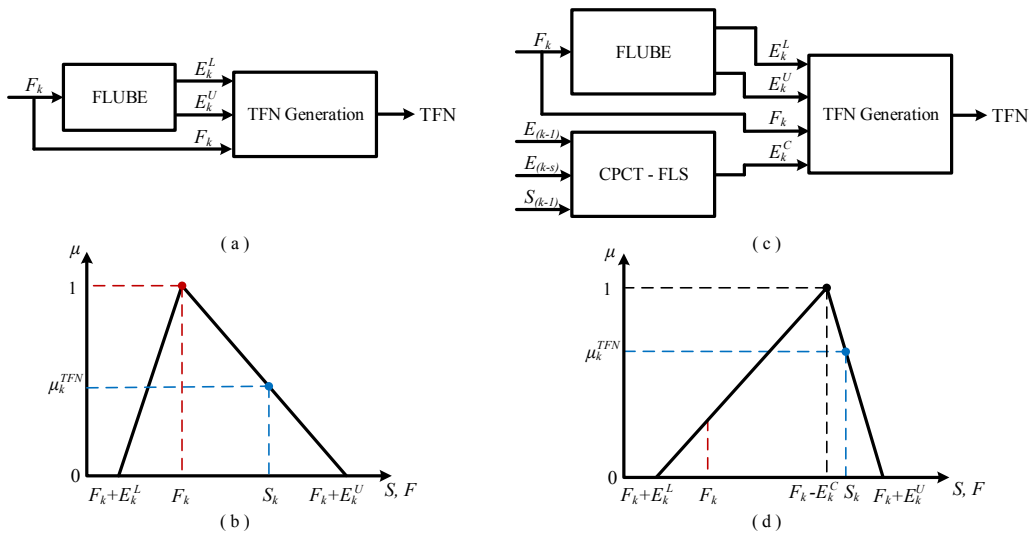$$L_k = F_k + E_k^L, \quad C_k = F_k, \quad U_k = F_k + E_k^U \quad (9)$$

Fig. 6. Illustration of the TFN generation methods and generated TFNs. (a)LinGRA, (b)TFN construction according to LinGRA, (c) ElinGRA, (d) TFN construction according to ElinGRA.

where E_k^L and E_k^U are the lower and upper error forecasts generated from the FLUBE. Thus, the decision maker has the opportunity to evaluate the success of the forecast by using linguistic terms. This can be accomplished by calculating the membership degree of the TFN. Thus, for a crisp target value (S'), the success of the forecast can be defined as follows [13]:

$$\mu_k^{TFN}(S') = \begin{cases} 0, & \text{if } S' \notin [L_k \ U_k] \\ \dfrac{S' - L_k}{|E_k^L|}, & \text{if } S' \in [L_k \ C_k] \\ \dfrac{U_k - S'}{|E_k^U|}, & \text{if } S' \in [C_k \ U_k] \end{cases} \quad (10)$$

The membership degree will represent the success of the forecast that is represented by TFNs. Thus, for a μ value of the TFNs that is close to 1, the success of the forecast will be relatively high.

2) Enhanced LinGRA

As it has been asserted in LinGRA, the support of the TFN will be generated by using the outputs of the FLUBE. Furthermore, the center of the TFN is determined by the forecast value, which is generated by forecast model, since the forecast value is expected to be equal to the Target value at k^{th} sample. However, the error of the forecast is not always equal to zero since forecasting model cannot overcome with uncertainness and nonlinearities in the data to be forecasted. Likewise, this fact is not considered in the LinGRA, so a correcting mechanism for the center of the TFN is needed. To overcome the above mentioned problem, the proposed ElinGRA inherits an extra fuzzy logic system called CPCT-FLS to define the center of the TFN with a correction term (E^C) as shown in Fig. 6c.

The CPCT-FLS is constructed by choosing the inputs to be target data (S_{k-1}) and the error term (E_{k-1}) at the $(k-1)^{th}$ sample and the error term E_{k-s} which is defined with respect to the seasonality of the target data. (E_k^C), which is the output of the CPCT-FLS, will be the extra input to the TFN generation block as shown in Fig. 6c. The fuzzy rule base of the CPCT-FLS is as follows:

$$R_\omega^C : \text{If } S_{k-1} \text{ is } A_\omega^C, E_{k-1} \text{ is } B_\omega^C, E_{k-s} \text{ is } C_\omega^C$$
$$\text{Then } E^C \text{ is } D_\omega^C \quad (11)$$

where A_ω^C, B_ω^C and C_ω^C, are the antecedent MFs, D_ω^C is the consequent singleton and W $(\omega = 1, \dots W)$ is the total number of rules. The CPCT-FLS will be designed via the ANFIS toolbox/MATLAB in an offline manner.

The extra information about the forecast error will be used to evaluate the linguistic term to represent the forthcoming target value. In this context, we will use the output of the Extra FLS E^C to redefine the center of the TFN (C_k) while keeping the L_k and U_k as given in Equation (9) in order to provide an identical PINAW and PICP values of the LinGRA. Thus, the triplet of the TFN will be defined as follows:

$$L_k = F_k + E_k^L, \quad C_k = F_k - E_k^C, \quad U_k = F_k + E_k^U \quad (12)$$

where E_k^C is the error forecast value at the kth sample generated from the extra FLS. Note that, since the CPCT-FLS trained with all error terms distinctively from other FLS ($E_{max} > 0, E_{min} < 0$), the signal of the error forecast terms are different from each other at Equation (12).

In a similar manner, the success of the forecasts will be evaluated with the membership degrees according to Equation (10) as mentioned for LinGRA.

IV. EXPERIMENTAL RESULTS AND DISCUSSIONS

In this section, we will present the experimental results where the LinGRA and ElinGRA methods are evaluated on the electrical consumption data set (Data Set-1) and air passenger data set (Data Set-2). As it has been asserted in previous section, the FLUBE mechanism is the common operator in the LinGRA and ElinGRA. Thus, we will first design and evaluate the performance of the FLUBE, and then we will compare the performances of the proposed ElinGRA with LinGRA.

A. Performance Evaluation of the FLUBE

Here, the FLUBE will be designed and evaluated to define the uncertainties of the single point forecast. As mentioned in the FLUBE design, the bound characteristics are determined via the error terms E_k which are calculated from forecast model. Accordingly, since the handled data sets inherit trend and seasonality characteristics, we firstly have designed a SARIMA $(2,1,2)(15,1,2)_{12}$ and a SARIMA $(0,1,2)(13,1,14)_{12}$ forecast model for the Data Set-1 and Data Set-2, respectively. The performance values of the SARIMA models are tabulated in Table 1. It can be observed that the performances of the SARIMA models are satisfactory since they resulted with relatively low MAPE values while providing high POA values [20]. Consequently, to design the FLUBE, the error terms E_k are calculated via SARIMA forecast model. The error terms (E_k) with respect to the consumption values (S_k) will be used for the error model training. As it has been mentioned in the Remark, there is need to shift the error terms by their mean values which are calculated as -3.04 and -0.26 for the Data Set-1 and Data Set-2, respectively. Then, the SA has been employed for both data sets with the tuning parameter of the SA (P) as 60 to collect the training data sets $[S_{min}, E_{min}]$ and $[S_{max}, E_{max}]$ for LFLS and UFLS, respectively. The LFLS and UFLS are constructed with 5 Gaussian antecedent MFs and 5 linear consequent MFs ($W = 5$) to provide an approximate bounds on the uncertainty (E^L and E^U). The outputs of the FLUBE are given in Fig. 5a and Fig.7 for the Data Set-1 and Data Set-2, respectively.

TABLE I. ACCURACY OF THE FORECASTING MODELS AND PERFORMANCE VALUES OF THE FLUBE

Data Set	Accuracy of the forecasting model		Performance values of the FLUBE	
	MAPE	POA	PICP	PINAW
Data Set-1	1.95%	99.98%	90.52%	4.46%
Data Set-2	3.13%	99.91%	85.48%	6.75%

The FLUBE bounds coverage property is illustrated in Fig. 8 and Fig. 9 for Data Set-1 and Data Set-2, respectively. The forecast coverage performances of the FLUBE for each data set are evaluated with respect to the performance indexes; PICP and PINAW which are also presented in Table 1. For both data sets, it can be concluded that the FLUBE provided a satisfactory performance since they resulted with a relatively small PINAW value and a high PICP value [6].
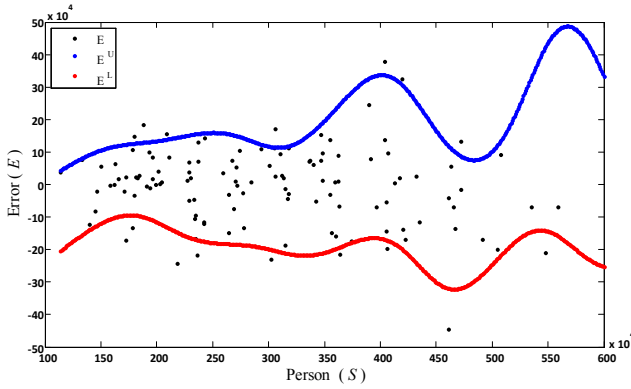
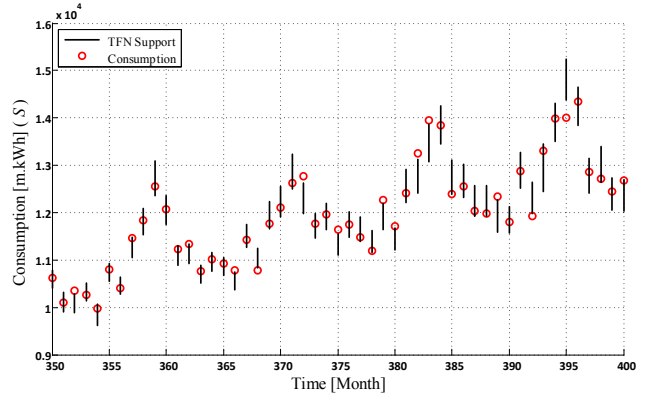Fig. 7.Illustration of the FLUBE bounds on the Data Set-2

Fig. 8. 2-D representation of the method for 50 samples of the Data Set-1.

Fig. 9. 2-D representation of the method for 50 samples of the Data Set-2.

B. Performance Evaluations of the LinGRA and ElinGRA

In this subsection, we will compare the performance of the LinGRA and ElinGRA for the data sets. As it has been asserted in Subsection 3.B, both approaches are using the outputs of the FLUBE to the generate TFN representation. The main difference between these two approaches is the representation of the center point of the TFN. In the LinGRA, the center point will be assigned directly to the forecast value, while in the ElinGRA the center point of the TFN is redefined with the center point correction term (E^C) which generated from the Extra FLS. For the handled data sets, the Extra FLS of the ElinGRA is constructed by defining the antecedent parts of its fuzzy rules with 4 Gaussian antecedent MFs while their consequent part with 4 linear MFs.

In the TFN generation of LinGRA, we will use the outputs of the FLUBE at the kth sample (E_k^L and E_k^U) to define the uncertainty bounds of the forecast value via Equation (9). In a similar manner, are used for ElinGRA with Equation (12) instead of Equation (9) to generate the TFNs. For instance, the uncertainty bounds of the forecast (L_k and U_k) are shown in Fig. 8 and Fig. 9 for of the handled data sets. Now, the TFN can be generated by using Equation (8). Since the TFN will naturally results with an extra dimension (μ), the forecast is illustrated in a 3-D plot as shown in Fig. 10 for Data Set-1 and Fig.11 for Data Set-2 where the generated TFNs for each sample can be clearly seen for both approaches.

It can be observed that at each sample (k) the interval of the forecast uncertainty (i.e. the support of the TFN) varies with respect to the current forecast value and thus nonsymmetrical TFNs will are generated for both the LinGRA and ElinGRA. However, since their center definition is not identical, both approaches resulted with unique TFN representation for the same forecast uncertainty. In order to clearly illustrate this concept, we have presented the generated TFNs of the LinGRA and ElinGRA for certain samples.
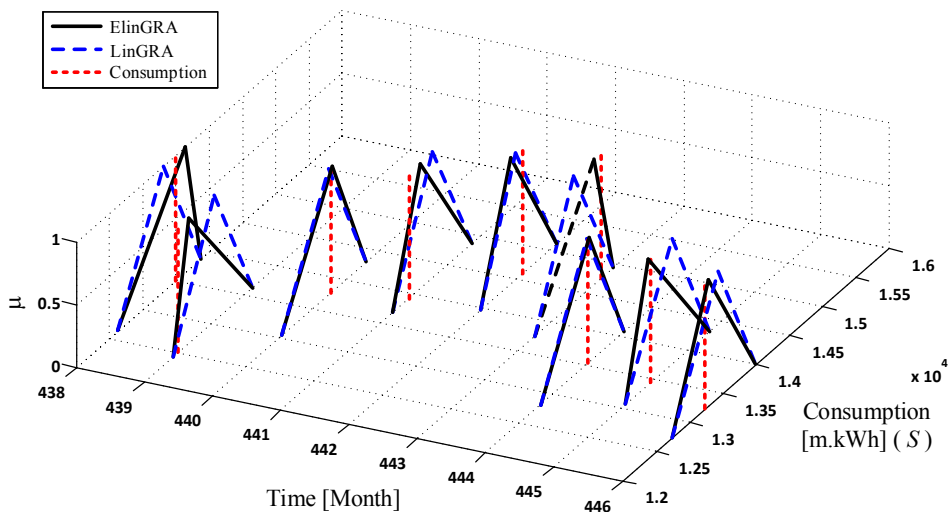

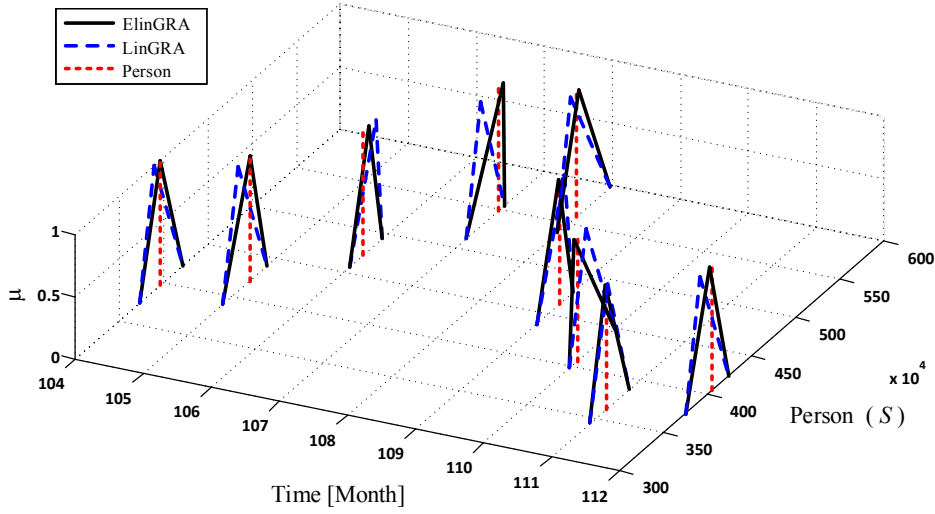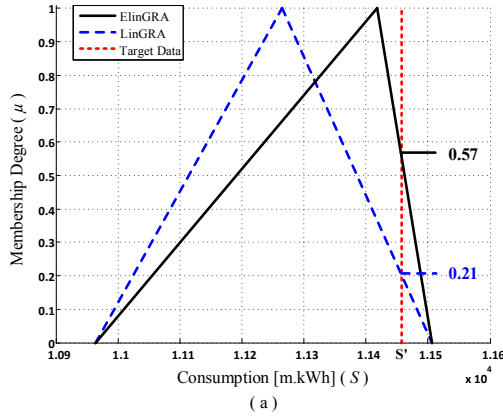Fig. 10. 3-D representation of the LinGRA and ElinGRA for the Data Set-1 (selected 9 samples).
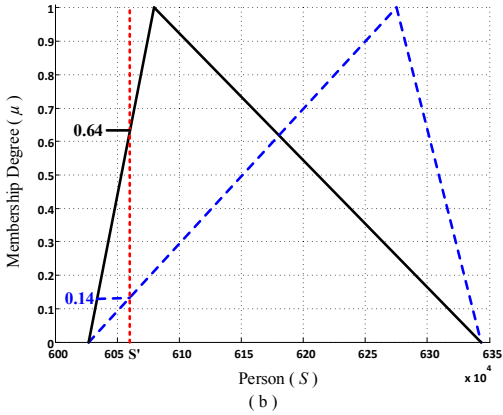
Fig. 11. 3-D representation of the LinGRA and ElinGRA for the Data Set-2 (selected 9 samples).

(a)

(b)

Fig. 12. Enhancement of the membership degree with ElinGRA: (a) 357th sample on the Data set-1 (b) 120th sample on the Data set-2.

The TFNs generated for the 357th sample on the Data Set-1 and the 120th sample on the Data Set-2 are illustrated in Fig.12. For the 357th sample of the Data Set-2, the current forecast and consumption values are $F_k = 11265$ and $S_k = 1145$, respectively. Since the success of the forecast is measured with respect to the μ value, and as mentioned previously a value close to one shows the success of the forecast which is represented with TFN. For the studied samples, the μ values of the TFNs generated from the LinGRA and ElinGRA are calculated as $\mu_{LinGRA} = 0.21$ and $\mu_{ElinGRA} = 0.57$,

respectively. It can be concluded that ElinGRA was able to represent the success of the forecast about 2.8 times better in comparison to the LinGRA. A similar analysis can be also done for the 120th sample of the Data Set-2 given in Fig.12b where the ElinGRA resulted with a 4.5 times better representation of forecast success in comparison with the LinGRA.

TABLE II. SUCCESS OF THE LINGRA AND ELINGRA

Data Set	Mean μ values	
	LinGRA	ElinGRA
Data Set-1	0.56	0.61
Data Set-2	0.52	0.67

Moreover, to analyze the overall forecast quality of the LinGRA and ElinGRA, the distribution of the membership grades (μ) of the studied data sets are illustrated in Fig. 13 and Fig. 14, respectively. It can be clearly observed that the membership grades of both approaches are mostly distributed around the maximum membership grade value 1. However, in comparison to the LinGRA, the count numbers of the μ values of the ElinGRA around 1 are relatively higher. In order to make a fair comparison, we have calculated the mean values of the μ_{LinGRA} and $\mu_{ElinGRA}$ and presented them in Table 2. It can be clearly observed that the ElinGRA approach was able to represent overall forecast quality almost by 9% and 29% better in comparison to LinGRA.
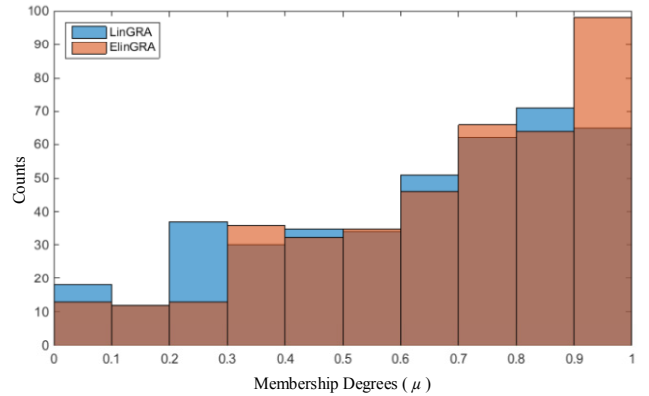

Fig. 13. The histogram of the membership degrees generated from the LinGRA and ElinGRA for the Data Set-1.
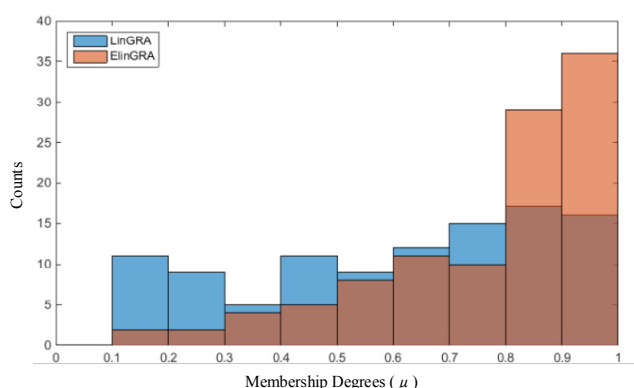
Fig. 14. The histogram of the membership degrees generated from the LinGRA and ElinGRA for the Data Set-2.

V. Conclusions and Future Work

In this study, a novel approach for fuzzy linguistic term generation for time series forecasting via TFNs is proposed. This new approach called ElinGRA is an enhanced version of the previously studied methodology (LinGRA), so it also used FLUBE, which is proposed to estimate the uncertainty in the forecast. Since ElinGRA does not depend on the method chosen for the forecasting, it can easily be applied to and time series forecasting process. Compared to the LinGRA, ElinGRA has a CPCT which benefits to improve the accuracy of the forecasting method, for the TFN representation.

To illustrate the superiority of the proposed approach, experimental results are presented on the Australian monthly electrical consumption and the air passenger data sets. The results of the experiments, which are presented with tables and histograms, show ElinGRA increases the information about the accuracy and success of the single point forecast by providing a relative membership degree (μ) in comparison with the LinGRA. Thus, the linguistic forecast representation will give the opportunity to the decision maker to quantify the uncertainty of the point forecasts with the linguistic terms which might increase the interpretability for further assessments.

Future work will focus on different datasets which inherit higher level of uncertainties and complexities.

References

[1] W. C. Wang, K. W. Chau, C.T. Cheng and L. Qiu, "A comparison of performance of several artificial intelligence methods for forecasting monthly discharge time series," *Journal of Hydrology*, vol. 374 (3), pp. 294-306, 2009.

[2] A. Khosravi and S. Nahavandi, "An interval type-2 fuzzy logic system-based method for prediction interval construction," *Appl. Soft Computing*, vol. 24 (9), pp. 222-231, 2014.

[3] R. J. Hyndman and A. B. Koehler, "Another look at measures of forecast accuracy," *Int. J. of Forecasting*, vol. 22, pp. 679-688, 2006.

[4] T. Heskes, "Practical confidence and prediction intervals," *Adv. Neural Inf. Process. Syst.*, T. P. M. Mozer and M. Jordan, Eds. Cambridge, MA: MIT Press, 1997, vol. 9, pp. 176-182.

[5] A. Khosravi, S. Nahavandi, D. Creighton, and A. F. Atiya, "Comprehensive review of neural network-based prediction intervals and new advances," *IEEE Trans. Neural Netw.*, vol. 22 (9), pp. 1341-1356, 2011.

[6] A. Khosravi, S. Nahavandi, and D. Creighton, "Construction of optimal prediction intervals for load forecasting problems," *IEEE Trans. Power Syst.,* vol. 25 (3), pp. 1496-1503, 2010.

[7] G. Chryssolouris, M. Lee and A. Ramsey, "Confidence interval prediction for neural network models," *IEEE Trans. Neural Netw.*, vol. 7 (1), pp. 229-232, 1996.

[8] J. T. G. Hwang and A. A. Ding, "Prediciton intervals for artificial neural networks," J. Amer. Stat. Assoc., vol. 92 (438), pp. 748-757, 1997.

[9] D. J. C. MacKay, "The evidence framework applied to classification networks," *Neural Comput.,* vol. 4 (5), pp. 720-736, 1992.

[10] R. Dybowski and S. J. Roberts, "Confidence intervals and prediction intervals for feed-forward neural networks,"in *Clinical Applications of Artificial Neural Networks*, R. Dybowski and V. Gant, Eds. Cambridge, U.K.:Cambridge Univ. Press, 2001, pp. 298-326

[11] D. A. Nix and A. S. Weigend, "Estimating the mean and variance of the target probability distribution," in *Proc. IEEE Int. Conf. Neural Netw. World Congr. Comput. Intell.*, vol. 1, Orlando, FL, 1994, pp. 55-60.

[12] A. Khosravi, S. Nahavandi, D. Creighton, and A. F. Atiya, "Lower upper bound estimation method for construction of neural network based prediction intervals," *IEEE Trans. Neural Netw.*, vol. 22 (3), pp. 337-346, 2011.

[13] A. Sahin, T. Kumbasar, E. Yesil, M. F. Dodurka and O. Karasakal, " An approach to represent time series forecasting via fuzzy numbers," in *International Conference on Artificial Intelligence, Modelling and Simulation (AIMS),* Madrid, Spain, 2014, pp. 51-56

[14] G. E. P. Box, G. M. Jenkins and G. C. Reinsel, Time Series Analysis Forecasting and Control, 4th ed., John Wiley & Sons, New Jersey, 2008.

[15] P. J. Brockwell and R. A. Davis, Introduction to Time Series and Forecasting, 2nd ed., Springer-Verlag, New York, 2002.

[16] H. Quan, D. Srinivasan and A. Khosravi, "Short-term load and wind power forecasting using neural network-based prediction intervals," *IEEE Trans. Neural Netw. and Learning Syst.*, vol. 25 (2), pp. 303-3315, 2014.

[17] R. Ak, Y. Li, V. Vitelli, E. Zio, E. L. Droguett and C. M. C. Jacinto, "NSGA-‖-trained neural network approach to the estimation of prediction intervals of scale deposition rate in oil & gas equipment," *Expert Syst. with Applicat.*, vol. 40 (4), pp. 1205-1212, 2013

[18] S. Makridakis, S. C. Wheelwright, and R. J. Hyndman, Forecasting - Methods and Applications, 3rd ed., John Wiley & Sons, New York, 1998.

[19] E. Yesil, M. F. Dodurka and L. Urbas, "Triangular fuzzy number represantation of relations in Fuzzy Cognitive Maps," Proceedings of the FUZZ-IEEE 2014 - *International Conference on Fuzzy Systems*, July 6-11, Beijing, China, 2014.

[20] D. Frechtling, Forecasting Tourism Demand: Methods and Strategies, 1st ed., Butterworth Heinemann, Oxford, 2001.