

ORIGINAL RESEARCH PAPER

An Interval Type-2 Fuzzy LSTM Algorithm for Modeling Environmental Time-Series Prediction

Aref Safari, Rahil Hosseini*

Department of Computer Engineering, Shahr-e-Qods Branch, Islamic Azad University, Tehran, Iran

ARTICLE INFORMATION

Received: 2022.07.12
Revised: 2022.11.03
Accepted: 2022.11.27
Published online: 2022.12.17

DOI: [10.22034/AP.2022.1963124.1133](https://doi.org/10.22034/AP.2022.1963124.1133)

KEYWORDS

Deep learning
Type-2 fuzzy logic
LSTM network
Air Pollution Prediction

ABSTRACT

The statistical attributes of the non-stationary problems such as air quality and other natural phenomena frequently changed. Type-2 fuzzy logic is a robust and capable model to cope with high-order uncertainties associated with non-stationary time-dependent features. This research's main objective is to present a novel Fuzzy Deep LSTM (IT2FLSTM) model to predict air quality for Tehran and Beijing in a short and long time series scale. The proposed model has been evaluated on a real dataset that contains the one-decade information about outdoor pollutants from April 2011 to November 2020 in Tehran and Beijing. The IT2FLSTM model was evaluated using a ROC curve analysis and validated using 10-fold cross-validation. The results confirm the IT2FLSTM model's superiority with an average area under the ROC curve (AUC) of 97 % and a 95% confidence interval of [95-98] %. The proposed IT2FLSTM model promises to predict complex problems to make strategic prevention decisions to save more lives.

How to Site: Safari A., Hosseini R., An Interval Type-2 Fuzzy LSTM Algorithm for Modeling Environmental Time-Series Prediction, Anthropogenic Pollution Journal, Vol 6 (2), 2022: 62-72, DOI: [10.22034/AP.2022.1963124.1133](https://doi.org/10.22034/AP.2022.1963124.1133).

1. Introduction

Nowadays, most megacities suffer from air pollution problems. Besides the effects on human health, some issues occur if the pollutions last for a long time (Marlier et al., 2016; Sekhavati et al., 2021; Khajeh Hoseini et al., 2021; Hoseini et al., 2022). Handling such problems requires intelligent models to address the high-order uncertainty in the air pollution issue's characteristics. There are non-stationary time series features in air

quality patterns whose statistical attributes such as means and variances change over time.

This research aims at modeling uncertainty sources associated with non-stationary time series features in real-world applications such as air pollution (Sekhavati et al., 2022). Also, the main innovation of this study is to present an Interval Type-2 Fuzzy LSTM Algorithm for Modeling Environmental Time-Series Prediction for a real-world and global challenge such as air pollution

Corresponding author:



This work is licensed under the Creative Commons Attribution 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

prediction. In this work, the proposed model aimed at overcoming the drawbacks of the existing methods and presenting a robust, reliable, and accurate model for prediction of AQI time-series data through type-2 fuzzy logic and deep learning models. On the other hand, precise and reliable prediction can help reasonable strategies and assist the specialists in planning the best policies to model an event in uncertain circumstances.

1.1. Related Works

During the recent decade, applications of type-2 fuzzy logic systems in problems with high-order uncertainty have been grown, especially for prediction problems with dynamic and non-stationary problems. Many types of research (Georgescu, 2019; Silva et al., 2020) have been reported for modeling uncertainty using fuzzy logic. Also, there have been several practical, real-world applications for type-2 fuzzy systems, mostly modeling uncertainty, control, and predictions (Gaxiola et al., 2019; Almanza et al., 2020). However, fuzzy systems suffer from a lack of learning and adaption mechanisms for their membership function parameters and fuzzy rule. Various intelligent models such as neuro-fuzzy and recurrent neural networks (RNNs), have been applied to model and predict the time series data in uncertain environments (Safari et al., 2017; Ibarra-Berastegi et al., 2008; Chen et al., 2020; Zhou et al., 2020; Biancofiore et al., 2017; Anh et al., 2019).

It has been proven that the RNNs, especially the LSTM networks, have an excellent capability for solving complex problems, where the model must learn the multi-layered inter-relationships between two-time series (Zhou et al., 2020; Lin et al., 2013). An LSTM cell's benefit compared to a regular recurrent unit is its cell memory (Wang et al., 2019). The cell vector of an LSTM has the capability of the information earlier stored memory and the part of its new information (Liu et al., 2020). These capabilities can be used in the prediction of non-stationary problems (Smagulova et al., 2019). In addition to the general benefits of using RNNs for time series prediction, the LSTM network can also automatically learn the data's temporal dependencies (Guo et al., 2017). On the other hand, the LSTM is a non-linear prediction method to learn the arbitrary complex mapping from inputs to outputs, the LSTM networks can model both short and long time-series predictions. However, LSTM model cannot handle the uncertainty associated with non-stationary features. This study takes advantage of type-2 fuzzy logic to develop a new architecture for handling uncertainty through a novel deep interval type-2 fuzzy LSTM (IT2FLSTM) model to obtain a reliable result in time-series prediction. This research aimed to propose an intelligent method to predict the Air Quality Index (AQI) in Tehran and Beijing to produce a reliable prediction for different pollutants described in Table-1 in Environmental Protection Agency (EPA) format.

Table 1. AQI levels by EPA

No	AQI Range	Level
1	0-50	Healthy
2	51-100	Moderate
3	101-150	Unhealthy for sensitive groups
4	151-200	Unhealthy
5	201-300	Very Unhealthy
6	301-500	Hazardous

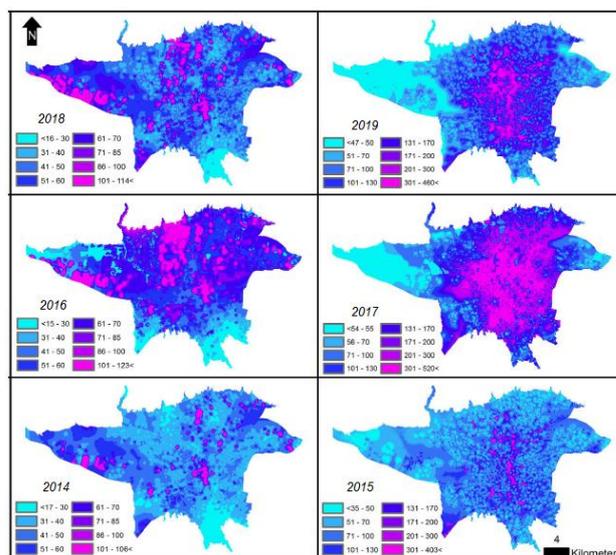


Figure 1. An architecture of the LSTM cell (Liu et al., 2020)

2. Research Background

This section presents a brief overview of the LSTM network. It follows a review of interval type-2 fuzzy sets (IT2FS) concepts and their mathematic definitions.

2.1. A Review of LSTM network

The LSTMs have been developed to address classic RNNs' limitations by enhancing the network structure's gradient vanishing. The use of a cell state attains this c_t , which stores long-term information as follows (Liu et al., 2020):

Input gate:

$$i_t = \sigma(W_{i1}z_{t-1} + W_{i2}h_t + W_{i3}x_t + W_{i4}s + b_i) \quad (1)$$

Output gate:

$$o_t = \sigma(W_{o1}z_{t-1} + W_{o2}h_t + W_{o3}x_t + W_{o4}s + b_o) \quad (2)$$

Forget Gate:

$$f_t = \sigma(W_{f1}z_{t-1} + W_{f2}h_t + W_{f3}x_t + W_{f4}s + b_f) \quad (3)$$

Where z_{t-1} is the hidden state of the LSTM at time $t-1$ and W is the weight matrices. As well as, t index is the time step, x is the input, and h is the output variables at time t , and σ is the sigmoid activation function.

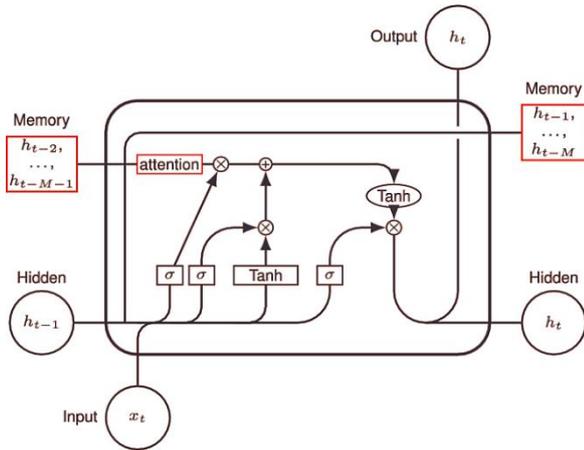


Figure 2. An architecture of the LSTM cell (Liu et al., 2020)

The gates adjust the states and hidden cell of the LSTM using the following equations:

Hidden state:

$$z_t = o_t \odot \text{Tanh}(c_t) \tag{4}$$

Cell state:

$$c_t = f_t \odot c_{t-1} + i_t \odot \text{tanh}(W_{c1}z_{t-1} + \dots + W_{c4}s + b_c) \tag{5}$$

Where \odot is the element-wise product, and tanh is the activation function.

2.2. Interval Type-2 Fuzzy Sets

Type-2 fuzzy sets are known as fuzzy-fuzzy sets [43]. Membership function (MF) of a type-2 fuzzy set (T2FS) of a given element is itself a type-1 fuzzy set (T1FS) (Wu, 2012). A T2FS \tilde{A} , is characterized through a type-2 MF $\mu_{\tilde{A}}(x,u)$ where $x \in X$ and $u \in J_x \subseteq [0,1]$ as follows:

$$\tilde{A} = \{ \{(x,u), \mu_{\tilde{A}}(x,u)\} \mid \forall x \in X, \forall u \in J_x \subseteq [0,1] \} \tag{6}$$

Where, $0 \leq \mu_{\tilde{A}}(x,u) \leq 1$, X is the domain of fuzzy set and J_x is the domain of the secondary MF at x . \tilde{A} is as:

$$\tilde{A} = \frac{\int_{x \in X} \int_{u \in J_x} \mu_{\tilde{A}}(x,u)}{x, u J_x} \subseteq [0,1] \tag{7}$$

Where \int represents union over all admissible x and u and,

$$FOU(\tilde{A}) = \bigcup_{x \in X} J_x \tag{8}$$

where x is the primary variable, J_x , an interval in $[0,1]$, is the primary MF of x , u is the secondary variable, and $\int_{u \in J_x}$ is the secondary MF at x . Uncertainty about \tilde{A} is addressed by the union of all of the primary memberships, called the footprint of uncertainty (FOU) of \tilde{A} , i.e., $[FOU\tilde{A}]$, (Mendel et al, 2006) as:

$$FOU(\tilde{A}) = \bigcup_{x \in X} J_x \tag{9}$$

The FOU for a Gaussian primary MF with uncertain standard deviation is shown in Figure.3. The FOU is bounded by upper bound membership functions (UMF) $\bar{\mu}_{\tilde{A}}(x)$ and lower bound membership function (LMF) $\underline{\mu}_{\tilde{A}}(x)$, which are type-1 fuzzy sets; consequently, the membership grade of each element of an IT2FS is identified by an interval $[\underline{\mu}_{\tilde{A}}(x), \bar{\mu}_{\tilde{A}}(x)]$. In the IT2FLS, the UMF and LMF can better represent input variables' uncertainties than type-1 fuzzy sets. Similarly, the FOU in the IT2FLS provides more degrees of freedom when designing a fuzzy system (Mendel and John, 2006).

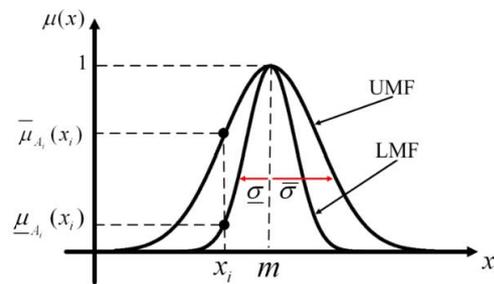


Figure 3. Gaussian primary MF (Mendel, 2020)

An interval type-2 fuzzy system (IT2FS) architecture contains four components; fuzzifier, fuzzy rule base, inference engine, type-reducer, and defuzzifier (Mendel, 2017). In general, the type-reducer is required in the output processing block before the defuzzification in the IT2FLS (Sumati and Patvardhan, 2018). The fuzzifier can be categorized into two types, singleton and non-singleton, according to the number of non-zero MF values and defines the membership grade of input. In this paper, the TSK fuzzy rule type is considered more precise than Mamdani rules. The product and minimum t-norms were used in inference methods (Chen and Zou, 2020). In this work, the singleton fuzzifier was implemented. The output processing, including the type-reducer and defuzzifier, generates the crisp output. For this step, the Karnik-Mendel (KM) algorithm (Mendel, 2013) was applied in this paper. The next section presents the IT2FLSTM model.

3. The Proposed IT2FLSTM Model

This section presents the details of the architecture IT2FLSTM. Then, in 3.2, the mathematical model of the IT2FLSTM is presented. Finally, the cell structure of the proposed IT2FLSTM has been discussed in 3.3

3.1 The Architecture of the IT2FLSTM

As shown in this architecture, the IT2FLSTM is fuzzified by the IT2F sets. The framework of the proposed IT2FLSTM includes five layers, i.e., the input layer, encoder layer, hidden layers, the decoder layer, and the output layer. The details of each layer are illustrated in Figure. 4.

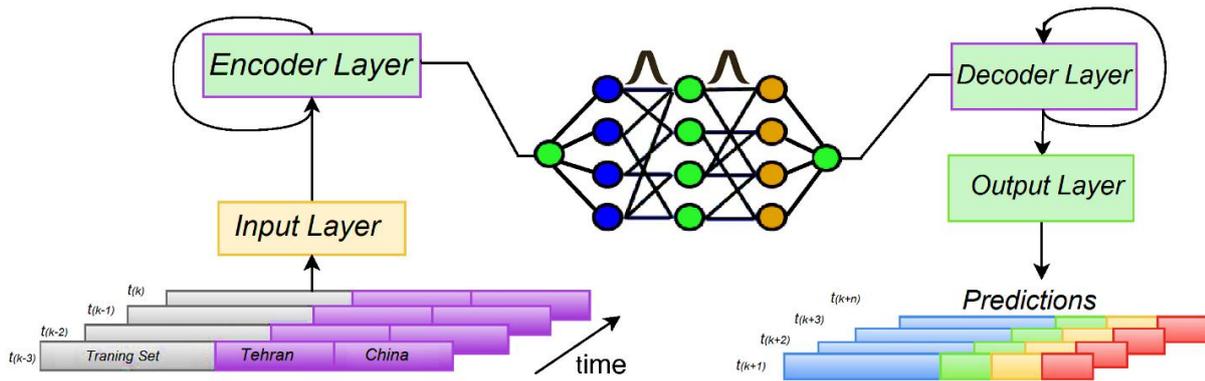


Figure 4. The architecture of the proposed model

Input Layer: The first layer is called the input layer, which directly handles the original data. The output of this layer feeds the inputs of the deep layers.

Encoder, Hidden, and Decoder layers: The main idea is to map the entire input sequence to a vector and then use an encoder to generate the output sequence. In this layer, the encoder represents the whole input sequence in the hidden layer activities. The proposed structure can reform according to the dimension of the training time-series data set.

Output Layer: The final layer is considered as the prediction layer to make the decision and perform the prediction based on the input features received from the previously hidden layers.

3.2.1. Mathematical Model of the Proposed IT2FLSTM

Time series prediction models specify future values of a target y_{it} for a given entity i at time t . The next step-ahead prediction is as follows:

$$\hat{y}_{i,t+1} = P_f(y_{i,t-e:t}, x_{i,t-e:t}, t) \tag{10}$$

Where $\hat{y}_{i,t+1}$ is the next step predicted value, t is the time step at T , $y_{i,t-e:t}$ and $x_{i,t-e:t}$ are observations of the target and observed inputs, respectively, over a look-back window e , and P_f is the prediction function.

And the final prediction is produced by Z_t :

$$P_f(y_{i,t-e:t}, x_{i,t-e:t}, t) = g_{dec}(z_t) \tag{11}$$

$$Z_t = g_{enc}(y_{i,t-k:e}, x_{i,t-e:t}, t) \tag{12}$$

Where g_{enc} and g_{dec} denote the encoder and decoder functions, respectively. The mathematical model of different parts of the LSTM network in the proposed IT2FLSTM architecture are given as follows:

Definition 1: Let N be the number of memory units of the model. In time-step t , i.e., the current time, the network keeps in memory a set of vectors by the following equations:

$$i_t = \sigma(W_{ix}x_t + W_{im}m_{t-1} + W_{ic}c_{t-1} + b_i) \tag{13}$$

$$f_t = \sigma(W_{fx}x_t + W_{fm}m_{t-1} + W_{fc}c_{t-1} + b_f) \tag{14}$$

$$c_t = f_t \odot c_{t-1} + i_t \odot g(W_{cx}x_t + W_{cm}m_{t-1}) \tag{15}$$

$$o_t = \sigma(W_{ox}x_t + W_{om}m_{t-1} + W_{oc}c_t + b_o) \tag{16}$$

$$m_t = o_t \odot h(c_t) \tag{17}$$

$$f_\phi = (W_{hm}m_t + b_y) \tag{18}$$

Where σ is the sigmoid function, W is the weight matrices, W_{ix} is a matrix of fuzzy weights from the input cell to the output gate. b is the bias vector, and i, f, o , and c are the input gates, the hidden (forget) gate, the output gate, and cell activation function, respectively. The cell output is represented by o_t , and \odot is the element-wise product in equation (15). f_ϕ is the activation function of the network. The hidden unit in this architecture is represented in memory blocks. Each block contains one or a large number of memory cells. This procedure is a way for these cells to preserve information for a specific time in an uncertain time-series and decide which piece of information should be stored and when to use it.

3.2. The Mathematical Model of the IT2FLSTM

The input variables of the IT2FLSTM model can be defined as p where p is the inputs of the proposed IT2FLSTM model as follows:

$$P \left\{ \begin{array}{l} R_p^{j_1, j_2}: \text{ if } p_t = \tilde{F}_p^{j_1} \text{ and } p_{t+1} = \tilde{F}_p^{j_2} \\ \text{ then } y_p = [c_p^{j_1, j_2}, \tilde{c}_p^{j_1, j_2}] \end{array} \right\} \tag{19}$$

where $\tilde{F}_p^{j_1}$ and $\tilde{F}_p^{j_2}$ are the interval type-2 fuzzy sets of the inputs x_p and x_{p+1} in IT2FLSTM, respectively, M is the number of applied fuzzy rules, $c_i^{j_1, j_2}$ and $\tilde{c}_i^{j_1, j_2}$ are the consequents of a fuzzy rule. The product t -norm operator in equation (20) is applied to compute the

membership interval as follows:

$$\begin{cases} \underline{f}^l(p') = \mu_{p_1^l}(p_1^l) \times \dots \times \mu_{p_x^l}(p_x^l) \\ \overline{f}^l(p') = \bar{\mu}_{p_1^l}(p_1^l) \times \dots \times \bar{\mu}_{p_x^l}(p_x^l) \end{cases} \quad (20)$$

The inference part of the IT2FLSTM model can be entirely characterized by M fuzzy rules employed in the inference process as:

$$\tilde{R}^l: \text{if } x_1 \text{ is } \tilde{F}^l \dots \text{and } x_p \text{ is } F_p^l \text{ Then } y \text{ is } \tilde{G}^l \quad (21)$$

Where ($l=1, \dots, M$)

The relation for each fuzzy rule is as:

$$\tilde{R}^l: \tilde{F}^l \times \dots \times F_p^l \rightarrow \tilde{G}^l = \tilde{A}^l \rightarrow \tilde{G}^l \quad (22)$$

The membership function of the rule is as,

$$\mu_{\tilde{R}^l}(x, y) = \mu_{\tilde{F}^l}(x_1) \cap \dots \cap \mu_{F_p^l}(x_p) \cap \mu_{\tilde{G}^l}(y) \quad (23)$$

Where \cap signifies the product t -norm operation [43]. The output of each inference procedure is, $\tilde{B}^l = \tilde{A}_x \circ \tilde{R}^l$ with membership functions of $\mu_{\tilde{B}^l}(y)$ as:

$$\mu_{\tilde{B}^l}(y) = \bigcup_{x \in X} [\mu_{A_x}(x) \cap \mu_{\tilde{A} \rightarrow \tilde{G}}(x, y)] \quad (24)$$

Where \circ defines the composition operation and \bigcup represents the maximum t -conorm operation [43], and \tilde{F}^l is the membership interval for the fuzzy rule, where $x=x'$ and F^l is as

$$F^l(x') = [\underline{f}^l(x'), \overline{f}^l(x')] \quad (25)$$

The firing output set $B^{\wedge l}$ is produced through a fuzzy rule and the aggregation of the consequent of the IT2FLSTM model as:

$$\tilde{B}^l: \begin{cases} FOU(\tilde{B}^l) = [\underline{\mu}_{\tilde{B}^l}(y | x'), \overline{\mu}_{\tilde{B}^l}(y | x')] \\ \underline{\mu}_{\tilde{B}^l}(y | x') = \underline{f}^l(x') * \underline{\mu}_{\tilde{G}^l}(y) \\ \overline{\mu}_{\tilde{B}^l}(y | x') = \overline{f}^l(x') * \overline{\mu}_{\tilde{G}^l}(y) \end{cases} \quad (26)$$

Where $*$ represents the product t -norm operation.

The final output $B^{\vee l}$ is considered as the integration of all rule firing sets $B^{\wedge l}$ on the output:

$$\tilde{B}^l: \begin{cases} FOU(\tilde{B}) = [\underline{\mu}_{\tilde{B}}(y | x'), \overline{\mu}_{\tilde{B}}(y | x')] \\ \underline{\mu}_{\tilde{B}}(y | x') = \underline{\mu}_{\tilde{B}^1}(y | x') \vee \dots \vee \underline{\mu}_{\tilde{B}^M}(y | x') \\ \overline{\mu}_{\tilde{B}}(y | x') = \overline{\mu}_{\tilde{B}^1}(y | x') \vee \dots \vee \overline{\mu}_{\tilde{B}^M}(y | x') \end{cases} \quad (27)$$

Where \vee is the max operation.

Then the type reduced set $Y_C(x')$ is computed using the centroid $C_{\tilde{B}}$ of \tilde{B} :

$$Y_C(x') = C_{\tilde{B}}(x') = \frac{1}{[l_{\tilde{B}}(x'), r_{\tilde{B}}(x')]} \quad (28)$$

Where the two points $l_b(x')$ and $r_b(x')$ are computed through the KM algorithm (Mendel, 2013).

3.3 Measuring Uncertainty in Time-Series

Uncertainty in a model affects the confidence of a prediction model and its accuracy. This paper presents a model to manage the uncertainty associated with time-series prediction considering a distribution over the IT2FLSTM predicted sample data points. In this model, ω is the weights of data-points from a short-term to a long-term time-series. This distribution depends on the data points as $D=\{X, Y\}$, where D is the distribution of time steps and X, Y are the samples on 2-D measurements in whole distribution D , where

$$X=\{x_1, x_2, \dots, x_n\} \text{ and,}$$

$Y=\{y_1, y_2, \dots, y_n\}$, respectively. Therefore, the weight distribution after predicting the time-series can be written as $p(\omega|X, Y)$. To approximate this distribution, a Monte-Carlo based approach collects weights by using the Bernoulli rat as computed follows:

$$p(\omega|X, Y) = Bern(\omega; \alpha) \quad (29)$$

Where α is the Bernoulli rate on the weights. Hence, the model uncertainty is the variance of T Monte-Carlo samples (Data points) as follows (Loquercio, 2020):

$$Var_p^{model}(Y|X)(y) = \frac{1}{T} \sum_{t=1}^T (y_t - \bar{y})^2 \quad (30)$$

Where $\{y_t\}_{t=1}^T$ is a set of T sampled outputs of the IT2FLSTM model for weights instances given by:

$$\bar{y} = \frac{1}{T} \sum_t y_t \quad (31)$$

4. Performance Evaluation and Experimental Results

In this section, the evaluations of the proposed model have been presented. Firstly, the datasets and metrics for performance measurements used in this study have been explained. Then the statistical results, comparative study and experimental results have been discussed.

4.1 Applied Data Set on IT2FLSTM model

This study applied well-known existing time-series datasets, including the latest public outdoor pollutant data, including NO2, SO3, PM10, PM2.5, and CO2. This dataset contains the one-decade information about outdoor pollutants from April 2011 to November 2020. The dataset is available on the following link:

<https://data.world/datasets/air-pollution>.

4.2 An ROC Curve Analysis

An ROC curve analysis is conducted to have a reliable estimate of the IT2FLSTM performance. The results were statistically verified. The following equations were used for assessing the performance through an ROC curve analysis of the proposed model. Also, the standard metrics, such as precision, recall, and the F-measure, were applied to evaluate the proposed model as follows:

$$\text{Precision} = \frac{TP}{(TP + FP)} \times 100\% \tag{32}$$

$$\text{Recall} = \frac{TP}{(TP + FN)} \times 100\% \tag{33}$$

$$F - \text{measure} = \frac{2 \text{ Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \tag{34}$$

$$\text{Accuracy} = \frac{(TP + TN)}{(TP + FP + FN + TN)} \tag{35}$$

$$\mu_i = \frac{1}{10} \sum_{k=1}^{10} AUC_j \tag{36}$$

Where μ (i) is the means of the ROC curve's accuracy for the 10-fold cross-validation. The cross-validation technique was applied to the Tehran and Beijing AQI dataset to validate the results. The results have been reported in Table 2.

Table 2. 10-fold cross-validation results

Fold #	Tehran	Beijing
1	90.89	92.09
2	91.25	91.85
3	93.07	93.07
4	94.00	94.01
5	91.98	91.98
6	93.81	93.81
7	91.09	92.09
8	90.45	92.45
9	91.24	91.24
10	93.21	93.21
Mean	92.09	92.58

Table 3 presents the proposed model's accuracy based on different datasets with different cells and depths of layers to compare the proposed IT2FLSTM model's training process. Table 3 summarizes the information by selecting and reporting the main features, including the number of cells, depth, and accuracy during the training process. It shows different layers and cells configuration. The results reveal that the IT2FLSTM with 20 Layers and 3000 cells is the most robust configuration, which reported the best performance during the prediction process as depicted in Figure 5.

Table 3. Training accuracy of the IT2FLSTM model applied to AQI features

Cell	Depth	CO	SO ₂	NO ₃	PM	Mean
10000	5	90	93	91	95	92
7000	10	91	93	92	95	93
5000	15	93	94	92	96	94
3000	20	95	96	95	97	96

4.3. Statistical Evaluation

To represent the proficiency of the IT2FLSTM method and its robustness, a two-sample t-test (left tailed) was

applied. The null hypothesis is defined as $H_0 = \mu_i > \mu_j$ and $H_1: \mu_i < \mu_j$, where μ_i and μ_j are the means of the area under the ROC curve (AUC) of IT2FLSTM and IT2FLS for ten different runs of the cross-validation technique (Equation 39), respectively. The t-test results in Table 4 reveals the superiority of the proposed IT2FLSTM model for Air quality time-series prediction, compared to IT2FLS. The t-test (according to the defined hypothesis testing) failed to reject the defined null hypothesis.

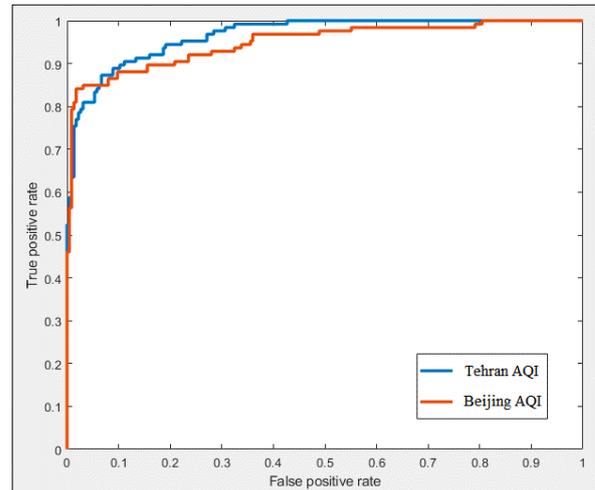


Figure 5. ROC curve analysis of the IT2FLSTM model

Table 4. T-test analysis of the DIT2FLSM and IT2FLS

Fold#	IT2FLSTM	IT2FLS
1	0.9179	0.7724
2	0.9022	0.7111
3	0.9122	0.6351
4	0.8919	0.6221
5	0.9324	0.7128
6	0.9319	0.7815
7	0.965	0.7541
8	0.9446	0.7401
9	0.9669	0.8001
10	0.9871	0.8111
Mean	0.9321	0.73513

5. Discussion

Experimental results reveal that the proposed IT2FLSTM model has reported reliable results based on the real Air quality datasets obtained from the official Harvard Dataset. Moreover, the ROC curve analysis shows the results obtained by the proposed model with different configurations, including different numbers of layers and cells, indicates the performance of 97% (AUC) for the IT2FLSTM model for Air quality prediction. The proposed IT2FLSTM model has reported the highest performance than the standard LSTM, IT1FLS, IT2FLS, T1FLSTM, and IT2ANFIS for AQI time-series prediction. The proposed IT2FLSTM model can be applied to the uncertain and chaotic time-series in short and long-series.

The obtained results with the time-series in Air quality have been reported in Table 5. The proposed IT2FLSTM model is proficient in coping with the uncertainty to model phenomena with long-term dependencies and high-order uncertainties such as AQI prediction.

Figure 6 represents the IT2FLSTM's one day in ahead prediction results on the AQI testing samples for Tehran in (a) and Beijing in (b). The prediction results also show the robustness of the proposed IT2FLSTM model in the short-time series from 2 to 20 Nov 2020. Figure.7 illustrates the seasonal comparison of the original observations with the IT2FLSTM model predictions for Tehran and Beijing from Spring 2019 to Winter 2019. The bar chart of the actual observations in Figure 8 shows the prediction values of the IT2FLSTM model. The results confirm that the proposed model outputs are close to real indexes for Tehran and Beijing AQI prediction in short and long time-series.

The Monte-Carlo uncertainty measurement with multi-step sampling has been applied to the AQI prediction for Beijing and Tehran's time-series. The Monte-Carlo measurement technique requires multiple forward passes of each time-series data to estimate the associate uncertainty of the IT2FLSTM model. A Monte-Carlo

method was applied to provide a quantitative analysis of the associated uncertainty of the proposed model. To show the proposed model's proficiency for time-series prediction, uncertainty measure, in Equations (29 to 31), was applied to AQI prediction. The results have been presented in Table 5.

5.1 Comparison Analysis

The proposed IT2FLSTM model has more design degrees of freedom than a type-1 fuzzy, RNN and LSTM methods or other related models because of the FOU parameters in type-2 fuzzy sets and its potential to model inequality of time intervals. In this work, the problem of distortion in the long time series can be addressed by LSTM algorithm alongside a Type-2 fuzzy logic approach for modeling high-order uncertainties in time-series prediction using footprint of uncertainty and unequal length of the time intervals. According to obtained results in ROC curve analysis, the proposed IT2FLSTM model is 8% better than the RNN Machine learning method (Liu et al., 2019), 4 % better than the earlier work of authors: T2Fuzzy-ANN (Safari et al., 2017) and 6% better than Fuzzy Time Series model in (Lima et al., 2020) in average both scenarios and terms of the AUC.

Table 5. Quantitative comparisons of IT2FLSTM uncertainty measure

Range	Sample Points (T)	Tehran			Beijing		
		RMSE	STD	Mean	RMSE	STD	Mean
Short	100	0.0890	0.0398	0.053	0.0822	0.0364	0.073
Long	1000	0.0701	0.0360	0.069	0.0779	0.0391	0.063
Long	5000	0.0430	0.0270	0.039	0.0390	0.0248	0.047
Long	10000	0.0206	0.0124	0.024	0.0123	0.0239	0.029

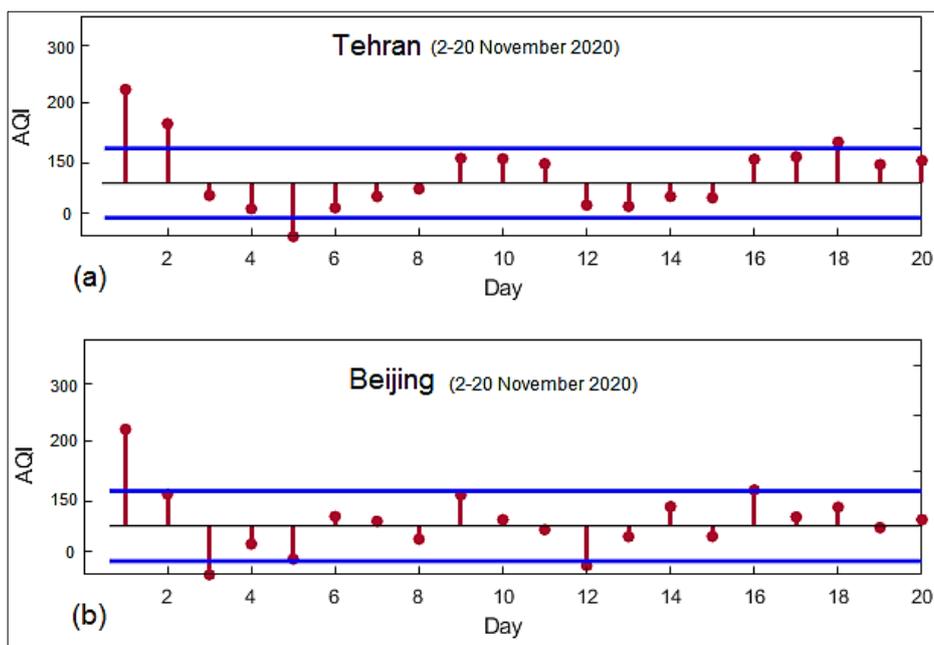


Figure 6. Daily AQI prediction for Tehran (a) and (b) Beijing (from 2 to 20 Nov 2020)

Similarly, the experiment results confirmed that the proposed IT2FLSTM model has lower error rates than current and counterpart methods including T1FLS, IT2FLS and the LSTM models. Also, the experimental results

confirmed the superiority of the proposed IT2FLSTM model in terms of the RMSE, MAE, and MPE, according to the Table 5 and the obtained results. The further are reported in Tables 6 and 7.

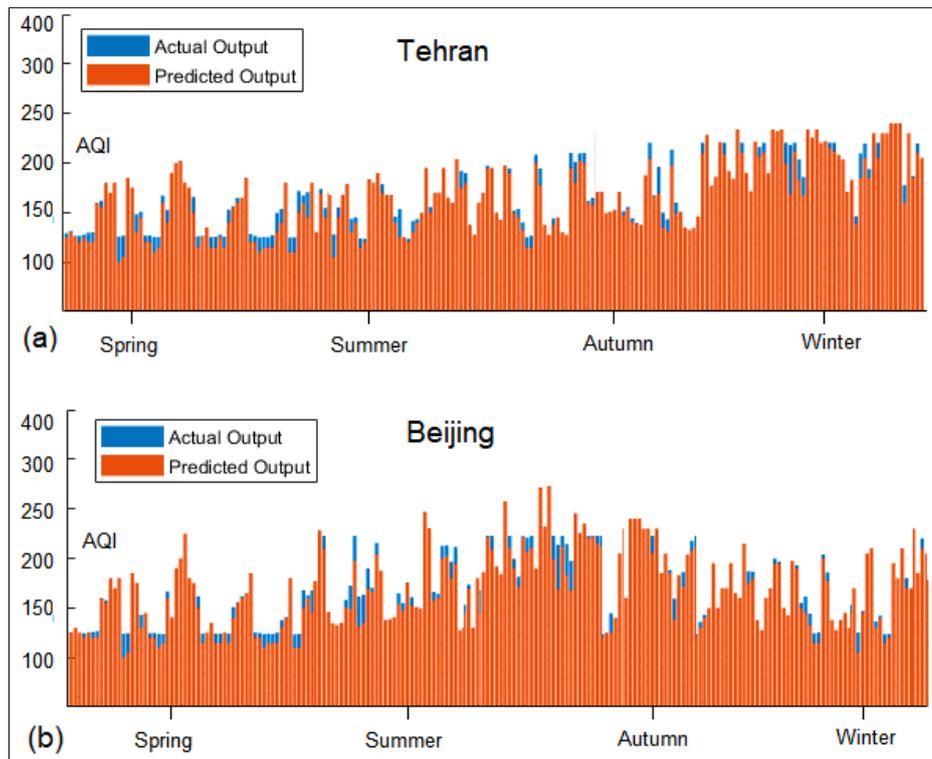


Figure 7. The IT2FLSTM in seasonal time-series prediction of Air quality pollutants for (a) Tehran and (b) Beijing

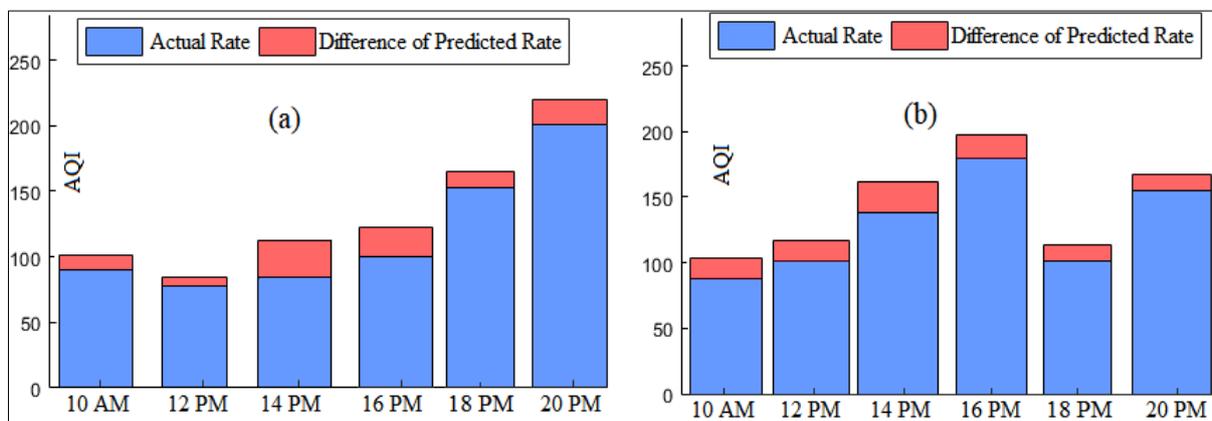


Figure 8. The IT2FLSTM applied to Hourly AQI prediction for Tehran (a) and Beijing (b) (20 Nov 2020)

Table 6. Performance comparison for IT2FLSTM model applied to Tehran AQI time-series

Method	TEHRAN				
	AUC%	95% CI	Recall	Precision	F-Measure
T1FLS	76	[73-78]	87%	89%	87%
IT2FLS	79	[74-80]	89%	90%	89%
LSTM	86	[84-87]	91%	92%	90%
RNN [19]	88	[84-89]	88%	87%	89%
Fuzzy Time-Series [22]	91	[99-91]	90%	91%	90%
T2-ANFIS [26]	93	[90-93]	92%	92%	91%
IT2FLSTM (this work)	97	[95-98]	95%	98%	96%

Table 7. Performance comparison for IT2FLSTM model applied to Beijing AQI time-series

Method	BEIJING				
	AUC%	95% CI	Recall	Precision	F-Measure
T1FLS	75	[74-78]	85%	88%	87%
IT2FLS	79	[77-81]	87%	89%	88%
LSTM	85	[83-86]	90%	91%	90%
RNN [19]	89	[84-90]	89%	88%	89%
Fuzzy Time-Series [22]	90	[88-93]	91%	90%	90%
T2-ANFIS [26]	93	[89-93]	93%	92%	91%
IT2FLSTM (this work)	97	[94-98]	96%	98%	96%

6. Conclusion

In this study, a deep interval type-2 fuzzy learning (IT2FLSTM) model was proposed to predict the air pollutant indexes in Tehran and Beijing. The experimental results reveal that the performance of the proposed IT2FLSTM model is better than its counterparts. The results show that the proposed IT2FLSTM model is 21% greater than T1FLS, 18% greater than IT2FLS, 11% greater than LSTM, in terms of AUC. The proposed IT2FLSTM model has an average AUC of 97% with a 95% confidence interval [95-98]%. Additionally, the model can easily get updated because of its deep architecture when new cases are reported.

6.1. Future Research

Computational intelligence methods, such as fuzzy logic and deep neural networks such as RNN and LSTM, are robust models to solve real-world problems. Nevertheless, the deep learning methods are incapable of modelling long-term dependencies in temporal data, and its learning using gradient descent is a complex and difficult task in order to obtain a reliable prediction in uncertain time series circumstances for a long-term dependencies' scenarios. For the future works and extension of this research, the authors suggest a framework for the dependent time series in a long-term prediction, especially in the real-world applications such as climate changes and global pollutions based on novel autoencoder and fuzzy-stacked deep learning methods. Also, a tuning procedure on the IT2FLSTM cell structure parameters based on optimization algorithms can be used to improve the performance of the current proposed IT2FLSTM model.

Disclosure statement

The authors reported no potential conflict of interest.

References

- Loquercio, M. Segu, and D. Scaramuzza. (2020). A General Framework for Uncertainty Estimation in Deep Learning, *IEEE robotics and automation letters*.
- Safari, R. Hosseini and M. Mazinani, "A Type-2 Fuzzy Time Series Model for Pattern Similarity Analysis: A Case Study on Air Quality Forecasting," in *IEEE Intelligent Systems*, (2021).
- Safari, R. Hosseini, M. Mazinani, "A novel deep interval type-2 fuzzy LSTM (DIT2FLSTM) model applied to COVID-19 pandemic time-series prediction", *Journal of Biomedical Informatics*, Vol 123, (2021).
- Biancofiore, F., Busilacchio, M., Verdecchia, M., Tomassetti, B., Aru o, E., Bianco, S., Di Tommaso, S., Colangeli, C., Rosatelli, G., Di Carlo, P. (2017). Recursive neural network model for analysis and forecast of PM10 and PM2.5. *Atmos. Pollut. Res.* 8, 652–659.
- Li, C. Zhou, W. Peng, Y. Lv, and X. Luo. (2020). Accurate prediction of short-term photovoltaic power generation via a novel double input-rule-modules stacked deep fuzzy method, *Energy*, 11(7) 118700. <https://doi.org/10.1016/j.energy.2020.118700>
- Confalonieri, U., et al. (2007) Human Health. In: Parry, M.L., Canziani, O.F., Palutikof, J.P., van der Linden, P.J. and Hanson, C.E., Eds., *Climate Change 2007: Impacts, Adaptation and Vulnerability, Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge, 391-431.
- Wu and J. M. Mendel. (2019). Recommendations on designing practical interval type-2 fuzzy systems, *Engineering Applications of Artificial Intelligence*, 85 182–193. <https://doi.org/10.1016/j.engappai.2019.06.012>
- Wu. (2013). approaches for reducing the computational cost of interval type-2 fuzzy logic systems Overview and comparisons, *IEEE Transactions on Fuzzy Systems*, 21(1) 80–99. doi: 10.1109/TFUZZ.2012.2201728.
- Wu. (2012). On the Fundamental Differences Between Interval Type-2 and Type-1 Fuzzy Logic Controllers, *IEEE Transactions on Fuzzy Systems*, 20(5) 832-848 doi: 10.1109/TFUZZ.2012.2186818.
- Gaxiola F., Melin P., Valdez F., Castro J.R. (2018). Ensemble Neural Network with Type-2 Fuzzy Weights Using Response Integration for Time Series Prediction. *Studies in Fuzziness and Soft Computing*, vol 361. https://doi.org/10.1007/978-3-319-75408-6_15
- Ibarra-Berastegi, G., Elias, A.; Barona, A., Saenz, J., Ezcurra, A., de Argandoña, J.D. (2008). From diagnosis to prognosis for forecasting air pollution using neural networks: Air pollution monitoring in Bilbao. *Environ. Model. Softw.* 23, 622–637.
- M. Mendel. (2013). on km algorithms for solving type-2 fuzzy set problems, *IEEE Transactions on Fuzzy Systems*, 21(3) 426–446. doi: 10.1109/TFUZZ.2012.2227488.

13. J. M. Mendel, R. I. John, and F. Liu. (2006). Interval type-2 fuzzy logic systems made simple, *IEEE Transactions on Fuzzy Systems*, 14 (6) doi: 808–821. 10.1109/TFUZZ.2006.879986.
14. Safari A, Hosseini R, Mazinani M. Dynamic Type-2 Fuzzy Time Warping (DT2FTW): A Hybrid Model for Uncertain Time-Series Prediction. *IJFIS* 2021; 21:338-348.
15. K. Smagulova, O. Krestinskaya, and A. P. James. (2018). A memristor-based long short-term memory circuit. *Analog Integr. Circuits Signal Process.* 95(3) 467–472. doi: <https://doi.org/10.1007/s10470-018-1180-y>
16. Li, C., Wang, Z. and Rao, M Long. (2019). short-term memory networks in memristor crossbar arrays. *Nat Mach Intell*, 1(1) 49–57. <https://doi.org/10.1038/s42256-018-0001-4>
17. Li, R., Chen, X. and Balezentis, T. (2020). Multi-step least squares support vector machine modeling approach for forecasting short-term electricity demand with application. *Neural Comput & Applic.* 21(1) (2020) 1–20. <https://doi.org/10.1007/s00521-020-04996-3>
18. Liang Guo, Naipeng Li, Feng Jia, Yaguo Lei, Jing Lin. (2017). A recurrent neural network-based health indicator for remaining useful life prediction of bearings, *Neurocomputing*, 24098-109 doi: <https://doi.org/10.1016/j.neucom.2017.02.045>.
19. Liu, H.; Li, Q.; Yu, D.; Gu, Y. (2019). Air Quality Index and Air Pollutant Concentration Prediction Based on Machine Learning Algorithms. *Appl. Sci.* 9, 4069.
20. Marlier, M.E., Jina, A.S., Kinney, P.L. (2016). Extreme Air Pollution in Global Megacities. *Curr Clim Change Rep* 2, 15–27. <https://doi.org/10.1007/s40641-016-0032-z>
21. N. Anh, S. Suresh, M. Pratama, and N. Srikanth. (2019). Interval prediction of wave energy characteristics using meta-cognitive interval type-2 fuzzy inference system, *Knowledge-Based Systems*, 169 doi:28-38. <https://doi.org/10.1016/j.knosys.2019.01.025>.
22. P. C. de Lima Silva, H. J. Sadaei, R. Ballini, and F. G. Guimarães. (2020). Probabilistic Forecasting with Fuzzy Time Series, *IEEE Transactions on Fuzzy Systems*, 28(8) 1771-1784 doi: 10.1109/TFUZZ.2019.2922152.
23. P. Pahlavani, H. Sheikhan and B. Bigdeli. (2017). Assessment of an air pollution monitoring network to generate urban air pollution maps using Shannon information index, fuzzy overlay, and Dempster-Shafer theory, A case study: Tehran, Iran, *Atmospheric Environment*, (167), 254-269, <https://doi.org/10.1016/j.atmosenv.2017.08.039>.
24. R. N. Almanza, R. J. Ramirez, G. Licea, J. R. Castro. (2020). Automated Ontology Extraction from Unstructured Texts using Deep Learning, Intuitionistic and Type-2 Fuzzy Logic Enhancements in Neural and Optimization Algorithms: Theory and Applications, *Springer cham.* 727-755. <https://doi.org/10.1007/978-3-030-35445-9>
25. Safari, A., Hosseini, R., Mazinani, M. (2022). An Adaptive Intelligent Type-2 Fuzzy Logic Model to Manage Uncertainty of Short and Long Time-Series in Covid-19 Patterns Prediction: A Case Study on Iran. *Computational Intelligence in Electrical Engineering*.
26. Safari, A., Mazinani, M., Hosseini, R. (2017). A Novel Type-2 Adaptive Neuro-Fuzzy Inference System Classifier for Modelling Uncertainty in Prediction of Air Pollution Disaster (RESEARCH NOTE). *International Journal of Engineering*, 30(11), 1746-1751.
27. V. Georgescu. (2014). Joint propagation of ontological and epistemic uncertainty across risk assessment and fuzzy time series models, *Computer Science and Information Systems.* 11(2) doi: <https://doi.org/10.2298/CSIS121215048G>
28. V. Sumati and C. Patvardhan. (2018). Interval Type-2 Mutual Subsethood Fuzzy Neural Inference System (IT2MSFuNIS), *IEEE Transactions on Fuzzy Systems*, 26(1)
29. W. Chen and Y. Zou. (2020). Group decision making under generalized fuzzy soft sets and limited cognition of decision makers, *Engineering Applications of Artificial Intelligence*, 87 103344. Doi: <https://doi.org/10.1016/j.engappai.2019.103344>.
30. Xiaoyang Liu, Zhigang Zeng, Donald C. Wunsch. (2020). Memristor-based LSTM network with in situ training and its applications, *Neural Networks*, 131, 300-311. <https://doi.org/10.1016/j.neunet.2020.07.035>
31. Y. Lin, J. Chang, N. R. Pal and C. Lin. (2013). A Mutually Recurrent Interval Type-2 Neural Fuzzy System (MRIT2NFS) With Self-Evolving Structure and Parameters, *IEEE Transactions on Fuzzy Systems.* 21(3) 492-509, 10.1109/TFUZZ.2013.2255613.
32. Sekhavati E, Jalilzadeh Yengejeh R, (2021) Assessment Optimization of Safety and Health Risks Using Fuzzy TOPSIS Technique (Case Study: Construction Sites in the South of Iran). *Journal of Environmental Health and Sustainable Development*, 6(4): 1494-1506.
33. Sekhavati E, Jalilzadeh R, (2022) Optimizing the Risk of Building Environments Using Multi-Criteria Decision Making. *Anthropogenic Pollution*, 6(1):1-7.
34. Hoseini LK, Yengejeh RJ, Rouzbehani MM, Sabzalipour S, (2022) Health risk assessment of volatile organic compounds (VOCs) in a refinery in the southwest of Iran using SQRA method. *Frontiers in Public Health*, 10:978354
35. Khajeh Hoseini L, Jalilzadeh Yengejeh R, Mahmoudie A, Mohammadi Rouzbehani M, Sabzalipour S, (2021) Prioritization of Effective Strategic Parameters in the Removal of VOCs from the ROP System by Using AHP: A Case Study of Abadan Oil Refinery. *Journal of Health Sciences & Surveillance System*, 9(3):199-205.
36. Guo, H., Gu, W., Khayatnezhad, M. & Ghadimi, N. 2022. Parameter extraction of the SOFC mathematical model based on fractional order version of dragonfly algorithm. *International Journal of Hydrogen Energy*, 47, 24059-24068.
37. Lin, H. 2022. Levafix blue color's visible light degradation utilizing Fenton and photo-Fenton procedures. *Water and Environmental Sustainability*, 2, 1-8.
38. Mobar, S. & Bhatnagar, P. 2022. ling women by

- Greenhouse plan as illustrated in the Post-Feminist Tamil Film 36 Vayadhinile. *Water and Environmental Sustainability*, 2, 9-12.
39. Zhang, J., Khayatnezhad, M. & Ghadimi, N. 2022. Optimal model evaluation of the proton-exchange membrane fuel cells based on deep learning and modified African Vulture Optimization Algorithm. *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*, 44, 287-305.