# Neural Network Approach to Solving Fuzzy Nonlinear Equations using Z-Numbers

Raheleh Jafari, Sina Razvarz, and Alexander Gegov, *Member, IEEE*

***Abstract*—In this work, the fuzzy property is described by means of the Z-number as the coefficients and variables of the fuzzy equations. This alteration for the fuzzy equation is appropriate for system modeling with Z-number parameters. In this paper, the fuzzy equation with Z-number coefficients and variables is tended to be used as the models for the uncertain systems. The modeling issue related to the uncertain system is to obtain the Z-number coefficients and variables of the fuzzy equation. Nevertheless, it is extremely hard to get the Z-number coefficients of the fuzzy equations.**

**In this paper in order to model the uncertain nonlinear systems, a novel structure of the multilayer neural network is utilized in such a manner that it is able to obtain the Z-number coefficients of the fuzzy equation. The suggested technique is validated by some examples with applications.**

***Index Terms*—Uncertain nonlinear system, fuzzy equation, Z-number, multilayer neural network.**

## I. Introduction

Fuzzy polynomial interpolation can be considered as a special case of fuzzy system modeling. Polynomials with fuzzy coefficients can be utilized for interpolating fuzzy data [9]. Interpolation technique is extensively applied for function estimation [24]. In [31], the modeling of the system by utilizing the fuzzy polynomial interpolation is described. In [35], a systolic algorithm in order to interpolate and evaluate the polynomials is proposed. In [46], two-dimensional polynomial interpolation is suggested. In [37], smooth function estimation is utilized. In [9], [42] smooth function estimation causes a model by implementing Lagrange interpolating polynomials at the points of product grids. Nevertheless, these approaches may not work well when the interpolation points have uncertainties.

The fuzzy polynomial is taken to be as a special form of the fuzzy equation. Various techniques exist in order to make the fuzzy equations. In [17] the fuzzy number with parametric shape is utilized, also the crisp linear system is implemented instead of the original fuzzy equation. In [1] the homotypic analysis method is suggested. The Newton technique is studied in [2]. The numerical solution of the fuzzy equation by utilizing the fixed point method is investigated in [8]. Iterative method [25], interpolation method [41] as well as the Runge-Kutta method [32] are applied for finding the numerical solutions of fuzzy equations. Neural network technique is also utilized for resolving fuzzy equations. In [11], the neural network technique is utilized in order to solve the fuzzy quadratic equation. In [23] the results of [11] has been extended to the fuzzy polynomial equation. In [22], the neural network technique is used for finding the solution of the dual fuzzy equation. Evaluation of fully fuzzy matrix equations by the fuzzy neural network is investigated in [30]. However, the issue of obtaining the Z-number coefficients of nonlinear fuzzy equations using neural network approach has not been touched yet. Studying such an issue is a challenge since there are some difficult points, such as accuracy, convergence and construction issues regarding to the selected model.

Real-world information is incomplete and is generally expressed in natural language. Furthermore, this information is oftentimes partially reliable and a degree of reliability is also described in natural language. Accordingly, the concept of a Z-number is a more adequate concept for the explanation of real-world information [26], [44]. Various fields connected to the analysis of the decisions utilize the concept of Z-numbers. Z-number involves less complexity in calculation when compared with nonlinear system modeling techniques. Also in comparison with fuzzy numbers, the Z-numbers are more accurate. To fully utilize the Z-information in real life scenarios, more deep studies on Z-number are required [6]. There exist few works concerned with the theoretical concept of Z-numbers [18]. [4] was a starting point in the extension of the Z-numbers. In [45] a general framework of calculation of a Z-number-valued function based on the Zadehs extension principle is suggested. A theorem for transferring the Z-number into the fuzzy number is proposed in [27]. Also, in [44] the author has proposed a new method in order to transfer the Z-number into the fuzzy number. An approach to use Z-numbers for answering questions and decisions making is considered in [26]. A distance-based measure of linguistic Z-numbers is proposed in [40]. In [38] several techniques of approximate evaluation of a Z-number for reducing calculational complexity is suggested. In [29] decision making under interval, set-valued, fuzzy and Z-number uncertainty is considered. In [14] numerical solution of linear regression based on Z-Numbers by the improved neural network is proposed. In [3] a Z-number-based fuzzy inference system for control of the omnidirectional soccer robot is suggested. A general and computationally effective method to calculation with discrete Z-numbers is proposed in [5].

This paper builds on a recent work of the authors that presents a detailed study for finding a numerical solution of fuzzy equations using neural networks [21]. In particular, this paper aims to find the coefficients of a fuzzy equation in

R. Jafari is with the Centre for Artificial Intelligence Research (CAIR), University of Agder, Grimstad, Norway, e-mail: raheleh.jafari@uia.no.

S. Razvarz is with the Departamento de Control Automático, CINVESTAV-IPN (National Polytechnic Institute), Mexico City, Mexico.

A. Gegov is with the School of Computing, University of Portsmouth, Buckingham Building, Portsmouth PO1 3HE, UK.

the context of modeling as opposed to [21] whose aim is to find a Z-number solution to a fuzzy equation in the context of control. Also, the structure of the neural network along with the fuzzy equations proposed in this paper is novel and different from the one in [21]. Finally, this paper introduces novel theorems in addition to the ones in [21].

The researches that have been done so far on modeling the nonlinear systems are mostly based on the regular fuzzy numbers. However, this leads to information loss that affects decision making. In this paper, the fuzzy equation is utilized in order to model the uncertain nonlinear systems, where the coefficients and variables are Z-numbers. Studying of previous works by other researchers shows that no study has been done for obtaining the Z-number coefficients of the fuzzy equations, so this paper can be considered as one of the first attempts at obtaining the coefficients of fuzzy equations on the basis of Z-numbers. We introduce a novel multilayer neural network architecture in order to estimate the Z-number coefficients of the fuzzy equations. The backpropagation technique is utilized for training the neural network. Some important theorems are developed in order to estimate the upper bounds of the modeling errors with fuzzy equations. The suggested technique is validated by some examples with applications. The remaining of the article is organized as follows. In Section 2, some basic definitions related to the Z-numbers are given. The proposed method for obtaining the Z-number coefficients of the fuzzy equations is demonstrated in Section 3. Some important theorems are given in Section 4. Some examples with applications in mechanics are given in Section 5. Section 6 concludes the work and provides discussions on further work.

## II. Nonlinear system modeling with fuzzy equations and Z-numbers

A common discrete-time nonlinear system is defined as

$$\vartheta_{r+1} = f\left[\vartheta_r, q_r\right], \quad w_r = g\left[\vartheta_r\right] \tag{1}$$

where $q_r \in \Re^u$ is the input vector, $\vartheta_r \in \Re^l$ is an internal state vector, also $w_r \in \Re^m$ is the output vector. $f$, as well as $g$, are generalized nonlinear smooth functions $f, g \in C^\infty$. Define $W_r = \left[w_{r+1}^T, w_r^T, \cdots\right]^T$ as well as $Q_r = \left[q_{r+1}^T, q_r^T, \cdots\right]^T$. Assume $\frac{\partial W}{\partial \vartheta}$ is non-singular at the instance $\vartheta_r = 0$, $Q_r = 0$, so the following model is extracted

$$w_r = \Upsilon[w_{r-1}^T, w_{r-2}^T, \cdots q_r^T, q_{r-1}^T, \cdots] \tag{2}$$

in which $\Upsilon(\cdot)$ is a nonlinear difference equation representing the plant dynamics, $q_r$ as well as $w_r$ are calculable scalar input and output respectively. The nonlinear system (2) is a NARMA model. The input of the nonlinear system is defined as

$$\vartheta_r = [w_{r-1}^T, w_{r-2}^T, \cdots q_r^T, q_{r-1}^T, \cdots]^T \tag{3}$$

the output as $w_r$.

The nonlinear system explained in (2), can be written as the below mentioned linear-in-parameter model

$$w_r = \sum_{\iota=1}^{\gamma} \sum_{\kappa=1}^{\delta} a_{\iota\kappa} f_\iota\left(\vartheta_r\right) g_k\left(\upsilon_r\right) \tag{4}$$

where $a_{\iota\kappa}$ is the linear parameter, $f_\iota\left(\vartheta_r\right)$ as well as $g_k\left(\upsilon_r\right)$ are nonlinear functions. The variables of these functions are quantifying input and output.

The uncertain nonlinear systems can be modeled using the linear-in-parameter models with uncertain parameters. In this paper, it has been assumed that the model of the nonlinear systems (4) contains uncertainties in the $a_{\iota\kappa}$, $\vartheta_r$ as well as $\upsilon_r$. These uncertainties have been stated in the form of Z-numbers [45].

**Definition 1.** Suppose $c$ is: 1) normal, there is $\varsigma_0 \in \Re$ where $c(\varsigma_0) = 1$, 2) convex, $c(\beta\varsigma + (1-\beta)\varsigma) \geq \min\{c(\varsigma), c(\varrho)\}$, $\forall\varsigma, \varrho \in \Re, \forall\beta \in [0,1]$, 3) upper semi-continuous on $\Re$, $c(\varsigma) \leq c(\varsigma_0) + \varepsilon$, $\forall\varsigma \in N(\varsigma_0)$, $\forall\varsigma_0 \in \Re$, $\forall\varepsilon > 0$, $N(\varsigma_0)$ is a neighborhood, 4) $c^+ = \{\varsigma \in \Re, c(\varsigma) > 0\}$ is compact, so $c$ is a fuzzy variable, $c \in E : \Re \to [0,1]$.

The fuzzy variable $c$ is demonstrated as

$$c = (\underline{c}, \overline{c}) \tag{5}$$

in which $\underline{c}$ is the lower-bound variable also, $\overline{c}$ is the upper-bound variable.

**Definition 2.** The $Z$-number is made up of two components $Z = [c(\varsigma), p]$. The first component $c(\varsigma)$ is the restriction on a real-valued uncertain variable $\varsigma$. The second component $p$ is a measure of the reliability of $c$. $p$ can be reliability, strength of belief, probability or possibility. The $Z$-number can be stated as $Z^+$-number, in a case $c(\varsigma)$ be a fuzzy number also $p$ be the probability distribution of $\varsigma$. If $c(\varsigma)$, as well as $p$, are fuzzy numbers, the $Z$-number can be stated as $Z^-$-number.

The $Z^+$-number contains more information when compared with the $Z^-$-number. In this paper, the definition of $Z^+$-number is utilized, i.e., $Z = [c, p]$, $c$ is a fuzzy number also, $p$ is a probability distribution.

The most common membership functions which define the fuzzy numbers are the triangular function

$$\mu_c = H(s, u, v) = \begin{cases} \frac{\varsigma - s}{u - s} & s \leq \varsigma \leq u \\ \frac{v - \varsigma}{v - u} & u \leq \varsigma \leq v \end{cases} \quad otherwise\ \mu_c = 0 \tag{6}$$

and trapezoidal function

$$\mu_c = H(s, u, v, w) = \begin{cases} \frac{\varsigma - s}{u - s} & s \leq \varsigma \leq u \\ \frac{w - \varsigma}{w - v} & v \leq \varsigma \leq w \\ 1 & u \leq \varsigma \leq v \end{cases} \quad otherwise\ \mu_c = 0 \tag{7}$$

If $c$ represents a fuzzy event in $\Re$, the real line, then the probability measure can be stated as

$$P(c) = \int_{\Re} \mu_c(\varsigma) p(\varsigma) d\varsigma \tag{8}$$

where $p$ is the probability density of $\varsigma$. For discrete $Z$-numbers we have

$$P(c) = \sum_{\iota=1}^{n} \mu_c(\varsigma_\iota) p(\varsigma_\iota) \tag{9}$$

**Definition 3.** The $\alpha$-level for fuzzy number $c$ is stated as

$$[c]^\alpha = \{\varsigma \in \Re : c(\varsigma) \geq \alpha\} \tag{10}$$

where $0 < \alpha \leq 1$, $c \in E$.

Therefore $[c]^0 = c^+ = \{\varsigma \in \Re, c(\varsigma) > 0\}$. As $\alpha \in [0,1]$, $[c]^\alpha$ is bounded, $\underline{c}^\alpha \leq [c]^\alpha \leq \overline{c}^\alpha$. The $\alpha$-level of $c$ between $\underline{c}^\alpha$ and $\overline{c}^\alpha$ can be defined as

$$[c]^\alpha = (\underline{c}^\alpha, \overline{c}^\alpha) \tag{11}$$

$\underline{c}^\alpha$, as well as $\overline{c}^\alpha$, are the function of $\alpha$. We define $\underline{c}^\alpha = d_A(\alpha)$, $\overline{c}^\alpha = d_B(\alpha)$, $\alpha \in [0,1]$.

**Definition 4.** The $\alpha$-level of the $Z$-number $Z = (c,p)$ is defined as follows

$$[Z]^\alpha = ([c]^\alpha, [p]^\alpha) \tag{12}$$

where $0 < \alpha \leq 1$. $[p]^\alpha$ is computed by the Nguyen's theorem

$$[p]^\alpha = p([c]^\alpha) = p([\underline{c}^\alpha, \overline{c}^\alpha]) = \left[\underline{P}^\alpha, \overline{P}^\alpha\right] \tag{13}$$

where $p([c]^\alpha) = \{p(\varsigma) | \varsigma \in [c]^\alpha\}$. Therefore, $[Z]^\alpha$ is stated as

$$[Z]^\alpha = \left(\underline{Z}^\alpha, \overline{Z}^\alpha\right) = \left(\left(\underline{c}^\alpha, \underline{P}^\alpha\right), \left(\overline{c}^\alpha, \overline{P}^\alpha\right)\right) \tag{14}$$

where $\underline{P}^\alpha = \underline{c}^\alpha p(\underline{\varsigma}_\iota^\alpha)$, $\overline{P}^\alpha = \overline{c}^\alpha p(\overline{\varsigma}_\iota^\alpha)$, $[\varsigma_\iota]^\alpha = (\underline{\varsigma}_\iota^\alpha, \overline{\varsigma}_\iota^\alpha)$.

The same as with the fuzzy numbers [13], three main operations are defined for the $Z$-numbers: $\oplus$, $\ominus$ and $\odot$, which indicate sum, subtract and multiply respectively. In this paper, the proposed operations are different from the ones in [44].

Suppose $Z_1 = (c_1, p_1)$ as well as $Z_2 = (c_2, p_2)$ are two discrete $Z$-numbers expressing the uncertain variables $\varsigma_1$ and $\varsigma_2$, so, $\sum_{\iota=1}^n p_1(\varsigma_{1\iota}) = 1$, $\sum_{\iota=1}^n p_2(\varsigma_{2\iota}) = 1$. The following operation is defined

$$Z_{12} = Z_1 * Z_2 = (c_1 * c_2, p_1 * p_2) \tag{15}$$

where $* \in \{\oplus, \ominus, \odot\}$.

The operations used for the fuzzy numbers $[c_1]^\alpha = [c_{11}^\alpha, c_{12}^\alpha]$ and $[c_2]^\alpha = [c_{21}^\alpha, c_{22}^\alpha]$ are stated as [13],

$$\begin{array}{l} [c_1 \oplus c_2]^\alpha = [c_1]^\alpha + [c_2]^\alpha = [c_{11}^\alpha + c_{21}^\alpha, c_{12}^\alpha + c_{22}^\alpha] \\ [c_1 \ominus c_2]^\alpha = [c_1]^\alpha - [c_2]^\alpha = [c_{11}^\alpha - c_{22}^\alpha, c_{12}^\alpha - c_{21}^\alpha] \\ [c_1 \odot c_2]^\alpha = \left( \begin{array}{c} \min\{c_{11}^\alpha c_{21}^\alpha, c_{11}^\alpha c_{22}^\alpha, c_{12}^\alpha c_{21}^\alpha, c_{12}^\alpha c_{22}^\alpha\} \\ \max\{c_{11}^\alpha c_{21}^\alpha, c_{11}^\alpha c_{22}^\alpha, c_{12}^\alpha c_{21}^\alpha, c_{12}^\alpha c_{22}^\alpha\} \end{array} \right) \end{array} \tag{16}$$

For the discrete probability distributions, the following relation is defined for all $p_1 * p_2$ operations

$$p_1 * p_2 = \sum_\iota p_1(\varsigma_{1,\iota}) p_2(\varsigma_{2,(n-\iota)}) = p_{12}(\varsigma) \tag{17}$$

Let us consider the procedures underlying computation of gH-difference $Z_3 = Z_1 \ominus_{gH} Z_2$ of Z-numbers $Z_1 = (c_1, p_1)$ and $Z_2 = (c_2, p_2)$, where $Z_3 = (c_3, p_3)$. The Hukuhara difference of two fuzzy numbers $c_1$ and $c_2$ is defined as [7],

$$\begin{array}{c} c_1 \ominus_H c_2 = c_3 \\ c_1 = c_2 \oplus c_3 \end{array} \tag{18}$$

Let $[c_1]^\alpha = [c_{11}^\alpha, c_{12}^\alpha]$ and $[c_2]^\alpha = [c_{21}^\alpha, c_{22}^\alpha]$. In a case that $c_1 \ominus_H c_2$ exists, the $\alpha$-level can be defined as

$$[c_1 \ominus_H c_2]^\alpha = [\underline{c}_{11}^\alpha - \underline{c}_{22}^\alpha, \overline{c}_{12}^\alpha - \overline{c}_{21}^\alpha] \tag{19}$$

Clearly, $c_1 \ominus_H c_1 = 0$, $c_1 \ominus c_1 \neq 0$.

Moreover, the generalized Hukuhara difference is defined as [10],

$$c_1 \ominus_{gH} c_2 = c_3 \Longleftrightarrow \left\{ \begin{array}{l} 1)\ c_1 = c_2 \oplus c_3 \\ \text{or } 2)\ c_2 = c_1 \oplus (-1) c_3 \end{array} \right. \tag{20}$$

By taking into consideration the $\alpha$-level, we have $[c_1 \ominus_{gH} c_2]^\alpha = [\min\{\underline{c}_{11}^\alpha - \underline{c}_{22}^\alpha, \overline{c}_{12}^\alpha - \overline{c}_{21}^\alpha\}, \max\{\underline{c}_{11}^\alpha - \underline{c}_{22}^\alpha, \overline{c}_{12}^\alpha - \overline{c}_{21}^\alpha\}]$ and if $c_1 \ominus_{gH} c_2$ also, if $c_1 \ominus_H c_2$ exists, $c_1 \ominus_H c_2 = c_1 \ominus_{gH} c_2$. Let $[c_3]^\alpha = [c_{31}^\alpha, c_{32}^\alpha]$. The conditions for the existence of $c_3 = c_1 \ominus_{gH} c_2 \in E$ are

$$\begin{array}{l} 1) \left\{ \begin{array}{c} \underline{c}_{31}^\alpha = \underline{c}_{11}^\alpha - \underline{c}_{22}^\alpha \ and\ \overline{c}_{32}^\alpha = \overline{c}_{12}^\alpha - \overline{c}_{21}^\alpha \\ with\ \underline{c}_{31}^\alpha\ increasing,\ \overline{c}_{32}^\alpha\ decreasing,\ \underline{c}_{31}^\alpha \leq \overline{c}_{32}^\alpha \end{array} \right. \\ 2) \left\{ \begin{array}{c} \underline{c}_{31}^\alpha = \overline{c}_{12}^\alpha - \overline{c}_{21}^\alpha \ and\ \overline{c}_{32}^\alpha = \underline{c}_{11}^\alpha - \underline{c}_{22}^\alpha \\ with\ \underline{c}_{31}^\alpha\ increasing,\ \overline{c}_{32}^\alpha\ decreasing,\ \underline{c}_{31}^\alpha \leq \overline{c}_{32}^\alpha \end{array} \right. \end{array} \tag{21}$$

where $\forall \alpha \in [0,1]$. Now we proceed to compute the gH-difference $Z_3 = Z_1 \ominus_{gH} Z_2$ which is defined as

$$Z_3 = (c_1 \ominus_{gH} c_2, p_1 - p_2) \tag{22}$$

where $p_1$ and $p_2$ are represented by discrete probability distributions [33],

$$\begin{array}{l} p_1 = \frac{p_1(\varsigma_{11})}{\varsigma_{11}} + \frac{p_1(\varsigma_{12})}{\varsigma_{12}} + \ldots + \frac{p_1(\varsigma_{1n})}{\varsigma_{1n}} \\ p_2 = \frac{p_2(\varsigma_{21})}{\varsigma_{21}} + \frac{p_2(\varsigma_{22})}{\varsigma_{22}} + \ldots + \frac{p_2(\varsigma_{2n})}{\varsigma_{2n}} \end{array} \tag{23}$$

also, $\sum_{\iota=1}^n p_1(\varsigma_{1\iota}) = 1$, $\sum_{\iota=1}^n p_2(\varsigma_{2\iota}) = 1$.

Suppose $c$ is a triangular function, the absolute value of the $Z$-number $Z = (c,p)$ is defined as

$$|Z(\varsigma)| = (|s_1| + |u_1| + |v_1|, p(|s_2| + |u_2| + |v_2|)) \tag{24}$$

Let $c_1$ as well as $c_2$ are triangular functions, the supremum metric for $Z$-numbers $Z_1 = (c_1, p_1)$ and $Z_2 = (c_2, p_2)$ is expressed as

$$D(Z_1, Z_2) = d(c_1, c_2) + d(p_1, p_2) \tag{25}$$

where $d(\cdot,\cdot)$ is the supremum metrics for fuzzy sets [20]. $D(Z_1, Z_2)$ has the below-mentioned properties,

$$\begin{array}{l} D(Z_1 + Z, Z_2 + Z) = D(Z_1, Z_2) \\ D(Z_2, Z_1) = D(Z_1, Z_2) \\ D(bZ_1, kZ_2) = |b| D(Z_1, Z_2) \\ D(Z_1, Z_2) \leq D(Z_1, Z) + D(Z, Z_2) \end{array} \tag{26}$$

where $b \in \Re$, $Z = (c,p)$ is $Z$-number, also $c$ is a triangle function.

**Definition 5.** Suppose $\widetilde{Z}$ is the space of $Z$-numbers. The $\alpha-$level of the $Z$-number valued function $H : [0,s] \to \widetilde{Z}$ is defined as

$$H(c,\alpha) = [\underline{H}(c,\alpha), \overline{H}(c,\alpha)] \tag{27}$$

where $c \in \widetilde{Z}$, for each $\alpha \in [0,1]$.

Using the definition of Generalized Hukuhara difference, the gH-derivative of $H$ at $c_0$ is defined as

$$\frac{d}{dt} H(c_0) = \lim_{\zeta \to 0} \frac{1}{\zeta} [H(c_0 + \zeta) \ominus_{gH} H(c_0)] \tag{28}$$

In (28), $H(c_0 + \zeta)$ as well as $H(c_0)$ represent symmetric pattern with $Z_1$ and $Z_2$ respectively given in (20).

In this paper, the fuzzy equation (4) is used in order to model the uncertain nonlinear system (1). Modeling with fuzzy equation (or fuzzy polynomial ) is named as fuzzy interpolation. Here, the fuzzy equation (4) is utilized in order to model the uncertain nonlinear system (1), in such a way that the output of the plant $w_r$ approaches to the desired output $w_r^*$,

$$\min_{q_r} \|w_r - w_r^*\| \tag{29}$$

The aim of the modeling is to find the $a_{\iota\kappa}$ for the below fuzzy equation

$$w_r^* = \sum_{\iota=1}^{\gamma}\sum_{\kappa=1}^{\delta} a_{\iota\kappa} f_\iota\left(\vartheta_r\right) g_k\left(\upsilon_r\right) \tag{30}$$

where $\vartheta_r = [w_{r-1}^T, w_{r-2}^T, \cdots q_r^T, q_{r-1}^T, \cdots]^T$, $a_{\iota\kappa}$ is the linear parameter, $f_\iota\left(\vartheta_r\right)$ as well as $g_k\left(\upsilon_r\right)$ are nonlinear functions.

## III. Neural network approach for Z-number parameter approximation

In this section, a neural network is designed in order to demonstrate the fuzzy equation (4), see Fig. 1. The inputs of the neural network are the Z-numbers $(\vartheta_r, p)$ and $(\upsilon_r, p)$, also, the output is the Z-number $(w_r, p)$. The proposed neural network finds $(a_{\iota\kappa}, p)$ in such a way that the output of the neural network $(w_r, p)$ approaches to the desired output $(w_r^*, p)$. Based on Definition 2, (14) and (30), the results (31)-(34) are achieved.

The Z-number inputs $(\vartheta_r, p)$ as well as $(\upsilon_r, p)$ are initially implemented to $\alpha$-level as in (12), when the $\alpha-$level sets of $\vartheta_r$ and $\upsilon_r$ are nonnegative, i.e., $0 \le \underline{\vartheta}_r^\alpha \le \overline{\vartheta}_r^\alpha$ and $0 \le \underline{\upsilon}_r^\alpha \le \overline{\upsilon}_r^\alpha$,

$$\begin{array}{c} \left([\vartheta_r]^\alpha, [p]^\alpha\right) = \left((\underline{\vartheta}_r^\alpha, \underline{\vartheta}_r^\alpha p(\underline{\varsigma}_1^\alpha)), (\overline{\vartheta}_r^\alpha, \overline{\vartheta}_r^\alpha p(\overline{\varsigma}_1^\alpha))\right) \\ \left([\upsilon_r]^\alpha, [p]^\alpha\right) = \left((\underline{\upsilon}_r^\alpha, \underline{\upsilon}_r^\alpha p(\underline{\varsigma}_2^\alpha)), (\overline{\upsilon}_r^\alpha, \overline{\upsilon}_r^\alpha p(\overline{\varsigma}_2^\alpha))\right) \end{array} \tag{31}$$

where $[\varsigma_1]^\alpha = (\underline{\varsigma}_1^\alpha, \overline{\varsigma}_1^\alpha), \varsigma_1 \in [\vartheta_r]^\alpha, [\varsigma_2]^\alpha = (\underline{\varsigma}_2^\alpha, \overline{\varsigma}_2^\alpha), \varsigma_2 \in [\upsilon_r]^\alpha$. Afterward, we have the following relation in the first hidden unit for $0 \le \alpha \le 1$,

$$\begin{array}{c} \left([\Omega_\iota]^\alpha, [p]^\alpha\right) = h_1(f_\iota(\underline{\vartheta}_r^\alpha, \underline{\vartheta}_r^\alpha p(\underline{\varsigma}_1^\alpha)), f_\iota(\overline{\vartheta}_r^\alpha, \overline{\vartheta}_r^\alpha p(\overline{\varsigma}_1^\alpha))) \\ \iota = 1 \cdots \gamma \\ \left([\Omega_\kappa]^\alpha, [p]^\alpha\right) = h_2(g_\kappa(\underline{\upsilon}_r^\alpha, \underline{\upsilon}_r^\alpha p(\underline{\varsigma}_2^\alpha)), g_\kappa(\overline{\upsilon}_r^\alpha, \overline{\upsilon}_r^\alpha p(\overline{\varsigma}_2^\alpha))) \\ \kappa = 1 \cdots \delta \end{array} \tag{32}$$

where $h_1$ and $h_2$ are identity activation functions. Also, we have the following relation in the second hidden unit for $0 \le \alpha \le 1$,

$$\begin{array}{c} \left([\Omega_{\iota,\kappa}]^\alpha, [p]^\alpha\right) = h_3(\{\sum_{\iota,\kappa\in A}(\underline{\Omega}_\iota^\alpha, \underline{\Omega}_\iota^\alpha p(\underline{\varsigma}_\iota^\alpha))(\underline{\Omega}_\kappa^\alpha, \underline{\Omega}_\kappa^\alpha p(\underline{\varsigma}_\kappa^\alpha)) \\ + \sum_{\iota,\kappa\in B}(\overline{\Omega}_\iota^\alpha, \overline{\Omega}_\iota^\alpha p(\overline{\varsigma}_\iota^\alpha))(\overline{\Omega}_\kappa^\alpha, \overline{\Omega}_\kappa^\alpha p(\overline{\varsigma}_\kappa^\alpha)) \\ + \sum_{\iota,\kappa\in C}(\underline{\Omega}_\iota^\alpha, \underline{\Omega}_\iota^\alpha p(\underline{\varsigma}_\iota^\alpha))(\overline{\Omega}_\kappa^\alpha, \overline{\Omega}_\kappa^\alpha p(\overline{\varsigma}_\kappa^\alpha)), \\ \sum_{\iota,\kappa\in A'}(\overline{\Omega}_\iota^\alpha, \overline{\Omega}_\iota^\alpha p(\overline{\varsigma}_\iota^\alpha))(\overline{\Omega}_\kappa^\alpha, \overline{\Omega}_\kappa^\alpha p(\overline{\varsigma}_\kappa^\alpha)) \\ + \sum_{\iota,\kappa\in B'}(\underline{\Omega}_\iota^\alpha, \underline{\Omega}_\iota^\alpha p(\underline{\varsigma}_\iota^\alpha))(\underline{\Omega}_\kappa^\alpha, \underline{\Omega}_\kappa^\alpha p(\underline{\varsigma}_\kappa^\alpha)) \\ + \sum_{\iota,\kappa\in C'}(\overline{\Omega}_\iota^\alpha, \overline{\Omega}_\iota^\alpha p(\overline{\varsigma}_\iota^\alpha))(\underline{\Omega}_\kappa^\alpha, \underline{\Omega}_\kappa^\alpha p(\underline{\varsigma}_\kappa^\alpha))\}) \end{array} \tag{33}$$

where $h_3$ is identity activation function. Also, $A = \{\iota, \kappa | \ \underline{\Omega}_\iota^\alpha \ge 0, \ \underline{\Omega}_\kappa^\alpha \ge 0\}$, $B = \{\iota, \kappa | \ \overline{\Omega}_\iota^\alpha < 0, \ \overline{\Omega}_\kappa^\alpha < 0\}$, $C = \{\iota, \kappa | \ \underline{\Omega}_\iota^\alpha < 0, \ \overline{\Omega}_\kappa^\alpha \ge 0\}$, $A' = \{\iota, \kappa | \ \overline{\Omega}_\iota^\alpha \ge 0, \ \overline{\Omega}_\kappa^\alpha \ge 0\}$, $B' = \{\iota, \kappa | \ \underline{\Omega}_\iota^\alpha < 0, \ \underline{\Omega}_\kappa^\alpha < 0\}$, $C' = \{\iota, \kappa | \ \overline{\Omega}_\iota^\alpha < 0, \ \underline{\Omega}_\kappa^\alpha \ge 0\}$, $[\varsigma_\iota]^\alpha = (\underline{\varsigma}_\iota^\alpha, \overline{\varsigma}_\iota^\alpha), \varsigma_\iota \in [\Omega_\iota]^\alpha$, $[\varsigma_\kappa]^\alpha = (\underline{\varsigma}_\kappa^\alpha, \overline{\varsigma}_\kappa^\alpha), \varsigma_\kappa \in [\Omega_\kappa]^\alpha$.

For $0 \le \alpha \le 1$ the output of the neural network is

$$\begin{array}{c} \left([w_r]^\alpha, [p]^\alpha\right) \\ = O(\{\sum_{\iota,\kappa\in A}(\underline{\Omega}_{\iota,\kappa}^\alpha, \underline{\Omega}_{\iota,\kappa}^\alpha p(\underline{\varsigma}_{\iota,\kappa}^\alpha))(\underline{a}_{\iota,\kappa}^\alpha, \underline{a}_{\iota,\kappa}^\alpha p(\underline{\varsigma}_{\iota,\kappa}^\alpha)) \\ + \sum_{\iota,\kappa\in B}(\overline{\Omega}_{\iota,\kappa}^\alpha, \overline{\Omega}_{\iota,\kappa}^\alpha p(\overline{\varsigma}_{\iota,\kappa}^\alpha))(\overline{a}_{\iota,\kappa}^\alpha, \overline{a}_{\iota,\kappa}^\alpha p(\overline{\varsigma}_{\iota,\kappa}^\alpha)) \\ + \sum_{\iota,\kappa\in C}(\underline{\Omega}_{\iota,\kappa}^\alpha, \underline{\Omega}_{\iota,\kappa}^\alpha p(\underline{\varsigma}_{\iota,\kappa}^\alpha))(\overline{a}_{\iota,\kappa}^\alpha, \overline{a}_{\iota,\kappa}^\alpha p(\overline{\varsigma}_{\iota,\kappa}^\alpha)), \\ \sum_{\iota,\kappa\in A'}(\overline{\Omega}_{\iota,\kappa}^\alpha, \overline{\Omega}_{\iota,\kappa}^\alpha p(\overline{\varsigma}_{\iota,\kappa}^\alpha))(\overline{a}_{\iota,\kappa}^\alpha, \overline{a}_{\iota,\kappa}^\alpha p(\overline{\varsigma}_{\iota,\kappa}^\alpha)) \\ + \sum_{\iota,\kappa\in B'}(\underline{\Omega}_{\iota,\kappa}^\alpha, \underline{\Omega}_{\iota,\kappa}^\alpha p(\underline{\varsigma}_{\iota,\kappa}^\alpha))(\underline{a}_{\iota,\kappa}^\alpha, \underline{a}_{\iota,\kappa}^\alpha p(\underline{\varsigma}_{\iota,\kappa}^\alpha)) \\ + \sum_{\iota,\kappa\in C'}(\overline{\Omega}_{\iota,\kappa}^\alpha, \overline{\Omega}_{\iota,\kappa}^\alpha p(\overline{\varsigma}_{\iota,\kappa}^\alpha))(\underline{a}_{\iota,\kappa}^\alpha, \underline{a}_{\iota,\kappa}^\alpha p(\underline{\varsigma}_{\iota,\kappa}^\alpha))\}) \end{array} \tag{34}$$

where $O$ is identity activation function. Also, $A = \{\iota, \kappa | \ \underline{\Omega}_{\iota,\kappa}^\alpha \ge 0, \ \underline{a}_{\iota,\kappa}^\alpha \ge 0\}$, $B = \{\iota, \kappa | \ \overline{\Omega}_{\iota,\kappa}^\alpha < 0, \ \overline{a}_{\iota,\kappa}^\alpha < 0\}$, $C = \{\iota, \kappa | \ \underline{\Omega}_{\iota,\kappa}^\alpha < 0, \ \overline{a}_{\iota,\kappa}^\alpha \ge 0\}$, $A' = \{\iota, \kappa | \ \overline{\Omega}_{\iota,\kappa}^\alpha \ge 0, \ \overline{a}_{\iota,\kappa}^\alpha \ge 0\}$, $B' = \{\iota, \kappa | \ \underline{\Omega}_{\iota,\kappa}^\alpha < 0, \ \underline{a}_{\iota,\kappa}^\alpha < 0\}$, $C' = \{\iota, \kappa | \ \overline{\Omega}_{\iota,\kappa}^\alpha < 0, \ \underline{a}_{\iota,\kappa}^\alpha \ge 0\}$, $[\varsigma_{\iota,\kappa}]^\alpha = (\underline{\varsigma}_{\iota,\kappa}^\alpha, \overline{\varsigma}_{\iota,\kappa}^\alpha), \varsigma_{\iota,\kappa} \in [\Omega_{\iota,\kappa}]^\alpha$, $[\varsigma_{\iota,\kappa}]^\alpha = (\underline{\varsigma}_{\iota,\kappa}^\alpha, \overline{\varsigma}_{\iota,\kappa}^\alpha), \varsigma_{\iota,\kappa} \in [a_{\iota,\kappa}]^\alpha$.

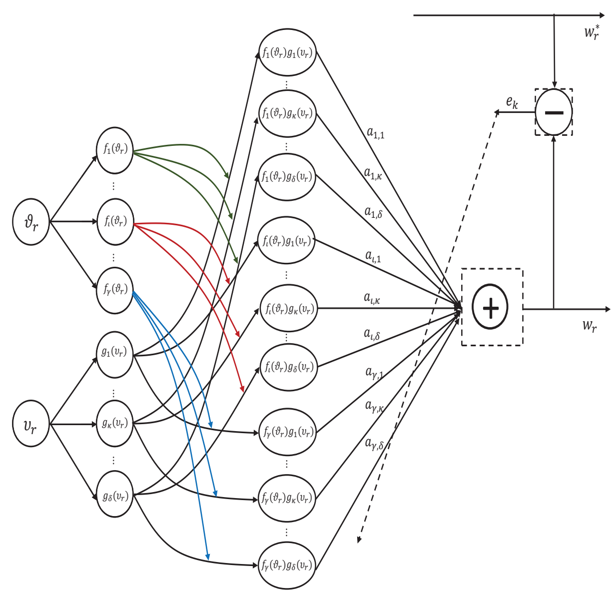

Fig. 1. Neural network structure representing the fuzzy equation

For training the weights, a cost function is defined for the Z-numbers. The training error is defined as follows

$$(e_r, p) = (w_r^*, p) - (w_r, p) \tag{35}$$

For $0 \le \alpha \le 1$ we have,

$$\begin{array}{c} \left([w_r^*]^\alpha, [p]^\alpha\right) = \left((\underline{w}_r^{*\alpha}, \underline{w}_r^{*\alpha} p(\underline{\varsigma}_r^{*\alpha})), (\overline{w}_r^{*\alpha}, \overline{w}_r^{*\alpha} p(\overline{\varsigma}_r^{*\alpha}))\right) \\ \left([w_r]^\alpha, [p]^\alpha\right) = \left((\underline{w}_r^\alpha, \underline{w}_r^\alpha p(\underline{\varsigma}_r^\alpha)), (\overline{w}_r^\alpha, \overline{w}_r^\alpha p(\overline{\varsigma}_r^\alpha))\right) \\ \left([e_r]^\alpha, [p]^\alpha\right) = \left((\underline{e}_r^\alpha, \underline{e}_r^\alpha p(\underline{\varsigma}_r^\alpha)), (\overline{e}_r^\alpha, \overline{e}_r^\alpha p(\overline{\varsigma}_r^\alpha))\right) \end{array} \tag{36}$$

where $[\varsigma_r^*]^\alpha = (\underline{\varsigma}_r^{*\alpha}, \overline{\varsigma}_r^{*\alpha}), \varsigma_r^* \in [w_r^*]^\alpha, [\varsigma_r]^\alpha = (\underline{\varsigma}_r^\alpha, \overline{\varsigma}_r^\alpha), \varsigma_r \in [w_r]^\alpha, [\varsigma_r]^\alpha = (\underline{\varsigma}_r^\alpha, \overline{\varsigma}_r^\alpha), \varsigma_r \in [e_r]^\alpha$.

The cost function is defined as

$$\begin{array}{c} (\Phi_r, p) = (\underline{\Phi}^\alpha, \underline{P}^\alpha(\Phi)) + (\overline{\Phi}^\alpha, \overline{P}^\alpha(\Phi)) \\ (\underline{\Phi}^\alpha, \underline{P}^\alpha(\Phi)) = \frac{1}{2}\left((\underline{w}_r^{*\alpha}, \underline{w}_r^{*\alpha} p(\underline{\varsigma}_r^{*\alpha})) - (\underline{w}_r^\alpha, \underline{w}_r^\alpha p(\underline{\varsigma}_r^\alpha))\right)^2 \\ (\overline{\Phi}^\alpha, \overline{P}^\alpha(\Phi)) = \frac{1}{2}\left((\overline{w}_r^{*\alpha}, \overline{w}_r^{*\alpha} p(\overline{\varsigma}_r^{*\alpha})) - (\overline{w}_r^\alpha, \overline{w}_r^\alpha p(\overline{\varsigma}_r^\alpha))\right)^2 \end{array} \tag{37}$$

where $0 \le \alpha \le 1$. $\Phi_r$ is a scalar function. $\Phi_r \to 0$ signifies that $\left([w_r]^\alpha, [p]^\alpha\right) \to \left([w_r^*]^\alpha, [p]^\alpha\right)$.

The gradient technique is utilized in order to train the Z-number weight $(a_{\iota,\kappa}, p) = \left((\underline{a}_{\iota,\kappa}, \overline{a}_{\iota,\kappa}), p\right)$. $\frac{\partial(\Phi_r, p)}{\partial(\underline{a}_{\iota,\kappa}, p)}$ as well as $\frac{\partial(\Phi_r,p)}{\partial(\overline{a}_{\iota,\kappa},p)}$ are computed as follows

$$\begin{aligned}\frac{\partial(\Phi_r,p)}{\partial(\underline{a}_{\iota,\kappa},p)} &= \frac{\partial(\underline{\Phi}^\alpha,\underline{P}^\alpha(\Phi))}{\partial(\underline{w}_r^\alpha,\underline{w}_r^\alpha p(\underline{\varsigma}_r^\alpha))}\frac{\partial(\underline{w}_r^\alpha,\underline{w}_r^\alpha p(\underline{\varsigma}_r^\alpha))}{\partial(\underline{a}_{\iota,\kappa}^\alpha,\underline{a}_{\iota,\kappa}^\alpha p(\underline{\varsigma}_{\iota,\kappa}^\alpha))}\frac{\partial(\underline{a}_{\iota,\kappa}^\alpha,\underline{a}_{\iota,\kappa}^\alpha p(\underline{\varsigma}_{\iota,\kappa}^\alpha))}{\partial(\underline{a}_{\iota,\kappa},p)}\\
&+\frac{\partial(\overline{\Phi}^\alpha,\overline{P}^\alpha(\Phi))}{\partial(\overline{w}_r{}^\alpha,\overline{w}_r^\alpha p(\overline{\varsigma}_r^\alpha))}\frac{\partial(\overline{w}_r{}^\alpha,\overline{w}_r^\alpha p(\overline{\varsigma}_r^\alpha))}{\partial(\underline{a}_{\iota,\kappa}^\alpha,\underline{a}_{\iota,\kappa}^\alpha p(\underline{\varsigma}_{\iota,\kappa}^\alpha))}\frac{\partial(\underline{a}_{\iota,\kappa}^\alpha,\underline{a}_{\iota,\kappa}^\alpha p(\underline{\varsigma}_{\iota,\kappa}^\alpha))}{\partial(\underline{a}_{\iota,\kappa},p)}\\
&= -\left((\underline{w}_r^{*\alpha},\underline{w}_r^{*\alpha}p(\underline{\varsigma}_r^{*\alpha}))-(\underline{w}_r^\alpha,\underline{w}_r^\alpha p(\underline{\varsigma}_r^\alpha))\right)\textstyle\sum_{\iota,\kappa\in A}(\underline{\Omega}_{\iota,\kappa}^\alpha,\\
&\underline{\Omega}_{\iota,\kappa}^\alpha p(\underline{\varsigma}_{\iota,\kappa}^\alpha))\Gamma-((\overline{w}_r^{*\alpha},\overline{w}_r^{*\alpha}p(\overline{\varsigma}_r^{*\alpha}))\\
&-(\overline{w}_r^\alpha,\overline{w}_r^\alpha p(\overline{\varsigma}_r^\alpha)))(\textstyle\sum_{\iota,\kappa\in B'}(\underline{\Omega}_{\iota,\kappa}^\alpha,\underline{\Omega}_{\iota,\kappa}^\alpha p(\underline{\varsigma}_{\iota,\kappa}^\alpha))\\
&+\textstyle\sum_{\iota,\kappa\in C'}(\overline{\Omega}_{\iota,\kappa}^\alpha,\overline{\Omega}_{\iota,\kappa}^\alpha p(\overline{\varsigma}_{\iota,\kappa}^\alpha)))\Gamma\end{aligned}\tag{38}$$

where $\Gamma=(1-\alpha)$. Also,

$$\begin{aligned}\frac{\partial(\Phi_r,p)}{\partial(\overline{a}_{\iota,\kappa},p)} &= \frac{\partial(\underline{\Phi}^\alpha,\underline{P}^\alpha(\Phi))}{\partial(\underline{w}_r^\alpha,\underline{w}_r^\alpha p(\underline{\varsigma}_r^\alpha))}\frac{\partial(\underline{w}_r^\alpha,\underline{w}_r^\alpha p(\underline{\varsigma}_r^\alpha))}{\partial(\overline{a}_{\iota,\kappa}^\alpha,\overline{a}_{\iota,\kappa}^\alpha p(\overline{\varsigma}_{\iota,\kappa}^\alpha))}\frac{\partial(\overline{a}_{\iota,\kappa}^\alpha,\overline{a}_{\iota,\kappa}^\alpha p(\overline{\varsigma}_{\iota,\kappa}^\alpha))}{\partial(\overline{a}_{\iota,\kappa},p)}\\
&+\frac{\partial(\overline{\Phi}^\alpha,\overline{P}^\alpha(\Phi))}{\partial(\overline{w}_r^\alpha,\overline{w}_r^\alpha p(\overline{\varsigma}_r^\alpha))}\frac{\partial(\overline{w}_r^\alpha,\overline{w}_r^\alpha p(\overline{\varsigma}_r^\alpha))}{\partial(\overline{a}_{\iota,\kappa}^\alpha,\overline{a}_{\iota,\kappa}^\alpha p(\overline{\varsigma}_{\iota,\kappa}^\alpha))}\frac{\partial(\overline{a}_{\iota,\kappa}^\alpha,\overline{a}_{\iota,\kappa}^\alpha p(\overline{\varsigma}_{\iota,\kappa}^\alpha))}{\partial(\overline{a}_{\iota,\kappa},p)}\\
&= -\left((\underline{w}_r^{*\alpha},\underline{w}_r^{*\alpha}p(\underline{\varsigma}_r^{*\alpha}))-(\underline{w}_r^\alpha,\underline{w}_r^\alpha p(\underline{\varsigma}_r^\alpha))\right)(\textstyle\sum_{\iota,\kappa\in B}(\overline{\Omega}_{\iota,\kappa}^\alpha,\\
&\overline{\Omega}_{\iota,\kappa}^\alpha p(\overline{\varsigma}_{\iota,\kappa}^\alpha))+\textstyle\sum_{\iota,\kappa\in C}(\underline{\Omega}_{\iota,\kappa}^\alpha,\underline{\Omega}_{\iota,\kappa}^\alpha p(\underline{\varsigma}_{\iota,\kappa}^\alpha)))\Gamma_1\\
&-((\overline{w}_r^{*\alpha},\overline{w}_r^{*\alpha}p(\overline{\varsigma}_r^{*\alpha}))-(\overline{w}_r{}^\alpha,\overline{w}_r^\alpha p(\overline{\varsigma}_r^\alpha)))\textstyle\sum_{\iota,\kappa\in A'}(\overline{\Omega}_{\iota,\kappa}^\alpha,\\
&\overline{\Omega}_{\iota,\kappa}^\alpha p(\overline{\varsigma}_{\iota,\kappa}^\alpha))\Gamma_1\end{aligned}\tag{39}$$

where $\Gamma_1=(1-\alpha)$.

The Z-number coefficient $(a_{\iota,\kappa},p)$ is updated as

$$\begin{aligned}(\underline{a}_{\iota,\kappa},p)(r+1)&=(\underline{a}_{\iota,\kappa},p)(r)-\eta\frac{\partial(\Phi_r,p)}{\partial(\underline{a}_{\iota,\kappa},p)}\\
(\overline{a}_{\iota,\kappa},p)(r+1)&=(\overline{a}_{\iota,\kappa},p)(r)-\eta\frac{\partial(\Phi_r,p)}{\partial(\overline{a}_{\iota,\kappa},p)}\end{aligned}\tag{40}$$

where $\eta$ is the training rate $\eta>0$. In order to increase the training process, a momentum term is added as

$$\begin{aligned}(\underline{a}_{\iota,\kappa},p)(r+1)&=(\underline{a}_{\iota,\kappa},p)(r)-\eta\frac{\partial(\Phi_r,p)}{\partial(\underline{a}_{\iota,\kappa},p)}\\
&+\gamma\left[(\underline{a}_{\iota,\kappa},p)(r)-(\underline{a}_{\iota,\kappa},p)(r-1)\right]\\
(\overline{a}_{\iota,\kappa},p)(r+1)&=(\overline{a}_{\iota,\kappa},p)(r)-\eta\frac{\partial(\Phi_r,p)}{\partial(\overline{a}_{\iota,\kappa},p)}\\
&+\gamma\left[(\overline{a}_{\iota,\kappa},p)(r)-(\overline{a}_{\iota,\kappa},p)(r-1)\right]\end{aligned}\tag{41}$$

where $\gamma>0$.

**Remark 1.** One of the primary advantages of the least mean square index (37) is having a self-correcting feature which allows operating for an arbitrarily long period without deviating from its constraints. The corresponding gradient algorithm is susceptible to cumulative round off errors and is appropriate for long runs without an extra error-correction process. It is more robust in statistics, identification as well as signal processing [36].

**Learning algorithm**

1) Step 1: Select the training rates $\eta>0$, $\gamma>0$, also the stopping criterion $\overline{\Phi}>0$. The initial Z-number vector $A=(a_{1,1},...,a_{\iota,\kappa})$ is chosen arbitrarily. The initial learning iteration is $r=1$, and the initial learning error is $\Phi=0$.
2) The following steps should be repeated for $\alpha=\alpha_1,...,\alpha_m$, till all the training data are implemented
   a) Forward computation: Compute the $\alpha$-level of the Z-number output $w_r$ using the $\alpha$-level of the Z-number input vectors $(\vartheta_r,\upsilon_r)$, and the Z-number connection weight $A$.
   b) Back-propagation: Adjust Z-number parameters $a_{\iota,\kappa}$, $\iota=1,...,\gamma$, $\kappa=1,...,\delta$, utilizing the cost function for the $\alpha$-level of the Z-number output $w_r$, and the Z-number target output $w_r^*$.
   c) Stopping criterion: Compute the cycle error $\Phi_r$, $\Phi=\Phi+\Phi_r$. $r=r+1$. In a case that $\Phi>\overline{\Phi}$, let $\Phi=0$, a new training cycle is begun. Go to (a).

## IV. Upper bounds of the modeling errors

This section extends some important estimation theorems into fuzzy equation modeling. Initially, the modeling error is defined for Z-numbers.

**Definition 6.** The distance between two Z-numbers, $\psi,\varphi\in\widetilde{Z}$, is described as the Hausdorff metric $D_H[\psi,\varphi]$,

$$\begin{aligned}D_H[\psi,\varphi]=\max\{&\sup_{(a_1,c_1)\in\psi}\inf_{(a_2,c_2)\in\varphi}(d(a_1,a_2)\\
&+d(c_1,c_2)),\sup_{(a_1,c_1)\in\varphi}\inf_{(a_2,c_2)\in\psi}(d(a_1,a_2)+d(c_1,c_2)\}\end{aligned}\tag{42}$$

$d(a,c)$ is the supremum metrics defined for fuzzy sets.

**Lemma 1.** Let $\varpi\subset\widetilde{Z}$ is a compact set, in this case, $\varpi$ is uniformly support-bounded, i.e. there exists a compact set $\Psi\subset\Re$, in such a way that $\forall\psi\in\varpi$,

$$Supp(\psi)\subset\Psi.\tag{43}$$

**Lemma 2.** Suppose $\psi,\varphi\in\widetilde{Z}$, also $\alpha\in(0,1]$, in this case the following is concluded: (i) if $f,g:\Re\to\Re$ are continuous, $[f(\psi)g(\varphi)]^\alpha=f([\psi^\alpha])g([\varphi^\alpha])$ holds; (ii) if $f,g:\Re\to\Re$ are continuous, so $f(Supp(\psi))g(Supp(\varphi))=Supp(f(\psi))Supp(g(\varphi))$.

**proof.** As (i) concludes from [43], so just (ii) will be proved. Initially, $\overline{f(D_1)g(D_2)}=f(\overline{D}_1)g(\overline{D}_2)$ is shown for $D_1,D_2\subset\Re$. As $f(D_1)g(D_2)\subset f(\overline{D}_1)g(\overline{D}_2)$, also $f(\overline{D}_1)g(\overline{D}_2)$ is closed by the continuity of $f$ as well as $g$, therefore, $\overline{f(D_1)g(D_2)}\subset f(\overline{D}_1)g(\overline{D}_2)$. Also, for randomly given $\theta\in f(\overline{D}_1)g(\overline{D}_2)$, there exists a sequence $\{a_nc_n|n\in N\}\subset\Re$, $a,c\in\Re$, in such a way that $a_nc_n\to ac$ $(n\to+\infty)$, $\theta=f(a)g(c)$. Since $f$ as well as $g$ are continues so, $lim_{n\to+\infty}f(a_n)g(c_n)=f(a)g(c)=\theta$. However, $f(a_n)g(c_n)\in f(D_1)g(D_2)$, therefore, $\theta\in\overline{f(D_1)g(D_2)}$. So $f(\overline{D}_1)g(\overline{D}_2)\subset\overline{f(D_1)g(D_2)}$. Hence $\overline{f(D_1)g(D_2)}=f(\overline{D}_1)g(\overline{D}_2)$.

Since,

$$\begin{aligned}Supp(f(\psi))Supp(g(\varphi))&=\overline{\{\theta\in\Re|(f(\psi)g(\varphi))(\theta)>0\}}\\
f(Supp(\psi))g(Supp(\varphi))&=fg\left(\overline{\{ac\in\Re|\psi\varphi(ac)>0\}}\right)\end{aligned}\tag{44}$$

the following is concluded

$$\begin{aligned}f(Supp(\psi))g(Supp(\varphi))&=\overline{fg(\{ac\in\Re|\psi\varphi(ac)>0\})}\\
&=\overline{\{fg(ac)\in\Re|\psi\varphi(ac)>0\}}\end{aligned}\tag{45}$$

As, $\{\theta\in\Re|(f(\psi)g(\varphi))(\theta)>0\}=\{fg(ac)|\psi\varphi(ac)>0\}$, so

$$Supp(f(\psi))Supp(g(\varphi))=f(Supp(\psi))g(Supp(\varphi))\tag{46}$$

**Lemma 3.** Suppose $G\subset\Re$ is a compact set, also $f_1g_1$ as well as $f_2g_2$ are continuous on $G,\zeta>0$, in addition

$$|f_1(a)g_1(c)-f_2(a)g_2(c)|<\zeta,\quad\forall a,c\in G\tag{47}$$

Hence $|\sup_{a,c\in G_1}f_1(a)g_1(c)-\sup_{a,c\in G_1}f_2(a)g_2(c)|<\zeta$ is valid for each compact set $G_1\subset G$.

**proof.** As $G_1$ is a compact set also, $f_1g_1$ as well as $f_2g_2$ are continuous on $G_1$, so there exist $a_0, c_0 \in G_1$, $\hat{a}_0, \hat{c}_0 \in G_1$, in such a way that

$$\begin{aligned} f_1(a_0)g_1(c_0) = \sup_{a,c\in G_1} f_1(a)g_1(c) \\ f_2(\hat{a}_0)g_2(\hat{c}_0)) = \sup_{a,c\in G_1} f_2(a)g_2(c) \end{aligned} \quad (48)$$

Let $|f_1(a_0)g_1(c_0) - f_2(\hat{a}_0)g_2(\hat{c}_0))| \geq \zeta$, so

$$\begin{aligned} f_1(a_0)g_1(c_0) - f_2(\hat{a}_0)g_2(\hat{c}_0)) \leq -\zeta, \\ or \quad f_1(a_0)g_1(c_0) - f_2(\hat{a}_0)g_2(\hat{c}_0)) \geq \zeta \end{aligned} \quad (49)$$

In the first case of (49), since $f_1(\hat{a}_0)g_1(\hat{c}_0) \leq f_1(a_0)g_1(c_0)$,

$$\begin{aligned} f_1(\hat{a}_0)g_1(\hat{c}_0) - f_2(\hat{a}_0)g_2(\hat{c}_0)) \\ \leq f_1(a_0)g_1(c_0) - f_2(\hat{a}_0)g_2(\hat{c}_0)) \leq -\zeta \\ \Rightarrow |f_1(\hat{a}_0)g_1(\hat{c}_0) - f_2(\hat{a}_0)g_2(\hat{c}_0))| \geq \zeta \end{aligned} \quad (50)$$

holds, that contradicts (47). In the second case (49), as $f_2(a_0)g_2(c_0)) \leq f_2(\hat{a}_0)g_2(\hat{c}_0))$, the following is obtained

$$\begin{aligned} f_1(a_0)g_1(c_0) - f_2(a_0)g_2(c_0)) \\ \geq f_1(a_0)g_1(c_0) - f_2(\hat{a}_0)g_2(\hat{c}_0)) \geq \zeta \\ \Rightarrow |f_1(a_0)g_1(c_0) - f_2(a_0)g_2(c_0))| \geq \zeta \end{aligned} \quad (51)$$

that contradicts (47). Hence, (49) is not correct, so $-\zeta < f_1(a_0)g_1(c_0) - f_2(\hat{a}_0)g_2(\hat{c}_0)) < \zeta$, hence $|f_1(a_0)g_1(c_0) - f_2(a_0)g_2(c_0))| < \zeta$, i.e. $|\sup_{a,c\in G_1} f_1(a)g_1(c) - \sup_{a,c\in G_1} f_2(a)g_2(c)| < \zeta$. The proof is finalized.

**Theorem 1.** Suppose $fg : \Re \to \Re$ is a continuous function, so for each compact set $\varpi \subset \widetilde{Z}_0$ (the set of all the bounded Z-number set) also $\chi > 0$, there exists $a_{\iota\kappa} \in \widetilde{Z}_0$, $\iota = 1, 2, ..., \gamma$, $\kappa = 1, 2, ..., \delta$, in such a manner that

$$\begin{aligned} d(f(\tilde{\vartheta})g(\tilde{\upsilon}), \textstyle\sum_{\iota=1}^{\gamma} \sum_{\kappa=1}^{\delta} f_\iota(\vartheta)g_\kappa(\upsilon)a_{\iota\kappa}) < \chi \\ \forall\vartheta, \upsilon \in \varpi, \ \forall\tilde{\vartheta}, \tilde{\upsilon} \in \Re \end{aligned} \quad (52)$$

in which $\chi$ is a finite number.

**proof.** The subsequent results will lead to the proof of the Theorem.

Suppose $fg : \Re \to \Re$, $fg$ can be extended by the extension principle to the Z-number function that is stated as $fg : \widetilde{Z}_0 \to \widetilde{Z}$, also

$$\begin{aligned} f(\psi_1)(\omega_1)g(\psi_2)(\omega_2) = \textstyle\bigvee_{f(\vartheta)=\omega_1} \bigvee_{g(\upsilon)=\omega_2} \{\psi_1(\vartheta)\}\{\psi_2(\upsilon)\}, \\ \forall\psi_1, \psi_2 \in \widetilde{Z}_0, \quad \omega_1, \omega_2 \in \Re \end{aligned} \quad (53)$$

$fg$ is termed as the extended function. Furthermore, $cc(\Re)$ is the set of bounded closed intervals of $\Re$. Clearly

$$\psi_1, \psi_2 \in \widetilde{Z}_0 \Longrightarrow \forall\alpha \in (0, 1], \quad [\psi_1]^\alpha, [\psi_2]^\alpha \in cc(\Re) \quad (54)$$

Also,

$$Supp(\psi_1), Supp(\psi_2) \in cc(\Re) \quad (55)$$

Hence,

$$\begin{aligned} Supp(\psi_1) = [s_1(\psi_1), s_2(\psi_1)] \\ Supp(\psi_2) = [s_1(\psi_2), s_2(\psi_2)] \end{aligned} \quad (56)$$

**Theorem 2.** If $fg : \Re \to \Re$ is a continuous function, so for each compact set $\varpi \subset \widetilde{Z}_0$, $\tau > 0$ also arbitrary $\varepsilon > 0$, there exists $a_{\iota\kappa} \in \widetilde{Z}_0$, $\iota = 1, 2, ..., \gamma$, $\kappa = 1, 2, ..., \delta$, in such a way that

$$d(f(\vartheta)g(\upsilon), \sum_{\iota=1}^{\gamma} \sum_{\kappa=1}^{\delta} f_\iota(\vartheta)g_\kappa(\upsilon)a_{\iota\kappa}) < \tau, \quad \forall\vartheta, \upsilon \in \varpi \quad (57)$$

in which, $\tau$ is termed as a finite number. The lower as well as the upper limits of the $\alpha$-level set of Z-number function approach to $\tau$, however, the center approaches to $\varepsilon$.

**proof.** Since $\varpi \subset \widetilde{Z}_0$ is a compact set, so by implementing Lemma 1, $\Psi \subset \Re$ is the compact set corresponding to $\varpi$. Furthermore, $\forall\varepsilon > 0$, also using the results of [12], there exists $a_{\iota\kappa} \in \Re$, $\iota = 1, 2, ..., \gamma$, $\kappa = 1, 2, ..., \delta$, in such a way that

$$|f(\vartheta)g(\upsilon) - \sum_{\iota=1}^{\gamma} \sum_{\kappa=1}^{\delta} f_\iota(\vartheta)g_\kappa(\upsilon)a_{\iota\kappa}| < \varepsilon, \quad \forall\vartheta, \upsilon \in \Psi \quad (58)$$

If $\tilde{f}(\vartheta)\tilde{g}(\upsilon) = \sum_{\iota=1}^{\gamma} \sum_{\kappa=1}^{\delta} f_\iota(\vartheta)g_\kappa(\upsilon)a_{\iota\kappa}$, where $\vartheta, \upsilon \in \Re$, hence

$$|f(\vartheta)g(\upsilon) - \tilde{f}(\vartheta)\tilde{g}(\upsilon)| < \varepsilon, \quad \forall\vartheta, \upsilon \in \Psi \quad (59)$$

Theorem 3 will lead to the validation of (57).

**Theorem 3.** Suppose $\varpi \subset \widetilde{Z}_0$ is compact, also $\Psi$ is the corresponding compact set of $\varpi$ and $fg, \hat{f}\hat{g} : \Re \to \Re$ are the continuous functions that,

$$|f(\vartheta)g(\upsilon) - \hat{f}(\vartheta)\hat{g}(\upsilon)| < \zeta, \quad \forall\vartheta, \upsilon \in \Psi \quad (60)$$

where $\zeta > 0$. So, $\forall\psi_1, \psi_2 \in \varpi$, $d(f(\psi_1)g(\psi_2) - \hat{f}(\psi_1)\hat{g}(\psi_2)) \leq \zeta$.

**proof.** Assume $\psi \in \widetilde{Z}$ also $\alpha \in (0, 1]$. Since $fg$ as well as $\hat{f}\hat{g}$ are continuous, then $[f(\psi_1)g(\psi_2)]^\alpha = f([\psi_1^\alpha])g([\psi_2^\alpha])$, and $[\hat{f}(\psi_1)\hat{g}(\psi_2)]^\alpha = \hat{f}([\psi_1^\alpha])\hat{g}([\psi_2^\alpha])$ are valid by Lemma 2. Hence, the following result is obtained using [34],

$$\begin{aligned} D_H([f(\psi_1)g(\psi_2)]^\alpha - [\hat{f}(\psi_1)\hat{g}(\psi_2)]^\alpha) \\ = D_H(f([\psi_1^\alpha])g([\psi_2^\alpha]) - \hat{f}([\psi_1^\alpha])\hat{g}([\psi_2^\alpha])) \\ = \sup_{|n|=1}\{|s(n, f([\psi_1^\alpha])g([\psi_2^\alpha])) - s(n, \hat{f}([\psi_1^\alpha])\hat{g}([\psi_2^\alpha]))|\} \end{aligned} \quad (61)$$

Since $n \in \Re$: $|n| = 1$, the following is valid,

$$\begin{aligned} |s(n, f([\psi_1^\alpha])g([\psi_2^\alpha])) - s(n, \hat{f}([\psi_1^\alpha])\hat{g}([\psi_2^\alpha]))| \\ = |\sup\{nj | j \in f([\psi_1^\alpha])g([\psi_2^\alpha])\} \\ - \sup\{nj | j \in \hat{f}([\psi_1^\alpha])\hat{g}([\psi_2^\alpha])\}| \\ = |\sup\{nf(\vartheta)g(\upsilon) | \vartheta \in [\psi_1]^\alpha, \upsilon \in [\psi_2]^\alpha\} \\ - \sup\{n\hat{f}(\vartheta)\hat{g}(\upsilon) | \vartheta \in [\psi_1]^\alpha, \upsilon \in [\psi_2]^\alpha\}| \end{aligned} \quad (62)$$

Also, taking into consideration the conditions in the theorem, the following is obtained

$$\begin{aligned} |nf(\vartheta)g(\upsilon) - n\hat{f}(\vartheta)\hat{g}(\upsilon)| = |f(\vartheta)g(\upsilon) - \hat{f}(\vartheta)\hat{g}(\upsilon)| \\ < \zeta, \quad \forall\vartheta \in [\psi_1]^\alpha, \upsilon \in [\psi_2]^\alpha \end{aligned} \quad (63)$$

Hence, using (61), (62) as well as Lemma 3, the following relation is extracted which proves the theorem,

$$\begin{aligned} D_H([f(\psi_1)]^\alpha[g(\psi_2)]^\alpha, [\hat{f}(\psi_1)]^\alpha[\hat{g}(\psi_2)]^\alpha) < \zeta \\ \Rightarrow d(f(\psi_1)g(\psi_2), \hat{f}(\psi_1)\hat{g}(\psi_2)) \\ = \sup_{\alpha\in(0,1]}\{D_H([f(\psi_1)]^\alpha[g(\psi_2)]^\alpha, [\hat{f}(\psi_1)]^\alpha)[\hat{g}(\psi_2)]^\alpha)\} \\ \leq \zeta \end{aligned} \quad (64)$$

## V. NUMERICAL EXAMPLES

In this section, some examples with applications are given to show how to apply fuzzy equations to model uncertain nonlinear systems.

**Example 1** In order to produce electricity from sun the photo voltaic cells (PV cells) are used. The PV cells are designed in a parallel form. According to the changes in the position of sun and climate situation the sun radiation changes and causes the changes in the produced current and voltage by PV cells. The total produced power of power station is defined as [19],

$$O = \varphi_1 I_1 V_1 \oplus \varphi_2 I_2 V_2 \oplus \varphi_3 I_3 V_3 \oplus \varphi_4 I_4 V_4 \quad (65)$$

where $I_1 = \sqrt{2}\vartheta$, $I_2 = \vartheta^2$, $I_3 = \frac{\vartheta}{2}$, $I_4 = \vartheta$ are the currents and $V_1 = e^{\upsilon}$, $V_2 = 2\upsilon$, $V_3 = \sqrt{\upsilon}$, $V_4 = \sqrt{3}\upsilon$ are the voltages produced by PV cells. $\vartheta$ and $\upsilon$ are the elapsed time. $\varphi_1, \varphi_2, \varphi_3$, and $\varphi_4$ are the characteristic coefficients of the PV cells, which are satisfied in the triangular Z-number uncertainty,

$$\begin{aligned} \varphi_1 &= [(2,5,7), p(0.81,0.9,0.95)], \\ \varphi_2 &= [(1,4,8), p(0.72,0.85,0.9)], \\ \varphi_3 &= [(3,6,8), p(0.82,0.86,0.91)], \\ \varphi_4 &= [(6,8,11), p(0.82,0.85,0.92)] \end{aligned} \quad (66)$$

The diodes and buck-boost convector are utilized in order to connect the PV cell to ultra capacitor and energy storage, see Figure 2.

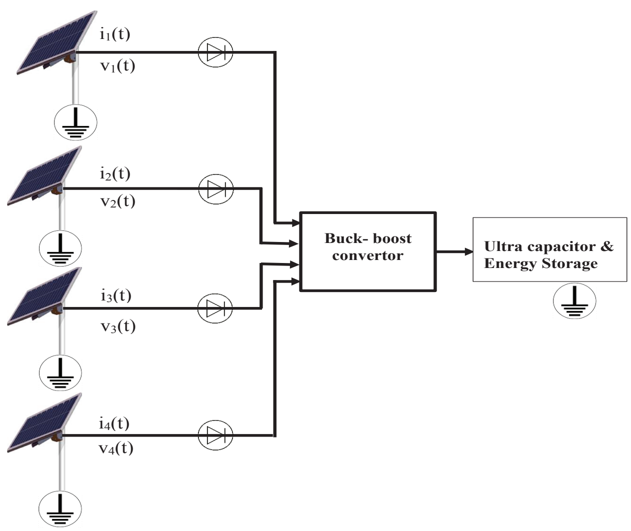


Fig. 2. PV cells in power station

Four types of inputs are utilized for training the neural network, where weights are considered as the Z-number parameters of (65). The input data are given as

$$\vartheta = \left\{ \begin{array}{c} [(2,6,7), p(0.7,0.8,0.9)], \\ [(11,14,19), p(0.7,0.8,0.91)], \\ [(9,11,13), p(0.8,0.86,0.9)], \\ [(1,2,3), p(0.8,0.9,0.95)] \end{array} \right\}$$

$$\upsilon = \left\{ \begin{array}{c} [(1,2,5), p(0.7,0.8,0.89)], \\ [(10,12,14), p(0.8,0.9,0.94)], \\ [(9,11,15), p(0.7,0.8,0.92)], \\ [(3,6,7), p(0.8,0.9,0.95)] \end{array} \right\}$$

The inputs, $\vartheta$, and $\upsilon$ are implemented to (65), and the corresponding outputs data are obtained as

$$O = \left\{ \begin{array}{c} [(607.54, 643.87, 695.12), p(0.7,0.8,0.91)], \\ [(3788.76, 4011.86, 4457.16), p(0.7,0.82,0.9)], \\ [(2453.52, 2874.83, 3287.77), p(0.8,0.86,0.9)], \\ [(887.76, 921.65, 966.88), p(0.8,0.9,0.95)] \end{array} \right\}$$

**Table 1.** *Estimation of Z-number coefficients with neural network method*

| $r$ | $\varphi_1$ |
|---|---|
| 1 | $[(6.27, 9.17, 11.19), p(0.6, 0.8, 0.85)]$ |
| 2 | $[(5.91, 8.75, 10.64), p(0.72, 0.8, 0.87)]$ |
| ⋮ | ⋮ |
| 70 | $[(2.02, 5.01, 7.03), p(0.81, 0.9, 0.95)]$ |
| $r$ | $\varphi_2$ |
| 1 | $[(5.36, 8.28, 12.35), p(0.71, 0.8, 0.86)]$ |
| 2 | $[(4.89, 7.77, 11.85), p(0.8, 0.86, 0.9)]$ |
| ⋮ | ⋮ |
| 70 | $[(1.06, 4.05, 8.03), p(0.85, 0.9, 0.94)]$ |
| $r$ | $\varphi_3$ |
| 1 | $[(7.43, 10.49, 12.37), p(0.6, 0.81, 0.87)]$ |
| 2 | $[(6.82, 9.91, 11.87), p(0.7, 0.8, 0.92)]$ |
| ⋮ | ⋮ |
| 70 | $[(3.04, 6.06, 8.05), p(0.82, 0.86, 0.91)]$ |
| $r$ | $\varphi_4$ |
| 1 | $[(10.23, 12.19, 15.29), p(0.6, 0.7, 0.8)]$ |
| 2 | $[(9.81, 11.74, 14.79), p(0.71, 0.8, 0.85)]$ |
| ⋮ | ⋮ |
| 70 | $[(6.04, 8.08, 11.07), p(0.82, 0.85, 0.92)]$ |

These input-output pairs have been utilized for training the neural network. The weights are $\varphi_1$, $\varphi_2$, $\varphi_3$, and $\varphi_4$. The learning rates are $\eta = 0.01$ and $\gamma = 0.01$. The neural network starts from,

$$\begin{aligned} \varphi_1(0) &= [(6.50, 9.50, 11.50), p(0.26, 0.27, 0.28)], \\ \varphi_2(0) &= [(5.50, 8.50, 12.50), p(0.28, 0.29, 0.298)], \\ \varphi_3(0) &= [(7.50, 10.50, 12.50), p(0.16, 0.18, 0.20)], \\ \varphi_4(0) &= [(10.50, 12.50, 15.50), p(0.32, 0.33, 0.38)] \end{aligned} \quad (67)$$

Using equation (14) for the first iteration ($r = 1$) and at $\alpha$-cut$=$ 0, the reliability of $(6.50, 9.50, 11.50)$ is obtained as below,

$$\begin{aligned} \underline{P}^{\alpha} &= (6.50)(0.01) = 0.065 \\ \overline{P}^{\alpha} &= (11.50)(0.03) = 0.345 \end{aligned} \quad (68)$$

where the probability density of 6.50 is 0.01 and the probability density of 11.50 is 0.03 as a supposition. Table 1 demonstrates the training outcomes. It can be seen that the Z-number coefficients approach to their actual quantity after 70 iterations. In this Table, $r$ is taken to be the number of iterations. In order to obtain a minor estimated error, the number of iterations should be increased. Table 2 shows the accuracy of the neural network method in obtaining the Z-number coefficients (for $\alpha = 0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1$). In order to compare our outcomes, the fuzzy cubic spline (FCS) method is used [16], [28]. In general, the major advantage of FCS technique is its calculation simplicity. The results of the FCS method is demonstrated in Table 3. Both the neural network based algorithm (the proposed method in this paper) and the FCS technique can estimate the solutions of the fuzzy equations. The approximated errors of the neural network based algorithm are less than the FCS method.

**Table 2.** *The accuracy of the neural network method in obtaining the Z-number coefficients for r=70*

| $\alpha$ | $Error$ |
|---|---|
| 0 | $[(0.00406, 0.00677), p(0.7, 0.8, 0.9)]$ |
| 0.1 | $[(0.00341, 0.00691), p(0.8, 0.85, 0.9)]$ |
| 0.2 | $[(0.00346, 0.00527), p(0.8, 0.9, 1)]$ |
| 0.3 | $[(0.00685, 0.01252), p(0.75, 0.8, 0.9)]$ |
| 0.4 | $[(0.00573, 0.00716), p(0.8, 0.9, 1)]$ |
| 0.5 | $[(0.00481, 0.00702), p(0.7, 0.8, 0.9)]$ |
| 0.6 | $[(0.00352, 0.00572), p(0.75, 0.8, 0.9)]$ |
| 0.7 | $[(0.00333, 0.00532), p(0.75, 0.8, 0.9)]$ |
| 0.8 | $[(0.00252, 0.00513), p(0.8, 0.92, 1)]$ |
| 0.9 | $[(0.00281, 0.00551), p(0.8, 0.92, 1)]$ |
| 1 | $[(0.00385, 0.00524), p(0.75, 0.8, 0.9)]$ |

**Table 3.** *Estimation of Z-number coefficients with FCS method*

| $r$ | $\varphi_1$ |
|---|---|
| 1 | $[(6.07, 9.02, 11.11), p(0.6, 0.81, 0.87)]$ |
| 2 | $[(5.84, 8.91, 10.81), p(0.7, 0.8, 0.92)]$ |
| ⋮ | ⋮ |
| 110 | $[(2.11, 5.13, 7.09), p(0.81, 0.9, 0.95)]$ |
| $r$ | $\varphi_2$ |
| 1 | $[(5.11, 8.01, 12.09), p(0.71, 0.82, 0.92)]$ |
| 2 | $[(4.76, 7.69, 11.79), p(0.72, 0.8, 0.92)]$ |
| ⋮ | ⋮ |
| 110 | $[(1.16, 4.15, 8.13), p(0.8, 0.85, 0.95)]$ |
| $r$ | $\varphi_3$ |
| 1 | $[(7.11, 10.09, 12.06), p(0.6, 0.8, 0.85)]$ |
| 2 | $[(6.83, 9.81, 11.76), p(0.8, 0.86, 0.9)]$ |
| ⋮ | ⋮ |
| 110 | $[(3.19, 6.14, 8.16), p(0.82, 0.9, 0.94)]$ |
| $r$ | $\varphi_4$ |
| 1 | $[(10.03, 12.02, 15.05), p(0.7, 0.8, 0.87)]$ |
| 2 | $[(9.83, 11.81, 14.74), p(0.71, 0.8, 0.9)]$ |
| ⋮ | ⋮ |
| 110 | $[(6.16, 8.18, 11.19), p(0.82, 0.85, 0.92)]$ |

In order to change the Z-number into the fuzzy number, the following formula is utilized,

$$\nu = \frac{\int \varphi \pi_P(\varphi) d\varphi}{\int \pi_P(\varphi) d\varphi} \qquad (69)$$

Let $Z = [(6.27, 9.17, 11.19), p(0.6, 0.8, 0.85)]$, so $Z^{\nu} = (6.27, 9.17, 11.19; 0.7)$ therefore, $Z' = (\sqrt{0.7}\ 6.27, \sqrt{0.7}\ 9.17, \sqrt{0.7}\ 11.19)$. The results of Table 1 and Table 3 based of fuzzy numbers are demonstrated in Table 4 and Table 5, respectively.

**Table 4.** *Estimation of fuzzy number coefficients with neural network method*

| $r$ | $\varphi_1$ | $\varphi_2$ |
|---|---|---|
| 1 | $(5.24, 7.67, 9.36)$ | $(4.48, 6.92, 10.33)$ |
| 2 | $(4.94, 7.32, 8.91)$ | $(4.09, 6.51, 9.91)$ |
| ⋮ | ⋮ | ⋮ |
| 70 | $(1.69, 4.19, 5.88)$ | $(0.88, 3.38, 6.71)$ |
| $r$ | $\varphi_3$ | $\varphi_4$ |
| 1 | $(6.21, 8.77, 10.34)$ | $(8.55, 10.19, 12.79)$ |
| 2 | $(5.71, 8.29, 9.93)$ | $(8.21, 9.82, 12.37)$ |
| ⋮ | ⋮ | ⋮ |
| 70 | $(2.54, 5.07, 6.73)$ | $(5.05, 6.76, 9.26)$ |

**Table 5.** *Estimation of fuzzy number coefficients with FCS method*

| $r$ | $\varphi_1$ | $\varphi_2$ |
|---|---|---|
| 1 | $(5.07, 7.54, 9.29)$ | $(4.27, 6.71, 10.11)$ |
| 2 | $(4.88, 7.45, 9.04)$ | $(4.01, 6.43, 9.86)$ |
| ⋮ | ⋮ | ⋮ |
| 110 | $(1.76, 4.29, 5.93)$ | $(0.97, 3.47, 6.81)$ |
| $r$ | $\varphi_3$ | $\varphi_4$ |
| 1 | $(5.94, 8.44, 10.09)$ | $(8.39, 10.05, 12.59)$ |
| 2 | $(5.71, 8.21, 9.83)$ | $(8.22, 9.88, 12.33)$ |
| ⋮ | ⋮ | ⋮ |
| 110 | $(2.66, 5.13, 6.82)$ | $(5.15, 6.84, 9.36)$ |

The Z-number $Z = (c, p) = [(6.27, 9.17, 11.19), p(0.6, 0.8, 0.85)]$ and the fuzzy number $(5.24, 7.67, 9.36)$ are compared and the results are shown in Figure 3. It can be noticed from this figure that the Z-number has an advantage to fuzzy number in having more and various information, also the solution generated by Z-number is more precise. The membership function for the restriction in the Z-number is in probability form and is stated as $\mu_{c_Z} = (6.27, 9.17, 11.19)$.

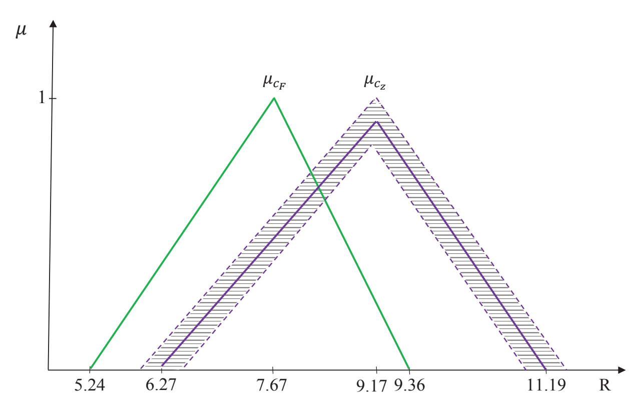


Fig. 3. Comparison between Z-number and fuzzy number

**Example 2** In order to transfer the food and astronaut to space station two space shuttles are used. During the traveling, the weights and velocities of the shuttles change. By reaching the shuttles to the space station, the collision occurs between shuttles and station. The produced momentum after the collision is defined as [15],

$$S = A_1 M_1 V_1 \oplus A_2 M_2 V_2 \quad (70)$$

where $M_1 = \sqrt{\vartheta}, M_2 = \vartheta^2$ are the masses, and $V_1 = \sqrt{2\upsilon}, V_2 = \upsilon$ are the velocities of shuttles. $A_1$, and $A_2$ are the characteristic coefficients of the momentum of shuttles.

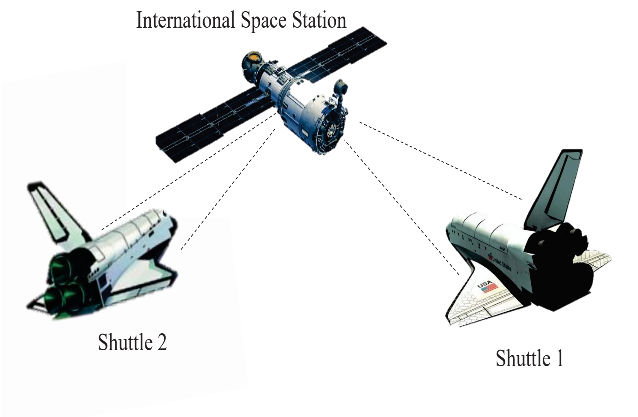


Fig. 4. International space station

The real parameters are stated as

$$\begin{array}{c} A_1 = [(1,5,8), p(0.71,0.8,0.9)], \\ A_2 = [(7,13,16), p(0.8,0.9,95)] \end{array} \quad (71)$$

Two types of inputs are utilized for training the neural network, where weights are considered as the Z-number parameters of (70). The input data are given as

$$\vartheta = \left\{ \begin{array}{c} [(2,4,8), p(0.8,0.9,0.95)], \\ [(8,10,12), p(0.7,0.8,0.9)] \end{array} \right\}$$

$$\upsilon = \left\{ \begin{array}{c} [(3,5,9), p(0.7,0.8,0.9)], \\ [(7,10,13), p(0.8,0.9,0.96)] \end{array} \right\}$$

The inputs, $\vartheta$, and $\upsilon$ are implemented to (70), and the corresponding outputs data are obtained as

$$S = \left\{ \begin{array}{c} [(401.32, 452.87, 500.01), p(0.7,0.8,0.91)], \\ [(1307.76, 1798.86, 2065.23), p(0.7,0.82,0.9)] \end{array} \right\}$$

These input-output pairs have been utilized for training the neural network. The weights are $A_1$, and $A_2$. The learning rates are $\eta = 0.01$, and $\gamma = 0.01$, also $\alpha = 0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1$. The neural network starts from,

$$\begin{array}{c} A_1(0) = [(5.50, 9.48, 12.37), p(0.20, 0.22, 0.23)], \\ A_2(0) = [(11.46, 17.50, 20.39), p(0.26, 0.27, 0.28)] \end{array} \quad (72)$$

We use neural network method to approximate $A_1$ and $A_2$. The comparison results between the neural network method and the FCS method are shown in Figure 5. In this figure, it can be seen that the estimated errors of the neural network method are less when compared with the FCS method. Furthermore, the neural network method approaches to the real solution more faster than the FCS method. The FCS method initially is not robust.

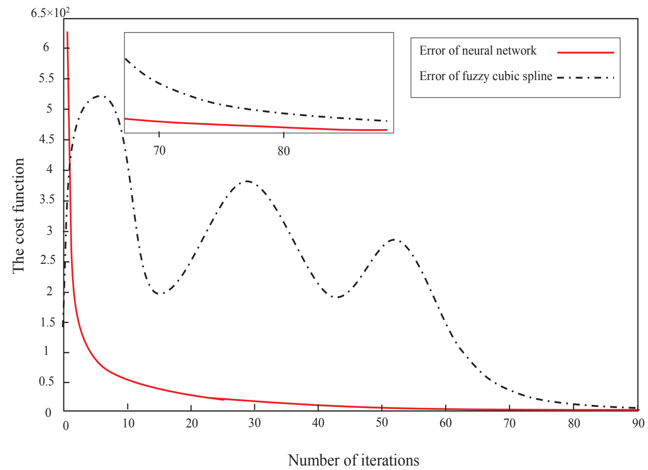


Fig. 5. Estimated errors of neural network method and FCS method

**Example 3** Consider the following system where all the input and output data are designed as Z-numbers [39],

$$\begin{array}{c} e(r) = \gamma cos(\zeta \Delta t r) \\ y(r+1) = \frac{\Delta t^2 [e(r) - \beta y^3(r)] - y(r-1) + \delta y(r)}{(1+\kappa \Delta t)} \end{array} \quad (73)$$

with $\delta = \kappa \Delta t - \alpha \Delta t^2 + 2$, where $\Delta t, \kappa, \alpha, \beta, \gamma$ are Z-number parameters, and $\zeta$ is a random variable. (73) is the discrete-time version of the well-known Duffing equation with,

$$\begin{array}{c} \Delta t = [(0.03, 0.05, 0.06), p(0.6, 0.8, 0.86)] \\ \kappa = [(0.1, 0.3, 0.5), p(0.6, 0.7, 0.87)] \\ \alpha = [(-4.2, -4, -3.8), p(0.6, 0.8, 86)] \\ \beta = [(0.8, 1, 1.2), p(0.7, 0.8, 0.85)] \\ \gamma = [(0.2, 0.5, 0.7), p(0.7, 0.8, 0.85)] \end{array} \quad (74)$$

$\zeta$ being a random variable uniformly distributed in the interval $[0.1, 2.9]$ with mean $E\{\zeta\} = 1.5$, and the initial conditions being $y(0) = y(1) = 1$. By substituting the value of Z-number parameters in (73) and choosing $\zeta$ randomly, $\zeta = 2.1053$, the following relation is extracted

$$y(r+1) = C_1 cos(pr) - C_2 y^3(r) - C_3 y(r-1) + C_4 y(r) \quad (75)$$

where,

$$\begin{array}{l} p = \zeta \Delta t = [(0.06315, 0.10526, 0.12631), p(0.7, 0.8, 0.9)] \\ C_1 = \frac{\Delta t^2 \gamma}{1+\kappa \Delta t} = [(0.00017, 0.00123, 0.00251), p(0.8, 0.85, 0.9)] \\ C_2 = \frac{\Delta t^2 \beta}{1+\kappa \Delta t} = [(0.00069, 0.00246, 0.00430), p(0.7, 0.8, 0.9)] \\ C_3 = \frac{1}{1+\kappa \Delta t} = [(0.97087, 0.98522, 0.99700), p(0.8, 0.85, 0.9)] \\ C_4 = \frac{\delta}{1+\kappa \Delta t} = [(1.94798, 1.99507, 2.03900), p(0.7, 0.8, 0.9)] \end{array} \quad (76)$$

The neural network is performed with training of $r = 50$ and starts from,

$$\begin{array}{l} C_1(0) = [(2.00036, 2.00141, 2.00271), p(0.16, 0.18, 0.20)] \\ C_2(0) = [(3.00193, 3.00372, 3.00555), p(0.32, 0.33, 0.38)] \\ C_3(0) = [(3.11021, 3.12451, 3.13619), p(0.26, 0.27, 0.28)] \\ C_4(0) = [(4.13722, 4.18437, 4.22831), p(0.20, 0.22, 0.23)] \end{array} \quad (77)$$

The learning rates are $\eta = 0.01$ and $\gamma = 0.01$. The accuracy of the neural network method in obtaining the Z-number coefficients is demonstrated in Figure 6 which is compared with the FCS method. In this figure, it can be seen that the estimated errors of the neural networks based algorithm are less when compared with the FCS method.

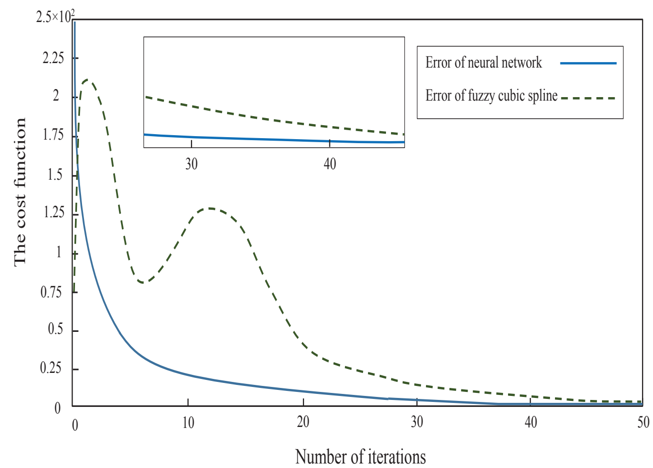


Fig. 6. Estimated errors of neural network method and FCS method

## VI. Conclusion

In this paper, the classical fuzzy equation is modified in such a way that its coefficients and variables are Z-numbers. In order to obtain the Z-number coefficients of the fuzzy equations, a novel structure of the multilayer neural network is utilized. The structure of the suggested multilayer neural network is based on the fuzzy equation. For training the neural network, the backpropagation learning approach is used. Also, some important theorems in order to support the proposed method are given. Future work is to study the stability of the training algorithm.

## References


[1] S. Abbasbandy, The application of homotopy analysis method to nonlinear equations arising in heat transfer, *Phys. Lett. A,* Vol.360, pp.109-113, 2006.

[2] S. Abbasbandy, R. Ezzati, Newton's method for solving a system of fuzzy nonlinear equations, *Appl. Math. Comput.* Vol.175, pp.1189-1199, 2006.

[3] R.H. Abiyev, N. Akkaya, I. Gunsel, Control of Omnidirectional Robot Using Z-Number-Based Fuzzy System, *IEEE Transactions on Systems, Man, and Cybernetics: Systems,* Vol.49, pp.238-252, 2019, DOI: 10.1109/TSMC.2018.2834728.

[4] R.A. Aliev, A.V. Alizadeh, O.H. Huseynov, The arithmetic of discrete Z-numbers, *Inform. Sci.* Vol.290, pp.134-155, 2015.

[5] R.A. Aliev, A.V. Alizadeh, O.H. Huseynov, K.I. Jabbarova, Z-number-based linear programming, *International Journal of Intelligent Systems,* vol.30, pp.563-589, 2015.

[6] R.A. Aliev, W. Pedrycz, O.H. Huseynov, S.Z. Eyupoglu, Approximate Reasoning on a Basis of Z-Number-Valued If-Then Rules, *IEEE Trans. Fuzzy Syst.* Vol.25, pp.1589-1600, 2017.

[7] R.A. Aliev, W. Pedrycz, V. Kreinovich, O.H. Huseynov, The general theory of decisions, *Inform. Sci.* Vol.327, pp.125-148, 2016.

[8] T. Allahviranloo, M. Otadi, M. Mosleh, Iterative method for fuzzy equations, *Soft Computing,* Vol.12, pp.935-939, 2007.

[9] V. Barthelmann, E. Novak, K. Ritter, High dimensional polynomial interpolation on sparse grids, *Adv Comput Math.* Vol.12, pp.273-88, 2000.

[10] B. Bede, L. Stefanini, Generalized differentiability of fuzzy-valued functions, *Fuzzy Sets Syst.* Vol.230, pp.119-141, 2013.

[11] J. Buckley, E. Eslami, Neural net solutions to fuzzy problems: The quadratic equation, *Fuzzy Sets Syst.* Vol.86, pp.289-298, 1997.

[12] G. Cybenko, Approximation by Superposition of Sigmoidal Activation Function, *Math.Control, Sig Syst*, Vol.2, pp.303-314, 1989.

[13] L.C. De Barros, R.C. Bassanezi, W.A. Lodwick, The Extension Principle of Zadeh and Fuzzy Numbers, *A First Course in Fuzzy Logic, Fuzzy Dynamical Systems, and Biomathematics. Studies in Fuzziness and Soft Computing, Springer, Berlin, Heidelberg,* Vol.347, pp.23-41, 2017.

[14] S. Ezadi, T. Allahviranloo, Numerical Solution of Linear Regression Based on Z-Numbers by Improved Neural Network, *Intelligent Automation and Soft Computing,* 2017, https://doi.org/10.1080/10798587.2017.1328812.

[15] R.P. Feynman, R.B. Leighton, M. Sands, The Feynman lectures on physics, *Volume III: Quantum Mechanics (Definitive ed.). New York: BasicBooks,* 2005. ISBN 978-0805390490.

[16] M.C. Floreno, G. Novelli, Implementing fuzzy polynomial interpolation (FPI) and fuzzy linear regression (LFR), *Le Matematiche.* Vol.51, pp.59-76, 1996.

[17] M. Friedman, M. Ming, A. Kandel, Fuzzy linear systems, *Fuzzy Sets Syst.* Vol.96, pp.201-209, 1998.

[18] L.A. Gardashova, Application of operational approaches to solving decision making problem using Z-Numbers, *Journal of Applied Mathematics.* Vol.5 pp.1323-1334, 2014.

[19] P. Gevorkian, Sustainable energy systems engineering: the complete green building design resource, *McGraw Hill Professional, New York,* 2007. ISBN 978-0-07-147359-0.

[20] R. Jafari, W. Yu, Fuzzy Control for Uncertainty Nonlinear Systems with Dual Fuzzy Equations, *J. Intell. Fuzzy. Syst.* Vol.29, pp.1229-1240, 2015.

[21] R. Jafari, W. Yu, X. Li, Numerical solution of fuzzy equations with Z-numbers using neural networks, *Intelligent automation and Soft Computing,* pp.1-7, 2017.

[22] A. Jafarian, R. Jafari, Approximate solutions of dual fuzzy polynomials by feed-back neural networks, *J. Soft Comput. Appl.* 2012, doi:10.5899/2012/jsca-00005.

[23] A. Jafarian, R. Jafari, A. Khalili, D. Baleanud, Solving fully fuzzy polynomials using feed-back neural networks, *International Journal of Computer Mathematics*, Vol.92, No.4, pp.742-755, 2015.

[24] A. Jafarian, R. Jafari, M. Mohamed Al Qurashi, D. Baleanud, A novel computational approach to approximate fuzzy interpolation polynomials, *SpringerPlus.* 2016, 5:1428. doi:10.1186/s40064-016-3077-5.

[25] M. Kajani, B. Asady, A. Vencheh, An iterative method for solving dual fuzzy nonlinear equations. *Appl. Math. Comput.* Vol.167, pp.316-323, 2005.

[26] B. Kang, D. Wei, Y. Li, Y. Deng, A method of converting Z-number to classical fuzzy number, *Journal of Information and Computational Science.* Vol.9, pp.703-709, 2012.

[27] B. Kang, D. Wei, Y. Li, Y. Deng, Decision making using Z–Numbers under uncertain environemnt, *Journal of Computational Information Systems.* Vol.8, pp.2807-2814, 2012.

[28] W. Lodwick, J. Santos, Constructing consistent fuzzy surfaces from fuzzy data, *Fuzzy Sets and Systems,* Vol.135, pp.259-277, 2003.

[29] J. Lorkowski, R. Aliev, V. Kreinovich, Towards Decision Making under Interval, Set-Valued, Fuzzy, and Z-Number Uncertainty: A Fair Price Approach, *FUZZ-IEEE,* pp.2244-2253, 2014.

[30] M. Mosleh, Evaluation of fully fuzzy matrix equations by fuzzy neural network, *Appl. Math. Model.* Vol.37, pp.6364-6376, 2013.

[31] R.D. Neidinger, Multivariable interpolating polynomials in newton forms, *in Proceedings of the JointmMathematics Meetings,* pp.5-8, Washington, DC,USA, 2009.

[32] S. Pederson, M. Sambandham, The Runge-Kutta method forhybrid fuzzy differential equation, *Nonlinear Anal. Hybrid Syst.* Vol.2, pp.626-634, 2008.

[33] D. Qiu, H. Jiang, Y. Yu, On computing generalized Hukuhara differences of Z-numbers, *Journal of Intelligent and Fuzzy Systems,* Vol.36, pp.1-11, 2019.

[34] H. Radstrom, An embedding theorem for spaces of convex sets, *Proc. Amer. Math. Soc.* Vol.3, pp.165-169, 1952.

[35] H. Schroeder, V.K. Murthy, E.V. Krishnamurthy, Systolic algorithm for polynomial interpolation and related problems, *Parallel Computing.* pp.493-503, 1991.

[36] J.A.K. Suykens, J. De Brabanter, L. Lukas, J. Vandewalle, Weighted least squares support vector machines: robustness and sparse approximation, *Neurocomputing,* Vol.48, pp.85-105, 2002.

[37] J. Szabados, P. Vertesi, Interpolation of functions, *World Scientific Publishing Co., Singapore,* 1990.

[38] S. Tadayon, B. Tadayon, Approximate Z-number evaluation based on categorical sets of probability distributions, *Proc. 2nd World Conf. on Soft Computing, WConSC,* 2012.
[39] C. Turchetti, G. Biagetti, F. Gianfelici, P. Crippa, Nonlinear System Identification: An Effective Framework Based on the KarhunenLove Transform, *IEEE Trans. Signal Processing,* Vol.57, no.2, pp.536-550, 2009.
[40] J. Wang, Y. Cao, and H. Zhang, Multi-Criteria Decision-Making Method Based on Distance Measure and Choquet Integral for Linguistic Z-Numbers, *Cognit. Comput.,* vol. 9, no. 6, pp. 827842, 2017.
[41] M. Waziri, Z. Majid, A new approach for solving dual fuzzy nonlinear equations using Broyden's and Newton's methods, *Advances in Fuzzy Systems*, Vol.2012, Article 682087, 5 pages, 2012.
[42] D. Xiu, J.S. Hesthaven, High-order collocation methods for differential equations with randominputs, *SIAMJournal on Scientific Computing,* Vol.27, pp.11181139, 2005.
[43] L.B. Yang, Y.Y. Gao, Fuzzy Mathematics-Theory and its Application, *published by South China University of Technology, Guangzhou China*, 1993.
[44] L.A. Zadeh, Toward a generalized theory of uncertainty (GTU) an outline, *Inform. Sci.* vol.172, pp.1-40, 2005.
[45] L.A. Zadeh, A note on Z-numbers, *Information Sciences,* Vol. 181, pp.2923-2932, 2011.
[46] A. Zolic, Numerical mathematics, *Faculty of mathematics: Belgrade.* pp.91-97, 2008.