Desalination and Water Treatment
www.deswater.com
doi: 10.5004/dwt.2018.22311

133 (2018) 237–244
November

Application of quadratic membership functions to hydrological cases

C. Tzimopopoulos[a,*], K. Papadopoulos[b], B. Papadopoulos[c]

[a]*Department of Rural and Surveying Engineering, Aristotle University, 54124 Thessaloniky, Greece, Tel. +30 2310 996141; Fax: +302310996049; email: tzimop@eng.auth.gr*
[b]*College of Engineering and Technology, American University of the Middle East, Egaila, Kuwait, email: kyriakos.papadopoulos1981@gmail.com*
[c]*Department of Civil Engineering, Democritus University of Thrace, Xanthi, Greece, Tel. +25410 79747; email: papadob@civil.duth.gr*

Received 1 February 2018; Accepted 4 April 2018

ABSTRACT

In the design of water management projects, there is vagueness in the measured data, and the classical linear regression becomes problematic in cases where the data set is too small, or there is difficulty in verifying that the error is normally distributed, or if there is vagueness in the relationship between the independent and dependent variables, or if there is ambiguity associated with the event, or if the linearity assumption is inappropriate. Therefore, the classical regression gives imprecise and inaccurate output data, and in such cases, the fuzzy set theory method provides the alternative means of treating the uncertainty in water management problems. In the present article, two cases are presented. Firstly, a fuzzy linear relation between runoff and precipitation is considered for 25 storms in Monocacy River at Jug Bridge, Maryland, USA, and a possibilistic linear fuzzy model runoff–precipitation is presented, considering the precipitation as crisp data, the runoff as fuzzy data and the coefficients as fuzzy. Secondly, a linear fuzzy model for rainfall data between two meteorological stations located in the region of Central Macedonia (Northern Greece) is also presented, with crisp data in one station and fuzzy data in the other.

Keywords: Fuzzy regression; Runoff–precipitation; Triangular numbers; Rainfall data

1. Introduction

Rainfall measurement models have been extensively used in the design process of water resources projects such as hydrological prediction, spillway design, climatic change studies, rainfall and runoff correlation and so on. Rainfall measurements in a specific area are commonly displayed in the form of time series where recorded values can be either continuous or discrete. In many instances, there is a correlation between runoff and precipitation that belongs to the same or different stations and comprises measurements with differing range. The same correlation exists among rainfall data belonging to different stations. Generally, a linear relation is assumed between them, and we can conclude in a certain relation [1].

In classical linear regression, the difference between measurement values and estimated values is a random variable, which is normally distributed and is considered to be caused by measurement errors. According to this, the classical regression is considered to be probabilistic and has many uses but can be rendered problematic if the data set is small, if it is hard to prove that error distribution is normal, if there is fuzziness between dependent and independent variables or if linearity acceptance is not proper.

In the last few decades, new regression models have been introduced based on fuzzy logic [2–10]. In fuzzy regression, the difference between measurement values and estimated values is attributed to the inherent fuzziness of the system, as well as to the fuzziness of input and output data. In contrast with classical regression analysis, fuzzy regression analysis

* Corresponding author.

Presented at the 6th International Conference on Environmental Management, Engineering, Planning and Economics (CEMEPE-2017), 25–30 June 2017, Thessaloniki, Greece.

1944-3994/1944-3986

uses fuzzy functions for the regression factors. The above problem usually has three cases [11]: crisp input values x_{ij} and crisp output values y_j, crisp input values x_{ij} and fuzzy output values $\tilde{y}_j$ and fuzzy input values $\tilde{x}_{ij}$ and fuzzy output values $\tilde{y}_j$. In all of these cases, estimated values $\tilde{Y}_j$ are fuzzy.

The adjustment of a fuzzy regression model can be achieved through two general methods:

- The possibilistic model [2,11,12]. Fuzzy regression is possibilistic, and the membership function $\mu_{\tilde{A}}(x)$ of a fuzzy number $\tilde{A}$ is considered equal to the possibility distribution function $\pi_x(x)$. The fuzziness of the model is minimized by taking into account the minimum of the spreads around the centre of the fuzzy parameters, while considering that the experimental values of every sample are within a specific interval of possible values. It is to point out that possibilistic parameters in the models are non-interactive, that is, the joint possibilistic distribution of parameters is defined by minimum operators.
- The least squares model [13–15]. The distance between the estimated output value of the model $\tilde{Y}_i$ and the observed output value $\tilde{y}_i$ is minimized. This method of Diamond [14] is considered to be an extension of the classical linear regression method, based on the notion of model efficiency optimization depending on data.

In this article, quadratic membership functions as defined by [6,16] are considered to propose a method of interactive fuzzy parameters in possibilistic linear hydrological systems. Two cases are presented: a fuzzy linear relation between runoff and precipitation is considered for 25 storms in Monocacy River at Jug Bridge, Maryland, USA, borrowed from [1], and the method can be reduced to linear programming; and a linear fuzzy model for rainfall data between Aggistron and Ano Vrontou meteorological stations, located in the region of Central Macedonia (Northern Greece), is also presented. In this model, the independent observed rainfall values are crisp, and the dependent observed rainfall values, as well as the parameters of the model, are fuzzy.

Celmiņš [16] wanted to maximize the membership values of the observations by minimizing the sum of squares of the deviations of the membership values from one. He obtained relatively simple algorithms if the membership functions of the data vectors belonged to a particular class of conical functions. Tanaka and Ishibuchi [6] used a method similar to one considered by Celmiņš, but the proposed approach is simpler and more understandable than Celmiņš approach and leads to a linear programming problem. They used quadratic membership functions as defined by Celmiņš and proposed an identification method of interactive fuzzy parameters in possibilistic linear systems.

Here, we use the proposed method by Tanaka and Ishibuchi [6] in order to obtain interval inclusion between measured and estimated values. Input measured data were considered crisp (x_i = rainfall measurements), and output values were considered fuzzy ($\tilde{y}_i$ = runoff or rainfall measurements). Quadratic membership functions were used for measured output values. For the second case, a modified version of Tanaka and Ishibuchi [6] model is also applied.

2. Mathematic model

2.1. Definitions

2.1.1. Definition 1

A fuzzy set $\tilde{A}$ on a universe set X is a mapping $\tilde{A}$:$X \to [0, 1]$, assigning to each element $x \in X$ a degree of membership $0 \le \tilde{A}(x) \le 1$. The membership function is also defined as $\mu_{\tilde{A}}(x)$ with the properties:

(i) $\mu_{\tilde{A}}(x)$ is upper semicontinuous; (ii) there are real numbers $c \le a \le b \le d$, such that $\mu_{\tilde{A}}$ is increasing on $[c, a]$, decreasing on $[b, d]$ and $\mu_{\tilde{A}}(x) = 1$ on $[a, b]$; (iii) for each $x \in R$, $\mu_{\tilde{A}}(x) = 0$, outside of the interval $[c, d]$ and (iv) $\tilde{A}$ is a convex fuzzy set, that is, $\mu_{\tilde{A}}(\lambda x + (1-\lambda)x) \ge \mu_{\tilde{A}}(\lambda x) \wedge \mu_{\tilde{A}}((1-\lambda)x)$.

2.1.2. Definition 2

Let X being a Banach space and $\tilde{A}$ being a fuzzy set on X. We define the α-cuts of $\tilde{A}$ as $[\tilde{A}]^{\alpha} = \{x \in R \mid \tilde{A}(x) \ge \alpha\}$, $\alpha \in [0, 1]$, and for $\alpha = 0$, we define the closure $[\tilde{A}]^0 = \overline{\{x \in R \mid \tilde{A}(x) > 0\}}$.

2.1.3. Definition 3

Let $\mathcal{K}(X)$ the family of all non-empty compact convex subsets of a Banach space. A fuzzy set $\tilde{A}$ on X is called compact if $[\tilde{A}]^{\alpha} \in \mathcal{K}(X)$, $\forall\alpha \in [0, 1]$. The space of all compact and convex fuzzy sets on X is denoted as $\mathcal{F}(X)$.

2.1.4. Definition 4

Let $[\tilde{A}] \in \mathcal{F}(R)$. The α-cuts of $\tilde{A}$ are: $[\tilde{A}]^{\alpha} = [A_{\alpha}^{-}(x), A_{\alpha}^{+}(x)]$. According to representation theorem of [17] and the theorem of [18], the membership function and the α-cut form of a fuzzy number $\tilde{A}$ are equivalent and in particular the α-cuts $[\tilde{A}]^{\alpha} = [A_{\alpha}^{-}(x), A_{\alpha}^{+}(x)]$ uniquely represent $\tilde{A}$, provided that the two functions are monotonic (A_{α}^{-} increasing and A_{α}^{+} decreasing) and $A_{\alpha}^{-} \le A_{\alpha}^{+}$, for $\alpha = 1$.

2.2. Model development

Consider a fuzzy dependent variable $\tilde{Y}_j$ and x_{ij}, the independent variables influencing the variable $\tilde{Y}_j$. The result of fuzzy linear regression is an equation of the form [2,6]:

$$\tilde{Y}_j = \tilde{A}_0 + \tilde{A}_1 x_{1j} + \tilde{A}_2 x_{2j} + \ldots + \tilde{A}_n x_{nj} = \sum_{i=0}^{n} \tilde{A}_i x_{ij}, \quad x_{0j} = 1 \tag{1}$$

where the measured input values x_{ij} are crisp numbers, and the measured output values $\tilde{y}_j$ are fuzzy numbers. The parameters $\tilde{A} = (A_0 \ldots A_n)$ are considered as fuzzy vectors. If the solutions of the h-level spreads of the fuzzy parameters contain all non-negative elements for $j = 0, 1 \ldots n$, then, the fuzzy parameters are called non-interactive and the trends between the spreads and the mode of the fuzzy estimated values are compatible. If not all the estimated spreads are non-negative, then, interactive fuzzy parameters exist [6,13]. The selection of Tanaka and Ishibushi model is considered necessary in order to prevent the case of interactive fuzzy parameters. According to [6], in that case, a membership function of fuzzy parameters $\tilde{A}_i = (\vec{r}^*, C)$ is defined by:

$$\mu_A(\vec{r}) = \max\{1 - (\vec{r} - \vec{r}^*)^t \cdot C^{-1} \cdot (\vec{r} - \vec{r}^*), 0\} \tag{2}$$

where $\vec{r}^* = (r_1^*, ..., r_n^*)$ is a centre vector of fuzzy parameter $\tilde{A}_i$, and C is a symmetrical positive definite matrix.

Tanaka and Ishibuchi [6] propose for the membership function of $\tilde{Y}$ the expression:

$$\mu_{\tilde{Y}} = \max_{(\hat{r} | y = \hat{r}^T x)} \left(1 - (\vec{r} - \vec{r}^*)^t \cdot C^{-1} \cdot (\vec{r} - \vec{r}^*), 0\right), \tag{3}$$

and they prove it as equal to:

$$\mu_{\tilde{Y}}(y) = \max\left(1 - \frac{(\vec{r}^{*t} \cdot \vec{x} - y)^2}{\vec{x}^t \cdot C \cdot \vec{x}}, 0\right) \tag{4}$$

According to their theory, the following steps are followed:

- Given the input–output data (x_i, $\tilde{y}_i$, $i = 1,\ldots,n$) and a threshold h, it must hold an inclusion [19]:

$$[\tilde{y}_i]^h \subset [\tilde{Y}_i^*]^h, \quad i = 1, ..., n$$

where $[\tilde{Y}_i^*]^h$ is a h-level defined by:

$$\left[\tilde{y}_i\right]^h \subset \left[\tilde{Y}_i^*\right]^h = \left\{\tilde{y} \middle| \mu_{\tilde{Y}_i}, (\tilde{y} \geq h)\right\}$$

- The following sum of spreads of the estimated fuzzy $[\tilde{Y}_j^*]$, $i = 1\ldots n$ should be:

$$J(C) = \sum_{i=0}^{N} x_i^T C x_i \tag{5}$$

where the matrix C is proved to be a positive semi-definite matrix.

The problem now is formulated as follows:

$$\begin{aligned} &\min J = \sum_{i=0}^{N} x_i^T C x_i \\ &\text{subject to} \\ &r^T x_i - \{(1-h) x_i^T C x_i\}^{1/2} \leq y_i - \{(1-h) e_i\}^{1/2} \\ &r^T x_i + \{(1-h) x_i^T C x_i\}^{1/2} \geq y_i + \{(1-h) e_i\}^{1/2} \\ &x_i^T C x_j = 0, \quad \forall i \neq j, \quad \forall i, j \in I, \; i = 0, 1, ..., N \end{aligned} \tag{6}$$

where:

$$\begin{aligned} &[\tilde{Y}_i]^h = [r_i x_i - \{(1-h) x_i^T C x_i\}^{1/2}, \quad r_i x_i + \{(1-h) x_i^T C x_i\}^{1/2}] \\ &[\tilde{y}_i]^h = [y_i - \{(1-h) e_i\}^{1/2}, \; y_i + \{(1-h) e_i\}^{1/2}] \end{aligned}$$

and the membership function of the output data $\tilde{y}_i$ is defined as:

$$\mu_{\tilde{y}}(y) = \max\{1 - (y - y_i)^2 / e_i, 0\} \tag{7}$$

In the above relations, i defines the sum of the fuzziness of the model, which should be minimized according to [6]. The h value ($h \in [0, 1]$) is referred to as the degree of fit of the estimated fuzzy linear model to the given data and is subjectively selected by a decision maker as an input to the model. Besides $x_i C_x$ for any x is essential to the above inclusion according to definition [6], which means that the matrix C is positive semi-definite.

The solution of this problem according to [6,20] is as follows.

2.2.1. First phase

An optimum vector r^* is found that minimizes the expression:

$$\sum_{i=1}^{m} (\tilde{y}_i - r^{*T} x_i)^2 \tag{8}$$

This solution constitutes the classical linear regression solution.

2.2.2. Second phase

The following optimization problem is solved with linear programming:

$$\min J = \sum_{i=0}^{N} x_i^T C x_i, \quad i = 0, 1, ..., n$$

subject to:

$$\begin{aligned} &[\tilde{y}_i]^h \subseteq [\tilde{Y}_i^*]^h \Leftrightarrow 1 - \frac{(\tilde{y}_i - r^{*T} x_i)^2}{x_i^T C x_i} \geq h \\ &\Leftrightarrow \{(1-h) x_i^T C x_i\} \geq ([\tilde{y}_i]^h - r^{*T} x_i)^2 \\ &\Leftrightarrow \{(1-h) x_i^T C x_i\}^{1/2} \geq k_1 = r^{*T} x_i - y_i + \{(1-h) e_i\}^{1/2} \\ &\quad \{(1-h) x_i^T C x_i\}^{1/2} \geq k_2 = y_i - r^{*T} x_i + \{(1-h) e_i\}^{1/2} \\ &\Leftrightarrow \{(1-h) x_i^T C x_i\} \geq \{\max(k_1, k_2)\}^2, \end{aligned} \tag{9}$$

$$k_1 = r^{*T} x_i - y_i + \{(1-h) e_i\}^{1/2}, \qquad k_2 = y_i - r^{*T} x_i + \{(1-h) e_i\}^{1/2}$$

This problem is called Min. problem according to [6]. If the optimum solution C^* is a positive semi-definite matrix, then (r^*, C^*) is the solution of the problem. Otherwise, the third phase follows.

2.2.3. Third phase

The following orthogonal constraints are added to the above problem:

$$x_i^T C x_j = 0, \quad \forall i \neq j, \quad \forall i, j \in I, \; i = 1, 2, ..., N \tag{10}$$

and the problem is solved including these conditions. The solution is (r^*, C^*). In relation to Eq. (10), I is the set of subscripts of the independent vectors $\{I = (1, 2\ldots n)\}$.

Remark. The estimated membership function of parameters can be computed by the following equation:

$$1 - (\vec{r} - \vec{r}^*)^t \cdot C^{-1} \cdot (\vec{r} - \vec{r}^*) = h \tag{11}$$

for different h-levels. According to [16], this boundary of the supports of the membership function is a hyperellipsoid, and the support principal axes are not parallel to the coordinate system.

3. Application

3.1. Runoff–precipitation model

We consider the data for 25 storms on the Monocacy River at Jug Bridge, Maryland, USA [1]:

In Table 1, the precipitation X is crisp, and the runoff Y is fuzzy, with e the fuzziness of runoff taken equal to $0.15Y$.

For this case, Eq. (1) becomes:

$$\tilde{Y}_j = \tilde{A}_0 + \tilde{A}_1 x_{1j},\ x_{0j} = 1$$

Utilizing classical statistics, the vector $r^* = (-0.1248, 0.4294)^T$ is produced. The problem now is formulated as follows:

$$\min J = \sum_j (x_{ji}^T C x_{ji}) = 25c_{11} + 53.89c_{21} + 53.89c_{12} + 153.42c_{22}$$

subject to:

$$\begin{bmatrix} 1 & 1.11 & 1.11 & 1.20 \\ 1 & 1.17 & 1.20 & 1.40 \\ \cdots & \cdots & \cdots & \cdots \\ 1 & 2.93 & 2.90 & 8.60 \\ 1 & 1.16 & 1.20 & 1.30 \end{bmatrix} \cdot \begin{bmatrix} c_{11} \\ c_{21} \\ c_{12} \\ c_{22} \end{bmatrix} \geq \begin{bmatrix} 0.199 \\ 0.071 \\ \cdots \\ 0.180 \\ 0.331 \end{bmatrix}$$

Table 1
Number of storms

a/a	Precipitation, X (in)	Runoff, Y (in)	Fuzziness, $e = 0.15Y$ (in)
1	1.11	0.52	0.078
2	1.17	0.40	0.060
3	1.79	0.97	0.146
4	5.62	2.92	0.438
5	1.13	0.17	0.026
6	1.54	0.19	0.029
7	3.19	0.76	0.114
8	1.73	0.66	0.099
9	2.09	0.78	0.117
10	2.75	1.24	0.186
11	1.20	0.39	0.059
12	1.01	0.30	0.045
13	1.64	0.70	0.105
14	1.57	0.77	0.116
15	1.54	0.59	0.089
16	2.09	0.95	0.143
17	3.54	1.02	0.153
18	1.17	0.39	0.059
19	1.15	0.23	0.035
20	2.57	0.45	0.068
21	3.57	1.59	0.239
22	5.11	1.74	0.261
23	1.52	0.56	0.084
24	2.93	1.12	0.168
25	1.16	0.64	0.096

where the numbers (0.199, 0.071…0.180, 0.331) mean $\max(k_1, k_2)^2$, and k_1 and k_2 are:

$$k_1 = r_i^* x_i - y_i + \{(1-h)e_i\}^{1/2}, \quad k_2 = y_i - r_i^* x_i + \{(1-h)e_i$$

The following matrix is produced from the solution:

$$C = \begin{bmatrix} 0.0407 & 0.1205 \\ 0.1205 & 0.0086 \end{bmatrix}$$

This matrix is not positive semi-definite, and thus, the following restriction is added to the problem:

$$x_i^T C x_j = 0,\ \ \forall i \neq j,\ \ \forall i, j \in I,\ i = 1,2,....,N, \text{ with}$$
$$x_{12} = (1,\ 1.01)^T,\ x_4 = (1,\ 5.62)$$

The new matrix that is produced is:

$$C^* = \begin{bmatrix} 1.032 & -0.307 \\ -0.307 & 0.177 \end{bmatrix}$$

This matrix is positive semi-definite. Table 2 shows the values of J, for various combinations of x_i and x_j vectors. This matrix means that as the first matrix is not positive semi-definite, we have added for the above case the equation:

$$x_1^T C x_2 = 0, \text{ or } \begin{bmatrix} 1 \\ 1.01 \end{bmatrix}^T \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} \begin{bmatrix} 1 \\ 1.17 \end{bmatrix} =$$
$$0 \rightarrow c_{11} + 1.01c_{12} + 1.17c_{21} + 1.1817c_{22} = 0$$

with a first combination (x_1, x_2), and the new matrix was positive semi-definite with $J = 11{,}840.25$. The same is repeated for different combinations of (x_i, x_j) vectors in order to obtain the optimal value of J, which is 19.89 with the combination (x_{12}, x_4).

Fig. 1 shows the estimated fuzzy outputs with $h = 0$, and the supports include fully the observed measurements of the Monocacy River storms [1] with their deviations, in the entire study region (points $y \pm e^{1/2}$). According to [19], the α-cuts for $h = 0$, 0.5, and 1 are the intervals of confidence for the fuzzy number and show the fuzziness of the estimated fuzzy number $\tilde{Y}$ for every h. Fig. 2 illustrates the inclusion of predicted and measured data at $x = 5.62$. For $h = 0$ we have:

$$[\tilde{y}_4]^{h=0} = [2.92 - 0.438^{1/2}, 2.92 + 0.438^{1/2}]$$
$$= [2.26,\ 3.58], [\tilde{Y}_4]^{h=0} = [0.51,\ 4.07]$$

and $[\tilde{y}_4]^{h=0} \subset [\tilde{Y}_4]^{h=0}$. For $h = 0.5$ we have:

$$[\tilde{y}_4]^{h=0.5} = [2.92 - (0.5 \times 0.438)^{1/2},\ 2.92 + (0.5 \times 0.438)^{1/2}]$$
$$= [2.45,\ 3.39], [\tilde{Y}_4]^{h=0.5} = [1.03,\ 3.55]$$

and $[\tilde{y}_4]^{h=0.5} \subset [\tilde{Y}_4]^{h=0.5}$. Fig. 3 illustrates the estimated membership function of parameters in the Min. problem. In this figure, the ellipse is obtained by applying Eq. (11):

Table 2
The optimal value of J vs. x_i, x_j

Combination	x_1x_2	x_1x_3	x_1x_4	x_1x_7	x_1x_{10}	x_1x_{22}	$\mathbf{x_{12}x_4}$	$x_{12}x_{17}$	$x_{12}x_1$	$\mathbf{x_4x_{12}}$	x_2x_4
Optimal value J	11,840.25	116.99	20.50	28.95	36.86	21.17	**19.89**	25.02	1,716.21	**19.89**	20.92

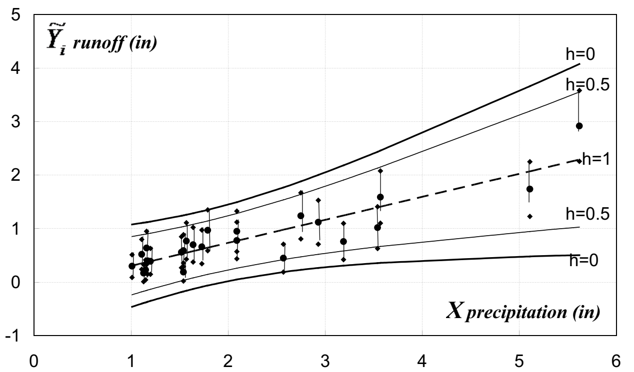

Fig. 1. The estimated fuzzy output and the given data for the Runoff–Precipitation application.

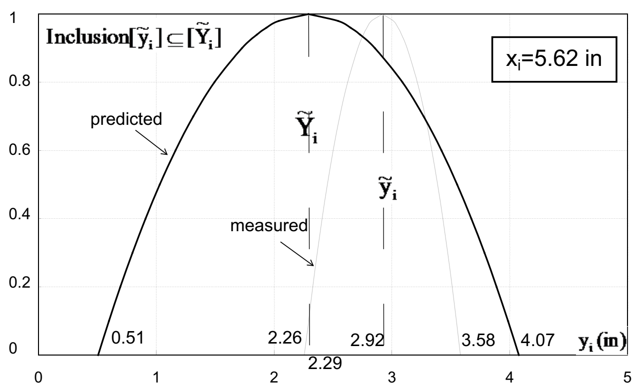

Fig. 2. Inclusion of predicted and measured data at x = 5.62 for the runoff–precipitation application.

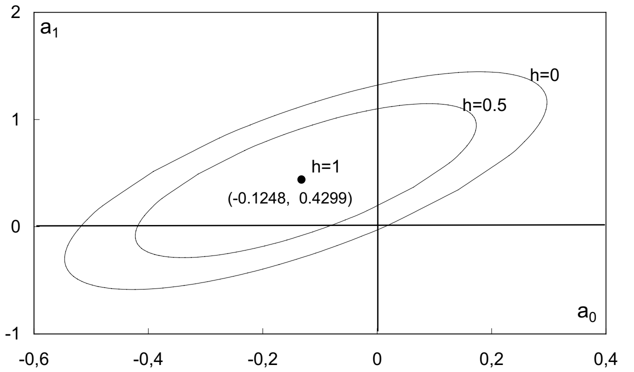

Fig. 3. The estimated membership function of parameters in the Min. problem for the runoff–precipitation application.

Table 3
Mean monthly rainfalls: crisp input and fuzzy output data

T	x_i	Aggistro, X (mm)	Vrontou, Y (mm)	Fuzziness, $e = 0.2Y$ (mm)
1929–1930	x_1	47.6	65.8	13.16
1930–1931	x_2	65.6	118.9	23.78
1931–1932	x_3	40.3	57	11.4
1932–1933	x_4	30.8	52.5	10.5
1933–1934	x_5	45.1	61.1	12.22
1934–1935	x_6	54.1	73.9	14.78
1935–1936	x_7	58.7	104.4	20.88
1936–1937	x_8	60.6	83.1	16.62
1937–1938	x_9	55.1	78.8	15.76
1938–1939	x_{10}	45.9	79	15.8
1939–1940	x_{11}	54.7	106.5	21.3
1940–1941	x_{12}	47.5	77.5	15.5

3.2. Rainfall data model

We consider the rainfall measurement stations of Aggistro and Ano Vrontou with the data in Table 3:

In Table 3, the Aggistro data X are crisp, and the Ano Vrodou data Y are fuzzy, with e the fuzziness of Y taken equal to $0.2Y$.

For this case, Eq. (1) becomes: $\tilde{Y}_j = \tilde{A}_0 + \tilde{A}_1 x_{1j}$, $x_{0j} = 1$. Utilizing classical statistics the vector $r^* = (-11.941, 1.8181)^T$ is produced. The problem now is formulated as follows:

$$\min J = \sum_j (x_{ji}^T C x_{ji}) = 12c_{11} + 606c_{21} + 606c_{12} + 31612c_{22}$$

subject to:

$$\begin{bmatrix} 1 & 47.6 & 47.6 & 2265 \\ 1 & 65.6 & 65.6 & 4303 \\ \cdots & \cdots & \cdots & \cdots \\ 1 & 54.7 & 54.7 & 2292 \\ 1 & 47.5 & 47.5 & 2256 \end{bmatrix} \cdot \begin{bmatrix} c_{11} \\ c_{21} \\ c_{12} \\ c_{22} \end{bmatrix} \geq \begin{bmatrix} 154 \\ 271 \\ \cdots \\ 557 \\ 49 \end{bmatrix}$$

where the numbers (154, 271…557, 49) mean $\max(k_1, k_2)^2$, and k_1 and k_2 are:

$$k_1 = r_i x_i - \{(1-h) x_i^T C x_i\}^{1/2}, \quad k_2 = r_i x_i + \{(1-h) x_i^T C x_i\}^{1/2}$$

The following matrix is produced from the solution:

$$C = \begin{bmatrix} -2531 & 62.73 \\ 62.73 & 1.26 \end{bmatrix}$$

This matrix is not positive semi-definite, and therefore, the following restriction is added to the problem:

$$x_i^T C x_j = 0, \quad \forall i \neq j, \quad \forall i,j \in I, \ i = 1,2,....,N, \text{ with}$$
$$x_2 = (1, \ 65.6)^T, \ x_4 = (1, \ 30.80)$$

The new matrix that is produced is:

$$C^* = \begin{bmatrix} 1388 & -36.72 \\ -36.72 & 1.06 \end{bmatrix}$$

This matrix is positive semi-definite. Table 3 shows the values of J, for various combinations of x_i and x_j vectors. Again as in Table 2, this matrix means that as the first matrix is not positive semi-definite, we have added for the above case the equation:

$$x_1^T C x_{12} = 0, \text{ or } \begin{bmatrix} 1 \\ 47.6 \end{bmatrix}^T \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} \begin{bmatrix} 1 \\ 47.5 \end{bmatrix} =$$
$$0 \rightarrow c_{11} + 47.6c_{12} + 47.5c_{21} + 22610c_{22} = 0$$

with a first combination of (x_1, x_{12}), and the new matrix was positive semi-definite with J = 22717852. The same procedure is repeated for different combinations of (x_i and x_j) vectors in order to obtain the optimal value of J, which is 5,817 with combination (x_2 and x_4).

Fig. 4 shows the estimated fuzzy outputs with h = 0, 0.5, and 1, and the supports include the observed measurements with their deviations, in the entire study region (points $y \pm e^{1/2}$). Fig. 5 illustrates the inclusion of predicted and measured data at x = 54.1. For h = 0, we have $[\tilde{y}_6]^{h=0}$ = [70.08, 77.7], $[\tilde{y}_6]^{h=0}$ = [63.4, 109.5] and $[\tilde{y}_6]^{h=0} \subset [\tilde{y}_6]^{h=0}$. For h = 0.5, we have $[\tilde{y}_6]^{h=0.5}$ = [71.2, 76.6], $[\tilde{y}_6]^{h=0.5}$ = [70.01, 102.7] and $[\tilde{y}_6]^{h=0.5} \subset [\tilde{y}_6]^{h=0.5}$. Fig. 6 illustrates the estimated membership function of parameters in the Min. problem for the original [6] model.

Remark: As shown in Figs. 1 and 4, the supports for h = 0 do not include the observed measurements data with their deviations, in the entire study region (points $y \pm e$). The inclusion contains only the points $y \pm e^{1/2}$. This result is due

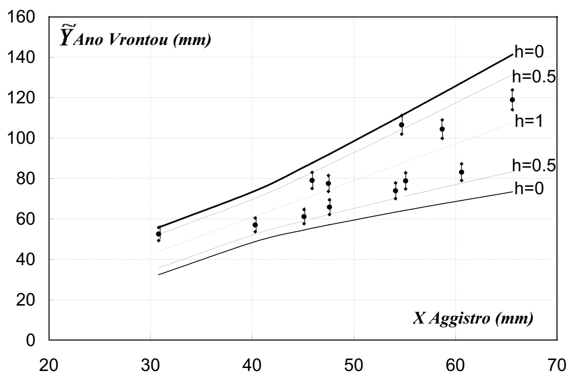

Fig. 4. The estimated fuzzy output and the given data for the original [6] model.

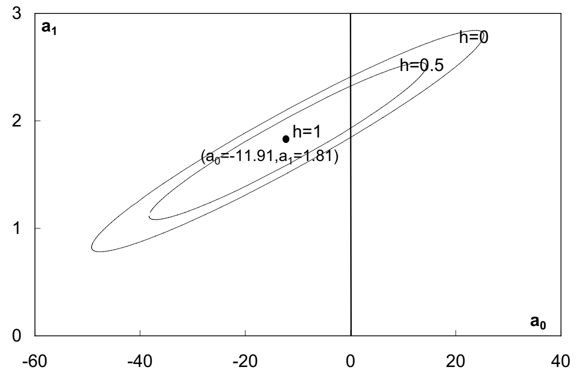

Fig. 6. The estimated membership function of parameters for the original [6] model.

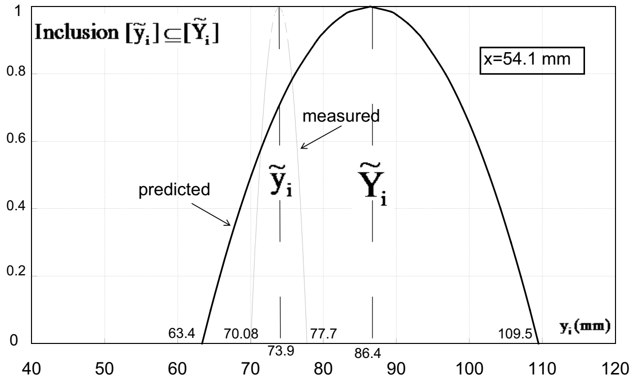

Fig. 5. Inclusion of predicted and measured data at x = 54.1 mm for the original [6] model.

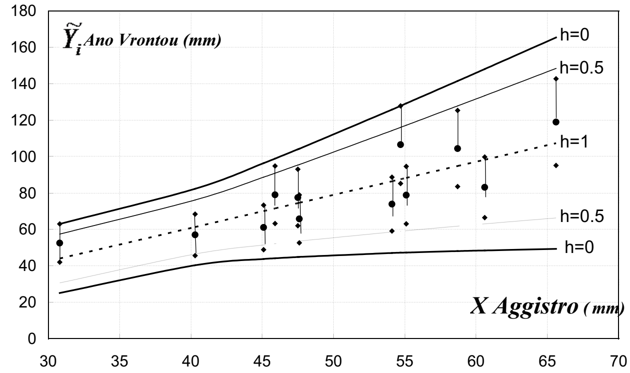

Fig. 7. The estimated fuzzy output and the given data using the new modified method.

Table 4
The optimal value of J vs. x_i, x_j

Combination	x_1, x_{12}	x_1, x_2	x_1, x_3	x_1, x_4	x_1, x_5	x_2, x_3	x_2, x_4
Optimal value J	22,717,852	12,806	15,743	5,822	77,985	6,627	5,817

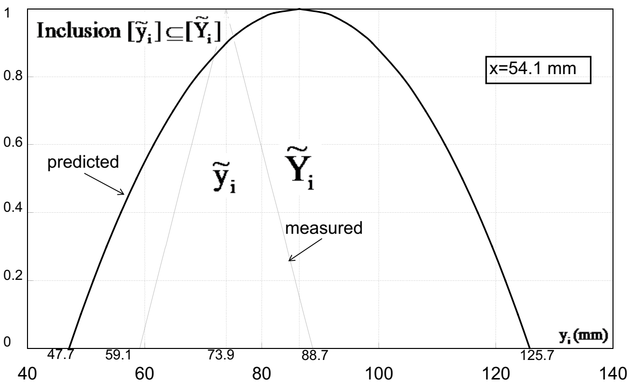

Fig. 8. Inclusion of predicted and measured data at x = 54.1 mm using the new modified method.

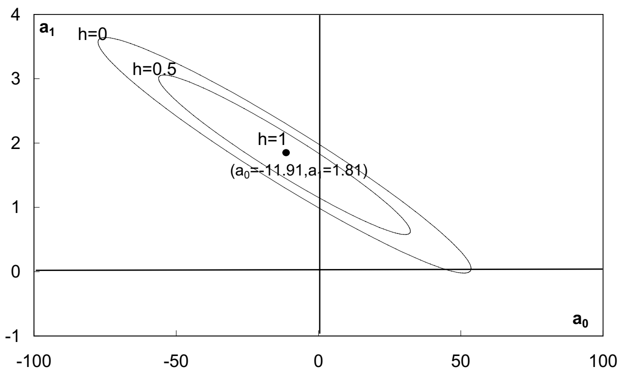

Fig. 9. The estimated membership function of parameters using the new modified method.

to the fact that Tanaka-Ishibuchi model uses as membership function of the data y: $\mu_{Yi}(y) = \max\{1 - (y - y_i)^2/e_i, 0\}$. In order, for the inclusion, to contain the points $y \pm e$, the membership function is modified by [20] and takes the following form: $\mu_{Yi}(y) = \max\{1 - |y - y_i|/e_i, 0\}$.

With this new modification, Fig. 7 shows the estimated fuzzy outputs with h = 0, 0.5, and 1, and the supports include the observed measurements with their deviations, in the entire study region (points $y \pm e$). Fig. 8 illustrates the inclusion of predicted and measured data at x = 54.1. In the modified model for h = 0, we have $[\tilde{y}_6]^{h=0}$ = [73.9 – 14.8, 73.9+14.8] = [59.12, 88.68], $[\tilde{y}_6]^{h=0}$ = [47.1, 125.8] and $[\tilde{y}_6]^{h=0} \subset [\tilde{y}_6]^{h=0}$. For h = 0.5, we have $[\tilde{y}_6]^{h=0.5}$ = [66.51, 81.29], $[\tilde{y}_6]^{h=0.5}$ = [58.6, 114.3] and $[\tilde{y}_6]^{h=0.5} \subset [\tilde{y}_6]^{h=0.5}$. Fig. 6 illustrates the estimated membership function of parameters in the Min. problem for the [6] modified model.

4. Conclusions

In this paper, the approach of Tanaka and Ishibuchi is considered, which could identify interactive fuzzy parameters in a possibilistic linear model of precipitation–runoff measurements and rainfall station measurements. This model is more general and can be used also for cases with interactive distribution parameters, where not all the estimated spreads are non-negative.

In the case of rainfall and runoff measurement observations, fuzzy correlation is achieved, even for small samples, and we can extend the shorter time series, due to fuzzy correlation. The estimated fuzzy outputs with h = 0 include the observed measurements with their deviations, in the entire study region.

In cases in which the original [6] approach does not insure data inclusion inside estimated supports, a modified version of [6] model is applied, which insures fully inclusion.

In the case of rainfall measurement observations, station association is achieved, even for small samples, and we can extend the shorter time series, due to fuzzy correlation of two rainfall stations.

The model of Tanaka and Ishibuchi has a membership function of parameters, which is quadratic, but it can be reduced to linear programming, easy to apply.

Finally fuzzy regression can be a useful tool for managers and researchers in hydrology (water management), for estimating relationships among variables with fuzzy, incomplete, and limited information. It may be more effective than statistical regression with rigid assumptions, when the last ones are either violated or cannot be properly employed.

References

[1] R. Linsley, J. Franzini, Water Resources Engineering, McGraw-Hill Book Co. Inc., New York, 1964, p. 654.

[2] H. Tanaka, S.Uejima, K. Asai, Linear regression analysis with fuzzy models, IEEE Trans. Syst. Man Cybern., 12 (1982) 903–907.

[3] A. Bárdossy, I. Bogardi, L. Ducstein, Fuzzy regression in hydrology, Water Resour. Res., 26 (1990) 1497–1508.

[4] V. Sahin, M.J. Hall, The effects of afforestation and deforestation on water yields, J. Hydrol., 178 (1996) 293–309.

[5] C.-G. Wen, C.-S. Lee, Development of a cost function for wastewater treatment systems with fuzzy regression, Fuzzy Sets Syst., 106 (1999) 143–153.

[6] H. Tanaka, H. Ishibuchi, Identification of possibilistic linear systems by quadratic membership functions of fuzzy parameters, Fuzzy Sets Syst., 41 (1991) 145–160.

[7] R.R. Shrestha, A. Bárdossy, F. Nestmann, Analysis and propagation of uncertainties due to the stage–discharge relationship: a fuzzy set approach, Hydrol. Sci. J., 52 (2007) 595–610.

[8] U.T. Khan, C. Valeo, A new fuzzy linear regression approach for dissolved oxygen prediction, Hydrol. Sci. J., 60 (2015) 1096–1119.

[9] C. Tzimopoulos, B. Papadopoulos, Fuzzy Logic with Application in Engineering Science, Ziti, Ed. (in Greek) 2013.

[10] C. Tzimopoulos, K. Papadopoulos, B. Papadopoulos, Fuzzy Regression with Applications in Hydrology, Int. J. Eng. Innov. Technol., 5 (2016) 69–75.
[11] D.T. Redden, W.H. Woodall, Further examination of fuzzy linear regression, Fuzzy Sets Syst., 79 (1996) 203–211.
[12] D.A. Savic, W. Pedrycz, Evaluation of fuzzy linear regression models, Fuzzy Sets Syst., 39 (1991) 51–63.
[13] P.-T. Chang, S.E. Lee, A generalized fuzzy least-squares regression, Fuzzy Sets Syst., 82 (1996) 289–298.
[14] P. Diamond, Higher level fuzzy numbers arising from fuzzy regression models, Fuzzy Sets Syst., 36 (1990) 265–275.
[15] M.-S. Yang, H.-H. Liu, Fuzzy least-squares algorithms for interactive fuzzy linear regression models, Fuzzy Sets Syst., 135 (2003) 305–316.
[16] A. Celmiņš, Least squares model fitting to fuzzy vector data, Fuzzy Sets Syst., 22 (1987) 245–269.
[17] C.V. Negoita, D.A. Ralescu, Representation theorems for fuzzy concepts, Kybernetes, 4 (1975) 169–174.
[18] R. Goetschel, W. Voxman, Elementary fuzzy calculus, Fuzzy Sets Syst., 18 (1986) 31–43.
[19] A. Kaufmann, M. Gupta, Introduction to Fuzzy Arithmetic, International Thomson Computers Press, London, 1991.
[20] C. Tzimopoulos, C. Evangelides, B. Papadopoulos, Rainfall Data Regression Model Using Fuzzy Set Logic, Sustainable Hydraulics in the Era of Global Change, Proc. 4th IAHR Europe Congress, Liege, Belgium, 2016.