Estimation of a Fuzzy Regression Model Using Fuzzy Distances

Antonio Francisco Roldán López de Hierro, Juan Martínez-Moreno, Concepción Aguilar- Peña, and Concepción Roldán López de Hierro

***Abstract*—Regression analysis is a powerful statistical tool that has many applications in different areas. The problem of regression analysis under a fuzzy environment has been treated in the literature from different points of view and considering a variety of input/output data (crisp or fuzzy). However, we realize that, in general, most research papers have a conflict between the solution of the fuzzy regression problem using crisp distances (minimizing a real error function) and the interpretation of fuzzy data as possibility distributions. The main aim of this paper is to develop a methodology to solve this problem introducing a fuzzy partial order and a family of fuzzy distance measures on the whole set of fuzzy numbers. The new approach allows us to obtain linear and nonlinear models that reach the lowest fuzzy error; the estimation process, in general, can be considered easier to apply in practice, and it is not limited to triangular fuzzy numbers. Numerical examples are provided to illustrate the usefulness and applicability of these results, and comparisons with existing methodologies show that the performance of the proposed solution is very satisfactory.**

***Index Terms*—Fuzzy distance, fuzzy number (FN), fuzzy regression, least-squares method, minimization problem.**

I. Introduction

REGRESSION analysis is a powerful statistical tool that has many applications in different areas, such as engineering, environmental sciences, finance and economics, social sciences, biology, etc. This popular methodology is used to find the relationship between two or more quantitative random variables. For the conventional regression analysis, several methods have been outlined in the literature, of which the method of least squares is probably the most widely used in empirical research. The main aim of this paper is to solve fuzzy regression problems considering a wide range of genuine fuzzy metrics (or semimetrics). In this context, five problems naturally arise in order to deal with the uncertainty inherent in the model.

Manuscript received August 27, 2014; revised January 25, 2015; accepted May 26, 2015. Date of publication July 13, 2015; date of current version March 29, 2016. This work was supported by Junta de Andalucía through projects FQM-268, FQM-245, and FQM-178 of the Andalusian CICYE.

A. F. Roldán López de Hierro is with the Department of Quantitative Methods for Economics and Business, University of Granada, 18071 Granada, Spain (e-mail: aroldan@ugr.es).

J. Martínez-Moreno is with the Department of Mathematics, University of Jaén, Jaén 23071, Spain (e-mail: jmmoreno@ujaen.es).

C. Aguilar-Peña is with the Department of Statistics and Operations Research, University of Jaén, Jaén 23071, Spain (e-mail: caguilar@ujaen.es).

C. Roldán López de Hierro is with the Department of Statistics and Operations Research, University of Granada 18071 Granada, Spain (e-mail: iroldan@ugr.es).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TFUZZ.2015.2455533

1) First, the problem of regression analysis under a fuzzy environment has been treated in the literature from different points of view and considering a variety of input/output data (crisp or fuzzy). We realize that, in general, most of research papers consider crisp distances, which leads to a conflict between the solution of the fuzzy regression problem and the interpretation of fuzzy data as possibility distributions (see, for instance, [9], [26], and [30] and references therein, which provided extensive reviews of the main approaches to fuzzy regression).
2) Second, since crisp distances are not coherent in this framework, we need to consider a fuzzy distance to carry out the regression. In this sense, due to its possible applications in several areas, the notion of *metric* plays a key role in many fields of study that interpret the distance measure between two points as the difference between them. Traditionally, the distance between two points has always been a real number. Even in a fuzzy context, the distance between two *fuzzy numbers* (FNs) is usually interpreted as a crisp number considering, for instance, the area between them (see [7] and [30]). However, this conception does not capture both imprecision and uncertainty and, consequently, is not consistent with factors such as vagueness and ambiguity, which affect the behavior of the phenomenon studied in the fuzzy setting. Therefore, it would be more reasonable to use an FN rather than a real number to measure the distance between two FNs.
   Within the field of fuzzy algebra, it is not clear how to consider a canonical way to associate a FN to a pair of FNs that can be interpreted as their distance. Although Mizumoto and Tanaka [23] established on this set an arithmetic (that extends the real arithmetic), the usual arithmetic with fuzzy values determines a semilinear structure, which implies that there is no generally applicable definition for the difference of fuzzy values preserving the connection with the sum in the numerical case (see [32]). Thus, many researchers have worked on finding a variety of ways to determine the difference between FNs (see [28], [33] and references therein) and have interpreted their methodologies as a distance measure rather than a classical distance (see, for instance, [2], [3], [6], [17], and [29]). However, they have only considered simple shapes of FNs, perhaps avoiding more general representations that could involve complex calculations.
3) Closely related to the above problem, we need to introduce a partial order on the dataset (to enable us) to find

the model that reaches the lowest fuzzy error. Although different partial orders can be considered on the set of all FNs, no one is universally accepted (see [5] and [22]).

4) When fuzzy random variables (FRVs) are considered in regression analysis for constructing the relationship between explanatory and response variables, this problem becomes, in general, more complex and we can find difficulties in searching for optimal solutions for nonlinear problems especially. This has led most researchers to consider the problem in some particular cases (see [19]–[21], [25], and [34]–[36]). For simplification, they considered FNs with very simple shapes of membership functions, i.e., data are restricted to numeric intervals or triangular FNs and/or the model is restricted to the linear case.
5) Finally, taking into account that an empirical study can involve, at the same time, fuzzy and crisp variables, a fuzzy regression methodology must be able to treat, in a unified way, both cases. The methodologies that only consider noncrisp FNs [for instance, triangular or trapezoidal fuzzy numbers (TFNs)] cannot give an appropriate response to this problem. Our methodology provides a procedure to cover all possibilities. In this sense, from our point of view, fuzzy metrics must extend real metrics, which are more intuitively for researchers. If two FNs are very similar to real numbers, then the distance between them must be approached by the considered distance over real numbers.

In this paper, we consider these five problems from a new and broad perspective. In order to answer the first three raised questions, we introduce a partial order and a family of distance measures on the whole set of FNs (other authors only consider triangular or TFNs) depending on a variety of concepts closely related to the fuzzy setting, and we study their metric properties. Thus, we provide several ways to determine an FN that can be interpreted as the nearness between two FNs. Our notion involves some of the most important possibilistic and geometric elements of any FN (its center, its spreads, its level cuts, and some real pseudosemimetrics) so as to define the corresponding concepts for the distance measure. The main advantages of this family are the following: They are defined on the whole set of all FNs (not only on TFNs); some of the most useful subsets of the set of all FNs (including TFNs) are closed under these distance measures; in some cases, they are genuine fuzzy metrics over TFNs; finally, following classical techniques, it is possible to prove that this set is provided with a Hausdorff, first (or second) countable topology. With respect to the fourth raised problem, our model considers TFNs, and we can obtain linear and nonlinear solutions. On the one hand, we have considered TFNs because they have more intuitive and more natural interpretations. Moreover, it is always possible to consider an approximation operator that produces a TFN that is the closest to the given original FN among all TFNs having an expected interval identical to the original (see [16]). On the other hand, the aim of this paper is to find the optimal model, and it naturally leads to nonlinear solutions in the examples.

The main contribution of the new formulation developed in this paper is to be able to consider fuzzy semidistances to obtain easily linear and nonlinear fuzzy regression models and consider a wide range of fuzzy data that are appropriate to handle either vagueness or uncertainty in real-life studies. The proposed solution in this paper, which considers fuzzy distance measures, is consistent with fuzzy data type, and the associated numerical estimation process can be considered easy to apply, in contrast with other methods which become quite complex in practice. In this paper, we focus on the observational situation where the response variable is fuzzy, and the explanatory variables are crisp quantitative characters. We also explain how other fuzzy regression problems may be reduced to this case.

This paper is organized as follows. In Section II, some notations and preliminary results are presented. In Section III, we introduce a partial order and a family of new fuzzy distance measures between FNs and study its metric properties. After that, in Section IV, the fuzzy regression model is introduced and the problem is formulated. Then, we propose an estimation procedure to solve this problem and provide some real-life examples to illustrate the usefulness and applicability of the developed theory. Finally, some remarks and conclusions are presented in Section V.

II. Preliminaries

There exist many different notions of *FN*, in which its maximum of the membership function can be less than 1 or in which its level sets are not necessarily intervals (see [38] and [39]). For our purpose, we will use the following one. Let $\mathbb{I} = [0,1]$. A *fuzzy set on* $\mathbb{R}$ is a map $\mathcal{A} : \mathbb{R} \to \mathbb{I}$. A *FN* on $\mathbb{R}$ is a fuzzy set $\mathcal{A}$ on $\mathbb{R}$ such that, for all $\alpha \in]0,1]$, the *α-level set* (or *α-cut*) $\mathcal{A}_\alpha = \{x \in \mathbb{R} : \mathcal{A}(x) \geq \alpha\}$ is a nonempty, closed subinterval of $\mathbb{R}$. For the details of FNs, we refer the reader to [11]–[13], [23], and [41].

The *kernel* (or *core*) of an FN $\mathcal{A}$ is $\ker \mathcal{A} = \mathcal{A}_1$, and its *support* is the closure $\operatorname{supp} \mathcal{A} = \overline{\{x \in \mathbb{R} : \mathcal{A}(x) > 0\}}$. In the following, we will only consider FNs with compact support, and we will write $\mathcal{A}_0 = \operatorname{supp} \mathcal{A}$ and $\mathcal{A}_\alpha = [\, \underline{a}_\alpha, \overline{a}_\alpha \,]$ for all $\alpha \in \mathbb{I}$. Let $\mathcal{F}$ be the family of all FNs (with compact support). The number $\mathcal{D}_c\mathcal{A} = (\underline{a}_1 + \overline{a}_1)/2$ is the *center* of the FN $\mathcal{A}$, and its *radius* is $\operatorname{rad} \mathcal{A} = (\overline{a}_1 - \underline{a}_1)/2 \geq 0$. The mappings $\underline{a}, \overline{a} : \mathbb{I} \to \mathbb{R}$ completely determine the FN $\mathcal{A}$: a fuzzy set $\mathcal{A}$ is an FN if, and only if, there exist two left continuous mappings $\underline{a}, \overline{a} : \mathbb{I} \to \mathbb{R}$ such that $\underline{a}$ is nondecreasing, $\overline{a}$ is nonincreasing, and $\mathcal{A}_\alpha = [\, \underline{a}_\alpha, \overline{a}_\alpha \,]$ for all $\alpha \in \mathbb{I}$ (see [15]).

Following [29], we will say that a mapping is *finite* (respectively, *countable*) if its image is a finite (respectively, countable) set. In this sense, it is possible to consider *finite* (respectively, *countable*) *FNs*. It is not difficult to prove that an FN $\mathcal{A}$ is finite (respectively, countable) if, and only if, $\underline{a}$ and $\overline{a}$ are finite (respectively, countable) mappings.

A *(generalized) left right fuzzy number* (*LRFN*) *of Dubois and Prade* [11], [12] is an FN $\mathcal{A} = (a_1/a_2/a_3/a_4)_{LR}$, where $a_1, a_2, a_3, a_4 \in \mathbb{R}$ (*corners* of $\mathcal{A}$), $a_1 \leq a_2 \leq a_3 \leq a_4$, defined by

$$\mathcal{A}(x) = \begin{cases} L\left(\dfrac{x-a_1}{a_2-a_1}\right), & \text{if } a_1 < x < a_2 \\ 1, & \text{if } a_2 \leq x \leq a_3 \\ R\left(\dfrac{a_4-x}{a_4-a_3}\right), & \text{if } a_3 < x < a_4 \\ 0, & \text{in any other case} \end{cases}$$

where $L, R : \mathbb{I} \to \mathbb{I}$ are strictly increasing, continuous mappings such that $L(0) = R(0) = 0$ and $L(1) = R(1) = 1$. Clearly, the kernel of $\mathcal{A}$ is $[a_2, a_3]$ and its support is $[a_1, a_4]$. Note that an LRFN can only be a discontinuous mapping at $x = a_2$ and $x = a_3$.

TFN are special cases of LRFN with $L(x) = R(x) = x$ for all $x \in \mathbb{I}$ (we will denote them by $\mathcal{A} = (a_1/a_2/a_3/a_4)$). TFNs can be uniquely determined by their kernels and their supports, i.e., if $\mathcal{A}$ and $\mathcal{B}$ are TFNs, then $\mathcal{A} = \mathcal{B}$ if, and only if, $\ker \mathcal{A} = \ker \mathcal{B}$ and $\operatorname{supp} \mathcal{A} = \operatorname{supp} \mathcal{B}$.

A TFN is *real* (respectively, *rectangular*, *triangular*) if $a_1 = a_2 = a_3 = a_4$ (respectively, $a_1 = a_2 \leq a_3 = a_4$, $a_1 \leq a_2 = a_3 \leq a_4$). A real TFN is known as a *crisp FN*. The number $\mathcal{D}_c\mathcal{A} = A^c = (a_2 + a_3)/2$ is its *center* and its *spreads* are $\operatorname{rad} \mathcal{A} = A^m = (a_3 - a_2)/2 \geq 0$, $A^\ell = a_2 - a_1 \geq 0$ and $A^s = a_4 - a_3 \geq 0$. The center and the spreads also determine the TFN. This way, we can write $\mathcal{A} = \operatorname{Tra}(A^c, A^m, A^\ell, A^s)$. If $\mathcal{A}$ is triangular, we will denote it by $\operatorname{Tri}(A^c, A^\ell, A^s)$, and if it also is symmetric, by $\operatorname{Tri}(A^c, A)$, where A is its only spread.

Let $\mathcal{T}$ be the family of all (generalized) TFNs on $\mathbb{R}$. Clearly, $\mathbb{R}$ can be embedded in $\mathcal{T}$: Each $r \in \mathbb{R}$ is associated with $\tilde{r} \in \mathcal{T}$, where $\tilde{r}(x) = 1$ if $x = r$ and $\tilde{r}(x) = 0$ if $x \neq r$.

There exists an arithmetic between FNs (see [23]) that extends the real arithmetic. However, the most usual way to operate with FNs is throughout the intervalar arithmetic with the α-level sets and the functions $\underline{a}$ and $\overline{a}$. Thus, we define, for every $\mathcal{A}, \mathcal{B} \in \mathcal{F}$ and all $r > 0$, $\mathcal{C} = \mathcal{A} + \mathcal{B}$ by $\underline{c}_\alpha = \underline{a}_\alpha + \underline{b}_\alpha$ and $\overline{c}_\alpha = \overline{a}_\alpha + \overline{b}_\alpha$, and $\mathcal{E} = r\mathcal{A}$ by $\underline{e}_\alpha = r\underline{a}_\alpha$ and $\overline{e}_\alpha = r\overline{a}_\alpha$. When $r \in \mathbb{R}$ and $\mathcal{A} \in \mathcal{F}$, we will write $r + \mathcal{A} = \tilde{r} + \mathcal{A}$.

Let $(\Omega, \mathfrak{A}, P)$ be a probability space, and let $(\mathbb{R}^p, \mathfrak{B})$ be a measurable space and its Borel σ-field. A measurable function $X : \Omega \to \mathbb{R}^N$ is a *random vector* (RV). Puri and Ralescu [27] introduced the concept of FRV as an extension of both random variables and fuzzy sets. In this context, a mapping $\mathcal{X} : \Omega \to \mathcal{T}$ is a *trapezoidal FRV* if the representation of $\mathcal{X}$, $(X_1, X_2, X_3, X_4) : \Omega \to \mathbb{R}^4$, is an RV.

A *fuzzy negation* is a nonincreasing mapping $q : \mathbb{I} \to \mathbb{I}$ such that $q(0) = 1$ and $q(1) = 0$. Examples of continuous fuzzy negations are the *standard negation* $q(\alpha) = 1 - \alpha$, the *cosine negation* $q(\alpha) = (1 + \cos(\pi\alpha))/2$, the *Sugeno negation* $q(\alpha) = (1 - \alpha)/(1 + \lambda\alpha)$, and the *Yager negation* $q(\alpha) = (1 - \alpha^\lambda)^{1/\lambda}$.

For our purpose, a *defuzzification* will be a process in which fuzzy quantities are approximated by scalars or crisp FNs, i.e., a mapping $\mathcal{D} : \mathcal{F} \to \mathbb{R}$. There exist lots of different *valuation methods* with a variety of names: *mean value*, *ambiguity*, *expected value*, *center of gravity*, *area under FNs*, etc. (see, for instance, [4] and [39]). We will not assume any additional condition on the defuzzification $\mathcal{D}$.

Lemma 1: If $\mathcal{D}$ is a defuzzification and $\mathcal{A} \in \mathcal{T}$, then there exists a unique $\mathcal{D}^*\mathcal{A} \in \mathcal{T}$ such that $\mathcal{A} = \mathcal{D}\mathcal{A} + \mathcal{D}^*\mathcal{A}$.

Note that $\mathcal{D}_c\mathcal{D}_c^*\mathcal{A} = 0$ for all $\mathcal{A} \in \mathcal{T}$. In particular, taking $\mathcal{D} = \mathcal{D}_c$, if $\mathcal{A} \in \mathcal{T}$, then there exists a unique $r \in \mathbb{R}$ and $\mathcal{B} \in \mathcal{T}$ such that $\mathcal{A} = r + \mathcal{B}$ and $\mathcal{D}_c\mathcal{B} = 0$.

Let $\mathbb{R}_0^- = (-\infty, 0]$, $\mathbb{R}_0^+ = [0, \infty)$ and $\mathbb{R}^+ = (0, \infty)$. Following [31], a *metric on a set* S is a mapping $d : S \times S \to \mathbb{R}_0^+$ verifying, for all $x, y, z \in S$, (i) $d(x, x) = 0$; (ii) if $d(x, y) = 0$, then $x = y$; (iii) $d(x, y) = d(y, x)$; (iv) $d(x, z) \leq d(x, y) + d(y, z)$.

A mapping $d : S \times S \to \mathbb{R}_0^+$ is a *pseudometric* (respectively, a *semimetric*; *pseudosemimetric*) *on* S if it satisfies (i), (iii), and (iv) [respectively, (i), (ii), and (iii); (i) and (iii)].

The *average rate of change* of a mapping $f : S \to \mathbb{R}$ on an interval $I = [a, b] \subseteq S \subseteq \mathbb{R}$, with $a < b$, is $\Delta f_{[a,b]} = (f(b) - f(a))/(b - a)$.

A *partition* of the interval $\mathbb{I}$ is a set $\mathcal{P} = \{\delta_0, \delta_1, \ldots, \delta_n\}$ such that $0 = \delta_0 < \delta_1 < \cdots < \delta_n = 1$. The simplest partition of $\mathbb{I}$ is $\mathcal{P}_0 = \{0 = \delta_0 < \delta_1 = 1\}$. If $\mathcal{P} = \{\delta_i\}_{i=0}^n$ is a partition of $\mathbb{I}$ and $f : S \to \mathbb{R}$ is a mapping defined on $S \supseteq \mathbb{I}$, we will denote, for all $i \in \{1, 2, \ldots, n\}$,

$$\Delta f_i = \Delta f_{[\delta_{i-1}, \delta_i]} = \frac{f(\delta_i) - f(\delta_{i-1})}{\delta_i - \delta_{i-1}}.$$

III. Fuzzy Distance Measures Between Fuzzy Numbers and a Partial Order

In this section, we introduce a family of distance measures between arbitrary FNs depending on a variety of concepts closely related to the fuzzy setting: an arbitrary defuzzification, the radius of the kernel of an FN, a partition of $\mathbb{I}$, and some real pseudosemimetrics. Then, we study its metric properties, which justify why we are interested in this collection. First, we introduce a partial order on the set of all FNs (which, in general, is not constrained to a particular class of FNs). However, its best properties are verified on the set of all trapezoidal FNs. Throughout this section, let $\mathcal{D} : \mathcal{F} \to \mathbb{R}$ be a defuzzification, and let $\mathcal{P} = \{\delta_i\}_{i=0}^n$ be a partition of $\mathbb{I}$.

Definition 2: For all FNs $\mathcal{A}, \mathcal{B} \in \mathcal{F}$, define $\mathcal{A} \preccurlyeq \mathcal{B}$ *w.r.t.* $(\mathcal{D}, \mathcal{P})$ if, and only if, $\mathcal{D}\mathcal{A} \leq \mathcal{D}\mathcal{B}$, $\operatorname{rad} \mathcal{A} \leq \operatorname{rad} \mathcal{B}$, $\Delta\underline{a}_i \leq \Delta\underline{b}_i$ and $\Delta\overline{a}_i \geq \Delta\overline{b}_i$ for all $i \in \{1, 2, \ldots, n\}$. If we consider $(\mathcal{D}_c, \mathcal{P})$, we will only write $\mathcal{A} \preccurlyeq \mathcal{B}$.

Theorem 3: The relationship $\preccurlyeq$ w.r.t. $(\mathcal{D}, \mathcal{P})$ is reflexive and transitive. Furthermore, if there exist $\alpha, \beta \in \mathbb{R}$, verifying $\alpha + \beta \neq 0$, such that $\mathcal{D}\mathcal{C} = \alpha\,\underline{c}_1 + \beta\,\overline{c}_1$ for all $\mathcal{C} \in \mathcal{T}$, then $\preccurlyeq$ is a partial order on $\mathcal{T}$. Moreover, if $\mathcal{A} \preccurlyeq \mathcal{B}$, then $\mathcal{A} + \mathcal{C} \preccurlyeq \mathcal{B} + \mathcal{C}$ and $r\mathcal{A} \preccurlyeq r\mathcal{B}$ for all $r > 0$.

In the following, we will handle metrics valued in the set of all FNs. Then, we need to extend the notion of real metric to metric valued in the own set.

Definition 4: Let 0_S be a point of a set S provided with a partial order $\sqsubseteq$. Consider the set $S_{0,\sqsubseteq}^+ = \{x \in S : 0_S \sqsubseteq x\}$ and let $s : S_{0,\sqsubseteq}^+ \times S_{0,\sqsubseteq}^+ \to S_{0,\sqsubseteq}^+$ be a mapping. A *distance function on* $(S, 0_S, \sqsubseteq, s)$ (or a *metric*) is a mapping $d : S \times S \to S_{0,\sqsubseteq}^+$

TABLE I
SOME DISTANCE MEASURES $D(\mathcal{A},\mathcal{B})$, WHERE $q_0 = q = 1$, $\phi_i(x,y) = \psi(x,y) = |x-y|^p$ AND $\varphi_i(x,y) = ||x|-|y||^p$

Trapezoidal	$\mathcal{D} = \mathcal{D}_c,\, p = 1$	$\mathcal{D}\mathcal{A} = \overline{a}_1,\, p = 0.8$	$\mathcal{D}\mathcal{A} = \int_{\mathbb{R}} \mathcal{A}(t)dt,\, p = 0.8$
FNs	$q_1(\alpha) = q_2(\alpha) = 1-\alpha$		$q_1(\alpha) = (1-\alpha^{2.5})^{0.4},\, q_2(\alpha) = (1-\alpha)/(1+2\alpha)$
$\mathcal{A} = (5.1/7/11.3)$ $\mathcal{B} = (7/8/9.8)$	$(0.1/1/3.5)$	$(0.08083/1/3.0814)$	$L(t) \approx 2.70/(7.78-3.91t) - 0.5,\, 0.61 \le t \le 1.53$ $R(t) \approx (1-0.16(t-1.53)^{2.5})^{0.4},\, 1.53 \le t \le 3.61$
$\mathcal{A} = (7/8/9/10)$ $\mathcal{B} = (3/5/7/9)$	$(1/2/3/4)$	$(0.167/1.167/2.315/3.315)$	$L(t) \approx 3/(6.667-4t) - 0.5,\, 0.1667 \le t \le 1.1667$ $R(t) \approx (1-(t-2.315)^{2.5})^{0.4},\, 2.315 \le t \le 3.315$
$\mathcal{A} = (-2.6/0/2.5/8.5)$ $\mathcal{B} = (-3.5/-1.5/0.5/4.1)$	$(0.9/1.5/2/4.4)$	$(0.747/1.411/2.071/4.085)$	$L(t) \approx 4.46/(15.61-8.955t) - 0.5,\, 0.747 \le t \le 1.411$ $R(t) \approx (1-0.174(t-2.071)^{2.5})^{0.4},\, 2.071 \le t \le 4.085$

verifying, for all $x, y, z \in S$,

$$\text{(i) } d(x,x) = 0_S; \quad \text{(ii) if } d(x,y) = 0_S, \text{ then } x = y;$$
$$\text{(iii) } d(x,y) = d(y,x); \quad \text{(iv) } d(x,z) \sqsubseteq s(d(x,y), d(y,z)).$$

We also say that $(S, 0_S, s)$ is a *metric space w.r.t. the partial order* $\sqsubseteq$. The function d is:

1) a *pseudometric* if it satisfies (i), (iii), and (iv);
2) a *semimetric* (*on* $(S, 0_S)$) if it satisfies (i)–(iii);
3) a *pseudosemimetric*(*on* $(S, 0_S)$) if it satisfies (i) and (iii).

Next, we describe a mapping $D : \mathcal{F} \times \mathcal{F} \to \mathcal{F}$ which assigns, to each pair $(\mathcal{A},\mathcal{B})$ of FNs, another FN $D(\mathcal{A},\mathcal{B})$. Our definition is based in the following geometrical considerations.

1) The center of $D(\mathcal{A},\mathcal{B})$ must only depends on the points that researcher considers they are more important for $\mathcal{A}$ and $\mathcal{B}$, that is, the researcher must consider a defuzzification $\mathcal{D} : \mathcal{F} \to \mathbb{R}$ that highlights a concrete point of each FN (its mean, its ambiguity, etc.). The distance between these points will be determinant to calculate the center of $D(\mathcal{A},\mathcal{B})$. As a consequence, the property $\mathcal{D}_c(D(\mathcal{A},\mathcal{B})) = \phi_0(\mathcal{D}\mathcal{A}, \mathcal{D}\mathcal{B})$ seems to be reasonable if ϕ_0 measures distances between real numbers.
2) The spread of $D(\mathcal{A},\mathcal{B})$ must essentially depends on the spreads of $\mathcal{A}$ and $\mathcal{B}$. In this sense, the larger are the radius of $\mathcal{A}$ and $\mathcal{B}$, the greater should be the radius of their similarity measure. As a consequence, we will understand that $spr D(\mathcal{A},\mathcal{B}) = \psi(spr\mathcal{A}, spr\mathcal{B})$ is the best possibility when ψ measures the distance between nonnegative real numbers.
3) The researcher may have the control on the type of FN he/she wants to obtain. Then, we will introduce some functions q_1 and q_2 that allow the researcher to decide whether the obtained distance if an LR, trapezoidal, triangular, rectangular, or crisp FN, depending on his/her study.
4) In the support of $D(\mathcal{A},\mathcal{B})$, we will have some influence the kernels and the supports of $\mathcal{A}$ and $\mathcal{B}$, and the way in which we measure the distance between them. In fact, we will take into account a distance measure between n of their α-cuts.
5) If the researcher considers that it is important, the way in which we determine the distance between two FNs will be able to extend the real notions in two senses: 1) if a and b are real numbers, then $D(\tilde{a}, \tilde{b})$ must be coherent with the way in which the researcher measures the real distance between a and b (which, for instance, could be $|a-b|$); and 2) some researchers used, in the family of all triangular FNs, real distance measures as $D[(a/b/c), (a'/b'/c')] = (a-a')^2 + (b-b')^2 + (c-c')^2$. As the reader can easily check, this distance measure is not a true distance in the classical sense because it does not satisfy the triangular inequality. However, it has such good properties (for instance, in fuzzy regression, it is considered in the *least squares method*) that our definition must cover this case. As the previous expression defines a pseudosemimetric in the set of all triangular FNs, we will consider this kind of similarity measures [1].

As a consequence of the previous considerations, we propose the following family of distance measures between FNs. Notice that, in some particular cases, a similar metric was already considered in [37].

Definition 5: Let $q, q_0 \ge 0$, let $q_1, q_2 : \mathbb{I} \to [0, \infty[$ be two left continuous nonincreasing mappings on $\mathbb{I}$ and let $\phi_0 : \mathbb{R} \times \mathbb{R} \to \mathbb{R}_0^+$, $\psi : \mathbb{R}_0^+ \times \mathbb{R}_0^+ \to \mathbb{R}_0^+$, $\{\phi_i : \mathbb{R}_0^+ \times \mathbb{R}_0^+ \to \mathbb{R}_0^+\}_{i=1}^n$ and $\{\varphi_i : \mathbb{R}_0^- \times \mathbb{R}_0^- \to \mathbb{R}_0^+\}_{i=1}^n$ be pseudosemimetrics on their respective domains. For all FNs $\mathcal{A},\mathcal{B} \in \mathcal{F}$ and all $\alpha \in \mathbb{I}$, define

$$\underline{D(\mathcal{A},\mathcal{B})}_\alpha = q_0\phi_0(\mathcal{D}\mathcal{A}, \mathcal{D}\mathcal{B}) - q\psi(\text{rad}\,\mathcal{A}, \text{rad}\,\mathcal{B}) - q_1(\alpha)\sum_{i=1}^{n}\phi_i(\Delta\underline{a}_i, \Delta\underline{b}_i) \tag{1}$$

$$\overline{D(\mathcal{A},\mathcal{B})}_\alpha = q_0\phi_0(\mathcal{D}\mathcal{A}, \mathcal{D}\mathcal{B}) + q\psi(\text{rad}\,\mathcal{A}, \text{rad}\,\mathcal{B}) + q_2(\alpha)\sum_{i=1}^{n}\varphi_i(\Delta\overline{a}_i, \Delta\overline{b}_i). \tag{2}$$

Let $D(\mathcal{A},\mathcal{B})$ be the only FN determined by its α-cuts (1), (2).

Table I shows how the mapping D acts in different cases. In the first column, we use $\mathcal{D}_c$; in the second one, we introduce $\mathcal{D}\mathcal{A} = \overline{a}_1$; finally, we study the area defuzzification $\mathcal{D}\mathcal{A} = \int_{\mathbb{R}} \mathcal{A}(t)dt$ using the Sugeno negation $q_1(\alpha) = (1-\alpha^{2.5})^{0.4}$ and

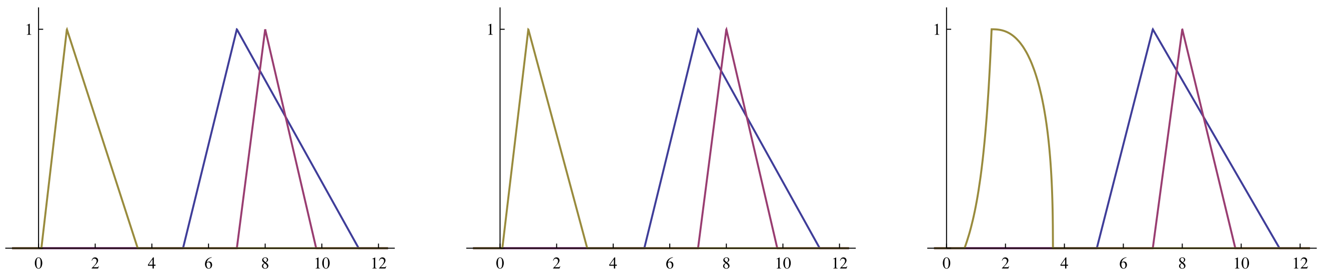

Fig. 1. Action of the mapping D in different cases, when $\mathcal{A} = (5.1/7/11.3)$ and $\mathcal{B} = (7/8/9.8)$.

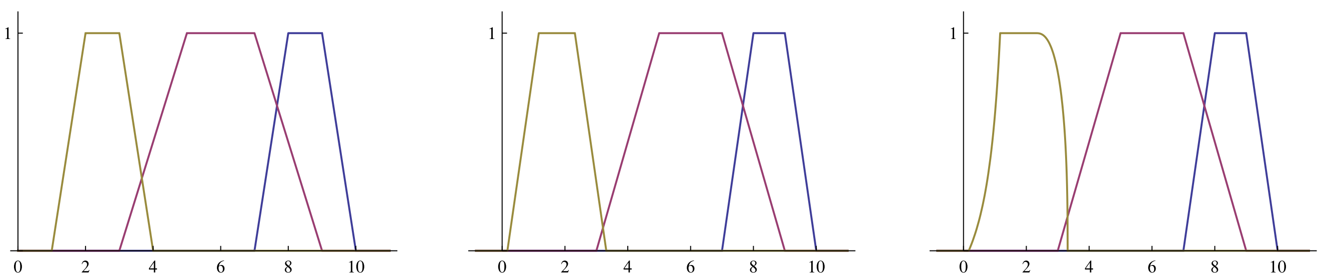

Fig. 2. Idem when $\mathcal{A} = (7/8/9/10)$ and $\mathcal{B} = (3/5/7/9)$.

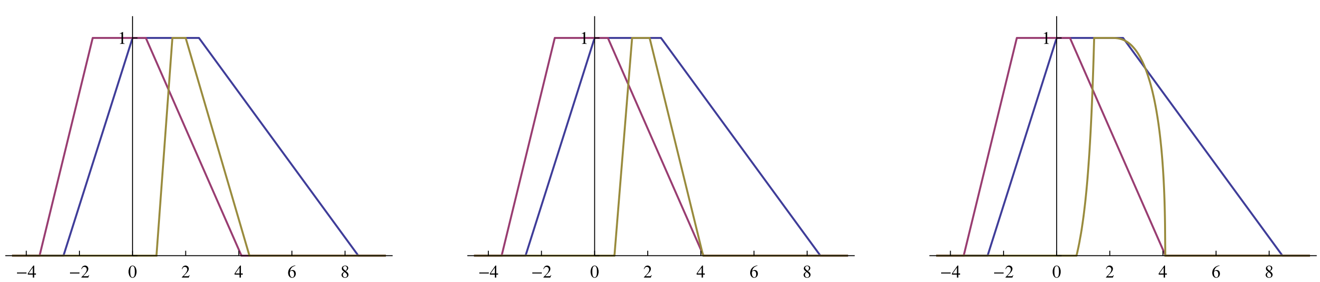

Fig. 3. Idem when $\mathcal{A} = (-2.6/0/2.5/8.5)$ and $\mathcal{B} = (-3.5/-1.5/0.5/4.1)$.

the Yager negation $q_2(\alpha) = (1-\alpha)/(1+2\alpha)$. The fuzzy distance measures in each case are plotted on Figs. 1–3.

Taking into account the applications, we are deeply interested in the defuzzification $\mathcal{D}_c$, the mappings $q_i(\alpha) = 1 - \alpha$ for all $\alpha \in \mathbb{I}$, and the partial order $\preccurlyeq$ on $\mathcal{T}$. However, many properties are true in a more general context. Therefore, we are going to deal with arbitrary fuzzy distance measures.

Lemma 6: If $\mathcal{A}, \mathcal{B} \in \mathcal{F}$, then $D(\mathcal{A}, \mathcal{B}) \in \mathcal{F}$, $D(\mathcal{B}, \mathcal{A}) = D(\mathcal{A}, \mathcal{B})$, and $D(\mathcal{A}, \mathcal{A}) = \tilde{0}$.

The first property why we consider the distance measure D is that some of the most useful subsets of $\mathcal{F}$ (including $\mathcal{T}$) are closed under D. Thus, if a researcher is handling, for instance, triangular FNs, then their distance is also triangular.

Theorem 7: 1. If q_1 and q_2 are the standard negation and $\mathcal{A}$ and $\mathcal{B}$ are trapezoidal (respectively, triangular, rectangular, real) FNs, then $D(\mathcal{A}, \mathcal{B})$ is also trapezoidal (respectively, triangular, rectangular, real). In addition to this, if $q_0, q > 0$ and $\mathcal{D} = \mathcal{D}_c$, then $D : \mathcal{T} \times \mathcal{T} \to \mathcal{T}$ is a semimetric on $(\mathcal{T}, \tilde{0})$.

2. If q_1 and q_2 are finite (respectively, countable) mappings, then $D(\mathcal{A}, \mathcal{B})$ is also finite (respectively, countable) for all $\mathcal{A}, \mathcal{B} \in \mathcal{F}$.

Proof: It is clear that if q_1 and q_2 are the standard negation, then $D(\mathcal{A}, \mathcal{B}) \in \mathcal{T}$ for all $\mathcal{A}, \mathcal{B} \in \mathcal{F}$. In particular, if $\mathcal{A}, \mathcal{B} \in \mathcal{T}$, then $D(\mathcal{A}, \mathcal{B}) \in \mathcal{T}$.

Suppose that $\mathcal{A}$ and $\mathcal{B}$ are crisp FNs. Then $\underline{a}_\alpha = \overline{a}_\alpha = a$ for all $\alpha \in \mathbb{I}$; therefore, rad $\mathcal{A} = \Delta\underline{a}_i = \Delta\overline{a}_i = 0$ for all i. As the same is true for $\mathcal{B}$, then $\underline{D(\mathcal{A}, \mathcal{B})}_\alpha = \overline{D(\mathcal{A}, \mathcal{B})}_\alpha = q_0\phi_0\left(\mathcal{D}\mathcal{A}, \mathcal{D}\mathcal{B}\right)$ for all $\alpha \in \mathbb{I}$; therefore, $D(\mathcal{A}, \mathcal{B})$ is a crisp FN.

Suppose that $\mathcal{A}$ and $\mathcal{B}$ are rectangular FNs. Then, $\underline{a}_\alpha = a$ and $\overline{a}_\alpha = a'$ for all $\alpha \in \mathbb{I}$ (being $a \leq a'$); therefore, $\Delta\underline{a}_i = \Delta\overline{a}_i = 0$ for all i. As the same is true for $\mathcal{B}$, then $\underline{D(\mathcal{A}, \mathcal{B})}_\alpha = q_0\phi_0\left(\mathcal{D}\mathcal{A}, \mathcal{D}\mathcal{B}\right) - q\psi(\text{rad } \mathcal{A}, \text{rad } \mathcal{B})$ and $\overline{D(\mathcal{A}, \mathcal{B})}_\alpha = q_0\phi_0\left(\mathcal{D}\mathcal{A}, \mathcal{D}\mathcal{B}\right) + q\psi(\text{rad } \mathcal{A}, \text{rad } \mathcal{B})$ for all $\alpha \in \mathbb{I}$. Since these functions are constant on $\mathbb{I}$, then $D(\mathcal{A}, \mathcal{B})$ is a rectangular FN.

Suppose that $\mathcal{A}$ and $\mathcal{B}$ are triangular FNs. Then, rad $\mathcal{A} =$ rad $\mathcal{B} = 0$. Since rad $D(\mathcal{A},\mathcal{B}) = q\psi(\text{rad}\,\mathcal{A}, \text{rad}\,\mathcal{B}) = 0$, then $D(\mathcal{A},\mathcal{B})$ is also triangular. The trapezoidal case is obvious.

Next, we prove the triangular inequality. Since ϕ_0 is a pseudometric, by the second ball item on page 6

$$\begin{aligned}\mathcal{D}_c(D(\mathcal{A},\mathcal{C})) &= q_0\phi_0(\mathcal{D}_c\mathcal{A},\mathcal{D}_c\mathcal{C})\\ &\leq q_0\phi_0(\mathcal{D}_c\mathcal{A},\mathcal{D}_c\mathcal{B}) + q_0\phi_0(\mathcal{D}_c\mathcal{B},\mathcal{D}_c\mathcal{C})\\ &= \mathcal{D}_c(D(\mathcal{A},\mathcal{B})) + \mathcal{D}_c(D(\mathcal{B},\mathcal{C}))\\ &= \mathcal{D}_c(D(\mathcal{A},\mathcal{B}) + D(\mathcal{B},\mathcal{C})).\end{aligned}$$

In a similar way, since ψ is a pseudometric, by the third ball item on page 6

$$\begin{aligned}\text{rad}\, D(\mathcal{A},\mathcal{C}) &= q\,\psi(\text{rad}\,\mathcal{A}, \text{rad}\,\mathcal{C})\\ &\leq q\,\psi(\text{rad}\,\mathcal{A}, \text{rad}\,\mathcal{B}) + q\psi(\text{rad}\,\mathcal{B}, \text{rad}\,\mathcal{C})\\ &= \text{rad}\, D(\mathcal{A},\mathcal{B}) + \text{rad}\, D(\mathcal{B},\mathcal{C})\\ &= \text{rad}\,(D(\mathcal{A},\mathcal{B}) + D(\mathcal{B},\mathcal{C})).\end{aligned}$$

Now, define $\mathcal{G} = D(\mathcal{A},\mathcal{C})$ and $\mathcal{H} = D(\mathcal{A},\mathcal{B}) + D(\mathcal{B},\mathcal{C})$. This means that

$$\begin{aligned}\underline{h}_\alpha &= \underline{D(\mathcal{A},\mathcal{B})}_\alpha + \underline{D(\mathcal{B},\mathcal{C})}_\alpha\\ &= q_0\left[\,\phi_0(\mathcal{D}_c\mathcal{A},\mathcal{D}_c\mathcal{B}) + \phi_0(\mathcal{D}_c\mathcal{A},\mathcal{D}_c\mathcal{B})\,\right]\\ &\quad - q\left[\,\psi(\text{rad}\,\mathcal{A}, \text{rad}\,\mathcal{B}) + \psi(\text{rad}\,\mathcal{B}, \text{rad}\,\mathcal{C})\,\right]\\ &\quad - q_1(\alpha)\sum_{i=1}^{n}\left[\,\phi_i\,(\Delta\underline{a}_i, \Delta\underline{b}_i) + \phi_i\,(\Delta\underline{b}_i, \Delta\underline{c}_i)\,\right].\end{aligned}$$

Therefore

$$\begin{aligned}\Delta\underline{h}_i &= \frac{\underline{h}_{\delta_i} - \underline{h}_{\delta_{i-1}}}{\delta_i - \delta_{i-1}} = \frac{q_1(\delta_{i-1}) - q_1(\delta_i)}{\delta_i - \delta_{i-1}}\\ &\quad \cdot \sum_{j=1}^{n}\left[\,\phi_j\left(\Delta\underline{a}_j, \Delta\underline{b}_j\right) + \phi_j\left(\Delta\underline{b}_j, \Delta\underline{c}_j\right)\,\right].\end{aligned}$$

Similarly, since $\underline{g}_\alpha = \underline{D(\mathcal{A},\mathcal{C})}_\alpha = q_0\phi_0(\mathcal{D}_c\mathcal{A},\mathcal{D}_c\mathcal{C}) - q\psi(\text{rad}\,\mathcal{A}, \text{rad}\,\mathcal{C}) - q_1(\alpha)\sum_{i=1}^{n}\phi_i\,(\Delta\underline{a}_i, \Delta\underline{c}_i)$, it is not difficult to prove that

$$\Delta\underline{g}_i = \frac{\underline{g}_{\delta_i} - \underline{g}_{\delta_{i-1}}}{\delta_i - \delta_{i-1}} = \frac{q_1(\delta_{i-1}) - q_1(\delta_i)}{\delta_i - \delta_{i-1}}\sum_{j=1}^{n}\phi_j\left(\Delta\underline{a}_j, \Delta\underline{c}_j\right).$$

Since q_1 is nonincreasing, then $q_1(\delta_{i-1}) - q_1(\delta_i) \geq 0$. If $q_1(\delta_{i-1}) - q_1(\delta_i) = 0$ for some i, then $\Delta\underline{g}_i = \Delta\underline{h}_i$. If $q_1(\delta_{i-1}) - q_1(\delta_i) > 0$, since each ϕ_j is a pseudometric, then $\Delta\underline{g}_i \leq \Delta\underline{h}_i$. Repeating this argument, we can prove

$$\Delta\overline{g}_i = \frac{q_2(\delta_i) - q_2(\delta_{i-1})}{\delta_i - \delta_{i-1}}\sum_{j=1}^{n}\varphi_j(\Delta\overline{a}_j, \Delta\overline{c}_j)$$

$$\Delta\overline{h}_i = \frac{q_2(\delta_i) - q_2(\delta_{i-1})}{\delta_i - \delta_{i-1}}\sum_{j=1}^{n}[\varphi_j(\Delta\overline{a}_j, \Delta\overline{b}_j) + \varphi_j(\Delta\overline{b}_j, \Delta\overline{c}_j)].$$

Since q_2 is nonincreasing, then $q_2(\delta_i) - q_2(\delta_{i-1}) \leq 0$. If $q_2(\delta_i) = q_2(\delta_{i-1})$ for some i, then $\Delta\overline{g}_i = \Delta\overline{h}_i$. If $q_2(\delta_i) - q_2(\delta_{i-1}) < 0$, then $\Delta\overline{g}_i \geq \Delta\overline{h}_i$. This proves that $D(\mathcal{A},\mathcal{C}) \preccurlyeq D(\mathcal{A},\mathcal{B}) + D(\mathcal{B},\mathcal{C})$. ■

The following properties let us to interpret the fuzzy distance measure D as a fuzzy metric on $(\mathcal{T}, \tilde{0}, +)$ w.r.t. the partial order $\preccurlyeq$. Indeed, if $q_1(1) = q_2(1) = 0$, then

1) $\ker D(\mathcal{A},\mathcal{B}) = [q_0\phi_0(\mathcal{D}\mathcal{A},\mathcal{D}\mathcal{B}) - q\psi(\text{rad}\,\mathcal{A}, \text{rad}\,\mathcal{B}),$ $q_0\phi_0(\mathcal{D}\mathcal{A},\mathcal{D}\mathcal{B}) + q\psi(\text{rad}\,\mathcal{A}, \text{rad}\,\mathcal{B})]$.
2) $\mathcal{D}_c(D(\mathcal{A},\mathcal{B})) = q_0\phi_0(\mathcal{D}\mathcal{A},\mathcal{D}\mathcal{B})$.
3) rad $D(\mathcal{A},\mathcal{B}) = q\psi(\text{rad}\,\mathcal{A}, \text{rad}\,\mathcal{B})$.
4) $\tilde{0} \preccurlyeq D(\mathcal{A},\mathcal{B})$.
5) If ϕ_0, ψ, ϕ_i, and φ_i are pseudometrics, then $D(\mathcal{A},\mathcal{C}) \preccurlyeq D(\mathcal{A},\mathcal{B}) + D(\mathcal{B},\mathcal{C})$ w.r.t. $(\mathcal{D}_c, \mathcal{P})$.

Theorem 8: If $q_0, q > 0$, $\mathcal{D} = \mathcal{D}_c$ and q_1 and q_2 are the standard negation, then D verifies the following properties for all $\mathcal{A},\mathcal{B},\mathcal{C} \in \mathcal{T}$:

(i) $D(\mathcal{A},\mathcal{A}) = \tilde{0}$;

(ii) if $D(\mathcal{A},\mathcal{B}) = \tilde{0}$, then $\mathcal{A} = \mathcal{B}$;

(iii) $D(\mathcal{A},\mathcal{B}) = D(\mathcal{B},\mathcal{A})$;

(iv) $D(\mathcal{A},\mathcal{C}) \preccurlyeq D(\mathcal{A},\mathcal{B}) + D(\mathcal{B},\mathcal{C})$

whatever the metrics ϕ_0, ψ, ϕ_i and φ_i and the partition $\mathcal{P}$.

Proof: The fourth ball item on page 6 assures us that $D(\mathcal{A},\mathcal{B}) \succcurlyeq \tilde{0}$ for all $\mathcal{A},\mathcal{B} \in \mathcal{F}$. Lemma 6 implies that $D(\mathcal{B},\mathcal{A}) = D(\mathcal{A},\mathcal{B})$ and $D(\mathcal{A},\mathcal{A}) = \tilde{0}$ for all $\mathcal{A},\mathcal{B} \in \mathcal{F}$. The fifth ball item on page 6 shows that $D(\mathcal{A},\mathcal{C}) \preccurlyeq D(\mathcal{A},\mathcal{B}) + D(\mathcal{B},\mathcal{C})$ w.r.t. $(\mathcal{D}_c, \mathcal{P})$ for all $\mathcal{A},\mathcal{B},\mathcal{C} \in \mathcal{F}$. Now suppose that $\mathcal{A},\mathcal{B} \in \mathcal{T}$ verify $D(\mathcal{A},\mathcal{B}) = \tilde{0}$. Then, $q_0\phi_0(\mathcal{D}_c\mathcal{A},\mathcal{D}_c\mathcal{B}) = 0$, $q\psi(\text{rad}\,\mathcal{A}, \text{rad}\,\mathcal{B}) = 0$ and $\phi_i\,(\Delta\underline{a}_i, \Delta\underline{b}_i) = \varphi_i\left(\Delta\overline{a}_i, \Delta\overline{b}_i\right) = 0$ for all $i \in \{1, 2, \ldots, n\}$. Since $q, q_0 > 0$ and ϕ_0, ψ, ϕ_i, and φ_i are metrics, then $\mathcal{D}_c\mathcal{A} = \mathcal{D}_c\mathcal{B}$, rad $\mathcal{A} =$ rad $\mathcal{B}$, $\Delta\underline{a}_i = \Delta\underline{b}_i$ and $\Delta\overline{a}_i = \Delta\overline{b}_i$ for all $i \in \{1, 2, \ldots, n\}$. This means that $\mathcal{A} \preccurlyeq \mathcal{B}$ and $\mathcal{B} \preccurlyeq \mathcal{A}$. Therefore, Theorem 3 concludes that $\mathcal{A} = \mathcal{B}$ (since $\preccurlyeq$ is a partial order). ■

Using this fuzzy metric, it is possible to prove that the family $\{\mathcal{N}_r(\mathcal{A})\}_{r>0}$, where

$$\mathcal{N}_r(\mathcal{A}) = \left\{\mathcal{B} \in \mathcal{T} : \begin{bmatrix} q_0\phi_0\,(\mathcal{D}\mathcal{A},\mathcal{D}\mathcal{B}) < r\\ q\psi(\text{rad}\,\mathcal{A}, \text{rad}\,\mathcal{B}) < r\\ \phi_1\,(\Delta\underline{a}, \Delta\underline{b}) < r\\ \varphi_1\left(\Delta\overline{a}, \Delta\overline{b}\right) < r \end{bmatrix}\right\}$$

is a neighborhood system at $\mathcal{A}$ in a Hausdorff, first countable topology on $\mathcal{T}$.

IV. Procedure to Solve a Fuzzy Regression Problem and Illustrative Examples

In this section, the regression model with real explanatory variables and a trapezoidal response FRV is established using the fuzzy distance measures between TFNs introduced in the previous section. Next, the fuzzy total error is calculated, and the optimal solution is obtained considering the partial order $\preccurlyeq$. Finally, we explain the estimation procedure distinguishing three cases: The first two cases consider whether the spreads depend on the explanatory variables or not and the third case considers as input an RV. We describe each one in detail and provide illustrative examples to clarify the uses of these. We point out that, since different researchers can be interested in different defuzzifications, the first part of the methodology is very general. However, for simplicity, in studied cases, we will only consider the defuzzification $\mathcal{D}_c$.

Let $\underline{X}, \mathcal{Y}$ be two variables where $\underline{X} = (X_1, ..., X_N)'$ is an RV and $\mathcal{Y}$ is a trapezoidal[1] FRV. We are interested in analyzing the relationship between $\mathcal{Y}$ and $\underline{X}$. The regression model we consider can be formalized as

$$\mathcal{Y} = \text{Tra}\,(Y^c_{\underline{X},\underline{a}_c} + \varepsilon^c, M_{\underline{X},\underline{a}_m} + \varepsilon^m, I_{\underline{X},\underline{a}_\ell} + \varepsilon^\ell, S_{\underline{X},\underline{a}_s} + \varepsilon^s) \tag{3}$$

where $\varepsilon^c, \varepsilon^m, \varepsilon^\ell$, and ε^s are the residuals (i.e., real-valued RVs such that $E[\varepsilon^c|\underline{X}] = E[\varepsilon^m|\underline{X}] = E[\varepsilon^\ell|\underline{X}] = E[\varepsilon^s|\underline{X}] = 0$ and whose variances are finite) and $\underline{a}_c = (a_{c1}, \ldots, a_{cN})'$, $\underline{a}_m = (a_{m1}, \ldots, a_{mN})'$, $\underline{a}_\ell = (a_{\ell 1}, \ldots, a_{\ell N})'$, and $\underline{a}_s = (a_{s1}, \ldots, a_{sN})'$ are the $(N \times 1)$-vectors of the parameters related to the vector $\underline{X}$. Therefore, the conditional expectation, that is, the population regression function, is

$$E[\mathcal{Y}|\underline{X}] = \text{Tra}\,(Y^c_{\underline{X},\underline{a}_c}, M_{\underline{X},\underline{a}_m}, I_{\underline{X},\underline{a}_\ell}, S_{\underline{X},\underline{a}_s}).$$

The function $\mathcal{Z} : \text{dom}\,\underline{X} \to \mathcal{T}$ that we are interested in obtaining to predict $\mathcal{Y}$ from $\underline{X}$ must be defined as $\mathcal{Z}_{\underline{X}} = \text{Tra}\,(Y^c_{\underline{X}}, M_{\underline{X}}, I_{\underline{X}}, S_{\underline{X}})$, where $Y^c : \text{dom}\,\underline{X} \to \mathbb{R}$ and $M, I, S : \text{dom}\,\underline{X} \to \mathbb{R}^+_0$ are arbitrary functions.

Consider a random experiment in which we observe the variable $(\underline{X}, \mathcal{Y})$ on n statistical units, i.e., suppose that we have a random sample $\{\underline{X}_i, \mathcal{Y}_i = \text{Tra}\,(Y_i, M_i, I_i, S_i)\}_{i=1}^n$ obtained from $(\underline{X}, \mathcal{Y})$. If we consider a distance measure D defined as in (1) and (2), the objective function in terms of the parameters $\underline{a}_c$, $\underline{a}_m$, $\underline{a}_\ell$, and $\underline{a}_s$ of model (3) is

$$\mathcal{E} = \sum_{i=1}^n D(\mathcal{Y}_i, \mathcal{Z}_{\underline{X}_i}) \tag{4}$$

that is, we are looking for a function $\mathcal{Z}_{\underline{X}} = E[\mathcal{Y}|\underline{X}]$ such that the fuzzy total error (4) is as small as possible w.r.t. the partial order $\preccurlyeq$ on $\mathcal{T}$.

Next, we justify that the least-squares method can be seen as a particular case of our estimation process.

Theorem 9: If $\mathcal{D} = \mathcal{D}_c$, $\mathcal{P} = \mathcal{P}_0$, q_1 and q_2 are the standard negation, $q_0 = q = 1$, and $\phi_0(x,y) = \psi(x,y) = \phi_i(x,y) = \varphi_i(x,y) = |x-y|^2$ for all x, y in their respective domains, then the solution of the corresponding regression problem coincides with the solution of the least-squares method.

This result encourages us to consider the widely used least-squares method for several reasons: Linear and nonlinear least-squares software is available in many statistical software packages, it is easy to use and understand, one should be able to apply it to find a good approximation for any functions, it has advantages in computation, it is more suitable for learning, etc. Our aim is not to improve this method, but using this criterion to develop a fuzzy methodology that is more useful and powerful.

[1]Note that if the response variable $\mathcal{Y}$ is not trapezoidal, it is always possible to consider an approximation operator which produces a trapezoidal FN $\mathcal{Y}'$ that is the closest to the given original FN among all trapezoidal FNs having an expected interval identical to the original (see [16]).

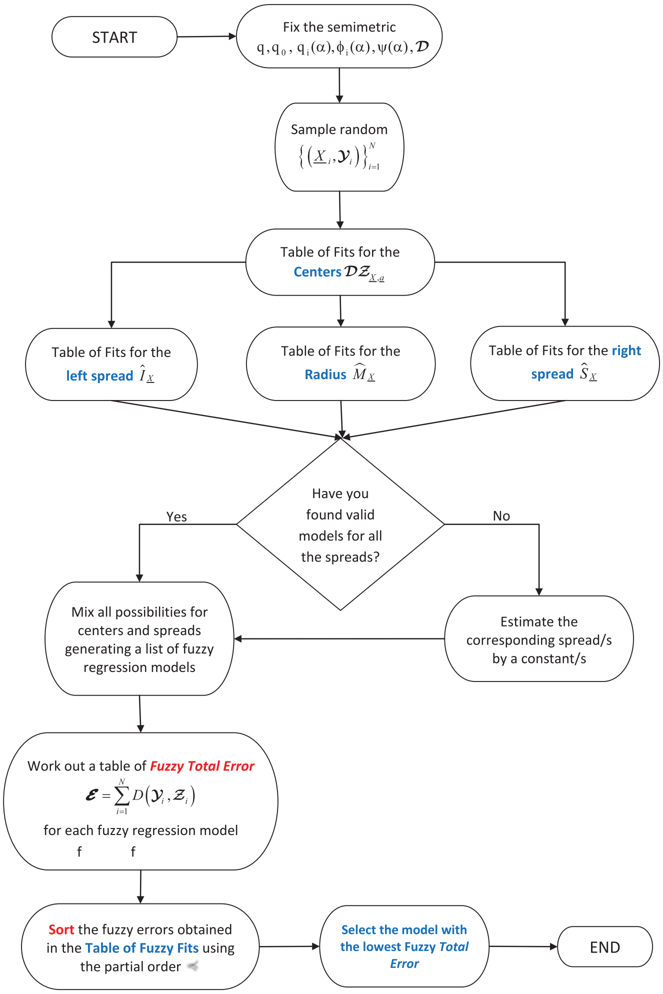

Fig. 4. Brief flow diagram to summarize the fuzzy regression method.

Next, we explain our methodology to solve the problem. Using Lemma 1, we may decompose $\mathcal{Z} = \mathcal{DZ} + \mathcal{D}^*\mathcal{Z}$, where $\mathcal{DZ}_{\underline{x}} \in \mathbb{R}$ and $\mathcal{D}^*\mathcal{Z}_{\underline{x}} \in \mathcal{T}$. We consider the real random sample $\{(\underline{X_i}, \mathcal{DY}_i)\}_{i=1}^n$ and the least-squares method in order to fit a model $\widehat{\mathcal{DZ}}$ as follows: We propose using statistical packages that may fit a variety of functional forms, listing the models in decreasing order of R-squared in a table, usually called *Table of fits*.

Therefore, after the comparison between the goodness-of-fit statistics in the *Table of fits* obtained by the statistical package, we can determine some of the best fits, $\widehat{\mathcal{DZ}}$, for $\mathcal{DZ}$ (in the sense of Remark 11). Next, we have to estimate the trapezoidal regression function $\mathcal{D}^*\mathcal{Z}$. However, we need to distinguish whether $\mathcal{D}^*\mathcal{Z}$ depends on the explanatory variable or it is sufficient to consider a constant. For clarity, before considering both cases, a summary of the methodology we propose is the following (see also Fig. 4):

Step 1: Generate a *Table of fits* with a statistical package listing the models for the center and the spreads in decreasing order of some goodness-of-fit statistics, such as R-squared (see, for instance, Table II).

Step 2: Use the metric $\mathcal{D}$ (selected by the researcher from the beginning; for instance, $\mathcal{D}_c$) and calculate the fuzzy errors $\mathcal{E}$

TABLE II
TABLE OF FITS FOR THE CENTER (LEFT SIDE) AND FOR THE LEFT (OR RIGHT) SPREAD (RIGHT SIDE) IN EXAMPLE 12

Model	Function	R-squared	Model	Function	R-squared
Squared Root-Y	$(29.9912 + 0.000107839x)^2$	94.80%	Multiplicative	$0.016733x^{0.878572}$	65.26 %
Squared-X	$1807.91 + 1.83701 \cdot 10^{-8}x^2$	94.26%	Squared Root-Y	$(18.2688 + 0.00004375x)^2$	64.01%
Multiplicative	$0.0155958x^{0.999421}$	91.58%	Exponential	$\exp(5.9244 + 0.00000248x)$	62.47%
Linear	$0.0180685x - 672.731$	91.46%	Double reciprocal	$1/(0.000237192 + 246.094/x)$	61.45%
Exponential	$\exp(7.2127 + 0.000002872x)$	90.62%	Linear	$0.0034698x + 185.714$	60.75%
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

TABLE III
TABLE OF ERRORS $\mathcal{E} = \sum_{i=1}^{n} D(\mathcal{Y}_i, \mathcal{Z}_{\underline{x_i}})$, SCALED BY 10^{-7}

Spreads \ Centers	$(29.99 + 0.0001x)^2$	$1807 + 1.84 \cdot 10^{-8}x^2$	$0.0156x^{0.999}$	$0.018x - 672.7$	$e^{2.87\times10^{-6}x+7.21}$
$0.0167x^{0.879}$	$(2.74/4.56/6.39)$	$(2.6/4.43/6.25)$	$(6.51/8.34/10.16)$	$(4.76/6.59/8.41)$	$(4.60/6.42/8.24)$
$(18.27 + 0.000044x)^2$	$(2.78/4.56/6.35)$	$(2.64/4.43/6.21)$	$(6.55/8.34/10.12)$	$(4.80/6.59/8.37)$	$(4.64/6.42/8.20)$
$e^{2.48\times10^{-6}x+5.92}$	$(2.46/4.56/6.67)$	$(2.32/4.43/6.53)$	$(6.23/8.34/10.44)$	$(4.48/6.59/8.69)$	$(4.31/6.42/8.53)$
$\frac{x}{0.000237x + 246.1}$	$(1.79/4.56/7.34)$	$(1.65/4.43/7.20)$	$(5.56/8.34/11.11)$	$(3.81/6.59/9.36)$	$(3.65/6.42/9.20)$
$185.7 + 0.0035x$	$(2.88/4.56/6.25)$	$(2.75/4.43/6.11)$	$(6.66/8.34/10.02)$	$(4.91/6.59/8.27)$	$(4.74/6.42/8.10)$

for each of the possible pairings of the centers and spreads, that is, for each of the fuzzy regression models. These results can be presented in a table (see, for instance, Table III).

Step 3: Sort the fuzzy models with the partial order $\preccurlyeq$ introduced in Definition 2 taking into account the fuzzy errors obtained in Step 2 for each one. In the examples, these results are shown in a table named *Table of fuzzy fits*. Note that the order established between the models using real numbers in step 1, which would seem logical to keep it, do not necessarily coincide with the ranking obtained when models are arranged in this step according to the fuzzy order above introduced (see for instance, Table V).

Step 4: To find the optimal solution of the fuzzy regression problem, we choose the functions corresponding to the lowest fuzzy error, i.e., once we have determined $\widehat{\mathcal{D}^*\mathcal{Z}_{\underline{x}}}$ following the previous considerations, we will estimate the regression model by $\widehat{\mathcal{Z}}_{\underline{x}} = \widehat{\mathcal{D}\mathcal{Z}} + \widehat{\mathcal{D}^*\mathcal{Z}_{\underline{x}}} = (\widehat{\mathcal{Z}}^1_{\underline{x}}/\widehat{\mathcal{Z}}^2_{\underline{x}}/\widehat{\mathcal{Z}}^3_{\underline{x}}/\widehat{\mathcal{Z}}^4_{\underline{x}}) \in \mathcal{T}$.

Note that steps 2 and 3 are the key steps to understand that our methodology is clearly different from those used previously by other authors: fuzzy errors are calculated and sorted according to a fuzzy partial order on the set of all trapezoidal FNs. Therefore, our technique is not limited to classical multivariate analysis, but, as we shall see in examples, the best estimate is achieved with fuzzy functions that could not be obtained with the simple use of classical real techniques.

Remark 10: Applying the previous methodology, we will obtain the fuzzy regression models $\mathcal{Z}_{\underline{x}} = \widehat{\mathcal{D}_c\mathcal{Z}_{\underline{x}}} + \widehat{\mathcal{D}^*_c\mathcal{Z}_{\underline{x}}} = (\mathcal{Z}^1_{\underline{x}}/\mathcal{Z}^2_{\underline{x}}/\mathcal{Z}^3_{\underline{x}}/\mathcal{Z}^4_{\underline{x}}) \in \mathcal{T}$. To evaluate the goodness-of-fit of the different models, some authors have mainly considered two numerical estimation of the following statistics.

1) On the one hand, we can calculate the total sum of squares due to error $\text{SSE} = \sum_{j=1}^{4} \text{SSE}_j$ (see [7] and [30]), where $\{\text{SSE}_j\}_{j=1}^{4}$ are given by

$$\text{SSE}_j = \sum_{i=1}^{n} \left(\mathcal{Y}^j_i - \mathcal{Z}^j_{\underline{x}_i}\right)^2, \quad 1 \le j \le 4$$

being $\mathcal{Y}_i = (\mathcal{Y}^1_i/\mathcal{Y}^2_i/\mathcal{Y}^3_i/\mathcal{Y}^4_i)$ for all $i \in \{1, 2, \ldots, n\}$. Some authors (see [7]) also used as comparative measure the total error sum $\sum_{j=1}^{4} E_j$, where $E_j = \frac{\text{SSE}_j}{4}$, for all $1 \le j \le 4$.

2) On the other hand, other authors (see, for instance, [9]) considered an R-squared (or adjusted R-squared) coefficient given, using our distance measure, by

$$\widetilde{R}^2 = 1 - \frac{\text{SSE}}{\text{SST}} = 1 - \frac{\sum_{j=1}^{4}\sum_{i=1}^{n}\left(\mathcal{Y}^j_i - \mathcal{Z}^j_{\underline{x}_i}\right)^2}{\sum_{j=1}^{4}\sum_{i=1}^{n}\left(\mathcal{Y}^j_i - \overline{\mathcal{Y}^j_i}\right)^2}$$

where $\overline{\mathcal{Y}^j_i} = \frac{1}{n}\sum_{i=1}^{n}\mathcal{Y}^j_i$.

Remark 11: The simplest regression models involve a single response variable and a single predictor variable. In this case, statistical packages fit a variety of functional forms and support different goodness-of-fit statistics for parametric models: 1) the sum of squares due to error (SSE) and (2) R-square or Adjusted R-square. Most statistical packages list the models in decreasing order of R-squared. R-squared values range from 0% to 100%. A higher R-squared value will indicate a more useful model. A low R-squared means you should ignore the corresponding model. Note that if we increase the number of fitted coefficients in our model, R-squared will increase although the fit may not

TABLE IV
TABLE OF ERRORS $10^{-7}\,\mathcal{E} = 10^{-7}\sum_{i=1}^{n} D(\mathcal{Y}_i, \mathcal{Z}_{\underline{x_i}})$

Spreads \ Centers	$(29.99 + 0.0001x)^2$	$1807 + 1\ .84 \cdot 10^{-8}x^2$	$0.0156x^{0.999}$	$0.018x - 672.7$	$e^{2.87\times10^{-6}x+7.21}$
$0.0167x^{0.879}$	Tri $(4.56, 1.82)$	Tri $(4.43, 1.82)$	Tri $(8.34, 1.82)$	Tri $(6.59, 1.82)$	Tri $(6.42, 1.82)$
$(18.27 + 0.000044x)^2$	Tri $(4.56, 1.78)$	Tri $(4.43, 1.78)$	Tri $(8.34, 1.78)$	Tri $(6.59, 1.78)$	Tri $(6.42, 1.78)$
$e^{2.48\times10^{-6}x+5.92}$	Tri $(4.56, 2.11)$	Tri $(4.43, 2.11)$	Tri $(8.34, 2.11)$	Tri $(6.59, 2.11)$	Tri $(6.42, 2.11)$
$\dfrac{x}{0.000237x + 246.1}$	Tri $(4.56, 2.78)$	Tri $(4.43, 2.78)$	Tri $(8.34, 2.78)$	Tri $(6.59, 2.78)$	Tri $(6.42, 2.78)$
$185.7 + 0.0035x$	Tri $(4.56, 1.68)$	Tri $(4.43, 1.68)$	Tri $(8.34, 1.68)$	Tri $(6.59, 1.68)$	Tri $(6.42, 1.68)$

TABLE V
TABLE OF FUZZY FITS (ORDERED BY THE FUZZY CRITERION)

Spreads \ Centers	$1807 + 1.84 \cdot 10^{-8}x^2$		$(29.99 + 0.0001x)^2$		$e^{2.87\times10^{-6}x+7.21}$		$0.018x - 672.7$		$0.0156x^{0.999}$
$185.7 + 0.0035x$	Tri $(4.43, 1.68)$	$\preccurlyeq$	Tri $(4.56, 1.68)$	$\preccurlyeq$	Tri $(6.42, 1.68)$	$\preccurlyeq$	Tri $(6.59, 1.68)$	$\preccurlyeq$	Tri $(8.34, 1.68)$
	$\curlywedge$		$\curlywedge$		$\curlywedge$		$\curlywedge$		$\curlywedge$
$(18.27 + 0.000044x)^2$	Tri $(4.43, 1.78)$	$\preccurlyeq$	Tri $(4.56, 1.78)$	$\preccurlyeq$	Tri $(6.42, 1.78)$	$\preccurlyeq$	Tri $(6.59, 1.78)$	$\preccurlyeq$	Tri $(8.34, 1.78)$
	$\curlywedge$		$\curlywedge$		$\curlywedge$		$\curlywedge$		$\curlywedge$
$0.0167x^{0.879}$	Tri $(4.43, 1.82)$	$\preccurlyeq$	Tri $(4.56, 1.82)$	$\preccurlyeq$	Tri $(6.42, 1.82)$	$\preccurlyeq$	Tri $(6.59, 1.82)$	$\preccurlyeq$	Tri $(8.34, 1.82)$
	$\curlywedge$		$\curlywedge$		$\curlywedge$		$\curlywedge$		$\curlywedge$
$e^{2.48\times10^{-6}x+5.92}$	Tri $(4.43, 2.11)$	$\preccurlyeq$	Tri $(4.56, 2.11)$	$\preccurlyeq$	Tri $(6.42, 2.11)$	$\preccurlyeq$	Tri $(6.59, 2.11)$	$\preccurlyeq$	Tri $(8.34, 2.11)$
	$\curlywedge$		$\curlywedge$		$\curlywedge$		$\curlywedge$		$\curlywedge$
$\dfrac{x}{0.000237x + 246.1}$	Tri $(4.43, 2.78)$	$\preccurlyeq$	Tri $(4.56, 2.78)$	$\preccurlyeq$	Tri $(6.42, 2.78)$	$\preccurlyeq$	Tri $(6.59, 2.78)$	$\preccurlyeq$	Tri $(8.34, 2.78)$

improve in a practical sense. To avoid this situation, we should use the degrees of freedom adjusted R-square statistic.

Next, we examine the three cases considered in detail and provide illustrative examples.

A. *Estimation Procedure When the Spreads Depend on the Explanatory Variables*

Suppose that using a statistical package, we have determined some of the best fits, $\widehat{\mathcal{DZ}}$, for $\mathcal{DZ}$. Next, we describe how to estimate $\mathcal{D}^*\mathcal{Z}$. For simplicity, henceforth in this section, we will consider the case $\mathcal{D} = \mathcal{D}_c$ and q_1 and q_2 are the standard negation, because in this case Theorem 8 assures us that D is a fuzzy metric. Actually, in practice, the well-known regression methodologies (for instance, the least square method) use semimetrics. Therefore, we will assume that $\phi_0(x, y) = |x - y|^{p_0}$, $\psi(x, y) = |x^k - y^k|^p$, $\phi_i(x, y) = |x^{k_i} - y^{k_i}|^{p_i}$ and $\varphi_i(x, y) = ||x|^{k_i'} - |y|^{k_i'}|^{p_i'}$ for all x, y in their respective domains.

The parameters in the above expressions must be chosen taking into account some properties such as robustness, stability, computational efficiency, etc. For example, if we consider $|x - y|$, the statistical optimization technique is known as *least absolute deviations* or L_1*-norm problem*, and if we choose its square, $(x - y)^2$, we are considering the popular least-squares technique. Both of them attempt to find a function that closely approximates the set of data. However, the method of least absolute deviations finds applications in many areas, due to its robustness (it is resistant to outliers in the data) compared with the least-squares method. In contrast, the least-squares solutions are stable and can be computed efficiently, and the problem always has one solution. For instance, the proposed methodology lets researchers use the least-squares method to fit a model for the center and least absolute deviation for the spreads. They are not forced to keep the same semimetric in all the fitting procedures. Furthermore, recently, Wu and Yang [43] replaced the Euclidean distance in their objective function by the exponential-type of the Euclidean distance because of its property of robustness to noise and outliers.

In this case, the function $\mathcal{D}_c^*\mathcal{Z}_{\underline{x}} = \text{Tra}\ (0, M_{\underline{x}}, I_{\underline{x}}, S_{\underline{x}})$ can also be estimated using some criteria of least squares. However, we must take into consideration that the estimated functions for the spreads, $\hat{M}_{\underline{x}}$, $\hat{I}_{\underline{x}}$ and $\hat{S}_{\underline{x}}$ should all take nonnegative values (note that since the function $\mathcal{DZ}$ takes values on $\mathbb{R}$, no sign constraints were considered for the centers). Therefore, among the different functions that we could calculate, we should be careful to choose the best fitted models (see Remark 11) among those that only take nonnegative values and are well defined in the set of all points for which we are interested in obtaining an prediction. This procedure is illustrated in the following example.

Example 12: Ferraro *et al.* [14] analyzed the dependence relationship of the Retail Trade Sales (in millions of dollars) of the U.S. in 2002 by kind of business on the number of employees (see Table II in [14]). In this example, we compare their results (which consider transforming the spreads by means of the logarithmic transformation) with those obtained with our methodology. Each interval can be seen, as usual, as a symmetric

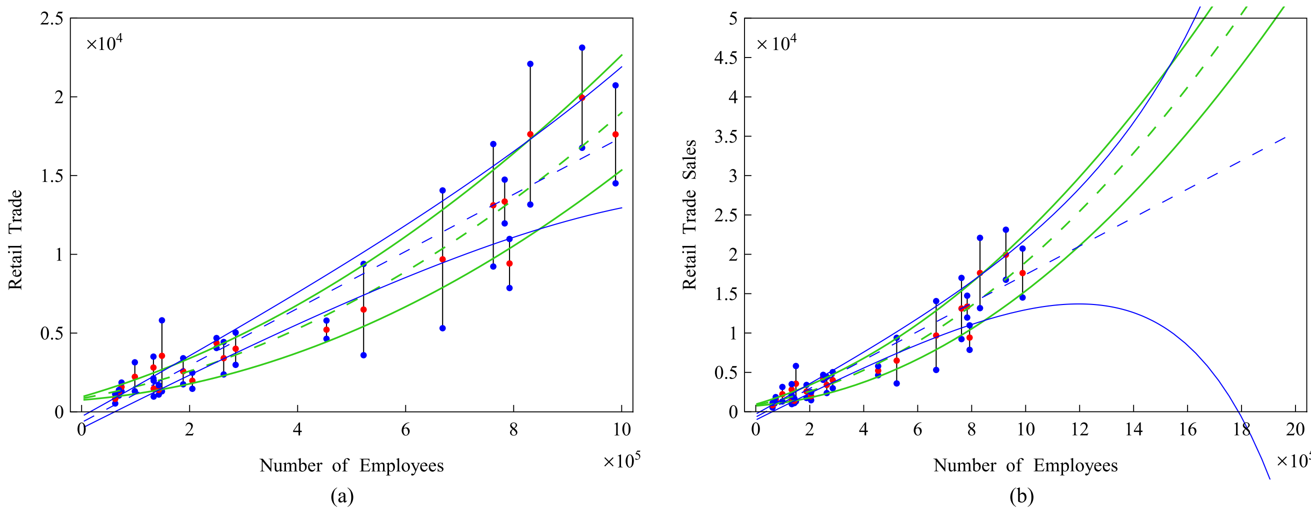

Fig. 5. Comparison between FeM (in blue color) and RoM (in green color). (a) Comparison in the domain of the data. (b) Large-scale prediction.

TABLE VI
COMPARISON OF SSE AND R^2 FOR MODELS FEM AND ROM

Fuzzy regression model	SSE$_1$	SSE$_2$ = SSE$_3$	SSE$_4$	SSE	$\Sigma_j E_j$	$\widetilde{R}^2$
Ferraro *et al.* [14]	7.99227×10^7	6.58725×10^7	9.39716×10^7	3.05639×10^8	7.64099×10^7	90.36%
Our methodology (RoM)	3.68159×10^7	4.42727×10^7	8.53534×10^7	2.10715×10^8	5.26787×10^7	93.99%

TFN considering its center and its spread. By symmetry, in this case, we will only have a unique spread (the same value is valid for the left and the right side).

Table II lists the different models obtained by a statistical package for the center and the spread using terms and notations that the program considers. Next, we calculate the fuzzy errors $\mathcal{E}$ for each of the possible pairings of the centers and spreads (see Table III). In this case, the error is a symmetric triangular FN that will be denoted by Tri $(a, b) =$ Tra $(a, 0, b, b) = (a - b/a/a/a + b))$ to reduce space in the table. Table IV shows these results. It is easy to sort the models for the centers and the spread, and Table V shows the order between them. Thus, our fuzzy regression model is $\widehat{\mathcal{Z}}_x =$ Tri $(\widehat{Y}_x, \widehat{I}_x, \widehat{S}_x)$, where

$$\widehat{Y}_x = 1807.91 + 1.83701 \times 10^{-8} x^2 \quad \text{and}$$
$$\widehat{I}_x = \widehat{S}_x = 0.0034698x + 185.714.$$

Next, we compare the results we have obtained with those obtained in [14]. The methodology developed in [14], denoted by FeM, uses for the center and the spread the functions called *Linear* and *Exponential*, respectively, in Table II. The methodology developed in this paper, denoted by RoM, indicates that the most appropriate models for the center and the spreads are those called squared-X and linear, respectively. As shown in Fig. 5, FeM presents several problems. When the number of employees is less than 6000, retail trade sales are negative. We also note that, due to the natural behavior of the exponential function, the estimated model FeM would have problems when the number of employees increases [see Fig. 5(b)].

TABLE VII
TANAKA'S DATA AND SPREADS

X	Y^I	Y^c	Y^S	I_i	S_i
1	6.2	8	9.8	1.8	1.8
2	4.2	6.48	6	2.2	2.2
3	6.9	9.5	12.1	2.6	2.6
4	10.9	13.5	16.1	2.6	2.6
5	10.6	13	15.4	2.4	2.4

Table VI includes the results of a comparison of the fuzzy regression models FeM and RoM using SSE and $\widetilde{R}^2$ (see Remark 10). Finally taking into account any goodness-of-fit measures considered, model RoM is more appropriate than FeM.

Remark 13: Note that some authors establish *a priori* non-negativity conditions on the functions $M_{\underline{x}}$, $I_{\underline{x}}$, and $S_{\underline{x}}$. In this context, for the problem of analyzing linear regression relationships, Coppi *et al.* [9] proposed a model with crisp input and LR fuzzy response. The authors considered an iterative least-squares estimation procedure and imposed a non-negative condition for the numerical minimization problem to avoid negative estimated spreads. Ferraro *et al.* [14] proposed an alternative model to overcome the non-negativity condition, because the inferences for the models with these restrictions are, commonly, more complex and less efficient (see [14] and

TABLE VIII
TABLE OF FITS FOR THE CENTER (LEFT SIDE) AND FOR THE LEFT (OR RIGHT) SPREAD (RIGHT SIDE) IN EXAMPLE 14

Model	Function	R-squared	Model	Function	R-squared
Double squared	$\sqrt{45.4676 + 5.80222x^2}$	78.84%	Double reciprocal	$1/(0.344065 + 0.208324/x)$	90.24%
Squared-X	$6.97118 + 0.28262x^2$	77.74%	Multiplicative	$1.87501 \cdot x^{0.212984}$	78.01%
Squared-Y	$\sqrt{3.905 + 35.129x}$	77.27%	Squared Root-Y - log X	$(1.37031 + 0.156125 \log x)^2$	77.04%
Root Y-Squared X	$(2.65312 + 0.0446295x^2)^2$	76.42%	Log X	$1.8806 + 0.458902 \log x$	75.94%
Linear	$4.95 + 1.71x$	76.09%	Linear	$1.84 + 0.16x$	57.14%
⋮	⋮	⋮	⋮	⋮	⋮

TABLE IX
TABLE OF ERRORS $\mathcal{E} = \sum_{i=1}^{n} D(\mathcal{Y}_i, \mathcal{Z}_{\underline{x_i}})$ IN EXAMPLE 14

Spreads \ Centers	$\sqrt{45.47 + 5.8x^2}$	$6.97 + 0.283x^2$	$\sqrt{3.9 + 35.1x}$	$(2.65 + 0.045x^2)^2$	$4.95 + 1.71x$
$1/(0.344 + 0.2/x)$	$(8.01/8.09/8.16)$	$(8.48/8.56/8.63)$	$(11.15/11.23/11.31)$	$(9.22/9.29/9.37)$	$(9.11/9.19/9.26)$
$1.87501 \cdot x^{0.212984}$	$(7.96/8.09/8.21)$	$(8.43/8.56/8.68)$	$(11.11/11.23/11.36)$	$(9.17/9.29/9.42)$	$(9.06/9.19/9.31)$
$(1.37 + 0.156 \log x)^2$	$(7.97/8.09/8.20)$	$(8.44/8.56/8.67)$	$(11.12/11.23/11.35)$	$(9.18/9.29/9.41)$	$(9.07/9.19/9.30)$
$1.88 + 0.46 \log x$	$(7.98/8.09/8.19)$	$(8.45/8.56/8.66)$	$(11.12/11.23/11.34)$	$(9.19/9.29/9.40)$	$(9.08/9.19/9.29)$
$1.84 + 0.16x$	$(7.89/8.09/8.28)$	$(8.36/8.56/8.75)$	$(11.04/11.23/11.42)$	$(9.10/9.29/9.48)$	$(9.00/9.19/9.38)$

TABLE X
TABLE OF ERRORS $\mathcal{E} = \sum_{i=1}^{n} D(\mathcal{Y}_i, \mathcal{Z}_{\underline{x_i}})$ IN EXAMPLE 14

Spreads \ Centers	$\sqrt{45.47 + 5.8x^2}$	$6.97 + 0.283x^2$	$\sqrt{3.9 + 35.1x}$	$(2.65 + 0.045x^2)^2$	$4.95 + 1.71x$
$1/(0.344 + 0.2/x)$	Tri $(8.09, 0.08)$	Tri $(8.56, 0.08)$	Tri $(11.23, 0.08)$	Tri $(9.29, 0.08)$	Tri $(9.19, 0.08)$
$1.875 \cdot x^{0.213}$	Tri $(8.09, 0.124)$	Tri $(8.56, 0.124)$	Tri $(11.23, 0.124)$	Tri $(9.29, 0.124)$	Tri $(9.19, 0.124)$
$(1.37 + 0.156 \log x)^2$	Tri $(8.09, 0.115)$	Tri $(8.56, 0.115)$	Tri $(11.23, 0.115)$	Tri $(9.29, 0.115)$	Tri $(9.19, 0.115)$
$1.88 + 0.46 \log x$	Tri $(8.09, 0.108)$	Tri $(8.56, 0.108)$	Tri $(11.23, 0.108)$	Tri $(9.29, 0.108)$	Tri $(9.19, 0.108)$
$1.84 + 0.16x$	Tri $(8.09, 0.19)$	Tri $(8.56, 0.19)$	Tri $(11.23, 0.19)$	Tri $(9.29, 0.19)$	Tri $(9.19, 0.19)$

references therein), introducing functions to transform the spreads into real numbers such as logarithmic transformations. We emphasize that, for this reason, we are against of imposing any sign constraints from the beginning.

The applicability of the proposed method is now shown with a comparison study.

Example 14 (Tanaka et al.): In 1987, Tanaka *et al.* designed an example to illustrate their fuzzy regression methodology for dealing with the problem of one numeric exploratory variable and one fuzzy dependent variable that took symmetric triangular values. They considered five pairs of observations listed in Table VII. Notice that, although the following dataset do not have the level of detail and complexity than those used in other studies, this example has the advantage that it has been considered in the literature by many researchers (see [7], [10], [19]–[21], [25], [30], [36], and [42]) for the experimental evaluation and comparison of their proposed methodology. This is why, we will use it as an comparison study.

Table VIII lists some of the best models for centers and spreads, ordered by an R-squared criterion. Table IX presents the fuzzy errors, calculated as $(a/b/c)$. To ranking these fuzzy errors by the introduced partial order, we show in Table X their corresponding expressions as Tri (A^c, A^s) (note that they are symmetric). Then, we order in Table XI all models by our fuzzy criterion: It is very important to notice that this ranking does not coincide with the R-squared criterion.

Finally, the model proposed by our methodology using the previous FNs ordering is, for all $x \in$ dom X,

$$\widehat{\mathcal{Z}}_x = \text{Tri}\left(\sqrt{45.47 + 5.8x^2}\ ,\ \frac{1}{0.344 + 0.21/x}\right).$$

Table XII lists a comparative study between our model and previous models by other authors, using SSE, $\Sigma_j E_j$ and $\widetilde{R}^2$.

Therefore, we deduce that the proposed purely fuzzy approach to the regression problem improves significantly the obtained results by other authors.

B. Estimation Procedure When Some Spreads Do Not Depend on the Explanatory Variables

There exist some cases in which we may suppose that the spreads do not depend on the explanatory variables, or in which the method described above does not provide good results since it does not lead to a significant model for some spread or because significant models always take negative values in the domain of $\underline{X}$. In these cases, we should fit a nonnegative constant considering an error function as described below. We emphasize that this problem has not been considered by other researchers using such general fuzzy semimetric D.

TABLE XI
TABLE OF FUZZY FITS (MODELS HAVE BEEN ORDERED BY THE FUZZY CRITERION) IN EXAMPLE 14

Spreads \ Centers	$\sqrt{45.47 + 5.8x^2}$		$6.97 + 0.283x^2$		$4.95 + 1.71x$		$(2.65 + 0.045x^2)^2$		$\sqrt{3.9 + 35.1x}$
$1/(0.344 + 0.2/x)$	Tri $(8.09, 0.08)$	$\preccurlyeq$	Tri $(8.56, 0.08)$	$\preccurlyeq$	Tri $(9.19, 0.08)$	$\preccurlyeq$	Tri $(9.29, 0.08)$	$\preccurlyeq$	Tri $(11.23, 0.08)$
	$\curlyvee$		$\curlyvee$		$\curlyvee$		$\curlyvee$		$\curlyvee$
$1.88 + 0.46 \log x$	Tri $(8.09, 0.108)$	$\preccurlyeq$	Tri $(8.56, 0.108)$	$\preccurlyeq$	Tri $(9.19, 0.108)$	$\preccurlyeq$	Tri $(9.29, 0.108)$	$\preccurlyeq$	Tri $(11.23, 0.108)$
	$\curlyvee$		$\curlyvee$		$\curlyvee$		$\curlyvee$		$\curlyvee$
$(1.37 + 0.156 \log x)^2$	Tri $(8.09, 0.115)$	$\preccurlyeq$	Tri $(8.56, 0.115)$	$\preccurlyeq$	Tri $(9.19, 0.115)$	$\preccurlyeq$	Tri $(9.29, 0.115)$	$\preccurlyeq$	Tri $(11.23, 0.115)$
	$\curlyvee$		$\curlyvee$		$\curlyvee$		$\curlyvee$		$\curlyvee$
$1.875 \cdot x^{0.213}$	Tri $(8.09, 0.124)$	$\preccurlyeq$	Tri $(8.56, 0.124)$	$\preccurlyeq$	Tri $(9.19, 0.124)$	$\preccurlyeq$	Tri $(9.29, 0.124)$	$\preccurlyeq$	Tri $(11.23, 0.124)$
	$\curlyvee$		$\curlyvee$		$\curlyvee$		$\curlyvee$		$\curlyvee$
$1.84 + 0.16x$	Tri $(8.09, 0.19)$	$\preccurlyeq$	Tri $(8.56, 0.19)$	$\preccurlyeq$	Tri $(9.19, 0.19)$	$\preccurlyeq$	Tri $(9.29, 0.19)$	$\preccurlyeq$	Tri $(11.23, 0.19)$

TABLE XII
TABLE OF FUZZY FITS (MODELS HAVE BEEN ORDERED BY THE FUZZY CRITERION) IN EXAMPLE 14

Author	Fuzzy regression model	*SSE*	$\Sigma_j E_j$	$\widetilde{R}^2$
Tanaka *et al.* (1987)	$\widehat{\mathcal{Y}}_{THW}(x) = (0, 3.85, 7.7) + 2.1x$	67.235	16.8088	56.51%
Diamond [10] (1988)	$\widehat{\mathcal{Y}}_{DM}(x) = (3.11, 4.95, 6.79) + (1.55, 1.71, 1.87)x$	37.132	9.283	75.98%
Kim and Bishu [21] (1998)	$\widehat{\mathcal{Y}}_{KB}(x) = (3.11, 4.95, 6.84) + (1.55, 1.71, 1.82)x$	37.207	9.30175	75.93%
Wu and Tseng [42] (2002)	$\widehat{\mathcal{Y}}_{WT}(x) = (3.11, 4.95, 6.79) + (1.55, 1.71, 1.87)x$	37.132	9.283	75.98%
Kao and Chyu [20] (2003)	$\widehat{\mathcal{Y}}_{KC}(x) = 4.81 + 1.72x + (-2.2, 0.12, 2.44)$	37.65	9.4125	75.65%
Nasrabadi and Nasrabadi [25] (2004)	$\widehat{\mathcal{Y}}_{NN}(x) = (2.36, 4.86, 7) + 1.73x$	38.542	9.6355	75.07%
Hojati *et al.* [19] (2005)	$\widehat{\mathcal{Y}}_{HBS}(x) = (5.1, 6.75, 8.4) + (1.1, 1.25, 1.4)x$	49.61	12.4025	67.91%
Chen and Hsueh [7] (2009)	$\widehat{\mathcal{Y}}_{CH}(x) = 1.71x + (2.63, 4.95, 7.27)$	37.644	9.411	75.65%
Roldán *et al.* [30] (2012)	$\widehat{\mathcal{Y}}_{RRM}(x) = \text{Tri}\left(4.95 + 1.71x, \dfrac{1}{0.344 + 0.208/x}\right)$	36.9017	9.22542	76.13%
Proposed approach	$\widehat{\mathcal{Y}}_{RRMA}(x) = \text{Tri}\left(\sqrt{45.47 + 5.8x^2}, \dfrac{1}{0.344 + 0.208/x}\right)$	32.4945	8.1262	78.98%

Hence, we assume that for some of the spreads it has not been possible to obtain a good-fitting model and we are interested in estimating a constant. Since minimizing functions for any of the spreads leads to similar error functions, it is sufficient to consider

$$\mathsf{e} : [0, \infty[\to [0, \infty[, \qquad \mathsf{e}(x) = \sum_{i=1}^{n} \left| x^k - a_i^k \right|^p \tag{5}$$

where $p > 0, k > 0$ and $0 \leq a_1 < a_2 < \cdots < a_n$ ($n \geq 2$ because $n = 1$ is a trivial case). Clearly, e is a continuous function in $[0, \infty[$ and of class C^∞ in $[0, \infty[\setminus \{a_1, a_2, \ldots, a_n\}$. Moreover, e is strictly decreasing in $[0, a_1[$ and strictly increasing in $]a_n, +\infty[$. Then, e has, at least, one absolute minimum on $[0, \infty[$, which is on $[a_1, a_n]$. This proves that the fuzzy regression problem always has a solution. In order to determine it, we have to distinguish between $p > 1$, $p = 1$ or $0 < p < 1$ (see more details in [30]).

If $p > 1$, then e is of class C^1 in $[0, \infty[$ and e' has a unique zero, which is its only absolute minimum. To determine this point, we have to solve the equation $\mathsf{e}'(x_0) = 0$ in $]a_1, a_n[$, which is equivalent to the system

$$\left\{ \begin{array}{c} \displaystyle\sum_{i=1}^{r} \left(x_0^k - a_i^k\right)^{p-1} = \sum_{i=r+1}^{n} \left(a_i^k - x_0^k\right)^{p-1} \\ \text{in the interval }]a_r, a_{r+1}] \end{array} \right\}_{r=1}^{r=n-1}. \tag{6}$$

If $p = 1$ in (5), then e has an absolute minimum in the median of the data set $\{a_1, a_2, \ldots, a_n\}$ (a complete study of the median of a FRV can be found in [32]). Finally, if $0 < p < 1$, then the absolute minimum of e in $[a_1, a_n]$ must be in $\{a_1, a_2, \ldots, a_n\}$ since $\mathsf{e}'' < 0$ on each $]a_r, a_{r+1}[$.

In the interesting case $p = 2$, system (6) has as the unique solution $((a_1^k + a_2^k + \ldots + a_n^k)/n)^{1/k}$. In the case $p = 3$, system (6) is equivalent to the biquadratic system:

$$\left\{ \begin{array}{c} (2r - n)x^{2k} - 2x^k \left(\displaystyle\sum_{i=1}^{r} a_i^k - \sum_{i=r+1}^{n} a_i^k \right) \\ + \left(\displaystyle\sum_{i=1}^{r} a_i^{2k} - \sum_{i=r+1}^{n} a_i^{2k} \right) = 0 \\ \text{in the interval }]a_r, a_{r+1}] \end{array} \right\}_{r=1}^{r=n-1}.$$

Example 15: Academic institutions use student satisfaction data to better understand, change, and improve campus

TABLE XIII
MARKS (X) AND FUZZY SATISFACTION DATA ($\mathcal{Y}$) OBSERVED ON EACH STUDENT (LEFT) AND REGRESSION MODELS FOR CENTER (RIGHT) IN EXAMPLE 15

X	Y_1	Y_2	Y_3	Y_4	Model	Function	R-Squared
8.2	60	70	75	82	Y-square	$\sqrt{862.177x + 329.635}$	91.2962%
2.5	50	60	64	75	Square root of X	$27.6797\sqrt{x} + 6.21902$	91.1915%
2.2	37	43	50	60	Linear	$6.63892x + 32.9298$	90.7341%
5.0	57	63	65	75	Double Square root	$(1.76654\sqrt{x} + 4.24198)^2$	90.6618%
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	Multiplicative	$33.2329x^{0.44552}$	89.5718%
6	70	75	75	80	$\vdots$	$\vdots$	$\vdots$

TABLE XIV
TABLE OF ERRORS CONSIDERING DIFFERENT FUZZY ERROR MEASURES TO ESTIMATE THE SPREADS

	$p = 2, k = 1$	$p = 1, k = 3$	$p = 0.5, k = 2.5$
$\widehat{\mathcal{D}_c^* \mathcal{Z}}$	Tra $(0, 1.7453, 7.6415, 7.7358)$	Tra $(0, 1.5, 7, 8)$	Tra $(0, 1.5, 5, 7)$
$\sqrt{862.2x + 329.6}$	Tra $(1251.4, 59.3, 604.2, 472.3)$	Tra $(2.64 \cdot 10^6, 579.3, 31434, 25910)$	Tra $(3454.8, 91.15, 552, 554.5)$
$27.68\sqrt{x} + 6.22$	Tra $(1258.1, 59.3, 604.2, 472.3)$	Tra $(2.72 \cdot 10^6, 579.3, 31434, 25910)$	Tra $(3524.8, 91.15, 552, 554.5)$
$6.64x + 32.93$	Tra $(1323.4, 59.3, 604.2, 472.3)$	Tra $(2.65 \cdot 10^6, 579.3, 31434, 25910)$	Tra $(3454.6, 91.15, 552, 554.5)$
$(1.77\sqrt{x} + 4.24)^2$	Tra $(1234.9, 59.3, 604.2, 472.3)$	Tra $(2.57 \cdot 10^6, 579.3, 31434, 25910)$	Tra $(3409.1, 91.15, 552, 554.5)$
$33.23x^{0.446}$	Tra $(1280.1, 59.3, 604.2, 472.3)$	Tra $(2.81 \cdot 10^6, 579.3, 31434, 25910)$	Tra $(3570.7, 91.15, 552, 554.5)$

environments, thereby creating settings more conducive for student development. In this sense, student satisfaction is an indicator of the institution's responsiveness to students' needs and a measure of institutional effectiveness, success, and vitality. Most questionnaires that collect opinions and judgments of students are designed using a five-point Likert scale or a integer scale. However, since opinion/valuation assessments are intrinsically imprecise, several studies suggest that it is reasonable and feasible to use a scale of FNs to avoid the loss of information, and therefore, it is more convenient to use fuzzy regression methods instead of traditional regression methods. Thus, students have been requested to reply by using FNs on $[0, 100]$.

The research was carried out in the second term of course 2011/2012 at the Higher Technical School of Computer Sciences and Telecommunications of the University of Granada (Southern Spain). Student satisfaction in Statistics, a subject in the first course of the Computer Engineering degree, is an important issue that we are interested in explaining through the student's College knowledge in this subject. This knowledge was measured by a mark (a real number on $[0, 10]$) obtained in a test at the beginning of the course. Each test was associated with a number (not with the name of the student) used by the students to provide the corresponding satisfaction at the end of the course. Therefore, their individual opinions were private and confidential. To collect students' satisfaction data, the lecturer explained the purpose of the study at the beginning of the course and gave the instructions to measure the satisfaction in Statistics using a TFN. A week was given at the end of the course to collect the answers of students' satisfaction. The number of registered students in "Statistics" who have participated in the research was 54. 11.11% of the students were females, whereas 88.88% were males.

Applying the classical regression analysis, we can find a list of models for the center that can be considered significant. However, in this example, no model can be considered significant for the spreads, and consequently, they have been estimated using constants. Table XIV shows the fuzzy errors obtained varying the different models fitted for the center in Table XIII . To estimate the spreads, we have considered three cases: $p = 2$ and $k = 1$ in the second column, $p = 1$ and $k = 3$ in the third column, and $p = 0.5$ and $k = 2.5$ in the last column of Table XIV . In each case, we have calculated the best value for $\widehat{\mathcal{D}_c^* \mathcal{Z}}$, that is, $\widehat{M}$, $\widehat{I}$, and $\widehat{S}$. We conclude that the best fit for the center in all cases is $\widehat{\mathcal{D}_c \mathcal{Z}_x} = (1.76654\sqrt{x} + 4.24198)^2$. The spreads do not depend on the explanatory variables and can be estimated using different constants depending on the fuzzy error measure chosen by the researcher. Table XIV shows the results and highlights in gray the best fit obtained following the methodology described in this paper. Fig. 6 plots the fuzzy regression model obtained for $p = 2$ and $k = 1$ described by $\widehat{\mathcal{Z}}_x =$ Tra $(\widehat{Y}_x, \widehat{M}, \widehat{I}, \widehat{S})$, where

$$\widehat{Y}_x = (1.76654\sqrt{x} + 4.24198)^2, \quad \widehat{M} = 1.745283,$$
$$\widehat{I} = 7.641509, \quad \text{and} \quad \widehat{S} = 7.735849. \tag{7}$$

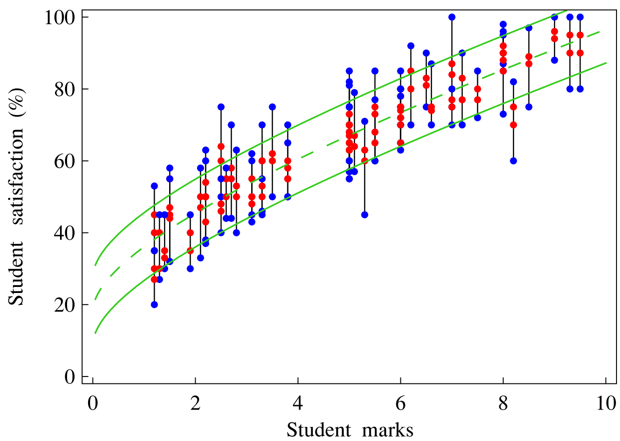

Fig. 6. Observed satisfaction data by the mark of the student and the estimated fuzzy regression model (7).

C. Extension of the Previous Methodology

Our fuzzy regression model can be generalized to the case in which the explanatory variables are also fuzzy, that is, the case of $\underline{\mathcal{X}}, \mathcal{Y}$ being two variables where $\underline{\mathcal{X}} = (\mathcal{X}_1, ..., \mathcal{X}_N)'$ is a trapezoidal FRV and $\mathcal{Y}$ is a trapezoidal FRV. Based on a random sample $\{\mathcal{X}_{1i}, ..., \mathcal{X}_{Ni}, \mathcal{Y}_i\}_{i=1}^n$, if we are interested in analyzing the relationship between $\mathcal{Y}$ and $\mathcal{X}_1, ..., \mathcal{X}_N$, then this problem can be reduced to the previous one by considering as input variable the crisp RV $\underline{X}' = (X_1^1, X_1^2, X_1^3, X_1^4, X_2^1, ..., X_{N-1}^4, X_N^1, X_N^2, X_N^3, X_N^4)'$. Indeed, if $\mathcal{X}_1, ..., \mathcal{X}_N$ are not trapezoidal, we can also consider an approximation operator as we have described above.

Example 16: Wu [40] proposed an approach for constructing a fuzzy regression model. Later, Chen and Hsueh [7] considered the same fuzzy data to illustrate their methodology and compared their results with the results obtained by Wu's approach; see [7, Table III]. In this example, a triangular FRV $\mathcal{Y}$ is studied depending on two triangular FRV $\mathcal{X}_1$ and $\mathcal{X}_2$. We are going to study $\mathcal{Y}$ depending on the six RVs X_1^ℓ, X_1^c, X_1^s, X_2^ℓ, X_2^c and X_2^s, i.e., depending on $\underline{X} = (X_1^\ell, X_1^c, X_1^s, X_2^\ell, X_2^c, X_2^s)$ (note that $X_1^m = X_2^m = 0$).

Following the methodology that we have explained above, the fuzzy regression model fitted is $\widehat{\mathcal{Z}}_{\underline{X}} = \text{Tri}\,(\hat{Y}_{\underline{X}}^c, \hat{I}_{\underline{X}}, \hat{S}_{\underline{X}})$, where

$$\begin{aligned}\hat{Y}_{\underline{X}}^c &= 2.5246 + 0.0068X_1^\ell + 0.5269X_1^c - 0.0248X_1^s \\ &\quad + 0.0006X_2^\ell + 0.0043X_2^c + 0.0034X_2^s \qquad (8)\end{aligned}$$

$$\begin{aligned}\hat{I}_{\underline{X}} &= 15.3986 + 0.0411X_1^\ell + 0.1892X_1^c - 0.03489X_1^s \\ &\quad - 0.0041X_2^\ell + 0.021X_2^c - 0.016X_2^s \qquad (9)\end{aligned}$$

$$\begin{aligned}\hat{S}_{\underline{X}} &= 4.5245 + 0.5813X_1^\ell - 0.4936X_1^c + 0.2551X_1^s \\ &\quad - 0.0092X_2^\ell + 0.0154X_2^c - 0.0080X_2^s \qquad (10)\end{aligned}$$

whose adjusted R-squared are $99.8234\,\%$, $91.9099\,\%$, and $88.9876\,\%$, respectively. Table XV shows that the total estimation error (SSE) of our proposed approach is the lowest: 1784.483. Finally, remark that the R-squared value of the model we have obtained using the proposed methodology is $\widetilde{R}^2 = 0.9909$.

TABLE XV
COMPARISON OF THE ESTIMATION ERRORS FROM VARIOUS MODELS USING SSE

Authors	Fitted fuzzy regression model	SSE
Wu	$3.453 + 0.496\mathcal{X}_1 + 0.009\mathcal{X}_2$	3674.819
Chen-Hsueh	$0.507\mathcal{X}_1 + 0.009\mathcal{X}_2 \ddot{-}$ $(-18.167/0.060/10.592)$	2217.535
Proposed methodology	$\widehat{\mathcal{Z}}_{\underline{X}} = \text{Tri}\,(\hat{Y}_{\underline{X}}^c, \hat{I}_{\underline{X}}, \hat{S}_{\underline{X}})$ (see (8)-(10))	1784.483

TABLE XVI
DATASET INTRODUCED IN [21]

	Inside control room experience (X_1)	Outside control room experience (X_2)	Education (X_3)	Response Time ($\mathcal{Y}$)
Team 1	2.0	0.0	15.25	Tri (5.83, 3.56)
Team 2	0.0	5.0	14.13	Tri (0.85, 0.52)
Team 3	1.13	1.5	14.13	Tri (13.93, 8.5)
Team 4	2.0	1.25	13.63	Tri (4, 2.44)
Team 5	2.19	3.75	14.75	Tri (1.65, 1.01)
Team 6	0.25	3.5	13.75	Tri (1.58, 0.96)
Team 7	0.75	5.25	15.25	Tri (8.18, 4.99)
Team 8	4.25	2.0	13.5	Tri (1.85, 1.13)

Example 17: In some cases, our methodology can produce extraordinary results. For instance, consider the following experimental case, which was introduced by Kim and Bishu in [21] and was later intensively studied, among others, in [7], [8], [18], [24], and [34]. In this example, we consider the cognitive response times of the nuclear power plant room crew to an abnormal event (more details can be found in [21]). A fuzzy (usually, linear) model is assumed between the cognitive response time and the crews experience inside a control room (in years) and experience outside a control room (in years) and education (in years). Table XVI lists the original data.

The observations of independent variables are crisp ($\underline{X} = (X_1, X_2, X_3)$), but the observations of the dependent variable ($\mathcal{Y}$) are presented as symmetric triangular FNs. Applying our methodology, we obtain the fuzzy regression model $\widehat{\mathcal{Z}}_{\underline{X}} = \text{Tri}\,(\hat{Y}_{\underline{X}}^c, \hat{S}_{\underline{X}})$, where

$$\begin{aligned}\hat{Y}_{\underline{X}}^c = {} & \exp(-788.318 - 9.20165\,X_1 - 3.9764\,X_2 \\ & + 111.28\,X_3 + 0.889021\,X_1^2 + 0.746288\,X_2^2 \\ & - 3.92647\,X_3^2 + 14.2274\sqrt{X_1})\end{aligned}$$

$$\begin{aligned}\hat{S}_{\underline{X}} = {} & \exp(-791.587 - 9.18783X_1 - 3.98627X_2 \\ & + 111.799X_3 + 0.887756X_1^2 + 0.748419\,X_2^2 \\ & - 3.94458X_3^2 + 14.2211\sqrt{X_1}).\end{aligned}$$

This model is not only good, but is the best possible since the data set is perfectly fitted by this model ($\widetilde{R}^2 = 100\%$). Therefore, it is not necessary to compare our model with obtained in [7], [8], [18], [21], [24], and [34].

V. Conclusion and Final Notes

In this paper, based on the idea that real techniques with fuzzy data may produce loss of information, we have developed fuzzy tools that extend well-known and widely used real techniques. In the first part of this paper, we have introduced a family of new fuzzy distance measures between FNs (not necessarily trapezoidals), depending on a wide range of real and fuzzy concepts, and we have studied its metric properties. Precisely, these properties justify its importance. Then, we have explored the problem of analyzing linear and nonlinear fuzzy regression relationships using fuzzy distances. The use of the fuzzy measures introduced has allowed us to be consistent with the type of data considered. The imprecision of the response variable was described by trapezoidal-valued FNs because simple shapes of FNs are easier to fix and handle and most researchers prefer to use them in real-life examples. We did not pay attention to nonnegativity conditions because if the fitted model is not valid in a value or set of values, we can always choose one that is also significant or apply some transformation that leads to a significant model, and in any case, it is always possible to consider a constant as a function for the spread. Results of the numerical examples show that the result of the proposed formulation is to obtain fuzzy models easily, giving results that are comparable (or improved) with those obtained with more complex techniques proposed by other researchers. Further investigation will be necessary in order to find out new properties of this family of fuzzy measures and how to apply it to real contexts.

Acknowledgment

The authors are very thankful to the anonymous reviewers for their careful reading of the manuscript and for their constructive reports, which have been very useful to improve the paper. The first author is grateful to the Department of Quantitative Methods for Economics and Business, University of Granada.

References

[1] M. A. Firozja, G. H. Fath-Tabar, and Z. Eslampia, "The similarity measure of generalized fuzzy numbers based in interval distance," *Appl. Math. Lett.*, vol. 25, pp. 1528–1534, 2012.

[2] T. Allahviranloo and M. A. Firozja, "Ranking of fuzzy numbers by a new metric," *Soft Comput.*, vol. 14, pp. 773–783, 2010.

[3] T. Allahviranloo, R. Nuraei, M. Ghanbari, E. Haghi, and A. A. Hosseinzadeh, "A new metric for L-R fuzzy numbers and its application in fuzzy linear systems," *Soft Comput.*, vol. 16, no. 10, pp. 1743–1754, 2012.

[4] A. I. Ban and L. Coroianu, "Nearest interval, triangular and trapezoidal approximation of a fuzzy number preserving ambiguity," *Int. J. Approx. Reason.*, vol. 53, pp. 805–836, 2012.

[5] J. J. Buckley and L. J. Jowers. *Monte Carlo Methods in Fuzzy Optimization* (Studies in Fuzziness and Soft Computing), vol. 222. New York, NY, USA: Springer, 2008.

[6] C. H. Cheng, "A new approach for ranking fuzzy numbers by distance method," *Fuzzy Sets Syst.*, vol. 95, pp. 307–317, 1988.

[7] L. H. Chen and C.-C. Hsueh, "Fuzzy regression models using the least-squares method based on the concept of distance," *IEEE Trans. Fuzzy Syst.*, vol. 17, no. 6, pp. 1259–1272, Dec. 2009.

[8] S. H. Choi and J. J. Buckley, "Fuzzy regression using least absolute deviation estimators," *Soft Comput.*, vol. 12, pp. 257–263, 2008.

[9] R. Coppi, P. D'Urso, P. Giordani, and A. Santoro, "Least squares estimation of a linear regression model with LR fuzzy response," *Comput. Statist. Data Anal.*, vol. 51, pp. 267–286, 2006.

[10] P. Diamond, "Fuzzy least squares," *Inf. Sci*, vol. 46, pp. 141–157, 1988.

[11] D. Dubois and H. Prade, "Operations on fuzzy numbers," *Int. J. Syst. Sci.*, vol. 9, pp. 613–626, 1978.

[12] D. Dubois and H. Prade.,*Possibility Theory*. New York, NY, USA: Plenum Press, 1988.

[13] D. Dubois and H. Prade, "Fuzzy elements in a fuzzy set," in *Proc. 10th Int. Fuzzy Syst. Assoc. Congr.*, Beijing, China, 2005, pp. 55–60.

[14] M. B. Ferraro, R. Coppi, G. González-Rodríguez, and A. Colubi, "A linear regression model for imprecise response," *Int. J. Approx. Reason.*, vol. 51, pp. 759–770, 2010.

[15] R. Goetschel and W. Voxman, "Elementary fuzzy calculus," *Fuzzy Sets Syst.*, vol. 18, pp. 31–43, 1986.

[16] P. Grzegorzewski and E. Mrówka, "Trapezoidal approximations of fuzzy numbers," *Fuzzy Sets Syst.*, vol. 153, pp. 115–135, 2005.

[17] D. Guha and D. Chakraborty, "A new approach to fuzzy distance measure and similarity measure between two generalized fuzzy numbers," *Appl. Soft Comput.*, vol. 10, pp. 90–99, 2010.

[18] H. Hassanpour, H. R. Maleki, and M. A. Yaghoobi, "A goal programming approach to fuzzy linear regression with non-fuzzy input and fuzzy output data," *Asia-Pacific J. Oper. Res.*, vol. 26, pp. 587–604, 2009.

[19] M. Hojati, C. R. Bector, and K. Smimou, "A simple method for computation of fuzzy linear regression," *Eur. J. Oper. Res.*, vol. 166, pp. 172–184, 2005.

[20] C. Kao and C. L. Chyu, "Least-squares estimates in fuzzy regression analysis," *Eur. J. Oper. Res.*, vol. 148, pp. 426–435, 2003.

[21] B. Kim and R. R. Bishu, "Evaluation of fuzzy linear regression models by comparing membership functions," *Fuzzy Sets Syst.*, vol. 100, pp. 343–352, 1998.

[22] G. J. Klir and B. Yuan.,*Fuzzy Sets and Fuzzy Logic*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1995.

[23] M. Mizumoto and J. Tanaka, "Some properties of fuzzy numbers," in *Advances in Fuzzy Set Theory and Applications*, M. M. Gupta *et al.*, Eds. New York, NY, USA: North Holland, 1979, pp. 153–164.

[24] M. Modarres, E. Nasrabadi, and M. M. Nasrabadi, "Fuzzy linear regression models with least square errors," *Appl. Math. Comput.*, vol. 163, pp. 977–989, 2005.

[25] M. M. Nasrabadi and E. Nasrabadi, "A mathematical-programming approach to fuzzy linear regression analysis," *Appl. Math. Comput.*, vol. 155, pp. 873–881, 2004.

[26] W. Näther, "Regression with fuzzy random data," *Comput. Statist. Data Anal.*, vol. 51, pp. 235–252, 2006.

[27] M. L. Puri and D. A. Ralescu, "Fuzzy random variables," *J. Math. Anal. Appl.*, vol. 114, pp. 409–422, 1986.

[28] A. Roldán, J. Martínez-Moreno, and C. Roldán, "On interrelationships between fuzzy metric structures," *Iran. J. Fuzzy Syst.*, vol. 10, no. 2, pp. 133–150, 2013.

[29] A. Roldán, J. Martínez-Moreno, and C. Roldán, "Some applications of the study of the image of a fuzzy number: Countable fuzzy numbers, operations, regression and a specificity-type ordering," *Fuzzy Sets Syst.*, vol. 257, pp. 204–216, 2014.

[30] C. Roldán, A. Roldán, and J. Martínez-Moreno, "A fuzzy regression model based on distances and random variables with crisp input and fuzzy output data: A case study in biomass production," *Soft Comput.*, vol. 16, pp. 785–795, 2012.

[31] B. Schweizer and A. Sklar.,*Probabilistic Metric Spaces*. New York, NY, USA: Dover Publications, 2005.

[32] B. Sinova, M. A. Gil, A. Colubi, and S. Van Aelst, "The median of a random fuzzy number. The 1-norm distance approach," *Fuzzy Sets Syst.*, vol. 200, pp. 99–115, 2012.

[33] L. Stefanini, "A generalization of Hukuhara difference and division for interval and fuzzy arithmetic," *Fuzzy Sets Syst.*, vol. 161, pp. 1564–1584, 2010.

[34] S. M. Taheri and M. Kelkinnama, "Fuzzy linear regression based on least absolutes deviations," *Iran. J. Fuzzy Syst.*, vol. 9, no. 1, pp. 121–140, 2012.

[35] H. Tanaka, I. Hayashi, and J. Watada, "Linear regression analysis with fuzzy model," *IEEE Trans. Syst. Man. Cybern.*, vol. SMC-12, no. 6, pp. 903–907, Nov. 1982.

[36] H. Tanaka, I. Hayashi, and J. Watada, "Possibilistic linear regression analysis for fuzzy data," *Eur. J. Oper. Res.*, vol. 40, pp. 389–396, 1989.

[37] L. Tran and L. Duckstein, "Comparison of fuzzy numbers using a fuzzy distance measure," *Fuzzy Sets Syst.*, vol. 130, pp. 331–341, 2002.

[38] J. V. Riera and J. Torrens, “Aggregation of subjective evaluations based on discrete fuzzy numbers,” *Fuzzy Sets Syst.*, vol. 191, pp. 21–40, 2012.
[39] W. Voxman, “Canonical representations of discrete fuzzy numbers,” *Fuzzy Sets Syst.*, vol. 118, pp. 457–466, 2001.
[40] H. C. Wu, “Linear regression analysis for fuzzy input and output data using the extension principle,” *Comput. Math. Appl.*, vol. 45, pp. 1849–1859, 2003.
[41] C. X. Wu and M. Ma.,*The Basic of Fuzzy Analysis*. Beijing, China: Nat. Defence Ind. Press, 1991.
[42] B. Wu and N. F. Tseng, “A new approach to fuzzy regression models with application to business cycle analysis,” *Fuzzy Sets Syst.*, vol. 130, pp. 33–42, 2002.
[43] K.-L. Wu and M.-S. Yang, “Alternative c-means clustering algorithms,” *Pattern Recog.*, vol. 35, pp. 2267–2278, 2002.

Antonio Francisco Roldán López de Hierro received the M.Sc. and Ph.D. degrees in mathematics from the University of Granada, Granada, Spain, in 1998 and 2003, respectively, and the Ph.D. degree in statistics from the University of Jaén, Jaén, Spain, in 2012.

He is currently a Teacher (government employee) with a secondary school and as a Lecturer with the Department of Quantitative Methods for Economics and Business, University of Granada. He is a coauthor of approximately 45 original research articles. His research interests include fuzzy sets theory, fuzzy regression, fuzzy ranking, decision making, and fixed point theory.

Juan Martínez-Moreno received the Ph.D. degree in mathematics from the University of Granada, Granada, Spain, in 2001.

He is currently a Member with the Department of Mathematics, University of Jaén, Jaén, Spain. His main research interests include fuzzy mathematics and fixed point theory.

Concepción Aguilar-Peña was born in Jaén, Spain, in 1975. She received the M.S. degree in mathematics with a speciality in statistics from the University of Granada, Granada, Spain, in 1981. She is currently working toward the Ph.D. degree in statistics with the University of Jaén, Jaén, Spain.

She started working with the Ministry of Economy and Finance as a Civil Servant. She is also a Lecturer with the University of Jaén. Her research interests include fuzzy theory and regression analysis.

Concepción Roldán López de Hierro received the B.S. and M.S. degrees in mathematics with a speciality in statistics from the University of Granada, Granada, Spain, in 1996, and the Ph.D. degree in statistics from the University of Granada in 2002.

She is currently an Associate Professor with the Department of Statistics and Operations Research, University of Granada. Her research interests include regression analysis, statistical modeling, applied statistics, computational statistics, fuzzy set theory, and its applications in the previous areas.