Journal of Uncertain Systems
Vol.4, No.1, pp.59-72, 2010
Online at: www.jus.org.uk

Linear Fuzzy Regression Using Trapezoidal Fuzzy Intervals

Amory Bisserier, Reda Boukezzoula*, Sylvie Galichet
Laboratoire d'Informatique, Systèmes, Traitement de l'Information et de la Connaissance - LISTIC. Université de Savoie
BP. 80439 – 74941 Annecy-le-vieux, Cedex France

Received 2 November 2008; Revised 25 April 2009

Abstract

Conventional Fuzzy regression using possibilistic concepts allows the identification of models from uncertain data sets. However, some limitations still exist. This paper deals with a revisited approach for possibilistic fuzzy regression methods. Indeed, a new modified fuzzy linear model form is introduced where the identified model output can envelop all the observed data and ensure a total inclusion property. Moreover, this model output can have any kind of spread tendency. In this framework, the identification problem is reformulated according to a new criterion that assesses the model fuzziness independently of the collected data. The potential of the proposed method with regard to the conventional approach is illustrated by simulation examples.

Keywords: fuzzy linear regression, fuzzy intervals, total inclusion, model identification

1 Introduction

Classically, in regression models, deviations between the observed and the estimated values are supposed to be due to measurement errors and/or random variations. In this case, statistical techniques are well suited for the model determination. However, in many real applications, the deviations are due to the imprecise observed data or the indefiniteness of the system structure. In this case, the uncertainty is not due to randomness but fuzziness. So, regression analysis on fuzzy data for dealing with fuzziness is usually called fuzzy regression analysis.

Fuzzy regression, a fuzzy type of classical regression analysis, has been proposed to evaluate the functional relationship between dependent and independent variables in a fuzzy environment. Indeed, unlike statistical regression modeling that is based on probability theory, fuzzy regression is based on possibility theory and fuzzy set theory [23, 24]. In the fuzzy literature, the regression problem with fuzzy data has been previously treated from different points of view and by considering different kinds of input/output data. In this context, the complete specification of regression problems highly depends on the nature of input-output data [8]. Some works are thus devoted to Crisp-Input Crisp-Output (CICO) data [16] while others [9, 15, 21] consider the regression problem using Fuzzy-Inputs and Fuzzy-Outputs (FIFO) data. Most commonly, a mixed approach Crisp-Inputs and Fuzzy-Outputs (CIFO) is chosen [6, 8, 10-12, 14, 17, 18]. In this framework, the fuzzy regression with CIFO data aroused a major interest. However, the FIFO model regression remains a little studied field. Indeed, only a few papers concerning this problem have been published. In this context, it is even more difficult to design a robust fuzzy regression methodology. First step towards the identification of such systems consists in studying and implementing a fuzzy regression for CIFO systems. So, in this paper, the CIFO regression problem is considered.

In the context of regression model identification, two main problems come up, namely, the model structure specification and the estimation of the given model. The first problem is focused on the choice of a suitable model structure for a data set. This problem is traditionally addressed *a priori*, in order to take advantage of the estimators found for the second one. In this paper, as commonly used, the model structure is assumed to be linear. In this case, the fuzzy regression problem is reduced to an estimation problem of the model parameters.

According to Diamond [6-8, 17, 18], fuzzy regression techniques can be classified into two distinct areas. The first proposed by Tanaka, which minimizes the total spread of the output, is named possibilistic regression. This method tries to minimize the whole fuzziness of the model by minimizing the total spread of its fuzzy parameters,

* Corresponding author. Email: reda.boukezzoula@univ-savoie.fr (R. Boukezzoula).

subject to including the data points of each sample within a specified feasible data interval. In other words, the regression problem is viewed as finding fuzzy parameters of a regression model according to a mathematical programming problem. The second approach developed by Diamond [6], which minimizes the total square error of the output is called the fuzzy least square method. In the present study, the possibilistic approach is employed, for which a new revisited methodology is proposed. In this case, the main objective is to revisit some theoretical works about fuzzy regression techniques and to propose some slight improvements for their limitations.

Since the fuzzy regression has been introduced by Tanaka *et al.* [17], several fuzzy regression approaches have been proposed. In this context, Tanaka *et al.* [18] propose different expressions of the criterion to be optimized and different formulations of the constraints to be satisfied for possibility and necessity estimation models. Still in a linear context, Tanaka and Ishibushi [19] extend their approach for dealing with interactive fuzzy parameters. In this paper, a mixed approach (CIFO) is adopted [17, 18] with the idea of keeping a simple model, possibly invertible (Boukezzoula *et al.* 2003, 2006). From most of these methods, three types of problems emerge:

- The assumption of symmetrical triangular fuzzy parameters is most frequently used. However, such parameters have some limitations, especially when a total inclusion of the observed data in the predicted model output must be ensured. Indeed, the identification is made at a chosen level considered as a degree of the fitting of the obtained model to the observed data. If this principle allows in simplifying the problem computation by using interval arithmetic to express the inclusion problem, after reconstruction of the parameters, the inclusion cannot be guaranteed anymore at any level α [1, 2].
- The obtained models are not able to represent any tendency of the output spread. In this case, the obtained models become more imprecise than necessary in some situations. Indeed, one weakness is the fact that the fuzziness of the model output varies in the same way than the absolute value of the inputs. It follows that it is impossible to have a decreasing (resp. increasing) spread of the model output for positive (resp. negative) inputs. This restriction is acceptable in a measurement context where it is usual to express percentage relative errors. However, when fuzziness is considered as an intrinsic characteristic of the system to be modeled, the assumption that the higher the input, the higher the fuzziness attached to the model output, is open to criticism [2].
- In conventional regression methods [10, 11, 14, 17, 18], the used criteria are only based on the available data, their minimization does not guarantee that the identified model has the least global fuzziness that could be achieved on a whole domain *D* (input domain of definition). If the identified model is to be used on the domain *D*, it may be more judicious to prefer a model with a lower global fuzziness, i.e. a less imprecise model. Actually, the global fuzziness of a model is an intrinsic characteristic of the latter that should be assessed independently of the identification data, already used to express the constraints of the linear program.

According to above-discussed points, the originality of the presented work is triple:

- A total inclusion property is ensured by assuming a trapezoidal fuzzy interval representation of the predicted output and the parameters in the regression model.
- A modified form of the fuzzy linear model is proposed so that any kind of spread tendency can be represented.
- The formulation of the linear program used to identify the model introduces a modified criterion that assesses the model fuzziness independently of the collected data.

The structure of this paper is as follows. In section 2, the concepts of intervals and fuzzy intervals are introduced. Section 3 is devoted to the conventional fuzzy linear regression. A revisited approach of the latter is detailed in section 4. Section 5 presents the identification process and its illustration by simulation examples. Finally, conclusions and perspectives are presented in Section 6.

2 Relevant Concepts and Notations

2.1 Conventional and Fuzzy Intervals

An interval *a* is defined by the set of elements lying between its lower and upper limits as:

$$a = \{x / a^- \le x \le a^+, x \in \Re\} . \quad (1)$$

Given an interval *a*, its Midpoint *M*(*a*) and its Radius *R*(*a*) are defined by:

$$M(a) = (a^- + a^+)/2 \text{ and } R(a) = (a^+ - a^-)/2 . \quad (2)$$

An interval *a* can be viewed as a particular case of a fuzzy number whose membership function μ_a takes the value 1 over the interval and 0 anywhere else (see Figure 1.a). In the same way, a fuzzy interval *A* is represented by its membership function μ_A. In order to specify the fuzzy interval shape, one has to consider two dimensions. The first one (horizontal dimension) is similar to that used in interval representation, that is the real line $\Re$. The second one (vertical dimension) is related to the handling of the membership degrees and thus restricted to the interval [0, 1] (see Figure 1.b).

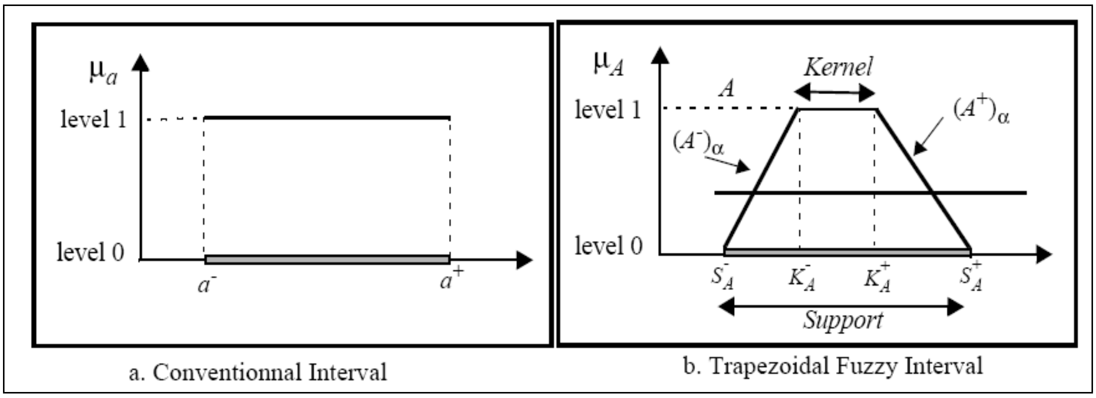


Figure 1: Conventional and fuzzy interval representation

In this context, two kinds of information are required for completely defining a fuzzy interval. Both pieces of information, called support and kernel intervals, are defined on the horizontal dimension, but are associated to two different levels (level 0 and level 1) on the vertical dimension (see Figure 1.b). To completely define the fuzzy interval, two additional functions (profiles) are used to link the support to the kernel:

$$(A^-)_\alpha = Inf\{x/\mu_A(x) \geq \alpha; x \geq S_{A^-}\}, (A^+)_\alpha = Sup\{x/\mu_A(x) \geq \alpha; x \leq S_{A^+}\}, \quad (3)$$

where $\alpha \in [0, 1]$ represents the vertical dimension.

In this case, for a given α-cut on the fuzzy interval *A*, a conventional interval $(A)_\alpha$ is obtained:

$$(A)_\alpha = [(A^-)_\alpha, (A^+)_\alpha]. \quad (4)$$

When the profile functions are assumed to be linear, the fuzzy interval *A* becomes a trapezoidal one and is fully defined by its support and its kernel intervals (see Figure 1.b):

$$\text{Support: } S_A = [S_{A^-}, S_{A^+}] \text{ and Kernel : } K_A = [K_{A^-}, K_{A^+}]. \quad (5)$$

Finally, in the same way that the conventional interval *a* is denoted $[a, a^+]$, the fuzzy trapezoidal interval *A* will be defined by its support and kernel, i.e.:

$$A = (K_A, S_A) = ([K_{A^-}, K_{A^+}], [S_{A^-}, S_{A^+}]). \quad (6)$$

It is obvious that a triangular symmetrical fuzzy interval is a particular case of a trapezoidal one. In this case, we have:

$$\begin{cases} K_{A^-} = K_{A^+} = K_A \\ S_{A^-} = K_A - R_A, \text{and } S_{A^+} = K_A + R_A. \end{cases} \quad (7)$$

2.2 Fuzzy Interval Inclusion

For two conventional intervals *a* and *b*, an inclusion relation (Boukezzoula *et al.* 2007) of *a* in *b* (see Figure 2.a) is defined as follows:

$$a \subseteq b \Leftrightarrow \begin{cases} b^- \leq a^- \\ a^+ \leq b^+ \end{cases} \Leftrightarrow \begin{cases} M(b) - R(b) \leq M(a) - R(a) \\ M(a) + R(a) \leq M(b) + R(b) \end{cases} \Leftrightarrow \begin{cases} M(b) - M(a) \leq R(b) - R(a) \\ M(a) + M(b) \leq R(b) - R(a). \end{cases} \quad (8)$$

From Eq. (8) it follows

$$a \subseteq b \Leftrightarrow |M(b) - M(a)| \leq R(b) - R(a). \quad (9)$$

In the particular case where *a* is a scalar value, the relation of Eq. (9) becomes

$$a \in b \Leftrightarrow |M(b) - a| \leq R(b). \quad (10)$$

For two fuzzy intervals *A* and *B*, an inclusion relation of *A* in *B* is defined by the following relation between their profiles:

$$A \subseteq B \Leftrightarrow \forall x : \mu_A(x) \leq \mu_B(x). \quad (11)$$

In the case of trapezoidal fuzzy intervals, Eq. (11) leads to

$$K_A \subseteq K_B \text{ and } S_A \subseteq S_B. \quad (12)$$

As the fuzzy interval profile functions are linear, it is clear that condition (12) guarantees inclusion for any level α. Moreover, when the particular case of triangular fuzzy intervals is assumed, the inclusion condition defined in Eq. (12) imposes an equality condition between the kernel values (see Figure 2(b)).

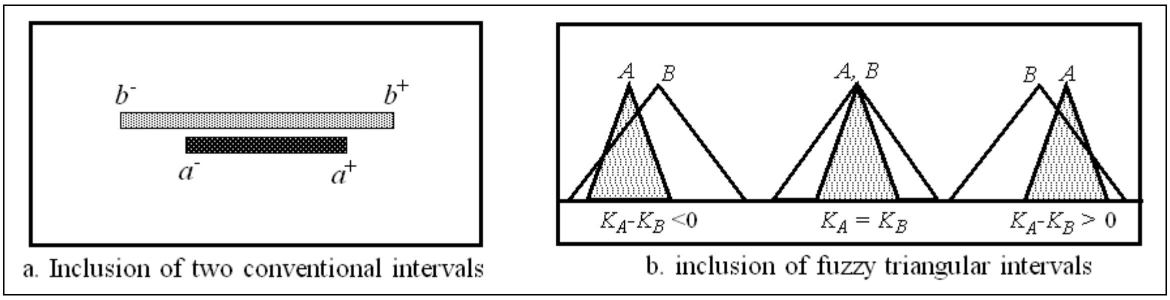


Figure 2: Conventional and fuzzy triangular interval inclusions

3 Conventional Fuzzy Linear Regression Methodology

The linear regression model is the most frequently used form in regression analysis for expressing the relationship between one or more explanatory variables and response. As commonly assumed in the fuzzy regression literature, the case of simple linear regression model involving a single independent variable is considered. The case of multiple inputs is a straightforward generalization of this methodology. The simple linear regression model can be mathematically expressed as follows:

$$Y = A_0 \oplus A_1 \otimes x \oplus E, \quad (13)$$

where Y is the output (dependant fuzzy variable), x is a non fuzzy (crisp) input, E is the fuzzy error associated with the regression model and A_0, A_1 are fuzzy intervals, the unknowns to be estimated from observed data.

Let us consider a set of M observed data samples defined on an interval D = [inf(D), sup(D)]. Let the j^{th} sample be represented by the couple (x_j, Y_j), j = 1,..., M where x_j are crisp inputs ranked in an increasing order and Y_j are the corresponding fuzzy interval outputs.

Like any regression technique, the fuzzy regression objective is to determine a predicted functional relationship:

$$\hat{Y} = A_0 \oplus A_1 \otimes x, \text{ where: } \hat{Y} \text{ is the predicted output.} \quad (14)$$

In this case, as the model structure is assumed to be fixed, using the formalism of fuzzy intervals we can express the regression problem as an estimation problem of the fuzzy interval parameters A_0 and A_1. In other words, the fuzzy regression problem is reduced to the estimation of the fuzzy parameters A_0 and A_1. As discussed in the paper introduction, a possibilistic fuzzy regression approach is adopted where the objective is to determine the fuzzy parameters such that the observed data are included in the predicted one.

In the fuzzy literature, several techniques have been proposed to deal with this problem [17, 18]. In this framework, whatever the methodology chosen, two points have to be considered for its implementation:

- The model form induced by the used parameters shape and implying constraints for the optimization problem.
- The used identification criterion in the optimization process.

3.1 Conventional Model Form

As proposed by (Tanaka *et al.* 1982, 1989), the parameters A_0, A_1 are assumed to be symmetrical triangular fuzzy intervals. In this case, as the model is linear and the inputs are considered as crisp values, the model output will be also a triangular fuzzy interval whose membership function is of the same kind as that of the model parameters.

So, when considering the j^{th} observed data, whose output is the triangular symmetrical fuzzy interval $Y_j = (K_{Y_j}, S_{Y_j})$, the corresponding predicted output is the triangular fuzzy interval given by:

$$\hat{Y}_j = (K_{\hat{Y}_j}, S_{\hat{Y}_j}) = (K_{\hat{Y}_j}, [S^-_{\hat{Y}_j}, S^+_{\hat{Y}_j}]) = (K_{\hat{Y}_j}, [K_{\hat{Y}_j} - R(\hat{Y}_j(x)), K_{\hat{Y}_j} + R(\hat{Y}_j(x))]) \quad (15)$$

Some remarks and design considerations can now be expressed concerning the used model form.

3.1.1 Model Tendency

According to the used model (see Eq. (14)), it can be stated that the model output varies in the same way that the absolute value of the input. So, it is impossible to have a decreasing (resp. increasing) spread of the model output for positive (resp. negative) inputs. To justify this assertion, the output modal value and the spread are determined. Indeed, as A_0 and A_1 are symmetrical triangular fuzzy intervals, and x a crisp input, $\hat{Y}(x)$ is also a symmetrical triangular one. In this case, the modal value $M(\hat{Y}(x))$ and the spread $R(\hat{Y}(x))$ are given by the following equation:

$$\begin{cases} M(\hat{Y}(x)) = K_{\hat{Y}_j} = M(A_0) + M(A_1).x = K_{A_0} + K_{A_1}.x \\ R(\hat{Y}(x)) = R(A_0) + R(A_1).|x|. \end{cases} \tag{16}$$

In this case, as x is varying on the domain D, the variation of Eq. (16) needs to be analyzed according to the sign of x. So, from Eq. (16), it means that the variation of $M(\hat{Y}(x))$ depends on the sign of K_{A_1} and can be increasing or decreasing for any value of the input x. In the same time, according to Eq. (16), it can be stated that the variation of $R(\hat{Y}(x))$ depends on the sign of the input. As $R(A_1)$ is always positive, it can be stated that if x is positive, the output radius will increase, whereas when x is negative, the output radius will decrease.

In this framework, it is possible to have any kind of variation of the output modal value, with an appropriate sign of K_{A_1}. However, the radius output variation is limited by the sign of the input x. As mentioned previously, classical fuzzy regression models are not able to represent any tendency of the output spread, they become more imprecise than necessary in several situations.

3.1.2 Constraint Definition

It is well-known that when the α-cut principle is used, a fuzzy interval can be viewed as a weighted family of nested conventional intervals. In other words, for a specified α-cut, a fuzzy interval becomes a conventional interval, which states that the fuzzy interval representation is a generalization of the conventional one.

This principle has been applied by Tanaka for defining the inclusion constraints in the optimization problem (the observed data are included in the predicted ones). In this case, for a set of observed data, the fuzzy model parameters A_0 and A_1 are determined so that all observed data are included in the predicted ones for a given and specified α-cut, i.e.,

$$[Y_j]_\alpha \subseteq [\hat{Y}_j]_\alpha \Leftrightarrow |M([\hat{Y}_j]_\alpha) - M([Y_j]_\alpha)| \le R([\hat{Y}_j]_\alpha) - R([Y_j]_\alpha) \cdot \tag{17}$$

After the optimization method is performed, the obtained parameters computed for a given -cut are assumed to be valid for all α∈[0, 1]. If this strategy (the α-cut representation) has the advantage to reduce the fuzzy computational complexity and makes easier its implementation, it can not ensure a total inclusion. Indeed, when the supports are included, the total fuzzy intervals inclusion is respected if and only if the modal values are equal (see Figure 2(b)). It means that the identification at α= 1 is impossible if the observed modal values are not strictly aligned (Diamond and Tanaka 1999). Moreover, the higher the α considered for identification is, the wider the support of the predicted fuzzy number is [15]. These drawbacks weaken the potential use of this method, especially in real identification problems.

3.2 The Used Identification Criterion

The first criterion proposed by (Tanaka *et al.* 1982) is the minimization of the sum of the radius of the model fuzzy parameters:

$$J = R(A_0) + R(A_1) \cdot \tag{18}$$

However, the optimization of this criterion leads to a model whose parameters are often crisp numbers, and whose output is too wide to be exploited [18].

That's why another criterion has been introduced by Tanaka *et al.* [18]. It consists in minimizing the fuzziness of the model. This characteristic is defined as the sum of the spread of all predicted intervals:

$$J = Sum = M.R(A_0) + R(A_1).\sum_{j=1}^{M} |x_j| \cdot \tag{19}$$

In the literature, several variants of this criterion [10, 11, 14] have been applied for the fuzzy linear regression problem. Although the expressions of these criteria are slightly different, their common characteristic resides in using a linear criterion expressing the observed output spread. In this case, a strong dependence of the criteria to the learning data is unavoidable.

4 Revisited Fuzzy Linear Regression: the Proposed Approach

In order to deal with the two drawbacks discussed in the previous section (the model form and the identification criterion), three evolutionary concepts are introduced into the conventional fuzzy regression model identification problems.

4.1. Proposed Model Form

In the proposed model, the identified fuzzy interval parameters A_0 and A_1 are assumed to be trapezoidal. In this case, the predicted output induced by these parameters is also trapezoidal and given by the following equation:

$$\hat{Y}_j = (K_{\hat{Y}_j}, S_{\hat{Y}_j}) = ([K^-_{\hat{Y}_j}, K^+_{\hat{Y}_j}], [S^-_{\hat{Y}_j}, S^+_{\hat{Y}_j}]) \,. \tag{20}$$

4.1.1 Model Tendency

As stated previously, the main drawback of the conventional approach is that the output model tendency is not taken into account. In order to solve this problem, a modified model expression is proposed. In this case, the model output can have any kind of spread variation for any sign of x by introducing a *shift* on the original model input. Doing so, it is possible to obtain the desired sign for the shifted input variable, and so to influence the spread variation of the output.

In this case, the fuzzy linear model defined on its domain D, becomes:

$$Y = A_0 \oplus A_1 \otimes (x - shift) = A_0 \oplus A_1 \otimes w \text{ where : } w = x - shift \,. \tag{21}$$

In the model of Eq. (21), the output spread is given by:

$$\forall x \in D : R(\hat{Y}(x)) = R(A_0) + R(A_1).|w| \,. \tag{22}$$

According to Eq. (22) and by tuning the value of *shift*, the model output can have any spread variation on D. Indeed:

1. If $x - shift \geq 0$, $\forall x \in D$ i.e., $shift \leq \inf(D)$, then the model output has an increasing spread on D.
2. If $x - shift < 0$, $\forall x \in D$ i.e., $shift > \sup(D)$, then the model output has a decreasing spread on D.

For the sake of simplicity, the value $shift = \inf(D)$ is chosen for a model whose output has an increasing radius. On the contrary, for decreasing radius, $shift = \sup(D)$ is taken (see Table 1).

Table 1: Used shifted model according to the output spread variation

Output spread variation	↗	↘
Used model	$A_0 \oplus A_1 \otimes (x - inf(D))$	$A_0 \oplus A_1 \otimes (x - sup(D))$

According to this model, the problem of more imprecision than necessary for the output is solved by considering all possible tendencies (two for a single variable model).

4.1.2 Constraint Definition

In this case, by using trapezoidal model form, it can be ensured that total inclusion of all observed inputs in the predicted ones at each level α is respected. As the fuzzy parameters are trapezoidal, the model output $\hat{Y}(x)$ is also trapezoidal.

In this situation, knowing that the observed output Y_j are symmetrical fuzzy triangular intervals and the model outputs $\hat{Y}_j$ are assumed to be trapezoidal, the problem of inclusion can be reduced to an inclusion of the kernels and the supports. Thus, in order to extend the Tanaka interval method and solving the inclusion problem, two inclusion constraints must be taken into account in the identification method:

Support inclusion $[Y_j]_{\alpha=0} \subseteq [\hat{Y}_j]_{\alpha=0}$, and Kernel inclusion

$$[Y_j]_{\alpha=1} \subseteq [\hat{Y}_j]_{\alpha=1} \,. \tag{23}$$

In this case, as a trapezoidal fuzzy interval shape is assumed, it is obvious that if Eq. (23) is respected, then the total inclusion is guaranteed for each level $\alpha \in [0, 1]$, i.e.,

$$\forall \alpha \in [0,1] : [Y_j]_\alpha \subseteq [\hat{Y}_j]_\alpha . \tag{24}$$

In this case, according Eq. (9) and (10) the inclusion constraints can be written as:

- Kernel inclusion constraints (α= 1):

$$[Y_j]_{\alpha=1} \subseteq [\hat{Y}_j]_{\alpha=1} \Leftrightarrow K_{Y_j} \in [K^-_{\hat{Y}_j}, K^+_{\hat{Y}_j}] \Leftrightarrow \left| M(K_{\hat{Y}_j}) - K_{Y_j} \right| \le R(K_{\hat{Y}_j}) . \tag{25}$$

- Support inclusion constraints (α = 0):

$$[Y_j]_{\alpha=0} \subseteq [\hat{Y}_j]_{\alpha=0} \Leftrightarrow [K_{Y_j} - R_{Y_j}, K_{Y_j} + R_{Y_j}] \subseteq [S^-_{\hat{Y}_j}, S^+_{\hat{Y}_j}] \Leftrightarrow \left| M(S_{\hat{Y}_j}) - K_{Y_j} \right| \le R(S_{\hat{Y}_j}) - R_{Y_j} . \tag{26}$$

- Constraints of inclusion of the kernel into the support:

In order to obtain a fuzzy interval, another inclusion constraint must be verified, i.e., the inclusion of the kernel into the support:

$$[K^-_{\hat{Y}_j}, K^+_{\hat{Y}_j}] \subseteq [S^-_{\hat{Y}_j}, S^+_{\hat{Y}_j}] \Leftrightarrow \left| M(K_{\hat{Y}_j}) - M(S_{\hat{Y}_j}) \right| \le R(S_{\hat{Y}_j}) - R(K_{\hat{Y}_j}) . \tag{27}$$

- Radius positivity constraints:

Obviously, the radius of the identified fuzzy intervals must be positive or null, i.e.,

$$R(K_{A_i}) \ge 0 \text{ and: } R(S_{A_i}) \ge 0 \text{, for } i = 0, 1. \tag{28}$$

To sum up, the identification method must be performed under the constraints presented in Eq. (25)-(28).

4.2 The New Used Identification Criterion

As discussed previously, this section aims at defining a criterion which is independent of the learning data. Knowing that the identified model is to be used on the whole domain D, it may be more judicious to prefer a model with a lower global fuzziness, i.e. a less imprecise model. As the vertical dimension must be taken into account, the global fuzziness of the model is in fact the volume delimited by its output on the domain D.

It can be stated that the output area represented by a trapezoidal fuzzy interval (Yager 2008) is given by the following expression:

$$area(w) = (K^+_{\hat{Y}} + S^+_{\hat{Y}})/2 - (K^-_{\hat{Y}} + S^-_{\hat{Y}})/2 \tag{29}$$

with:

$$\forall w \in D : \begin{cases} K^-_{\hat{Y}} = K^-_{A_0} + (M(K_{A_1}) - R(K_{A_1}).\Delta).w \\ K^+_{\hat{Y}} = K^+_{A_0} + (M(K_{A_1}) + R(K_{A_1}).\Delta).w \\ S^-_{\hat{Y}} = S^-_{A_0} + (M(S_{A_1}) - R(S_{A_1}).\Delta).w \\ S^+_{\hat{Y}} = S^+_{A_0} + (M(S_{A_1}) + R(S_{A_1}).\Delta).w \end{cases} \text{ where: } \Delta = sign(M(D)) . \tag{30}$$

The volume delimited by the model output on its whole domain D is given by:

$$Volume = \int_{w_{min}}^{w_{max}} area(w)dw . \tag{31}$$

By substitution of Eq. (29) and Eq. (30) in Eq. (31) it yields:

$$Volume = R(K_{A_0}) + R(S_{A_0}) + (R(K_{A_1}) + R(S_{A_1})).M(D).\Delta . \tag{32}$$

It is clear that this criterion is independent from the data. Thus, the optimization is performed not only at the learning points, but on the definition domain of the model, independently from the learning data distribution. This property guarantees a certain robustness of the proposed criterion in Eq. (32). So, identifying fuzzy models with the criterion of Eq. (32) leads to models whose fuzziness is possibly lower than usually.

5 Illustrative Examples

In this section, simulation results using the proposed regression methodology are presented. Three illustration examples are considered in order to emphasize the specific points discussed previously and to show the benefits of the proposed concepts.

The first illustration is dedicated to show the influence of the model form on the inclusion property. In this case, in order to be able to compare the results, the conventional and the proposed methods are implemented using a well-known and often used data set. The next illustration is used to show the proposed criterion potential and its robustness. The last illustration deals with the output tendency property and its advantage in fuzzy regression.

In each case, several performance indicators are introduced to express the viability and the performances of the identified model. In this paper, the most relevant indicators exploited in the fuzzy regression literature are considered.

5.1 Illustration 1

We consider the data set (Hojati *et al.* 2005, Tanaka and Lee 1998) presented in Table 2. The objective is to illustrate the different characteristics of the models presented in the previous part.

Table 2: Observed data set

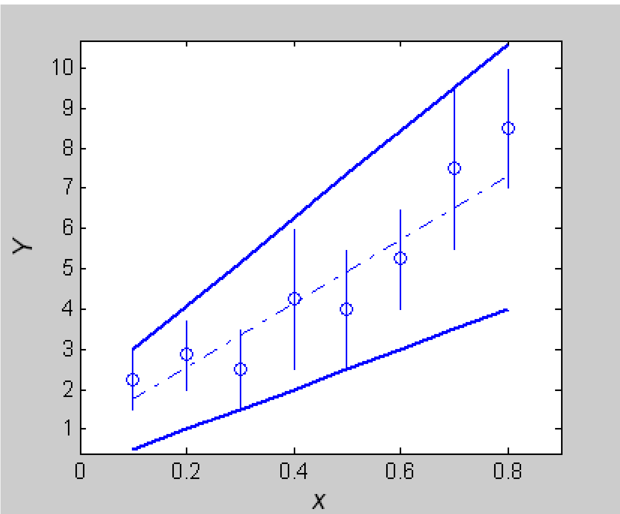
j	x_j	Y_j
1	0.1	(2.25, [1.5, 3])
2	0.2	(2.875, [2, 3.75])
3	0.3	(2.5, [1.5, 3.5])
4	0.4	(4.25, [2.5, 6])
5	0.5	(4.0, [2.5, 5.5])
6	0.6	(5.25, [4, 6.5])
7	0.7	(7.5, [5.5, 9.5])
8	0.8	(8.5, [7, 10])

First, we apply the identification method proposed by (Tanaka *et al.* 1989) in order to study the inclusion of the observed data in the predicted ones. So, the performance indicator considered here is the compatibility measure defined by the following equation:

$$Compatibility = (1/M)\sum_{j=1}^{M} \max_{x} \ \min(\mu_{\hat{Y}_j}(x), \mu_{Y_j}(x)) \,. \tag{33}$$

The identified model has the form given by Eq. (14) where the coefficients are symmetrical triangular fuzzy intervals.

The identification is performed minimizing the sum of the radius of the predicted intervals (see Eq. (19)) at = 0 under the inclusion constraints given by Eq. (17). The obtained results are illustrated in Figure 3.



a. Output Model Evolution

	Triangular Fuzzy Model
A_0	(0.96, [0, 1.92])
A_1	(7.92, [5, 10.83])
Compatibility	0.83

b. Fuzzy Model Parameters

Figure 3: Identified triangular fuzzy model

Let us consider the observed and predicted triangular fuzzy output for *j* = 1. They are respectively:

$$Y_1 = (2.25, [1.5, 3]) \text{ and } \hat{Y}_1 = (1.75, [0.5, 3]) . \quad (34)$$

As illustrated in Figure 4, it can be stated that if the inclusion constraint is respected for = 0, it is not the case for any $\alpha \in [0,1]$.

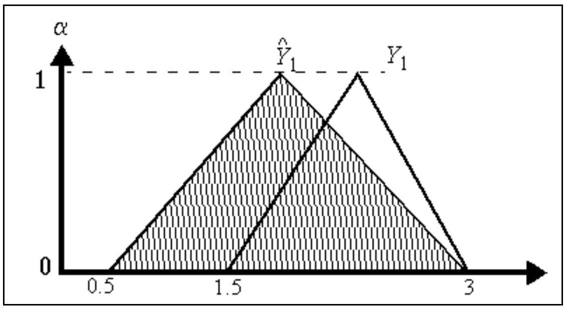


Figure 4: Observed and predicted output for *j* = 1

As proposed previously (see section 4), it is possible to improve the Tanaka's method by ensuring a total inclusion of the observed outputs in the predicted ones. This objective can be carried out by identifying a trapezoidal model (outputs and parameters are assumed to be trapezoidal). In this case, we identify trapezoidal fuzzy coefficients by minimizing, under the appropriate constraints of inclusion given by Eq. (25)-(28), the sum of the spreads given by the following equation:

$$Sum = M.\left(R(S_{A_0}) + R(K_{A_0})\right) + \left(R(S_{A_1}) + R(K_{A_1})\right)\sum_{j=1}^{M} | x_j | . \quad (35)$$

This criterion is an adaptation of the one given in Eq. (19) where the kernel and the support radius of the identified trapezoidal parameters are taken into account. The obtained results are shown in Figure 5.

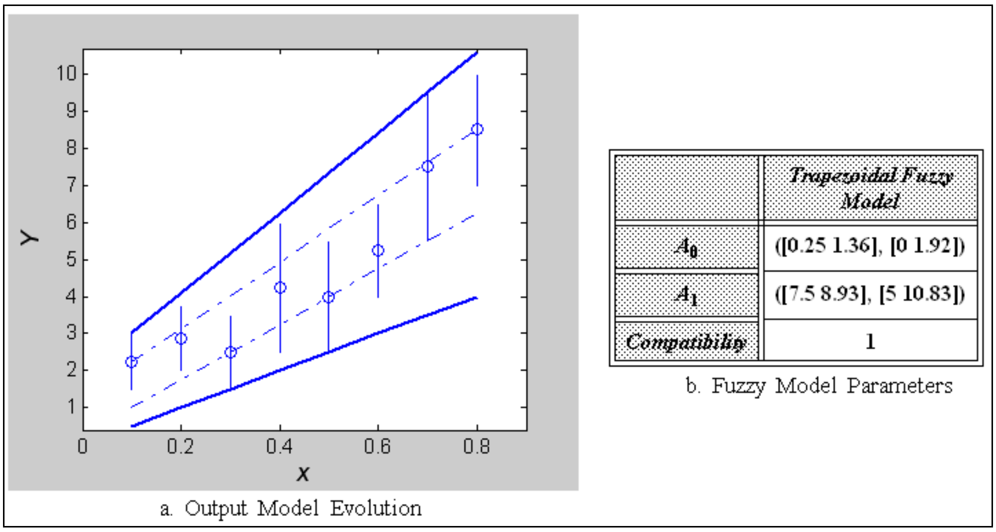


Figure 5: Identified trapezoidal fuzzy model

When considering the same observed fuzzy output used previously (for *j* = 1) and its corresponding predicted trapezoidal fuzzy output, we have:

$$Y_1 = (2.25, [1.5, 3]) \text{ and } \hat{Y}_1 = ([1, 2.25], [0.5, 3]) . \quad (36)$$

which proves that the inclusion of the observed output into the predicted one is ensured (see Figure 6).

In this case, inclusion is constrained in the identification process at the levels α = 0 and α= 1, which implies a total inclusion $\alpha \in [0,1]$. As a consequence the indicator *compatibility* has the best value, i.e. *compatibility* = 1, highlighting the fact that the inclusion of observed outputs in predicted ones is total (see Figure 7).

In the sequel, a trapezoidal model form is adopted. In other words, the model parameters are assumed to be trapezoidal.

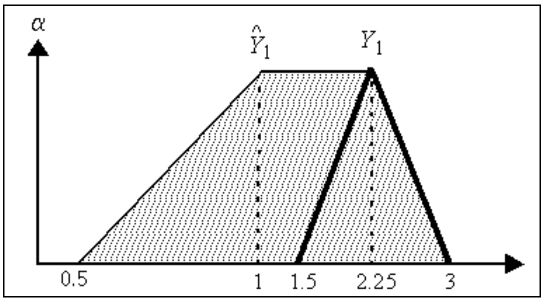


Figure 6: Observed and predicted output for j = 1

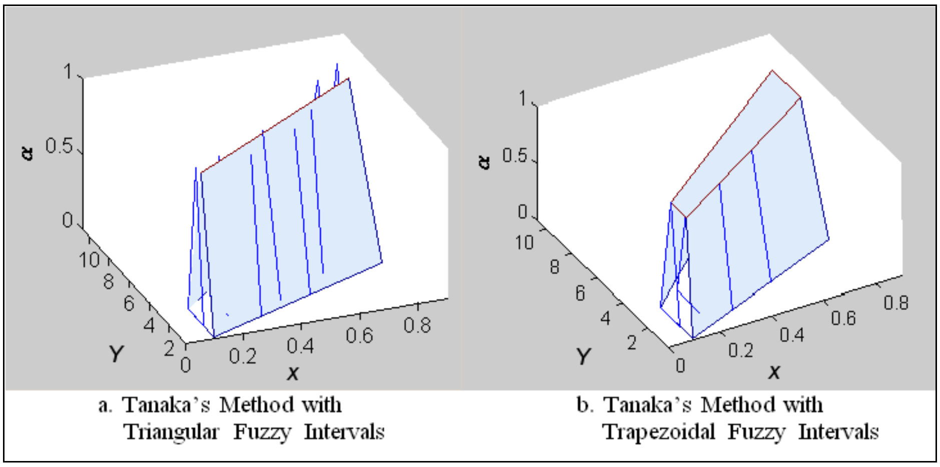


Figure 7: Identified triangular and trapezoidal fuzzy models in a 3-D view

5.2. Illustration 2

In this illustration, the robustness of the criterion used in the identification method is investigated. For this purpose, the influence of redundant learning data is considered. Indeed, the data at j = 8 in the observed output set is duplicated three times. In this case, two identification methods are applied.

- The first one is an adaptation of the conventional Tanaka's method for trapezoidal models, where the criterion is given by Eq. (35).
- The second one is the proposed strategy where the exploited criterion is the minimization of the volume, which is a criterion independent of the learning data.

The used performance indicators are:

- The Diamond Distance: for determining the fitting of the model to the data. It is defined as follows for symmetrical triangular fuzzy intervals (Eq. (37)) and trapezoidal ones (Eq. (38)).

$$Dist = \sum_{j=1}^{M} D(\hat{Y}_j - Y_j)^2 = \sum_{j=1}^{M} (S^-_{\hat{Y}_j} - S^-_{Y_j})^2 + (K_{\hat{Y}_j} - K_{Y_j})^2 + (S^+_{\hat{Y}_j} - S^+_{Y_j})^2 , \quad (37)$$

$$Dist = \sum_{j=1}^{M} D(\hat{Y}_j - Y_j)^2 = \sum_{j=1}^{M} (S^-_{\hat{Y}_j} - S^-_{Y_j})^2 + (K^-_{\hat{Y}_j} - K^-_{Y_j})^2 + (K^+_{\hat{Y}_j} - K^+_{Y_j})^2 + (S^+_{\hat{Y}_j} - S^+_{Y_j})^2 . \quad (38)$$

- The *Sum* and *Volume* criteria values: for giving indications on the fuzziness of the identified model.

The obtained results are shown in Table 3 and Figure 8.

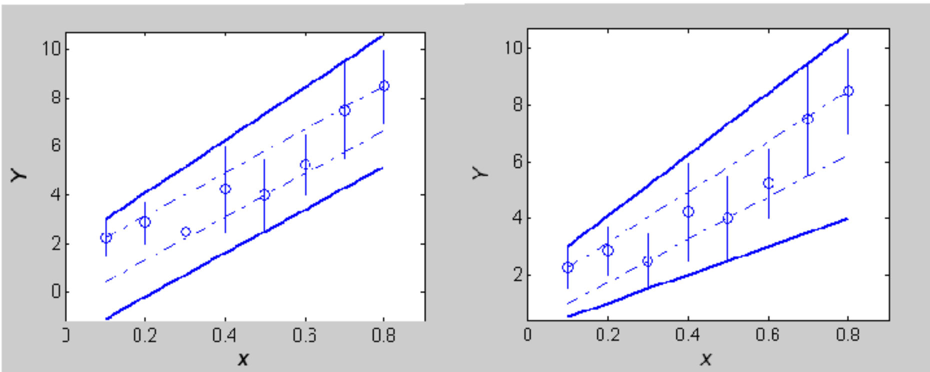
Several important points can be underlined on these results:

- The identified model on this data set with the minimization of the criterion (35) is different of the one given in Figure 5. This point highlights the lack of robustness of this criterion. Its structure is modified, as a greater importance is given to the redundant data. In this case, it is impossible to identify the initial model.
- The identified model with the proposed criterion *Volume*, independent of the learning data, remains the same as the one given in illustration 1 (see Figure 5), which is the optimal one in the sense of a lower global fuzziness. So, this criterion is more robust than the conventional one.
- The identified model with the proposed criterion *Volume*, independent of the learning data, remains the same as the one given in illustration 1 (see Figure 5), which is the optimal one in the sense of a lower global fuzziness. So, this criterion is more robust than the conventional one.

- Concerning the indicators *Distance* and *Sum*, the proposed model seems to be less efficient. However, these two indicators, less suited to our methodology, are defined on the whole learning data set. So, a greater importance is given to the redundant data, introducing a bias in the indicators values. If we consider the same indicators, but computed with the initial data set (i.e. only one representation of the data j = 8 is considered), it is obvious that the model identified with the criterion *Volume* is the best one for each value (see Table 4). In fact, the identified model presents a better fitting to the data and a lower sum of spreads of the outputs for each input x considered.

Table 3: Identified parameters and corresponding performance indicators

	Conventional Method with Sum Criterion	*Proposed Method with Volume Criterion*
A_0	([-0.46 1.36], [-1.96 1.92])	([0.25 1.36], [0 1.92])
A_1	([8.93 8.93], [8.93 10.83])	([7.5 8.93], [5 10.83])
Distance	71.05	91.29
Sum	37.08	38.42
Volume	3.28	3.15



a. Output Evolution Using the Conventional Method

b. Output Evolution Using the Proposed Method

Figure 8: Identified fuzzy models for conventional and proposed methods

Table 4: Performance indicators computed on the initial data set (no redundant data)

	Conventional Method with Sum Criterion	*Proposed Method with Volume Criterion*
Distance	50.13	48.08
Sum	26.24	25.17

The fuzzy trapezoidal model form and the volume criterion given by Eq. (32) are adopted in the illustration 3.

5.3 Illustration 3

This section aims at illustrating the tendency problem discussed previously. In this case, a comparison between conventional models and shifted proposed ones is achieved. In the previous examples, it can be stated that the observed data outputs have globally increasing radius. As the outputs are corresponding to positive inputs, conventional models can represent this tendency without any problem. This is no more the case for negative examples. In order to illustrate this point a new training set is built from the initial data by applying a constant translation (-0.9) to all input data without modifying the corresponding output (see Table 7).

In this case, two trapezoidal models (conventional and shifted one) are optimized according to the criterion *Volume* on the domain D = [-0.8, -0.1].

As data present an increasing spread tendency on the domain *D*, the value of *shift* is set to *inf*(*D*), i.e., *shift* = -0.8. The results are illustrated in Table 6 and Figure 9.

Table 5: Translated data set

j	x_j	Y_j
1	-0.8	(2.25, [1.5, 3])
2	-0.7	(2.875, [2, 3.75])
3	-0.6	(2.5, [1.5, 3.5])
4	-0.5	(4.25, [2.5, 6])
5	-0.4	(4.0, [2.5, 5.5])
6	-0.3	(5.25, [4, 6.5])
7	-0.2	(7.5, [5.5, 9.5])
8	-0.1	(8.5, [7, 10])

Table 6: Identified parameters and corresponding performance indicators

	Conventional Model	*Shifted Model*
A_0	([7.57 9.39],[0.07 11.29])	([1 2.25],[0.5 3])
A_1	8.93	([7.5 8.93],[5 10.83])
Distance	58.83	48.08
Sum	28.14	25.17
Volume	3.52	3.15

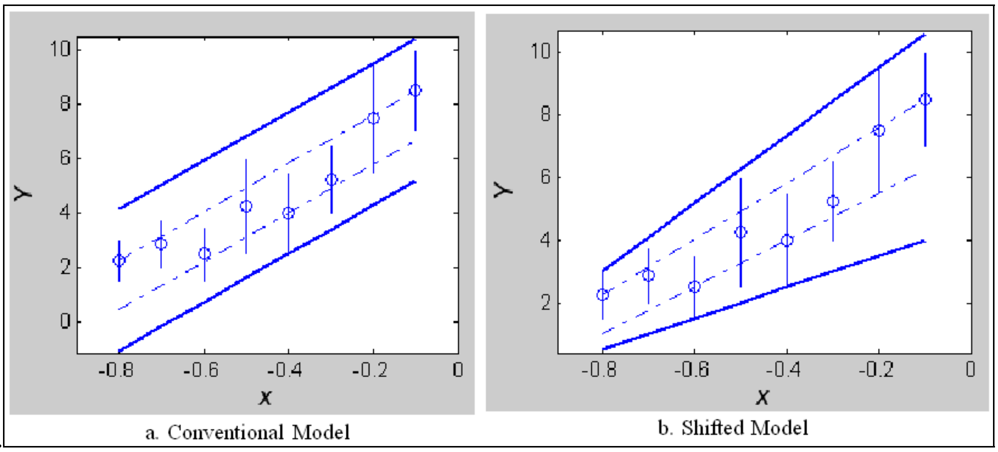


Figure 9: Identified conventional and shifted fuzzy models

Some important points can be underlined on this illustration:

- The conventional model cannot represent an increasing tendency of the observed output spread for negative inputs. As a consequence, the coefficient A_1 is crisp, and the output radius is constant on the domain *D*. On the contrary, with the appropriated *shift* value, the shifted model output well represents the increasing tendency of the observed output radius.
- All the indicators are better for the shifted model, i.e. this model has a better fitting to the data and a lower fuzziness on the domain *D*. This is due to the fact that the shifted model has an improved expressiveness, and can represent any kind of data tendencies.

These illustrations demonstrate the benefits of a trapezoidal shifted model, identified with the minimization of the global fuzziness in fuzzy linear regression. As a consequence, total inclusion property is ensured and the identified model can represent any kind of spread tendencies. Moreover, its robustness is improved.

6 Concluding Remarks

In this paper, a revisited fuzzy regression method has been proposed where a linear model is identified from Crisp-Inputs Fuzzy-Outputs (CISO) data. We have shown that the proposed methodology can be built according to two evolutionary concepts: a new trapezoidal model form and a new expression of the output fuzziness used as an optimization criterion.

The proposed regression strategy ensures a total inclusion property of the observed data in the predicted ones. Moreover, the use of shifted models allows the representation of any kind of fuzzy output spread tendency. In addition, we can say that the proposed criterion is independent of the learning data and presents a certain robustness performance.

The design methodology has been illustrated on simple first order simulated systems and can potentially be applied to a wide class of linear systems according to its generalization for multi-inputs systems. Future work will focus on the extension of this approach to Fuzzy-Inputs Fuzzy-Outputs (FIFO) linear systems.

References

[1] Bisserier A., R. Boukezzoula, and S. Galichet, Linear fuzzy regression using trapezoidal fuzzy intervals, *Proceedings of the 12th Int. Conf. on Information Processing and Management of Uncertainty in Knowledge Based Systems*, pp.181-188, 2008.

[2] Bisserier, A., S. Galichet, and R. Boukezzoula, Fuzzy piecewise linear regression, *Proceedings of the 5th IEEE World Congress on Computational Intelligence*, pp.2089-2094, 2008.

[3] Boukezzoula, R., S. Galichet, and L. Foulloy, Nonlinear internal model control: Application of inverse model based fuzzy control, *IEEE Transactions on Fuzzy Systems*, vol.11, no.6, pp.814-829, 2003.

[4] Boukezzoula, R., L. Foulloy, and S. Galichet, Inverse controller design for interval fuzzy systems, *IEEE Transactions On Fuzzy Systems*, vol.14, no.1, pp.111-124, 2006.

[5] Boukezzoula, R., S. Galichet, and L. Foulloy, MIN and MAX operators for fuzzy intervals and their potential use in aggregation operators, *IEEE Trans. on Fuzzy Systems*, vol.15, no.6, pp.1135-1144, 2007.

[6] Diamond, P., Fuzzy least squares, *Information Sciences*, vol.46, pp.141-157, 1988.

[7] Diamond, P., and R. Korner, Extended fuzzy linear models and least-squares estimates, *Comput. Math. Appl.*, vol.9, pp.15-32, 1997.

[8] Diamond, P., and H. Tanaka, *Fuzzy Regression Analysis, Fuzzy Sets in Decision Analysis, Operations Research and Statistics*, Kluwer Academic Publishers, Norwell, MA, 1999.

[9] D'Urso, P. Linear regression analysis for fuzzy/crisp input and fuzzy/crisp output data, *Computational Statistics & Data Analysis*, vol.42, no.1-2, pp.47-72, 2003.

[10] Ge, H.W., and S.T. Wang, Dependency between degree of fit and input noise in fuzzy linear regression using non-symmetric fuzzy triangular coefficients, *Fuzzy Sets and Systems*, vol.158, no.19, pp.2189-2202, 2007.

[11] Guo, P., and H. Tanaka, Dual models for possibilistic regression analysis, *Computational Statistics & Data Analysis*, vol.51, no.1, pp.253-266, 2006.

[12] Hao, P.Y., and J.H. Chiang, Fuzzy regression analysis by support vector learning approach, *IEEE Trans. Fuzzy Systems*, vol.16, pp.428-441, 2008.

[13] Hojati, M., C.R. Bector, and K. Smimou, A simple method for computation of fuzzy linear regression, *European Journal of Operational Research*, vol.166, pp.172-184, 2005.

[14] Hung, W.L., and M.S. Yang, An omission approach for detecting outliers in fuzzy regression models, *Fuzzy Sets and Systems*, vol.157, no.23, pp.3109-3122, 2006.

[15] Sakawa, M., and H. Yano, Multiobjective fuzzy linear regression analysis for fuzzy input-output data, *Fuzzy Sets and Systems*, vol.47, pp.173-181, 1992.

[16] Savic, D.A., and W. Pedrycz, Evaluation of fuzzy linear regression models, Fuzzy Sets and Systems, vol.39, pp.51-63, 1991.

[17] Tanaka, H., S. Uejima, and K. Asai, Linear regression analysis with fuzzy model, *IEEE, Systems, Trans. Systems Man Cybernet*, vol.2, pp.903-907, 1982.

[18] Tanaka, H., I. Hayashi, and J. Watada, Possibilistic linear regression analysis for fuzzy data, *European Journal of Operational Research*, vol.40, pp.389-396, 1989.
[19] Tanaka, H., and H. Ishibushi, Identification of possibilistic linear systems by quadratic membership functions of fuzzy parameters, *Fuzzy Sets and Systems*, vol.41, pp.145-160, 1991.
[20] Tanaka, H., and H. Lee, Interval regression analysis by quadratic programming approach, *IEEE Trans. Fuzzy Systems*, vol.6, pp.473-481, 1998.
[21] Wu, H.C., Fuzzy estimates of regression parameters in linear regression models for imprecise input and output data, *Computational Statistics & Data Analysis*, vol.42, no.1-2, pp.203-217, 2003.
[22] Yager, R., Using trapezoids for representing granular objects: Applications to learning and OWA aggregation, *Information Sciences*, vol.178, pp.363-380, 2008.
[23] Zadeh, L.A., Fuzzy sets, *Information and Control*, vol.8, pp.338-353, 1965.
[24] Zadeh, L.A., Fuzzy sets as a basis for a theory of possibility, *Fuzzy Sets and Systems*, vol.1, pp.3-28, 1978.