

# An Interval Approach for Fuzzy Linear Regression with Imprecise Data

Amory Bisserier<sup>1</sup> Reda Boukezzoula<sup>1</sup> Sylvie Galichet<sup>1</sup>

<sup>1</sup> Laboratoire d'Informatique, Systèmes, Traitement de l'Information et de la Connaissance (LISTIC)  
 Université de Savoie BP 80439 74941 Annecy le Vieux, Cedex France  
 Email: {amory.bisserier, reda.boukezzoula, sylvie.galichet}@univ-savoie.fr

**Abstract** In this paper, a revisited approach for fuzzy regression linear model representation and identification is introduced. By adopting the commonly used principle of  $\alpha$ -cuts, the fuzzy regression implementation is reduced to the handling of conventional intervals, for inputs, parameters and outputs. Using the Midpoint-Radius representation of intervals, the uncertainty attached to linear models becomes more interpretable. Actually, it is possible to determine the output uncertainty origin (model parameters and/or inputs). In this context, a possibilistic regression method is proposed to identify models of minimal global uncertainty, that is with respect to all possible inputs.

**Keywords** Interval Regression, Fuzzy Regression, Uncertainty Representation, Fuzzy Inputs-Fuzzy Outputs.

## 1 Introduction

Fuzzy regression, a fuzzy type of conventional regression analysis, has been proposed to evaluate the functional relationship between input and output variables in a fuzzy environment. Indeed, unlike statistical regression modeling based on probability theory, fuzzy regression is based on possibility theory and fuzzy set theory [14].

In the fuzzy literature, the regression problem with fuzzy data has been previously treated from different points of view. According to [5], fuzzy regression techniques can be classified into two distinct areas. The first proposed by Tanaka ([10], [12]) which minimizes the total spread of the output is named possibilistic regression. In this case, the problem is viewed as finding fuzzy coefficients of a regression model according to a mathematical programming problem. The second approach developed by Diamond [4], which minimizes the total square error of the output is called the fuzzy least square method. In this paper a possibilistic approach is adopted.

In the chosen possibilistic context, different kinds of input/output data can be considered. Actually, the complete specification of regression problems highly depends on the nature of input-output data [5]. Some researches are thus devoted to Crisp-Input Crisp-Output (CICO) data [9] while others [6], [7], [8], [13] consider the regression problem using Fuzzy-Inputs and Fuzzy-Outputs (FIFO) data. Most commonly, a mixed approach with Crisp-Inputs and Fuzzy-Outputs (CIFO) is used [1], [2], [10], [11], [12].

In this framework, the fuzzy regression with CICO and CIFO data aroused a major interest. However, the FIFO model regression remains a little studied field. Indeed, only a few pa-

pers concerning this problem have been published ([6], [7], [8], [13]). For instance, in [6] and [13], different distances are proposed in order to extend the least square method to FIFO data. In [8], a multiobjective optimization technique for FIFO systems is proposed, whereas in [7] the optimization is based on the minimization of the deviation between observations and predictions.

All above mentioned methods lead to interesting results but present a lack of representativity and an awkward illustration of the uncertainties in the regression model. In order to overcome this problem, an extension of the regression methodology initially developed for CIFO systems in [1], [2] to FIFO ones is proposed. The key idea of the proposed approach resides in using a Midpoint-Radius regression model representation. By doing so, the regression problem can be viewed as a conventional linear optimization problem where a new criterion is proposed.

From a practical point of view and as commonly used in fuzzy regression literature ([5], [7], [8], [12]), the  $\alpha$ -cut principle is adopted. In this case, for a specified  $\alpha$ -cut, the fuzzy interval becomes a conventional interval, which states that a fuzzy interval representation is a generalization of a conventional one. In order to take into account the maximum of uncertainty in the system (the pessimistic case), an  $\alpha$ -cut equal to 0 (0-cut) is considered. In this case, the fuzzy regression problem can be reduced to an interval regression one. That is the approach adopted in this paper where a new regression methodology is proposed.

The structure of this paper is as follows. In section 2, relevant concepts and notations are introduced. Section 3 is devoted to fuzzy regressive models representation. An identification method of such models based on a linear problem optimization is proposed in section 4. An application is shown in section 5. Finally, conclusion and perspectives are presented in section 6.

## 2 Relevant Concepts and Notations

For the sake of rigor and clarity, let us define the basic notions and notations used in this paper.

A fuzzy trapezoidal interval  $A$  is a particular case of a fuzzy number where a possibility distribution, represented by a membership function  $\mu_A$ , is associated with the conventional intervals of the support and the kernel (see Figure 1).

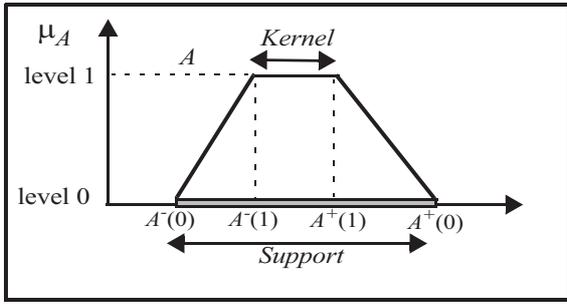


Figure 1: Trapezoidal fuzzy interval representation

In this case, for a given  $\alpha$ -cut ( $\alpha \in [0, 1]$ ) on the fuzzy interval  $A$ , a conventional interval is obtained:

$$A(\alpha) = [A^-(\alpha), A^+(\alpha)] \quad (1)$$

Using the  $\alpha$ -cut representation, a symmetrical fuzzy interval is viewed as a weighted family of nested intervals. In the sequel, in order to consider the maximum uncertainty, only  $\alpha$ -cuts at level 0 will be considered. Consequently, all computations are performed with conventional intervals.

In this framework, an interval  $A$  is defined by the set of elements lying between its lower and upper limits (endpoints) as:

$$A = [A^-, A^+] = \{x | A^- \leq x \leq A^+, x \in \mathfrak{R}\} \quad (2)$$

The interval  $A$  is called thick if  $A^- < A^+$  and thin (or point) if  $A^- = A^+$ .

In order to manipulate and exhibit the uncertainty with intervals, an efficient representation is needed. Indeed, the latter should be able to represent arbitrary intervals and their sets in a single and uniform graphical way. In this framework, it is obvious that to uniquely define an interval, two parameters are necessary, hence the space of intervals is two-dimensional. There are many possible representations of an interval. The most used one is the endpoints representation given in equation (2). However, the main drawback of this representation resides in the fact that the uncertainty (the Radius) is awkward and not directly exhibited in the interval. In order to overcome this difficulty, another representation is adopted. Indeed, in order to facilitate the uncertainty representation, Midpoint ( $M$ ) and Radius ( $R$ ) coordinates are used (see Figure 2). In this case, an interval is represented as a point in the space  $(M, R)$ .

For an interval  $A$ , its Midpoint and its Radius are defined by:

$$M_A = (A^- + A^+)/2 \text{ and: } R_A = (A^+ - A^-)/2 \quad (3)$$

Moreover, the set of all the intervals contained in the interval  $A$  can be represented by a domain  $C(A)$ . As shown in Figure 2, the domain  $C(A)$  is represented by the upper and lower bound lines of  $A$ , forming a triangle.

In the  $(M, R)$  representation, it is clear that the most imprecise element of  $C(A)$  is the interval  $A$  itself which corresponds to the highest value of Radius, i.e. the top of the triangle. On the opposite, the triangle base represents all the crisp values between the lower and the upper bounds (the set of all the thin intervals).

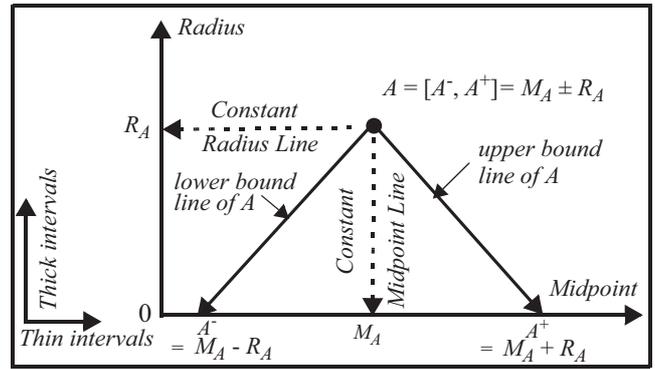


Figure 2: Midpoint-Radius representation of  $A$  and  $C(A)$

For two intervals  $A = [A^-, A^+]$  and  $B = [B^-, B^+]$ , the addition and multiplication operations are defined as follows:

$$\begin{cases} A \oplus B = [A^- + B^-, A^+ + B^+] \\ A \otimes B = [\min(Z), \max(Z)] \end{cases} \quad (4)$$

where:

$$Z = \{A^- B^-, A^- B^+, A^+ B^-, A^+ B^+\} \quad (5)$$

In this case, it can be stated that the Midpoint and Radius of these operations are given by:

$$\begin{cases} M_{A \oplus B} = M_A \oplus M_B \\ R_{A \oplus B} = R_A \oplus R_B \end{cases} \quad (6)$$

and:

$$\begin{cases} M_{A \otimes B} = M_A M_B + \varphi_{A \otimes B} \cdot \text{sign}(M_A M_B) \\ R_{A \otimes B} = R_A R_B + |M_A| R_B + R_A |M_B| - \varphi_{A \otimes B} \end{cases} \quad (7)$$

where:

$$\varphi_{A \otimes B} = \min\{|M_A| R_B, R_A |M_B|, R_A R_B\} \quad (8)$$

For two intervals [3]  $A$  and  $B$ , an inclusion relation of  $A$  in  $B$  (see Figure 3) is defined as follows:

$$A \subseteq B \Leftrightarrow \begin{cases} B^- \leq A^- \\ A^+ \leq B^+ \end{cases} \Leftrightarrow \begin{cases} M_B - R_B \leq M_A - R_A \\ M_A + R_A \leq M_B + R_B \end{cases} \quad (9)$$

$$\Leftrightarrow \begin{cases} M_B - M_A \leq R_B - R_A \\ M_A - M_B \leq R_B - R_A \end{cases} \quad (10)$$

From equation (10) it follows:

$$A \subseteq B \Leftrightarrow |M_B - M_A| \leq R_B - R_A \quad (11)$$

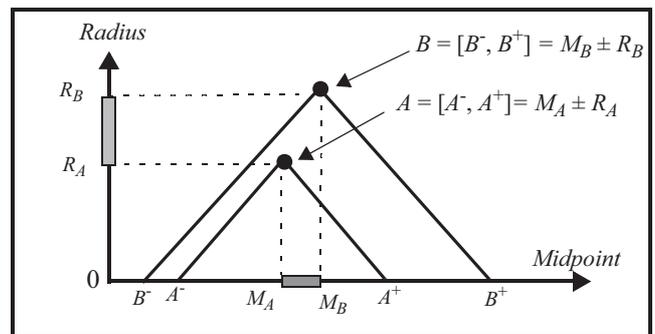


Figure 3: Inclusion of two intervals ( $A \subseteq B$ )

### 3 Fuzzy Regressive Model Representation

The fuzzy linear regression model is the most frequently used form in regression analysis for expressing the relationship between one or more explanatory variables and response. For the sake of simplicity and 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 but more complex in a computational point of view.

An interval regression model is mathematically expressed in the following form:

$$Y(X) = A_0 \oplus A_1 \otimes X \quad (12)$$

where  $Y$  is the interval output (dependant variable),  $A_0$  and  $A_1$  the interval parameters, i.e. unknowns to be estimated from observed data. According to (3), the parameters  $A_0$  and  $A_1$  are denoted as follows:

$$A_0 = [A_0^-, A_0^+] = [M_{A_0} - R_{A_0}, M_{A_0} + R_{A_0}] \quad (13)$$

$$A_1 = [A_1^-, A_1^+] = [M_{A_1} - R_{A_1}, M_{A_1} + R_{A_1}]$$

The input  $X$  can take any value in a definition domain to be chosen.

As mentioned in the introduction, the regression problem has been previously treated from different points of view and by considering different kinds of input/output data. Indeed, according to the nature of  $X$  (crisp or imprecise) two regression problems can be distinguished. The first one deals with Crisp-Input Imprecise-Output data ([12], [10], [1], [2]). In this case, as the model is linear and the inputs are considered as crisp values, the model output (12) will also be an interval. So, for a crisp input  $X = X_*$ , equation (12) becomes:

$$Y(X_*) = [Y_*^-, Y_*^+] = A_0 \oplus A_1 \cdot X_* \quad (14)$$

In this case, the definition domain is the interval  $D = [inf(D), sup(D)]$ , which means that the input  $X$  can take any crisp value in  $D$ .

For the sake of illustration simplicity, in the sequel, the particular case when the parameter  $A_1$  and the input are positive is considered. The other cases are a straightforward generalization of this particular case. The evolution of the model output (14) is illustrated in Figure 4. Actually, as the input  $X$  is a crisp value of  $D$ , the model output uncertainty is only due to the uncertainty of the two parameters  $A_0$  and  $A_1$ .

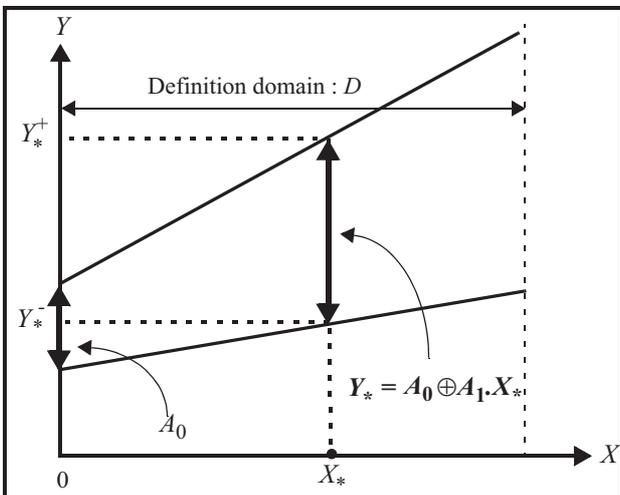


Figure 4: Interval model representation with a crisp input  
The second regression problem is devoted to systems with

Imprecise-Inputs and Imprecise-Outputs. For an interval input  $X = X_* = [X_*^-, X_*^+]$ , the interval model output is given by the following equation with  $Z$  defined according to (5):

$$\begin{aligned} Y(X_*) &= [Y_*^-, Y_*^+] = [A_0^- + \min(Z), A_0^+ + \max(Z)] \\ &= [A_0^- + A_1^- \cdot X_*^-, A_0^+ + A_1^+ \cdot X_*^+] \end{aligned} \quad (15)$$

In this case, the input  $X$  can be any interval included in the interval  $D$  previously defined for crisp inputs. Consequently, the model definition domain is now  $\Delta = C(D)$  (see Figure 2). The model output evolution is presented in Figure 5.

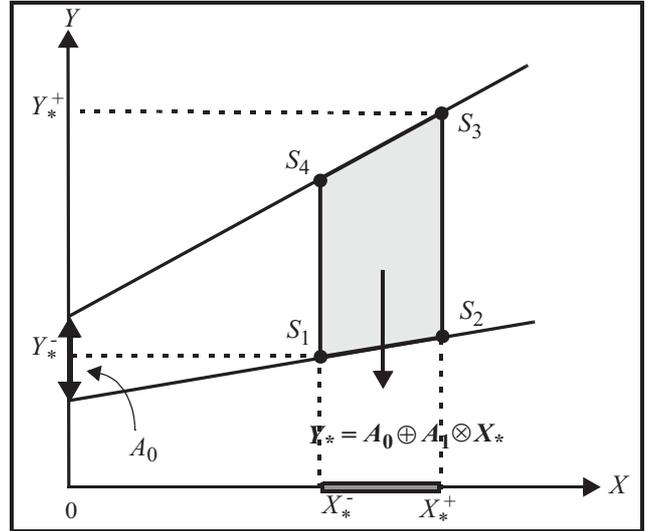


Figure 5: Conventional interval model representation

Several main drawbacks can be underlined:

- The interval input  $X_*$  is represented by its endpoints coordinates on a real line with only one dimension. It leads to a lack of information about the input uncertainty (Radius). In this case, neither the definition domain  $\Delta$  nor the output uncertainty on it can be exhibited in a conventional 2D space.
- According to Figure 5, the model output is defined by the 4-sided area delimited by the four vertices ( $S_1, S_2, S_3, S_4$ ). If the vertices  $S_1$  and  $S_3$  exhibit the model output bounds, the others ( $S_2$  and  $S_4$ ) have no interpretability in the model output. In this situation, two intervals are defined to represent the model output, which limits the interpretability of this kind of representation ([7], [8]).

In order to deal with these drawbacks, the Midpoint-Radius representation is used to define all handled intervals. In this context, the output  $Y$  is expressed by:

$$\begin{aligned} Y(X_*) &= [Y_*^-, Y_*^+] \\ &= [M_{Y(X_*)} - R_{Y(X_*)}, M_{Y(X_*)} + R_{Y(X_*)}] \end{aligned} \quad (16)$$

where:

$$\begin{cases} M_{Y(X_*)} = M_{A_0} + M_{A_1} \otimes X_* \\ R_{Y(X_*)} = R_{A_0} + R_{A_1} \otimes X_* \end{cases} \quad (17)$$

For positive parameter  $A_1$  and input  $X_*$ , we have:

$$M_{A_1} \otimes X_* = M_{A_1} M_{X_*} + \phi_{A_1} \otimes X_* \cdot \text{sign}(M_{A_1} M_{X_*}) \quad (18)$$

As:

$$\begin{aligned} \varphi_{A_1 \otimes X_*} &= \min\{M_{A_1}R_{X_*}, R_{A_1}M_{X_*}, R_{A_1}R_{X_*}\} \\ &= R_{A_1}R_{X_*} \end{aligned} \quad (19)$$

the equation (18) becomes:

$$M_{A_1 \otimes X_*} = M_{A_1}M_{X_*} + R_{A_1}R_{X_*} \quad (20)$$

By adopting the same principle, the following Radius expression is determined:

$$\begin{aligned} R_{A_1 \otimes X_*} &= R_{A_1}R_{X_*} + M_{A_1}R_{X_*} + R_{A_1}M_{X_*} - R_{A_1}R_{X_*} \\ &= M_{A_1}R_{X_*} + R_{A_1}M_{X_*} \end{aligned} \quad (21)$$

By substitution of equations (20) and (21) in (17), it follows:

$$\begin{cases} M_{Y(X_*)} = M_{A_0} + M_{A_1}M_{X_*} + R_{A_1}R_{X_*} \\ R_{Y(X_*)} = R_{A_0} + M_{A_1}R_{X_*} + R_{A_1}M_{X_*} \end{cases} \quad (22)$$

The Midpoint and Radius output evolutions on the domain  $\Delta$  are illustrated in Figure 6 and Figure 7.

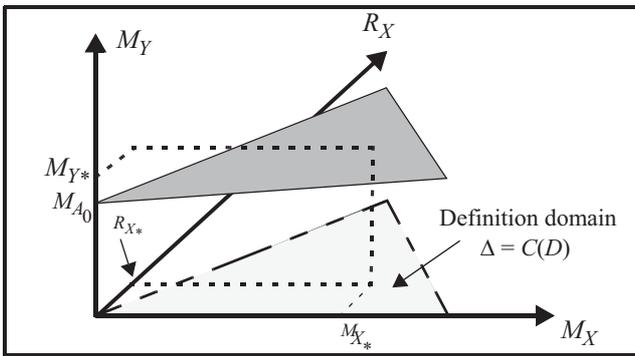


Figure 6: The midpoint output evolution

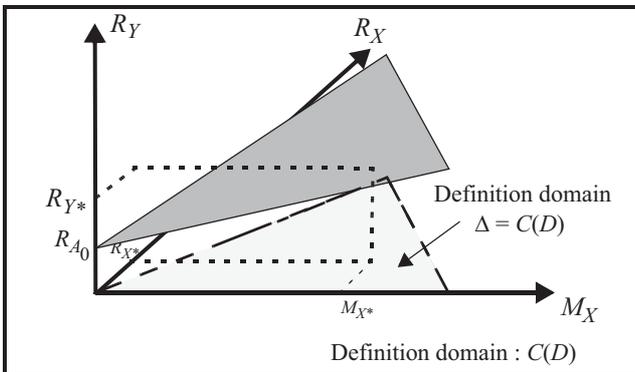


Figure 7: The radius output evolution.

From equations (16) and (22), it is obvious that:

$$\begin{cases} Y_*^- = M_{A_0} - R_{A_0} + (M_{A_1} - R_{A_1})(M_{X_*} - R_{X_*}) \\ Y_*^+ = M_{A_0} + R_{A_0} + (M_{A_1} + R_{A_1})(M_{X_*} + R_{X_*}) \end{cases} \quad (23)$$

A global representation of the model output evolution on the domain  $\Delta$  is illustrated in Figure 8.

This Midpoint-Radius representation presents several advantages:

- In this representation, only one interval output is defined for a considered input. Indeed, the output expression represents as well the contribution of the input Midpoint and Radius.
- According to this representation, the origin of uncertainty in the output can be precisely determined. In other words,

the contributions of the parameters  $A_0, A_1$  and the input  $X$  are well defined in the model output (22).

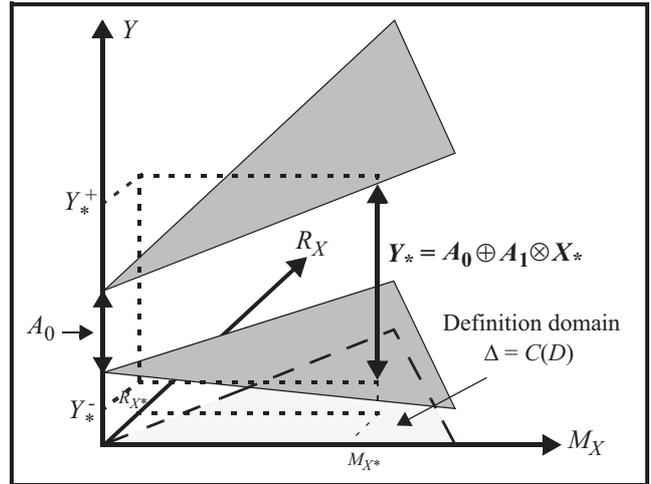


Figure 8: The model output evolution

### 4 Interval Regressive Model Identification

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 this paper, as commonly used, the model structure is assumed to be linear. In this case, the regression problem is reduced to an estimation problem of the model parameters.

Let us consider a set of  $N$  observed data samples. Let the  $j^{th}$  sample be represented by the couple  $(X_j, Y_j), j = 1, \dots, N$  where  $X_j$  are interval inputs and the  $Y_j$  are the corresponding interval output.

All the observed input intervals are assumed to be included in the interval  $D = [D^-, D^+]$  given by:

$$D^- = \min(X_j^-) \text{ and } D^+ = \max(X_j^+) \quad (24)$$

Like any regression technique, the fuzzy regression objective is to determine a predicted functional relationship between inputs  $X$  and outputs  $Y$ :

$$\hat{Y}(X) = A_0 \oplus A_1 \otimes X \quad (25)$$

In this case, the definition domain of the identified model is the set of all the intervals included in  $D$ , i.e.  $\Delta = C(D)$ .

As discussed in the paper introduction, a possibilistic regression approach is adopted where the objective is to determine the model parameters such that the observed data are included in the predicted ones (possibility model). In this framework, another kind of model (necessity one) could be identified [11].

In the proposed identification methodology, two points have to be considered for its implementation:

- the interval inclusion constraints to be introduced in the optimization problem,
- the identification criterion to be optimized.

#### 4.1 The optimization constraints

In this case, the inclusion constraints are given by:

$$Y_j \subseteq \hat{Y}_j \Leftrightarrow |M_{\hat{Y}_j} - M_{Y_j}| \leq R_{\hat{Y}_j} - R_{Y_j} \quad (26)$$

with:

$$M_{\hat{Y}_j} = M_{A_0} + M_{A_1} \otimes X_j \text{ and: } R_{\hat{Y}_j} = R_{A_0} + R_{A_1} \otimes X_j \quad (27)$$

By substitution of (27) in (26), the constraints inclusion can be written as:

$$\begin{aligned} |M_{A_0} + M_{A_1} M_{X_j} + R_{A_1} R_{X_j} - M_{Y_j}| \leq \\ R_{A_0} + M_{A_1} R_{X_j} + R_{A_1} M_{X_j} - R_{Y_j} \end{aligned} \quad (28)$$

Moreover, the optimization constraints must ensure the identification of intervals, whose bounds are ordered. Moreover, the sign of the parameter  $A_1$  must be constrained as well. In the case of a model whose parameter  $A_1$  is identified as positive, it follows:

$$R_{A_0} \geq 0, \text{ and } : M_{A_1} \geq R_{A_1} \geq 0 \quad (29)$$

It is important to note that all the considered constraints in the identification problem are linear according to the identification variables (Midpoint and Radius of  $A_0$  and  $A_1$ ).

#### 4.2 The used criterion

The objective is to identify the less imprecise model on its definition domain, as it has been proposed in our previous works [1]. The computation of the global uncertainty on the whole domain  $D$  proposed in [1] is here extended to the whole domain  $\Delta$ , whose elements are intervals. It follows that the criterion  $J$  to be optimized, i.e. the global uncertainty on  $\Delta$ , is expressed as::

$$\begin{aligned} J &= \iint_{\Delta} (\hat{Y}^+ - \hat{Y}^-) dM_X dR_X = 2 \iint_{\Delta} R_{\hat{Y}} dM_X dR_X \\ &= 2 \iint_{\Delta} (R_{A_0} + M_{A_1} R_X + R_{A_1} M_X) dM_X dR_X \end{aligned} \quad (30)$$

The substitution of equation (23) in (30) yields to:

$$J = 2R_{A_0} \cdot R_D^2 + 2R_{A_1} \cdot R_D^2 \cdot M_D + \frac{2}{3} M_{A_1} \cdot R_D^3 \quad (31)$$

To sum up, the identification method is performed by minimizing the criterion (31) under the constraints (28) and (29).

At this stage, several points can be discussed about the proposed approach:

- As the interval  $D$  is defined with the observed inputs, its Radius and Midpoint are known numerical values. So, the criterion  $J$  is linear according to the optimization variables. Moreover,  $J$  is independent of the observed inputs, i.e. the optimization is made on the whole domain  $\Delta$  of the model, not only on the observed data. This point contributes to improve the model robustness.
- It is obvious that the criterion must be optimized under constraints in order to ensure inclusion property. If the constraints are released, the identified model corresponds to the minimal possible uncertainty, achieved with null parameters, and has no sense.
- As the minimization of the model uncertainty is performed on  $\Delta$ , the input uncertainty is considered in the model output expression in the identification process.

## 5 Illustrative example

In this part, the proposed identification method is applied on the illustrative example given in [8], [7]. Given the data set [8], [7] illustrated in Table 1, corresponding of  $\alpha$ -cuts of level 0 of proposed fuzzy data set, the objective is to identify the model parameters  $A_0$  and  $A_1$ .

Table 1: Observed data set and predicted one

inputs $X_j$	observed outputs $Y_j$	predicted outputs $\hat{Y}_j$
[1.5 ; 2.5]	[3.5 ; 4.5]	[3.5 ; 6.5]
[3 ; 4]	[5 ; 6]	[4.16 ; 7.25]
[4.5 ; 6.5]	[6.5 ; 8.5]	[4.82 ; 8.5]
[6.5 ; 7.5]	[6 ; 7]	[5.69 ; 9]
[8 ; 9]	[8 ; 9]	[6.35 ; 9.75]
[9.5 ; 11.5]	[7 ; 9]	[7 ; 11]
[10.5 ; 11.5]	[10 ; 11]	[7.45 ; 11]
[12 ; 13]	[9 ; 10]	[8.1 ; 11.75]

The considered definition domain is  $D = [1.5, 13]$  and represented in Figure 9. This domain corresponds to a set of positive inputs, and, as discussed previously, the parameter  $A_1$  is assumed to be positive too.

The identified parameters corresponding to the less imprecise model according to the criterion  $J$  defined in (31) are:

$$A_0 = [2.844 ; 5.25] \text{ and } A_1 = [0.438 ; 0.5] \quad (32)$$

In this case, the optimal value of uncertainty obtained is  $J = 153.95$ .

The identified model output representation is given in Figure 9, whereas the predicted outputs are presented in Table 1.

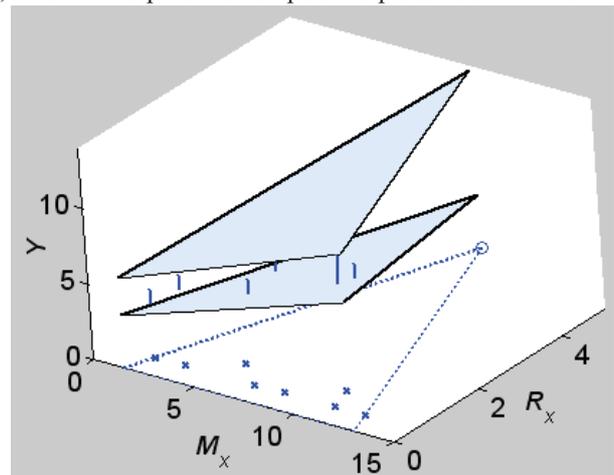


Figure 9: Identified model representation.

As presented in Table 1, the main advantage of the proposed approach compared to those presented in [7] and [8] is that only one interval output corresponds to a considered input. In Figure 9, it can be stated that the desired inclusion is respected.

Another advantage of the representation using the (Midpoint, Radius) space for the interval inputs concerns the exploitation

of the model characteristics. According to (22), the output Radius evolution on  $\Delta$  can be represented (see Figure 10). On the axe ( $M_X, R_X = 0$ ), i.e. the triangle basis, only the parameter uncertainty is exhibited, whereas at the top of the triangle, the output uncertainty is maximum, corresponding to the most uncertain input in  $\Delta$ , i.e.  $D$ . In this example (Figure 10), the output radius is nearly constant on the basis of the triangle, corresponding to the case of a nearly crisp parameter  $A_1$ . The output radius increasing on  $\Delta$  is essentially due to the input uncertainty.

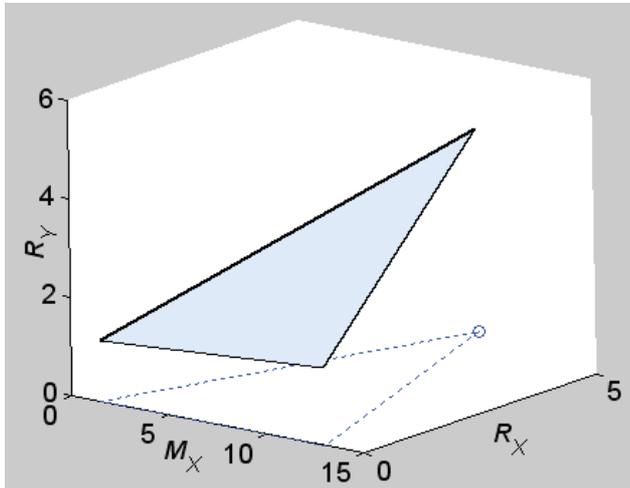


Figure 10: Identified model output Radius evolution.

The same analysis can be performed on the output Midpoint evolution on  $\Delta$ , corresponding to the global tendency of the output on the domain.

In [8], the identified model, whose output respects the inclusion statement, corresponds to a criterion value  $J = 281.03$ , according to (31). Its parameters are the following:

$$A_0 = [1.25 ; 6.75] \text{ and } A_1 = [0.4 ; 0.56] \quad (33)$$

In fact, the vertical dimension  $\alpha$  is introduced into the proposed multiobjective optimization problem. It is known that the higher the  $\alpha$  value is set for optimization, the more uncertain the identified model is and inclusion is respected for any lower  $\alpha$  levels. So, this compromise between higher level of inclusion and lower uncertainty leads to the identification of a model whose uncertainty is not minimal on the domain.

On the other hand, the proposed method in [7], which minimizes deviations between observations and predictions without constraining inclusion, leads to the identification of a model which corresponds to a criterion value  $J = 102.6$ , and whose parameters are:

$$A_0 = [3 ; 3.82] \text{ and } A_1 = [0.5 ; 0.54] \quad (34)$$

It is obvious that guaranteeing the inclusion increase the model uncertainty.

## 6 Conclusion

In this paper, a revisited representation of interval regressive model is proposed and used in an identification problem. By adopting the Midpoint-Radius formalism for intervals, uncertainty representation is improved. Applied to regressive interval models, this formalism allows a better handling of the input uncertainty in the definition domain establishment and in the model output definition. In an identification problem, based on the optimization of a linear expression of the model output global uncertainty on its definition domain, the desired

inclusion is ensured.

Further work will focus on the extension of this approach to fuzzy linear systems, by considering the vertical dimension  $\alpha$ . In the proposed space, a symmetrical trapezoidal fuzzy interval can be interpreted by a zooming operation as illustrated in the following Figure 11:

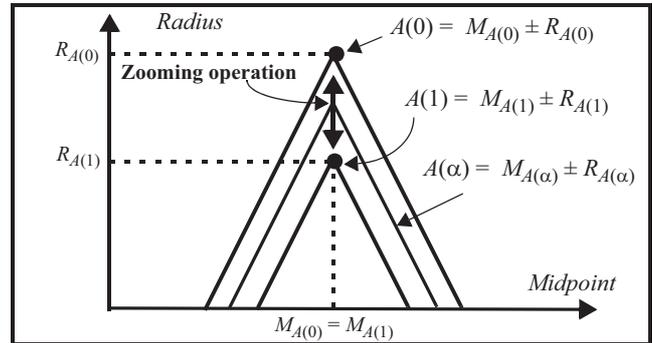


Figure 11: Midpoint-Radius representation of symmetrical trapezoidal fuzzy interval  $\alpha$ -cuts.

A particular attention should be paid to the loss of linearity of the membership function of the model output, due to the fuzzy product between the input  $X$  and the parameter  $A_1$ .

## References

- [1] A. Bissierier, S. Galichet, R. Boukezzoula, Fuzzy Piecewise Linear Regression, *5th IEEE World Congress on Computational Intelligence (WCCI 2008 / FUZZ-IEEE 2008)*, Hong-Kong, China, 2008, 2089-2094.
- [2] A. Bissierier, R. Boukezzoula, S. Galichet, Linear fuzzy regression using trapezoidal fuzzy intervals, *Information Processing and Management of Uncertainty (IPMU2008)*, Malaga, Spain, 2008, 181-188
- [3] R. Boukezzoula, S. Galichet and L. Foulloy, MIN and MAX Operators for Fuzzy Intervals and their Potential Use in Aggregation Operators, *IEEE Trans. on Fuzzy Systems* 15(6): 1135-1144, 2007
- [4] P. Diamond, Fuzzy Least Squares, *Information Sciences* 46: 141-157, 1988.
- [5] P. Diamond, H. Tanaka, Fuzzy regression analysis, *Fuzzy sets in decision analysis, operations research and statistics*, Kluwer Academic Publishers, Norwell, MA, 1999.
- [6] P. D'Urso, Linear regression analysis for fuzzy/crisp input and fuzzy/crisp output data, *Computational Statistics & Data Analysis* 42 (1-2), 2003, 47-72.
- [7] M. Hojati, C.R. Bector, K. Smimou, A simple method for computation of fuzzy linear regression, *European Journal of Operational Research* 166: 172-184, 2005.
- [8] M. Sakawa, H. Yano, Multiobjective fuzzy linear regression analysis for fuzzy input-output data, *Fuzzy Sets and Systems* 47: 173-181, 1992.
- [9] D.A. Savic, W. Pedrycz, Evaluation of fuzzy linear regression models, *Fuzzy Sets and Systems* 39, 1991, 51-63.
- [10] H. Tanaka, I. Hayashi, J. Watada, Possibilistic linear regression analysis for fuzzy data, *European Journal of Operational Research* 40: 389-396, 1989.
- [11] H. Tanaka, H. Lee, Interval regression analysis by quadratic programming approach, *IEEE Trans. on Fuzzy Systems* 6(4): 473-481, 1998.
- [12] H. Tanaka, S. Uejima, K. Asai, Linear regression analysis with fuzzy model, *IEEE, Systems, Trans. Systems Man Cybernet.* SMC-2: 903-907, 1982.
- [13] H.-C. Wu, Fuzzy estimates of regression parameters in linear regression models for imprecise input and output data, *Computational Statistics & Data Analysis* 42 (1-2), 2003, 203-217.
- [14] L.A. Zadeh, Fuzzy Sets, *Information and Control* 8: 338-353, 1965.