Estimating Fuzzy Linear Regression Model for Air Pollution Predictions in Baghdad City

Suhaila Najma Alsoltany and Iftikhar Abdulhamed Alnaqash
Department of Mathematics, College of Education, University of Al-Mustansiriyah.
Department of Statistic, College of Economic and Administration,
University of Al-Mustansiriyah.

Abstract

Regression analysis is one f the basic tools of scientific investigation of functional relationship between dependent and independent variables, For many years linear regression models has been used in almost every field of science. The purpose of regression analysis is to explain the variation of dependent variables in terms of the variation of explanatory variables, residuals are assumed to be due to random errors, however the residuals are sometimes due to the indefiniteness of the model structure or imprecise observations, the uncertainty in this type of regression model becomes fuzziness, not random.

The aim of this paper is to study and applied the method of estimation fuzzy linear regression parameters using fuzzy data collecting from (145) sample in three stations (Andalus square, jadiriya, alawi) in Bagdad city every day, In order to measurements the concentrations of airborne stuck which represents the response variable, and also the most important air pollutants, namely, (lead, zinc, copper, iron, nickel, chromium, cadmium) as independents variables the main result identify the best techniques to estimate the fuzzy linear regression parameters for this data and calculates the expected value of the concentrations of airborne stuck in Bagdad city for the next years.

Keywords**:** fuzzy data,fuzzy regression, regression analysis, air quality.

1-Introduction

For many years the efforts of humans focuses on fight air pollution caused by its group of activities and effectiveness in the life, where it is difficult to humans even use all his available capacity to control on the natural phenomena that will cause change in the air quality and its pollution, this efforts may be successes in mitigate the negative effects of these phenomena, the basic goal of science is a permanent quest to knowledge and interpretation the phenomena and different relationships between its, through studies the variables presented these phenomena and the interference between its effects.

Classical regression analysis is helpful in certainty the probable form of the relationship between variables, and usually the basic objective is to predict or estimate, the value of one variable corresponding to a given value of another variable, residuals are assumed to be due to random errors. In many world problems observation can be described only in fuzzy data this type of data is easy to find in natural language, social science, psychometrics, environments, and econometrics etc.), fuzzy set theory provides a means for modeling such data utilizing fuzzy membership function. Fuzzy regression was deal with fuzzy data; Regression is based probability theory whereas the fuzzy regression is based on probability theory & fuzz y set theory [1].

The concept of fuzziness in regression analysis leads us to the fuzzy linear regression (FLR) models[4], In general, the estimation problems of fuzzy uncertainty of dependent variables with the fuzziness of response function(regression coefficients)considered as the parameter estimations of (FLR) models, Methods of estimating the parameters of fuzzy linear regression models can be roughly divided into two categories, first is the adoption of Tanaka et al approach, (1982) [6], they formulated a linear regression model with fuzzy response data, crisp predictor data and fuzzy parameters as a mathematical programming problem (LP) under two limitations are appropriate degree model and the degree of uncertainty and then solve this problem, The second is based on a least squares Fuzzy approach, Diamond [3] proposed the fuzzy least squares approach to determine fuzzy parameters by defining a metric between two fuzzy numbers, various researchers attempted extending the regression analysis from crisp to fuzzy domain, which is a fuzzy extension of the ordinary least squares based on a new defined distance on the space of fuzzy numbers.

2- **Topic of the study**

As result of special circumstances that pass in Iraq, there are many negative situations deployed in the streets of Baghdad and other governorates which is practices contaminated environmental in general and air, in particular, it's appeared as a result of many phenomena, continuous power outages, poor services provided to peoples or absence it's, the weakness of the environmental awareness for the implications that caused by such practices, and the difficulty of application environmental laws against violators result from the weakness security situation.

The increase of air pollution that causes respiratory and eye diseases in add ion to increasing the concentration of some chemical compounds cause some types of cancer, which motivated us to study the problem of the deterioration of air quality, which is one of the most important problems facing all governorates in Iraq and a greater degree in Baghdad.

3- **Aim of the study**

The main goals of this paper are:

a. Identify the types of Fuzzy Linear Regression Model to obtain an appropriate linear relation between dependent variable and several independent variables in the cases when the values of (response variable, explanatory variable or both) are classes values, i.e. in a fuzzy environment, explain the methods for evaluating Fuzzy Parameters in Linear Regression models and display it's easy computationally that the researchers in the other scientific fields use these models in the data analysis.
b. Construct a Fuzzy Linear Regression model to predict the concentrations hazy airborne minutes of the city of Baghdad, which affect air quality Depending on the proportions of the heavy materials in the air (lead, zinc, copper, iron, nickel, chromium, cadmium).

4- **Theoretical Side**

a. *Fuzzy Set Theory* [8]

Zadeh 1965 describes the fuzzy uncertainty with ambiguity and haziness and introduces the theory of fuzzy to build such a system as

Let F(R) denote the set of normalized fuzzy number, denoted the set of fuzzy numbers with compact α-level sets $(\widetilde{X_\alpha})$ for α.>0, by $F_c(R)$, The fuzzy number is defined as a generalization of the real number in the sense that it does not refer to the number by one value but by connected set of possible values. So the fuzzy number is a special case of the fuzzy set of the convex real line [3].

Following the concept of fuzzy set construct by Zadeh where he define the fuzzy set $(\widetilde{A})$at the comprehensive set of elements (x) as a set of elements with the degree of membership function $\mu_{\tilde{A}}(x)$take values in therange from [0,1]. In the usual notation if respectively the left and right points of the triangular as shown below [1].

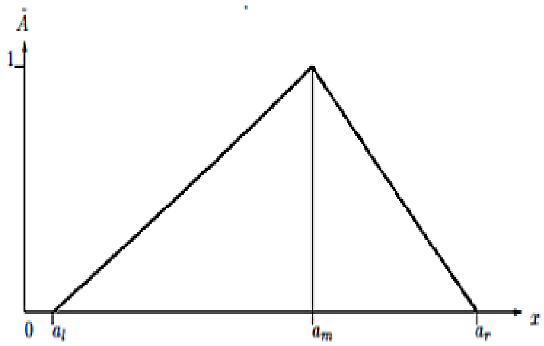

Fig.(1) Triangular fuzzy number.

b. *Fuzzy Linear Regression Models*

Fuzzy linear regression is a fuzzy type of classical regression analysis in which some elements of the model are represent by fuzzy numbers, it is used in evaluating the functional relationship between the dependent and independent variables in a fuzzy environment, The fuzzy regression model may be roughly classified by conditions of independent and dependent variables into three categories as follows [7].

I. Input and output data are both non-fuzzy number.
II. Input data is non-fuzzy number but output data is fuzzy number.
III. Input and output data are both fuzzy number.

Category I:

We may need to consider that the relationship between variables in linear regression model may be fuzzy. The basic model assumes a fuzzy linear function as [4], [6].

$$\hat{y}_i = \tilde{A}_0 x_{i0} + \tilde{A}_1 x_{i1} + \tilde{A}_2 x_{i2} + \cdots + \tilde{A}_p x_{ip},\quad i = 1,2,\dots,n \quad \dots\dots (1)$$

In matrix form we have

$$\hat{Y} = \tilde{A}\mathrm{X}$$

Where:

$X = [x_1, x_2, \dots\dots x_p]^T$ Vector Crisp Independent Variables.

$\widetilde{A} = [\tilde{A}_1, \tilde{A}_2, \dots, \tilde{A}_p]^T$ Vector Fuzzy Coefficient presented in the form of a symmetric triangular fuzzy number denoted by $\widetilde{A}_j = [\propto_j, r_j]$ with its membership function described as follow [1]:

$$\mu_{\tilde{A}_j}(a_j) = \begin{cases} 1 - \dfrac{|\propto_j - a_j|}{r_j}, & \propto_j - a_j \le a_j \le \propto_j + a_j \quad \forall j, \\ 0 & \text{other wise} \end{cases} \quad j = 1,2,\dots,p \quad \dots\dots (2)$$

Where:

$\propto_j$: RepresentCenter numbers $\tilde{A}_j$ $\forall j, j = 1,2,\dots,p$

r_j : Representspread numbers $\tilde{A}_j$ $\forall j, j = 1,2,\dots,p$

So we can write the regression model in (1) as follows:

$$\hat{y}_i = (\propto_0, r_0)x_{i0} + (\propto_1, r_1)x_{i1} + (\propto_2, r_2)x_{i2} + \cdots + (\propto_p, r_p)x_{pi} \quad, i = 1,2,\dots,n \quad \dots\dots (3)$$

The above fuzzy regression analysis assume the crisp input and output data while the relation between them is denoted by fuzzy function of which the distribution of the parameter is a possibility function, The membership function of the fuzzy number $(\tilde{y}_i)$ is given by [1]:

$$\mu_{\tilde{A}_j}(a_j) = \begin{cases} 1 - \dfrac{|y_i - X^t \propto_j|}{r^t |X|}, & X \ne 0 \\ 1 & X = 0, Y \ne 0 \quad \forall j, j = 1,2,\dots,p \\ 0 & X = 0, Y = 0 \end{cases} \quad \dots\dots (4)$$

Where each value of dependent variable can be estimated as a fuzzy $\tilde{y}_i = [\tilde{y}_i^l, \tilde{y}_i^{h=1}, \tilde{y}_i^u] \forall i$, $i = 1,2,\dots\dots n$ where

$\tilde{y}_i^l = \sum_{j=0}^{p} (\propto_j - r_j) x_{ij}$ Represent lower limit of dependent variable.

The center value ofdependent variable $(\tilde{y}_i)$ is $\tilde{y}_i^{h=1} = \sum_{j=0}^{p} (\propto_j) x_{ij}$

$\tilde{y}_i = \sum_{j=0}^{p} (\propto_j + r_j) x_{ij}$Represent upper limit of dependent variable

Note: if the relationship between dependent and independent variables in linear regression model may be not fuzzy the model will be classical linear regression model.

Category II:

Consider the following general fuzzy linear regression model:[2]

$$\tilde{y}_i = \tilde{A}_0 x_{i0} + \tilde{A}_1 x_{i1} + \tilde{A}_2 x_{i2} + \cdots + \tilde{A}_p x_{ip},\quad \forall i, i = 1,2,\dots,n \quad \dots\dots (5)$$

Where: x_{ij} are real numbersrepresent input variables

$y_i = [c_i - s_i,\ c_i s_i]$ are fuzzy numbers represent output variable, where:$c_i$is the center (meanor mode) value, $s_i$is the spread.

$A_j = [\alpha_j - r_j, \alpha_j r_j]$ are the fuzzy regression parameters, which has the same membership function as y_i,where: $\alpha_i$is the center (mean or mode) value, $r_i$is the spread.

The general fuzzy regression model in (5) isrewrite as:

$$(c_i, s_i) = (\propto_0, r_0)x_{i0} + (\propto_1, r_1)x_{i1} + \cdots + (\propto_p, r_p)x_{ip} \quad \forall i, i = 1,2,\dots.,n \quad \dots\dots (6)$$

Note: if the relationship between dependent and independent variables in linear regression model may be not fuzzy the model may be Multivariate Regression Model.

Category III:

Consider the following general fuzzy linear regression model [2],[8]

$$\tilde{y}_i = \tilde{A}_0 \tilde{x}_{i0} + \tilde{A}_1 \tilde{x}_{i1} + \tilde{A}_2 \tilde{x}_{i2} + \cdots + \tilde{A}_p \tilde{x}_{ip} \quad, = \widetilde{AX} \forall i, i = 1,2,\dots,n \quad \dots\dots (7)$$

Where: $x_{ij}= [d_{ij}- t_{ij},d_{ij}+ t_{ij}]$ are fuzzy numbers representing input variables,

Where: d_{ij}:is the center (mean or mode) value, $t_{ij}$is the spread.

$y_i= [c_i- s_i,c_i+ s_i]$ are fuzzy numbers represent output variable, where: $c_i$is the center (mean or mode) value, $s_i$is the spread.

$A_j= [\alpha_j- r_j,\alpha_j+ r_j]$ are the fuzzy regression parameters, which has the same membership function as y_i, where: $\alpha_i$is the center (mean or mode) value, $r_i$is the spread.

The general fuzzy regression model in (5) is rewrite as:

$$(c_i,s_i) = (\propto_0,r_0)(d_{i0},t_{i0}) + (\propto_1,r_1)(d_{i1},t_{i1}) + \cdots + (\propto_p,r_p)(d_{ip},t_{ip}) \forall i, i = 1,2, \dots n \quad \text{.............................. (8)}$$

Note: if the relationship between dependent and independent variables in linear regression model may be not fuzzy the model may be like a classical regression only the formal for estimated as follow

$$\hat{\beta}_i = \frac{\sum_{i=1}^{n}(x_{il}y_{il}+ x_{im}y_{im}+x_{iu}y_{iu})-3n\bar{x}\bar{y}}{\sum_{i=1}^{n}(x_{il}^2+ x_{im}^2+ x_{iu}^2)-3n\bar{x}^2} \quad \text{........... (9)}$$

$$\bar{x} = \frac{\sum_{i=1}^{n}(x_{il}+x_{im}+x_{iu})}{3n} \quad \text{.............................. (10)}$$

$$\bar{y} = \frac{\sum_{i=1}^{n}(y_{il}+y_{im}+y_{iu})}{3n} \quad \text{.............................. (11)}$$

c. *Estimation Fuzzy Linear Regression ParametersTecqhinecs*

1- First Method(Tanaka approach)[5],[6]

This method based on minimizing fuzziness for the model fitting, through transfer the problem to a linear programming problem, the objective function is to minimize the total spread of the fuzzy number $\hat{y}_i$.

$$\min r^t|x_{ij}| = \min r_j \sum_{j=0}^{p} x_{ij} \quad \text{.................. (12)}$$

The constraint problem is the limits of the observation to response variable $\hat{y}_i$ require that each observation $\hat{y}_i$ has the least (h) degree of belonging to $\hat{y}_i$ this lead to formulate a linear programming problem as follow [1]:

$$\left.\begin{array}{c} \min \quad (r_j \sum_{j=0}^{p}|x_{ij}|) \\ \textit{subject to} \\ \sum_{j=0}^{p} \propto_j x_{ij} + (1+h)\sum_{j=0}^{p} r_j|x_{ij}| \geq \tilde{y}_i \\ \sum_{j=0}^{p} \propto_j x_{ij} - (1+h)\sum_{j=0}^{p} r_j|x_{ij}| \leq \tilde{y}_i \\ c_j \geq 0 \end{array}\right\} \quad \text{.............................. (13)}$$

2. Second Method [3],[4]

In this method, using the least square errors as criteria to fined estimating parameters for fuzzy regression model by reducing the total square of the distance (square total error)

$$\min \sum_{i=1}^{n} d^2 (y_i, \tilde{y}_i) \quad \text{.................................. (14)}$$

To estimate the parameters above model there are two approaches:

a) Diamond approach [5]

In this approach we shall construct two regression models in (5) to the ends of the class$\tilde{y}_i$:

$$\tilde{y}_{li} = \tilde{l}_0 + \tilde{l}_1x_{i1} + \tilde{l}_2x_{i2} + \cdots + \tilde{l}_px_{ip} \forall i, \; i = 1,2, \dots, n \quad \text{.. (15)}$$

Where: $\tilde{y}_{li} = c_i - s_i$ and

$$\tilde{y}_{ui} = \tilde{u}_0 + \tilde{u}_1x_{i1} + \tilde{u}_2x_{i2} + \cdots + \tilde{u}_px_{ip} \forall i, \; i = 1,2, \dots, n \quad \text{....................................... (16)}$$

Where: $\tilde{y}_{ui} = c_i + s_i$

and you can write the following formula:

$$s(\tilde{l}_j) = \sum_{i=1}^{n} d^2 \left(\tilde{y}_{il}, \hat{\tilde{l}}_0 + \hat{\tilde{l}}_{ij}x_{ij}\right)$$

$$s(\tilde{l}_j) = \sum_{i=1}^{n} \left(\tilde{y}_{il} - (\hat{\tilde{l}}_0 + \hat{\tilde{l}}_{ij}x_{ij})\right)$$

$$\frac{\partial s(\tilde{l}_j)}{\partial \tilde{l}_j} = 0$$

$$\hat{\tilde{l}}_j = (x^tx)^{-1}x^t\tilde{y}_l \quad \text{...................................... (17)}$$

By the same way we find

$$\frac{\partial s(\tilde{u}_j)}{\partial \tilde{u}_j} = 0$$

$$\hat{\tilde{u}}_j = (x^tx)^{-1}x^t\tilde{y}_u \quad \text{.................................... (18)}$$

Then we get parameters of regression model $\tilde{A}_m = [\tilde{\alpha}_m - \tilde{r}_m, \tilde{\alpha}_m + \tilde{r}_m]$

Where:

$$\tilde{\alpha}_m = \frac{\hat{l}_i+\hat{u}_i}{2} \quad ......(19)$$

$$\tilde{r}_m = \frac{\left|\hat{u}_i-\hat{l}_i\right|}{2} \quad ......(20)$$

b. Least squares approach [4],[7]

In this approach we can use ordinary least-squares method to estimate the fuzzy parameters in the general fuzzy linear regression model (5), by using the formula of the model in (6):

The least-squares estimates of $a_i$and $r_i$are the values of a_i, r_i which minimize the value of$D^2$where

$$D^2 = \sum_{i=0}^{n}[\left(c_i - (\propto_0+\propto_1 x_{i1}+..+\propto_p x_{ip})\right)^2 + \left(s_i - (r_0 + r_1 x_{i1}+..+r_p x_{ip})\right)^2 \quad ......(21)$$

Let $\|\vec{v}\|$denote the length of vector v, then by using vector and matrix expressions D^2can be rewritten as

$$D^2 = \|x \propto -c\|^2 + \|xr - s\|^2 \quad ......(22)$$

Where:

$$\propto = (\propto_0, \propto_1, .., \propto_p)^t$$
$$r = (r_0, r_1, .., r_p)^t$$
$$c = (c_0, c_1, .., c_n)^t$$
$$s = (s_0, s, .., s_n)^t$$
$$\frac{\partial D^2}{\partial \propto}=0, \frac{\partial D^2}{\partial r}=0$$

Then the solutions of $\propto$and rwhich minimize$D^2$are as follows:

$$\tilde{\propto} = (x^t x)^{-1} x^t c \quad ......(24)$$
$$\tilde{r} = (x^t x)^{-1} x^t s \quad ......(25)$$

5. Practical side

5-1 Description of data

The data recorded by the adoption of the measure ambient air pollutants in the three stations of the governorate of Baghdad devices (Andalus Square, Jadiriya, Alawi Square) and published in the (state of the environment in Iraq Report, 2009) prepared by the Ministry of Environment, where collected (145) daily model as follows (57 daily Al-Andalus Square station model 0.43 daily Jadiriya station model 0.45 daily Al-Alawi station model) to measure the concentrations of airborne minutes (dispersant materials or scattered in the air (solid, liquid, gas), which released to the atmosphere from natural sources or be due to various human activities) units of measurement (Micro/ m 3) Which represents the response of the study variable, as well as the measurement of concentrations of heavy metals to air pollutants (lead, zinc, copper, iron, nickel, chromium, cadmium) of the airborne minutes models in three stations, which represent the explanatory variables for study table below shows the data of the study.

Table (1)
The minimum and maximum monthly rates of concentrations of airborne stuck and concentrations of heavy metals in units (Micro / m 3 shows) for the whole city of Baghdad.

i	*Values of variables*	*airborn e stuck*	*Lead*	*Zinc*	*Copper*	*Iron*	*Nickel*	*Chromium*	*Cadmium*
1	Lower limit	628	0.3	8	0.4	28.1	.9	2	0
	Average	3140	1.55	9.3	.5	90.9	1.1	2.6	.1
	Upper limit	5454	2.8	10.5	.5	153.8	1.3	3.4	.1
2	Lower limit	479	.7	3.8	.5	19.1	.1	.8	0
	Average	2171.5	1.35	6.8	.5	44.7	.8	1.55	0
	Upper limit	3864	2	9.7	.6	70.2	1.1	2.3	0
3	Lower limit	406	.3	6.7	.5	17.9	.3	1	.1
	Average	2634.5	.5	8.8	.6	29.6	.5	1.5	.1
	Upper limit	4863	.8	10.9	.6	40.8	.8	2.1	.1
4	Lower limit	657	.3	9.9	.2	8.2	.1	.1	0
	Average	1012	.55	11.2	.4	12.4	.4	.1	0
	Upper limit	1367	.8	12.4	.6	16.5	.7	.2	0
5	Lower limit	537	.5	8.8	.2	23.2	.2	.2	0
	Average	1074.5	.5	10.7	.3	25.9	.9	.2	0
	Upper limit	1612	.5	12.7	.4	28.7	1.3	.3	0
6	Lower limit	978	.1	3.8	.3	11.9	.4	.1	0
	Average	2091.5	.2	7.2	.3	17.7	1	.9	0
	Upper limit	3205	.4	10.5	.4	23.4	1.6	1.6	0
7	Lower limit	294	1.2	10.1	1.7	87.8	.1	2.6	0
	Average	1351.5	4.1	39.2	2.2	149.4	3.4	9.5	.1
	Upper limit	2409	6.9	68.2	3.1	211	14.9	16.5	.3
8	Lower limit	404	1.9	38.7	2.3	75.8	5.7	1.2	0
	Average	717.5	5.1	5655.2	2.6	127.2	8.8	8	.1
	Upper limit	1031	8.2	71.5	3.1	178.6	12	14.8	.2
9	Lower limit	897	1.5	34.5	2	100.8	3.9	6.4	0
	Average	1215	4.1	49.7	2.5	140.4	4.5	7.1	0
	Upper limit	1533	6.6	64.8	2.9	179.9	5.2	7.7	0
10	Lower limit	519	1.2	16.5	.9	37.7	.3	2.3	0
	Average	654.5	6.1	28.5	1.1	66.1	2.5	5.8	.1
	Upper limit	790	11	39.8	1.3	94.4	4.8	9.2	.3
11	Lower limit	768	6.4	6.4	.9	84.2	2.3	6.9	0
	Average	819	9.7	9.7	1.8	111	4.4	9.4	0
	Upper limit	870	13	13	2.7	137.8	6.5	12.5	0
12	Lower limit	818	3.3	27.9	1.2	70.8	1.5	3.6	0
	Average	1003.5	7.4	36.9	1.4	89	2.6	5.8	0
	Upper limit	1614	11.5	45.9	1.7	107.2	3.6	6.3	0
13	Lower limit	328	4.6	31.2	1.8	50.5	1.5	.8	.1
	Average	2003	7.5	47.5	2.2	115	5.2	4.9	.4
	Upper limit	3678	10.4	64.3	2.7	181.4	8.8	9	.8
14	Lower limit	414	12.2	6.9	2.7	64.3	2.1	2.3	.5
	Average	503	14.5	45.3	3.4	73.5	2.5	2.9	.5
	Upper limit	592	16.8	83.7	4.1	82.7	2.8	3.4	.6
15	Lower limit	334	11.5	68.9	2.7	221	18.4	16.3	.7
	Average	503	15.4	81.7	2.7	244.4	11	20.4	.8
	Upper limit	672	19.4	94.4	2.8	267.9	30.6	24.5	.8
16	Lower limit	345	.6	2	.2	9.1	.3	2.3	0
	Average	578	3.2	4.1	.5	21.1	2.5	9.8	0
	Upper limit	811	5.8	6.2	.7	33.1	4.8	9.2	0
17	Lower limit	362	.9	4.1	.4	11.4	2.3	6.9	0
	Average	516.5	.95	5.4	1.4	31.3	4.3	9.4	.1
	Upper limit	671	1	6.7	2.3	51.2	6.5	12.5	.3
18	Lower limit	250	1.6	3.5	.2	36	1.5	3.6	0
	Average	488	3.8	6.9	2.3	40.8	2.6	5.8	0
	Upper limit	726	6	10.3	4.4	45.6	3.6	6.3	0

5.2 Structures fuzzy regression model for airborne stuck

This paragraph has been allocated for the construction of a fuzzy regression model for airborne stuck to determine the factors affecting the air quality in a fuzzy conditions (the values of the explanatory and response variables are fuzzy)

The modern man has begun to realize the importance of maintaining air quality around him when he noticed the negative effects caused by the deterioration of air quality and the change in its basic components, and the emergence of new components and the accumulation of elements in it and make Work to improve air quality of gradually return to its previous status.

Since the air quality depends on the concentrations of heavy metals in it, so the general formula for the fuzzy regression model of this study is:

$$y_i = A_0 + A_1x_1 + A_2x_2 + A_3x_3 + A_4x_4 + A_5x_5 + A_6x_6 + A_7x_7 = A_0 + \sum_{j=1}^{7} A_iX_j$$

$$y_i = \langle a_0, r_0\rangle + \sum_{j=1}^{7}\langle a_i, r_i\rangle X_j \ \ldots\ldots, i = 1,2,3,\ldots,18$$

There is more than one way to estimate the above fuzzy regression model parameters $\langle \boldsymbol{a_i}, \boldsymbol{r_i}\rangle$ as follow:

1. Tanaka approach

Transfer the fuzzy regression model to a linear programming problem as in (13)

$$\text{Min } 18r_0 + 87.4r_1 + 498r_2 + 25.5r_3 + 1344.5r_4 + 75.8r_5 + 86.7r_6 + 2.3r_7$$

Number of constraint 2n =2*18=36

Subject to

$$\alpha_0 + 1.55\alpha_1 + 9.3\alpha_2 + .5\alpha_3 + 90.9\alpha_4 + 1.1\alpha_5 + 2.6\alpha_6 + .1\alpha_7 - r_0 - 1.55r_1 - 9.3r_2 - .5r_3 - 90.9r_4 - 1.1r_5 - 2.6r_6 - .1r_7 \le 3140$$

$$\alpha_0 + 1.35\alpha_1 + 6.8\alpha_2 + .5\alpha_3 + 44.7\alpha_4 + .8\alpha_5 + 1.55\alpha_6 + 0\alpha_7 - r_0 - 1.35r_1 - 6.8r_2 - .5r_3 - 44.7r_4 - .8r_5 - 1.55r_6 - 0r_7 \le 2171.5$$

$$\alpha_0 + 3.8\alpha_1 + 6.9\alpha_2 + 2.3\alpha_3 + 40.8\alpha_4 + 2.6\alpha_5 + 5.8\alpha_6 + 0\alpha_7 + r_0 + 3.8r_1 + 6.5r_2 + 1.6r_3 + 40.6r_4 + 2.3r_5 + 5.8r_6 + 0r_7 \le 488$$

$$\alpha_0, \alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_{5l}, \alpha_6, \alpha_7, r_0, r_1, r_2, r_3, r_4, r_5, r_6, r_7 \ge 0$$

We solve the linear programming problem above by using ready program win QSB to obtained the fuzzy regression parameters which represent Center classes of estimated parameters $(\boldsymbol{\alpha_m})$and the spread of the estimated parameters $(r_m)$and the results obtained are shown in Table (2) as :

Table (2)
Values of center classes and spreadofestimated parameters.

j	α_{mj}	r_{mj}
0	1679.5	1090.4
1	-57.3	63.7
2	-8.6	8
3	-260	185
4	13.5	11.5
5	-146.5	113.5
6	-61.4	58.8
7	2240	2555

The appropriate model of the data is:

$$y_i = \langle 1679.5, 1090.4\rangle + \langle -57.3, 63.7\rangle x_1 + \langle -8.6, 8\rangle x_2 + \langle -260, 185\rangle x_3 + \langle 13.5, 11.5\rangle x_4 + \langle -146.5, 113.5\rangle x_5 + \langle -61.4, 58.8\rangle x_6 + \langle 2240, 2555\rangle x_7$$

2. a) Diamond approach [1]

We construct two regression models to the ends of the class of the response variable $\tilde{y}_i$ as in (15,16), we solve the two models above by using Ready program **spss**, then applying the formulas (19.20),Results obtained are shown in Table (3):

Table (3)
Values of center classes and spread of estimated parameters.

j	α_{mj}	r_{mj}
0	1685.6	1085.7
1	-57.5	64.1
2	-8.6	9.1
3	-263.5	183.3
4	13.4	11.3
5	-145.8	110
6	-62.3	59
7	2245	2563.7

The appropriate model of the data is:

$y_i = \langle 1685.6, 1085.7\rangle + \langle -57.7, 64.1\rangle x_1 + \langle -8.6, 9.1\rangle x_2 + \langle -263.5, 183.3\rangle x_3 + \langle 13.4, 11.3\rangle x_4 + \langle -145.8, 110\rangle x_5 + \langle -62.3, 59\rangle x_6 + \langle 2245, 2563.7\rangle x_7$

b) Least- squares approach

Using the fuzzy regression model explain in (6), we calculate the center classes which are equal to (lower limit +upper limit ⁄2) and the spared which are equal to (upper limit - lower limit ⁄2) of response variable, we have the following results shown in Table (4).

Table (4)
Center classes and spread of response variable.

I	*center classes of response variable* c_i	*Spread of response variables* s_i
1	2634.5	2228.5
2	2171.5	1692.5
3	3140	2413
4	2091.5	1113.5
5	1074.5	537.5
6	1012	355
7	1215	318
8	717.5	313.5
9	1351.5	1057.5
10	1003.5	398
11	819	51
12	654.5	135.5
13	503	169
14	503	89
15	2003	1675
16	488	238
17	516.5	154.5
18	578	233

We construct two regression models, the first one depend on the center classes of response variable values and second depend on the spared of response variable values, we solve the two models above by using Ready program spss to obtained a parameters of non-fuzzy regression model, which represent of center classes parameters estimated ($\boldsymbol{\alpha}_m$) of general regression model, and the spread of estimated parameters (r_m) the results obtained are shown in Table (5).

Table (5)
Values of center classes and spreadof estimated parameters.

j	α_{mi}	$\boldsymbol{r}_{mi}$
0	1669.1	1094.3
1	-57.1	63.8
2	-8.8	7.2
3	-257	189
4	13.4	10.9
5	-143.6	124.4
6	-122.7	97.5
7	2204.7	2867.8

The appropriate model of the data is:

$y_i = \langle 1669.1, 1094.3\rangle + \langle -57.1, 63.8\rangle x_1 + \langle -8.8, 7.2\rangle x_2 + \langle -257, 189\rangle x_3 + \langle 13.4, 10.9\rangle x_4 + \langle -143.6, 124.4\rangle x_5 + \langle 122.7, 97.5\rangle x_6 + \langle 2204.7, 2467.8\rangle x_7$

To compare the approaches of estimating general fuzzy model parameters we calculated fuzzy spread class for each approach, the results obtained are shown in Table (6).

Table (6)
Upper and lower limits and the spread of the predictions for the variable response of for fuzzy regression model by three approaches.

	approachTanaka			*approachDiamond*			*Least square Approach*		
I	*Lower limits*	*upper limits*	*Spread*	*Lower limits*	*upper limits*	*spread*	*Lower limits*	*upper limits*	*Spread*
1	663.2	4360.1	3696.9	653.4	4294.4	3641	685.7	4324.1	3638.4
2	615.2	2995.1	2379.9	604.5	2965	2360.5	632.2	2974.8	2342.6
3	549.2	3195.4	2646.2	525.8	3172.3	2646.5	566.8	3177.7	2610.9
4	567.2	2534.5	1967.3	542.2	2521.1	1978.9	588.4	2519.2	1930.8
5	584.9	2788.8	22039	566.8	2761.5	2194.7	606	2772.3	2166.3
6	563.6	2567.8	2004.2	543.3	2551.4	2008.1	579.1	2557.7	1978.6
7	554.2	2885.4	2331.2	480.7	2836.2	2355.5	595.7	2793	2197.3
8	302.2	469.6	167.4	267.2	387.7	120.5	342.8	388.1	45.3
9	512.	1828.5	1316.5	452.8	1767.4	1314.6	561.6	1731.6	1170
10	521.8	1378.7	856.9	424.3	1403	978.7	551.2	1320.6	769.4
11	563.5	1137.2	573.7	581.1	1104.7	523.6	568.9	1091.8	522.9
12	587	1495	908	533	1472.3	939.3	629.5	1422.4	792.9
13	363.7	2952.1	2588.4	348.9	2851.3	2502.4	408.4	2875	2466.6
14	300.1	1988.3	1688.2	334.1	1912.8	1578.7	341.6	1912.1	1570.5
15	258.1	2992.8	2734.7	133	2882.5	2749.5	323	2841.5	2518.5
16	472.6	1272.5	799.9	395.5	1321	925.5	473.5	1260	786.5
17	412.6	655.8	243.2	355.68	685.91	330.23	409.3	645.3	236
18	415	868.5	453.5	425.5	868.12	442.62	414.7	847.2	432.5

6- Predict airborne stuckof the city of Baghdad

Dependence on the estimated fuzzy regression model and the averages of the study observation for each explanatory variable was predicted of fuzzy airborne stuck of the city of Baghdad, by using the least squares approach the results obtained is shown in below.

Lower limits	*upper limits*	*Spread*
486.5	2081	1603.5

7- Conclusions

1. From the results obtained by applying the three approaches for the fuzzy data of the airborne stuck we get least squares approach gave less Spread, which means it's more accurate.
2. The estimation of fuzzy models by Least-squares approach is similar to estimate the parameters in the classical models known so it is computationally simple.
3. Regression model based fuzzy data shows a very useful feature is the generalization of the regression models that are based on only real data. This is because the membership function associated fuzzy sets have great value in terms of determining the accuracy of the information. The fuzzy regression models are more suitable for various real-life situations.
4. It is noted that there is an increase in the concentrations of the outstanding minutes in the air on the specific percentages of world health organization, which means increased pollution in the air of the city of Baghdad, and therefore the public health risks.

References

[1] Arabpour A.R. and Tata M., “estimating the parameters of a Fuzzy Linear Regression Model”, Iranian Journal of Fuzzy Systems, 5(2), PP.1-19, 2008.

[2] Daniel PENA., "A New Statistics for Influence in Linear Regression", Technimetrics, 47(1), 1-12, 2005.

[3] Diamond P., "Fuzzy least squares" Information Sciences, 46, 141-157, 1988.

[4] Kao C. and Chyu C., "Least-squares estimates in fuzzy regression analysis", European J. Oper. Res, 148, 426-435, 2003.
[5] Tanaka H., "Fuzzy data analysis by possibility linear models", Fuzzy Sets and Systems, 24, 363-375, 1987.
[6] Tanaka H., Uegima S. and Asai K., "Linear regression analysis with fuzzy model", IEEE Trans. Systems Man Cybrnet., 12(6), 903-907, 1982.
[7] Yang M. and Lin T., "Fuzzy least-squares linear regression analysis for fuzzy input-output data", Fuzzy Sets and Systems,126, 389-399, 2002.
[8] ZadehL. A., "Fuzzy sets", Information and Control, 8, 338–353, 1965.

الخلاصة

تحليل الانحدار من الادوات الاساسية للفحص العلمي للعلاقة الدالية بين المتغير المعتمد والمتغيرات المستقلة، ولسنوات عديدة نماذج الانحدار الخطي استخدمت على الاكثر في كل حقل من الحقول العلمية. الغرض من تحليل الانحدار هو توضيح التغير الحاصل في المتغير المعتمد بدلالة التغيرات الحاصلة في المتغيرات التوضيحية بافتراض ان البواقي هي اخطاء عشوائية. على كل حال البواقي في بعض الاحيان تعود الى عدم تحديد هيكلية الأنموذج أو عدم دقة المشاهدات عدم التأكد من هذا النوع في نماذج الانحدار اصبح يعرف بالضبابية او عدم العشوائية.

الهدف من هذا البحث هو دراسة وتطبيق طرائق تقدير معلمات نموذج الانحدار الخطي الضبابي بأستعمال بيانات ضبابية جمعت من ١٤٥ عينة فيثلاث محطات هي (ساحة الاندلس، الجادرية، العلاوي) في بغداد على انواع نماذج الانحدار الضبابية من خلال تطبيق عملي على التلوث البيئي لهواء مدينة بغداد, حيث تم جمع ديتة بغداد كل يوم لغرض قياس تراكيز الدقائق العالقة في الهواء في والذي يمثل متغير الاستجابة كذلك اهم ملوثات الهواء المعروفة (الرصاص، الخارصين، النحاس، الحديد، النيكل، الكروم، الكادميوم) كمتغيرات التوضحية، النتيجة الرئيسية للبحث تحديد افضل اسلوب لتقدير معلمات الانحدار الخطي الضبابي لهذه البيانات وحساب القيم التنبؤية للدقائق العالقة في الهواء لمدينة بغداد للسنوات القادمة.