AL KUT JOURNAL OF ECONOMIC AND ADMINISTRATIVE SCIENCES
Publisher: College of Economics and Management - Wasit University

Estimating the Nonparametric Regression Function of the Fuzzy Phenomena by using Simulink

Zainb Hassan Radhy(1) **Mahdi Wahhab Neamah** (2)

zainb.h@s.uokerbala.edu.iq Mehdi.wahab@uokerbala.ude.iq

Karbala University,

Abstract

Statistical data is sometimes obtained from uncertain resources or fuzzy phenomenon therefore the conventional statistical analysis becomes unable to interpret the result of these data. And addition it is difficult to find function form or probability distribution for this kind of data So, must be using appropriate analysis model achieved assumption fuzzy data or phenomenon.

Concern has been focused on utilizing the fuzzy nonparametric regression models, which are convenient to deal with this data, in this paper presents a compare between to smoothing approaches to estimating the fuzzy nonparametric regression function by using crisp independent and fuzzy dependent variables within uncertain phenomena. A triangular membership function was adopted to generate the belonging of the elements within the fuzzy set. where applied the local linear smoothing and kernel smoothing, suggested two test functions were proposed to show the applied methods' The results of MATLAB simulations and the applied criteria of differentiating have shown the superiority of the local linear smoothing over kernel smoothing for the two proposed test functions. So we goal to modeling the fuzziness phenomena

Keyword: Fuzzy data, fuzzy nonparametric regression model, local linear smooth, kernel smooth, fuzzy Nadarya-Watson.

1. Introduction

Smoothing is a statistical method to estimate a function with real value through observation when not a variable model to represent this function, it is an approach to determine the effect independent variable on a dependent variable where we aim for an approximation nonparametric regression estimation function to the nonparametric regression real function.

Although many researchers studied fuzzy parametric regression, the resultant models were, sometimes, inefficient. Therefore, fuzzy nonparametric models have been proposed as an alternative solution to overcome this problem. This is because the nonparametric analysis is more reliable in finding the functional relationships among the variables. Moreover, it is not always to be able to measure the variables of a certain study with respect to specific values but in form of periods.

The fuzzy approach was firstly introduced by L.A.Zadeh [1] before it was improved to become a suitable method to deal with uncertain or fuzzy data, The fuzzy regression models of crisp set inputs and fuzzy outputs were under research by Ning Wang [2]. The research was based on the measure of Diamond distance which distance measure between two fuzzy numbers.

1.1. The Fuzzy and crisp Set

The Fuzzy set is set in all elements have a degree of membership defined as follows: X represents a set of elements denoting $\mathcal{X}$ then the fuzzy set $\widetilde{\mathfrak{M}}$ in X it is a set of order parts: [3]

$\widetilde{\mathfrak{m}} = \{\left(\mathcal{X}, \mathfrak{m}_{\widetilde{\mathfrak{m}}}(x) \,/\, \mathcal{X} \in X\right)\}$ where: $\mathfrak{M}_{\widetilde{\mathfrak{m}}}(x)$ represents the membership function to $\mathcal{X}$ in $\widetilde{\mathfrak{M}}$ In Crisp set or classical set collection of observation have a degree of belonging to the period [0,1] or not, there are many types of membership function Like triangular, trapezoidal, Gaussian were represented in this paper using the triangular membership function as the following[4]:

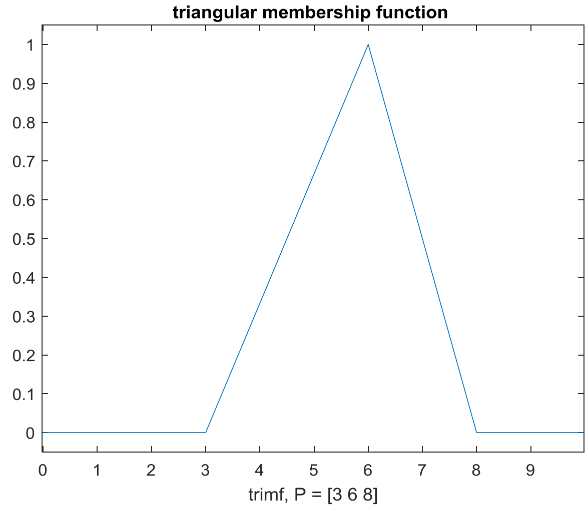

Figure 1- Represent triangular memberships function curve

$$\mathfrak{M}_k(x) = \begin{bmatrix} \mathcal{L}\dfrac{m_k - \mathcal{X}}{m_k - \ell_k} & if\ \ell_k \leq \mathcal{X} \leq m_k \\ \mathcal{R}\dfrac{\mathcal{X} - m_k}{u_k - m_k} & if\ m_k \leq \mathcal{X} \leq u_k \\ 0 & otherwise \end{bmatrix} \quad \dots\dots \ (1)$$

Where: $\tilde{k} = (\ell_k(x), m_k(x), u_k(x))$ represent fuzzy number from a type of (LR), $\mathcal{L}(.), \mathcal{R}(.)$ continuous in period [0, 1], $\mathcal{L}(1) = \mathcal{R}(1) = 0$, $\mathcal{L}(0) = \mathcal{R}(0) = 1$.

1.2. Define univariate fuzzy nonparametric regression model [2][5][6][7]

The fuzzy parametric regression model is prosaic because requiring some assumption, it is often indeterminate for this utilize the fuzzy nonparametric regression model with univariate crisp input and LR fuzzy output we describe this model by:

$$\mathcal{Y} = g(\mathcal{X}) + \varepsilon \dots\dots..(2)$$

$$g(\mathcal{X}) = \{(\ell_k(x), m_k(x), u_k(x)\} + \varepsilon$$

Where $(\mathcal{X})$ is crisp independent variables (input on domain R), $\mathcal{Y}$ is an (LR) fuzzy dependent Variable (output), $g(\mathcal{X})$ unknown fuzzy regression function and it is field from(D) to $\mathcal{R}_{LR}$it is limits are minimum, upper, and middle on respectively (ℓ_k, u_k, m_k), (ε) is a random error term.

2. Methodology

2.1. Local linear regression smooth

It is also called local linear polynomial and used to solve the problem of bandwidth effect on bias and variance because it is adaptive with point location [8]

Let $(X_1, Y_1) \ldots\ldots (X_n, Y_n)$ random sample with binary variable to estimate nonparametric regression function $g(x) = E[Y|X = x]$ by using Taylor series approximating $g(x)$ when X going to x_0

$$g(x) \approx g(x_0) + g^1(x_0)(X - x_0) + \frac{g^2(x_0)}{2!}(X - x_0)^2 \ldots\ldots \frac{g^P(x_0)}{P!}(X - x_0)^P$$

$$= g(x_0) + \beta_1(X - x_0) + \beta_2(X - x_0)^2 + \cdots + \beta_p(X - x_0)^p \quad \ldots\ldots (3)$$

g(x) represents the unknown function which we want to estimate at the point $X = x_0$, $g^i(x_0)$ represents derivative value from the rank(i)At the point x_0 where :

$(i = 1,2, \ldots . p)$, $(\beta_0, \beta_1, \ldots, \beta_p)$ represent parametric function under assumption Availability of all derivatives

We can find parametric estimation $(\beta_1, \beta_2, \ldots, \beta_p)$ by the weight least squared method upon kernel function $k(\frac{x-x_i}{h})$ by minimum The sum of the squares of errors' ratio to parametric $(\beta_1, \beta_2, \ldots, \beta_p)$ as the following:

$$Q = \sum_{i=1}^{n} \left\{Y_i - \beta_0 - \beta_1(X_i - x_0) - \cdots \beta_p(X_i - x_0)^p\right\}^2 k\left(\frac{x_i - x_0}{h}\right) \ldots . \quad (4)$$

We can rewrite the equation (4) by using the matrix as the following:

$$Q = (Y - X\beta)^T W (Y - X\beta) \quad \ldots\ldots\ldots \qquad (5)$$

Where:

$$X = \begin{bmatrix} 1 & (X_1 - x) & \ldots\ldots.. & (X_1 - x)^p \\ 1 & (X_2 - x) & \ldots\ldots.. & (X_2 - x)^p \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ 1 & (X_n - x) & \ldots & (X_n - x)^p \end{bmatrix}$$

$$y = (y_1, y_2, y_3, \ldots\ldots . y_n)^T \quad , \quad \beta = (\beta_1, \beta_2, \beta_3, \ldots\ldots . \beta_n)^T$$

As the Kernel function is sympatric, we can express the weight matrix as the following:

$$W = \begin{bmatrix} k(\frac{x_1 - x_0}{h}) & 0 & 0 & 0 \\ 0 & k(\frac{x_2 - x_0}{h}) & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & k(\frac{x_n - x_0}{h}) \end{bmatrix}$$

Using the matrix and simplification operation we get:

$$Q = Y^T WY - Y^T WX\beta - \beta^T X^T WY + \beta^T X^T WX\beta$$
$$= Y^T WY - 2\beta^T X^T WY + \beta^T X^T WX\beta \quad \ldots \quad (6)$$

And finding the derivation concerning β:

$$\frac{\partial \text{Q}}{\partial \beta} = -2\text{X}^\text{T}\text{WY} + 2\text{X}^\text{T}\text{WX}\beta^\text{T} \quad \ldots.. \qquad (7)$$

$\frac{\partial \text{Q}}{\partial \beta} = 0$ then the solution $\beta = (\beta_1, \beta_2, \beta_3, \ldots \ldots . \beta_n)^T$is weight least square estimation is the following:

$$\beta^{\Lambda} = (\text{X}^\text{T}\text{WX})^{-1} X^T WY \quad \ldots \ldots \qquad (8)$$

From defending the Tyler series, we can write function estimation $g(x)$ at the point x_0:

$$g(x)^{\Lambda} = \beta_0^{\Lambda} = e_1^T (\text{X}^\text{T}\text{WX})^{-1} X^T WY \quad \ldots \ldots \quad (9)$$

Where e_1^T unit vector with dimensions $(p+1) * 1$

Since we treatment with the Fuzziness theorem will be upon the univariate fuzzy model in equation (2) depend on the above theory to estimate g (x) Where:

$g(x) = (\ell(x), m(x), u(x))$

At $x \in D$ to all (x_i, y_i) where represent crisp set independent variable (input variable) and dependent variable from type LR (output) to (n) observations. the aim of fuzzy number estimation is close to fuzzy number observation we want to get beast fitting(match)

So, we must use a distance to measure the proximate between two fuzzy numbers (estimated and observed) so we use diamond distance to match [4][2][6]

Let C, D represent any fuzzy number as the following:

C= (m_C, a_c, β_c), D= (m_d, a_d, β_d), and there are of the type (LR) (that have one center) defined diamond fuzzy number (DFN) and the membership function his:

where diamond fuzzy as the following: $\mathcal{A}_{\tilde{D}} = \{\mathfrak{a}, \mathfrak{b}, \mathfrak{c}(\mathfrak{a}_\mathfrak{b}, \mathfrak{B}_\mathfrak{b}),$ the membership function is:

$$\mathfrak{m}_{\tilde{\mathcal{A}}_D}(x) = \left\{ \begin{array}{cc} 0 & x < \mathfrak{a} \\ \frac{x-\mathfrak{a}}{\mathfrak{b}-\mathfrak{a}} & \mathfrak{a} \le x \le \mathfrak{b} \\ \frac{x-\mathfrak{a}}{\mathfrak{b}-\mathfrak{a}} & \mathfrak{b} \le x \le \mathfrak{c} \\ \mathfrak{a}_{\mathfrak{b}} & base \\ \frac{\mathfrak{a}-x}{\mathfrak{a}-\mathfrak{b}} & \mathfrak{a} \le x \le \mathfrak{b} \\ \frac{x-\mathfrak{a}}{\mathfrak{b}-\mathfrak{c}} & \mathfrak{b} \le x \le \mathfrak{c} \\ 1 & x = \mathfrak{B}_{\mathfrak{b}} \\ 0 & etherwise \end{array} \right\} \quad \ldots\ldots\ldots. \quad (10)$$

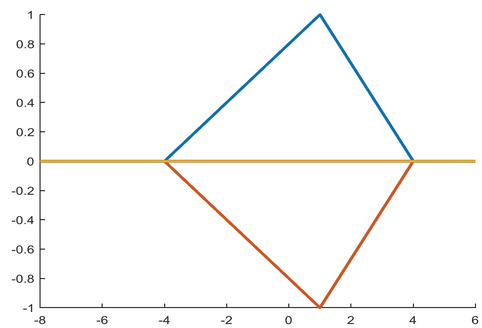

Figure 2- Graphical representation of the diamond fuzzy number

Upon on diamond distance, we can measure the convergence between two numbers fuzzy where the membership function is equal when $d^2(C, D) = 0$

$$d^2(C,D) = (\ell_c - \ell_d)^2 + (m_c - m_d)^2 + (u_c - u_d)^2 \quad \ldots \quad (11)$$

Where:

$C = (\ell_C, m_c, u_c)\ D = (\ell_d, m_d, u_d) \quad \ell_d, m_d, u_d \ge 0$

We can write the upper, lower, and mid of two fuzzy numbers as the following:

$$\ell_C = m_c - u_c \ , \ \ell_d = m_d - u_d$$

$$u_C = \ell_c + m_c \ , \ u_d = \ell_d + m_d$$

As we mentioned above the local linear is based on extended Taylor so will approximate the regression function formed from upper, lower and mid (ℓ_d, m_d, u_d). the function must have continuous derivatives in its domain so will be approximate function locally by linear functions below

$$\ell(x) \approx \tilde{\ell}(x) = \ell(x_0) + \tilde{\ell}(x_0)(x - x_0) \quad \dots \dots \quad (12)$$

$$m(x) \approx \widetilde{m}(x) = m(x_0) + \widetilde{m}(x_0)(x - x_0) \quad \dots . . \quad (13)$$

$$u(x) \approx \tilde{u}(x) = u(x_0) + \tilde{u}(x_0)(x - x_0) \quad \dots \dots \quad (14)$$

The linear functions above have limited regression function derivatives at the point(x_0),To clarify the theoretical idea let the fuzzy depend variable as the following:

$Y_i = (\ell_{yi}, m_{yi}, u_{yi})_{LR}, \quad i = 1,2, . . n$ and sample observation are crisp independent variable: $(X_i, Y_i) = ((x_1, x_2, \dots . . , x_n, ((\ell_{yi}, m_{yi}, u_{yi})_{LR})$ to estimation regression function at the point x_0 meaning:$g(x_0) = (\ell(x), m(x), u(x))$.

By using local linear smoothing and diamond distance in equation (11) will be minimize the locally weighted least squares below:[9][2]

$$min \sum_{i=1}^{n} d^2(((\ell_{yi}, m_{yi}, u_{yi})_{LR}, (\tilde{\ell}_{yi}, \widetilde{m}_{yi}, \tilde{u}_{yi})_{LR} k_h(|x_i - x_0|) \quad \dots \dots \ (15)$$

To applied the equations (12,13,14) in equation (15) we will get:

$min \ \sum_{i=1}^{n}((\ell_{yi} - \ell(x_0) - \tilde{\ell}(x_0)\,(x - x_0))^2 k_h(|x_i - x_0|) + \sum_{i=1}^{n}((m_{yi} - m(x_0) - \widetilde{m}(x_0)\,(x - x_0))^2 k_h(|x_i - x_0|) + \sum_{i=1}^{n}((u_{yi} - u(x_0) - \tilde{u}(x_0)\,(x - x_0))^2 k_h(|x_i - x_0|)$ $\dots \dots \dots \dots \dots$ (16)

Where $\{k_h(|x_i - x_0|) = \frac{k\frac{(x_i - x_0)}{h}}{h} \quad i = 1,2, . . , n$ is kernel function, it is a series of weights at the point $x_0$working as the control in process smooth with bandwidth parametric.to estimation fuzzy regression function $g(x)$ at the point (x_0)

$$g^{\Lambda}(x) = (\ell^{\Lambda}{}_{yi}, m^{\Lambda}{}_{yi}, u^{\Lambda}{}_{yi})_{LR}$$

$$= (m^{\Lambda}(x_0) - \alpha^{\Lambda}(x_0), m^{\Lambda}(x_0), m^{\Lambda}(x_0) + \beta^{\Lambda}(x_0)_{LR}$$

the equation (16) contains to add three parts and each part contain different unknown parametric. we will take partial derivatives to equation (16) respect to unknown parametric and then equal to zero to get three gropes from linear equations:

$$(\ell(x_0), \tilde{\ell}(x_0)), (m(x_0), \widetilde{m}(x_0)), (u(x_0), \tilde{u}(x_0))$$

By using matric based weight least square we get:[4][10]

$$\ell^{\Lambda}(x_0), \ell^{\Lambda'}(x_0) = (X^T(x_0)W(x_0; h))^{-1} X^T(x_0)W(x_0; h)L_y$$

$$m^{\Lambda}(x_0), m^{\Lambda'}(x_0) = (X^T(x_0)W(x_0; h))^{-1} X^T(x_0)W(x_0; h)M_y$$

$$u^{\Lambda}(x_0), u^{\Lambda'}(x_0) = \left(X^T(x_0)W(x_0;h)\right)^{-1}X^T(x_0)W(x_0;h)U_y$$

Where:

$$X(x_0) = \begin{bmatrix} 1 & x_1 - x_0 \\ 1 & x_2 - x_0 \\ 1 & x_3 - x_0 \\ . & . \\ . & . \\ 1 & x_n - x_0 \end{bmatrix}, L_y = \begin{bmatrix} \ell_{y1} \\ \ell_{y2} \\ \ell_{y3} \\ \vdots \\ \ell_{yn} \end{bmatrix}, M_y = \begin{bmatrix} m_{y1} \\ m_{y2} \\ m_{y3} \\ \vdots \\ m_{yn} \end{bmatrix}, U_y = \begin{bmatrix} u_{y1} \\ u_{y2} \\ u_{y3} \\ \vdots \\ u_{yn} \end{bmatrix}$$

The weight matric is:

$$W(x_0;h) = Diag(k_h(|x_1 - x_0|), k_h(|x_2 - x_0|), \dots k_h(|x_n - x_0|)$$

Where the diagonal matric elements will be equal to: $k_h(|x_i - x_0|) \quad i = 1,2,..n$

Then the regression function estimation as the following:

$$g^{\Lambda}(x) = e_1^T H(x_0;h)L_y, e_1^T H(x_0;h)M_y, e_1^T H(x_0;h)U_y \qquad \dots\dots \quad (17)$$

Where:

e_1^T *is once vector and* $H(x_0;h) = (X^T(x_0)W(x_0;h)^{-1}X^T(x_0)W(x_0,h) \quad \dots\dots \quad (18)$

the local linear smooth is status case from the local polynomials (which represent the weight regression about the point (x_0))when the degree of polynomial (d=1) the degree polynomial with bandwidth parametric will be working as the control smoothing in estimation). As the fuzzy dependent variable to type LR then we can write estimation $g(x)$as follow:

$$g^{\Lambda}(x) = (\ell^{\Lambda}(x_0), m^{\Lambda}(x_0), u^{\Lambda}(x_0))_{LR} \;, g^{\Lambda}(x) = \left(\frac{\sum_{i=1}^n w_i \ell_{yi}}{\sum_{i=1}^n w_i}, \frac{\sum_{i=1}^n w_i m_{yi}}{\sum_{i=1}^n w_i}, \frac{\sum_{i=1}^n w_i u_{yi}}{\sum_{i=1}^n w_i}\right)$$

We can define the weight:

$$w_i = k\frac{x - x_i}{h}\left[Z_{n,2} - (x - x_i)Z_{n,1}\right], Z_{n,j} = \sum_{i=1}^{n}\left(k\frac{x - x_i}{h}\right)(x - x_i)^j$$

In this paper the fuzzy détente variable (Y_i) Symmetrix triangular and $x \in [0,1]$

As the form $Y_i = \left(m_{yi}, \sigma_{yi}\right)$, $\sigma_{yi} = m_{yi} - \ell_{yi} = u_{yi} - m_{yi}$

Which σ_{yi} represent diffusion to Y_i,where diffusion vector to (n) observation from fuzzy dependent variable as follow:$\sigma = (\sigma_{y1}, \sigma_{y2}, \dots \sigma_{yi})^T = M_y - L_y = U_y - M_y$

From the equation (17) we get:$m^{\Lambda}(x_0) - \ell^{\Lambda}(x_0) = u^{\Lambda}(x_0) - m^{\Lambda}(x_0) = e_1^T H(x_0;h)\sigma$

Therefore, the regression function at the point x_0 will be symmetric fuzzy number from (LR) and we can express as follow

$$g^{\Lambda}(x) = e_1^T H(x_0;h)M_y, e_1^T H(x_0;h)\sigma)_L \qquad \dots.. \quad (19)$$

By Local Linear to Find smooth at each value of data, by A MATLAB m-file, version 2019b, was designated to apply the flowchart of Figure (3)

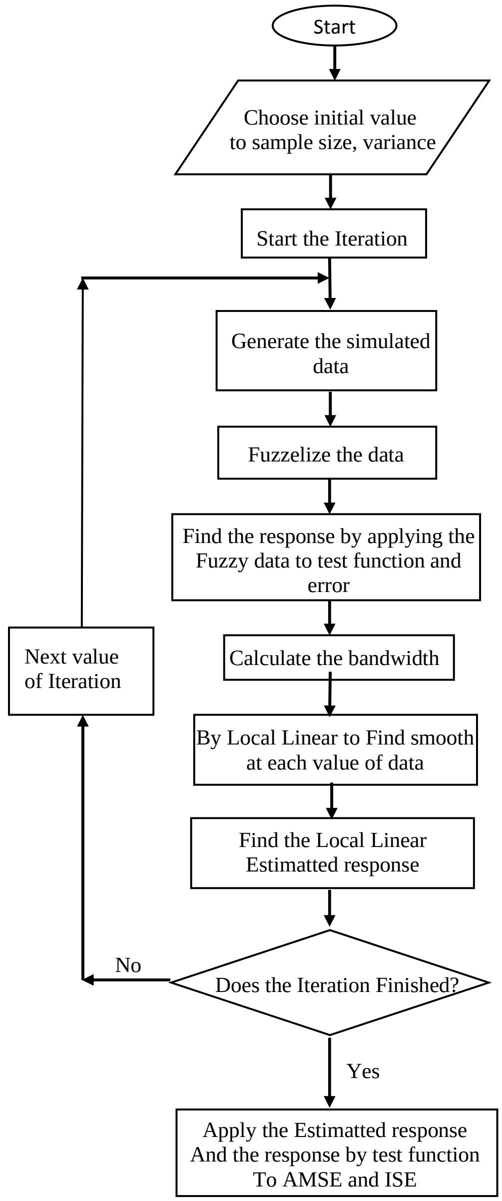

Figure 3- Flowchart to highlight steps of executing the local-linear smooth method

2.2. Kernel smoothing [11][12][13] [14][15]

Kernel smooth is a statistic approach, it is the special case from a local average approach which is based on the weight function that represents kernel function, it is a probability function because it is satisfied conditions probability theory of its characteristic's symmetries, continues, limited and real function.

the weight function is important because it is wakened to modification size and formula the weight according to the point location respectively the point estimation x, meaning it gives large weight to the points closed from x upon the bandwidth and distance between the observation and the point x. As the following some kernel function Where: $u = (x - x_i)/h \quad |u| \leq 1$, h=bandwidth parametric.

Table 1- Represent some kernel function formula

kernel	equation
normal	$k(u) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{u^2}{2}\right) \quad u < \infty$
Epanechnikov	$k(u) = \frac{3}{4}(1 - u^2)$
box	$k(u) = \frac{1}{2} \quad u \leq 1$
Triangular	$k(u) = (\ 1 - \lvert u \rvert) \quad u \leq 1$

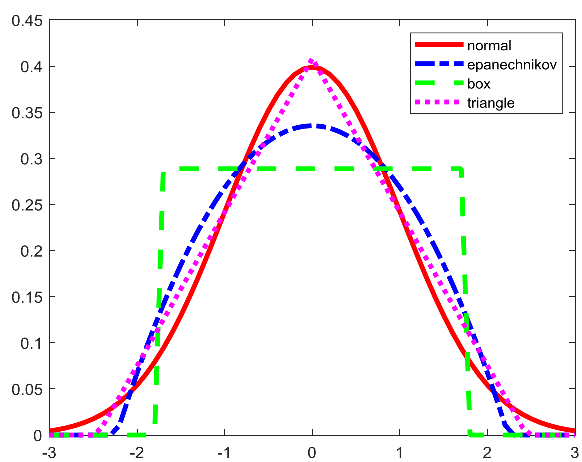

Figure 4- represent some kernel functions curve

We must choose the bandwidth parametric to be Careful because it is based on parts in converging nonparametric regression function to the main function, the small value to bandwidth effect in the curve estimation then formation low curve estimate (under smoothing curve) but if we choose a large value to bandwidth, it gives high curve estimate (over smoothing).

There are two methods for selecting bandwidth, first method used change bandwidth Upended location estimation called this method adaptive bandwidth which adopted in this paper, the second method used fixed bandwidth.

2.3. Fuzzy Nadarya- Watson [13][7][10]

Nadarya-Watson was first suggested kernel smooth (1964), it is a nonparametric statistical approach used to estimate function with real value through a sample of observation when not a variable parametric model for this function. The kernel is smooth as approximate to the curve regression. Nadarya-watson is a special case from local polynomial when $P = 0$ (in equation (3).

In the kernel method arrange the fixed design point $x_1, x_2, \dots, x_n$ in progressive order then calculate the distance between two points x_i و x_{i+1} such as the distance will be fixed and equal to all value from i,we can get polynomial estimation $g^{\wedge}(x)$ through the polynomial estimation $g(x)$ as the following:

$$g(x) = \mathcal{B}_0 + \mathcal{B}_1(.-x) + \cdots \mathcal{B}_p(.-x)^p \qquad \dots\dots\dots\dots \qquad (20)$$

Where: p degree of the polynomial at the point x,by using weight least square with kernel function as the following

$$min \sum_{i=1}^{n} \{y_i - m(x)\}^2 k(\frac{x - x_i}{h}) \qquad \dots\dots\dots \qquad (21)$$

H represents bandwidth parametric which locates the value or the local Neighbour size, when the polynomial degree $p = 0$ then $g^{\wedge}(x)$ represents $\mathcal{B}_0$ (fixed limit) and called (NW) estimation. We can find the least squared estimation by minimum the criterion as the following:

$$Q = \sum_{i=1}^{n} (y_i - \mathcal{B}_0)^2 k(\frac{x - x_i}{h}) \qquad \dots\dots. \qquad (22)$$

Then found derivation to according to $\mathcal{B}_0$ and equal to zero:

$$\sum_{i=1}^{n} \left(y_i - \mathcal{B}_0{}^{\wedge}\right)^2 k(\frac{x - x_i}{h}) = 0 \qquad \dots\dots \qquad (23)$$

$$\mathcal{B}_0{}^{\Lambda} = \frac{k(\frac{x-x_i}{h})y_i}{k(\frac{x-x_i}{h})} \quad \dots\dots\dots\dots\dots \quad (24)$$

The equation (24) represents (NW) smooth. If we used any function from kernel functions in a table (1) were satisfied assumptions theory We can write the formula (NW) estimation as follow:

$$g^{\wedge}(x) = \sum_{i=1}^{n} y_i \left[\frac{k(\frac{x-x_i}{h})y_i}{k(\frac{x-x_i}{h})}\right] = \sum_{i=1}^{n} y_i\, w_i \quad \dots\dots \quad (25)$$

In fuzzy nonparametric regression to type (LR) the estimation (x) at the point x_0 as follows:[2][5][16]

$g^{\wedge}(x) = (\, \mathcal{L}^{\Lambda}(x), \mathcal{M}^{\Lambda}(x), \mathcal{U}^{\Lambda}(x)_{LR}$

$g^{\wedge}(x)) = (\frac{\sum_{i=1}^{n} w_i \ell_{yi}}{\sum_{i=1}^{n} w_i}, \frac{\sum_{i=1}^{n} w_i m_{yi}}{\sum_{i=1}^{n} w_i}, \frac{\sum_{i=1}^{n} w_i \mathcal{U}_{yi}}{\sum_{i=1}^{n} w_i})_{LR} \quad \dots\dots. \quad (27)$

Meaning the three limits to regression function $g(x)$ will converge locally to the point neighbor by fixed unknown limited to regression function $(\mathcal{U}^{\Lambda}(x)\mathcal{M}^{\Lambda}(x)\mathcal{L}^{\Lambda}(x))$

2.4. Integral Square Error criterion (ISE) [10]

One of the important statistical criteria, is used to compare nonparametric estimators, for the difficult computation of these integral we use the formula mathematical approximate where the period of the variable x is divided into many subperiods relative to statistics as follow:

$ISE = \frac{1}{400}\sum_{i=1}^{400}\left(g(x_i) - g(x_i)\right)^2 \quad \dots\dots (28)$

Average Mean Square Error criterion (AMSE):[10]

The statistical measure used in most studies and research; the formula is:

$MASE = \frac{1}{n}\sum_{i=1}^{n}(\left(g(x_i) - g^{\Lambda}(x_i)\right)^2 \dots\dots (29)$

3. Implement by Simulink

For the application of the methods and comparing between them, we used Simulation help us to implement theorem, methods, choose different sizes and variances to add select an alternative model to the real model understudies we will build a fuzzy nonparametric regression model by MATLAB (2019) programing as follow:

generating randomly variable (x) crisp set, $x \sim u(0,1)$ and generate sizes samples (N=50,100,200), variance (SD=0.2,0.4,0.6), generating randomly error vector normally distribution $(\varepsilon \sim N(0,\sigma^2))$

generate the fuzzy depended variable (y) from type (LR) with triangular membership function by using cod in MATLAB

$$y = \max\left(\min(\frac{(x - X + 0.1)}{0.1}, \frac{(X + 0.1 - x)}{0.1}, 0\right)$$

Suggest two test functions to explained the thermotical part

1- [$y = x + 4\exp(.84x^3)$]
2- [$g(x) = 1 - 48x + 18x^2 - 15x^3 + 45x^4 - x^5$]

upon two criteria from comparation, ISE, AMSE

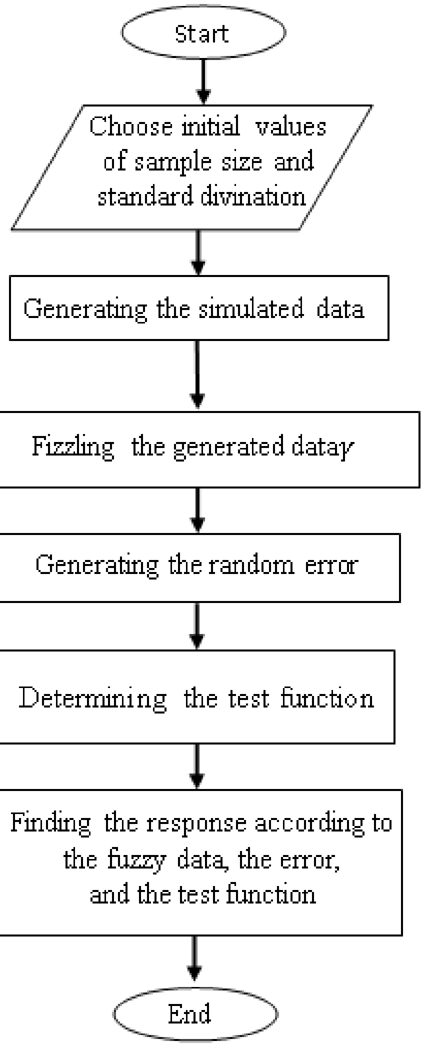

Figure 5- Flowchart to highlight steps of executing data fuzzing

4. Results

Below result where the table (1) represents the first test function suggest to apply the methods and some figures of different size and variance $y = x + 4\exp(0.84x^3)$

Suggest test function 1: $\boldsymbol{g(x) = x + 4\exp(0.84x^3)}$							
Method	Sample size	σ = 0.2		σ = 0.4		σ = 0.6	
		AMSE	ISE	AMSE	ISE	AMSE	ISE
Local-linear smooth	50	0.1024	0.0223	0.204	0.0892	0.3072	0.2008
	100	0.0685	0.0052	0.137	0.0207	0.2055	0.0465
	200	0.0412	0.00	0.082	0.0261	0.1236	0.0588
Nadarya-Watson smooth	50	0.1551	0.0449	0.3103	0.1797	0.4654	0.4004
	100	0.1579	0.0287	0.3157	0.1148	0.4736	0.2584
	200	0.1574	0.04	0.314	0.1663	0.4722	0.3742

Table 2- Represents the result the exponential function by two methods with two criteria (ISE, AMSE) for all different sizes and variances

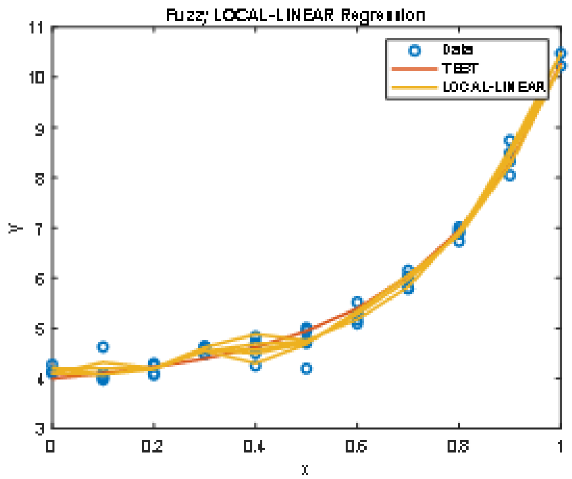

Figure 6- Represent (L-L-S) smooth at (SD = 0.4, n=100)

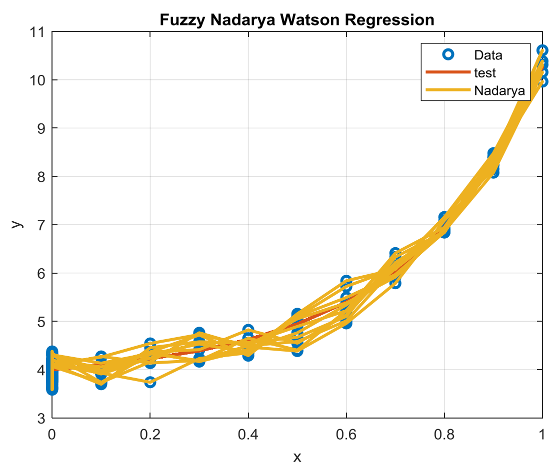

Figure 7- Represent (NW) smooth at (SD= 0.2, N=200)

From above result and graphical superiority local- linear method upon values of criterion to all sample sizes and variance and figures to the test function.

Below result where the table (2) represents the first test function suggest to apply the methods and some figures of different size and variance $\boldsymbol{g(x) = 1 - 48x + 18x^2 - 15x^3 + 45x^4 - x^5}$

Suggest test function 1: $\boldsymbol{g(x) = 1 - 48x + 18x^2 - 15x^3 + 45x^4 - x^5}$							
Methods	Sample size	σ = 0.2		σ = 0.4		σ = 0.6	
		AMSE	ISE	AMSE	ISE	AMSE	ISE
Local-linear smooth	50	0.1559	0.0669	0.4101	0.1878	0.3450	0.504
	100	0.1544	0.0657	0.3220	0.1569	0.4879	0.485
	200	0.1491	0.0644	0.3165	0.1343	0.4778	0.398
Nadarya-Watson smooth	50	0.0074	0.0234	0.2099	0.0789	0.2992	0.100
	100	0.00697	0.0224	0.1999	0.0627	0.2356	0.054
	200	0.0047	0.0212	0.1675	0.0619	0.1987	0.059

Table 3- Represent the result of the polynomial function by two methods with two criterions (ISE, AMSE) to all different size and variances

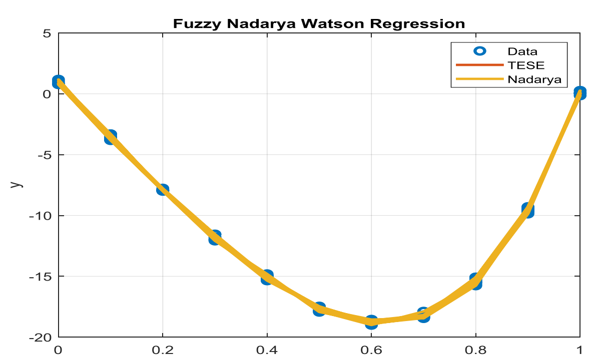

Figure 8- Represent (NW) smooth to (SD=0.2, N=100)

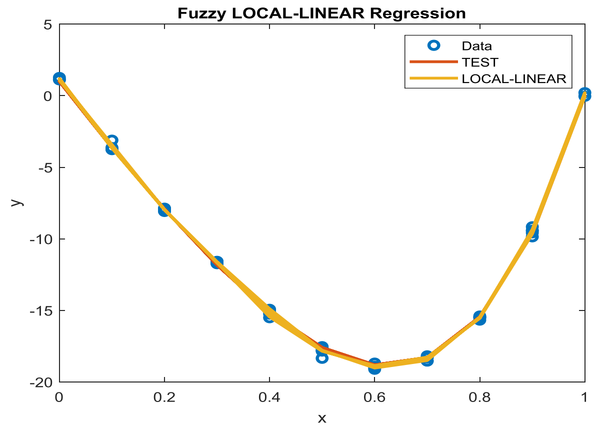

Figure 9- Represent (L-L-S) smooth at (SD=0.4, N=50)

5. Conclusion

We conclude by the methods are used in the simulating, the best local-liner smooth method because it has attractive bias properties and has the best convergence rate. to estimate fuzzy nonparametric regression of both the test functions upon value to the criterion and figures which explain convergent the test function to real function, the better is decreased when increasing degree of the polynomial and the good select bandwidth parametric as the choose optimal value to bandwidth the balance between bias and variance can be verified.

6. Reference

[1] L.A.Zadeh, “Fuzzy Sets,” *Inf. Control*, vol. 353, pp. 338–353, 1965.

[2] Ning Wang, “Fuzzy nonparametric regression based on local linear smoothing technique,” *Inf. Sci. (Ny).*, vol. 177, no. 18, pp. 3882–3900, 2007.

[3] Diamond, “Fuzzy least squares",” 1988.

[4] M. YILDIZ and M. MEMMEDLİ, “Comparative Analysis for Fuzzy Nonparametric Regression Models,” *Eskişehir Tech. Univ. J. Sci. Technol. A - Appl. Sci. Eng.*, vol. 22, no. 4, pp. 353–365, 2021, doi: 10.18038/estubtda.1033350.

[5] Jin Hee Yoon, “Fuzzy Rank Linear Regression Model,” 2009.

[6] H. Y. Jung, W. J. Lee, and S. H. Choi, “Fuzzy regression model using fuzzy partition,” *J. Phys. Conf. Ser.*, vol. 1334, no. 1, 2019, doi: 10.1088/1742-6596/1334/1/012019.

[7] J. M. Liu, R. Chen, and Q. Yao, “Nonparametric Regression with Trapezoidal Fuzzy Data,” *J. Econom.*, vol. 157, no. 1, pp. 151–164, 2010, doi: 10.1016/j.jeconom.2009.10.029.

[8] W. Xu, “Adaptive fuzzy kernel clustreing algorithm,” vol. 5, no. 4, pp. 51–58.

[9] J. Fan and I. Gijbels, “Local Linear Smoothers in Regression Function Estimation,” no. November, pp. 1–27, 2014.

[10] J. E. Chacón, “Data-driven choice of the smoothing parametrization for kernel density estimators,” *Can. J. Stat.*, vol. 37, no. 2, pp. 249–265, 2009, doi: 10.1002/cjs.10016.

[11] S. Demir and Ö. Toktamiş, “On the adaptive Nadaraya-Watson kernel regression estimators,” *Hacettepe J. Math. Stat.*, vol. 39, no. 3, pp. 429–437, 2010.

[12] C. B. Cheng and E. S. Lee, “Nonparametric fuzzy regression - k-NN and kernel smoothing techniques,” *Comput. Math. with Appl.*, vol. 38, no. 3, pp. 239–251, 1999, doi: 10.1016/S0898-1221(99)00198-4.

[13] K. De Brabanter, J. De Brabanter, J. A. K. Suykens, and B. De Moor, “Optimized fixed-size kernel models for large data sets,” *Comput. Stat. Data Anal.*, vol. 54, no. 6, pp. 1484–1504, 2010, doi: 10.1016/j.csda.2010.01.024.

[14] W. Pedrycz, J. C. Bezdek, R. J. Hathaway, and G. W. Rogers, “Two nonparametric models for fusing heterogeneous fuzzy data,” *IEEE Trans. Fuzzy Syst.*, vol. 6, no. 3, pp. 411–425, 1998, doi: 10.1109/91.705509.

[15] V. C. Raykar and R. Duraiswami, “Very fast optimal bandwidth selection for univariate kernel density estimation,” *Tech. Reports from UMIACS UMIACS-TR-2005-73*, pp. 522–526, 2006.

[16] M. J. Brewer, “A Bayesian model for local smoothing in kernel density estimation,” *Stat. Comput.*, 2000.