On the Comparison of Fuzzy Kernel Regression Estimator and Fuzzy Radial Basis Function Networks

Nimet YAPICI PEHLİVAN[1♠], Ayşen APAYDIN[2]

[1]*Selçuk University, Faculty of Science and Art, Department of Statistics, 42031, Konya, Turkey*
[2] *Ankara University, Faculty of Science, Department of Statistics, 06100, Ankara, Turkey*

ABSTRACT

In this paper, we suggest two fuzzy estimators in nonparametric regression: fuzzy kernel regression (FNPR) estimator and fuzzy radial basis function (FRBF) networks. Both FNPR estimator and FRBF networks are applied to original data taken from an experiment. We obtain MSE values of the FNPR estimator and FRBF networks and then compare them. We show that the FNPR estimator is more efficient than the FRBF networks.

Key Words: Fuzzy number, Fuzzy kernel regression estimator, Nonparametric regression, Neural networks, Fuzzy radial basis function networks.

1. INTRODUCTION

Regression analysis is one of the widely used statistical methods in research. In the least squares method, which is the most widely used one for estimating a regression model's parameters, the assumptions of "normality" and "homoscedasticity" on error terms must be satisfied. When these assumptions are not satisfied, it is appropriate to use nonparametric regression (NPR) techniques [8]. Various methods have been used in NPR, such as kernel regression estimators, k-nearest neighbourhood estimators, spline smoothing, median smoothing, regressograms, etc. Kernel regression estimators are the most widely used estimators in NPR [11].

Neural networks can be valuable when the functional relationship between dependent and independent variables is not known. Radial basis function (RBF) networks form one of the essential types of neural networks and are used for control, signal processing pattern recognition and time series analysis [10].

The concept "fuzziness" was first proposed by the American philosopher Black in 1937. The fundamentals of Fuzzy Set Theory (FST) were constituted by Zadeh in 1965. FST is applied to the fields of operation research, management sciences, artificial intelligence/expert systems, statistics and many other fields [14, 20].

Nonparametric regression relies on the experimenter to supply only qualitative information about the regression function. This qualitative information sometimes cannot be determined by precise number; in this case fuzzy numbers are employed. Cheng and Lee (2001) used RBF networks in fuzzy regression analysis without a predefined functional relationship between the input and the output. In that approach, only weights between the hidden unit and the output unit are considered as fuzzy numbers [3].

The main purpose of this study is to present the FNPR estimators and FRBF networks and to compare them. NPR are models which show the functional relationship between the dependent and the independent variables is not known but flexible.

In practice, the values of dependent and/or independent variables can not always be expressed by crisp numbers. In such cases, fuzzy numbers must be used. Therefore, in this study, at first we propose an FNPR estimator where both dependent and independent variables are fuzzy numbers. In the FNPR estimator, we used the Nadaraya-Watson kernel regression estimator which is the most commonly used type of NPR estimator. In the Nadaraya-Watson kernel regression estimator, dependent and independent variables are

---
♠Corresponding author, e-mail: nimet@selcuk.edu.tr

considered as fuzzy numbers and an FNPR estimator is constituted. Secondly, RBF Networks which is one of the types of neural networks are used as alternative methods to the Nadaraya-Watson estimator. In light of this idea, we proposed FRBF networks as an alternative method to the FNPR estimator. FRBF networks are RBF networks where all input, output and weights are fuzzy numbers.

This study is organized as follows: fuzzy arithmetic is considered in Section 2 and the FNPR estimator is given in Section 3. In Section 4, FRBF networks are suggested in detail. An application and results are given in Section 5. The conclusion part of this study is summarized in Section 6.

2. FUZZY ARITHMETIC

"Fuzziness" is the ambiguity that can be found in the definition of a concept or the meaning of a word. For example, the uncertainity in expressions like "old person", "high temperature" or "small number" can be called fuzziness [18]. Let X be a universe whose generic elements are denoted by x and A be a classical subset of X. The membership function $\mu_A(x)$ is the degree of membership of x in A. If the grade of the membership is allowed to be in the set {0,1}, then A is called "classical (crisp) set". Nevertheless, if the grade of the membership is allowed to be in the interval [0, 1], A is called "fuzzy set" and generally denoted by $\widetilde{A}$. For $x \in X$ and $A \subseteq X$, the membership function of the fuzzy set is denoted by $\mu_A(x) : X \rightarrow [0,1]$. Fuzzy set theory is a general form of the crisp set [14].

Some basic concepts of the fuzzy set $\widetilde{A}$ are, for example, weak α-level set (α-cut), strong α-level set, convexity and normality. For a normal and convex fuzzy set, $\forall \alpha \in [0,1]$: the weak α-level set is a closed interval, then it is called "fuzzy number". Triangular and trapezoidal fuzzy numbers are among the most important ones of the several types of fuzzy numbers. Generally, triangular fuzzy numbers are denoted by X=(a,c,b), where "c" is a center, "a" is a lower and "b" is an upper limit of fuzzy number. A triangular fuzzy number is called a symmetrical triangular fuzzy number, if c-a=b-c [14, 18]. α-level set of the fuzzy number A can be written as $[A]_\alpha = \{x \in A | \mu_A(x) \geq \alpha\}$. Since the α-level sets of fuzzy numbers become closed intervals, $[A]_\alpha$ is denoted by $[A]_\alpha = \left[[A]_\alpha^L, [A]_\alpha^U\right]$. Arithmetic operations on α-level set of fuzzy numbers are equal to the arithmetic operations on closed intervals [14].

İf A=[a,b] and B=[d,e] are two intervals, then some arithmetic operations on intervals are given as follows [1, 13, 16].

i) $A + B = [a,b] + [d,e] = [a+d, b+e]$

ii) $A - B = [a,b] - [d,e] = [a-e, b-d]$

iii) $A.B = [\min(ad, ae, bd, be), \max(ad, ae, bd, be)]$

iv) $A/B = [\min(a/d, a/e, b/d, b/e), \max(a/d, a/e, b/d, b/e)]$

v) $k.A = k.[a,b] = \begin{cases} [k.a, k.b], & k \geq 0 \\ [k.b, k.a], & k < 0 \end{cases}$

vi) $f(A) = f([a,b]) = [f(a), f(b)]$, when f is a continuous and monotonic non decreasing function.

$f(A) = f([a,b]) = [f(b), f(a)]$, when f is a continuous and monotonic non increasing function.

vii) $D(A,B) = D([a,b],[c,d]) = \max\{|a-c|, |b-d|\}$

where; D is a distance between two intervals.

viii) $\|[A]\| = \|[a,b]\| = D([a,b], 0) = \max(|a|, |b|)$

where; $\|[A]\|$ is an absolute value of the interval [A]

ix) $\|[A]\| = \max\{|A_1|, |A_2|, ..., |A_n|\}$

where; $\|[A]\|$ is a norm of the interval [A]. [A] is an interval vector and can be defined as $[A] = \begin{bmatrix} [A_1] & [A_2] & \cdots & [A_n] \end{bmatrix}^T$

3. FUZZY KERNEL REGRESSION ESTIMATOR

In this section, kernel regression estimators and fuzzy kernel regression estimators are introduced.

3.1. Kernel Regression Estimator

Regression analysis is a statistical technique for researching and modeling the relationship between the dependent and the independent variables. There is a functional relationship between these variables. If this relationship is unknown but flexible then it is called a "nonparametric regression model". The NPR model can be defined as,

$$y_i = m(x_i) + \varepsilon_i, \ i = 1,2,...,n \qquad (1)$$

In Eqn(1), "m" is an unknown regression function, "ε_i" is an independent and identically distributed (i.i.d.) random variable with $E(\varepsilon_i) = 0$ and $Var(\varepsilon_i) = \sigma^2$ [6, 11]. Nonparametric regression estimator is defined as,

$$\hat{m}(x) = \sum_{i=1}^{n} W_{hi}(x) Y_i \qquad (2)$$

where; $W_{hi}(x)$ denotes weights and h is a bandwith parameter. İf the weights $W_{hi}(x)$ are based on kernel functions then NPR is called "kernel regression". Kernel function K is a continuous, bounded and symmetrical real function. The most used kernel functions are uniform, normal, Epanechnikov and quartic kernel functions.

Several estimators have been used in kernel regression, such as the Nadaraya-Watson(NW) estimator, the Priestley-Chao estimator and the Gasser-Müller estimator [2 ,5, 11]. The Nadaraya-Watson estimator is

defined as follows,

$$\hat{m}_h(x) = \frac{\sum_{i=1}^{n} K_h(x - X_i)Y_i}{\sum_{i=1}^{n} K_h(x - X_i)}$$

$$= \frac{\sum_{i=1}^{n} K(\frac{x - X_i}{h})Y_i}{\sum_{i=1}^{n} K(\frac{x - X_i}{h})} \qquad (3)$$

In Eqn(3), K(.) is a kernel function and "h" is a bandwith parameter. The application of NPR always requires the choice of a bandwith parameter h. The bandwith parameter h depends on the sample size n. Usually, it is selected to balance the trade-off between variance and squared bias by minimizing some global measures of error [11, 17, 19].

3.2. Fuzzy Kernel Regression Estimator

In the fuzzy kernel regression estimator, the Nadaraya-Watson estimator is considered and fuzzified. The fuzzifying operation is performed by considering the dependent and the independent variables as symmetric triangular fuzzy numbers in the Nadaraya-Watson estimator. Independent and dependent variables are denoted by $X_i = (x_i, t_i)$ and $Y_i = (y_i, l_i)$, respectively. α-level sets of X_i, Y_i and h are expressed as, $[X_i]_\alpha = [X_i^L, X_i^U]$, $[Y_i]_\alpha = [Y_i^L, Y_i^U]$ and $[h]_\alpha = [h^L, h^U]$. We put their intervals instead of X_i, Y_i and h in Eqn(3) and apply arithmetic operations on these intervals. As a result, FNPR estimator is obtained as follows;

$$\hat{m}_h\left([x^L, x^U]\right) = \sum_{i=1}^{n} W_{hi}\left([x^L, x^U]\right)[Y_i^L, Y_i^U]$$

$$= \sum_{i=1}^{n} [W_{hi}(x^L), W_{hi}(x^U)][Y_i^L, Y_i^U]$$

$$= \sum_{i=1}^{n} [W_{hi}(x^L)Y_i^L, W_{hi}(x^U)Y_i^U]$$

$$= \left[\frac{\sum_{i=1}^{n} K\left(\frac{x^L - X_i^L}{h^L}\right)Y_i^L}{\sum_{i=1}^{n} K\left(\frac{x^L - X_i^L}{h^L}\right)}, \frac{\sum_{i=1}^{n} K\left(\frac{x^U - X_i^U}{h^U}\right)Y_i^U}{\sum_{i=1}^{n} K\left(\frac{x^U - X_i^U}{h^U}\right)} \right]$$

(4)

where; K(.) is a kernel function, h^L and h^U are the lower and the upper limits of bandwith parameter h, respectively. X_i^L and X_i^U are the lower and upper limits of X_i and also Y_i^L and Y_i^U are the lower and the upper limits of Y_i, respectively.

4. FUZZY RADIAL BASIS FUNCTION NETWORKS

In this section, RBF networks, FRBF networks and the training of FRBF networks are presented.

4.1. Radial Basis Function Networks

Radial basis function networks were developed by Broomhead&Lawe in 1988 and proposed as an alternative method to multi layer perceptrons [10]. The net input to the hidden layer is the distance from the input vector to the weight vector and computed as follows;

$$h_{pj} = \exp\left\{ -\frac{\left\| X_{pi} - w_{ij} \right\|^2}{2\sigma_{pj}^2} \right\} \qquad (5)$$

where; X_{pi} is an input vector, w_{ij}'s are weights between any input unit and hidden unit j. σ_{pj}^2 is a normalization factor for the hidden unit j. The distance is usually computed in Euclidean metric. w_{ij}'s are called "RBF centers" and determined using "c-means clustering algorithm or fuzzy c-means clustering algorithm" [7]. Normalization factor σ_{pj}^2's are determined by

$$\sigma_{pj}^2 = \frac{1}{m} \sum_{p=1}^{m} \left\| X_{pi} - w_{ij} \right\|^2 \qquad (6)$$

where; X_{pi} is a training pattern in the cluster, w_{ij}'s are the centers of the cluster associated with the hidden unit j and m is the number of training patterns in that cluster.

Output unit is calculated by

$$\hat{Y}_p = \sum_{j=1}^{n_H} v_j h_{pj} \qquad (7)$$

where; v_j's are the weights from the hidden unit j to the output unit [7]. In Eqn.s(5), (6) and (7); i $(i=1,2,...,n_I)$ denotes the number of input units, p $(p=1,2,...,n)$ shows the observation number and j $(j=1,2,...,n_H)$ indicates the number of hidden units.

RBF networks are trained as hybrid networks. The weights w_{ij} and v_j and also the normalization factor

σ_{pj}^2 are updated by backpropagation algorithm. If the error criterion is minimum, then the process will be stopped; otherwise it must be repeated [4, 7].

4.2. Fuzzy Radial Basis Function Networks

Fuzzy radial basis function networks are RBF networks with fuzzy inputs and/or fuzzy outputs and/or fuzzy weights. We introduce the FRBF networks whose inputs X_p, outputs Y_p and also weights w_{ij} and v_j are fuzzy. In the proposed FRBF Networks, the FCM algorithm is modified due to the fact that both X_i and w_{ij} are fuzzy numbers. The training algorithms of FRBF networks are constituted by fuzzifying Choi et al.(2003)'s RBF backpropagation algorithm and Ishibuchi et.al.(1995)'s backpropagation algorithm.

α-level sets of fuzzy inputs X_{pi} and fuzzy outputs Y_p are expressed as $[X_{pi}]_\alpha = [X_{pi}^L, X_{pi}^U]$ and $[Y_p]_\alpha = [Y_p^L, Y_p^U]$, respectively. The weights between the input units and the hidden units are considered as symmetrical triangular fuzzy numbers and denoted by $w_{ij} = (w_{ij}^L, w_{ij}^C, w_{ij}^U)$. Where, w_{ij}^L is the lower limit, w_{ij}^C is the center and w_{ij}^U is the upper limit of w_{ij}. α-level sets of w_{ij} are written as;

$$[W_{ij}]_\alpha = \left[[W_{ij}]_\alpha^L, [W_{ij}]_\alpha^U \right] \qquad (8)$$

where;

$$[W_{ij}]_\alpha^L = w_{ij}^L (1 - \frac{\alpha}{2}) + w_{ij}^U (\frac{\alpha}{2})$$

$$[W_{ij}]_\alpha^U = w_{ij}^L (\frac{\alpha}{2}) + w_{ij}^U (1 - \frac{\alpha}{2})]$$

The weights between the hidden units and the output unit are considered as symmetrical triangular fuzzy numbers and denoted by $v_j = (v_j^L, v_j^C, v_j^U)$. α-level sets of V_j is written as;

$$[v_j]_\alpha = [[v_j]_\alpha^L, [v_j]_\alpha^U] \qquad (9)$$

where;

$$[v_j]_\alpha^L = v_j^L (1 - \alpha/2) + v_j^U (\alpha/2)$$

$$[v_j]_\alpha^U = v_j^L (\alpha/2) + v_j^U (1 - \alpha/2)$$

Hidden unit j is calculated by,

$$\overline{[h_{pj}]}_\alpha = \exp\left(-\frac{1}{2} \left(\frac{\left\| [X_{pi}]_\alpha - [W_{ij}]_\alpha \right\|}{[\sigma_{pj}]_\alpha} \right)^2 \right)$$

$$= \exp\left[\left[-\frac{1}{2} \frac{\left| \max\left\{ \left| [X_{pi}]_\alpha^L - [W_{ij}]_\alpha^L \right|, \left| [X_{pi}]_\alpha^U - [W_{ij}]_\alpha^U \right| \right\} \right|^2}{[\sigma_{pj}^2]_\alpha} \right] \right] \qquad (10)$$

In Eqn(10), $p=1,2,...,n$; $i=1,2,...,n_I$; $j=1,2,...,n_H$, where, $\overline{[h_{pj}]}_\alpha$ is obtained as a crisp number because of arithmetic operations on the interval number.

The normalization factor of the hidden unit j is determined by;

$$\overline{[\sigma_{pj}^2]}_\alpha = \frac{1}{m} \sum_{p=1}^{m} \left\| [X_{pi}]_\alpha - [W_{ij}]_\alpha \right\|^2$$

$$= \frac{1}{m} \left\{ \begin{array}{l} \left(\max\left\{ \left| [X_{1i}]_\alpha^L - [W_{ij}]_\alpha^L \right|, \left| [X_{1i}]_\alpha^U - [W_{ij}]_\alpha^U \right| \right\} \right)^2 \\ + ... + \\ \left(\max\left\{ \left| [X_{mi}]_\alpha^L - [W_{ij}]_\alpha^L \right|, \left| [X_{mi}]_\alpha^U - [W_{ij}]_\alpha^U \right| \right\} \right)^2 \end{array} \right\} \qquad (11)$$

where; $\overline{[\sigma_{pj}^2]}_\alpha$ is a crisp number.

Fuzzy output unit for observation p is calculated by;

$$[\hat{Y}_p]_\alpha = \frac{\sum\limits_{j=1}^{n_H} [V_j]_\alpha \overline{[h_{pj}]}_\alpha}{\sum\limits_{j=1}^{n_H} \overline{[h_{pj}]}_\alpha}$$

$$= \left[\frac{\sum\limits_{j=1}^{n_H} [V_j]_\alpha^L \cdot \overline{[h_{pj}]}_\alpha}{\sum\limits_{j=1}^{n_H} \overline{[h_{pj}]}_\alpha}, \frac{\sum\limits_{j=1}^{n_H} [V_j]_\alpha^U \cdot \overline{[h_{pj}]}_\alpha}{\sum\limits_{j=1}^{n_H} \overline{[h_{pj}]}_\alpha} \right] \qquad (12)$$

Let Y_p be a fuzzy target corresponding to the fuzzy input X_p. The cost function for α-level sets of the fuzzy output $\hat{Y}_p$ and the corresponding fuzzy target Y_p is defined as follows [12].

$$E_{p,\alpha} = E_{p,\alpha}^L + E_{p,\alpha}^U \qquad (13)$$

where;

$$E_{p,\alpha}^L = \frac{\alpha}{2} \left([Y_p]_\alpha^L - [\hat{Y}_p]_\alpha^L \right)^2$$

$$E_{p,\alpha}^U = \frac{\alpha}{2} \left([Y_p]_\alpha^U - [\hat{Y}_p]_\alpha^U \right)^2$$

The cost function E is obtained by;

$$E = \sum_{p=1}^{n} \sum_{i=1}^{s} E_{p,\alpha_i} \qquad (14)$$

4.3. The Training of the Fuzzy Radial Basis Function Networks

In the training of FRBF networks, we rearrange BP algorithms proposed by Choi et.al.(2003) and Ishibuchi et.al.(1995) for fuzzy numbers. The purpose of FRBF networks is to minimize the cost function E. Then we use this rearranged BP algorithm for the training of v_j, w_{ij} and σ_{pj}^2

The fuzzy weights v_j's are updated by;

$$v_j^L(t+1) = v_j^L(t) + \Delta v_j^L(t) \qquad (15)$$

$$v_j^U(t+1) = v_j^U(t) + \Delta v_j^U(t) \qquad (16)$$

If $v_j^L > v_j^U$, then

$$v_j^L(t+1) = \min\{v_j^L(t+1), v_j^U(t+1)\}$$

$$v_j^U(t+1) = \max\{v_j^L(t+1), v_j^U(t+1)\}$$

In the Eqn(15)-(16), $\Delta v_j^L(t)$ and $\Delta v_j^U(t)$ are calculated by;

$$\Delta v_j^L(t) = -\eta \frac{\partial E_{p,\alpha}}{\partial v_j^L} + \lambda . \Delta v_j^L(t-1) \qquad (17)$$

$$\Delta v_j^U(t) = -\eta \frac{\partial E_{p,\alpha}}{\partial v_j^U} + \lambda . \Delta v_j^U(t-1) \qquad (18)$$

In the Eqn(17)-(18), $\frac{\partial E_{p,\alpha}}{\partial v_j^L}$ and $\frac{\partial E_{p,\alpha}}{\partial v_j^U}$ can be written as;

$$\frac{\partial E_{p,\alpha}}{\partial v_j^L} = -\alpha \left([Y_p]_\alpha^L - [\hat{Y}_p]_\alpha^L h_{pj}(1-\frac{\alpha}{2}) \right) - \alpha \left([Y_p]_\alpha^U - [\hat{Y}_p]_\alpha^U(\frac{\alpha}{2}) \right)$$

$$\frac{\partial E_{p,\alpha}}{\partial v_j^U} = -\alpha \left([Y_p]_\alpha^L - [\hat{Y}_p]_\alpha^L h_{pj}(\frac{\alpha}{2}) \right) - \alpha \left([Y_p]_\alpha^U - [\hat{Y}_p]_\alpha^U h_{pj}(1-\frac{\alpha}{2}) \right)$$

The fuzzy weights w_{ij}'s are updated by

$$w_{ij}^L(t+1) = w_{ij}^L(t) + \Delta w_{ij}^L(t)$$
(19)

$$w_{ij}^U(t+1) = w_{ij}^U(t) + \Delta w_{ij}^U(t) \qquad (20)$$

If $w_{ij}^L > w_{ij}^U$ then

$$w_{ij}^L(t+1) = \min\{w_{ij}^L(t+1), w_{ij}^U(t+1)\}$$

$$w_{ij}^U(t+1) = \max\{w_{ij}^L(t+1), w_{ij}^U(t+1)\}$$

In the Eqn(19)-(20), $\Delta w_{ij}^L(t)$ and $\Delta w_{ij}^U(t)$ are calculated by;

$$\Delta w_{ij}^L(t) = -\eta \frac{\partial E_{p,\alpha}}{\partial w_{ij}^L} + \lambda \Delta w_{ij}^L(t-1) \qquad (21)$$

$$\Delta w_{ij}^U(t) = -\eta \frac{\partial E_{p,\alpha}}{\partial w_{ij}^U} + \lambda \Delta w_{ij}^U(t-1) \qquad (22)$$

In the Eqn(21)-(22), $\frac{\partial E_{p,\alpha}}{\partial w_{ij}^L}$ and $\frac{\partial E_{p,\alpha}}{\partial w_{ij}^U}$ can be written as;

$$\frac{\partial E_{p,\alpha}}{\partial w_{ij}^L} = \frac{\partial E_{p,\alpha}}{\partial [W_{ij}]_\alpha^L} \frac{\partial [W_{ij}]_\alpha^L}{\partial w_{ij}^L} + \frac{\partial E_{p,\alpha}}{\partial [W_{ij}]_\alpha^U} \frac{\partial [W_{ij}]_\alpha^U}{\partial w_{ij}^L}$$

$$= \frac{\partial E_{p,\alpha}}{\partial [W_{ij}]_\alpha^L} (1-\frac{\alpha}{2}) + \frac{\partial E_{p,\alpha}}{\partial [W_{ij}]_\alpha^U} (\frac{\alpha}{2})$$

$$\frac{\partial E_{p,\alpha}}{\partial w_{ij}^U} = \frac{\partial E_{p,\alpha}}{\partial [W_{ij}]_\alpha^U} \frac{\partial [W_{ij}]_\alpha^U}{\partial w_{ij}^U} + \frac{\partial E_{p,\alpha}}{\partial [W_{ij}]_\alpha^L} \frac{\partial [W_{ij}]_\alpha^L}{\partial w_{ij}^U}$$

$$= \frac{\partial E_{p,\alpha}}{\partial [W_{ij}]_\alpha^U} (1-\frac{\alpha}{2}) + \frac{\partial E_{p,\alpha}}{\partial [W_{ij}]_\alpha^L} (\frac{\alpha}{2})$$

where; $\frac{\partial E_{p,\alpha}}{\partial [W_{ij}]_\alpha^L}$ and $\frac{\partial E_{p,\alpha}}{\partial [W_{ij}]_\alpha^U}$ can be written as follows;

i) If

$$\max\left\{ \left| [X_{pi}]_\alpha^L - [W_{ij}]_\alpha^L \right|, \left| [X_{pi}]_\alpha^U - [W_{ij}]_\alpha^U \right| \right\} = \left| [X_{pi}]_\alpha^L - [W_{ij}]_\alpha^L \right|$$

then

$$\frac{\partial E_{p,\alpha}}{\partial [W_{ij}]_\alpha^L} = -\alpha \left([Y_p]_\alpha^L - [\hat{Y}_p]_\alpha^L \right) h_{pj}(\sigma_{pj})^{-2} \left| [X_{pi}]_\alpha^L - [W_{ij}]_\alpha^L \right| v_j^L$$

$$-\alpha \left([Y_p]_\alpha^U - [\hat{Y}_p]_\alpha^U \right) h_{pj}(\sigma_{pj})^{-2} \left| [X_{pi}]_\alpha^L - [W_{ij}]_\alpha^L \right| v_j^U$$

$$\frac{\partial E_{p,\alpha}}{\partial [W_{ij}]_\alpha^U} = 0$$

ii) If

$$\max\left\{\left|[X_{pi}]_\alpha^L - [W_{ij}]_\alpha^L\right|, \left|[X_{pi}]_\alpha^U - [W_{ij}]_\alpha^U\right|\right\} = \left|[X_{pi}]_\alpha^U - [W_{ij}]_\alpha^U\right|$$

then

$$\frac{\partial E_{p,\alpha}}{\partial [W_{ij}]_\alpha^L} = 0$$

$$\frac{\partial E_{p,\alpha}}{\partial [W_{ij}]_\alpha^U} = \alpha\left([Y_p]_\alpha^L - [\hat{Y}_p]_\alpha^L\right) h_{pj}(\sigma_{pj})^{-2} \left|[X_{pi}]_\alpha^U - [W_{ij}]_\alpha^U\right| v_j^L$$

$$- \alpha\left([Y_p]_\alpha^U - [\hat{Y}_p]_\alpha^U\right) h_{pj}(\sigma_{pj})^{-2} \left|[X_{pi}]_\alpha^U - [W_{ij}]_\alpha^U\right| v_j^U$$

The normalization factors σ_{pj}^2's are updated by

$$\sigma_{pj}(t+1) = \sigma_{pj}(t) + \Delta\sigma_{pj}(t) \qquad (23)$$

In the Eqn(23), $\Delta\sigma_{pj}(t)$ is calculated by;

$$\Delta\sigma_{pj}(t) = -\eta \frac{\partial E_{p,\alpha}}{\partial \sigma_{pj}} + \lambda . \Delta\sigma_{pj}(t-1) \qquad (24)$$

In the Eqn (24), $\dfrac{\partial E_{p,\alpha}}{\partial \sigma_{pj}}$ can be written as follows;

$$\frac{\partial E_{p,\alpha}}{\partial \sigma_{pj}} = \zeta^L + \zeta^U$$

where; ζ^L and ζ^U are defined as;

i) If

$$\max\left\{\left|[X_{pi}]_\alpha^L - [W_{ij}]_\alpha^L\right|, \left|[X_{pi}]_\alpha^U - [W_{ij}]_\alpha^U\right|\right\} = \left|[X_{pi}]_\alpha^L - [W_{ij}]_\alpha^L\right|$$

then

$$\zeta^L = -\alpha([Y_p]_\alpha^L - [\hat{Y}_p]_\alpha^L) h_{pj}(\sigma_{pj})^{-3} \left|[X_{pi}]_\alpha^L - [W_{ij}]_\alpha^L\right|^2 v_j^L$$

$$\zeta^U = -\alpha([Y_p]_\alpha^U - [\hat{Y}_p]_\alpha^U) h_{pj}(\sigma_{pj})^{-3} \left|[X_{pi}]_\alpha^L - [W_{ij}]_\alpha^L\right|^2 v_j^U$$

ii) If

$$\max\left\{\left|[X_{pi}]_\alpha^L - [W_{ij}]_\alpha^L\right|, \left|[X_{pi}]_\alpha^U - [W_{ij}]_\alpha^U\right|\right\} = \left|[X_{pi}]_\alpha^U - [W_{ij}]_\alpha^U\right|$$

then

$$\zeta^L = -\alpha\left([Y_p]_\alpha^L - [\hat{Y}_p]_\alpha^L\right) h_{pj}(\sigma_{pj})^{-3} \left|[X_{pi}]_\alpha^U - [W_{ij}]_\alpha^U\right|^2 v_j^L$$

$$\zeta^U = -\alpha\left([Y_p]_\alpha^U - [\hat{Y}_p]_\alpha^U\right) h_{pj}(\sigma_{pj})^{-3} \left|[X_{pi}]_\alpha^U - [W_{ij}]_\alpha^U\right|^2 v_j^U$$

In Eqn.s (17), (18), (21), (22) and (24), η is a learning constant, λ is a momentum constant and t shows the number of repetitions.

An algorithm for the training of FRBF networks is given as follows;

Step 1: Fuzzy weights w_{ij}'s are determined by a modified FCM algorithm. Fuzzy weights v_j's are initialized to random fuzzy number.

Step 2: Step 3 must be repeated for α-level sets such as $\alpha_1, \alpha_2, ..., \alpha_s$.

Step 3: The following i), ii), iii) and iv) must be repeated for p=1,2,...,n.

i) h_{pj}, $\hat{Y}_p$ and $E_{p,\alpha}$ are calculated by Eqn.s(10)-(13).

ii) Fuzzy weights v_j's are updated by Eqn.s(15)-(18).

iii) Fuzzy weights w_{ij}'s are updated by Eqn.s(19)-(22).

iv) Normalization factors σ_{pj}^2's are updated by Eqn.s(23)-(24).

Step 4: If the iteration number reaches to a predefined value and then stop. Otherwise go to Step2.

5. APPLICATION

We have applied the FNPR estimator and the FRBF Networks to the data of the "Turkish Pastrami" production experiment. In this experiment, chemical, microbiological and organoleptic analyses are made. Pastrami production is made in three trials. In the production stage, measurements are taken on the 1st, 7th, 15th, 30th and 60th days. As a result of the chemical analysis, amounts of humidity, salt, value of pH and value of water activity(a_w) are determined. As a result of microbiological analysis, total number of mezophil aerob microorganism(PCA), number of staphylococcus-micrococcus microorganism(MSA), number of lactobacillus microorganism(RA) and number of yeast and mold microorganisms(PDA) are determined. As a result of organoleptic analysis, colour, taste, appearance and texture are evaluated by a test panel of six persons [9].

In this study, the functional relationship between chemical factors and microbiological factors and also the functional relationship between chemical factors and organoleptic factors are estimated. The efficiencies of these factors on pastrami production are evaluated. Independent variables are taken as salt, pH and water activity(a_w). Dependent variables are taken as PCA, MSA, RA, PDA, taste, texture and color.

Errors are computed by using mean squared error(MSE) defined in terms of the difference between two fuzzy numbers suggested by Ishibuchi et.al.(1995) and Lin et.al.(2000) as follows;

$$MSE = \frac{1}{n}\sum_{p=1}^{n}\sum_{i=1}^{s}\frac{\alpha_i}{2}\left\{\left([\hat{Y}_p]_{\alpha_i}^L - [Y_p]_{\alpha_i}^L\right)^2 + \left([\hat{Y}_p]_{\alpha_i}^U - [Y_p]_{\alpha_i}^U\right)^2\right\} \qquad (25)$$

The general functional structure between the dependent and independent variables is defined as $Y = f(X)$. Then, fifteen different functions are constituted and fuzzy intervals are formed by computing the values of mean $\pm$ standart deviation. Means and standart deviations of the measurements are taken at the end of three trials of the "Turkish Pastrami" experiment. Fuzzy estimates are from the FNPR estimator and the FRBF networks after forming fuzzy intervals for all fuzzy dependent and independent variables. MSE values belonging to 15 different models are evaluated by the FNPR estimator and the FRBF networks.

These 15 models are,

PCA $=f$(Salt), MSA $=f$(Salt)

RA$=f$(Salt), PDA$=f$(Salt)

Taste$=f$(Salt)

PCA $=f$(pH), MSA $=f$(pH)

RA $=f$(pH), PDA$=f$(pH)

Texture$=f$(pH)

PCA $=f$(a$_w$), MSA $=f$(a$_w$)

RA$=f$(a$_w$), PDA$=f$(a$_w$)

Color$=f$(a$_w$)

Bandwith parameters of FNPR estimator are determined by XPLORE package programme for each model and given in Table 1.

Table 1. The values of bandwith parameter h for FNPR estimator

Model	Parameter h
PCA $=f$(Salt)	[0.01, 0.01]
MSA $=f$(Salt)	[0.01, 0.01]
RA $=f$(Salt)	[0.01, 0.01]
PDA $=f$(Salt)	[0.01, 0.01]
Taste $=f$(Salt)	[0.01, 0.01]
PCA $=f$(pH)	[0.15, 0.15]
MSA $=f$(pH)	[0.01, 0.01]
RA $=f$(pH)	[0.01, 0.01]
PDA $=f$(Ph)	[0.01, 0.01]
Texture $=f$(Ph)	[10, 10]
PCA $=f$(a$_w$)	[0.01, 0.01]
MSA $=f$(a$_w$)	[0.01, 0.01]
RA $=f$(a$_w$)	[0.99, 0.99]
PDA $=f$(a$_w$)	[0.07, 0.07]
Color $=f$(a$_w$)	[0.01, 0.01]

FRBF networks with a single input unit, three hidden units and a single output unit are constituted for considered data. These networks were trained by the proposed training algorithm given in Section 4. In the training algorithm, the learning constant is η =0.01 and the momentum constant is λ =0.1.

Beginning values of weights w_{ij}'s between input-hidden units are determined by a modified FCM algorithm. In this algorithm, w_{ij}'s are calculated by independent variables taken as Salt, pH and a$_w$, respectively. Beginning values of normalization factor σ_j's of hidden units are calculated by Eqn(6). Beginning values of weights v_j's between hidden-output units are determined as arbitrary fuzzy numbers for all hidden units (j=1,2,3). The values of w_{ij}, σ_j and v_j's are summarized in Table 2.

Table 2. The values of weights and normalization factor for FRBF Networks

Model	The weights w_{ij} between input-hidden unit	The weights v_j between hidden-output unit	The normalization factor σ_j
PCA $=f$(Salt) MSA $=f$(Salt) RA$=f$(Salt) PDA$=f$(Salt) Taste$=f$(Salt)	w_{ij}^L =[5.99, 4.00, 5.84] w_{ij}^U =[7.87, 6.72, 6.20]	v_j^L =[1, 2, 1] v_j^U =[3, 3, 2]	σ_j =[1.57, 1.68, 1.12]
PCA $=f$(pH) MSA $=f$(pH) RA$=f$(pH) PDA$=f$(pH) Texture$=f$(pH)	w_{ij}^L =[5.80, 5.66, 5.45] w_{ij}^U =[5.84, 5.70, 5.63]	v_j^L =[1, 2, 1] v_j^U =[3, 3, 2]	σ_j =[0.28, 0.17,0.18]
PCA $=f$(a$_w$) MSA $=f$(a$_w$) RA$=f$(a$_w$) PDA$=f$(a$_w$) Color$=f$(a$_w$)	w_{ij}^L =[0.90, 0.81, 0.77] w_{ij}^U =[0.90, 0.83, 0.81]	v_j^L =[1, 2, 1] v_j^U =[3, 3, 2]	σ_j =[0.09, 0.04,0.07]

Estimated values of dependent variables for each model are calculated by FNPR estimator and FRBF networks. Then real values and estimated values are given in detail in [19]. MSE values of FNPR estimator and FRBF networks are shown in Table 3.

By investigating Table 3, it can be seen that the MSE values of the FNPR estimator are less than the MSE values of FRBF networks for all models. Therefore, the FNPR estimator is a more efficient method than the FRBF networks for these data.

Table 3. MSE values regarding fuzzy estimates evaluated by FNPR estimator and FRBF networks

Model	MSE_{FNPR}	MSE_{FRBF}
PCA = f(Salt)	0.0024	0.1973
MSA = f(Salt)	0.0791	1.2612
RA = f(Salt)	0.2382	12.5177
PDA = f(Salt)	0.0720	3.6713
Taste = f(Salt)	0.0071	0.1397
PCA = f(pH)	0.1270	0.2707
MSA = f(pH)	0.4701	1.2754
RA = f(pH)	2.1372	12.5428
PDA = f(PH)	1.1495	3.9400
Texture = f(PH)	0.1885	0.4029
PCA = f(a_w)	0.0864	0.2711
MSA = f(a_w)	0.2756	1.2766
RA = f(a_w)	6.6424	12.4789
PDA = f(a_w)	2.6516	3.9400
Color = f(a_w)	0.1749	0.2265

6. CONCLUSION

In practice, the values of dependent and/or independent variables can not always be expressed by crisp numbers. In such cases, fuzzy numbers must be used. Therefore, in this study, we proposed the fuzzy kernel regression estimator and also the fuzzy radial basis function networks. First of all, we proposed a FNPR estimator where both dependent and independent variables are fuzzy numbers. In the FNPR estimator, we used the Nadaraya-Watson kernel regression estimator. In a Nadaraya-Watson kernel regression estimator, dependent and independent variables are considered as fuzzy numbers and the FNPR estimator is constituted. The bandwith parameters of each model are given in Table 1. Secondly, we proposed FRBF networks as an alternative method to the FNPR estimator. FRBF networks are RBF networks where all input, output and weights are fuzzy numbers. Beginning values of weights w_{ij}'s between input-hidden units, beginning values of normalization factor σ_j's of hidden units and beginning values of weights v_j's between hidden-output units were summarized in Table 2 for FRBF networks. MSE values of the FNPR estimator and FRBF networks are given in Table 3. As can be seen from the results in Table 3, MSE values of the FNPR estimator are less than the MSE values of FRBF networks. Therefore, it can be concluded that the FNPR

estimator is more efficient than the FRBF networks for these data.

REFERENCES

[1] Alefeld, G., Mayer, G., "Interval Analysis: Theory and Applications", *Journal of Computational and Applied Mathematics*, 121: 421-464 (2000).

[2] Benedetti, J.K., "On the Nonparametric Estimation of Regression Functions", *Journal of the Royal Statistical Society Series B*, 39: 248-253 (1977).

[3] Cheng, C.B., Lee, E.S., "Fuzzy Regression with Radial Basis Function Network", *Fuzzy Sets and Systems*, 119: 291-301 (2001).

[4] Choi, S.W., Lee, D., Park, J.H., Lee, I.B., "Nonlinear Regression Using RBFN with Linear Submodels", *Chemometrics and Intelligent Laboratory Systems*, 65: 191-208 (2003).

[5] Chu, C.K., Marron, J.S., "Choosing a Kernel Regression Estimator", *Statistical Science*, 6(4): 404-436 (1991).

[6] Eubank, R.L., "Spline Smoothing and Nonparametric Regression", *Marcel Dekker Inc.,* New York, 1-8 (1988).

[7] Fu, L., "Neural Networks in Computer Intelligence", *McGraw Hill Inc.,* Singapure, 94-98 (1996).

[8] Gamgam, H., "Parametrik Olmayan İstatistiksel Teknikler", *Gazi Üniversitesi, Fen-Edebiyat Fakültesi Yayınları,* Ankara, 232-249 (1998).

[9] Gürbüz, Ü., Doğruer, Y., Nizamlıoğlu, M., "Pastırma Üretiminde Dumanlama İşleminin Uygulanabilme İmkanları ve Kaliteye Etkisi", *Veteriner Bilimleri Dergisi*, 13(2): 57-68 (1997).

[10] Györfi, L., Kohler, M., Krzyak, A., Walk, H., "Distribution Free Theory of Nonparametric Regression", *Springer-Verlag,* New York, 329-332 (2002).

[11] Hardle, W., "Applied Nonparametric Regression", *Cambridge University Press,* New York, 14-30 (1990).

[12] Ishibuchi, H., Kwon, K., Tanaka, H., "A Learning Algorithm of Fuzzy Neural Networks with Triangular Fuzzy Weights", *Fuzzy Sets and Systems*, 71: 277-293 (1995).

[13] Klir, G.J., Yuan, B., "Fuzzy Sets and Fuzzy Logic: Theory and Application", *Prentice Hall International Inc.,* New Jersey, 102-104 (1995).

[14] Lai, Y.J., Hwang, C.L., "Fuzzy Mathematical Programming: Method and Applications", *Springer-Verlag,* Berlin Heidelberg, 20-66 (1992).

[15] Lin, C.T., Duh, F.B., Liu, D.J., "A Neural Fuzzy Network for Word Information Processing", *Fuzzy Sets and Systems*, 127: 37-48 (2002).

[16] Moore, R.E., "Methods and Application of Interval Analysis", *SIAM*, Philadelphia, 9-14 (1979).

[17] Rice, J., "Bandwidth Choice for Nonparametric Regression", *The Annals of Statistics*, 12(4): 1215-1230 (1984).

[18] Terano, T., Asai, K., Sugeno, M., "Fuzzy Systems Theory and Its Applications", *Academic Press Inc.,* San Diego, 19-37 (1992).

[19] Yapıcı Pehlivan, N., "Parametrik Olmayan Regresyonda Bulanık Tahmin Ediciler", Doktora Tezi, *Selçuk Üniversitesi Fen Bilimleri Enstitüsü,* Konya (2005).

[20] Zadeh, L.A., "Fuzzy Sets", *Information and Control*, 8: 338-353 (1965).