Iranian Journal of Fuzzy Systems

Volume 23, Number 2, (2026), pp. 177-195

Original paper

Support vector weighted fuzzy regression

A. H. Khammar [1], S. M. Taheri [2], P. D'Urso [3] and M. Arefi [4]

[1] *Department of Statistics, School of Mathematical Sciences, University of Sistan and Baluchestan, Zahedan, Iran*
[2] *School of Engineering Science, College of Engineering, University of Tehran, Tehran, Iran*
[3] *Department of Social and Economic Sciences, Sapienza University of Rome, Piazzale Aldo Moro, 5-00185 Rome, Italy*
[4] *Department of Statistics, Faculty of Mathematical Sciences and Statistics, University of Birjand, Birjand, Iran*

amirhamzeh.kh@gmail.com, sm_ taheri@ut.ac.ir, pierpaolo.durso@uniroma1.it, Arefi@birjand.ac.ir

Abstract

Based on the idea of Support Vector Machine (SVM) methodology, a new robust support vector linear regression modelling known as Support Vector Weighted Fuzzy Regression (SVWFR) is introduced, for the case when the values of response variable are fuzzy rather than crisp. The extension of the proposed method to the nonlinear case is also investigated. In the proposed approach, a weighted operation is employed to improve the robustness of usual support vector fuzzy regression models by assigning weights to the support hyperplanes constraints. While the fuzzy machine learning-based models are typically sensitive to outliers, the advantages of the proposed models are their robustness with respect to outlier data. The efficiency and applicability of the proposed models are investigated by using three data sets: a synthetic dataset including outliers, a textile engineering data set, and a stress-test simulation with artificially introduced anomalies. Across all cases, the introduced models consistently outperformed current fuzzy regression also approaches, based on three well-known goodness of fit indices. Sensitivity analysis of nonlinear SVWFR parameters is examined.

Keywords: Kernel function, outlier, robustness, support vector machine.

1 Introduction and motivation

In many real-world applications of regression modelling, available data are imprecise and thus usually presented by fuzzy quantities rather than crisp numbers. In such situations, we need to employ fuzzy regression models which support the imprecise quantities. Following Tanaka et al's studies ([26] and [27]), the topic of fuzzy regression has expanded in two main directions: possibilistic approach and least-square approach. These approaches have been widely applied across numerous fields in recent decades. For further details, see e.g., D'Urso and Massari [9], D'Urso [10], and D'Urso and Gastaldi [11].

Meanwhile, in some situations we may encounter the types of outliers in the observed data. In such cases, it is usually preferred to use some robust approaches. The usual approach to achieve a robust regression model is to use certain distances during related optimization procedure of finding the parameters of the underlying model. In this topic, specially the absolute error distance is popular. Meanwhile, the weighted regression approach is another popular method in robust fuzzy regression analysis (see Chachi [5], Chang and Lee [7], D'Urso et al. [13], and D'Urso and Massari [12]). Readers interested in alternative robust frameworks may refer to see Arefi [2], Khammar et al. [21, 22], Taheri et al. [6], and Zeng et al. [33]. It should be noted that, based on three goodness of fit indices, we will compare our proposed models introduced in this paper with some of the mentioned works above, in Section 5.

On the other hand, the support vector machine (SVM) plays an important role in regression analysis. This method employs SVMs to approximate the underlying functional relationship in regression tasks (see Gu et al. [14], Gou et

Corresponding Author: M. Arefi
Received: December 2025; Revised: March 2026; Accepted: May 2026.
https://doi.org/10.22111/ijfs.2026.54371.9634

al. [19], Hastie et al. [17], Qin and Qiqi [25], and Vapnik [29]. The The core principle underlying these methods is structural risk minimization, which bounds the generalization error by the sum of the training error and a model complexity term. Hong and Hwang [18] were the first to apply SVM to the multivariate fuzzy linear regression model. Hao and Chiang [16] incorporated the concept of fuzzy set theory into the support vector regression machine, wherein the parameters to be identified in the model, such as the components of weight vector and bias term and the desired outputs in training samples, are set to be the fuzzy numbers. Asadollahi et al. [4] proposed a method for estimating fuzzy regression models based on a novel robust SVMs with exact predictors and fuzzy responses. They introduced a three-stage algorithm based on a modified robust loss function. Luo et al. [23] introduced a procedure to combine the fuzzy membership functions with support vector regression and support vector quantile regression models for short-term point and probabilistic load forecasting, respectively.

In spite of the usefulness various works in fuzzy repression by using the SVM approaches, however the problem of outlier data has been remained as a potential challenge. This article introduces a robust fuzzy regression model, based on SVM, for data with crisp predictors and fuzzy responses. For this purpose, we propose two new robust fuzzy support vector regression models (linear and nonlinear) known as support vector weighted fuzzy regression (SVWFR). Th main idea of the proposed approach, is to consider a weighted operation to guaranty the robustness of usual support vector fuzzy regression models by assigning weights to the support hyperplanes constraints. Meanwhile, the cross-validation grid search is used for regularization and smoothing parameters of the models. The effectiveness and advantages of the proposed methods are then compared with some common fuzzy regression models. Unlike the previous methods, this study addresses the problem of outliers data. Numerical results demonstrate that the proposed method yields sufficiently accurate estimates in fuzzy regression, even in the presence of outliers. Numerical examples are given to clarify the proposed model. The obtained models are justified by three well-known criterion.

This paper is organized as follows: In Section 2, a new SVFR (support vector fuzzy regression) approach is investigated to fit the linear fuzzy regression models when the response variable is a fuzzy number. In Section 3, we present a weighted version of SVFR for estimating the parameters in a fuzzy regression model. In Section 4, a nonlinear version of the SVWFR approach is proposed. To investigate the effectiveness of the proposed methods, four numerical comparison examples are presented in Section 5. In this section, the problem of sensitivity analysis of the parameters of th models is also investigated. In Section 6, a numerical example is provided to demonstrate the outliers detection capability of the proposed SVWFR approach. Finally, conclusions and new potential works are given in Section 7.

Throughout this article, the symbol $\tilde{N} = (N; l_N, r_N)_T$ is used for a triangular fuzzy number, wherein N, l_N, and r_N, are center, left spread, and right spread of $\tilde{N}$, respectively. The symmetric triangular fuzzy number is denoted by $\tilde{N} = (N; s_N)_T$. Details regarding fuzzy numbers and fuzzy arithmetic, can be found in [34].

2 A new linear support vector fuzzy regression

In this section, we introduce a new approach in support vector fuzzy regression (SVFR) to evaluate linear fuzzy regression models for crisp input-fuzzy output data, when the parameters of model are crisp quantities. To this end, we treat the lower and upper bounds of the fuzzy regression model as boundary lines within the SVR approach. Then, we maximize the distance between boundary lines in such a way that they include all the fuzzy data to find the optimal fuzzy regression model.

Suppose we are given a fuzzy training set $\{\mathbf{x_i}, \tilde{y}_i\}$, $i = 1, 2, ..., n$, where $\mathbf{x_i} = (x_{1i}, x_{2i}, ..., x_{pi})$, and $\tilde{y}_i = (y_i; l_{y_i})_T$. We consider the following fuzzy regression model

$$\tilde{Y} = b \oplus (\mathbf{wx}) \oplus \tilde{\delta} = (b + \mathbf{wx}; l_\delta)_T, \tag{1}$$

where, b is the bias term, $\mathbf{w} = (w_1, w_2, ..., w_p)$ is the weight vector and $\tilde{\delta} = (0; l_\delta)_T$ is the fuzzy error term (as a symmetric trangular fuzzy number), wherein l_δ is spread of $\tilde{\delta}$. Consequently, the output function $\tilde{Y}$ and the error term $\tilde{\delta}$ are, triangular fuzzy numbers whereas the input variables and the SVFR model parameter are crisp values.

In the following, we introduce a novel SVFR method for estimating the parameters of the fuzzy regression model defined in Eq. (1). Using arithmetic of fuzzy numbers [34], the values of estimated responses are obtained as

$$\hat{\tilde{Y}}_i = \hat{b} \oplus (\hat{\mathbf{w}}\mathbf{x_i}) \oplus \hat{\tilde{\delta}} = (\hat{b} + \hat{\mathbf{w}}\mathbf{x_i}; \hat{l}_\delta)_T. \tag{2}$$

Model (2) can be denoted by its lower limit $\hat{\hat{Y}}^L$, its center $\hat{\hat{Y}}^C$ and its upper limit $\hat{\hat{Y}}^U$ as

$$\hat{\hat{Y}}^L = \hat{b} + \hat{\mathbf{w}}\mathbf{x} - |\hat{\mathbf{w}}|\mathbf{l_x} - \hat{l}_\delta,$$
$$\hat{\hat{Y}}^C = \hat{b} + \hat{\mathbf{w}}\mathbf{x_i},$$
$$\hat{\hat{Y}}^U = \hat{b} + \hat{\mathbf{w}}\mathbf{x} + |\hat{\mathbf{w}}|\mathbf{l_x} + \hat{l}_\delta.$$

The center, lower limit and upper limit can be formulated as lower, center and upper hyperplanes as follows by introducing the technique of the SVR.

$$\hat{\hat{Y}}^L = 0,$$
$$\hat{\hat{Y}}^C = 0,$$
$$\hat{\hat{Y}}^U = 0.$$

In SVR approach, the optimal solution is determined by maximizing the margin of separation which is bounded by two support hyperplanes. The optimal fuzzy regression model, therefore, can be obtained by minimizing the margin bounded by lower and upper hyperplanes ($\hat{\hat{Y}}^L = 0$ and $\hat{\hat{Y}}^U = 0$). Figure 1 shows a typical visualization of such a procedure. Two support hyperplanes are the lower and upper bounds of the model, and the separating hyperplane is the center of the model.

To find the best fuzzy regression model of the form (2), using SVR method, we maximize $\hat{\hat{Y}}^L$ and $\hat{\hat{Y}}^U$, i.e the distance between lower and upper bounds of the model as boundary lines in the way that they include all the data i.e. $\hat{\hat{Y}}_i^L \leq \tilde{y}_i^L = y_i - l_{y_i}$ and $\hat{\hat{Y}}_i^U \geq \tilde{y}_i^U = y_i + l_{y_i}$, $i = 1, 2, ..., n$ (see Figure 1). Since $\hat{\hat{Y}}^L$ and $\hat{\hat{Y}}^U$ are two parallel boundary lines, we have

$$D(\hat{\hat{Y}}^L, \hat{\hat{Y}}^U) = \frac{|2\hat{l}_\delta|}{\sqrt{||\hat{\mathbf{w}}||^2}}.$$

Since maximization of $D^2(\hat{\hat{Y}}^L, \hat{\hat{Y}}^U)$ is equivalent to minimization of $\frac{||\hat{\mathbf{w}}||^2}{4\hat{l}_\delta^2}$, by modifying the basic idea of SVR for linear regression, we achieve the following minimization problem for model (1).

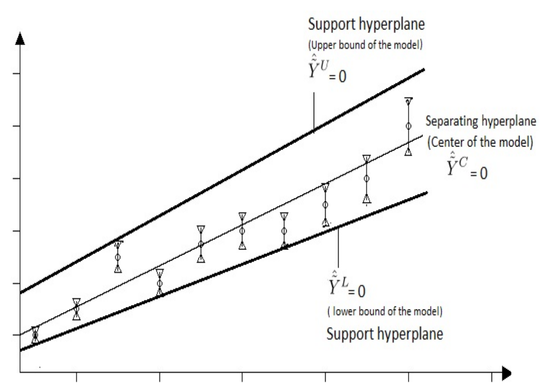

Figure 1: Explanation of SVFR.

$$\min_{\hat{b},\hat{\mathbf{w}},\hat{l}_\delta,\xi_i,\xi_i^*} \frac{||\hat{\mathbf{w}}||^2}{4\hat{l}_\delta^2} + C\sum_{i=1}^{n}(\xi_i + \xi_i^*),$$

$$s.t.$$

$$\begin{cases} \hat{b} + \hat{\mathbf{w}}\mathbf{x_i} - \hat{l}_\delta \leq (y_i - l_{y_i}) + \xi_i, \\ \hat{b} + \hat{\mathbf{w}}\mathbf{x_i} + \hat{l}_\delta \geq (y_i + l_{y_i}) - \xi_i^*, \\ 0 \leq \hat{l}_\delta \leq max(\mathbf{l_y}), \\ \xi_i \geq 0,\ \xi_i^* \geq 0,\ i = 1, 2, ..., n, \end{cases} \tag{3}$$

where, $C > 0$ is a fixed penalty parameter chosen by the user (a larger C corresponds to assigning a higher penalty to errors), ξ_i and ξ_i^* $(i = 1, 2, ..., n)$ are surplus variables that measure the amount of variation of the constraints for each point, and $\mathbf{l_y} = (l_{y_1}, l_{y_2}, ..., l_{y_n})$. By solving such a problem, the crisp coefficients and the parameter (spread) of the fuzzy error term would be estimated.

3 Linear support vector weighted fuzzy regression (Linear SVWFR)

In this section, we present a weighted SVFR method for estimating the parameters of the fuzzy linear regression model in Eq. (1) that is not sensitive to outliers. For this purpose, we extend the proposed SVFR in Sec. 2 by considering weights ω_i to support hyperplanes constraints.

We first consider the averages of differences between the lower and upper α-cuts of $\tilde{y}_i$ and $\hat{\tilde{Y}}_i$ as follows:

$$E_i^L = \int_0^1 |\tilde{y}_i^L[\alpha] - \hat{\tilde{Y}}_i^L[\alpha]| d\alpha,$$

and

$$E_i^U = \int_0^1 |\tilde{y}_i^U[\alpha] - \hat{\tilde{Y}}_i^U[\alpha]| d\alpha.$$

From the above equations, we consider minimum limitation of lower bound and maximum limitation of upper bound of the ith fuzzy output of SVFR as follows:

$$\hat{\tilde{Y}}_i^{L'} = \hat{\tilde{Y}}_i^L - MED(E_i^L),$$

and

$$\hat{\tilde{Y}}_i^{U'} = \hat{\tilde{Y}}_i^U + MED(E_i^U),$$

where, MED stands for the median absolute deviation as

$$MED(E_i^L) = median|E_i^L - median(\mathbf{E}^L)|,$$

and

$$MED(E_i^U) = median|E_i^U - median(\mathbf{E}^U)|,$$

where, $\mathbf{E}^L = (E_1^L, E_2^L, ..., E_n^L)$, and $\mathbf{E}^U = (E_1^U, E_2^U, ..., E_n^U)$. Now, for the ith case, its weight can be expressed as

$$\omega_i = \begin{cases} 1 & \text{if } \hat{\tilde{Y}}_i^{L'} \leq y_i \leq \hat{\tilde{Y}}_i^{U'}, \\ 0 & \text{if } y_i < \hat{\tilde{Y}}_i^{L'}, \\ 0 & \text{if } y_i > \hat{\tilde{Y}}_i^{U'}, \end{cases} \tag{4}$$

where, y_i is the center of fuzzy number $\tilde{y}_i$.

Therefore the minimization problem for the proposed method, can be formulated as follows

$$\min_{\hat{b}, \hat{\mathbf{w}}, \hat{l}_\delta, \xi_i, \xi_i^*} \frac{||\hat{\mathbf{w}}||^2}{4\hat{l}_\delta^2} + C\sum_{i=1}^n (\xi_i + \xi_i^*),$$

$$s.t.$$

$$\begin{cases} \omega_i(\hat{\tilde{Y}}_i^L - \tilde{y}_i^L) \leq \xi_i, \\ \omega_i(\hat{\tilde{Y}}_i^U - \tilde{y}_i^U) \geq -\xi_i^*, \\ 0 \leq \hat{l}_\delta \leq max(\mathbf{l_y}), \\ \xi_i \geq 0, \ \xi_i^* \geq 0, \ i = 1, 2, ..., n, \end{cases} \tag{5}$$

where, $\omega_i \in \{0, 1\}$ $(i = 1, 2, ..., n)$ denotes the weight assigned to the ith observation, determind via Eq. (4) to effectively mitigate the influence of outliers (4).

The following theorem, establishes how to estimate the parameters of the SVWFR model by solving minimization problem (5).

Theorem 3.1. *Let $\{\mathbf{x_i}, \tilde{y}_i\}$, $i = 1, 2, ..., n$ denote a fuzzy training set, where $\tilde{y}_i = (y_i; l_{y_i})_T$. The corresponding estimated fuzzy responses are*

$$\hat{\tilde{Y}}_i = (\hat{b} + \hat{\mathbf{w}}\mathbf{x_i}; \hat{l}_\delta)_T.$$

Then, based on the Lagrange multipliers α_{1i}, α_{2i}, p_i, p_i^*, η_1, and η_2, the parameters of model (1) are obtained as follows

$$\begin{cases} \hat{\mathbf{w}} = \frac{\sum_{i=1}^{n}(\alpha_{2i}-\alpha_{1i})\omega_i\mathbf{x_i}}{2}\Big(\frac{\sum_{i=1}^{n}(\alpha_{1i}+\alpha_{2i})+\eta_1-\eta_2}{\sum_{j=1}^{n}\sum_{i=1}^{n}(\alpha_{2i}-\alpha_{1i})(\alpha_{2j}-\alpha_{1j})\omega_i\omega_j\mathbf{x_i}\mathbf{x_j}}\Big)^2, \\ \hat{l}_\delta = \frac{-\sum_{i=1}^{n}(\alpha_{1i}+\alpha_{2i})-\eta_1+\eta_2}{2\sum_{j=1}^{n}\sum_{i=1}^{n}(\alpha_{2i}-\alpha_{1i})(\alpha_{2j}-\alpha_{1j})\omega_i\omega_j\mathbf{x_i}\mathbf{x_j}}, \\ \hat{b} = \frac{-1}{2}\big(\hat{\mathbf{w}}\omega_i\mathbf{x_i} + \hat{\mathbf{w}}\omega_j\mathbf{x_j} - \omega_i(y_i - l_{y_i}) - \omega_j(y_j + l_{y_j})\big), \end{cases} \tag{6}$$

for some values α_{1i}, $\alpha_{2j} \in [0, C]$.

Proof. 1) First, we find the solution to the minimization problem in the dual variables by finding the saddle point of the Lagrangian function as

$$\begin{aligned} L = \frac{||\hat{\mathbf{w}}||^2}{4\hat{l}_\delta^2} &+ C\sum_{i=1}^{n}(\xi_i + \xi_i^*) - \sum_{i=1}^{n}\alpha_{1i}\omega_i(y_i - l_{y_i} - \hat{b} - \hat{\mathbf{w}}\mathbf{x_i} + \hat{l}_\delta + \xi_i) \\ &- \sum_{i=1}^{n}\alpha_{2i}\omega_i(\hat{b} + \hat{\mathbf{w}}\mathbf{x_i} + \hat{l}_\delta - y_i - l_{y_i} + \xi_i^*) \\ &- \sum_{i=1}^{n}p_i\xi_i - \sum_{i=1}^{n}p_i^*\xi_i^* - \eta_1\hat{l}_\delta - \eta_2(max(\mathbf{l_y}) - \hat{l}_\delta), \end{aligned} \tag{7}$$

where, α_{1i}, α_{2i}, p_i, p_i^*, η_1, and η_2 are the Lagrange multipliers.
2) In the second step, differentiating the objecting function with respect to the parameters and setting the results to zero, yields

$$\begin{cases} \hat{\mathbf{w}} = \frac{\sum_{i=1}^{n}(\alpha_{2i}-\alpha_{1i})\omega_i\mathbf{x_i}}{2}\Big(\frac{\sum_{i=1}^{n}(\alpha_{1i}+\alpha_{2i})+\eta_1-\eta_2}{\sum_{j=1}^{n}\sum_{i=1}^{n}(\alpha_{2i}-\alpha_{1i})(\alpha_{2j}-\alpha_{1j})\omega_i\omega_j\mathbf{x_i}\mathbf{x_j}}\Big)^2, \\ \hat{l}_\delta = \frac{-\sum_{i=1}^{n}(\alpha_{1i}+\alpha_{2i})-\eta_1+\eta_2}{2\sum_{j=1}^{n}\sum_{i=1}^{n}(\alpha_{2i}-\alpha_{1i})(\alpha_{2j}-\alpha_{1j})\omega_i\omega_j\mathbf{x_i}\mathbf{x_j}}, \\ \sum_{i=1}^{n}(\alpha_{1i} - \alpha_{2i}) = 0, \\ \alpha_{1i} = C - p_i \Rightarrow \alpha_{1i} \le C, \\ \alpha_{2i} = C - p_i^* \Rightarrow \alpha_{2i} \le C. \end{cases} \tag{8}$$

The values α_{1i} and α_{2i} can be determined by solving the following dual problem, which is obtained by substituting (8) into (7) (we used "Mathematica" software [30] to do this)

$$\begin{aligned} &\max_{\alpha_{1i},\alpha_{2i}} \\ &L^* = \frac{10\big(\sum_{i=1}^{n}(\alpha_{1i}+\alpha_{2i})+\eta_1-\eta_2\big)^2}{\sum_{j=1}^{n}\sum_{i=1}^{n}(\alpha_{1i}-\alpha_{2i})(\alpha_{1j}-\alpha_{2j})\omega_i\omega_j\mathbf{x_i}\mathbf{x_j}} - \sum_{i=1}^{n}(\alpha_{1i}-\alpha_{2i})\omega_i y_i - \sum_{i=1}^{n}(\alpha_{1i}+\alpha_{2i})\omega_i l_{y_i} - \eta_2\max(\mathbf{l_y}), \\ &\text{s.t.} \\ &\sum_{i=1}^{n}(\alpha_{1i}-\alpha_{2i}) = 0, \quad \alpha_{1i},\ \alpha_{2i} \in [0, C]. \end{aligned} \tag{9}$$

3) Finally, determine $\hat{b}$, we apply the Karush–Kuhn–Tucker (KKT) conditions [23] to Eqs. (7) and (9), which yields

$$\hat{b} = \frac{-1}{2}\big(\hat{\mathbf{w}}\omega_i\mathbf{x_i} + \hat{\mathbf{w}}\omega_j\mathbf{x_j} - \omega_i(y_i - l_{y_i}) - \omega_j(y_j + l_{y_j})\big),$$

for some $\alpha_{1i}, \alpha_{2j} \in [0, C]$. This completes the proof. □

Remark 3.2. *The SVWFR estimation method is reduced to the ordinary SVFR estimation method if we put $\omega_i = 1$ for $i = 1, 2, ..., n$.*

Remark 3.3. *The proposed SVWFR model includes a regularization parameter C. To select its optimal value, we employed cross-validation combined with a grid search over an exponentially spaced range, specifically from*

$$\{10, 25, 50, 100, 500, 10^3, 10^4, 10^5\}.$$

We then select the parameter value that minimizes the following sum of squared scores (SSS) that between $\tilde{y}_i$ and $\hat{\tilde{Y}}_{(-i)}$

$$SSS = \frac{1}{2}\sum_{i=1}^{n}(l_{y_i} - \hat{l}_{Y_{(-i)}})^2 + \sum_{i=1}^{n}(y_i - \hat{Y}_{(-i)})^2, \tag{10}$$

where, $\hat{\tilde{Y}}_{(-i)} = (\hat{Y}_{(-i)}, \hat{l}_{Y_{(-i)}})_T$, $i = 1, 2, ..., n$ is the predicted fuzzy response value after leaving the ith observation from the data set.

The algorithm of the proposed SVWFR is summarized as follows.

Algorithm 1. Computational algorithm for the fuzzy linear regression model (1).

Input: Given fuzzy data set $\{\mathbf{x_i}, \tilde{y}_i\}$, $i = 1, 2, ..., n$.
Step 1. Determine the optimal parameter C of SVFR by using an exponential grid search from $\{10, 25, 50, 100, 500, 10^3, 10^4, 10^5\}$.
Step 2. Compute the estimated fuzzy responses $\hat{\tilde{Y}}_i$ of SVFR with the resulting optimal parameter C, by Theorem 3.1, for $\omega = 1$.
Step 3. Use Equation (4) as a weight for ith observation.
Step 4. Determine the new optimal parameter C of SVWFR by using an exponential grid search from $\{10, 25, 50, 100, 500, 10^3, 10^4, 10^5\}$.
Step 5. Compute the estimated fuzzy responses $\hat{\tilde{Y}}_i$ of SVWFR with resulting optimal parameter C, by Theorem 3.1.
Output: Estimated fuzzy responses $\hat{\tilde{Y}}_i = (\hat{Y}_i, \hat{l}_{Y_i})_T$.

---

4 Nonlinear support vector weighted fuzzy regression (Nonlinear SVWFR)

Now, we treat nonlinear SVWFR without assuming any underlying model function. This is achieved by mapping $\mathbf{x_i} \to \phi(\mathbf{x_i})$ to embed the original R^n features into a Hilbert feature space F, $\phi : R^n \to F$ with nonlinear kernel $k_h(\mathbf{x_i}, \mathbf{x_j}) = \phi(\mathbf{x_i})\phi(\mathbf{x_j})$. Then, by applying the SVR algorithm, we obtain, the following fuzzy nonlinear regression model

$$\tilde{Y} = (b + \mathbf{w}\phi(\mathbf{x_i}); l_\delta)_T.$$

where, b is the bias term, $\mathbf{w} = (w_1, w_2, ..., w_p)$ is the weight vector, and $\tilde{\delta} = (0; l_\delta)_T$ is the error term.

In the following theorem, we show how to obtain estimates for the parameters of the nonlinear SVWFR based on optimization problem (5). The proof is similar to that of Theorem 3.1, and is therefore omitted.

Theorem 4.1. *Suppose that $\{\mathbf{x_i}, \tilde{y}_i\}$, $i = 1, 2, ..., n$ is a set of fuzzy training set, where, $\tilde{\mathbf{y}}_\mathbf{i} = (\mathbf{y_i}; \mathbf{l_{y_i}})_T$. Also, assume that the values of estimated fuzzy responses are*

$$\hat{\tilde{Y}}_i = (\hat{b} + \hat{\mathbf{w}}\phi(\mathbf{x_i}); \hat{l}_\delta)_T. \tag{11}$$

Then, based on the Lagrange multipliers α_{1i}, α_{2i}, p_i, p_i^*, η_1, and η_2, the parameters in (11) are obtained as follows

$$\begin{cases} \hat{\mathbf{w}} = \frac{\sum_{i=1}^{n}(\alpha_{2i}-\alpha_{1i})\omega_i\phi(\mathbf{x_i})}{2}\Big(\frac{\sum_{i=1}^{n}(\alpha_{1i}+\alpha_{2i})+\eta_1-\eta_2}{\sum_{j=1}^{n}\sum_{i=1}^{n}(\alpha_{2i}-\alpha_{1i})(\alpha_{2j}-\alpha_{1j})\omega_i\omega_j k_h(\mathbf{x_i},\mathbf{x_j})}\Big)^2, \\ \hat{l}_\delta = \frac{-\sum_{i=1}^{n}(\alpha_{1i}+\alpha_{2i})-\eta_1+\eta_2}{\sum_{j=1}^{n}\sum_{i=1}^{n}(\alpha_{2i}-\alpha_{1i})(\alpha_{2j}-\alpha_{1j})\omega_i\omega_j k_h(\mathbf{x_i},\mathbf{x_j})}, \\ \hat{b} = \frac{-1}{2}\big(\hat{\mathbf{w}}\omega_i\phi(\mathbf{x_i}) + \hat{\mathbf{w}}\omega_j\phi(\mathbf{x_j}) - \omega_i(y_i - l_{y_i}) - \omega_j(y_j + l_{y_j})\big), \end{cases} \tag{12}$$

for some α_{1i}, $\alpha_{2j} \in [0, C]$.

Remark 4.2. *Based on the parameters given in equation* (12), *the estimation of the response variable is obtained as*

$$\hat{\tilde{Y}}_i = \Bigg(\hat{b} + 8\sum_{j=1}^{n}(\alpha_{2j}-\alpha_{1j})\omega_j k_h(\mathbf{x_i},\mathbf{x_j})\big(\frac{\sum_{i=1}^{n}(\alpha_{1i}+\alpha_{2i})+\eta_1-\eta_2}{\sum_{j=1}^{n}\sum_{i=1}^{n}(\alpha_{1i}-\alpha_{2i})(\alpha_{1j}-\alpha_{2j})\omega_i\omega_j k_h(\mathbf{x_i},\mathbf{x_j})}\big)^2;$$
$$\frac{2\sum_{i=1}^{n}(\alpha_{1i}+\alpha_{2i})+\eta_1-\eta_2}{\sum_{j=1}^{n}\sum_{i=1}^{n}(\alpha_{1i}-\alpha_{2i})(\alpha_{1j}-\alpha_{2j})\omega_i\omega_j k_h(\mathbf{x_i},\mathbf{x_j})}\Bigg)_T.$$

Remark 4.3. *In order to implement the above calibrated procedure of nonlinear SVWFR, we have to determine the kernel function $k_h(.)$. Some common kernel functions are as follows:*

$$\begin{aligned}
&\textit{Gaussian kernel,} && k_h(\mathbf{x_i},\mathbf{x_j}) = \frac{1}{\sqrt{2\pi}}\exp\Big(-\frac{t^2}{2h^2}\Big), \\
&\textit{Epanecknikov kernel,} && k_h(\mathbf{x_i},\mathbf{x_j}) = \frac{3}{4}(1-\frac{t^2}{h^2}), \qquad |t/h| \le 1, \\
&\textit{Triweight kernel,} && k_h(\mathbf{x_i},\mathbf{x_j}) = \frac{35}{32}(1-\frac{t^2}{h^2})^3, \qquad |t/h| \le 1.
\end{aligned}$$

where, $t^2 = \sum_{k=1}^{p}(x_{ki}-x_{kj})^2$. Note that we use a cross-validation method as expressed in Remark 3.3 for optimal choice of smoothing parameter h. The values of $\{0.01, 0.1, 0.5, 1, 3, 5, 10, 25, 50, 100, 500, 1000\}$ were considered for the smoothing parameter h. Then, we select the value of h which minimizes the SSS function in (10).

4.1 Model evaluation criteria

Here, we recall some well-known criteria to evaluate the proposed models.

Definition 4.4. *Suppose that $\tilde{y}_i$ and $\hat{\tilde{Y}}_i$ are the values of the observed and estimated fuzzy responses. Based on the distance $d(\tilde{y}_i, \hat{\tilde{y}}_i)$, a measure for goodness of fit of the model is defined as*

$$I = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{1+d(\tilde{y}_i,\hat{\tilde{Y}}_i)},$$

where

$$d(\tilde{y}_i,\hat{\tilde{Y}}_i) = \frac{1}{3}\big(|y_i-\hat{Y}_i| + |(y_i-l_{y_i})-(\hat{Y}_i-\hat{l}_{Y_i})| + |(y_i+l_{y_i})-(\hat{Y}_i+\hat{l}_{Y_i})|\big).$$

Definition 4.5. [3] *Consider the assumptions in Definition 3. The mean of similarity measures (MSM) between $\tilde{y}_i$ and $\hat{\tilde{Y}}_i$, $i = 1, 2, ..., n$ are defined as*

$$MSM = \frac{1}{n}\sum_{i=1}^{n}SM(\tilde{y}_i,\hat{\tilde{Y}}_i),$$

where,

$$SM(\tilde{y}_i,\hat{\tilde{Y}}_i) = 1 - \frac{\int_R |\hat{\tilde{Y}}_i(x)-\tilde{y}_i(x)|\,dx}{\int_R \hat{\tilde{Y}}_i(x)dx + \int_R \tilde{y}_i(x)dx},$$

and $\hat{\tilde{Y}}_i(x)$ and $\tilde{y}_i(x)$ are the membership functions of $\hat{\tilde{Y}}_i$ and $\tilde{y}_i$, respectively.

Definition 4.6. *Suppose that $\tilde{y}_i = (y_i; l_{y_i})_T$ and $\hat{\tilde{Y}}_i = (\hat{Y}_i, \hat{l}_{Y_i})_T$, $i = 1, ..., n$, are the observed values and the estimated values of response variables, respectively. The measure of central tendency for all data is defined as*

$$\varphi_Y = \frac{1}{n}\sum_{i=1}^{n}(y_i-\hat{Y}_i)^2.$$

The values of indices I and MSM are in $[0, 1]$. The optimal model is the model with maximum values of I and MSM. On the other hand, The smaller the value of φ_Y, the better the central tendency of the fuzzy regression model.

5 Four numerical comparisons

In this section, we present four examples to demonstrate the effectiveness of the proposed linear and nonlinear SVWFR models.

Example 5.1. *Consider the data set given in Table 1 (Arefi [2]). In this data set, the observations of the independent variable are crisp numbers and the observations of the dependent variable are symmetric triangular fuzzy numbers. By using such data, we compare our proposed method with those of Hong and Hwang [18], Yongqi [32], Yao and Yu [31], and Khammar et al. [21].*

Based on the proposed method in Section 3, we fit a fuzzy regression model as follows

$$\hat{\tilde{Y}}_i = \hat{b} \oplus (\hat{w} \otimes x_i) \oplus \hat{\tilde{\delta}},$$

where, $\hat{\tilde{\delta}} = (0, \hat{l}_\delta)_T$, $i = 1, 2, ..., 10$. *Applying the proposed linear SVWFR approach to the given data set, we obtain the fuzzy regression model with* $C = 100$, *as*

$$\hat{\tilde{Y}}_i = 1.900 \oplus ((0.675\ x_i) \oplus (0, 1.000))_T,$$

where, the pairs $(\hat{b}, \hat{w}, \hat{l}_\delta)$ *are determined from Theorem 3.1.*

We now apply the proposed nonlinear SVWFR approach to the data in Table 1. Considering $C = 100$ *and Gaussian kernel with* $h = 1$, *we obtain the fuzzy nonlinear regression model with* $\hat{b} = 5.865$ *and* $\hat{l}_\delta = 1$. *The values of predicted response variables and the obtained weight value for each observation are reported in Table 3. The weight values for all points are* $\omega_i = 1$. *Thus, in data set there is no outlier.*

The results of optimal models are given in Table 2. The results show that the proposed nonlinear SVWFR model performs better than the best model (among the other four models) with 22 *percent increase in the index* MSM, 16 *percent increase in the index* I, *and* 39 *decrease in the index* φ_Y. *See Figure 2 for the performances of the fuzzy regression models considered in this example.*

Table 1: Data set in Example 5.1.

x_i	$\tilde{y}_i = (y_i, l_{y_i})_T$	x_i	$\tilde{y}_i = (y_i, l_{y_i})_T$
1	$(2.0, 0.2)_T$	6	$(6.0, 0.6)_T$
2	$(3.0, 0.3)_T$	7	$(6.0, 0.6)_T$
3	$(5.0, 0.5)_T$	8	$(7.0, 0.7)_T$
4	$(4.0, 0.4)_T$	9	$(8.0, 0.8)_T$
5	$(5.5, 0.6)_T$	10	$(10.0, 1.0)_T$

Table 2: Comparison results for numerical data in Table 1.

Model	Parameters	MSM	I	φ_Y
SVWFR with $C = 100$	$\hat{b} = 1.900$, $\hat{w} = 0.675$, $\hat{l}_\delta = 1.000$	0.454	0.654	0.427
Nonlinear SVWFR with $C = 100$, $h = 1$	$\hat{b} = 5.865$, $\hat{l}_\delta = 1.000$	0.525	0.718	0.235
Hong and Hwang [18]	$\hat{\tilde{b}} = (3.555, 0.111)_T$, $\hat{\tilde{w}} = (0.444, 0.088)_T$	0.367	0.619	1.266
Yao and Yu [31]	$\hat{\tilde{b}} = (0.550, 0.000)_T$, $\hat{\tilde{w}} = (1.010, 0.480)_T$	0.269	0.410	1.152
Yongqi [32]	$\hat{\tilde{b}} = (1.070, 0.000)_T$, $\hat{\tilde{w}} = (0.810, 0.280)_T$	0.431	0.563	0.427
Khammar et al. [21] with $h = 1.5$	$\hat{\tilde{b}} = (1.700, 0.950)_T$, $\hat{\tilde{w}} = (0.720, 0.217)_T$	0.361	0.460	0.387

Table 3: The values of predicted response variable $\hat{\tilde{Y}}_i$ and the obtained weight values based on our proposed models in Example 5.1.

	SVWFR with $C = 100$	Nonlinear SVWFR with $C = 100$	
$No.$	$\hat{\tilde{Y}}_i$	$\hat{\tilde{Y}}_i$	ω_i
1	$(2.57, 1.0)_T$	$(1.20, 1.0)_T$	1
2	$(3.25, 1.0)_T$	$(2.32, 1.0)_T$	1
3	$(3.93, 1.0)_T$	$(4.50, 1.0)_T$	1
4	$(4.60, 1.0)_T$	$(3.44, 1.0)_T$	1
5	$(5.27, 1.0)_T$	$(5.11, 1.0)_T$	1
6	$(5.95, 1.0)_T$	$(6.43, 1.0)_T$	1
7	$(6.62, 1.0)_T$	$(6.43, 1.0)_T$	1
8	$(7.30, 1.0)_T$	$(7.31, 1.0)_T$	1
9	$(7.97, 1.0)_T$	$(8.20, 1.0)_T$	1
10	$(8.65, 1.0)_T$	$(10.01, 1.0)_T$	1

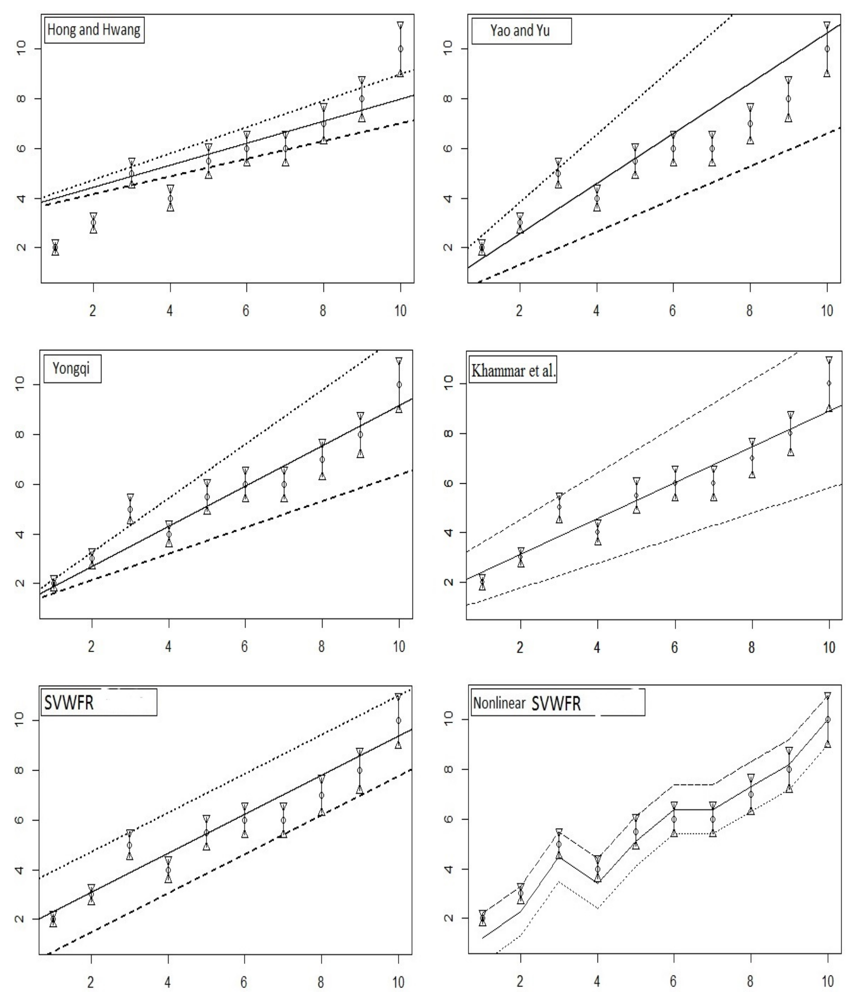

Figure 2: Fuzzy regression models in Example 5.1.

Example 5.2. *In this example, we use a crisp input-fuzzy output data provided by Gunn* [15] *(see Table 4). The* *proposed linear and nonlinear SVWFR approaches are applied to this data set. We compare our proposed models with the methods introduced by Kao and Chyu* [20]*, Hong and Hwang* [18]*, and Arabpour and Tata* [1]*.*

Thus, we fit a fuzzy regression model as follows

$$\hat{\tilde{Y}}_i = \hat{b} \oplus \big((\hat{w}x_i) \oplus \hat{\tilde{\delta}}\big),$$

where, $\hat{\tilde{\delta}} = (0, \hat{l}_\delta)_T$*,* $i = 1, 2, ..., 9$*. Applying the proposed linear SVWFR, we obtain the fuzzy regression model with* $C = 50$ *as (see Figure 3)*

$$\hat{\tilde{Y}}_i = -5.274 \oplus \big((1.210\ x_i) \oplus (0, 2.964)_T\big),$$

in which, the pairs $(\hat{b}, \hat{w}, \hat{l}_\delta)$ *are determined from Theorem 3.1.*
We now apply the proposed nonlinear SVWFR approach to the data in Table 4. In this case, with $C = 50$ *we obtain the fuzzy nonlinear regression model with parameters* $\hat{b} = 3.151$ *and* $\hat{l}_\delta = 1.000$ *(see Figure 3). For this data set, we use Gaussian kernel with* $h = 3$*. These parameters are determined concerning Theorem 4.1, Remark 4.3.*

The values of predicted response variables and the obtained weight value for each observation are reported in Table 5. The weight values for all points are $\omega_i = 1$*. Thus, there is no outliers in this data set.*

The results of optimal models are given in Table 6. The results show that the proposed nonlinear SVWFR model performs better than the best model (among the other four models) with 230 *percent increase in the index* MSM*,* 162 *percent increase in the index* I*, and* 94 *decrease in the index* φ_Y*. See Figure 3 illustrates the advantages of the fuzzy regression models examined in this example.*

Table 4: The data set in Example 5.2.

No.	x_i	$\tilde{y}_i$
1	1.0	$(-1.6, 0.5)_T$
2	3.0	$(-1.8, 0.5)_T$
3	4.0	$(-1.0, 0.5)_T$
4	5.6	$(1.2, 0.5)_T$
5	7.8	$(2.2, 1.0)_T$
6	10.2	$(6.8, 1.0)_T$
7	11.0	$(10.0, 1.0)_T$
8	11.5	$(10.0, 1.0)_T$
9	12.7	$(10.0, 1.0)_T$

Table 5: The values of predicted response variable $\hat{\tilde{Y}}_i$ and the obtained weight values for each observation in Example 5.2.

	SVWFR with $C = 50$	Nonlinear SVWFR with $C = 50$, and $h = 3$	
No.	$\hat{\tilde{Y}}_i$	$\hat{\tilde{Y}}_i$	ω_i
1	$(-4.064, 2.964)_T$	$(-1.100, 1.000)_T$	1
2	$(-1.644, 2.964)_T$	$(-1.300, 1.000)_T$	1
3	$(-0.434, 2.964)_T$	$(-0.551, 1.000)_T$	1
4	$(1.502, 2.964)_T$	$(1.699, 1.000)_T$	1
5	$(4.164, 2.964)_T$	$(2.200, 1.000)_T$	1
6	$7.068, 2.964)_T$	$(6.800, 1.000)_T$	1
7	$(8.036, 2.964)_T$	$(10.000, 1.000)_T$	1
8	$(8.641, 2.964)_T$	$(10.000, 1.000)_T$	1
9	$(10.093, 2.964)_T$	$(10.000, 1.000)_T$	1

Table 6: Comparison results for different models in Example 5.2.

Model	Parameters	MSM	I	φ_Y
SVWFR with $C = 50$	$\hat{b} = -5.274$, $\hat{w} = 1.210$, $\hat{l}_\delta = 2.964$	0.315	0.362	1.794
Nonlinear SVWFR with $C = 50$, $h = 3$	$\hat{b} = 3.151$, $\hat{l}_\delta = 1.000$	0.758	0.853	0.106
Kao and Chyu [20]	$\hat{b} = -4.895$, $\hat{w} = 1.195$, $\hat{l}_\delta = 1.800$	0.330	0.456	1.718
Hong and Hwang [18] with $C = 500$	$\hat{b} = -5.451$, $\hat{\tilde{w}} = (1.217, 0.508)_T$	0.289	0.321	1.882
Arabpour and Tata [1]	$\hat{\tilde{b}} = (-4.895, 0.354)_T$, $\hat{\tilde{w}} = (1.195, 0.057)_T$	0.227	0.526	1.718

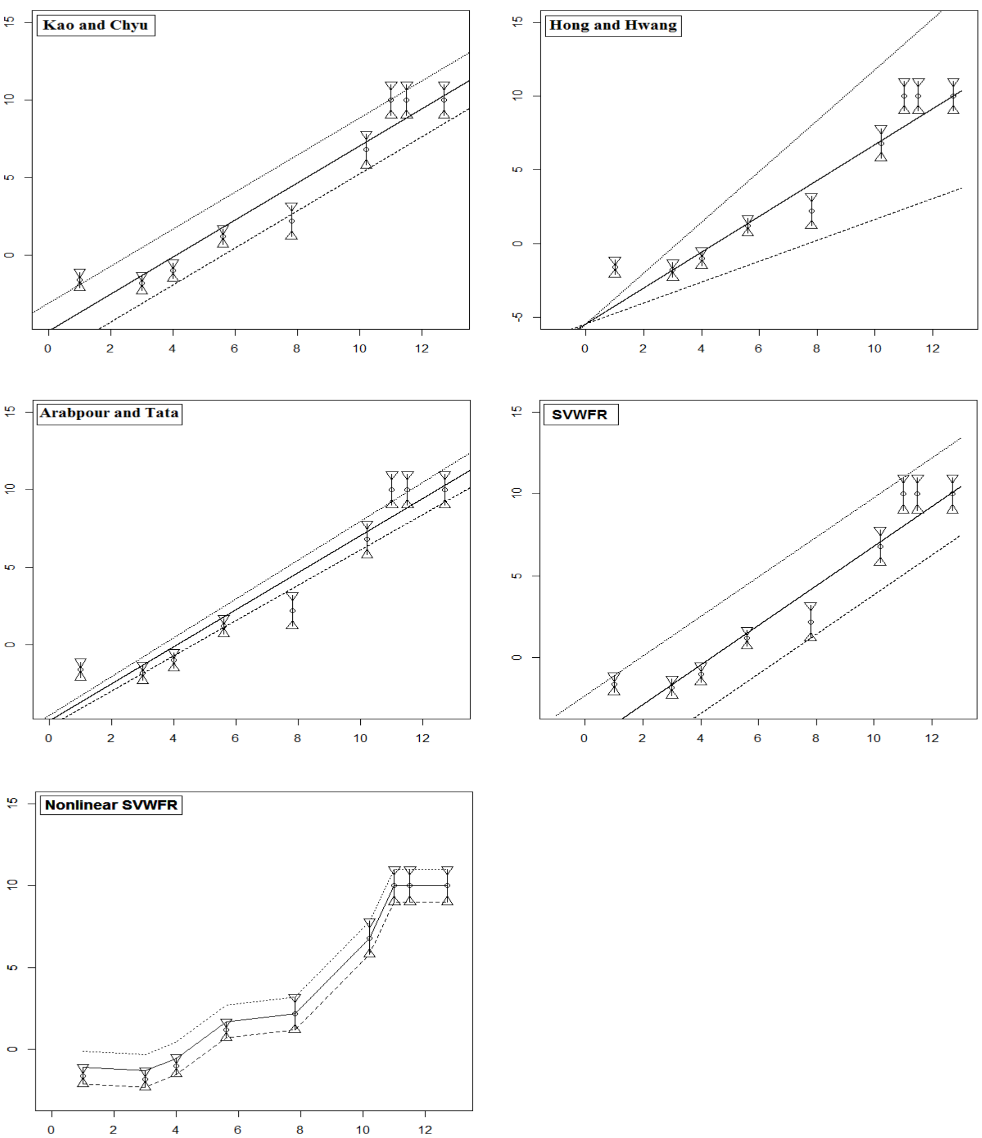

Figure 3: Fuzzy regression models in Example 5.2.

Example 5.3. *(Application in textile engineering) Data set in Table 7 shows a coloration process in loom industrial (see Tavanai et al.* [28]*). The variables* x_1*,* x_2 *and* x_3 *(as the explanatory/independent variables) are the color density(g/l), the time of process(m) and temperature (*^0C*), respectively, and* y' *represents the value of color suction. Because of some impreciseness in experimental conditions, the values of color suction are considered as triangular fuzzy numbers (as the response variable with the help of logarithm i.e.* $\tilde{Y} = (\log(y'), 0.1\log(y'))_T$*).*

We apply the linear and nonlinear SVWFR models to analyze this data set. For this, we assume that the fuzzy regression model is

$$\hat{\tilde{Y}}_i = \hat{b} \oplus \left((\hat{w}_1 x_{i1} + \hat{w}_2 x_{i2} + \hat{w}_3 x_{i3}) \oplus \hat{\tilde{\delta}}\right), \quad i = 1, 2, ..., 48.$$

Table 7: Data set in Example 5.3.

$No.$	x_{i1}	x_{i2}	x_{i3}	y_i'	$No.$	x_{i1}	x_{i2}	x_{i3}	y_i'
1	0.75	24.00	100	1.014	25	0.75	24.00	120	7.157
2	1.50	24.00	100	1.104	26	1.50	24.00	120	11.876
27	3.00	24.00	100	1.148	3	3.00	24.00	120	15.878
4	4.50	24.00	100	1.178	28	4.50	24.00	120	18.878
5	0.75	36.00	100	1.421	29	0.75	36.00	120	7.223
6	1.50	36.00	100	1.518	30	1.50	36.00	120	13.697
7	3.00	36.00	100	1.651	31	3.00	36.00	120	20.012
8	4.50	36.00	100	1.741	32	4.50	36.00	120	17.189
9	0.75	48.00	100	1.610	33	0.75	48.00	120	7.878
10	1.50	48.00	100	1.790	34	1.50	48.00	120	12.547
11	3.00	48.00	100	1.928	35	3.00	48.00	120	19.597
12	4.50	48.00	100	1.867	36	4.50	48.00	120	21.194
13	0.75	24.00	110	4.459	37	0.75	24.00	130	8.969
14	1.50	24.00	110	4.799	38	1.50	24.00	130	16.833
15	3.00	24.00	110	5.023	39	3.00	24.00	130	24.352
16	4.50	24.00	110	5.422	40	4.50	24.00	130	27.678
17	0.75	36.00	110	5.797	41	0.75	36.00	130	10.198
18	1.50	36.00	110	4.974	42	1.50	36.00	130	18.605
19	3.00	36.00	110	6.025	43	3.00	36.00	130	27.130
20	4.50	36.00	110	6.687	44	4.50	36.00	130	29.496
21	0.75	48.00	110	5.268	45	0.75	48.00	130	10.858
22	1.50	48.00	110	6.702	46	1.50	48.00	130	18.598
23	3.00	48.00	110	7.325	47	3.00	48.00	130	25.676
24	4.50	48.00	110	7.288	48	4.50	48.00	130	26.257

Based on our proposed SVWFR method, the optimal fuzzy linear regression model is fitted for $C = 100$ as follows

$$\hat{\tilde{Y}}_i = -7.785 \oplus \left((0.136 x_{i1} + 0.008 x_{i2} + 0.079 x_{i3}) \oplus (0, 0.338)_T\right),$$

where, the values of $(\hat{b}, \hat{w}_1, \hat{w}_2, \hat{w}_3, \hat{l}_\delta)$ are estimated based on Theorem 3.1.

Moreover, based on our proposed nonlinear SVWFR with Gaussian kernel, the optimal fuzzy nonlinear regression model is obtained for $C = 100$, $\hat{b} = -3.012$, $\hat{l}_\delta = 0.454$ and $h = 1$. The results show that, the values of MSM, I, and φ_Y indices for the nonlinear SVWFR model (with parameters $C = 100$ and $h = 1$) are better than those of the linear SVWFR model (Table 8). For instance, the performances of the linear and nonlinear SVWFR models are shown in Figure 4, where the values of $\hat{y}_i$s in nonlinear SVWFR model are more closer to the values y_is for $i = 1, 2, ..., 48$.

Table 8: Comparison results for linear and nonlinear models in Example 5.3.

Model	Parameters	MSM	I	φ_Y
SVWFR with $C = 100$	$\hat{b} = -7.785$, $\hat{w}_1 = 0.136$, $\hat{w}_2 = 0.008$, $\hat{w}_3 = 0.079$, $\hat{l}_\delta = 0.338$	0.236	0.787	0.113
Nonlinear SVWFR with $C = 100$, $h = 1$	$\hat{b} = -3.012$ $\hat{l}_\delta = 0.454$	0.422	0.866	0.036

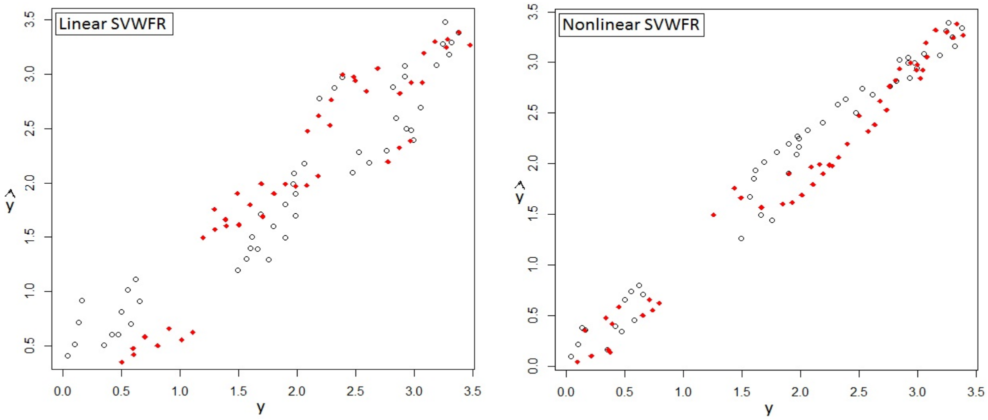

Figure 4: Plots of values $\hat{y}_i$s (red points) and values y_is (black points) for the proposed linear and nonlinear SVWFR models based in Example 5.3.

5.1 Sensitivity analysis of nonlinear SVWFR parameters

In this section, we investigate the sensitivity of the nonlinear SVWFR model parameters to variations in the hyperparameters C and h.

Example 5.4. *Consider the data set in Example 5.2. We check the sensitivity analysis in the proposed nonlinear SVWFR with Gaussian kernel based on the different values of*

$$C \in \{10, 25, 50, 100, 500, 10^3, 10^4, 10^5\} \text{ and } h \in \{0.01, 0.1, 0.5, 1, 3, 5, 25, 50, 100, 500, 1000\}.$$

Based on the results in Table 9, for a fixed value of C, the values of φ_Y decreases and the values of I increases when h increases (see Figures 5). Hence, The best performances for the proposed nonlinear SVWFR are the values with $h \geq 1$ and $C \geq 25$. Note that for different values of C and h, the results are very similar, and therefore in this example, the nonlinear SVWFR regression model does not show much sensitivity to choice the values of C and h.

Table 9: Sensitivity analysis based on the different values of C and h in Example 5.4.

		$h = 0.01$	$h = 0.10$	$h = 0.50$	$h = 1$	$h > 1$
$C = 10$	I	0.647	0.772	0.846	0.851	0.851
	φ_Y	0.863	0.181	0.104	0.108	0.111
$C = 25$	I	0.645	0.785	0.851	0.851	0.851
	φ_Y	0.843	0.165	0.118	0.111	0.111
$C = 50$	I	0.643	0.784	0.846	0.851	0.851
	φ_Y	0.832	0.784	0.158	0.111	0.111
$C = 100$	I	0.636	0.796	0.851	0.851	0.851
	φ_Y	0.822	0.166	0.111	0.111	0.111
$C = 500$	I	0.626	0.806	0.851	0.851	0.851
	φ_Y	0.788	0.181	0.111	0.111	0.111
$C = 1000$	I	0.625	0.790	0.851	0.851	0.851
	φ_Y	0.779	0.168	0.111	0.111	0.111
$C = 10^4$	I	0.624	0.799	0.851	0.851	0.851
	φ_Y	0.693	0.149	0.111	0.111	0.111
$C = 10^5$	I	0.621	0.810	0.851	0.851	0.851
	φ_Y	0.687	0.140	0.111	0.111	0.111

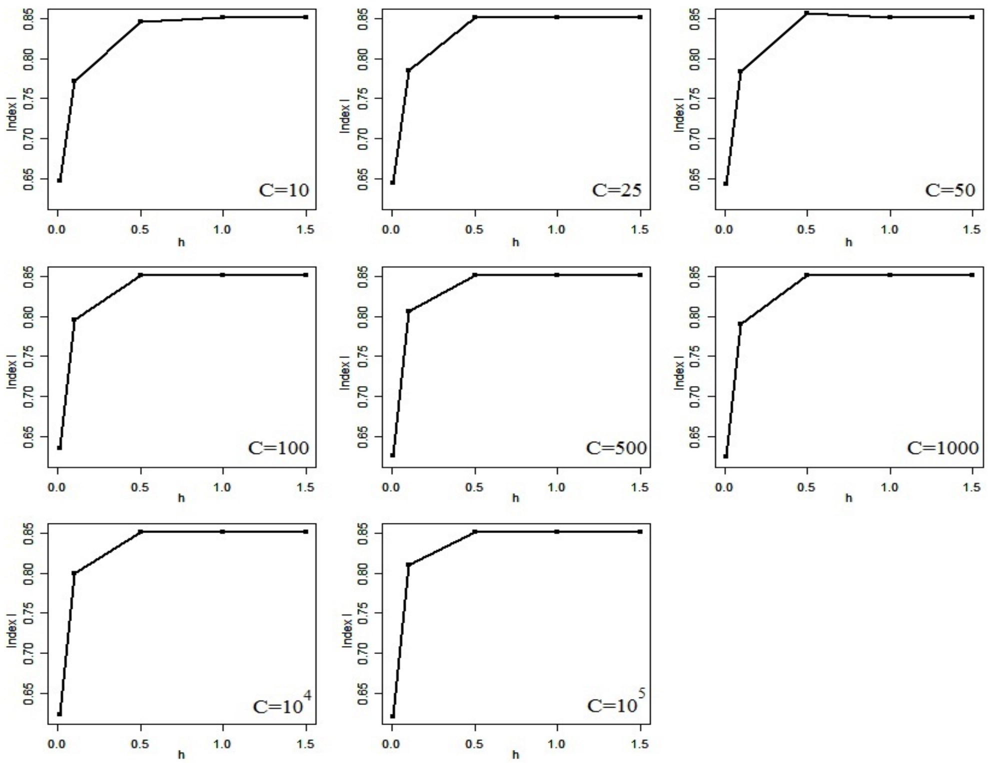

Figure 5: Sensitivity analysis in Example 5.4.

6 Robustness of the proposed model

In this section, we present an example to demonstrate the outlier detection capability of the approaches proposed in the preceding sections.

Example 6.1. *Consider the data set given in Table 4. Based on Figure 3, there is no outlier in this data set. To examine the robustness of the presented models, we create two data sets with changes in some centers and spreads of data (see Table 10). In anomalous data set 1, we change the centers of two data* $\tilde{y}_5$ *and* $\tilde{y}_8$ *as* $(20.2, 1.0)_T$ *and* $(-10, 1.0)_T$*, respectively. Also, in anomalous data set 2, we change the spreads of two data* $\tilde{y}_3$ *and* $\tilde{y}_7$ *as* $(-1, 5)_T$ *and* $(10, 10)_T$*, respectively.*

Using relation (4)*, the weight values are calculated for anomalous data sets (see Table 10). Based on the weights* $\omega_5 = \omega_8 = 0$ *in anomalous data set 1 and* $\omega_3 = \omega_7 = 0$ *in anomalous data set 2, these points are outliers. We also calculate the fuzzy regression models for anomalous data sets 1 and 2. The results are summerized in Table 11. Based on the goodness of fit indics, we can observe that:*

1) *Based on all goodness of fit indices, the proposed nonlinear SVWFR model has a suitable performance rather than the linear SVWFR model.*

2) *The proposed linear and nonlinear SVWFR models are robust with the presence of different types of outliers. Note that the estimated parameters of models with original data set and also with the anomalous data sets are approximately similar.*

Table 10: Fuzzy output data set and the obtained weight value for each observation in presence of outlier in Example 6.1.

	Original output data	Anomalous output data 1	Weights	Anomalous output data 2	Weights
$No.$	$\tilde{y}_i$	$\tilde{y}_i$	ω_i	$\tilde{y}_i$	ω_i
1	$(-1.6, 0.5)_T$	$(-1.6, 0.5)_T$	1	$(-1.6, 0.5)_T$	1
2	$(-1.8, 0.5)_T$	$(-1.8, 0.5)_T$	1	$(-1.8, 0.5)_T$	1
3	$(-1.0, 0.5)_T$	$(-1.0, 0.5)_T$	1	$(-1.0, 5.0)_T$	0
4	$(1.2, 0.5)_T$	$(1.2, 0.5)_T$	1	$(1.2, 0.5)_T$	1
5	$(2.2, 1.0)_T$	$(20.2, 1.0)_T$	0	$(2.2, 1.0)_T$	1
6	$(6.8, 1.0)_T$	$(6.8, 1.0)_T$	1	$(6.8, 1.0)_T$	1
7	$(10, 1.0)_T$	$(10, 1.0)_T$	1	$(10, 10)_T$	0
8	$(10, 1.0)_T$	$(-10, 1.0)_T$	0	$(10, 1.0)_T$	1
9	$(10, 1.0)_T$	$(10, 1.0)_T$	1	$(10, 1.0)_T$	1

Table 11: Comparison of results for anomalous data sets in Example 6.1.

Data	Model	Parameters	MSM	I	φ_Y
Anomalous data set 1	SVWFR with $C = 1000$	$\hat{b} = -4.846$, $\hat{w} = 1.169$, $\hat{l}_\delta = 1.001$	0.318	0.493	67.673
	Nonlinear SVWFR with $C = 1000$, $h = 3$	$\hat{b} = 3.536$, $\hat{l}_\delta = 1.000$	0.531	0.641	73.314
Anomalous data set 2	SVWFR with $C = 100$	$\hat{b} = -5.020$, $\hat{w} = 1.152$, $\hat{l}_\delta = 2.768$	0.380	0.357	1.947
	Nonlinear SVWFR with $C = 100$, $h = 5$	$\hat{b} = 3.829$, $\hat{l}_\delta = 1.123$	0.515	0.597	0.954

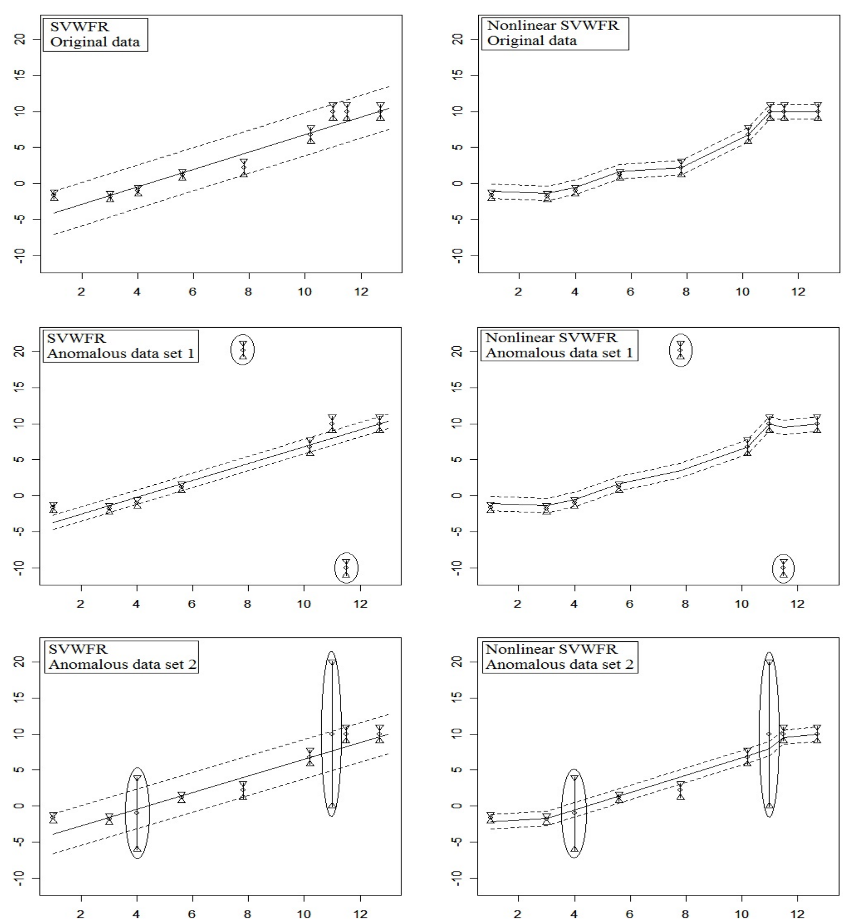

Figure 6: Robust Fuzzy regression models with anomalous data sets in Example 6.1.

7 Conclusions

A new approach was presented to fit a robust SVFR model, termed Support Vector Weighted Fuzzy Regression (SVWFR), for crisp inputs and fyzzy outputs. Furthermore, a nonlinear variant of SVWFR was proposed to capture nonlinear relationships. These approaches have the following key merits.

1) It is a new approach to fuzzy regression for crisp inputs-fuzzy output, and crisp parameters. In the proposed approach, a weighted operation is utilized to improve the robustness of SVFR by assigning weights to the support hyperplanes constraints.

2) We treat nonlinear SVWFR model for data without assuming any underlying function for the model.

3) By using different kernel functions, we can obtain different nonlinear fuzzy prediction models.

4) The proposed linear and nonlinear fuzzy regression models are robust with respect to the outlier data.

The proposed SVWFR model is marked by a set of novel contributions to the field of fuzzy regression and robust machine learning; i.e.:

- Integrated Robustness via Fuzzy Weighting: The introduction of binary weights $\omega_i \in \{0, 1\}$, determined through a median-based error envelope criterion, represents a fundamental shift from traditional penalty-based outlier suppression. It systematically filters anomalous observations before their influence enters the optimization space, making the model both resistant and interpretable in terms of robust decision rules.

- Hybridization of Fuzzy Set Theory and Structural Risk Minimization: By embedding fuzzy logic principles within the support vector machine framework, the SVWFR model advances a hybrid paradigm where epistemic uncertainty (fuzziness) and statistical learning principles are jointly leveraged.

- Extension to Nonlinear Spaces Without A Priori Model Specification: Through kernelization, the proposed nonlinear SVWFR enables effective learning in high-dimensional feature spaces without assuming a predefined functional form. This offers a model-free fuzzy regression solution, adaptable to complex nonlinear relationships.

Some methodological advances of the proposal are remarked below:

▷ Dual Optimization with Weighted Constraints: The derivation of the SVWFR parameters through dual optimization, integrating nonnegative Lagrange multipliers and a novel structure of slack variables constrained by weights, results in a generalized and modular learning algorithm. This allows for automatic relevance determination of observations in the learning process.

▷ Sequential Model Calibration: The proposed SVWFR adopts a two-phase model fitting strategyfirst calibrating a standard SVFR, and then computing weights ω_i based on the discrepancy between model and data using fuzzy α-cut differences. This algorithmic innovation introduces dynamic model feedback, enhancing robustness and empirical fit iteratively.

▷ Cross-validated Grid Search for Regularization and Smoothing Parameters: The estimation of critical parameters such as C (regularization) and h (kernel smoothing) is achieved via a systematic leave-one-out validation strategy using the sum of squares score (SSS). This enhances the generalization capability of the models and aligns with best practices in statistical learning.

The effectiveness -in term of empirical impact and practical value- of the SVWFR and its nonlinear extension has been rigorously validated across multiple numerical examples and real-world data, including:

☐ Synthetic datasets with controlled outliers (Examples 5.1 and 5.2).

☐ A textile engineering application involving process variables and fuzzy responses (Example 5.3).

☐ A stress-test simulation with artificially introduced anomalies (Example 6.1).

Across all cases, the proposed models consistently outperformed existing fuzzy regression approaches in terms of goodness-of-fit indices, while also successfully identifying and neutralizing outlier influence.

Importantly, the SVWFR approach provides interpretable weights that serve not only to increase robustness but also act as diagnostic tools for anomaly detection-offering dual utility in both prediction and insight.

In spite of advantages of our proposed models, such models seems to have a few limitations. First, the introduced weights is binary $\omega_i \in \{0, 1\}$. We can obtain the better results to introduce some methods with the soft weights $\omega_i \in [0, 1]$. Second, the calculations for large datasets and with more variables are heavy and the time to obtain the optimal estimations of paremetrs of SVWFR models is increased.

Future works: Building on the mentioned foundational advances of the proposed models, promising avenues for future research include:

1. Extending the SVWFR framework to handle fuzzy input-fuzzy output data structures.
2. Generalizing the weighting mechanism to soft (continuous) weights for finer granularity of robustness.
3. Incorporating interval-valued or intuitionistic fuzzy sets into the regression formulation.
4. Adapting the SVWFR framework for online or streaming fuzzy data environments.

Acknowledgement

The authors would like to thank referees for exact reading of the manuscript and for their valuable comments and suggestions which lead to improve the paper. This article was provided while the second author was on a research fellowship at the Sapienza-University of Rome, Italy. He would like to thank the university for hosting him and for its hospitality. Authors would like to acknowledge the financial support of **Iran National Science Foundation** for this research under grant number 4032212.

References

[1] A. R. Arabpour, M. Tata, *Estimating the parameters of a fuzzy linear regression model*, Iranian Journal of Fuzzy Systems, **5**(2) (2008), 1-19.

[2] M. Arefi, *Quantile fuzzy regression based on fuzzy outputs and fuzzy parameters*, Soft Computing, **24** (2020), 311-320. https://doi.org/10.1007/s00500-019-04424-2

[3] M. Arefi, S. M. Taheri, *Weighted similarity measure on interval-valued fuzzy sets and its application to pattern recognition*, Iranian Journal of Fuzzy Systems, **11**(5) (2014), 67-79.

[4] M. Asadollahi, M. G. Akbari, G. Hesamian, M. Arefi, *A robust support vector regression with exact predictors and fuzzy responses*, International Journal of Approximate Reasoning, **132** (2021), 206-225. https://doi.org/10.1016/j.ijar.2021.02.006

[5] J. Chachi, *A weighted least-squares fuzzy regression for crisp input-fuzzy output data*, IEEE Transactions on Fuzzy Systems, **27**(4) (2019), 739-748. https://doi.org/10.1109/TFUZZ.2018.2868554

[6] J. Chachi, S. M. Taheri, P. D'Urso, *Fuzzy regression analysis based on M-estimates*, Expert Systems with Applications, **187** (2022), 115891. https://doi.org/10.1016/j.eswa.2021.115891

[7] P. T. Chang, E. Stanley Lee, *A generalized fuzzy weighted least-squares regression*, Fuzzy Sets and Systems, **82** (1996), 289-298. https://doi.org/10.1016/0165-0114(95)00284-7

[8] P. Diamond, *Fuzzy least squares*, Information Sciences, **46**(3) (1988), 141-157. https://doi.org/10.1016/0020-0255(88)90047-3

[9] P. D'Urso, *Linear regression analysis for fuzzy/crisp input and fuzzy/crisp output data*, Computational Statistics and Data Analysis, **42**(1-2) (2003), 47-72. https://doi.org/10.1016/S0167-9473(02)00117-2

[10] P. D'Urso, J. Chachi, *OWA fuzzy regression*, International Journal of Approximate Reasoning, **142** (2022), 430-450. https://doi.org/10.1016/j.ijar.2021.12.009

[11] P. D'Urso, T. Gastaldi, *A least-squares approach to fuzzy linear regression analysis*, Computational Statistics and Data Analysis, **34**(4) (2000), 427-440. https://doi.org/10.1016/S0167-9473(99)00109-7

[12] P. D'Urso, R. Massari, *Weighted least squares and least median squares estimation for the fuzzy linear regression analysis*, Metron, **71** (2013), 279-306. https://doi.org/10.1007/s40300-013-0025-9

[13] P. D'Urso, R. Massari, A. Santoro, *Robust fuzzy regression analysis*, Information Sciences, **181**(19) (2011), 4154-4174. https://doi.org/10.1016/j.ins.2011.04.031

[14] B. Gu, W. S. Sheng, Z. Wang, D. Ho, S. Osman, S. Li, *Incremental learning for v-support vector regression*, Neural Networks, **67** (2015), 140-150. https://doi.org/10.1016/j.neunet.2015.03.013

[15] S. R. Gunn, *Support vector machines for classification and regression*, ISIS Technical Report, **14**(1) (1998), 5-16.

[16] P. Y. Hao, J. H. Chiang, *Fuzzy regression analysis by support vector learning approach*, IEEE Transactions on Fuzzy Systems, **16**(2) (2008), 428-441. https://doi.org/10.1109/TFUZZ.2007.896359

[17] T. Hastie, R. Tibshirani, J. Friedman, *The elements of statistical learning*, Springer New York, NY, (2008). https://doi.org/10.1007/978-0-387-84858-7

[18] D. H. Hong, C. Hwang, *Support vector fuzzy regression machines*, Fuzzy Sets and Systems, **138**(1) (2003), 271-281. https://doi.org/10.1016/S0165-0114(02)00514-6

[19] G. Jin, F. Z. Wen, W. Cheng, G. W. Ping, L. X. Ming, C. M. Zhen, *A minimum-of-maximum relative error support vector machine for simultaneous reverse prediction of concrete components*, Computers and Structures, **172** (2016), 59-70. https://doi.org/10.1016/j.compstruc.2016.05.003

[20] C. Kao, C. Chyu, *A fuzzy linear regression model with better explanatory power*, Fuzzy Sets and Systems, **126**(3) (2002), 401-409. https://doi.org/10.1016/S0165-0114(01)00069-0

[21] A. H. Khammar, M. Arefi, M. G. Akbari, *A robust least squares fuzzy regression model based on kernel function*, Iranian Journal of Fuzzy Systems, **17**(4) (2020), 105-119. https://doi.org/10.22111/ijfs.2020.5230

[22] A. H. Khammar, M. Arefi, M. G. Akbari, *A general approach to fuzzy regression models based on different loss functions*, Soft Computing, **25** (2021), 835-849. https://doi.org/10.1007/s00500-020-05441-2

[23] J. Luo, Y. Zheng, T. Hong, A. Luo, X. Yang, *Fuzzy support vector regressions for short-term load forecasting*, Fuzzy Optimization and Decision Making, **23** (2024), 363-385. https://doi.org/10.1007/s10700-024-09425-x+

[24] L. Nguyen, *Tutorial on support vector machine*, Applied and Computational Mathematics, **6**(4-1) (2017), 1-15. https://doi.org/10.11648/j.acm.s.2017060401.11

[25] Z. Qin, Q. Li, *An uncertain support vector machine with imprecise observations*, Fuzzy Optimization and Decision Making, **22** (2023), 611-629. https://doi.org/10.1007/s10700-022-09404-0

[26] H. Tanaka, S. Uejima, K. Asai, *Fuzzy linear regression model*, Proc. Int. Conf. Appl. Syst. Res. Cybern., Acapulco, Mexico, (1980), 12-15.

[27] H. Tanaka, S. Uejima, K. Asai, *Linear regression analysis with fuzzy model*, IEEE Transactions on Systems Man Cybernet, **12**(6) (1982), 903-907. https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=4308925

[28] H. Tavanai, S. M. Taheri, M. Nasiri, *Modeling of color yield in polyethylene terephthalate dyeing with statistical and fuzzy regression*, Iranian Polymer Journal, **14**(11) (2005), 954-968.

[29] V. N. Vapnik, *Statistical learning theory*, Springer New York, NY, (1998). https://doi.org/10.1007/978-1-4757-3264-1

[30] Wolfram Research, Inc., *Mathematica* 10.2, Wolfram Research, Inc. Champaign, Illinois, (2015).

[31] C. C. Yao, P. T. Yu, *Fuzzy regression based on asymmetric support vector machines*, Applied Mathematics and Computation, **182**(1) (2006), 175-193. https://doi.org/10.1016/j.amc.2006.03.046

[32] C. Yongqi, *Least squares support vector fuzzy regression*, Energy Procedia, **17**(Part A) (2012), 711-716. https://doi.org/10.1016/j.egypro.2012.02.160

[33] W. Zeng, Q. Feng, J. Li, *Fuzzy least absolute linear regression*, Applied Soft Computing, **52** (2017), 1009-1019. https://doi.org/10.1016/j.asoc.2016.09.029

[34] H. J. Zimmermann, *Fuzzy set theory and its applications*, 4th ed., Kluwer Nihoff, Boston, Springer Dordrecht, (2001). https://doi.org/10.1007/978-94-010-0646-0