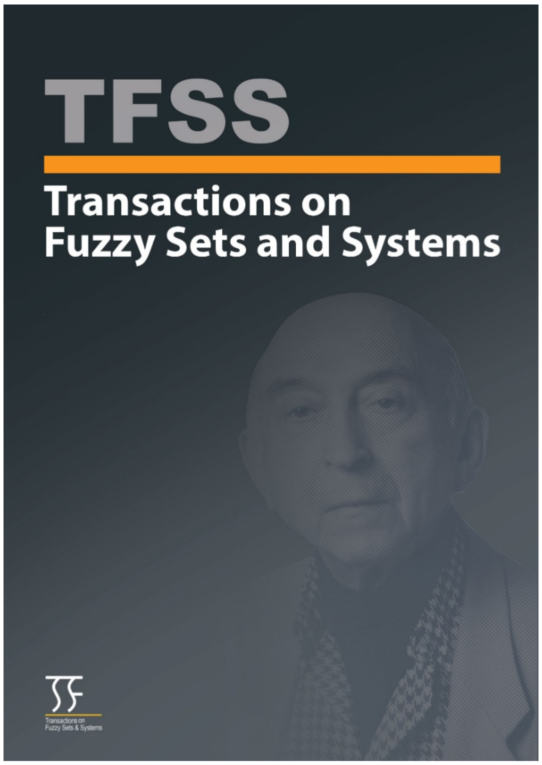
Transactions on Fuzzy Sets and Systems



Transactions on Fuzzy Sets and Systems

ISSN: **2821-0131**

https://sanad.iau.ir/journal/tfss/

Quantile-Based Fuzzy Semi-Parametric Support Vector Regression Model Based on Autoregressive Fuzzy Terms: An Estimation Approach

Vol.5, No.2, (2026), 59-84. DOI: https://doi.org/10.71602/tfss.2026.1202291

Author(s):

Batoul Salamah, Department of Applied Mathematics, Faculty of Mathematical Sciences, University of Guilan, Rasht, Iran. E-mail: batoul.salamah91@gmail.com

Reza Zarei, Department of Statistics, Faculty of Mathematical Sciences, University of Guilan, Rasht, Iran. E-mail: r.zarei@guilan.ac.ir, rezazarei.r@gmail.com

Farshid Mehrdoust, Department of Applied Mathematics, Faculty of Mathematical Sciences, University of Guilan, Rasht, Iran. E-mail: fmehrdoust@guilan.ac.ir

Mohammad Ghasem Akbari, Department of Statistics, Faculty of Mathematical Sciences, University of Birjand, Birjand, Iran. E-mail: g_z_akbari@birjand.ac.ir

Transactions on Fuzzy Sets and Systems (TFSS)
URL: https://sanad.iau.ir/journal/tfss/
Online ISSN: 2821-0131
Vol.5, No.2, (2026), 59-84
DOI: https://doi.org/10.71602/tfss.2026.1202291

Article Type: Original Research Article

Quantile-Based Fuzzy Semi-Parametric Support Vector Regression Model Based on Autoregressive Fuzzy Terms: An Estimation Approach

Batoul Salamah, Reza Zarei*, Farshid Mehrdoust, Mohammad Ghasem Akbari

Abstract. The semi-parametric regression model is one of the most useful statistical tools that has gained significant attention recently due to its key capability of combining both parametric and nonparametric features in one model. In practical applications, however, the recorded information or the relationship between one or more independent variables and the dependent variable is typically imprecise. Additionally, in certain situations, the error terms exhibit heteroscedasticity or the data distribution is skewed, which leads to inaccurate results when using conventional least squares models. In this regard, this paper introduces a semi-parametric quantile-based regression model using the support vector machine technique, along with precise regressors and fuzzy outcomes. We also employ the classic Durbin-Watson test to explore the existence of correlation among fuzzy residual expressions. We propose a mixed procedure that incorporates a mean absolute error and a cross-validation measure to calculate fuzzy multipliers in addition to the unknown autocorrelation criteria. We illustrate the efficacy of the suggested method via three numerical examples, including two applied examples and a simulation study. To this end, we use some common goodness-of-fit measurements to assess the performance of the suggested method in comparison with other approaches. The numerical results showed that when the fuzzy error terms are correlated, the suggested fuzzy semi-parametric quantile-based regression model performs better than the other methods.

AMS Subject Classification 2020: 62A86; 62J86
Keywords and Phrases: Fuzzy semi-parametric model, Support vector machine, Quantile regression, Fuzzy correlated errors, Goodness-of-fit measure, Kernel function.

1 Introduction

The most crucial method in statistics for evaluating and depicting the relationships between an outcome variable and multiple predictor variables is regression modeling. Regression models are also employed in many other fields, including economics, engineering, social sciences, ecological studies, and medicine. Generally speaking, there are two fundamental categories of regression techniques: parametric and nonparametric techniques. The choice of a regression approach, whether parametric or nonparametric, relies on the distribution of the stochastic error expression and the previous understanding of the functional statement for the relationship between the independent and dependent variables. Therefore, parametric regression is ideal for modeling data when the functional formula is known and correct. However, a large bias can occur when the functional formula of the relation is misdefined. Nonparametric regression models do not require prior assumptions about the distributions or a functional formula for the regression model. They also offer the

*Corresponding Author: Reza Zarei, Email: r.zarei@guilan.ac.ir, rezazarei.r@gmail.com, ORCID: 0000-0002-4636-4729
Received: 19 March 2025; **Revised**: 4 July 2025; **Accepted**: 5 July 2025; **Available Online**: 1 August 2025; **Published Online**: 7 November 2026.

How to cite: Salamah B, Zarei R, Mehrdoust F, Akbari MG. Quantile-based fuzzy semi-parametric support vector regression model based on autoregressive fuzzy terms: An estimation approach. *Transactions on Fuzzy Sets and Systems.* 2026; 5(2): 59-84. DOI: https://doi.org/10.71602/tfss.2026.1202291

flexibility to represent and analyze complex models. Non-parametric regression models can, however, occasionally lead to problems such as the curse of dimensionality, computational complexity, over-adjustment, and lack of extrapolation. Other regression models widely used recently include support vector regression (**SVR**) models. Using the principles of support vector machine (**SVM**) [1, 2, 3, 4, 5], **SVR** is a machine learning technique that fits the functional relation in regression analysis. **SVR** techniques are founded on the idea of minimizing constructional risk, where the total of the training error limits the generalization error. In contrast to classical regression models, which aim to estimate undetermined model parameters by reducing the observed training error, aims to attain overall performance by reducing the generalization error limit. In contrast to classical regression models, which aim to estimate undetermined model parameters by reducing the observed training error, **SVR** aims to attain overall performance by reducing the generalization error limit. Furthermore, **SVR** models have several advantages over conventional regression models, including robustness to outliers, efficiency in high-dimensional spaces, effective modeling of complex and nonlinear relationships, improved prediction accuracy by evaluating classification confidence, and lower computing power requirements than other regression approaches.

Recently, quantile regression has received widespread attention from both theoretical and practical perspectives. Quantile regression is a statistical method that aims to estimate the conditioned quantile functions of an outcome variable. Quantile regression is also a robust technique for understanding the influences of the entire distribution of predictor variables on a response variable. As a result, in contrast to conventional least squares regression techniques, which compute the conditional mean functions by minimizing the sum of the squares of the residuals, quantile regression methods depend on minimizing a weighted sum of absolute residuals to estimate specified conditional quantile functions. The quantile regression model has some advantages over traditional regression models, such as 1- robustness against outliers by focusing on quantiles rather than the mean. 2- Offering a deeper analysis by examining the full distribution of the outcome variable. 3- Not being based on postulations about the distribution of the dependent variable. 4- Unlike standard least squares, which assumes constant error variance, quantile regression effectively handles heteroscedasticity and provides accurate estimates, even when the variation of the outcome variable changes with the predictors [6, 7].

Although quantile regression models have many advantages and applications, classical quantile regression models are based on precise data as well as the exact relation between the predictors and the outcome variable. The assumption of a fuzzy relation among variables in the regression model is actually preferable when the phenomenon under investigation is imprecise or shows ambiguous variability instead of random variability [8]. Thus, using regression methods based on accurate data under the aforementioned conditions will lead to errors in the analysis and, as a result, to the presentation of an inappropriate model. Different approaches and theories have been proposed for treating uncertainty and imprecision during the past decades, among them the fuzzy set theory [9] have a key role and several researchers have concentrated on applying this theory to various fields, especially in probability and statistics [10, 11]. A type of classic regression modeling known as fuzzy regression modeling was first introduced by Tanaka et al. [12] in order to examine the functional relationship between the predictors and outcomes in a fuzzy setting [13]. After that, several methods were suggested for using the principles of fuzzy set theory in traditional regression frameworks. Naderkhani et al. [14] also proposed an adaptive neuro-fuzzy inference system ($ANFIS$), which aims to analyze and forecast the non-parametric fuzzy regression subject with precise input variables and a symmetrical trapezoidal fuzzy outcome variable. Chachi et al. [15] suggested the resistant M-estimation procedure to establish a fuzzy regression approach for precise predictors and a fuzzy response dataset, which provides consistent results when outliers are present. Yoon et al. [16] present a new method to improve fuzzy regression analysis by introducing an innovative fuzzy correlation metric as well as a distance-based variable selection technique. Notably, quantile regression techniques under fuzzy conditions outperform other fuzzy regression models because they simultaneously address uncertainty while offering robust and flexible modeling capabilities. Regarding the growing interest in this area, a recent review by Baha Alwan and Abdulmohsin Ali [17] provides a comprehensive overview of semi-parametric methodology for fuzzy quantile regression model estimation. However, few researchers have conducted studies on modeling quantile regression under imprecise conditions. Some studies conducted in this field include Hesamian and Akbari [18], who suggested a semi-parametric quantile-based regression procedure for cases with fuzzy regressors, a fuzzy smoothing function, precise multipliers, and fuzzy outcomes. Their manner depended upon a recent set of advanced kernel-driven signed-distance metrics in fuzzy number space. Arefi [19] proposed a regression model based on quantile estimation utilizing a fuzzy outcome variable and fuzzy multipliers, employing a new technique that integrates the loss function with fuzzy quantities. Also, Chachi and Chaji [20] suggested a method to predict the multipliers of the fuzzy regression procedure based on the quantile method with mathematical programming, relying upon some weights for sorted residues. Khammar et al. [21] introduced a quantile-based fuzzy varied parameter regression modeling that utilizes the quantile loss function and kernel function, specifically when the predictor variables and model constants are represented as fuzzy values. Hesamian and Akbari [22] examined a newly developed nonlinear quantile-based regression framework designed for cases where outcome variables are represented as triangular fuzzy quantities, while the regressors are precise data. Despite this, semi-parametric regression modeling is an extremely useful statistical tool, as it combines parametric and semi-parametric elements. Many studies have examined semi-parametric regression models. For example, Hesamian et al. [23] employed a semi-parametric partly linear method to improve traditional fuzzy linear regression approaches by incorporating fuzzy predictors, fuzzy outcomes, fuzzy smoothness function, and crisp multipliers. They also introduced a combined approach that utilizes curve-fitting algorithms and the least absolute errors to predict fuzzy smooth functions and imprecise parameters. Akbari and Hessamian [24] used a semi-parametric approach for fuzzy regressors and outcomes, extending the conventional adaptive grid multiple linear regression pattern. By combining kernel smoothing and adaptive grid regularization techniques, they developed a new variable selection approach within a fuzzy multiplex regression framework. As a result, they compute non-fuzzy parameters using a new robust technique and the ridge methodology. Notably, all the aforesaid fuzzy regression approaches depend upon uncorrelated error expressions. In various practical applications, fuzzy error terms are correlated due to common underlying influences or dependencies between data points. This, in turn, leads to incorrect results of the analysis and introduces the wrong regression model. Additionally, this contradicts the standard assumption that residuals are uncorrelated in the least squares or absolute deviation regression in a fuzzy situation. Hence, we must consider that the fuzzy error terms are correlated in some regression models.

The major objective of this paper is to present a recent quantile-based semi-parametric regression modeling with fuzzy outcomes and exact regressors, which utilizes a support vector regression technique to enhance the estimation process and improve prediction accuracy. Additionally, for this investigation, we expanded the proposed regression approach to include dependent error expressions rather than uncorrelated error expressions. To achieve this, we used a novel generalized subtraction technique to develop the semi-parametric **SVR** quantile formulation that is dependent on $LR{-}\mathbf{FN}s$ and exact regressors. We proposed a two-stage approach to evaluate the non-exact parameters as well as the autocorrelation criteria. For this reason, we employed the total of absolute residues and cross-validation measures to predict the parameters of the fuzzy regression procedure and the autocorrelation parameter. This approach reduces the summation of absolute residuals and cross-validation measures dependent on distance in the space of non-symmetrical $LR{-}\mathbf{FN}s$. Additionally, we check for serial correlation between the residuals using the conventional Durbin-Watson test. Using well-known model fit metrics, we compare the provided approach with several widely used fuzzy linear regression techniques. Additionally, we examined the performance and usefulness of the suggested semi-parametric quantile-based **SVR** method using several numerical instances and a simulation experiment. According to both comparative and numerical outcomes, the suggested model can produce sufficiently accurate fuzzy regression analysis results when the fuzzy error terms are correlated.

The structure of the remaining sections of this paper is as follows: Section 2 offers the foundational aspects of fuzzy numbers, generalized difference, and quantile regression. In Sect. 3, a semi-parametric **SVR** approach utilizing the quantile method is illustrated, with precise regressors and fuzzy outcomes, in situations where the fuzzy error expressions are dependent. Moreover, this section introduces a hybrid algorithm designed to calculate the imprecise parameters in addition to the corresponding accurate correlation measure. Sect. 4 offers a simulation investigation and two numeric examples to evaluate the quality and efficacy of the suggested method in contrast to other models using some commonly used goodness-of-fit criteria. Finally, Sect. 5 provides a summary of the major innovations discussed in this paper.

2 Introductory Concepts

We review the key terms and concepts that will be used in this paper in this section.

2.1 Fuzzy Numbers

Assume that $\widetilde{C}$ is a fuzzy subset [25] of $\mathbb{R}$. The definition of $\widetilde{C}$ is provided by a membership function, expressed as $\mu_{\widetilde{C}} : \mathbb{R} \to [0,1]$. For every $\beta \in (0,1]$, the β-cut of a fuzzy set $\widetilde{C}$, denoted by $\widetilde{C}[\beta]$, is represented as $\{z \in \mathbb{R} : \mu_{\widetilde{C}}(z) \geq \beta\}$. The crisp set $\widetilde{C}[0] = \overline{\{z \in \mathbb{R} : \mu_{\widetilde{C}}(z) > 0\}}$ identifies the support of the fuzzy set $\widetilde{C}$ [26]. The notation $\bar{C}$ represents the closure of C. By assuming that the domain of values for the β-cut can be shown as $\widetilde{C}^L_\beta = \inf\{z : z \in \widetilde{C}[\beta]\}$ and $\widetilde{C}^U_\beta = \sup\{z : z \in \widetilde{C}[\beta]\}$, $\widetilde{C}^L_\beta$ and $\widetilde{C}^U_\beta$ are named the lower and upper bounds of the β-cut, respectively. Moreover, a fuzzy set $\widetilde{C}$ of $\mathbb{R}$ is called a fuzzy number if: 1- it is normal, i.e. there is a unique element $z_0 \in \mathbb{R}$ that satisfies $\mu_{\widetilde{C}}(z_0) = 1$, 2- for every $\beta \in [0,1]$, $\widetilde{C}[\beta]$ is a non-empty, bounded, closed interval of $\mathbb{R}$. The most popular and intriguing technique for modeling fuzzy data is $LR-$**FN**s. The $LR-$fuzzy number or $LR-\mathbb{FN}$ $\widetilde{C}$ is represented as $\widetilde{C} = (c; l_c, r_c)_{LR}$ [27]. The membership function of an $LR-$**FN** $\widetilde{C}$ is determined by:

$$\mu_{\widetilde{C}}(z) = \begin{cases} L\big(\dfrac{c-z}{l_c}\big), & c - l_c \leq z \leq c, \\ R\big(\dfrac{z-c}{r_c}\big), & c \leq z \leq c + r_c, \end{cases}$$

so that $c \in \mathbb{R}$ is the center quantity, $l_c > 0$ is denoted by the left margin, and $r_c > 0$ is determined by the right margin. The reference or shape functions L and R define the left and right forms of the **FN**, respectively. The functions $L, R : [0,1] \to [0,1]$ must satisfy the following requirements:
1. $L(1) = R(1) = 0$, and
2. $L(0) = R(0) = 1$.
Additionally, L and R are continuous functions that are monotonically decreasing on the interval $[0,1]$. The symbol $\mathbb{F}(\mathbb{R})$ denotes the set of all $LR-$**FN**s. To address the inaccuracy in the dataset through numerical valuations, we also employed the most widely used (unimodal) $LR-$**FN**s, also known as the triangular fuzzy numbers **TFN**s, by the formula $L(z) = R(z) = 1 - z$. Also, **TFN**s can be represented as $\widetilde{C} = (c; l_c, r_c)_T$ and have the following membership function:

$$\mu_{\widetilde{C}}(z) = \begin{cases} \dfrac{z-(c-l_c)}{l_c}, & c - l_c \leq z \leq c, \\ \dfrac{(c+r_c)-z}{r_c}, & c \leq z \leq c + r_c, \\ 0, & z \in \mathbb{R} - (c - l_c, c + r_c). \end{cases}$$

In addition, several popular operations among two $LR-$**FN**s of $\widetilde{C} = (c; l_c, r_c)_{LR}$ and $\widetilde{D} = (d; l_d, r_d)_{LR}$ can be expressed as:

1. $\widetilde{C} \oplus \widetilde{D} = (c + d; l_c + l_d, r_c + r_d)_{LR}$.

2.

$$\gamma \otimes \widetilde{C} = \begin{cases} (\gamma c; \gamma l_c, \gamma r_c)_{LR}, & \text{if } \gamma > 0, \\ (\gamma c; -\gamma r_c, -\gamma l_c)_{RL}, & \text{if } \gamma < 0. \end{cases}$$

Definition 2.1. [28] *The operation* $(\widetilde{C} \ominus_{\boldsymbol{G}} \widetilde{D})$ *between two* **FN***s of* $\widetilde{C}$ *and* $\widetilde{D}$ *is known as the generalized difference operation. It is defined using the following* β*-cuts:*

$$(\widetilde{C} \ominus_{\boldsymbol{G}} \widetilde{D})[\beta] = \left[\inf_{\eta \in [\beta/2, 1-\beta/2]} (\widetilde{C}_\eta - \widetilde{D}_\eta), \sup_{\eta \in [\beta/2, 1-\beta/2]} (\widetilde{C}_\eta - \widetilde{D}_\eta) \right],$$

in which

$$\widetilde{C}_\eta = \begin{cases} \widetilde{C}^L[2\eta], & \eta \in [0, 0.5], \\ \widetilde{C}^U[2(1-\eta)], & \eta \in [0.5, 1]. \end{cases}$$

It should be noted that the major benefit of $\ominus_{\boldsymbol{G}}$ *on the Hukuhara subtraction [29] is its consistent existence [28]. The generalized difference* $\ominus_{\boldsymbol{G}}$ *satisfies the following features.*

Lemma 2.2. [28] *The key features of the generalized difference* $\ominus_{\boldsymbol{G}}$ *for three LR−***FN***s of* $\widetilde{C}$*,* $\widetilde{D}$*, and* $\widetilde{F}$ *are listed as follows:*

1. $\widetilde{C} \ominus_{\boldsymbol{G}} \widetilde{C} = \widetilde{0}$ where $\mu_{\widetilde{0}}(\mathrm{z}) = \begin{cases} 1 & z = 0, \\ 0 & z \neq 0. \end{cases}$

2. $\widetilde{C} \ominus_{\boldsymbol{G}} \widetilde{D} = (-1) \otimes \big(\widetilde{D} \ominus_{\boldsymbol{G}} \widetilde{C}\big) = \big((-1) \otimes \widetilde{D}\big) \ominus_{\boldsymbol{G}} ((-1) \otimes \widetilde{C})$.

3. $(\widetilde{C} \oplus \widetilde{D}) \ominus_{\boldsymbol{G}} \widetilde{D} = \widetilde{C}$.

4. $\widetilde{C} \ominus_{\boldsymbol{G}} \widetilde{D} = \widetilde{D} \ominus_{\boldsymbol{G}} \widetilde{C}$ if and only if $\widetilde{C} = \widetilde{D}$.

5. $(\widetilde{C} \oplus \widetilde{F}) \ominus_{\boldsymbol{G}} (\widetilde{D} \oplus \widetilde{F}) = \widetilde{C} \ominus_{\boldsymbol{G}} \widetilde{D}$.

6. $\widetilde{0} \ominus_{\boldsymbol{G}} (\widetilde{C} \ominus \widetilde{D}) = \widetilde{D} \ominus_{\boldsymbol{G}} \widetilde{C}$.

7. $n \otimes (\widetilde{C} \ominus_{\boldsymbol{G}} \widetilde{D}) = (n \otimes \widetilde{C}) \ominus_{\boldsymbol{G}} (n \otimes \widetilde{D})$, for any $n \in \mathbb{R}$.

*It is straightforward to verify that the generalized subtraction of two LR−***FN***s yields another LR−***FN***.*

Lemma 2.3. *Assume that* $\widetilde{C} = (c; l_c, r_c)_T$ *and* $\widetilde{D} = (d; l_d, r_d)_T$ *are two* **TFN***s. Then, we can define the generalized subtraction for* $\widetilde{C}$ *and* $\widetilde{D}$ *as follows:*

$$\widetilde{C} \ominus_{\boldsymbol{G}} \widetilde{D} = \big(c - d; |l_c - l_d|, |r_c - r_d|\big)_T.$$

Definition 2.4. *Let* $\widetilde{C} \in \mathbb{F}(\mathbb{R})$*, and* $\widetilde{C}_\beta : [0, 1] \to \mathbb{R}$ *be known as the* β*-values of* $\widetilde{C}$ *which has the following form:*

$$\widetilde{C}_\beta = \begin{cases} \widetilde{C}^L[2\beta] & 0 \leq \beta \leq 0.5, \\ \widetilde{C}^U[2(1-\beta)] & 0.5 \leq \beta \leq 1. \end{cases}$$

where $\widetilde{C}^L[\beta]$ *stands for the lower bound of the* β*-cuts of* $\widetilde{C}$ *and* $\widetilde{C}^U[\beta]$ *for the upper bound.*

Example 2.5. If we consider that $\widetilde{C} = (c; l_c, r_c)_{LR}$ is an $LR-$**FN**. Then, we can find from Definition 2.4 that:

$$\widetilde{C}_\beta = \begin{cases} c - l_c L^{-1}(2\beta), & 0 \leq \beta \leq 0.5, \\ c + r_c R^{-1}\left(2(1-\beta)\right), & 0.5 \leq \beta \leq 1. \end{cases}$$

For example,
1) Assuming that $\widetilde{C} = (c; l_c, r_c)_T$ is a **TFN**, after that:

$$\widetilde{C}_\beta = \begin{cases} (c - l_c) + 2l_c\beta, & 0 \leq \beta \leq 0.5, \\ c + r_c - 2r_c(1-\beta), & 0.5 \leq \beta \leq 1. \end{cases}$$

2) Assume that $\widetilde{C} = (c; l_c, r_c)_{LR}$ with the functions $L(z) = 1 - z$ and $R(z) = 1 - z^2$ then:

$$\widetilde{C}_\beta = \begin{cases} c - l_c(1 - 2\beta), & 0 \leq \beta \leq 0.5, \\ c + r_c(\sqrt{2\beta - 1}), & 0.5 \leq \beta \leq 1. \end{cases}$$

Remark 2.6. *If $\widetilde{C}_\beta$ is a non-increasing function of β, then β-values and β-cuts are related as follows:*

$$\widetilde{C}[\beta] = [\widetilde{C}^L[\beta], \widetilde{C}^U[\beta]] = [\widetilde{C}_{\beta/2}, \widetilde{C}_{1-\beta/2}], \quad \beta \in [0,1].$$

Definition 2.7. *If $\widetilde{C} = (c^L, c, c^U)_{LR}$ and $\widetilde{F} = (f^L, f, f^U)_{LR}$ are two $LR-$**FN**s. Then, the following definition of an absolute error distance criterion among $\widetilde{C}$ and $\widetilde{F}$ can be used:*

$$D(\widetilde{C}, \widetilde{F}) = |c - f| + r_1|c^L - f^L| + r_2|c^U - f^U|,$$

where $r_1 = \int_0^1 L^{-1}(\beta)d\beta$ and $r_2 = \int_0^1 R^{-1}(\beta)d\beta$. It can be shown that the distance D meets the requirements listed below:

1. $D(\widetilde{C}, \widetilde{F}) = 0 \Leftrightarrow \widetilde{C} = \widetilde{F}$,
2. $D(\widetilde{C}, \widetilde{F}) = D(\widetilde{F}, \widetilde{C})$,
3. $D(\widetilde{C}, \widetilde{N}) \leq D(\widetilde{C}, \widetilde{F}) + D(\widetilde{F}, \widetilde{N})$.

Definition 2.8. *Suppose that $\widetilde{C}$ is an* **FN**. *Next, using the triangular density function of* **h**, *we can calculate the expectation of $\widetilde{C}$ as follows.*

$$E_{\widetilde{C}} = \int_0^{0.5} \widetilde{C}^L_{2\beta} h(\beta)d\beta + \int_{0.5}^1 \widetilde{C}^U_{2(1-\beta)} h(\beta)d\beta,$$

where

$$h_{0,0.5,1}(\beta) = \begin{cases} 4\beta & \beta \in [0, 0.5], \\ 4(1-\beta) & \beta \in [0.5, 1]. \end{cases}$$

Example 2.9. Here are some special cases of $LR-$**FN** expectation that we utilized in our computational procedure.

1. Assume that $\widetilde{C} = (c; l_c, r_c)_T$ is a **TFN**. The expectation of $\widetilde{C}$ can then be computed as follows:

$$\begin{aligned} E_{\widetilde{C}} =& E_{h_{0,0.5,1}}(\widetilde{C}) = \int_0^1 h_{0,0.5,1}(\beta)\widetilde{C}_\beta \, d\beta = \int_0^{0.5} 4\beta(c - (1-2\beta)l_c) \, d\beta \\ &+ \int_{0.5}^1 4(1-\beta)(c - (1-2\beta)r_c) \, d\beta = c + \frac{r_c - l_c}{6}. \end{aligned}$$

2. Assume that $\widetilde{C} = (c; l_c)_{LL}$ where $L(z) = \sqrt{1 - z^3}$. Then

$$E_{\widetilde{C}} = E_{h_{0,0.5,1}}(\widetilde{C}) = \int_0^{0.5} 4\beta(c - l_c\sqrt[3]{1 - 4\beta^2})\, d\beta + \int_{0.5}^{1} 4(1-\beta)(c + l_c\sqrt[3]{1 - 4(1-\beta)^2})d\beta = \beta.$$

3. Assume that $\widetilde{C} = (c; l_c, r_c)_{LR}$ where $L(z) = 1 - z$ and $R(z) = 1 - z^2$. After that.

$$\begin{aligned} E_{\widetilde{C}} &= E_{h_{0,0.5,1}}(\widetilde{C}) \\ &= \int_0^{0.5} 4\beta(c - l_c(1 - 2\beta))\, d\beta + \int_{0.5}^{1} 4(1-\beta)(c + r_c(\sqrt{2\beta - 1}))\, d\beta \\ &= c - 0.166667 l_c + 0.266667 r_c. \end{aligned} \tag{1}$$

2.2 Quantile Regression

Quantile regression is a technique that addresses the limitations of the least squares regression model, including issues such as non-normality of the error distribution and the presence of outliers. Koenker and Bassett [6] first proposed this model in 1978, and it has since developed into a comprehensive method for the statistical analysis of both linear and nonlinear models of response variables across various domains. Using quantile regression and estimating a family of conditional quantile functions yield more complete patterns of the effects of explanatory variables across all parts of the distribution. Quantile regression is a generalization of quantile to conditional quantile when one or more explanatory variables are present. Assume that Z is a real number stochastic variable. Also, assume that $F(z)$ is the distribution function of the variable Z, which is defined as $F_Z(z) = F(z) = P(Z \le z)$. The following is the ηth quantile of Z:

$$Q_Z(\eta) = Q(z) = F_Z^{-1}(\eta) = \inf\{z : F(z) > \eta\}$$

where $\eta \in [0, 1]$. The loss function can be expressed in the following way.

$$\begin{aligned} \rho_\eta(z) &= [\eta - I(z < 0)]z \\ &= [(1-\eta)I(z \le 0) + \eta I(z > 0)]|z|. \end{aligned}$$

where I is the characteristic function. The ηth sample quantile based on a random sample $\{z_1, ..., z_n\}$ can be found by solving the following equation:

$$q_\eta = \arg\min_b E[\rho_\eta(Z - b)].$$

The prior minimization problem is transformed as follows when a discrete variable Z has a probability distribution $f(z) = P(Z = z)$:

$$\begin{aligned} q_\eta &= \arg\min_b E[\rho_\eta(Z - b)] \\ &= \arg\min_b \left\{ (1-\eta)\sum_{z \le b} |z - b| h(z) + \eta \sum_{z > b} |z - b| h(z) \right\}. \end{aligned}$$

The analogous measure is employed for a continuous stochastic variable by replacing the sum with an integral:

$$\begin{aligned} q_\eta &= \arg\min_b E[\rho_\eta(Z - b)] \\ &= \arg\min_b \left\{ (1-\eta)\int_{-\infty}^{b} |z - b| h(z) d(z) + \eta \int_{b}^{+\infty} |z - b| h(z) d(z) \right\}. \end{aligned}$$

Additionally, the sample data is used in the preceding formula to acquire the sample estimate $\hat{q}_\eta$ for $\eta \in [0,1]$. The ηth conditioned quantile of N considering Z is shown using the following formula:

$$Q_{N|Z}(\eta) = \inf\{n : F_{N|Z}(n) \geq \eta\}.$$

Consequently, the quantile-based regression approach is illustrated as follows:

$$Q_{N|Z}(\eta) = Z\beta_\eta.$$

This minimization problem determines β_η by considering the distribution function of N:

$$\beta_\eta = \arg\min_{\beta} E[\rho_\eta(N - Z\beta)].$$

By solving the above equation for the sample, we obtain the estimator of β as follows:

$$\hat{\beta}_\eta = \arg\min_{\beta} \sum_{i=1}^{n} (\rho_\eta(N_i - Z_i\beta)).$$

3 Quantile SVR-Based Fuzzy Semi-Parametric Modeling with Autoregressive Fuzzy Errors (FSPQSVR)

This section outlines the standard quantile approach for modeling semi-parametric support vector regression, incorporating autoregressive fuzzy errors based on precise inputs and fuzzy outputs. Let $\{(m_p, \widetilde{n}_p)\}_{p=1,2,\dots,j}$ represent a collection of training patterns, in which $\widetilde{n}_p \in \mathcal{F}(\mathbb{R})$ is the seen value for every input data $m_p \in \mathbb{R}^j$. Consider that all input data are sorted in matrix $C \in \mathbb{R}^{i\times j}$, where m_p^t is the p-th row. Moreover, if $K(.,.)$ is a kernel function [30, 31], then a kernel matrix $K(C, C^t)$ with rank j can be displayed so that its pq-th entry is $(K(C, C^t))_{pq} = K(m_p, m_q)$. Assume that $K(m, C^t) = (K(m, m_1), \dots, K(m, m_j))^T$ is a row vector for each $m \in \mathbb{R}^j$. Using the data mentioned above, we can implement the fuzzy semi-parametric **SVR** procedure with autocorrelated fuzzy error terms (**FSPSVR**):

$$\widetilde{n}_p^{*\rho} = \widetilde{z}_p^{\rho}(\widetilde{\mathbf{w}}) \oplus \widetilde{\mathbf{v}}_{\mathbf{p}},$$

where:

1.

$$\begin{aligned}\widetilde{z}_p^{\rho}(\widetilde{\mathbf{w}}) = \Bigg(&\big(K(m_p, C^t) - \rho K(m_{p-1}, C^t)\big)w; \\ &|K(m_p, C^t)|l_w + \Big|\big|(\rho s l_{n_{p-1}} - \rho(1-s) r_{n_{p-1}}) - \big|\rho K(m_{p-1}, C^t)\big| l_w\big| - (\rho s l_{n_{p-1}} - \rho(1-s) r_{n_{p-1}})\Big|, \\ &|K(m_p, C^t)|r_w + \Big|\Big|\big(\rho s r_{n_{p-1}} - \rho(1-s) l_{n_{p-1}}\big) - \big|\rho K(m_{p-1}, C^t)\big| r_w\Big| - \big(\rho s r_{n_{p-1}} - \rho(1-s) l_{n_{p-1}}\big)\Big|\Bigg)_{LR}.\end{aligned}$$

Here, $\widetilde{z}_p^{\rho}(\widetilde{\mathbf{w}})$ is the fuzzy function of unknown multipliers $\widetilde{w}$ in the regression model relative to the correlation parameter ρ, which is an $LR-$**FN**.

2. $\widetilde{\mathbf{w}} = (\mathbf{w}; l_{\mathbf{w}}, r_{\mathbf{w}})_{LR}$, $(p = 1, 2, \dots, j)$ stands for the fuzzy coefficients, which are fuzzy numbers of the LR type where $\mathbf{w} = (w_1, w_2, \dots, w_j)^T$, $l_{\mathbf{w}} = (l_{w_1}, l_{w_2}, \dots, l_{w_j})^T$, $r_{\mathbf{w}} = (r_{w_1}, r_{w_2}, \dots, r_{w_j})^T$,

3.

$$\begin{aligned}\widetilde{n}_p^{*\rho} &= \widetilde{n}_p \ominus_{\boldsymbol{G}} (\rho \otimes \widetilde{n}_{p-1}) \\ &= \Big(n_p - \rho n_{p-1}; \Big| l_{n_p} - \big(\rho s l_{n_{p-1}} - \rho(1-s) r_{n_{p-1}}\big) \Big|, \Big| r_{n_p} - \big(\rho s r_{n_{p-1}} - \rho(1-s) l_{n_{p-1}}\big) \Big| \Big)_{LR}.\end{aligned}$$

are fuzzy responses.

4. $\boldsymbol{\widetilde{v}_p = (v_p; l_{v_p}, r_{v_p})_{LR}}$, $p = 1, 2, ..., j$ are fuzzy error expressions whose $v_1, v_2, ..., v_j$ are stochastic samples drawn from a Gaussian distribution with a mean of 0 and a variance of σ^2. Moreover, $l_{v_1}, l_{v_2}, ..., l_{v_j}$ and $r_{v_1}, r_{v_2}, ..., r_{v_j}$ are independent and identically distributed positive stochastic variables.

Therefore, we can formulate the quantile function for a fuzzy random variable $\widetilde{n}_p^{*\rho}$ based on Theorem 3.7 as follows:

$$Q_{\widetilde{n}_p^{*\rho}}(\eta) = \widetilde{z}_p^{\rho}(\widetilde{\mathbf{w}}) \oplus Q_{\widetilde{v}_p}(\eta) \tag{2}$$

Therefore, we can show the quantile function for $\widetilde{n}_p^{*\rho}$ as an LR-fuzzy random variable as follows:

$$\begin{aligned}\widetilde{Q}_{\widetilde{n}_p^{*\rho}}(\eta) = \Bigg(& \Big(K(m_p, C^t) - \rho K(m_{p-1}, C^t) \Big) w + Q_{v_p}(\eta); \\ & |K(m_p, C^t)| l_w + \Bigg| \Big| (\rho s l_{n_{p-1}} - \rho(1-s) r_{n_{p-1}}) - |\rho K(m_{p-1}, C^t)| l_w \Big| \\ & - (\rho s l_{n_{p-1}} - \rho(1-s) r_{n_{p-1}}) \Bigg| + Q_{l_{v_p}}(\eta), \\ & |K(m_p, C^t)| r_w + \Bigg| \Big| (\rho s r_{n_{p-1}} - \rho(1-s) l_{n_{p-1}}) - \big|\rho K(m_{p-1}, C^t)\big| \ r_w \Big| \\ & - (\rho s r_{n_{p-1}} - \rho(1-s) l_{n_{p-1}}) \Bigg| + Q_{r_{v_p}}(\eta) \Bigg)_{LR}.\end{aligned} \tag{3}$$

Remark 3.1. *Interestingly, the classic regression method can be expressed using autoregressive modeling as follows:*

$$n_p^{*\rho} = z_p^{\rho}(\mathbf{w}) + \mathbf{v_p}, \quad \mathbf{p = 1, 2, ..., j},$$

in which ρ points out the correlation coefficient between v_p and v_{p-1} where $\rho = corr(v_p, v_{p-1})$ and $\rho \in [-1, 1]$ [32]. However, when $\rho = 0$, it means that the residual expressions in the regression model are not correlated. Conversely, $\rho \neq 0$ indicates that the residual expressions of the regression modeling may be either positively or negatively correlated. The Durbin-Watson test can be employed in regression modeling to determine whether there is a correlation between the residuals (errors) [33, 34]. The Durbin-Watson test assesses whether the residual expressions from one observation are correlated with those from the previous observation, specifically testing for first-order autocorrelation. Using the Durbin-Watson statistic, we can evaluate the null hypothesis $H_0 : \rho = 0$ as follows:

$$D = \frac{\sum_{p=2}^{j} (v_p - v_{p-1})^2}{\sum_{p=2}^{j} {v_p}^2}, \tag{4}$$

in which $v_p = n_p - \widehat{n}_p$ are the residuals derived from the standard least squares regression method, and the quantity of observations is denoted by j. In addition, when j is big enough, D is nearly equivalent to $2(1 - \widehat{\rho})$.

The test compares the calculated D value to critical values found in statistical tables at two significance levels (0.05 and 0.01) (Kutner et al. [32]). D is always between 0 and 4. When the value of D is 2, autocorrelation is absent, whereas values above 2 indicate negative serial correlation, and values below 2 indicate positive autocorrelation. Evaluating positive autocorrelation at the α significance level requires comparing the test statistic D with the critical thresholds, $d_{L,\alpha}$ and $d_{U,\alpha}$, which represent the lower and upper limits, respectively. Thus, it is necessary to verify the following rule to evaluate the hypothesis: $H_0 : \rho = 0$ against $H_1 : \rho > 0$. When D is smaller than $d_{L,\alpha}$, H_0 should be rejected. If D is greater than $d_{U,\alpha}$, H_0 is not rejected. However, if D lies within the range of $d_{L,\alpha}$ and $d_{U,\alpha}$, the test is indecisive. To test for negative autocorrelation at an α significance level, the critical values $d_{L,\alpha}$ and $d_{U,\alpha}$, which represent the lower and upper limits, are compared to the statistic $(4 - D)$. The residual expressions are negatively autocorrelated if $(4 - D) < d_{L,\alpha}$, which indicates that $\rho < 0$. If $d_{L,\alpha} < (4 - D) < d_{U,\alpha}$, the test is ineffective, but if $(4 - D) > d_{U,\alpha}$, then H_0 cannot be rejected. Additionally, two one-sided tests can be applied independently to construct a two-sided test for $H_0 : \rho = 0$ versus $H_1 : \rho \neq 0$. The assumption is that $\alpha = 0.05$ is maintained throughout the paper.

3.1 Algorithm Applied to Choosing Model Constituents:

To evaluate the constituents of the proposed **FSPQSVR**, we use the approach presented by Akbari et al.[35]. They introduced a two-stage method for calculating both the multipliers and the correlation coefficient for the proposed regression technique 3. Consider that in the proposed method 3, ρ and $\widetilde{w}$ are both undefined values. It is possible to estimate the undefined vector of fuzzy multipliers of $\widetilde{w}$ and ρ by simultaneously minimizing the following goal functions:

1) **The first goal function**:

$$
\begin{aligned}
g(\rho, \widetilde{\mathbf{w}}) =& \sum_{p=2}^{j} \rho_\eta(D(\widetilde{n}_p^{*\rho}, \widetilde{Q}_{\widetilde{n}_p^{*\rho}}(\eta))) = \sum_{p=2}^{j} \left(\left| \eta - I\left(D(\widetilde{n}_p^{*\rho}, \widetilde{Q}_{\widetilde{n}_p^{*\rho}}(\eta)) \leq 0 \right) \right| \left| D\left(\widetilde{n}_p^{*\rho}, \widetilde{Q}_{n_p^{*\rho}}(\eta) \right) \right| \right) \\
=& \sum_{p=2}^{j} \left(\left| \eta - I\left(D(\widetilde{n}_p^{*\rho}, \widetilde{Q}_{\widetilde{n}_p^{*\rho}}(\eta)) \leq 0 \right) \right| \right) \\
& * \left| \left| \left(n_p - \rho n_{p-1} \right) - \left(\left(K(m_p, C^t) - \rho K(m_{p-1}, C^t) \right) w + Q_{v_p}(\eta) \right) \right| \right. \\
& + \frac{1}{2} \left| \left(\left| l_{n_p} - \left(\rho s l_{n_{p-1}} - \rho(1-s) r_{n_{p-1}} \right) \right| \right) - \left(|K(m_p, C^t)| l_w \right. \right. \\
& + \left| \left| \left(\rho s l_{n_{p-1}} - \rho(1-s) r_{n_{p-1}} \right) - |\rho K(m_{p-1}, C^t)| l_w \right| \right. \\
& \left. \left. - \left(\rho s l_{n_{p-1}} - \rho(1-s) r_{n_{p-1}} \right) \right| + Q_{l_{v_p}}(\eta) \right) \Bigg| + \frac{2}{3} \left| \left(\left| r_{n_p} - \left(\rho s r_{n_{p-1}} - \rho(1-s) l_{n_{p-1}} \right) \right| \right) \right. \\
& - \left(\left| K(m_p, C^t) \right| r_w + \left| \left| \left(\rho s r_{n_{p-1}} - \rho(1-s) l_{n_{p-1}} \right) - |\rho K(m_{p-1}, C^t)| r_w \right| \right. \right. \\
& \left. \left. \left. - \left(\rho s r_{n_{p-1}} - \rho(1-s) l_{n_{p-1}} \right) \right| + Q_{r_{v_p}}(\eta) \right) \right| \Bigg| .
\end{aligned}
$$

Therefore we have:

$$\begin{aligned}
g(\rho, \widetilde{\mathbf{w}}) =&\eta \sum_{p=2: D(\widetilde{n}_p^{*\rho}, \widetilde{Q}_{\widetilde{n}_p^{*\rho}}(\eta)) \geq 0}^{p} \Bigg|\Big|(n_p - \rho n_{p-1}) - \Big(\Big(K(m_p, C^t) - \rho K(m_{p-1}, C^t)\Big) w + Q_{v_p}(\eta)\Big)\Big| \\
&+ \frac{1}{2}\Bigg|\Big(\Big|l_{n_p} - \Big(\rho s l_{n_{p-1}} - \rho(1-s) r_{n_{p-1}}\Big)\Big|\Big) - \Big(\Big|K(m_p, C^t)\Big| l_w + \Big|\big|(\rho s l_{n_{p-1}} - \rho(1-s) r_{n_{p-1}}) \\
&- |\rho K(m_{p-1}, C^t)| l_w\big| - \Big(\rho s l_{n_{p-1}} - \rho(1-s) r_{n_{p-1}}\Big)\Big| + Q_{l_{v_p}}(\eta)\Big)\Bigg| \\
&+ \frac{2}{3}\Bigg|\Big(\Big|r_{n_p} - \Big(\rho s r_{n_{p-1}} - \rho(1-s) l_{n_{p-1}}\Big)\Big|\Big) - \Big(\Big|K(m_p, C^t)\Big| r_w \\
&+ \Big|\big|(\rho s r_{n_{p-1}} - \rho(1-s) l_{n_{p-1}}) - \big|\rho K(m_{p-1}, C^t)\big| r_w\big| - \Big(\rho s r_{n_{p-1}} - \rho(1-s) l_{n_{p-1}}\Big)\Big| + Q_{r_{v_p}}(\eta)\Big)\Bigg|\Bigg| \\
&+ (1-\eta) \sum_{p=2: D(\widetilde{n}_p^{*\rho}, \widetilde{Q}_{\widetilde{n}_p^{*\rho}}(\eta)) < 0}^{j} \Bigg|\Big|(n_p - \rho n_{p-1}) - \Big(\big(K(m_p, C^t) - \rho K(m_{p-1}, C^t)\big) w + Q_{v_p}(\eta)\Big)\Big| \\
&+ \frac{1}{2}\Bigg|\Big(\Big|l_{n_p} - (\rho s l_{n_{p-1}} - \rho(1-s) r_{n_{p-1}})\Big|\Big) - \Big(|K(m_p, C^t)| l_w + \Big|\big|(\rho s l_{n_{p-1}} - \rho(1-s) r_{n_{p-1}}) \\
&- |\rho K(m_{p-1}, C^t)| l_w\big| - (\rho s l_{n_{p-1}} - \rho(1-s) r_{n_{p-1}})\Big| + Q_{l_{v_p}}(\eta)\Big)\Bigg| \\
&+ \frac{2}{3}\Bigg|\Big(\big|r_{n_p} - \Big(\rho s r_{n_{p-1}} - \rho(1-s) l_{n_{p-1}}\big)\big|\Big) - \Big(\Big|K(m_p, C^t)| r_w + \Big|\big|(\rho s r_{n_{p-1}} - \rho(1-s) l_{n_{p-1}}) \\
&- |\rho K(m_{p-1}, C^t)| r_w\big| - \Big(\rho s r_{n_{p-1}} - \rho(1-s) l_{n_{p-1}}\Big)\Big| + Q_{r_{v_p}}(\eta)\Big)\Bigg|\Bigg|,
\end{aligned}$$

where:

1. $g(\rho, \widetilde{\mathbf{w}})$ represents the sum of the absolute error in the goal function.

2. $\rho_\eta(v) = |\eta - I(v \leq 0)||v|, \quad \eta \in (0,1)$, is the quantile loss function (quantile cost function) [7].

Remark 3.2. *The absolute error distance criterion among two fuzzy quantities that was mentioned in Definition 2.7, i.e., $D(\widetilde{C}, \widetilde{F})$, was utilized in the calculations of the first goal function $g(\rho, \widetilde{\mathbf{w}})$ for the purpose of estimating the undefined multipliers $\widetilde{w}$ in the regression model 3.*
where: $r_1 = \int_0^1 L^{-1}(\beta) d\beta$ *and* $r_2 = \int_0^1 R^{-1}(\beta) d\beta$.

Remark 3.3. *We assume that $L(z) = 1 - z$ and $R(z) = 1 - z^2$ have these forms throughout the article, where we intend to compute the absolute error distance measure $D(\widetilde{C}, \widetilde{F})$. Therefore, in this case, we can deduce that:* $r_1 = \int_0^1 L^{-1}(\beta) d\beta = \frac{1}{2}$ *and* $r_2 = \int_0^1 R^{-1}(\beta) d\beta = \frac{2}{3}$.

Remark 3.4. *The two fuzzy numbers $\widetilde{C} = (c, l_c, r_c)$ and $\widetilde{F} = (f, l_f, r_f)$ in the absolute error distance criterion*

$D(\widetilde{C},\widetilde{F})$ *used in calculations of the first goal function* $g(\rho,\widetilde{\mathbf{w}})$ *are defined as follows:*

$$\begin{aligned}\widetilde{C}\equiv\widetilde{n}_p^{*\rho}=&\widetilde{n}_p\ominus_{\boldsymbol{G}}(\rho\otimes\widetilde{n}_{p-1})\\ =&\Big(n_p-\rho n_{p-1};\Big|l_{n_p}-\Big(\rho s l_{n_{p-1}}-\rho(1-s)r_{n_{p-1}}\Big)\Big|,\Big|r_{n_p}-\Big(\rho s r_{n_{p-1}}-\rho(1-s)l_{n_{p-1}}\Big)\Big|\Big)_{LR}.\\ \widetilde{F}\equiv\widetilde{Q}_{\widetilde{n}_p^{*\rho}}(\eta)=&\Big(\Big(K(m_p,C^t)-\rho K(m_{p-1},C^t)\Big)w+Q_{v_p}(\eta);\\ &|K(m_p,C^t)|l_w+\Big||(\rho s l_{n_{p-1}}-\rho(1-s)r_{n_{p-1}})-|\rho K(m_{p-1},C^t)|l_w|\\ &-(\rho s l_{n_{p-1}}\rho(1-s)r_{n_{p-1}})\Big|+Q_{l_{v_p}}(\eta),\\ &|K(m_p,C^t)|r_w+\Big||(\rho s r_{n_{p-1}}-\rho(1-s)l_{n_{p-1}})-|\rho K(m_{p-1},C^t)|r_w|\\ &-\Big(\rho s r_{n_{p-1}}-\rho(1-s)l_{n_{p-1}}\Big)\Big|+Q_{r_{v_p}}(\eta)\Big)_{LR}.\end{aligned}\tag{5}$$

2) **The second goal function**:

$$CV(\rho,\widehat{\mathbf{w}}^*)=\sum_{p=2}^{j}\Big(E_{n_p^{*\rho}}-E_{Q_{n_p^{*\rho}}(\widehat{w}^{*(-p)})}\Big)^2,\tag{6}$$

where $\widehat{w}^{*(-p)}$ shows the estimated value of $\widehat{w}$ dependent on the hold-out observation in matrix $C\in\mathbb{R}^{i\times j}$ whose $\mathbf{m_p^t}$ is the p-th row. It is important to note that CV represents the cross-validation measurement, while $g(\rho,\widetilde{\mathbf{w}})$ denotes the sum total of the absolute error goal function. In addition, quantities $E_{n_p^{*\rho}}$ and $E_{Q_{n_p^{*\rho}}(\widehat{w}^{*(-p)})}$ present the expectation of $\widetilde{n}_p^{*\rho}$ and $\widetilde{Q}_{\widetilde{n}_p^{*\rho}}(\widehat{w}^{*(-p)})$, respectively.
In the following, to find the second goal function, we need to calculate the expectation of $\widetilde{n}_p^{*\rho}$ and $\widetilde{Q}_{\widetilde{n}_p^{*\rho}}(\widehat{w}^{*(-p)})$ via Definition 2.8.

Corollary 3.5. *We conclude from Definition 2.8 and formula 1 in Example 2.9 that the expectation of* $\widetilde{n}_p^{*\rho}$ *and* $\widetilde{Q}_{\widetilde{n}_p^{*\rho}}(\widehat{w}^{*(-p)})$ *is respectively as follows:*

$$\begin{aligned}E_{\widetilde{n}_p^{*\rho}}=&(n_p-\rho n_{p-1})-0.166667\Big(\Big|l_{n_p}-(\rho s l_{n_{p-1}}-\rho(1-s)r_{n_{p-1}})\Big|\Big)\\ &+0.266667\Big(\Big|r_{n_p}-\big(\rho s r_{n_{p-1}}-\rho(1-s)l_{n_{p-1}}\big)\Big|\Big).\\ E_{\widetilde{Q}_{\widetilde{n}_p^{*\rho}}}=&\Big(\Big(K(m_p,C^t)-\rho K(m_{p-1},C^t)\Big)w+Q_{v_p}(\tau)\Big)-0.166667\Big(|K(m_p,C^t)|l_w\\ &+||(\rho s l_{n_{p-1}}-\rho(1-s)r_{n_{p-1}})-|\rho K(m_{p-1},C^t)|l_w|-(\rho s l_{n_{p-1}}-\rho(1-s)r_{n_{p-1}})|+Q_{l_{v_p}}(\tau)\Big)\\ &+0.266667\Big(|K(m_p,C^t)|r_w+\Big||\big(\rho s r_{n_{p-1}}-\rho(1-s)l_{n_{p-1}}\big)-|\rho K(m_{p-1},C^t)|r_w|\\ &-\Big(\rho s r_{n_{p-1}}-\rho(1-s)l_{n_{p-1}}\Big)\Big|+Q_{r_{v_p}}(\tau)\Big).\end{aligned}$$

Therefore, we can write the second goal function in formula 6 based on formula 1 in Example 2.9 and Corollary 3.5 as follows:

$$
\begin{aligned}
CV(\rho,\widehat{\mathbf{w}}^{*}) =& \sum_{p=2}^{j}\left(E_{n_p^{*\rho}} - E_{Q_{n_p^{*\rho}(\widehat{w}^{*(-p)})}}\right)^2 \\
=& \sum_{p=2}^{j}\Bigg(\Big(\Big(n_p-\rho n_{p-1}\Big) - 0.166667\Big(\big|l_{n_p} - \big(\rho s l_{n_{p-1}} - \rho(1-s) r_{n_{p-1}}\big)\big|\Big) \\
&+ 0.266667\Big(\big|r_{n_p} - \big(\rho s r_{n_{p-1}} - \rho(1-s) l_{n_{p-1}}\big)\big|\Big)\Big) - \Big(\Big(\big(K(m_p,C^t) - \rho K(m_{p-1},C^t)\big)\widehat{w}^{*(-p)} + Q_{v_p}(\eta)\Big) \\
&- 0.166667\Bigg(\big|K(m_p,C^t)\big| l_{\widehat{w}^{*(-p)}} + \Bigg|\Big|\Big(\rho s l_{n_{p-1}} - \rho(1-s) r_{n_{p-1}}\Big) - \big|\rho K(m_{p-1},C^t)\big| l_{\widehat{w}^{*(-p)}}\Big| \\
&- \Big(\rho s l_{n_{p-1}} - \rho(1-s) r_{n_{p-1}}\Big)\Bigg| + Q_{l_{v_p}}(\eta)\Bigg) + 0.266667\Big(\big|K(m_p,C^t)\big| r_{\widehat{w}^{*(-p)}} + \Big|\big|\big(\rho s r_{n_{p-1}} - \rho(1-s) l_{n_{p-1}}\big) \\
&- |\rho K(m_{p-1},C^t)| r_{\widehat{w}^{*(-p)}}\big| - \big(\rho s r_{n_{p-1}} - \rho(1-s) l_{n_{p-1}}\big)\Big| + Q_{r_{v_p}}(\eta)\Big)\Big)\Bigg)^2.
\end{aligned}
$$

The algorithm that can be used to minimize the goals mentioned earlier is:
Stage (1): Allow

$$\widehat{\rho} = \arg \min_{\rho\in\{-1,\dots,-0.01,0,0.01,\dots,1\}} CV(\rho,\widehat{\mathbf{w}}^{*}_{\rho}).$$

Stage(2): Allow

$$
\begin{aligned}
\widetilde{\widetilde{\mathbf{w}}}_{\widehat{\rho}} =& \arg\min_{\widetilde{\mathbf{w}}} \mathrm{g}\left(\widehat{\rho},\widetilde{\mathbf{w}}\right) \\
=& \arg \min_{w_p, l_{w_p}>0, r_{w_p}>0} \sum_{p=2}^{j} \rho_\eta\left(D\left(\widetilde{n}_p^{*\widehat{\rho}}, \widetilde{Q}_{\widetilde{n}_p^{*\widehat{\rho}}}(\widetilde{w})\right)\right).
\end{aligned}
$$

Hence, $\widehat{\rho}$ and $\widetilde{\widetilde{w}}_{\widehat{\rho}}$ are the optimal values.

Remark 3.6. *In the proposed* **FSPQSVR** *model, we calculate the smoothing value in the kernel $K(C,C^t)$, h, using the iterative method or the generalized Wasserman cross-validation measure. Furthermore, we compute σ^2 using the trial-and-error method in Mathematica software.*

Theorem 3.7. *Let $\widetilde{Z} = (\widetilde{Z}^L, Z, \widetilde{Z}^U)_{\mathrm{LR}}$ be an LR-fuzzy stochastic variable. Then, for every fuzzy number $\tilde{b} = (\tilde{b}^L, b, \tilde{b}^U)_{\mathrm{LR}}$, we have that:*

$$\widetilde{Q}_{Z\oplus\widetilde{b}}(\eta) = Q_Z(\eta) \oplus \widetilde{b}.$$

Proof.

$$
\begin{aligned}
\widetilde{Q}_{Z\oplus\widetilde{b}}(\eta) =& \left(F^{-1}_{Z+\widetilde{b}^L}(\eta), F^{-1}_{Z+b}(\eta), F^{-1}_{Z+\widetilde{b}^U}(\eta)\right)_{\mathrm{LR}} \\
=& \left(Q_Z(\eta)+\widetilde{b}^L, Q_Z(\eta)+b, Q_Z(\eta)+\widetilde{b}^U\right)_{\mathrm{LR}} \\
=& Q_Z(\eta) \oplus \left(\widetilde{b}^L, b, \widetilde{b}^U\right)_{\mathrm{LR}} \\
=& Q_Z(\eta) \oplus \widetilde{b}.
\end{aligned}
$$

□

Remark 3.8. *Several performance measures were applied to assess the performance of the suggested model and to compare it with other methods.*

1. *Mean absolute errors:*

$$MAE = \frac{1}{j}\sum_{p=1}^{j}|\widetilde{n}_p - \widetilde{\widetilde{n}}_p|,$$

2. *Mean absolute relative errors:*

$$MARE = \frac{1}{j}\sum_{p=1}^{j}\frac{\int_0^1|\widetilde{n}_p(m) - \widetilde{\widetilde{n}}_p(m)|\,d(m)}{\int_0^1\widetilde{n}_p(m)\,d(m)}.$$

3. *Mean similarity criterion:*

$$MSM = \frac{1}{j}\sum_{p=1}^{j}S_{UI}(\widetilde{\widetilde{n}}_p, \widetilde{n}_p),$$

Where

$$S_{UI}(\widetilde{\widetilde{n}}_p, \widetilde{n}_p) = \frac{\int \min\{\widetilde{\widetilde{n}}_p(m), \widetilde{n}_p(m)\}dm}{\int \max\{\widetilde{\widetilde{n}}_p(m), \widetilde{n}_p(m)\}dm},$$

It is worth noting that $MSM \in [0, 1]$. Additionally, MSM is zero if and only if there is no intersection between $\widetilde{\widetilde{n}}_p$ and $\widetilde{n}_p$, whereas MSM is one if and only if $\widetilde{\widetilde{n}}_p$ is equal to $\widetilde{n}_p$. As a result, an MSM value close to 1 indicates a strong match between outcome variables and their identical estimates.

4. *Determination coefficient:*

$$COD = 1 - \frac{\sum_{p=1}^{j} D^2(\widetilde{n}_p, \widetilde{\widetilde{n}}_p)}{\sum_{p=1}^{j} D^2(\widetilde{n}_p, \widetilde{\overline{n}}_p)},$$

in which $\widetilde{\overline{n}} = (\frac{1}{j}) \otimes (\oplus_{p=1}^{j}\widetilde{n}_p)$. A crucial point to remember is that $COD \in [0, 1]$. A COD of 1 indicates that the regression model perfectly fits the data points, while a COD of 0 means that the regression model does not fit the data points at all.

4 Numeric Examples

In this section, we use two numeric examples and a simulation study to assess the efficacy and applicability of the suggested **FSPQSVR** model and compare it with different models. The primary focus is on the linear fuzzy regression techniques created by Yoon and Choi [36], Choi and Buckley [37], Zeng et al. [38], Taheri and Kelkinnama [39], Kula and Apaydin [40], and Icen and Demirhan [41]. The measurements outlined in Remark 3.8 are used to compute the goodness-of-fit measures.

Example 4.1. (Zhou et al. [42]) proposed a fuzzy linear regression procedure to estimate affordable levels of house prices. These models identify six key factors: three related to policy (the interest rate on a mortgage, the tax on real estate, and the ratio of down payment) and three non-policy factors (size of the house, yearly income of the household, and family population). These factors aid in decision-making for both the government and realty extenders. An interview questionnaire is used in this study to collect data for the variables mentioned above. To enhance the credibility of our treatment, we distributed questionnaires to wage earners aged 20 to 60 who were working in trade office buildings in Shanghai, including the Green Building, Global Finance Building, Air $SOHO$, and Golden Bridge Software Park. This group consists of employees who purchased their properties within the last five years (2012-2016) or have agreed to buy a property. At first, 200 employees were invited to participate in the survey. Nevertheless, 53 of them either did not complete it correctly or gave answers that did not match the facts. Thus, the remaining 147 questionnaires are the main focus of this investigation. In this example, the fuzzy outcomes are presented as non-symmetrical triangular **FN**s, and the proposed **FSPQSVR** model is examined. The main goal is to evaluate $H_0 : \rho = 0$ in contrast to $H_1 : \rho > 0$ using the classic Durbin-Watson test. The amounts of v_p, shown in Table 1, are obtained by minimizing the MSE criterion. We can easily determine the test statistic from formula 4 as follows:

$$D = \frac{\sum_{p=2}^{147}(v_p - v_{p-1})^2}{\sum_{p=2}^{147} {v_p}^2} = 0.054,$$

where $v_p = n_p - \widehat{n}_p$. To test $H_0 : \rho = 0$ against $H_1 : \rho > 0$, it is necessary to compare the test statistic D with the lower critical quantity $d_{L,\alpha}$ and the upper critical quantity $d_{U,\alpha}$. From the Durbin-Watson tables, we find that $d_{L,\alpha} = 0.46$. Since $D = 0.054 < d_{L\alpha} = 0.46$ then we reject H_0 and H_1 is accepted. Thus, there is a positive correlation between the error expressions. Using the suggested algorithm discussed in Sub Sect. 3.1 of Sect. 3, we can estimate ρ and $\widetilde{\widehat{w}}_\rho$ as follows:

Stage(1) Allow:

$$\begin{aligned}\widehat{\rho} &= \arg \min_{\rho\in\{-1,\ldots,-0.01,0,0.01,\ldots,1\}} CV(\rho, \widehat{w}^*_\rho)\\ &= \arg \min_{\rho\in\{-1,\ldots,-0.01,0,0.01,\ldots,1\}} \sum_{p=2}^{147} \left(E_{n_p^{*\rho}} - E_{Q_{n_p^{*\rho}}(\widehat{w}^{*(-p)})} \right)^2,\end{aligned}$$

Stage(2) Allow:

$$\begin{aligned}\widetilde{\widehat{\mathbf{w}}}_{\widehat{\rho}} &= \arg \min_{\widetilde{\mathbf{w}}} \mathrm{g}\left(\widehat{\rho}, \widetilde{\mathbf{w}}\right)\\ &= \arg \min_{w_p, l_{w_p}>0, r_{w_p}>0} \sum_{p=2}^{147} \rho_\eta \left(D\left(\widetilde{n}_p^{*\widehat{\rho}}, \widetilde{Q}_{\widetilde{n}_p^{*\widehat{\rho}}}(\widetilde{w}) \right) \right).\end{aligned}$$

in which $\widetilde{z}_p^{*\rho}$ and $\widetilde{Q}_{\widetilde{z}_p^{*\rho}}(\widetilde{w})$ are given in formulas 5. The evaluated parameters of the suggested regression model for $\widetilde{w}_0, \widetilde{w}_1, ..., \widetilde{w}_6$ and the correlation parameter ρ, along with its proficiency measurements, are presented in Table 2. The estimated values of h, τ, and σ^2, as well as the proficiency measurements of several fuzzy regression methods, are summarized in this Table. Table 2 compares the proficiency measurements of the suggested procedure with those of other methods. When comparing the various methods using proficiency measurements, the results can be summarized as follows: the suggested method shows higher MSM and COD values than the other models ($MSM = 0.812$ and $COD = 0.95$), while its MAE and $MARE$ values ($MAE = 17.25$ and $MARE = 5.914$) are smaller than those of the alternative approaches. The results indicate that the suggested method outperformed the other approaches when the fuzzy error terms were correlated, in contrast to when they were uncorrelated with the dataset.

Table 1: The amounts of v_p in Example 4.1

$*$	47.174	−12.925	−41.593	13.052	35.119	−2.808	11.567
−9.760	50.384	12.772	−48.126	19.108	14.843	−1.504	−19.494
−33.408	7.217	−56.827	32.306	34.693	14.854	73.328	−6.147
−64.662	−12.512	6.525	−11.195	−61.664	5.318	11.951	−2.911
1.037	1.0349	4.701	−7.334	37.145	−8.946	−4.958	−30.636
−13.755	−66.279	−26.242	7.459	−22.402	17.641	102.521	−16.688
9.0367	63.256	−38.072	−18.297	−4.139	−22.634	2.596	7.107
−10.944	−4.097	21.689	−9.537	−64.139	−7.572	5.700	
5.363	6.979	−9.066	33.172	15.726	−14.145	−21.523	
−5.903	−68.833	−7.868	−8.272	19.320	37.813	41.055	
14.133	−41.902	31.805	−1.756	−14.177	43.868	11.641	
−12.083	17.737	−34.047	−39.840	−23.661	−5.278	−16.336	
−0.966	−16.708	15.892	−13.151	24.896	−15.475	−15.009	
−8.042	−55.928	−49.741	6.401	22.916	7.723	−33.547	
9.782	12.459	1.188	35.319	86.028	−17.089	−40.837	
−4.638	0.034	−11.541	212.598	11.041	−22.593	−21.538	
15.903	8.745	6.704	−19.141	16.523	−20.818	−22.114	
−12.300	13.273	−17.47	−37.597	−20.431	34.868	−11.495	
−10.505	−6.835	−1.971	12.994	−13.373	−4.711	72.699	
−34.227	11.022	44.273	24.077	−12.956	28.668	−26.118	

Table 2: The evaluated parameters of the model, along with its analogous proficiency criteria for the suggested method and various fuzzy regression methods in Example 4.1

Method	**Cofficient**	MSM	MAE	COD	$MARE$
Proposed	$\rho = 0.298$	0.812	17.25	0.95	5.914
	$h = 0.5$				
	$\tau = 0.2$				
	$\sigma^2 = 7.58$				
Icen and Demirhan	$\tilde{\tilde{w}}_0 = (-585.95; 250.564)_T$				
	$\tilde{\tilde{w}}_1 = (7.016; 2.693)_T$	0.675	25.928	0.842	13.674
	$\tilde{\tilde{w}}_2 = (-6.925; 1.055)_T$				
	$\tilde{\tilde{w}}_3 = (-52.515; 20.12)_T$				
	$\tilde{\tilde{w}}_4 = (-0.404; 0.056)_T$				
	$\tilde{\tilde{w}}_5 = (14.456; 3.391)_T$				
	$\tilde{\tilde{w}}_6 = (-3.528; 0.895)_T$				
Choi and Buckly	$\tilde{\tilde{w}}_0 = (-498.873; 208.564)_T$				

Method	**Cofficient**	MSM	MAE	COD	$MARE$
	$\tilde{\tilde{w}}_1 = (6.015; 2.693)_T$	0.604	31.251	0.702	18.041
	$\tilde{\tilde{w}}_2 = (-5.744; 2.714)_T$				
	$\tilde{\tilde{w}}_3 = (-64.345; 14.456)_T$				
	$\tilde{\tilde{w}}_4 = (0.084; 0.107)_T$				
	$\tilde{\tilde{w}}_5 = (16.834; 2.543)_T$				
	$\tilde{\tilde{w}}_6 = (-2.562; 0.527)_T$				
	$\tilde{\tilde{w}}_0 = (-604.54; 179.08)_T$				
	$\tilde{\tilde{w}}_1 = (8.002; 3.265)_T$	0.619	30.125	0.76	16.021
	$\tilde{\tilde{w}}_2 = (-5.865; 1.347)_T$				
Kula and Apaydin	$\tilde{\tilde{w}}_3 = (-58.243; 13.845)_T$				
	$\tilde{\tilde{w}}_4 = (1.045; 0.0984)_T$				
	$\tilde{\tilde{w}}_5 = (11.007; 4.45)_T$				
	$\tilde{\tilde{w}}_6 = (-2.934; 1.178)_T$				
	$\tilde{\tilde{w}}_0 = (-384.85; 301.005)_T$				
	$\tilde{\tilde{w}}_1 = (10.776; 3.25)_T$	0.618	29.289	0.762	16.10
	$\tilde{\tilde{w}}_2 = (8.543; 2.143)_T$				
Taheri and Kelkinnama	$\tilde{\tilde{w}}_3 = (-45.976; 14.85)_T$				
	$\tilde{\tilde{w}}_4 = (-0.533; 0.184)_T$				
	$\tilde{\tilde{w}}_5 = (11.967; 4.445)_T$				
	$\tilde{\tilde{w}}_6 = (-2.395; 0.748)_T$				
	$\tilde{\tilde{w}}_0 = (-541.37; 242.78)_T$				
	$\tilde{\tilde{w}}_1 = (6.244; 2.936)_T$	0.663	26.851	0.830	15.025
	$\tilde{\tilde{w}}_2 = (-7.036; 1.317)_T$				
Zeng et al.	$\tilde{\tilde{w}}_3 = (-48.966; 17.745)_T$				
	$\tilde{\tilde{w}}_4 = (-0.327; 0.036)_T$				
	$\tilde{\tilde{w}}_5 = (11.966; 2.995)_T$				
	$\tilde{\tilde{w}}_6 = (-3.895; 1.045)_T$				
	$\tilde{\tilde{w}}_0 = (-620.74; 266.94)_T$				

Choi and Yoon

Method	Cofficient	MSM	MAE	COD	$MARE$
	$\tilde{\tilde{w}}_1 = (7.536; 1.927)_T$	0.68	27.086	0.827	13.559
	$\tilde{\tilde{w}}_2 = (-8.045; 0.925)_T$				
	$\tilde{\tilde{w}}_3 = (-60.829; 15.932)_T$				
	$\tilde{\tilde{w}}_4 = (-0.638; 0.027)_T$				
	$\tilde{\tilde{w}}_5 = (12.953; 4.008)_T$				
	$\tilde{\tilde{w}}_6 = (-3.932; 0.637)_T$				

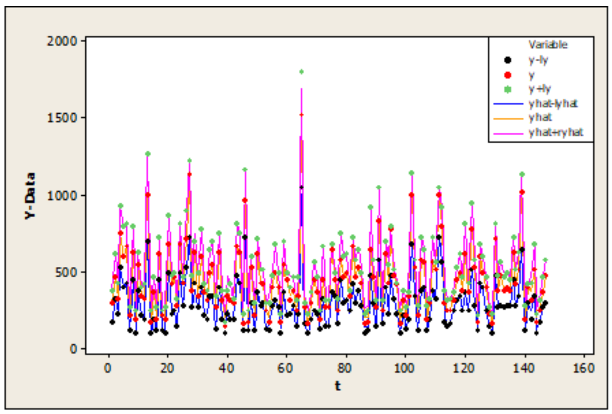


Figure 1: $\widetilde{n}$ and $\widetilde{\widetilde{n}}$ values of the suggested model in Example 4.1

Example 4.2. (House Price Example). The proposed **FSPQSVR** is applied to the house cost model in this example. The house cost data presented in Table 4 is sourced from Hao and Chiang [43]. In the aforesaid dataset, m_1 represents the quality of the materials, m_2 indicates the space on the first floor, m_3 indicates the space on the second floor, and n represents the selling price (in $10,000$ yen). Note that m_1 can have three values, namely, 1, 2, and 3, which represent low, medium, and high grades of materials, respectively. By considering the proposed **FSPQSVR** model in formula 2 and using the proposed algorithm, the currently unidentified parameters associated with ρ and $\widetilde{\widetilde{w}}_\rho$ can be determined as described below:

Stage(1) Allow:

$$\widehat{\rho} = \arg \min_{\rho \in \{-1,\ldots,-0.01,0,0.01,\ldots,1\}} CV(\rho, \widehat{w}^*_\rho)$$
$$= \arg \min_{\rho \in \{-1,\ldots,-0.01,0,0.01,\ldots,1\}} \sum_{p=2}^{15} \left(E_{n_p^{*\rho}} - E_{Q_{n_p^{*\rho}}(\widehat{w}^{*(-p)})} \right)^2,$$

Stage(2) Allow:

$$\widetilde{\widetilde{\mathbf{w}}}_{\widehat{\rho}} = \arg \min_{\widetilde{\mathbf{w}}} \mathrm{g}\left(\widehat{\rho}, \widetilde{\mathbf{w}}\right)$$
$$= \arg \min_{w_p, l_{w_p}>0, r_{w_p}>0} \sum_{p=2}^{15} \rho_\eta \left(D\left(\widetilde{n}_p^{*\widehat{\rho}}, \widetilde{Q}_{\widetilde{n}_p^{*\widehat{\rho}}}(\widetilde{w}) \right) \right).$$

The values of v_p obtained from the ordinary least squares method are presented in Table 3, where $v_p = n_p - \widehat{n}_p$. We obtain the Durbin-Watson test statistic as $D = 0.86034$. Furthermore, we want to test the hypothesis $H_0 : \rho = 0$ in contrast to the hypothesis $H_1 : \rho > 0$. However, the value of D is not smaller than $d_{L,\alpha} = 0.82$ and is not greater than $d_{U,\alpha} = 1.75$. Therefore, the test for positive correlation is ineffective. Then, we examine $H_0 : \rho = 0$ against $H_1 : \rho < 0$. Since $(4 - D) = 3.13966 > d_{U,\alpha} = 1.75$, the error terms are uncorrelated. Table 5 summarizes the performance measurements for several fuzzy regression models. The estimated parameters of the suggested regression model for $\widetilde{w}_0, \widetilde{w}_1, \widetilde{w}_2, \widetilde{w}_3$, h, τ, σ^2, and ρ in addition to its proficiency measurements, are summarized in the same Table. In this example, the proficiency measures are $MSM = 0.89$, $MAE = 11.927$, $COD = 0.962$ and $MARE = 14.108$, respectively. While the MAE value of 11.927 might appear numerically significant, it is important to consider the inherent variability and fuzziness of the house price data, as well as the nature of the fuzzy error expressions in this real-world application. The data's characteristics, including its range and the presence of diverse material qualities (n_1) and floor spaces (n_2, n_3), contribute to the overall scale of the prediction errors. Nevertheless, as indicated by the high COD value of 0.962, our model explains a substantial portion of the variance in the selling prices, demonstrating its strong predictive capability within this context. Furthermore, as summarized in Table 5, the proposed method consistently achieves a lower MAE compared to other established fuzzy regression methods, indicating its superior performance in addressing this specific estimation problem. Hence, the suggested method performs better than other fuzzy regression methods in situations with correlated fuzzy error expressions.

Table 3: The amounts of v_p in Example 4.2

Observation	v_p	Observation	v_p
1	48.994	9	−57.445
2	46.901	10	−49.116
3	4.202	11	−13.025
4	−13.841	12	−33.764
5	−3.366	13	67.344
6	−14.226	14	36.094
7	−23.373	15	26.240
8	−21.615	-	-

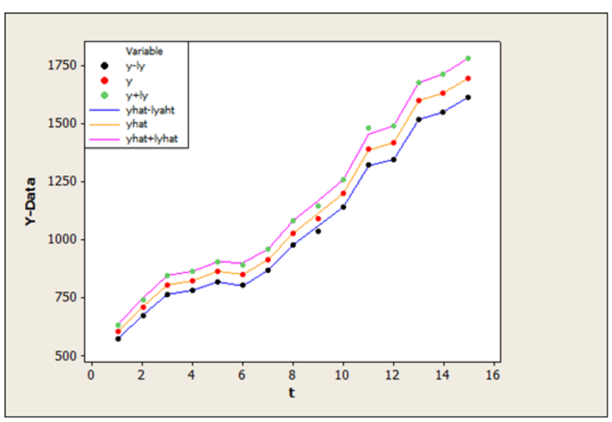


Figure 2: $\widetilde{n}$ and $\widetilde{\widehat{n}}$ values of the suggested model in Example 4.2

Table 4: Dataset used in Example 4.2

Observation	$\tilde{n}$	m_1	m_2	m_3
1	$(606; 30.30)_T$	1	38.09	36.43
2	$(710; 35.50)_T$	1	62.10	25.50
3	$(808; 40.40)_T$	1	63.76	44.71
4	$(826; 41.30)_T$	1	74.52	38.09
5	$(865; 43.25)_T$	1	75.38	41.40
6	$(852; 42.60)_T$	2	52.99	26.49
7	$(917; 45.85)_T$	2	62.93	26.49
8	$(1031; 51.55)_T$	2	72.04	33.12
9	$(1092; 54.60)_T$	2	76.12	43.06
10	$(1203; 60.15)_T$	2	90.26	42.64
11	$(1394; 69.70)_T$	3	85.70	31.33
12	$(1420; 71.00)_T$	3	95.27	27.64
13	$(1601; 80.05)_T$	3	105.98	27.64
14	$(1632; 81.60)_T$	3	79.25	66.81
15	$(1699; 84.95)_T$	3	120.50	32.25

Table 5: The evaluated parameters of the model, along with its analogous proficiency criteria for the suggested method and various fuzzy regression methods in Example 4.2

Method	Cofficient	MSM	MAE	COD	$MARE$
Proposed	$\rho = 0.9$	0.89	11.927	0.962	14.108
	$h = 5$				
	$\tau = 0.08$				
	$\sigma^2 = 5.5$				
Icen and Demirhan	$\tilde{\tilde{w}}_0 = (-258.07; 51.778)_T$				
	$\tilde{\tilde{w}}_1 = (278.554; 6.879)_T$	0.684	16.372	0.820	30.528
	$\tilde{\tilde{w}}_2 = (7.987; 0.574)_T$				
	$\tilde{\tilde{w}}_3 = (5.255; 0.285)_T$				
Choi and Yoon	$\tilde{\tilde{w}}_0 = (-258.07; 42.773)_T$				
	$\tilde{\tilde{w}}_1 = (255.82; 5.772, 1233498)_T$	0.725	15.268	0.836	28.386
	$\tilde{\tilde{w}}_2 = (10.554; 0.623)_T$				
	$\tilde{\tilde{w}}_3 = (8.224; 0.232)_T$				
Kula and Apaydin	$\tilde{\tilde{w}}_0 = (-233.38; 7.551)_T$				
	$\tilde{\tilde{w}}_1 = (266.82; 4.156)_T$	0.63	18.113	0.729	33.201
	$\tilde{\tilde{w}}_2 = (9.262; 0.573)_T$				

Method	Cofficient	MSM	MAE	COD	$MARE$
	$\tilde{\tilde{w}}_3 = (5.7128; 0.427)_T$				
Taheri and Kelkinnama	$\tilde{\tilde{w}}_0 = (-24325; 12.556)_T$				
	$\tilde{\tilde{w}}_1 = (239.4526.82; 4.156)_T$	0.627	18.923	0.722	33.982
	$\tilde{\tilde{w}}_2 = (11.852; 0.652)_T$				
	$\tilde{\tilde{w}}_3 = (7.526; 0.329)_T$				
Zeng et al.	$\tilde{\tilde{w}}_0 = (-268.65; 153)_T$				
	$\tilde{\tilde{w}}_1 = (248.225; 166)_T$	0.763	14.605	0.863	25.336
	$\tilde{\tilde{w}}_2 = (725858; 0.296)_T$				
	$\tilde{\tilde{w}}_3 = (492368; 0.554)_T$				
Choi and Buckley	$\tilde{\tilde{w}}_0 = (-236.884; 20.334)_T$				
	$\tilde{\tilde{w}}_1 = (220.884; 6.656)_T$	0.643	17.957	0.767	31.925
	$\tilde{\tilde{w}}_2 = (10.565; 0.366)_T$				
	$\tilde{\tilde{w}}_3 = (8.256; 0.321)_T$				

Example 4.3. (Simulation study). In this case, we examine the following fuzzy nonlinear regression model via 10 simulated data sets of size $j = 150$:

Stage 1:

1. Generate a random sample of size $j = 150$ using $LR-$**FN**s listed below:
2. Assume that $\widetilde{s_p} = (|m_p - 1|/4 + 2|\sin(\pi(1 + (m_p - 1)/4))| + 1 + \epsilon_p; |0.1\exp(-|m_p|/50) + \epsilon'_p|)_L$ in which

(a) $\epsilon_p \sim N(0, 0.01)$ and $\epsilon'_p \sim N(0, 0.50)$.

(b) A random sample denoted as m_p is taken from the uniform distribution defined as $U(-10, 10)$.

(c) $L(m) = \sqrt{1 - m^2}$.

Stage 2:

1. For a randomly chosen $P_1 = \{p_1^1, p_2^1, ..., p_{10}^1\} \subseteq \{1, 2, 3, ..., 150\}$, Allow $\widetilde{n}_p = \widetilde{s_p} \oplus P\{10\}$, $p \in P_1$.
2. For a randomly chosen $P_2 = \{p_1^2, p_2^2, ..., p_{10}^2\} \in \{1, 2, 3, ..., 150\} - P_1$, Allow $\widetilde{n}_p = \widetilde{s_p} \oplus (0; 1)_L$, $p \in P_2$.

Stage 3: Assume that $\widetilde{n}_p = \widetilde{s_p}$ for $p \in \{1, 2, 3, ..., 150\} - (P_1 \cup P_2)$.

We use the classic Durbin-Watson test to evaluate $H_0 : \rho = 0$ against $H_1 : \rho > 0$. From formula 4, we can easily deduce that the test statistic D is equal to 0.478744. Since D falls below $d_{L,\alpha} = 1.65$ with level of significance $\alpha = 0.05$, the correlation between the error expressions is positive. We compare the suggested method with various popular fuzzy regression techniques, and the results of their performance measurements are summed up in Table 6. Moreover, the performance measures of the suggested method and the estimated quantities of $\widetilde{w}_p$ and ρ are represented in Table 6. The evaluated values of h, τ, and σ^2 are also given in this Table. Regarding the $\overline{MAE}$ of 22.508 for the average model and 19.254 for the best model in the simulation study, it is crucial to recognize that the simulated data, generated with specific $LR-$**FN**s and defined error terms ($\epsilon_p \sim N(0, 0.01)$ and $\epsilon'_p \sim N(0, 0.50)$), inherently contains a certain level of uncertainty and fuzziness as described in Stage 1 and Stage 2 of the data generation process. This intrinsic variability naturally contributes to the magnitude of the observed MAE values. Despite these intrinsic data characteristics, the high $\overline{COD}$

values (e.g., 0.722 for the average model) demonstrate that our model effectively captures the underlying relationships in the simulated fuzzy nonlinear regression model. More critically, as illustrated in Table 6, our proposed method consistently shows improved MAE values compared to other fuzzy regression techniques across the simulated datasets, confirming its robust performance under these challenging conditions. As a result, it is evident that the newly implemented procedure outperformed the other approaches in this instance.

Table 6: The evaluated parameters of the model, along with its analogous proficiency criteria for the suggested method and various fuzzy regression methods in Example 4.3.

Method	**Cofficient**	$\overline{MSM}$	$\overline{MAE}$	$\overline{COD}$
Proposed	**Average model**	0.513	22.508	0.722
	Best Model $\rho = 0.412$ $h = 5$ $\tau = 0.08$ $\sigma^2 = 5.5$	$MSM = 0.597$	$MAE = 19.254$	$COD = 0.746$

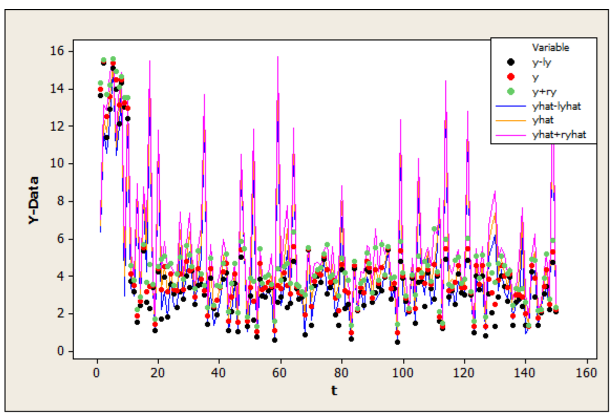


Figure 3: $\widetilde{n}$ and $\widetilde{\widetilde{n}}$ values of the suggested model in Example 4.3

5 Conclusion

In this paper, the main goal was to introduce a new approach for a semiparametric quantile-based regression model, which integrates a support vector machine approach to enhance its estimation capabilities, precise regressors, and fuzzy outcomes. Since the assumption that fuzzy error terms are always independent in fuzzy regression models is unrealistic, we study the proposed method in situations where there is a correlation among the fuzzy residual expressions. To do this, we applied a generalized difference and absolute error distance criterion to LR-fuzzy quantities that are not symmetric. Depending on the sum of absolute residuals and cross-validation measurements, we suggested a hybrid approach to estimate the optimum amounts of the autocorrelation measure and fuzzy multipliers. In order to detect the presence of correlation between fuzzy error terms, we employed the traditional Durbin-Watson test. Additionally, we examined the efficiency of the suggested regression model based on the quantile method and contrasted it with other well-founded methods using various widely used goodness-of-fit metrics. According to the numerical results, the suggested fuzzy semi-parametric quantile-based regression method produces more accurate results than alternative methods. Additionally, the outcomes demonstrated that when there was autocorrelation among fuzzy error expressions, the recommended strategy performed better. Future work could focus on exploring the existence of extreme points in the dataset. Additionally, expanding the suggested approach to an **FSPQSVR** procedure when the regressors and outcomes are fuzzy datasets could be another subject for future research.

Acknowledgements: The authors would like to thank referee for his constructive comments and suggestions which have enhanced the quality of the paper.

Conflict of Interest: The authors declare no conflict of interest.

References

[1] Hastie T, Tibshirani R, Friedman J. *The Elements of Statistical Learning.* Springer; 2009.

[2] Vapnik VN. *The Nature of Statistical Learning Theory.* Springer; 1995.

[3] Vapnik VN. *Statistical Learning Theory.* Wiley; 1998.

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

[5] Gu B, Sheng VS, Tay KY, Romano W, Li S. Cross validation through two-dimensional solution surface for cost-sensitive SVM. *IEEE Transactions on Pattern Analysis and Machine Intelligence.* 2016; 39(6): 1103-1121. DOI: https://doi.org/10.1109/TPAMI.2016.2578326

[6] Koenker R, Bassett JRG. Regression quantiles. *Econometrica: Journal of the Econometric Society.* 1978; 46(1): 33-50. DOI: https://doi.org/10.2307/1913643

[7] Koenker R. *Quantile Regression.* Cambridge University Press; 2005.

[8] Coppi R. Management of uncertainty in statistical reasoning: The case of regression analysis. *International Journal of Approximate Reasoning.* 2008; 47(3), 284-305. DOI: https://doi.org/10.1016/j.ijar.2007.05.011

[9] Zadeh LA. Fuzzy sets. *Information and Control.* 1965; 8(3): 338-353. DOI: https://doi.org/10.1016/S0019-9958(65)90241-X

[10] Taheri SM. Trends in fuzzy statistics. *Austrian Journal of Statistics.* 2003; 32(3): 239-257. DOI: https://doi.org/10.17713/ajs.v32i3.459

[11] Zendehdel J, Rezaei M, Akbari MG, Zarei R, Alizadeh Noughabi H. Testing exponentiality for imprecise data and its application. *Soft Computing.* 2018; 22: 3301-3312. DOI: https://doi.org/10.1007/s00500-017-2566-y

[12] Tanaka H, Uejima S, Asai K. Linear regression analysis with fuzzy model. *IEEE Transaction on Systems, Man, and Cybernetics.* 1982; 12(6): 903-907. DOI: https://doi.org/10.1109/TSMC.1982.4308925

[13] Kacprzyk J, Fedrizzi M. *Fuzzy Regression Analysis.* Warsaw : Omnitech Press; 1992.

[14] Naderkhani R, Behzad MH, Razzaghnia T, Farnoosh R. Fuzzy regression analysis based on fuzzy neural networks using trapezoidal data. *International Journal of Fuzzy Systems.* 2021; 23(5): 1267-1280. DOI: https://doi.org/10.1007/s40815-020-01033-2

[15] Chachi J, Taheri SM, DUrso P. Fuzzy regression analysis based on M-estimates. *Expert Systems with Applications.* 2022; 187: 115891. DOI: https://doi.org/10.1016/j.eswa.2021.115891

[16] Yoon JH, Kim DJ, Koo YY. Novel fuzzy correlation coefficient and variable selection method for fuzzy regression analysis based on distance approach. *International Journal of Fuzzy Systems.* 2023; 25(8): 2969-2985. DOI: https://doi.org/10.1007/s40815-023-01546-6

[17] Baha Alwan, E., Abdulmohsin Ali, O. About semi-parametric methodology for fuzzy quantile regression model estimation: A review. *Journal of Economics and Administrative Sciences.* 2024; 29(138): 127-146.

[18] Hesamian G, Akbari MG. Fuzzy quantile linear regression model adopted with a semi-parametric technique based on fuzzy predictors and fuzzy responses. *Expert Systems with Applications.* 2019; 118: 585-597. DOI: https://doi.org/10.1016/j.eswa.2018.10.026

[19] Arefi M. Quantile fuzzy regression based on fuzzy outputs and fuzzy parameters. *Soft Computing.* 2020; 24(1): 311-320. DOI: https://doi.org/10.1007/s00500-019-04424-2

[20] Chachi J, Chaji A. An OWA-based approach to quantile fuzzy regression. *Computers & Industrial Engineering.* 2021; 159: 107498. DOI: https://doi.org/10.1016/j.cie.2021.107498

[21] Khammar AH, Arefi M, Akbari MG. Quantile fuzzy varying coefficient regression based on kernel function. *Applied Soft Computing.* 2021; 107: 107313. DOI: https://doi.org/10.1016/j.asoc.2021.107313

[22] Hesamian G, Akbari MG. A fuzzy empirical quantile-based regression model based on triangular fuzzy numbers. *Computational and Applied Mathematics.* 2022; 41(6): 267. DOI: https://doi.org/10.1007/s40314-022-01974-4

[23] Hesamian G, Akbari MG, Asadollahi M. Fuzzy semi-parametric partially linear model with fuzzy inputs and fuzzy outputs. *Expert Systems with Applications.* 2017; 71: 230-239. DOI: https://doi.org/10.1016/j.eswa.2016.11.032

[24] Akbari MG, Hesamian G. Elastic net oriented to fuzzy semiparametric regression model with fuzzy explanatory variables and fuzzy responses. *IEEE Transactions on Fuzzy Systems.* 2019; 27(12): 2433-2442. DOI: https://doi.org/10.1109/TFUZZ.2019.2900603

[25] Zadeh LA. Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems.* 1978; 1(1): 3-28. DOI: https://doi.org/10.1016/0165-0114(78)90029-5

[26] Dubois D, Prade H. Operations on fuzzy numbers. *International Journal of Systems Science.* 1978; 9(6): 613-626. DOI: https://doi.org/10.1080/00207727808941724

[27] Lai YJ, Hwang CL. Fuzzy mathematical programming. In: Fuzzy Mathematical Programming. *Lecture Notes in Economics and Mathematical Systems.* Springer, Berlin, Heidelberg; 1992. p.74-186. DOI: https://doi.org/10.1007/978-3-642-48753-8_3

[28] Hesamian G, Akbari MG. Fuzzy absolute error distance measure based on a generalised difference operation. *International Journal of Systems Science.* 2018; 49(11): 2454-2462. DOI: https://doi.org/10.1080/00207721.2018.1505002

[29] Puri ML, Ralescu DA. Differentials of fuzzy functions. *Journal of Mathematical Analysis and Applications.* 1983; 91(2): 552-558. DOI: https://doi.org/10.1016/0022-247X(83)90169-5

[30] Gyrfi L, Kohler M, Krzyzak A, Walk H. *A Distribution-Free Theory of Nonparametric Regression.* Springer Series in Statistics; 2006.

[31] Wasserman L. *All of Nonparametric Statistics.* Springer New York; 2006.

[32] Kutner MH, Nachtsheim CJ, Neter J. *Applied Linear Regression Models.* New York: McGraw-Hill Irwin; 2004.

[33] Durbin J, Watson GS. Testing for serial correlation in least squares regression. III. *Biometrika.* 1971; 58(1): 1-19. DOI: https://doi.org/10.1093/biomet/58.1.1

[34] Chatterjee S, Simonoff JS. *Handbook of Regression Analysis.* John Wiley & Sons; 2013.

[35] Akbari MG, Hesamian G. A fuzzy linear regression model with autoregressive fuzzy errors based on exact predictors and fuzzy responses. *Computational and Applied Mathematics.* 2022; 41(6): 284. DOI:https://doi.org/10.1007/s40314-022-01994-0

[36] Choi SH, Yoon JH. General fuzzy regression using least squares method. *International Journal of Systems Science.* 2010; 41(5): 477-485. DOI: https://doi.org/10.1080/00207720902774813

[37] Choi SH, Buckley JJ. Fuzzy regression using least absolute deviation estimators. *Soft computing.* 2008; 12: 257-263. DOI: https://doi.org/10.1007/s00500-007-0198-3

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

[39] Taheri SM, Kelkinnama M. Fuzzy linear regression based on least absolutes deviations. *Iranian Journal of Fuzzy Systems.* 2012; 9(1):121-140. DOI: https://doi.org/10.22111/ijfs.2012.230

[40] Kula KS, Apaydin A. Fuzzy robust regression analysis based on the ranking of fuzzy sets. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems.* 2008; 16(05): 663-681. DOI: https://doi.org/10.1142/S0218488508005558

[41] Icen D, Demirhan H. Error measures for fuzzy linear regression: Monte Carlo simulation approach. *Applied Soft Computing.* 2016; 46: 104-114. DOI: https://doi.org/10.1016/j.asoc.2016.04.013

[42] Zhou J, Zhang H, Gu Y, Pantelous AA. Affordable levels of house prices using fuzzy linear regression analysis: The case of Shanghai. *Soft Computing.* 2018; 22: 5407-5418. DOI: https://doi.org/10.1007/s00500-018-3090-4

[43] Hao PY, Chiang JH. Fuzzy regression analysis by support vector learning approach. *IEEE Transactions on Fuzzy Systems.* 2008; 16(2): 428-441. DOI: https://doi.org/10.1109/TFUZZ.2007.896359

Batoul Salamah
Department of Applied Mathematics
Faculty of Mathematical Sciences
University of Guilan
Rasht, Iran
E-mail: batoul.salamah91@gmail.com

Reza Zarei
Department of Statistics
Faculty of Mathematical Sciences
University of Guilan
Rasht, Iran
E-mail: r.zarei@guilan.ac.ir, rezazarei.r@gmail.com

Farshid Mehrdoust
Department of Applied Mathematics
Faculty of Mathematical Sciences
University of Guilan
Rasht, Iran
E-mail: fmehrdoust@guilan.ac.ir

Mohammad Ghasem Akbari
Department of Statistics
Faculty of Mathematical Sciences
University of Birjand
Birjand, Iran
E-mail: g_z_akbari@birjand.ac.ir