Journal of Data Science and Modeling, Vol. 3, No. 2, 87-114, June 2025

Research Manuscript

Nonparametric Modeling of Fuzzy Time Dependent Data Based on Support Vector Machine

Reza Zarei[1], Mohammad Ghasem Akbari[2]*

[1]Department of Statistics, Faculty of Mathematical Sciences, University of Guilan, Rasht, Iran.

[2]Department of Statistics, Faculty of Mathematical Sciences, University of Birjand, Birjand, Iran.

Recieved: 27/03/2026 Accepted: 30/08/2026

Abstract:

This paper proposes a novel nonparametric support vector machine (SVM)-based approach for modeling and predicting fuzzy time-dependent data. Unlike traditional parametric methods such as ARIMA and linear regression, the proposed framework leverages SVM flexibility to capture complex, nonlinear temporal patterns in observations represented as fuzzy numbers. The nonparametric approach provides flexibility, robustness to model misspecification, and the ability to accommodate irregular patterns and outliers common in fuzzy time series. A practical example from software development quality monitoring motivates the approach, where conventional SVM methods cannot directly handle fuzzy observations or uncertainty captured by spreads. The model is evaluated using both simulated fuzzy time series and a real-world dataset based on three criteria: a similarity-based measure (MSM) for overall fuzzy similarity, root mean squared error (RMSE) for center accuracy, and average spread error (ASE) for spread accuracy. Results demonstrate that the proposed method consistently outperforms existing methods across all three metrics. Furthermore, comparison with a linear SVM confirms the necessity of the nonlinear kernel for capturing complex temporal patterns, while the BDS test confirms nonlinear structure in the real dataset. The findings also indicate that the fitted models remain stable in the presence of outliers. Overall, the proposed SVM-based nonparametric framework provides a powerful, accurate, and robust tool for modeling fuzzy time-dependent data.

Mathematics Subject Classification (2010): 91G70, 03E72, 62P05.

*Corresponding Author: g_z_akbari@birjand.ac.ir

1. Introduction

The field of time series analysis comprises methods to analyze the characteristics of a response variable with respect to time. It takes into consideration the fact that observations made over time may have an internal structure (such as autocorrelations, trends, seasonal and/or cyclic variations) that should be accounted for. Time series analysis has two main objectives: (1) to identify the nature of the phenomenon represented by the sequence of observations, and (2) to predict future values of the response variable (Box and Jenkins, 1976; Brockwell and Davis, 2009).

Modeling and prediction of time series have a wide range of applications across various disciplines. In business, economics, and finance, time series methods are essential for forecasting and decision-making (see, e.g., Box and Jenkins (1976); Palma (2016); Shumway and Stoffer (2017)). In engineering and the physical sciences, these methods are used for signal processing and system identification (Brockwell and Davis (2009); Tong (1990)). Additionally, applications in medicine (Woodward et al. (2012)), environmental studies, and interdisciplinary research have also been extensively developed (Efromovich (1999); Friedman and Stuetzle (1981); Byun and Lee (2002)).

Traditional parametric time series models, such as ARIMA Box and Jenkins (1976) and linear regression, rely on exact observations and lead to crisp predictions. However, due to various conditions of uncertainty, it is often preferable to predict a quantity via imprecise values modeled via methods of fuzzy statistics Viertl (2011); Taheri (2003). For instance, as noted by Yang and Liu (2019), we usually observe imprecise observations in carbon emissions, social benefits, and oil reserves, among others. Classical statistical time series models fail to address prediction problems based on ambiguous or vague information represented by fuzzy data Viertl (2011). This shortcoming can be overcome by time series models that are based on techniques of fuzzy statistics. In this respect, fuzzy time series models originally introduced by Song and Chissom (1993) have been successfully replacing traditional approaches when there is uncertainty in the observations.

The support vector machine (SVM) is a learning technique that helps to mitigate common issues associated with artificial neural networks, such as over-learning, under-learning, and convergence to local minima of the network structure (Vapnik, 1995, 1998). One of the most popular research fields in recent years has been the prediction of time series and regression using SVM Asadolahi et al. (2021); Vilela et al. (2019). The SVMs presented for time series prediction include a wide range of real-world application domains, including forecasting financial markets, electric load forecasting, and other scientific disciplines, as surveyed by Sapankevych and Sankar (2009). SVM is based on the structural risk minimiza-

tion principle, which is implemented using the maximum margin concept Ozer et al. (2011).

The choice of a nonparametric approach in this paper is motivated by three key considerations. First, the relationship between time and fuzzy observations is often unknown and potentially nonlinear; parametric models impose restrictive functional forms that may fail to capture complex temporal patterns. Second, nonparametric methods are inherently more robust to model misspecification, as they do not rely on strong assumptions about the underlying data-generating process. Third, fuzzy time series data often exhibit irregular patterns, outliers, and varying uncertainty structures that are difficult to accommodate with simple parametric models. However, nonparametric methods also have limitations: they typically require larger sample sizes to achieve reliable estimates, involve higher computational costs due to optimization procedures, and may produce less interpretable results compared to parametric alternatives. In the context of fuzzy time-dependent data, we believe the advantages of flexibility, robustness, and adaptability outweigh these disadvantages, particularly when the primary goal is accurate prediction rather than parameter interpretation.

Our knowledge of the majority of physical processes is mostly reliant on inaccurate human reasoning and imprecise conceptions. There are many real-world applications that deal with inaccurate information rather than exact data. Consider a study that predicts the heavy metal concentration or pH of the groundwater in a specific area, for example. Since these amounts are frequently impacted by a variety of environmental conditions, it might be challenging to assess the precision of the heavy metal concentration or pH values in such a situation. This shows that representing data as fuzzy quantities is preferable. On the other hand, fuzzy set theory is a key technique that has been put forth during the past few decades for modeling uncertainty and imprecision in many statistical contexts. Foundational work in this area includes contributions by Akbari and Rezaei (2010) on bootstrap testing, Taheri (2003) on trends in fuzzy statistics, Viertl (2011) on statistical methods for fuzzy data, Zendehdel et al. (2018) on testing exponentiality with imprecise data, and recent advances in fuzzy time series forecasting by Wang et al. (2024).

To illustrate the need for the proposed methodology, consider a software development project where monthly evaluations are conducted by a team of beta testers. The testers provide imprecise assessments of performance, reliability, and usability, naturally represented as fuzzy numbers of the form $\tilde{y}_i = (m_i; l_i, r_i)_T$. These evaluations are collected over time, forming a fuzzy time series with temporal dependence, potential nonlinear patterns, and occasional outliers. Standard support vector machines (SVMs) are designed for real-valued inputs and outputs

and cannot directly handle fuzzy observations, fuzzy coefficients, or the uncertainty captured by spreads. Moreover, classical SVM theory assumes independent data, whereas time series data exhibit temporal dependence that requires special attention in bandwidth selection, asymptotic properties, and model validation.

To overcome these limitations, the present paper proposes a novel nonparametric SVM-based approach for modeling fuzzy time-dependent data. The proposed framework leverages the α-pessimistic representation to transform fuzzy data into a family of real-valued representations, employs a three-stage estimation procedure to separately model the center and spreads of fuzzy observations, and extends the SVM framework to handle fuzzy responses and fuzzy coefficients. Unlike conventional methods that treat observations as independent, our approach captures temporal dependence through a kernel-based regression structure with bandwidth selection via generalized cross-validation. The methodology is designed to be flexible, robust to outliers, and capable of capturing nonlinear temporal patterns without relying on restrictive parametric assumptions.

The rest of the paper is organized as follows: some primary concepts related to fuzzy random variables and a brief explanation of support vector machines are presented in Section 2. In order to describe the computational details and to assess the effectiveness and performance of the proposed methods, two numerical and real-world data analyses are provided in Section 4. Finally, in Section 5, the main contributions of this paper are summarized.

Summary of Notation

To facilitate reading, we provide a summary of the main notation used throughout this paper. Table 1 lists the key symbols, their definitions, and the sections where they first appear.

2. Preliminary concepts

In this section, we recall the main concepts and definitions about fuzzy numbers and fuzzy random variables that will be used throughout the paper.

2.1 Fuzzy numbers

A fuzzy subset $\tilde{A}$ of $\mathbb{R}$ characterized by membership function $\eta_{\tilde{A}} : \mathbb{R} \to [0, 1]$ is said to be a fuzzy number if $\tilde{A}$ is convex, normal, and upper semi-continuous. In this paper, we denote the class of fuzzy numbers by $F(\mathbb{R})$. In addition, the α-cut of $\tilde{A}$ for $\alpha \in (0, 1]$ is denoted by $\tilde{A}[\alpha] = \{y : \eta_{\tilde{A}}(y) \geq \alpha\}$.

For a fuzzy number $\tilde{N} \in F(\mathbb{R})$, let $\tilde{N}^L[\alpha]$ and $\tilde{N}^U[\alpha]$ denote the lower and upper endpoints of the α-cut $\tilde{N}[\alpha]$, respectively, where $\tilde{N}[\alpha] = [\tilde{N}^L[\alpha], \tilde{N}^U[\alpha]]$. These endpoints are well-defined for all $\alpha \in (0, 1]$ due to the convexity and upper semi-continuity of fuzzy numbers.

Throughout this paper, we use two related but distinct notations for representing fuzzy numbers at different levels:

- The α-cut of a fuzzy number $\tilde{N}$, denoted by $\tilde{N}[\alpha]$, is the crisp set of all elements with membership degree at least α: $\tilde{N}[\alpha] = \{y : \eta_{\tilde{N}}(y) \geq \alpha\}$.
- The α-pessimistic representation of a fuzzy number $\tilde{N}$, denoted by $\tilde{N}_\alpha$, is a transformation that maps the fuzzy number to a real number for each $\alpha \in [0, 1]$.

These two representations are related through the following identity, which follows from the definition of the α-pessimistic representation:

$$\tilde{N}[\alpha] = [\tilde{N}^L[\alpha], \tilde{N}^U[\alpha]] = [\tilde{N}_{\alpha/2}, \tilde{N}_{1-\alpha/2}].$$

In other words, the lower endpoint of the α-cut is obtained from the $\alpha/2$-pessimistic value, and the upper endpoint is obtained from the $(1-\alpha/2)$-pessimistic value. This relationship allows us to move between the two representations as needed.

Definition 2.1. *(Kratschmer (2006)). For a given $\widetilde{N} \in F(\mathbb{R})$, the mapping $\widetilde{N}_\alpha : [0, 1] \to \mathbb{R}$ is called the α-pessimistic representation of $\widetilde{N}$, defined by*

$$\begin{aligned} \widetilde{N}_\alpha &= \begin{cases} \widetilde{N}^L[2\alpha], & 0 \leq \alpha \leq 0.5, \\ \widetilde{N}^U[2(1-\alpha)], & 0.5 < \alpha \leq 1.0. \end{cases} \\ &= \begin{cases} \inf_{\beta \in I_{2\alpha}} \widetilde{N}_\beta, & 0 \leq \alpha \leq 0.5, \\ \sup_{\beta \in I_{2(1-\alpha)}} \widetilde{N}_\beta, & 0.5 < \alpha \leq 1.0. \end{cases} \end{aligned}$$

where $I_\alpha = [\frac{\alpha}{2}, 1 - \frac{\alpha}{2}]$.

The α-pessimistic representation transforms a fuzzy number into a family of real-valued functions indexed by α. This transformation is essential because it allows us to apply classical probabilistic and statistical methods to fuzzy data. By working with the α-pessimistic values, we can define expectations, variances, quantiles, and other statistical quantities for fuzzy random variables in a way that preserves the ordering structure of the fuzzy numbers.

Table 1: Summary of notation

Symbol	Definition	First Appearance
$\tilde{A}$	Fuzzy number	Section 2.1
$F(\mathbb{R})$	Class of all fuzzy numbers	Section 2.1
$\eta_{\tilde{A}}(x)$	Membership function of fuzzy number $\tilde{A}$	Section 2.1
$\tilde{A}[\alpha]$	α-cut of fuzzy number $\tilde{A}$	Section 2.1
$\tilde{A}^L[\alpha]$	Lower endpoint of α-cut	Section 2.1
$\tilde{A}^U[\alpha]$	Upper endpoint of α-cut	Section 2.1
$\tilde{N}_\alpha$	α-pessimistic representation of fuzzy number $\tilde{N}$	Definition 2.1
$(m; l, r)_{LR}$	LR-type fuzzy number with center m, left spread l, right spread r	Example 2.2
$\tilde{Y} : \Omega \to \mathcal{F}(\mathbb{R})$	Fuzzy random variable	Definition 2.4
$(\tilde{F}_{\tilde{X}}(x))_\alpha$	α-pessimistic fuzzy CDF	Definition 2.9
$\tilde{Q}_{\tilde{Y}}(\tau)$	Fuzzy quantile function	Definition 2.11
$K_h(t_j, t)$	Kernel function with bandwidth h	Equation (3.2)
$\tilde{w}_j$	Fuzzy coefficient in SVM model	Equation (3.2)
$\tilde{Z}_t$	Fuzzy observation at time t	Section 3.2
ϵ_t	Error term at time t	Equation (3.2)
c_1, c_2, c_3	Regularization parameters in SVM optimization	Section 3.2
τ	Quantile level	Section 3.2
MSM	Mean Similarity Measure	Section 4
$RMSE$	Root Mean Squared Error	Section 4
ASE	Average Spread Error	Section 4
h	Bandwidth (smoothing) parameter	Equation (3.2)

Example 2.2. *Consider the triangular fuzzy number* $\tilde{N} = (2; 0.5, 0.5)_{LR}$ *with* $L(x) = R(x) = \max(0, 1 - x)$. *Then:*

- *For* $\alpha = 0.25$*:* $\tilde{N}_{0.25} = 2 - 0.5 \times (1 - 2(0.25)) = 2 - 0.25 = 1.75$
- *For* $\alpha = 0.75$*:* $\tilde{N}_{0.75} = 2 + 0.5 \times (1 - 2(1 - 0.75)) = 2 + 0.25 = 2.25$

This shows how the α*-pessimistic representation maps the fuzzy number to specific real values based on the pessimism level* α.

Remark 2.1. *Note that for* $\alpha = 1$*, the interval* I_1 *degenerates to the single point* $I_1 = [0.5, 0.5]$*, which corresponds to the median of the* α*-pessimistic representation. In practice, one typically considers* $\alpha \in (0, 1)$ *for non-degenerate intervals, as this provides a proper interval of integration and avoids degeneracy issues.*

It is clear that $\tilde{N}_\alpha$ is a non-decreasing function of $\alpha \in (0, 1]$. Moreover, it is easy to check that

$$\tilde{N}[\alpha] = [\tilde{N}^L[\alpha], \tilde{N}^U[\alpha]] = [\tilde{N}_{\alpha/2}, \tilde{N}_{1-\alpha/2}],$$

where $\tilde{N}_{\alpha/2}$ and $\tilde{N}_{1-\alpha/2}$ denote the α-pessimistic values at levels $\alpha/2$ and $1 - \alpha/2$, respectively.

Remark 2.3. *Suppose that* $\widetilde{N} = (n; l_n, r_n)_{LR}$ *is a LR fuzzy number. Then, we can easily obtain the* α*-pessimistic representation of* $\widetilde{N}$ *as follows:*

$$\widetilde{N}_\alpha = \left\{ \begin{array}{ll} n - l_n L^{-1}(2\alpha), & 0 \leq \alpha \leq 0.5, \\ n + r_n R^{-1}(2(1-\alpha)), & 0.5 < \alpha \leq 1.0. \end{array} \right\}$$

2.2 Fuzzy random variables

In this section, we list some definitions and theorems for a fuzzy random variable and its main probabilistic concepts.

Definition 2.4. *(Kratschmer (2006)). Let* $(\Omega, \mathcal{B}, \mathcal{P})$ *be a probability space. Consider the fuzzy-valued mapping* $\tilde{Y} : \Omega \to \mathcal{F}(\mathbb{R})$*. The mapping* $\tilde{Y}$ *is a fuzzy random variable (FRV) if:*

1. $\tilde{Y}$ *is measurable with respect to the Borel* σ*-algebra on* $\mathcal{F}(\mathbb{R})$*, and*
2. *the* α*-pessimistic mapping* $\tilde{Y}_\alpha : \Omega \to \mathbb{R}$ *is a real-valued random variable on* $(\Omega, \mathcal{B}, \mathcal{P})$ *for every* $\alpha \in [0, 1]$.

(See Kratschmer (2006) for a comprehensive treatment of measurability conditions for fuzzy random variables).

This definition extends the classical concept of a random variable to the fuzzy setting. Instead of taking real values, a fuzzy random variable takes fuzzy values. The key insight is that if we can decompose the fuzzy-valued mapping into a family of real-valued random variables (via the α-pessimistic representation), then many classical probabilistic concepts can be extended to the fuzzy setting. The measurability condition ensures that the fuzzy random variable is well-behaved from a probabilistic perspective and that probabilities of events involving the fuzzy random variable can be meaningfully defined.

Example 2.5. *Let* $\Omega = \{\omega_1, \omega_2\}$ *with* $P(\omega_1) = P(\omega_2) = 0.5$. *Define* $\tilde{Y}(\omega_1) = (2; 0.5, 0.5)_{LR}$ *and* $\tilde{Y}(\omega_2) = (4; 0.3, 0.7)_{LR}$. *Then* $\tilde{Y}$ *is a fuzzy random variable because for each* α, $\tilde{Y}_\alpha$ *takes real values and is measurable on* Ω*:*

- *For* $\alpha = 0.5$*:* $\tilde{Y}_{0.5}(\omega_1) = 2$ *and* $\tilde{Y}_{0.5}(\omega_2) = 4$
- *For* $\alpha = 0.25$*:* $\tilde{Y}_{0.25}(\omega_1) = 2 - 0.5(1 - 2(0.25)) = 2 - 0.25 = 1.75$ *and* $\tilde{Y}_{0.25}(\omega_2) = 4 - 0.3(1 - 2(0.25)) = 4 - 0.15 = 3.85$

This example illustrates how a fuzzy random variable assigns fuzzy numbers to outcomes of a random experiment, and the α*-pessimistic representation extracts real-valued random variables at each level of pessimism* α.

Remark 2.2. *The case* $\alpha = 0$ *in Definition 2.4 requires special attention. The* α*-cut at* $\alpha = 0$ *corresponds to the support of the fuzzy number, i.e.,* $\tilde{Y}[0] = \{y : \eta_{\tilde{Y}}(y) > 0\}$. *Unlike* α*-cuts for* $\alpha > 0$, *the support may be unbounded and may not satisfy the standard measurability conditions required for a random variable. Therefore, in practice, we typically restrict attention to* $\alpha \in (0, 1]$ *for the standard* α*-cut representation. The case* $\alpha = 0$ *may be included only under appropriate regularity conditions, such as when the support is guaranteed to be compact (see Kratschmer (2006) for further details).*

The following definition extends the classical notion of independent and identically distributed (i.i.d.) random variables to the fuzzy setting. Independence and identical distribution are fundamental concepts in statistics that underpin many inferential procedures, including parameter estimation, hypothesis testing, and prediction intervals. By requiring independence and identical distribution at each α-level, we ensure that the fuzzy random variables share the same uncertainty structure and are probabilistically independent. This is particularly important for fuzzy time series analysis, where we need to assume that the error terms are i.i.d. for model identifiability and for establishing the theoretical properties of the proposed estimators.

Definition 2.6. *Let $\widetilde{X}$ and $\widetilde{Y}$ be fuzzy random variables. Then, $\widetilde{X}$ and $\widetilde{Y}$ are said to be independent and identically distributed if $\widetilde{X}_\alpha$ and $\widetilde{Y}_\alpha$ are independent and identically distributed random variables for each $\alpha \in [0,1]$.*

Example 2.7. *Let $\tilde{X}_1 = (0; 0.5, 0.5)_{LR} + \epsilon_1$ and $\tilde{X}_2 = (0; 0.5, 0.5)_{LR} + \epsilon_2$, where $\epsilon_1, \epsilon_2 \sim N(0,1)$ are independent. Then $\tilde{X}_1$ and $\tilde{X}_2$ are i.i.d. fuzzy random variables because for each α:*

- *$\tilde{X}_{1\alpha} = c_\alpha + \epsilon_1$ and $\tilde{X}_{2\alpha} = c_\alpha + \epsilon_2$, where c_α is the α-pessimistic value of the fuzzy number $(0; 0.5, 0.5)_{LR}$.*
- *$\tilde{X}_{1\alpha}$ and $\tilde{X}_{2\alpha}$ are independent because ϵ_1 and ϵ_2 are independent.*
- *$\tilde{X}_{1\alpha}$ and $\tilde{X}_{2\alpha}$ are identically distributed because both follow $N(c_\alpha, 1)$.*

This example illustrates how the definition of i.i.d. fuzzy random variables naturally extends the classical i.i.d. concept by requiring independence and identical distribution at each α-level.

Remark 2.8. *According to the relation between α-cuts and α-pessimistic concepts and Definition 2.4, we can consider the following two equivalent definitions for fuzzy random variables*

$$\tilde{X}_\alpha = \begin{cases} \tilde{X}^L_{(2\alpha)} & 0 < \alpha \leq 0.5, \\ \\ \tilde{X}^U_{(2(1-\alpha))} & 0.5 < \alpha \leq 1. \end{cases} , \tilde{X}_{[\alpha]} = [\tilde{X}_{\frac{\alpha}{2}}, \tilde{X}_{1-\frac{\alpha}{2}}] \quad \forall \alpha \in (0,1].$$

The cumulative distribution function (CDF) is fundamental to probability theory. It completely characterizes the distribution of a random variable and is used in many statistical procedures, including hypothesis testing, goodness-of-fit tests, and confidence interval construction. The following definition extends the classical CDF to the fuzzy setting by defining the α-pessimistic CDF. For a given α, the α-pessimistic CDF represents our pessimistic assessment of the probability that the fuzzy random variable is less than or equal to x. This is useful for statistical inference with fuzzy data, as it allows us to quantify the uncertainty in probabilistic assessments.

Definition 2.9. *Let $\tilde{X}$ be a fuzzy random variable. Then, the α-pessimistic of the fuzzy cumulative distribution function of $\tilde{X}$, denoted by $(\tilde{F}_{\tilde{X}}(x))_\alpha$, is defined as follows*

$$(\tilde{F}_{\tilde{X}}(x))_\alpha = \begin{cases} \inf\limits_{\beta\in I_{2\alpha}} F_{\tilde{X}_\beta}(x) & 0 \le \alpha \le 0.5, \\ & = F_{\tilde{X}_{1-\alpha}}(x). \\ \sup\limits_{\beta\in I_{2(1-\alpha)}} F_{\tilde{X}_\beta}(x) & 0.5 < \alpha \le 1. \end{cases}$$

Example 2.10. *Let $\tilde{Y} = (Y^L, Y, Y^U)_T$ be a triangular fuzzy random variable where $Y \sim N(0,1)$, $Y^L = Y - 0.5$, and $Y^U = Y + 0.5$. Then for $\alpha = 0$:*

$$(\tilde{F}_{\tilde{Y}}(t))[0] = [F_{Y^U}(t), F_{Y^L}(t)].$$

This means that the support of the fuzzy CDF at the 0-level is bounded by the CDF of the upper endpoint Y^U and the CDF of the lower endpoint Y^L. For example, at $t = 0$:

- $F_{Y^L}(0) = P(Y - 0.5 \le 0) = P(Y \le 0.5) \approx 0.6915$
- $F_{Y^U}(0) = P(Y + 0.5 \le 0) = P(Y \le -0.5) \approx 0.3085$

So $(\tilde{F}_{\tilde{Y}}(0))[0] = [0.3085, 0.6915]$, which captures the imprecision in the probabilistic assessment due to the fuzziness of the random variable.

Consider the fuzzy random variable $\tilde{Y} = (\tilde{Y}^L, Y, \tilde{Y}^U)_T$. Then, based on Definition 2.9, we have that

$$(\tilde{F}_{\tilde{Y}}(t))[0] = [F_{Y^U}(t), F_{Y^L}(t)].$$

Moreover,

$$\tilde{F}_{\tilde{Y}}(t) = (F_{Y^U}(t), F_Y(t), F_{Y^L}(t))_{LR},$$

where,

$$\begin{cases} F_{Y^U}(t) = P(Y^U \le t), \\ F_Y(t) = P(Y \le t), \\ F_{Y^L}(t) = P(Y^L \le t). \end{cases}$$

Quantiles are essential in many statistical applications, particularly in quantile regression and robust estimation. They provide a complete characterization of the distribution of a random variable and are more robust to outliers than moments such as the mean and variance. Now, we aim to define the fuzzy quantile of fuzzy random variables based on the obtained results.

Definition 2.11. *Suppose that $(\tilde{Y}^L, Y, \tilde{Y}^U)_T$ are fuzzy random variables. Then, the α-pessimistic of the fuzzy quantile for fuzzy random variables is defined as follows*

$$(\tilde{Q}_{\tilde{Y}}(\tau))_\alpha = \begin{cases} \inf\limits_{\beta \in I_{2\alpha}} F^{-1}_{\tilde{Y}_\beta}(\tau) & 0 \leq \alpha \leq 0.5, \\ \sup\limits_{\beta \in I_{2(1-\alpha)}} F^{-1}_{\tilde{Y}_\beta}(\tau) & 0.5 < \alpha \leq 1. \end{cases}$$

$$= \begin{cases} F^{-1}_{\tilde{Y}_\alpha}(\tau) & 0 \leq \alpha \leq 0.5, \\ F^{-1}_{\tilde{Y}_\alpha}(\tau) & 0.5 < \alpha \leq 1. \end{cases} = F^{-1}_{\tilde{Y}_\alpha}(\tau)$$

Noting that in Definition 2.11, the quantity $F^{-1}_{\tilde{Y}_\beta}$ is increasing in terms of β. Hence, $\tilde{Q}_{\tilde{Y}}(\tau)$ may be considered as a fuzzy number. Therefore,

$$(\tilde{Q}_{\tilde{Y}}(\tau))[0] = [F^{-1}_{Y^L}(\tau), F^{-1}_{Y^U}(\tau)],$$

where

$$\begin{cases} F^{-1}_{Y^L}(\tau) = inf\{y : F_{Y^L}(y) \geq \tau)\}, \\ F^{-1}_{Y}(\tau) = inf\{y : F_{Y}(y) \geq \tau)\}, \\ F^{-1}_{Y^U}(\tau) = inf\{y : F_{Y^U}(y) \geq \tau)\}. \end{cases}$$

Example 2.12. *Let $Y \sim N(0,1)$ and consider the triangular fuzzy random variable $\tilde{Y} = (Y - 0.5, Y, Y + 0.5)_T$. Then:*

- *For $\tau = 0.95$, the classical quantile is $Q_Y(0.95) \approx 1.645$*
- *For the fuzzy random variable, the fuzzy quantile $\tilde{Q}_{\tilde{Y}}(0.95)$ is a fuzzy number with:*
  - *Center: $Q_Y(0.95) \approx 1.645$*
  - *Left spread: 0.5 (from $Y - 0.5$)*
  - *Right spread: 0.5 (from $Y + 0.5$)*

This means the fuzzy quantile is $\tilde{Q}_{\tilde{Y}}(0.95) = (1.645; 0.5, 0.5)_T$. This captures both the quantile value and the uncertainty in its estimation due to the fuzziness of the random variable. For example, at $\alpha = 0$:

$$\begin{aligned} (\tilde{Q}_{\tilde{Y}}(0.95))[0] &= [Q_{Y^L}(0.95), Q_{Y^U}(0.95)] \\ &= [Q_Y(0.95) - 0.5, Q_Y(0.95) + 0.5] = [1.145, 2.145]. \end{aligned}$$

This interval represents the range of possible quantile values when accounting for the full uncertainty in the fuzzy random variable.

The following result establishes a fundamental property of fuzzy quantiles that will be essential for the proposed methodology. This result extends the classical quantile translation property (see Koenker and Bassett (1978), Theorem 1).

Theorem 2.13. *Let $\tilde{Y}^L$, Y and $\tilde{Y}^U$ be fuzzy random variables. Then, for an arbitrary fuzzy number $\tilde{b} = (\tilde{b}^L, b, \tilde{b}^U)_{LR}$ we have that*

$$\tilde{Q}_{Y\oplus\tilde{b}}(\tau) = Q_Y(\tau) \oplus \tilde{b}. \tag{2.1}$$

The theorem states that adding a fuzzy constant to a fuzzy random variable shifts its quantile function by that same constant. This extends the classical property $Q_{Y+c}(\tau) = Q_Y(\tau) + c$ to the fuzzy setting.

Proof: From Definition 2.11, the α-pessimistic representation of the fuzzy quantile of $Y \oplus \tilde{b}$ is given by:

$$(\tilde{Q}_{Y\oplus\tilde{b}}(\tau))_\alpha = F^{-1}_{(Y\oplus\tilde{b})_\alpha}(\tau).$$

By the definition of fuzzy addition and the α-pessimistic representation, we have:

$$(Y \oplus \tilde{b})_\alpha = Y_\alpha + \tilde{b}_\alpha.$$

Since the classical quantile function satisfies $F^{-1}_{Y+c}(\tau) = F^{-1}_Y(\tau) + c$ for any real-valued random variable Y and constant c, we obtain:

$$F^{-1}_{(Y\oplus\tilde{b})_\alpha}(\tau) = F^{-1}_{Y_\alpha}(\tau) + \tilde{b}_\alpha.$$

But $F^{-1}_{Y_\alpha}(\tau) = (\tilde{Q}_Y(\tau))_\alpha$ by Definition 2.11. Therefore:

$$F^{-1}_{(Y\oplus\tilde{b})_\alpha}(\tau) = (\tilde{Q}_Y(\tau))_\alpha + \tilde{b}_\alpha.$$

Applying the α-pessimistic representation in reverse, we conclude:

$$\tilde{Q}_{Y\oplus\tilde{b}}(\tau) = \tilde{Q}_Y(\tau) \oplus \tilde{b}.$$

This completes the proof. □

3. Fuzzy time dependent data modeling

In this section, we provide a new approach in order to model fuzzy time-dependent data based on support vector machines and quantile models.

3.1 Support Vector Machine

Vapnik (1995, 1998) first proposed the idea of a support vector machine (SVM). The SVM was initially developed for the classification of binary objects. The approach was later extended to support vector regression (SVR) for the prediction of numerical values (Vapnik, 1998). The algorithm, which searches for a hyperplane that optimally separates positive and negative training examples (SVM) or fits a regression function (SVR), projects training data into a pre-defined feature space to produce a model for qualitative or quantitative predictions. SVM has a distinguishing characteristic that sets it apart from other machine learning (ML) techniques: it searches for hyperplanes that linearly segregate positive and negative training data in feature spaces of increasing size. The data are then transferred into a higher-dimensional space where linear separation might be achievable if linear separation is not possible in the supplied feature space.

Consider the data set $\{\mathbf{x}, y_i\}$ where $\mathbf{x} = (x_1, x_2, ..., x_p)'$ and $y_i \in \{-1, +1\}$ are a feature vector and a label vector, respectively $(i = 1, ..., n)$. Consider the hyperplane formulated by

$$\mathbf{w}'.\mathbf{x} + w_0 = 0.$$

The classification problem is to obtain the hyperplane such that $\mathbf{w}'.\mathbf{x} + w_0 > 1$ for positive values and $\mathbf{w}'.\mathbf{x} + w_0 < -1$ for negative values. In order to find the hyperplane with the largest margin, the following minimization problem should be solved

$$\begin{aligned} &\min_{\mathbf{w},\mathbf{w_0}} \quad \frac{1}{2}\mathbf{w}'.\mathbf{w} \\ &s.t. \quad y_i(\mathbf{w}'.\mathbf{x_i} + \mathbf{w_0}) \geq \mathbf{1} \qquad \mathbf{i} = \mathbf{1}, ..., \mathbf{n} \end{aligned}$$

To solve the above optimization problem, its dual should be obtained. The dualization method is obtained by using Lagrange coefficients and based on the objective function and its corresponding constraints as follows

$$\mathcal{L}(\mathbf{w}, w_0, \boldsymbol{\alpha}) \quad = \tfrac{1}{2}\mathbf{w}'.\mathbf{w} - \textstyle\sum_{i=1}^{n} \alpha_i[y_i(\mathbf{w}'.\mathbf{x}_i + w_0) - 1].$$

In order to obtain the optimal points, we derive the function $\mathcal{L}$ with respect to the main variables $\mathbf{w}$ and w_0 and set them equal to zero. In this case, we have

$$\begin{cases} \frac{\partial \mathcal{L}}{\partial \mathbf{w}} = 0 \Rightarrow \mathbf{w} = \sum_{i=1}^{n} \alpha_i y_i \mathbf{x}_i, \\ \frac{\partial \mathcal{L}}{\partial w_0} = 0 \Rightarrow \sum_{i=1}^{n} \alpha_i y_i = 0. \end{cases}$$

By inserting $\mathbf{w}$ into $\mathcal{L}$, we get

$$\begin{aligned} L_D &= \frac{1}{2}\Big(\sum_{i=1}^{n}\alpha_i y_i \mathbf{x}_i'\Big).\Big(\sum_{j=1}^{n}\alpha_j y_j \mathbf{x}_j\Big) - \sum_{i=1}^{n}\alpha_i\Big[y_i\Big(\big(\sum_{j=1}^{n}\alpha_j y_j \mathbf{x}_j\big).\mathbf{x}_i' + w_0\Big) - 1\Big] \\ &= \sum_{i=1}^{n}\alpha_i - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i\alpha_j y_i y_j \mathbf{x}_i'.\mathbf{x}_j, \end{aligned}$$

As a result, the optimization problem is summarized as follows

$$\begin{aligned} &\max_{\alpha} \quad L_D \\ &s.t. \quad \sum_{i=1}^{n}\alpha_i y_i = 0, \quad \alpha_i \geq 0 \quad i = 1, ..., n \end{aligned}$$

This optimization problem can be solved by several methods mentioned in Byun and Lee (2002).

3.2 Proposed Method

In this subsection, we introduce a nonparametric support vector machine (SVM)–based framework for modeling fuzzy time-dependent data. Suppose that $\tilde{Z}_t,\ t = 1, 2, \ldots, T$, denote fuzzy observations recorded over time. Each observation is assumed to be represented by an LR-type fuzzy number consisting of a center and two spreads. The main idea of the proposed approach is to model the temporal structure of fuzzy observations using a kernel-based regression function while estimating the parameters through an SVM optimization framework. This formulation allows the model to capture nonlinear temporal patterns and maintain the fuzzy structure of the data.

The regression model for the fuzzy time-dependent observations can be expressed as

$$\tilde{Z}_t = \bigoplus_{j=1}^{T}\tilde{w}_j K_h(t_j, t) + \tilde{w}_0 + \varepsilon_t, \qquad t = 1, 2, \ldots, T, \tag{3.2}$$

where

$$\tilde{w}_j = (w_j; l_{w_j}, r_{w_j})_{LR}, \qquad j = 1, \ldots, T,$$

$$\tilde{Z}_t = (Z_t; l_{Z_t}, r_{Z_t})_{LR}, \qquad t = 1, \ldots, T,$$

and the kernel function is defined as

$$K_h(t_j, t) = K\left(\frac{\|t_j - t\|}{h}\right), \qquad j = 1, \ldots, T,$$

where h is the smoothing (bandwidth) parameter controlling the local influence of observations.

The optimal value of h can be determined using the generalized cross-validation (GCV) approach (Craven and Wahba, 1979; Golub et al., 1979). GCV is a popular data-driven method for bandwidth selection in nonparametric regression that seeks to minimize the leave-one-out prediction error without actually performing leave-one-out computations.

The error term is assumed to follow a normal distribution,

$$\varepsilon_t \sim N(0, \sigma^2), \qquad \tilde{\varepsilon}_t = \varepsilon_t \oplus \tilde{w}_0.$$

Using Equation (2.1), the quantile representation of the fuzzy response $\tilde{Z}_t$ can be written as

$$Q_{\tilde{Z}_t}(\tau) = \left(\bigoplus_{j=1}^{T} \tilde{w}_j K(t_j, t) + \tilde{w}_0\right) + Q_{\epsilon_t}(\tau), \qquad t = 1, \ldots, T.$$

Therefore, the LR representation of the quantile function can be expressed as

$$Q_{\tilde{Z}_t}(\tau) = \left(\left(\sum_{j=1}^{T} w_j K(t_j, t)\right) + w_0 + Q_{\epsilon_t}(\tau),\right.$$
$$\left(\sum_{j=1}^{T} l_{w_j} K(t_j, t)\right) + l_{w_0},$$
$$\left.\left(\sum_{j=1}^{T} r_{w_j} K(t_j, t)\right) + r_{w_0}\right)_{LR}.$$

Note that w_0, l_{w_0}, and r_{w_0} are the intercept terms for the center, left spread, and right spread, respectively. These intercept terms are not included in the summations over j.

Since the proposed model involves nonlinear kernel functions and potentially noisy observations, a support vector machine framework is adopted for parameter estimation. Before presenting the estimation procedure, we define the check function (also known as the pinball loss function) from quantile regression:

$$\rho_\tau(u) = u(\tau - I(u < 0)), \quad \tau \in (0, 1),$$

where $I(\cdot)$ is the indicator function. Equivalently, this can be written in piecewise form as:

$$\rho_\tau(u) = \begin{cases} \tau u, & u \geq 0, \\ (\tau - 1)u, & u < 0. \end{cases}$$

This function penalizes positive and negative residuals asymmetrically, with the asymmetry controlled by the quantile level τ. The check function is central to quantile regression and allows us to estimate conditional quantiles of the response variable (Koenker and Bassett, 1978).

The parameters $c_1, c_2, c_3 > 0$ appearing in the objective functions below are regularization parameters (also known as cost or penalty parameters). These parameters control the trade-off between model complexity (measured by the squared norm of the coefficient vector) and the fit to the data (measured by the sum of the check function losses). Larger values of c_i place more emphasis on fitting the data, potentially leading to overfitting, while smaller values emphasize model simplicity, potentially leading to underfitting. The optimal values of these parameters are typically selected using cross-validation or other data-driven model selection criteria (see, e.g., Hastie et al. (2009)).

Step 1: Estimation of the center parameters. First, the center coefficients w_j are estimated by minimizing the following objective function:

$$\hat{\mathbf{w}} = \arg\inf_{\mathbf{w}} \left[\frac{\|\mathbf{w}\|^2}{2} + \frac{c_1}{2} \sum_{t=1}^{T} \rho_\tau \left(Z_t - \left(\sum_{j=1}^{T} w_j K(t_j, t) + w_0 \right) - Q_{\epsilon_t}(\tau) \right) \right].$$

Step 2: Estimation of the left spread parameters.

Using the estimated values $\hat{w}_j$ obtained in Step 1, the parameters l_{w_j} are estimated by solving

$$\hat{l}_{\mathbf{w}} = \inf_{l_{\mathbf{w}}} \left\{ \frac{\|l_{\mathbf{w}}\|^2}{2} + \frac{c_2}{2} \sum_{t=1}^{T} \rho_\tau \Big(Z_t - l_{Z_t} - (\Sigma_{j=1}^{T} \hat{w}_j K(t_j, t) + \hat{w}_0) \right.$$
$$\left. - Q_{\varepsilon_t}(\tau) - (\Sigma_{j=1}^{T} l_{w_j} K(t_j, t)) + l_{w_0} \Big) \right\}.$$

Step 3: Estimation of the right spread parameters.

Finally, using the estimates obtained in the previous steps, the right spread parameters r_{w_j} are obtained by minimizing

$$\hat{r}_{\mathbf{w}} = \inf_{r_{\mathbf{w}}} \left\{ \frac{\|r_{\mathbf{w}}\|^2}{2} + \frac{c_3}{2} \sum_{t=1}^{T} \rho_\tau \Big(Z_t - r_{Z_t} - (\Sigma_{j=1}^{T} \hat{w}_j K(t_j, t) + \hat{w}_0) \right.$$
$$\left. - Q_{\varepsilon_t}(\tau) - (\Sigma_{j=1}^{T} r_{w_j} K(t_j, t)) + r_{w_0} \Big) \right\}.$$

Remark 3.1. *The proposed model has several important advantages:*

- *The SVM-based objective function provides a regularized optimization framework that controls model complexity and improves prediction accuracy.*
- *The kernel structure allows the model to capture nonlinear temporal relationships in fuzzy time-dependent data.*
- *The estimation procedure separately models the center and spreads of fuzzy observations, preserving the fuzzy structure of the data.*
- *The smoothing parameter h controls the local influence of observations and enhances the robustness of the model in the presence of outliers or abrupt changes in the data.*
- *The bandwidth parameter h can be selected using generalized cross-validation, providing a data-driven mechanism for model tuning.*

4. Experimental Results

In this section, the performance of the proposed SVM-based fuzzy time-dependent regression model is assessed using two numerical examples. The first example is based on a simulated data set designed to illustrate the behavior of the method under controlled conditions, including the presence of artificially generated outliers. This allows us to evaluate the robustness and accuracy of the proposed estimator when facing perturbations in both the center and spreads of fuzzy observations.

The second example uses a real data set involving fuzzy time-dependent measurements collected from a software development project. This real-world example demonstrates the applicability of the proposed model to practical situations in which the underlying data exhibit uncertainty, imprecision, and nonlinear temporal dynamics.

To compare the performance of the proposed approach with existing methods and to quantify the goodness-of-fit between the observed fuzzy responses and the estimated fuzzy values, two evaluation criteria are employed: a similarity-based

measure for fuzzy numbers and the classical root mean squared error (RMSE) computed on the centers.

To evaluate the fitting performance of the fuzzy regression model, we employ the similarity measure introduced by Kim and Bishu (1998). This index quantifies the degree of overlap between each observed fuzzy response and its corresponding estimated fuzzy value. It is defined as

$$MSM = \frac{1}{n}\sum_{j=1}^{n} S(\widehat{\tilde{\eta}}_j, \tilde{\eta}_j),$$

where

$$S(\widehat{\tilde{\eta}}_j, \tilde{\eta}_j) = \frac{Card(\widehat{\tilde{\eta}}_j \cap \tilde{\eta}_j)}{Card(\widehat{\tilde{\eta}}_j \cup \tilde{\eta}_j)},$$

and $Card(\tilde{\eta})$ denotes the cardinality of the fuzzy number $\tilde{\eta}$. The operators $\cup$ and $\cap$ represent the union and intersection of fuzzy sets, respectively. Larger values of MSM indicate a better match between fitted and observed fuzzy responses.

For additional comparison of the central tendency of fuzzy observations,
we also report the root mean squared error (RMSE) computed on the centers:

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{j=1}^{n}\left(Z_j - \hat{Z}_j\right)^2}.$$

For additional comparison of the accuracy of the estimated spreads,
we also report the Average Spread Error (ASE):

$$\text{ASE} = \frac{1}{2n}\sum_{t=1}^{n}\left[(l_{Z_t} - \hat{l}_{Z_t})^2 + (r_{Z_t} - \hat{r}_{Z_t})^2\right],$$

where l_{Z_t} and r_{Z_t} are the observed left and right spreads of the t-th fuzzy observation, and $\hat{l}_{Z_t}$ and $\hat{r}_{Z_t}$ are their estimated counterparts. The ASE measures the average squared error of the spread estimates, with smaller values indicating a better fit in terms of the spreads.

Example 4.1 (Simulated dataset)**.** *In this example, a simulated fuzzy time series dataset is generated in order to assess the performance of the proposed SVM-based modeling approach under controlled conditions. A sample of 150 fuzzy observations is produced according to the following mechanism:*

$$Z_t = te^{0.2t} + 0.5\log(t)\sin(0.1t) + 2\log(2t)\cos(0.2t) + \varepsilon_t, \quad \varepsilon_t \sim N(0, 0.7^2),$$
$$l_{Z_t} = 4 - 0.3\log(2t)\cos(0.3t) + \epsilon'_t, \quad \epsilon'_t \sim N(0, 0.05^2),$$

where $t = 1, 2, \ldots, 150$ is the time index. The membership function is defined as $L(x) = \max(0, 1 - x)$, which determines the shape of the triangular fuzzy numbers. The above data generation mechanism produces fuzzy observations with a clear nonlinear trend in the centers together with varying spreads. The resulting simulated fuzzy dataset is illustrated in Figure 1.

To further evaluate the robustness of the proposed method, the simulation experiment was repeated 10 times. In each run, artificial outliers were introduced in two different ways. Specifically, 10 units were added to the centers of 10 randomly selected fuzzy observations, while an additional 10 units were added to the widths of another 10 fuzzy observations. This procedure creates abnormal observations in both the location and the dispersion of the fuzzy data, allowing us to assess the robustness of the modeling procedure in the presence of contaminated observations.

The SVM-based fuzzy time series model was then fitted to the centers, left spreads, and right spreads of the fuzzy data using a Gaussian kernel. The tuning parameters used in the three-stage estimation procedure are summarized in Table 2. These parameters were selected to achieve a suitable balance between smoothness and goodness of fit.

The quantile levels $\tau = 0.7$ for the center estimation and $\tau = 0.5$ for the spread estimation were selected through a systematic sensitivity analysis. We evaluated the performance of the model over a grid of τ values, specifically $\tau \in \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9\}$, using 5-fold cross-validation on the training data. For each value of τ, we computed the cross-validated RMSE for the center and the cross-validated ASE for the spread estimates. The results of this sensitivity analysis indicated that $\tau = 0.7$ yielded the lowest cross-validated RMSE for the center estimates, while $\tau = 0.5$ yielded the lowest cross-validated ASE for the spread estimates. The choice of $\tau = 0.7$ for the center provides robustness against outliers and skewness in the data, as quantile regression at higher quantile levels is less sensitive to extreme observations than the mean. The choice of $\tau = 0.5$ for the spreads corresponds to the median (least absolute deviation) estimator, which is a natural and robust choice for estimating the central tendency of the spread parameters.

Table 2: SVM tuning parameters used in the simulation study

Stage	c	τ	h
First	2	0.7	0.7
Second	2	0.5	0.7
Third	2	0.5	0.7

For a fair comparison, the competing methods of Zarei et al. (2020) and Hesamian

et al. (2022) were implemented following the recommendations provided in their original papers. The implementation details are as follows:

- ***Zarei et al. (2020):*** *This semi-parametric autoregressive fuzzy time series model was implemented with the autoregressive lag order selected via the Akaike Information Criterion (AIC) from a maximum lag of* 5*. The non-parametric component was estimated using a Gaussian kernel with bandwidth selected by leave-one-out cross-validation (LOOCV). The model was fitted to the centers, left spreads, and right spreads separately, as described in the original paper.*
- ***Hesamian et al. (2022):*** *This fuzzy non-parametric time series model was implemented using a triangular kernel with bandwidth selected by cross-validation. The optimal bandwidth was chosen by minimizing the mean squared error (MSE) over a grid of candidate bandwidth values. The model was applied to the fuzzy observations directly, preserving the fuzzy structure of the data as in the original implementation.*

To evaluate the predictive performance of the model, the mean similarity measure (MSM) and the root mean square error (RMSE) between the observed and fitted fuzzy numbers were calculated. The average values (± standard deviations) of these criteria over the 10 simulation runs (including the introduced outliers) are reported in Table 3. The relatively small standard deviations indicate that the performance estimates are stable across simulation runs, and the proposed method consistently outperforms the competing approaches. To further assess statistical significance, one could construct confidence intervals for the differences in performance using the standard errors computed from the standard deviations across runs.

Table 3: Average performance indices (± standard deviations) over 10 simulation runs for Example 4.1 (data with outliers)

Method	MSM	RMSE	ASE
Proposed SVM-based approach	0.701 ± 0.052	0.876 ± 0.098	0.234 ± 0.031
Linear SVM model	0.534 ± 0.048	1.876 ± 0.142	0.512 ± 0.056
Zarei et al. (Zarei et al., 2020)	0.627 ± 0.061	1.124 ± 0.115	0.412 ± 0.045
Hesamian et al. (Hesamian et al., 2022)	0.603 ± 0.058	1.168 ± 0.108	0.385 ± 0.042

The results presented in Table 3 clearly demonstrate the advantage of the proposed SVM-based approach. In terms of similarity, the MSM value obtained by the proposed method (0.701) is substantially higher than those obtained by Zarei et al. (2020) (0.627) and Hesamian et al. (2022) (0.603), representing improvements of approximately 11.8% *and* 16.3%*, respectively.*

A similar pattern is observed for the RMSE criterion. The proposed method achieves an RMSE value of 0.876*, which is considerably smaller than the values reported for Zarei et al. (2020) (*1.124*) and Hesamian et al. (2022) (*1.168*). This corresponds to reductions of approximately* 22.1% *and* 25.0%*, respectively.*

To further assess the ability of the proposed method to capture nonlinear patterns, we compared its performance with a linear SVM model (i.e., using a linear kernel instead of the Gaussian kernel) applied to the same simulated data. The linear SVM model achieved an MSM of 0.534 ± 0.048 *and an RMSE of* 1.876 ± 0.142*, which are substantially worse than the proposed method's performance (*0.701 ± 0.052 *and* 0.876 ± 0.098*, respectively). This comparison demonstrates that the nonlinear kernel is essential for capturing the complex temporal patterns in the data, as the true data-generating mechanism contains nonlinear components (exponential and trigonometric functions).*

In addition to the center accuracy, the proposed method also achieves the smallest ASE (0.234 ± 0.031*), compared to the linear SVM model (*0.512 ± 0.056*) and the competing methods (*0.412 ± 0.045 *and* 0.385 ± 0.042*). This indicates that the three-stage estimation procedure effectively captures both the center and the spreads of the fuzzy observations.*

These improvements indicate that the proposed modeling approach provides a more accurate reconstruction of the fuzzy observations even in the presence of artificially introduced outliers. Therefore, the results of this simulation study confirm that the proposed SVM-based fuzzy time series model is capable of effectively capturing both the central trend and the uncertainty structure of fuzzy data while maintaining robustness against abnormal observations.

Example 4.2. *In this example, we analyze a real fuzzy time-dependent data set arising from the development of a complex software system for running a virtual shopping center. This application is particularly well-suited for the proposed methodology because software quality assessments are inherently imprecise: different testers may have varying opinions, and the evaluation criteria themselves (performance, reliability, usability, compatibility) are subjective and difficult to quantify exactly.*

During the entire development period (30 months), the programming activities were accompanied by a seven-member team of beta testers. This team continuously recorded users' feedback and the results of their interaction with the software. The testers were selected to represent diverse user profiles, ensuring that the aggregated evaluations captured a broad range of perspectives on software quality.

At the end of each month, after assessing criteria such as performance, reliability, usability, and compatibility, the testers provided a global evaluation score. This score was obtained by aggregating individual assessments and was reported

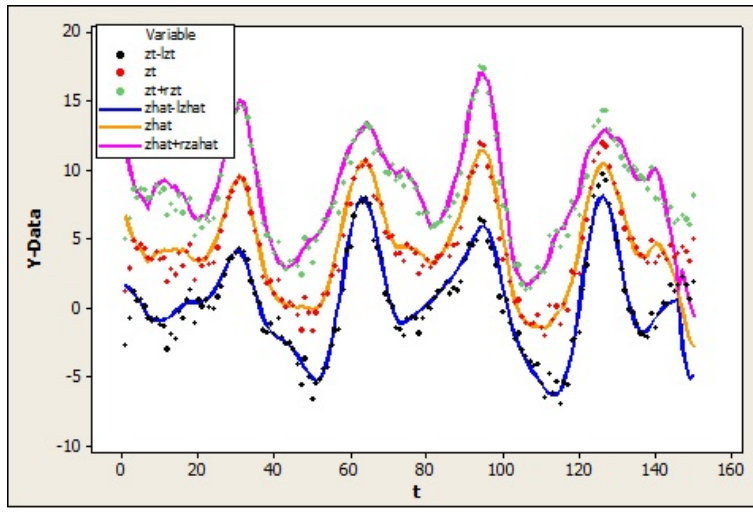

Figure 1: Scatter plot of the simulated fuzzy data in Example 4.1

as a triangular fuzzy number of the form $\tilde{y}_i = (m_i; l_i, r_i)_T$*, where* m_i *denotes the center (the most representative evaluation), and* l_i *and* r_i *are, respectively, the left and right spreads that capture the imprecision and uncertainty in the monthly evaluation. The corresponding time index* x_i *represents the month* i *(*$i = 1, \ldots, 30$*). A subset of the observed fuzzy time-dependent data is summarized in Table 4.*

Table 4: Fuzzy time-dependent data for Example 4.2.

$\tilde{y}_i$	x_i	$\tilde{y}_i$	x_i	$\tilde{y}_i$	x_i
$(29; 8, 9)_T$	21	$(26; 7, 6)_T$	11	$(6; 2, 8)_T$	1
$(30; 4, 7)_T$	22	$(25; 4, 6)_T$	12	$(5; 3, 6)_T$	2
$(35; 2, 3)_T$	23	$(5; 3, 3)_T$	13	$(8; 5, 8)_T$	3
$(44; 3, 5)_T$	24	$(8; 5, 3)_T$	14	$(10; 9, 9)_T$	4
$(45; 6, 6)_T$	25	$(9; 7, 8)_T$	15	$(13; 9, 8)_T$	5
$(43; 5, 4)_T$	26	$(10; 4, 3)_T$	16	$(19; 5, 7)_T$	6
$(47; 3, 9)_T$	27	$(12; 6, 8)_T$	17	$(20; 8, 5)_T$	7
$(48; 4, 5)_T$	28	$(13; 3, 9)_T$	18	$(23; 9, 5)_T$	8
$(50; 6, 4)_T$	29	$(20; 6, 7)_T$	19	$(25; 5, 7)_T$	9
$(49; 5, 4)_T$	30	$(28; 7, 9)_T$	20	$(27; 6, 8)_T$	10

The $\bar{X}$ *control chart of the fuzzy time-dependent observations, displayed in Figure 2, indicates the presence of two outlying points. These outliers represent atypical months in which the overall evaluation of the software deviated substantially from its general trend, possibly due to major changes in the system or unusual*

user behavior. This characteristic makes the data set suitable for assessing the robustness of the proposed modeling approach in the presence of contamination.

To investigate the nature of the temporal dynamics in this dataset, we applied the Brock-Dechert-Scheinkman (BDS) test (Brock et al., 1996) to the residuals of a linear AR model fitted to the center values. The BDS test is a widely used test for nonlinearity that examines whether the residuals are independent and identically distributed (i.i.d.). The test rejected the null hypothesis of linearity at the 5% significance level ($p = 0.023$), suggesting the presence of nonlinear temporal structure. This finding supports the use of a nonparametric approach capable of capturing nonlinear patterns.

For this data set, the fuzzy time series model is estimated separately for the center, left spread, and right spread of the triangular fuzzy observations using the SVM-based approach with a Gaussian kernel. The tuning parameters of the model in the three stages are chosen as follows: for the first stage (center) $c_1 = 2$, $\tau = 0.7$, $h = 0.7$; for the second stage (left spread) $c_2 = 2$, $\tau = 0.5$, $h = 0.7$; and for the third stage (right spread) $c_3 = 2$, $\tau = 0.5$, $h = 0.7$. The scatter plot of the fuzzy time-dependent data together with the fitted time series model for the center and spreads is illustrated in Figure 3.

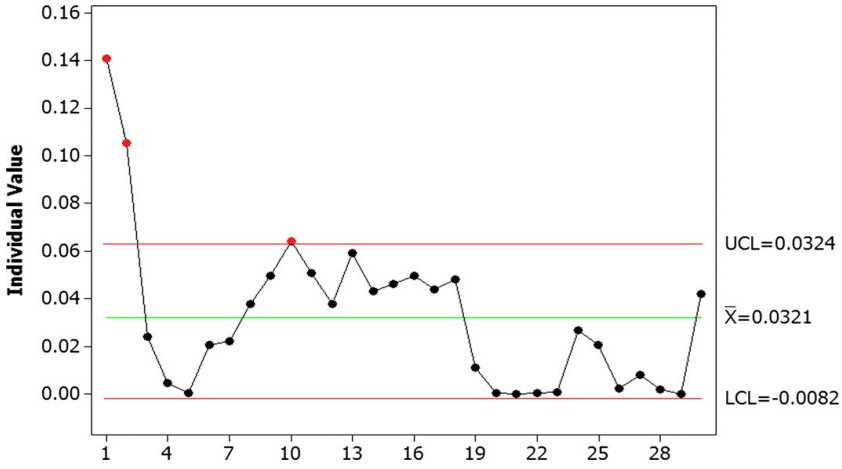

Figure 2: The $\bar{X}$ control chart for fuzzy time-dependent data in Example 4.2.

To quantitatively evaluate the performance of the proposed SVM-based fuzzy time series model, we use the mean similarity measure (MSM) and the root mean square error (RMSE) between the observed and fitted fuzzy numbers. The criteria are computed as averages over ten simulation runs based on the original data set, which includes the two outliers. Table 5 summarizes the average values of the performance indices for the proposed method and for two existing approaches in

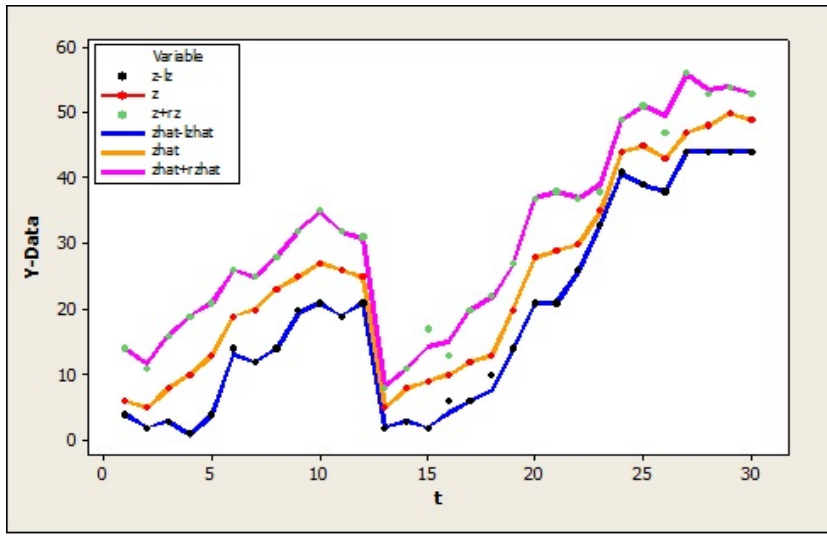

Figure 3: Scatter plot of the fuzzy time-dependent data and the fitted model in Example 4.2.

the literature, namely the methods of Zarei et al. (2020) and Hesamian et al. (2022). In addition, we include a comparison with a linear SVM model to assess the nonlinear modeling capability.

Table 5: Average performance indices ($\pm$ standard deviations) over ten simulations for Example 4.2 (data with outliers).

Method	MSM	RMSE	ASE
Proposed SVM-based approach	0.613 ± 0.048	9.530 ± 1.124	2.341 ± 0.287
Linear SVM model	0.486 ± 0.043	15.234 ± 1.654	3.987 ± 0.423
Zarei et al. (Zarei et al., 2020)	0.500 ± 0.055	16.128 ± 1.876	3.892 ± 0.412
Hesamian et al. (Hesamian et al., 2022)	0.482 ± 0.052	13.856 ± 1.543	3.445 ± 0.398

The results in Table 5 provide a clear comparison of the three modeling strategies. The proposed method attains the largest value of $\overline{MSM}$, indicating that the fitted fuzzy numbers are, on average, more similar to the observed ones than those obtained by the competing approaches. At the same time, it achieves the smallest $\overline{RMSE}$, which reflects a substantially lower prediction error.

More precisely, in terms of similarity, the proposed method improves $\overline{MSM}$ by about 22.6% compared to Zarei et al. (2020) (from 0.500 to 0.613) and by about 27.2% compared to Hesamian et al. (2022) (from 0.482 to 0.613). In terms of accuracy, the reduction in $\overline{RMSE}$ is also considerable: relative to Zarei et al. (2020), the error decreases by roughly 40.9% (from 16.128 to 9.530), and relative

to Hesamian et al. (2022) by about 31.2% *(from* 13.856 *to* 9.530*).*

To further validate the nonlinear modeling capability of the proposed approach on the real dataset, we again compared it with a linear SVM model. The linear SVM model yielded an MSM of 0.486 ± 0.043 *and an RMSE of* 15.234 ± 1.654*, compared to the proposed method's MSM of* 0.613 ± 0.048 *and RMSE of* 9.530 ± 1.124*. The superior performance of the nonlinear SVM model, combined with the BDS test result (*$p = 0.023$*) indicating the presence of nonlinear structure, provides strong evidence that the proposed method is capable of capturing nonlinear temporal patterns in real-world fuzzy time series data.*

Similarly, the proposed method yields the lowest ASE (2.341 ± 0.287*), outperforming the linear SVM model (*3.987 ± 0.423*) and the existing methods (*3.892 ± 0.412 *and* 3.445 ± 0.398*). This demonstrates that the proposed approach not only provides accurate center estimates but also reliably estimates the uncertainty (spreads) of the fuzzy responses.*

These improvements are obtained despite the presence of two outlier observations in the data set, which typically deteriorate the performance of classical fuzzy time series models. Therefore, the results of Example 4.2 indicate that the proposed SVM-based approach not only yields a closer fit to the fuzzy evaluations of the software over time but also exhibits a higher degree of robustness against atypical observations compared with existing methods. This suggests that, for real-world software quality monitoring problems involving imprecise and noisy assessments, the proposed methodology can provide more reliable and informative forecasts.

5. Summary and conclusions

The main objective of this paper was to introduce a novel nonparametric approach for modeling fuzzy time-dependent data based on support vector machines (SVMs). The proposed framework models fuzzy observations without relying on restrictive parametric assumptions and, as demonstrated by the simulation and real-data experiments, effectively captures nonlinear temporal patterns inherent in fuzzy time series. The comparison with a linear SVM model showed that the nonlinear kernel significantly improves predictive performance, while the BDS test confirmed the presence of nonlinear structure in the real dataset. To assess the performance of the proposed prediction method, a comprehensive set of evaluation criteria was employed: the mean similarity measure (MSM) for overall fuzzy similarity, the root mean squared error (RMSE) for center accuracy, and the average spread error (ASE) for spread accuracy. Both a comprehensive simulation study and a real-life application were conducted to examine the effectiveness and advantages of the model. In the simulation setting, fuzzy data were generated

under controlled scenarios that also included intentionally contaminated observations. This allowed us to examine the behavior of the proposed model in the presence of outliers. The real application further demonstrated the practicality of the approach in analyzing fuzzy assessments collected over time.

The results demonstrated that the SVM-based fuzzy modeling approach consistently provided a better fit to both simulated and real datasets compared with existing methods. This improvement was observed across all three evaluation metrics: MSM, RMSE, and ASE. For instance, in the real data example, the proposed method achieved improvements of approximately 22.6% in MSM, 40.9% in RMSE, and 39.8% in ASE compared to the best-performing competitor. Moreover, the proposed model maintained stable performance even in the presence of outlying or noisy observations, highlighting its robustness to abnormal data.

A limitation of the current study is that the ASE metric, while providing a useful measure of spread accuracy, is based on squared errors and may be sensitive to outliers. Future work could explore more robust spread-specific measures, such as the mean absolute spread error or interval-based criteria. Additionally, the current implementation focuses on triangular fuzzy numbers; extending the approach to other fuzzy number shapes would be a valuable direction for future research. Overall, the findings indicate that the proposed methodology offers a reliable, accurate, and robust alternative for the analysis of fuzzy time-dependent data in real-world applications.

Acknowledgment

We sincerely thank the reviewers for their valuable time, thorough reading, and constructive comments. Their insightful suggestions have significantly improved the quality and clarity of our manuscript.

References

Akbari, M. G. and Rezaei, A. H. (2010). Bootstrap testing fuzzy hypotheses and observations on fuzzy statistic. *Expert Systems with Applications*, **37**(8): 5782-5787.

Akbari, M. G. and Hesamian, G. (2018). A Semi-Parametric Model for Time Series based on Fuzzy data. *IEEE Transactions on Fuzzy Systems*, **26**(6): 3592-3602.

Asadolahi, M., Akbari, M. G., Hesamian, G. and Arefi, M. (2021). A robust support vector regression with exact predictors and fuzzy responses. *International Journal of Approximate Reasoning*, **132**: 812-829.

Box, G. E. P. and Jenkins, G. M. (1976). *Time Series Analysis: Forecasting and Control*. Holden-Day, San Francisco.

Brock, W. A., Dechert, W. D., Scheinkman, J. A. and LeBaron, B. (1996). A test for independence based on the correlation dimension. *Econometric Reviews*, **15**(3): 197-235.

Brockwell, P. J. and Davis, R. A. (2009). *Time Series: Theory and Methods*. Springer-Verlag, New York.

Byun, H. and Lee, S. W. (2002). Applications of support vector machines for pattern recognition: a survey. In: Lee S-W, Verri A (eds) *Pattern Recognition with Support Vector Machines*. Springer, Berlin, Heidelberg, pp 213-236.

Craven, P. and Wahba, G. (1979). Smoothing noisy data with spline functions: Estimating the correct degree of smoothing by the method of generalized cross-validation. *Numerische Mathematik*, **31**(4): 377-403.

Efromovich, S. (1999). *Non-parametric Curve Estimation: Methods, Theory and Applications*. Springer, New York.

Friedman, J. H. and Stuetzle, W. (1981). Projection pursuit regression. *Journal of the American Statistical Association*, **76**: 817-823.

Golub, G. H., Heath, M. and Wahba, G. (1979). Generalized cross-validation as a method for choosing a good ridge parameter. *Technometrics*, **21**(2): 215-223.

Hastie, T., Tibshirani, R. and Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. 2nd edn. Springer, New York.

Hesamian, G., Torkian, F. and Yarmohammadi, M. (2022). A fuzzy non-parametric time series model based on fuzzy data. *Iranian Journal of Fuzzy Systems*, **19**: 61-72.

Kim, B. and Bishu, R. R. (1998). Evaluation of fuzzy regression models by comparing membership functions. *Fuzzy Sets and Systems*, **100**: 343-352.

Koenker, R. and Bassett, G. (1978). Regression quantiles. *Econometrica*, **46**(1): 33-50.

Kratschmer, V. (2006). A unified approach to fuzzy random variables. *Fuzzy Sets and Systems*, **157**(13): 1818-1840.

Ozer, S., Chen, C. and Cirpan, H. A. (2011). A set of new Chebyshev kernel functions for support vector machine pattern classification. *Pattern Recognition*, **44**: 1435-1447.

Palma, W. (2016). *Time Series Analysis*. Wiley, New York.

Sapankevych, N. I. and Sankar, R. (2009). Time series prediction using support vector machines: a survey. *IEEE Computational Intelligence Magazine*, **4**(2): 24-38.

Shumway, R. H. and Stoffer, D. S. (2017). *Time Series Analysis and Its Applications*. Springer, New York.

Simonoff, J. (1996). *Smoothing Methods in Statistics*. Springer, New York.

Song, Q. and Chissom, B. S. (1993). Fuzzy time series and its models. *Fuzzy Sets and Systems*, **54**(3): 269-277.

Taheri, S. M. (2003). Trends in fuzzy statistics. *Austrian Journal of Statistics*, **32**: 239-257.

Tong, H. (1990). *Nonlinear Time Series: A Dynamical System Approach*. Oxford University Press, Oxford.

Vapnik, V. N. (1995). *The Nature of Statistical Learning Theory*. Springer, Berlin.

Vapnik, V. N. (1998). *Statistical Learning Theory*. Wiley, New York.

Viertl, R. (2011). *Statistical Methods for Fuzzy Data*. Wiley, New York.

Vilela, L. F. S., Leme, R. C., Pinheiro, C. A. M. and Carpinteiro, O. A. S. (2019). Forecasting financial series using clustering methods and support vector regression. *Artificial Intelligence Review*, **52**: 743-773.

Wang, B., Liu, X., Chi, M. and Li, Y. (2024). Bayesian network based probabilistic weighted high-order fuzzy time series forecasting. *Expert Systems with Applications*, **237**(Part C): 121430.

Woodward, W. A., Gray, H. L. and Elliott, A. C. (2012). *Applied Time Series Analysis*. CRC Press, Boca Raton.

Yang, X. and Liu, B. (2019). Uncertain time series analysis with imprecise observations. *Fuzzy Optimization and Decision Making*, **18**: 263-278.

Zarei, R., Akbari, M. G. and Chaci, J. (2020). Modeling autoregressive fuzzy time series data based on semi-parametric methods. *Soft Computing*, **24**: 7295-7304.

Zendehdel, J., Rezaei, M., Akbari, M. G., Zarei, R. and Alizadeh Noughabi, H. (2018). Testing exponentiality for imprecise data and its application. *Soft Computing*, **22**: 3301-3312.