International Journal of Approximate Reasoning 180 (2025) 109387

Contents lists available at ScienceDirect

International Journal of Approximate Reasoning

journal homepage: www.elsevier.com/locate/ijar

Fuzzy time series analysis: Expanding the scope with fuzzy numbers

Hugo J. Bello [a], Manuel Ojeda-Hernández [b], Domingo López-Rodríguez [b], Carlos Bejines [b],*

[a] *Departamento de Matemática Aplicada, Universidad de Valladolid, Campus Duques de Soria, Soria, 42004, Spain*
[b] *Departamento de Matemática Aplicada, Universidad de Málaga, Campus de Teatinos, Malaga, 29071, Spain*

ARTICLE INFO

Keywords:
Time series
Fuzzy numbers
Fuzzy time series
Fuzzy stochastic process

ABSTRACT

This article delves into the process of fuzzifying time series, which entails converting a conventional time series into a time-indexed sequence of fuzzy numbers. The focus lies on the well-established practice of fuzzifying time series when a predefined degree of uncertainty is known, employing fuzzy numbers to quantify volatility or vagueness. To address practical challenges associated with volatility or vagueness quantification, we introduce the concept of informed time series. An algorithm is proposed to derive fuzzy time series, and findings include the examination of structural breaks within the realm of fuzzy time series. Additionally, this article underscores the significance of employing topological tools in the analysis of fuzzy time series, accentuating the role of these tools in extracting insights and unraveling intricate relationships within the data.

1. Introduction

Fuzzy time series (FTS) emerged as a prominent paradigm in time series forecasting, utilizing principles of fuzzy logic to extend classical forecasting methods like ARIMA [1], exponential smoothing [2], and neural networks [3]. FTS employs fuzzy sets to represent linguistic terms associated with time series data, enabling the capture of uncertainty and imprecision inherent in the data.

Fuzzification entails transforming numerical time series data into fuzzy sets by assigning membership functions to linguistic terms [4]. These functions quantify the degree of membership of data points to specific fuzzy sets. Similarly, fuzzification of the forecasting data represents future values using the same linguistic terms and membership functions employed in the initial fuzzification process [5,6]. A comprehensive survey of fuzzy time series forecasting models can be found in the recent work [7].

The determination of fuzzy rules involves the application of fuzzy logic to extract rules from the fuzzy sets, enabling predictions regarding future time series values. Conversely, defuzzification entails converting fuzzy forecast results into crisp numerical values. Various defuzzification methods, such as centroid [8], height [9], and bisector [10], have been proposed in the literature, with their comparative analysis detailed in relevant studies (for instance, see [11]).

Overall, fuzzy time series provides a comprehensive approach to addressing uncertainty and imprecision in time series forecasting, offering a rich theoretical foundation for analysis through the utilization of fuzzy sets and fuzzy logic principles. The term fuzzy time series also refers to time-indexed fuzzy numbers as a generalization of crisp time series. In this sense, FTS can be studied using the theory of fuzzy random variables [12,13] and fuzzy stochastic processes [14].

---

* Corresponding author.
E-mail address: cbejines@uma.es (C. Bejines).

https://doi.org/10.1016/j.ijar.2025.109387
Received 21 May 2024; Received in revised form 4 October 2024; Accepted 17 February 2025

FTS has been successfully applied in various fields. For instance,

1. Financial Forecasting: FTS has been applied to stock market prediction, portfolio optimization, and other financial forecasting tasks. It allows for handling the uncertainty and non-linearity often present in financial data [15,16].
2. Decision Support Systems: FTS has been used to develop decision support systems for various domains, including healthcare, transportation, and energy management. These systems utilize fuzzy time series models to aid decision-making processes [17].
3. Pattern Recognition: FTS has been employed for pattern recognition tasks, such as speech recognition and image classification, where temporal relationships play a crucial role [18].

In summary, FTS is a powerful method for time series forecasting that uses fuzzy logic and fuzzy sets to capture the uncertainty and imprecision of time series data. It has several steps, including fuzzification, fuzzification of the forecasting data, determination of the fuzzy rules, and defuzzification, and has been successfully applied in various fields. It has been developed in depth in recent years to solve problems where crisp time series cannot provide a good solution.

As captured by fuzzy sets, the notion of “uncertainty” within databases draws from two distinct streams of thought that have long been recognized in the mathematical literature. Firstly, data imprecision may hinder processes related to it from achieving certainty. Secondly, uncertainty can also pertain to ambiguity in interpretation, which proves particularly beneficial in scenarios requiring the description of preferences among elements or objects. Consequently, fuzzy random variables can be interpreted in various ways over time (see [19–22]). In the first approach, the “ontic” interpretation [23], the variance of the fuzzy random variable is expressed as a crisp number, simplifying subsequent processing tasks. In the second one, the “epistemic” interpretation [24], the variance is defined as a fuzzy interval, providing a gradual depiction of our partial understanding regarding the variance of an underlying classical random variable that has been imprecisely observed. These two approaches have proven to be truly useful in practice [25–28]. As a pertinent article on this subject, we highly recommend article [29], where the authors elaborate on these different definitions and discuss the use of a new variance, demonstrating its practical utility through simple examples.

As we have already pointed out, in numerous practical scenarios the existence of a numerical measure that captures the uncertainty or volatility of observations is well established. However, effectively modeling problems characterized by such behavior, such as those found in the stock market or the economy, poses significant challenges, even within the framework of fuzzy logic. Our key contribution lies in the definition of informed time series, which is intended to encompass scenarios where the precision or importance of each observation is determined in advance. This concept is a tool in applications, as we show in several examples.

We begin by providing a brief overview of foundational concepts before introducing the novel notion of informed time series. The primary objective of this concept is to capture and represent uncertainty straightforwardly, with fuzzy numbers playing a pivotal role in this endeavor. Section 3 of our article is dedicated to a comprehensive exploration of the informed time series concept, by providing examples and highlighting applications of r-cuts and we can gain a deeper understanding of their practical significance. Next, section 4 focuses on its fuzzification process, showing an algorithm divided into two steps. By leveraging fuzzy number arithmetic, we can derive several meaningful results, as showcased in both sections. In section 5, we delve into the examination of stationary time series with structural breaks and present a noteworthy finding: under certain conditions, it is established that fuzzification also maintains stationarity in the presence of structural breaks. Section 6 is devoted to the use of fuzzy logic in dynamic time warping. We provide a computational algorithm and examples with real data (an Olive Oil dataset from four different countries). Lastly, we outline a series of future research avenues, which can be pursued based on the findings obtained in this study.

2. Preliminaries

In this section, we will present the fundamental definitions that will serve as the foundation for the subsequent discourse on fuzzy numbers, fuzzy time series, and crisp time series.

2.1. Fuzzy numbers

Firstly, let us recall the concept of fuzzy numbers. A fuzzy number is a mathematical construct that extends the notion of a crisp number to incorporate uncertainty. It is characterized by a membership function, which assigns a degree of membership to each element of its domain, thereby representing the degree of uncertainty associated with its value. The membership function is typically a continuous, bounded, and convex function that ranges between 0 and 1.

It is worth noting that [30] provides a comprehensive and accessible introduction to fuzzy numbers, highlighting their importance and applications in different areas of science and engineering. For more information on fuzzy sets, [31] provides a detailed overview of the theoretical foundations and applications of fuzzy sets and discusses the relationship between fuzzy set theory and other fields such as possibility theory.

Definition 1 *([32]).* A fuzzy set $p : \mathbb{R} \longrightarrow [0,1]$ is called a **fuzzy number** if it satisfies the following conditions:

1. The function $p(x), x \in \mathbb{R}$, is at least piece-wise continuous.
2. The fuzzy set p is convex, i.e., for each $0 \leq t \leq 1$ and $x, y \in \mathbb{R}$, we have

$$p(tx + (1-t)y) \geq tp(x) + (1-t)p(y).$$

3. There is exactly one $\overline{x} \in \mathbb{R}$ with $p(\overline{x}) = 1$.

The core of a fuzzy set p is the set of elements whose degree equals 1. Then, the core of a fuzzy number is exactly one value. The support of $p : \mathbb{R} \to [0,1]$ is the set

$$\mathrm{supp}(p) = \{x \in \mathbb{R} \mid \mu(x) > 0\}.$$

Definition 2 *([33]).* A fuzzy number p is called an L-R fuzzy number if it has compact support and can be expressed through parameterized **reference functions** or shape functions L and R in the form

$$p(x) = \begin{cases} p_1(x) = L((\overline{x} - x)/\alpha) & \text{for } x < \overline{x}, \\ p_2(x) = R((x - \overline{x})/\beta) & \text{for } x \geq \overline{x}, \end{cases}$$

where $\alpha, \beta > 0$, $Core(\mathrm{p}) = \{\overline{x}\}$ and the functions L and R are conventionally taken as decreasing functions satisfying $L(0) = R(0) = 1$ and $L(1) = R(1) = 0$. Consequently, p_1 and p_2 are increasing and decreasing functions, respectively.

As an abbreviated notation, we can denote the L-R fuzzy number p as

$$p = < \overline{x}, \alpha, \beta >_{L,R} \tag{1}$$

The set of all L-R fuzzy numbers is denoted by $\mathcal{F}(\mathbb{R})$.

Remark 1 *([33]).* A fuzzy number as in Equation (1) is determined by its r-cut intervals

$$[p]^r = [p^-(r), p^+(r)] \tag{2}$$

in the following way

$$\begin{aligned} p^-(r) &= x - L^{-1}(r) \cdot \alpha \\ p^+(r) &= x + R^{-1}(r) \cdot \beta \end{aligned} \tag{3}$$

2.2. Norms in $\mathcal{F}(\mathbb{R})$

According to [34], let us define the following scalar product in $\mathcal{F}(\mathbb{R})$, which plays a significant role throughout the proofs of the article.

$$\begin{array}{cccc} \langle \cdot, \cdot \rangle : & \mathcal{F}(\mathbb{R}) \times \mathcal{F}(\mathbb{R}) & \to & \mathbb{R} \\ & (u, v) & \mapsto & \int_0^1 (u^-(r)v^-(r) + u^+(r)v^+(r))dr \end{array}$$

and consider the following distances

$$d_p(u, v) = \left(\int_0^1 h^p([u]^r, [v]^r) dr \right)^{1/p},$$

and

$$d_\infty(u, v) = \sup_{0<r\leq 1} h([u]^r, [v]^r),$$

where h is the Hausdorff metric

$$h([u]^r, [v]^r) = max\{|u^-(r) - v^-(r)|, |u^+(r) - v^+(r)|\}.$$

Properties of the spaces $(\mathcal{F}(\mathbb{R}), d_p)$ and $(\mathcal{F}(\mathbb{R}), d_\infty)$ have been studied in [35]. In particular, the space $(\mathcal{F}(\mathbb{R}), d_\infty)$ is a complete metric space [36]. on the other hand, for two fuzzy sets, a way of measuring the difference between them was introduced in [37].

The norm $\|\cdot\|_\infty$ is defined using the distance d_∞ as follows

$$\|p\|_\infty = d_\infty(0, p).$$

Notice that for $p = < x, \alpha, \beta >_{L,R}$,

$$\begin{aligned} \|p\|_\infty &= d_\infty(0, p) \\ &= \sup_{0<r\leq 1} \max\left(|p^+(r)|, |p^-(r)|\right) \\ &= \max\left(|p^+(1)|, |p^-(0)|\right) \qquad (4) \\ &= \max\left(|x - R^{-1}(0) \cdot \beta|,\ |x + L^{-1}(1) \cdot \alpha|\right). \qquad (5) \end{aligned}$$

In Equation (4) we have used that p^+ is increasing and p^- is decreasing.

2.3. Fuzzy random variables

Regarding the concept of fuzzy random variables, the first definition of this concept was presented by R. Féron in [38]. Subsequently, several authors such as Kwakernaak [12,39], Puri et al. [13,40,41], Körner [42,23] and Couso et al. [43] have proposed different formal definitions. Each of these definitions has considered a different meaning of the concept of measurability. Krätschmer [44] studied the formal relations between some of them and proposed a unified approach. In this paper, we consider the variance from the "ontic" point of view, that is, the one studied by Körner. Lubiano [45] focuses on the problem of estimating the expected value of fuzzy random variables in random samples of finite sets from this approach. However, it is worth noticing that the "epistemic" point of view described by [24] must be applied in a similar study in future research.

Definition 3 *([23]).* Given a complete probability space $(\Omega, \mathcal{A}, P)$, we consider that a **fuzzy random variable** is a Borel-measurable function $X : (\Omega, \mathcal{A}) \to (\mathcal{F}(\mathbb{R}), d_\infty)$.

If X is a fuzzy random variable, denoted by f. r. v., then

$$[X]^r = [X^-(r), X^+(r)], r \in (0,1]$$

is a random closed interval set and $X^+(r), X^-(r)$ are real valued random variables. If X is a f. r. v., then $\|\cdot\|_\infty$ is also a fuzzy random variable [36]. On the other hand, an f. r. v. X is called **integrably bounded** if $E(\|X\|_\infty) < \infty$. Notice here that the choice of notation reflects that $E(\|X\|_\infty)$ is indeed the expectancy of a (crisp) random variable, since the norm transforms X into a conventional random variable.

For an integrably bounded f. r. v., the expectation $E(X)$ is defined [13] as the unique fuzzy number which satisfies the property:

$$[E(X)]^r = E\big([X]^r\big) = [E(X^-(r)), E(X^+(r))],\ 0 < r \le 1. \tag{6}$$

If $1 \le p < \infty$, X is called a p-order f. r. v. provided $E\big(\|X\|_p^p\big) < \infty$. The space $\mathcal{L}_p(\mathcal{F}(\mathbb{R}))$ is the set of all p-order f. r. v.

Now, we provide an approach for the covariance introduced by Feng et al.

Definition 4 *([36]).* Let X and Y be in $\mathcal{L}_2(\mathcal{F}(\mathbb{R}))$. The covariance of X and Y is defined as

$$\operatorname{cov}(X,Y) = \frac{1}{2}\int_0^1 \Big(\operatorname{cov}(X^-(r), Y^-(r)) + \operatorname{cov}(X^+(r), Y^+(r))\Big)dr. \tag{7}$$

2.4. Stochastic processes and fuzzy stochastic processes

We recall the crisp stochastic process considering expectations and covariances as in Equation (6) and Equation (7).

Definition 5 *([46]).* A **stochastic process** is simply a sequence of random variables $(X_t)_t$. A stochastic process is said to be **stationary** if it satisfies

1. X_t has constant expectation $E(X_t) = \mu$.
2. X_t has finite variance.
3. The covariance depends only on the distance of the time instants, i.e.,

$$\operatorname{cov}(X_t, X_{t'}) = \operatorname{cov}(X_{t+h}, X_{t'+h}).$$

Two stochastic processes X_t, Y_t are **cross-stationary**, or **jointly-stationary**, if the covariance $\operatorname{cov}(X_{t+h}, Y_t)$ is a function only of lag h.

We introduce a novel definition of a fuzzy stationary stochastic process.

Definition 6. A **fuzzy stochastic process** is a sequence $(\mathcal{X}_t)_t$ of f.r.v. The fuzzy stochastic process $(\mathcal{X}_t)_t$ is **stationary** if it satisfies:

1. $\mathcal{X}_t$ is integrably bounded and $E(\mathcal{X}_t) = \mu$.
2. $\mathcal{X}_t$ is 2-order and $\operatorname{var}(\mathcal{X}_t) < \infty$.
3. $\operatorname{cov}(\mathcal{X}_t, \mathcal{X}_{t'}) = \operatorname{cov}(\mathcal{X}_{t+h}, \mathcal{X}_{t'+h})$, $\forall h$.

Remark 2. Considering a stationary stochastic process simplifies both analysis and modeling because the statistical properties (like mean, variance, and autocovariance) remain constant over time. This stability makes it easier to estimate model parameters and generate reliable forecasts, as the relationships between past and future values are predictable. Models like ARMA and ARIMA are specifically designed for stationary series, allowing for a straightforward application of well-established techniques such as autocorrelation analysis and spectral methods. The consistency of a stationary process also ensures that estimators converge more quickly, making the modeling process more efficient and accurate.

The notion fuzzy stationarity introduced here is the natural extension of the well known and well studied crisp case and we expect that it will allow us to translate the good properties of these processes into the fuzzy realm.

One of the aims of this paper is to show that in many cases the fuzzy versions of crisp stationary process related to practical situations also satisfy stationarity (see Theorem 3).

2.5. Time series

Next, we define crisp time series. A crisp time series is a conventional time series that consists of a sequence of crisp or precise values recorded at regular time intervals. Unlike fuzzy time series, crisp time series do not explicitly account for uncertainty or vagueness in the data.

Definition 7. A (crisp) **time series** $(x_t)_t$ is a sequence of time-indexed real numbers.

Crisp time series are considered the realization of a sequence of random variables (i.e. a stochastic process in the sense of Section 2.4). Analogously, a fuzzy time series is a sequence of fuzzy sets, where each fuzzy set represents the value of a variable at a specific time point. Fuzzy time series provide a framework for modeling and analyzing data with inherent uncertainty, allowing for a easy representation of the underlying dynamics. The uncertainty associated with each data point is captured through the membership function of the respective fuzzy set [5]. Formally,

Definition 8. A **fuzzy time series** $(f_t)_t$ is a sequence of time-indexed where each $f_t : \mathbb{R} \longrightarrow [0,1]$.

2.5.1. Time series classification and dynamic time warping

This preliminary subsection provides the background for Section 6, where applications and computational examples are presented. A **classification dataset**, is a set

$$D = \{(x_1, c_1), \dots (x_n, c_n)\},$$

where $x_1, \dots x_n$ are time series and $c_1, \dots, c_n$ are tags or classes associated to them. A **time series classification algorithm** is a method that attempts, given new time series x to predict or associate its more likely class c.

Time series distances can be used to classify time series. This is the idea behind the **k-nearest neighbors algorithm** (kNN) [47] In kNN a time series x is classified by a plurality vote of its neighbors, with the time series being assigned to the most common class among its k nearest neighbors (that is the k time series in the dataset with the smallest distances to x, k being a positive integer, typically small). If $k = 1$, then the time series is simply assigned to the class of that single nearest neighbor.

One of the most used time series distances for classification is the so-called dynamic time warping. **Dynamic time warping** (DTW) is a time series algorithm for measuring similarity between two temporal sequences, which may vary in speed. This method calculates an optimal match between two time series $x = (x_t)_{t=0}^N$, $y = (y_t)_{t=0}^M$ with the following rules:

- Every index x must be matched with *one or more* indices y, and vice versa.
- The last index x_N must be matched with the last index y_M (but it does not have to be its only match).
- The first index x_0 must be matched with the first index y_0 (but it does not have to be its only match).
- The mapping of the indices from x to y must be monotonically increasing, and vice versa, i.e. if $j > i$ are indices from x, then there must not be two indices $l > k$ in y, such that index i is matched with index l and index j is matched with index k, and vice versa.

The match that satisfies all the above restrictions and has the minimal cost is considered the optimal match, where the cost is computed as the cumulative sum of distances between pairs of matched indices. The pseudocode of DTW is presented in Algorithm 1.

Algorithm 1 Dynamic time warping distance between time series.

```
  Input: $x = (x_t)_{t=0}^N, y = (y_t)_{t=0}^M$
  Output DTW distance

$D_{ij} \leftarrow 0$

for $i = 1$ to $N$ do
    for $j = 1$ to $M$ do
        $cost \leftarrow d(x_i, y_j)$
        $D_{ij} \leftarrow cost + \min\{D_{i-1j}, D_{ij-1}, D_{i-1j-1}\}$
    end for
end for
return $D_{NM}$
```

3. Informed time series

This section is devoted to the new notion of informed time series. This novel definition attempts to capture situations in which a level of the imprecision or importance of each observation is known beforehand. This is the case of many practical applications as we will explore in the examples in the following sections.

Definition 9. An **informed time series** is a sequence of real numbers with a degree of the form $(x_t, u_t)_{t=1}^N$, where $x_t \in \mathbb{R}$ and $u_t \in U \subset \mathbb{R}^+$.

The value x_t represents a time-indexed observation, whereas u_t represents a measure of the uncertainty or irregularity of the measured value x_t. We will call u_t the **uncertainty** at time t and U the **uncertainty set**.

As we see we are only considering non-negative values for the uncertainty set. The idea behind this is to simplify the scale of imprecision to fix 0 as minimal value of imprecision.

Remark 3. Even though an informed time series $(x_t, u_t)_{t=1}^N$ reflects a notion of the uncertainty via the series u_t, one must not confuse this notion with the idea of fuzzy-valued time series. In Definition 9 we are reflecting the practical situations described in Examples 1 and 2 in which we have a crisp value of a time indexed phenomena x_t and we also have managed to obtain a notion of the uncertainty of the measurement of x_t at time t.

Our approach at this stage stays away of the fuzzy realm since both x_t and u_t are crisp. But this allows us to tackle the problem of finding a way of defining a fuzzy time series $(\mathfrak{F}(x_t, u_t))_t$ that better reflects the information intrinsically gathered in the information time series.

Example 1. In financial time series, if x_t is the closing price of a stock on a given day t, we can define u_t as the total variation of the price at time t. The series (x_t, u_t) is an informed time series.

Example 2. An irregular time series $(x_{t_k})_{t_k \in T}$ is a sequence of observations in which the spacing of observation times is not constant. If we divide the time interval T into n intervals T_l we can define

$$u_l = \frac{1}{\# T_l} \sum_{t_k \in T_l} |t_k - t_{k+1}|$$

and

$$x_l = \frac{1}{\# T_l} \sum_{t_k \in T_l} x_{t_k}.$$

The value x_l is the average value of the series in the time interval T_l and u_t is a measure of the irregularity of the observations. The series $(x_l, u_l)_{l=1}^n$ is an informed time series that regularizes the original series.

Definition 10. A mapping $\mathcal{U} : \mathbb{R}^+ \to \mathcal{F}(\mathbb{R})$ is called an *uncertainty mapping* if it satisfies the following properties

($P1$) $\text{core}(\mathcal{U}(s)) = \{0\}$ for all $s \in \mathbb{R}^+$.
($P2$) The r-cut interval extremes $[\mathcal{U}(s)]^r = [\mathcal{U}(s)^+(r), \mathcal{U}(s)^-(r)]$ are linear i.e. the real valued mappings $\mathcal{U}(\cdot)^+(r), \mathcal{U}(\cdot)^-(r)$ are linear for every r.

Remark 4. Every uncertainty mapping $\mathcal{U} : \mathbb{R}^+ \to \mathcal{F}(\mathbb{R})$ satisfies the following properties

1. $\mathcal{U}$ can be written as $\mathcal{U}(s) =< 0, \alpha(x), \beta(x) >_{L,R}$ for increasing linear mappings α, β.
2. $\mathcal{U}(s + s') = \mathcal{U}(s) + \mathcal{U}(s')$
3. If $s_1 > s_2$, with $s_1, s_2 \in \mathbb{R}^+$ then $\text{supp}(\mathcal{U}(s_2)) \subseteq \text{supp}(\mathcal{U}(s_1))$.

Proof. By ($P1$), the core of every fuzzy number $\mathcal{U}(s)$ is $\{0\}$, so using L-R notation the mapping $\mathcal{U}$ can be written as $\mathcal{U}(s) =< 0, \alpha(s), \beta(s) >_{L,R}$. By (3)

$$\mathcal{U}(s)^+(r) = -L^{-1}(r) \cdot \alpha(s). \tag{8}$$

$$\mathcal{U}(s)^-(r) = +R^{-1}(r) \cdot \beta(s)$$

Then, by ($P2$) the previous mappings are linear and therefore α and β are also linear.

The second property is a direct consequence of the first one. The last property follows from the fact that α and β are non-negative increasing linear functions. ☐

Remark 5. Given an informed time series $(x_t, u_t)_t$ it is clear that there is not a unique way to define a uncertainty mapping $\mathcal{U}$. For instance, we could define $\mathcal{U}(u_t) = 0$ for all $t \in \mathbb{N}$ and the properties are trivially satisfied.

Below, we define the fuzzy informed time series.

Definition 11. Every uncertainty mapping $\mathcal{U}$ defines a **fuzzification mapping**

$$\begin{array}{rccc} \mathfrak{F}: & \mathbb{R}\times\mathbb{R}^+ & \to & \mathcal{F}(\mathbb{R}) \\ & (x,s) & \mapsto & x+\mathcal{U}(s) = < x, \alpha(s)(x), \beta(s)(x) >_{L,R} \end{array}$$

where the addition $x+\mathcal{U}(i)$ is understood as the standard addition of fuzzy numbers, based on the minimum operator.

On the other hand, given an informed time series $X=(x_t,u_t)_t$ we can define a **fuzzification of the informed time series** as the sequence of fuzzy numbers:

$$\left(\mathfrak{F}(x_t,u_t)\right)_t = \left(x_t+\mathcal{U}(u_t)\right)_t \tag{9}$$

Property $(P1)$ ensures that $\operatorname{core}(\mathfrak{F}(x_t,u_t)) = x_t$ and property $(P2)$ implies that higher uncertainty values of an observation x_t will produce fuzzifications $\mathfrak{F}(x_t,u_t)$ with bigger supports.

Remark 6. Note that an informed time series represents a specific instance within a fuzzy time series, where fuzzy numbers are considered. We particularly focus on informed time series due to their distinct properties and behaviors.

Proposition 1. *For any uncertainty mapping $\mathcal{U}$, its fuzzification mapping $\mathfrak{F}$ has the following arithmetic properties*

1. $\mathfrak{F}(x,u)=\mathfrak{F}(x,0)+\mathfrak{F}(0,u)$.
2. $\mathfrak{F}(kx,ku)=k\mathfrak{F}(x,u)$ *for every* $k\geq 0$.
3. $\mathfrak{F}(x+y,u+u')=\mathfrak{F}(x,u)+\mathfrak{F}(y,u')$.

Proof. The first property follows trivially from the definition. The other two follow from the fact that by 4 the mappings α and β are linear. $\square$

Example 3. Any informed time series $X=(x_t,u_t)_{t=1}^N \subset \mathbb{R}\times I$ can be trivially fuzzified taking the trivial uncertainty mapping $\mathcal{U}(u_t) = \langle 0,0,0\rangle_{LR}$ which associates each uncertainty with the zero crisp number. This results in a fuzzification that associates each observation x_t with its crisp fuzzy number $\langle x_t,0,0\rangle_{LR}$.

In Section 4, we will explore several possibilities for fuzzification algorithms under the previous definition. We will assume now that the informed time series (x_t,u_t) are realizations of a sequence of random variables (X_t,U_t), in other words, we will turn our attention towards stochastic processes.

In view of the previous algorithm, a natural question is to ask what properties of (X_t,U_t) are translated into the fuzzification $(\mathfrak{F}(X_t,U_t))_t$. The following results are aimed at answering the previous question.

We will start with the following lemma which is a generalization of [36, Theorem 3.1] for L-R fuzzy numbers.

Lemma 1. *Using notation as in Equation (2), the mappings*

$$\begin{array}{rccc} \Phi_r^+: & (\mathcal{F}(\mathbb{R}),\|\cdot\|_q) & \to & \mathbb{R} \\ & p & \mapsto & p^+(r) \\ \Phi_r^-: & (\mathcal{F}(\mathbb{R}),\|\cdot\|_q) & \to & \mathbb{R} \\ & p & \mapsto & p^-(r) \end{array}$$

Are linear and continuous for every $q\leq\infty$.

Proof. If $p=< x,\alpha,\beta >_{L,R}$, we find that $\Phi_r^+(p) = x+R^{-1}(r)\cdot\beta$ and $\Phi_r^- = x-L^{-1}(r)\cdot\alpha$ so the linearity follows trivially from the fuzzy arithmetic.

In view that the mappings are linear transformations of normed spaces, to prove the continuity it suffices to show that zero-convergent sequences transformed into zero-convergent sequences. Since $\|p\|_\infty > \|p\|_p\ \forall p$ so it suffices to prove the continuity for the norm $\|\cdot\|_\infty$. To that end consider a sequence $(p_n)_n = (< x_n,\alpha_n,\beta_n >_{L,R})_n$ and assume that it converges to 0 under the norm $\|\cdot\|_\infty$. By Equation (5) we deduce that

$$\| < x_n,\alpha_n,\beta_n >_{L,R}\|_\infty = \sup_{r\leq 1}\max\left(|x_n - R^{-1}(r)\cdot\beta_n|,\ |x_n+L^{-1}(r)\cdot\alpha_n|\right)\to 0$$

which implies that for every $r\leq 1$

$$|x_n - R^{-1}(r)\cdot\beta_n| \to 0$$
$$|x_n + L^{-1}(r)\cdot\alpha_n| \to 0$$

Since $\alpha_n \geq 0$, $\beta_n \geq 0$ we find that the sequences $(x_n)_n$, $(\beta_n)_n$ and $(\alpha_n)_n$ must also converge to 0. Therefore we have concluded that the sequences $(\Phi_r^+(p_n))_n = (x_n + R^{-1}(r) \cdot \beta_n)$ and $(\Phi_r^-(p_n))_n = (x_n - L^{-1}(r) \cdot \alpha_n)$ also converge to 0. $\square$

Theorem 1. *Given an informed time series $(x_t, u_t)_t$, for every uncertainty mapping $\mathcal{U}$, the induced fuzzification $\mathfrak{F}$ defined as in Equation (9)*

$$\begin{array}{cccc} \mathfrak{F}: & \mathbb{R} \times \mathbb{R}^+ & \rightarrow & (\mathcal{F}(\mathbb{R}), \| \cdot \|_q) \\ & (x, u) & \mapsto & x + \mathcal{U}(u) \end{array}$$

is continuous for every $q \leq \infty$.

Proof. As in the previous result, it suffices to prove it for $q = \infty$. To check the continuity take two convergent sequences $x_n \to x$, $i_n \to i$:

$$\begin{aligned}
&\|\mathfrak{F}(x_n, i_n) - \mathfrak{F}(x, i)\|_\infty = \|x_n + \mathcal{U}(i_n) - x + \mathcal{U}(i)\|_\infty \\
&\leq \|x_n - x\|_\infty + \|\mathcal{U}(i_n) - \mathcal{U}(i)\|_\infty \\
&= |x_n - x| + \|\mathcal{U}(i_n) - \mathcal{U}(i)\|_\infty \\
&= |x_n - x| + \sup_{r \leq 1} \max \left(|(\mathcal{U}(i_n) - \mathcal{U}(i))^+(r)|, |(\mathcal{U}(i_n) - \mathcal{U}(i))^-(r)| \right) \\
&= |x_n - x| + \sup_{r \leq 1} \max \left(|\mathcal{U}(i_n)^+(r) - \mathcal{U}(i)^+(r)|, |\mathcal{U}(i_n)^-(r) - \mathcal{U}(i)^-(r)| \right) \qquad (10) \\
&= |x_n - x| + \sup_{r \leq 1} \max \left(R^{-1}(r) \cdot |\beta(i_n) - \beta(i)|, L^{-1}(r) \cdot |\alpha(i_n) - \alpha(i)| \right) \\
&= |x_n - x| + \max \left(R^{-1}(0) \cdot |\beta(i_n) - \beta(i)|, L^{-1}(1) \cdot |\alpha(i_n) - \alpha(i)| \right) \qquad (11)
\end{aligned}$$

The absolute value $|x_n - x|$ appears because $x_n - x$ is in fact a crisp number. In (10) we used Lemma 1. On the other hand, notice that by the linearity of α, β the last expression converges to zero. $\square$

The following corollary shows that fuzzifications of informed time series are in fact continuous transformations when we consider them as operators between the space of time series and the space of fuzzy time series.

Corollary 1. *The following mapping is continuous for every $q \leq \infty$:*

$$\begin{array}{ccc} (\mathbb{R} \times \mathbb{R}^+)^{\mathbb{N}} & \rightarrow & (\mathcal{F}(\mathbb{R}), \| \cdot \|_q)^{\mathbb{N}} \\ (x_t, u_t)_t & \mapsto & \left(\mathfrak{F}(x_t, u_t)\right)_t \end{array}$$

Proof. Note that the composition of this mapping with any canonical projection is continuous, by Theorem 1. The result follows immediately. $\square$

Lemma 2. *If $(X_t)_t, (U_t)_t$ are stochastic processes, then for every fuzzification $\mathfrak{F}$, the sequence $(\mathcal{X}_t)_t = (\mathfrak{F}(X_t, U_t))_t$ is a fuzzy stochastic process.*

Proof. Given an instant t, consider the random variable $(X_t, U_t) : (\Omega, A, P) \to \mathbb{R} \times \mathbb{R}$. The composition

$$\begin{array}{ccccc} (\Omega, A, P) & \xrightarrow{(X_t, U_t)} & \mathbb{R} \times \mathbb{R} & \xrightarrow{\mathfrak{F}} & (\mathcal{F}(\mathbb{R}), \| \cdot \|_q) \\ \omega & \mapsto & (X_t(\omega), U_t(\omega)) & \mapsto & X_t(\omega) + \mathcal{U}(U_t(\omega)) \end{array} \qquad (12)$$

By Theorem 1, we conclude that $\mathfrak{F}$ is continuous which implies that it is also measurable. Since Equation (12) is a composition of measurable functions it must also be measurable. We have concluded that $\mathfrak{F}(X_t, U_t)$ is a fuzzy random variable. $\square$

Theorem 2. *Suppose that $(X_t)_t$ and $(U_t)_t$ are stochastic processes with finite variance and cross-covariance. Every fuzzification $\mathfrak{F}$ generates a fuzzy stochastic process $(\mathcal{X}_t)_t = (\mathfrak{F}(X_t, U_t))_t$ with finite variance and integrably bounded.*

Proof. Taking into account Equation (5)

$$\begin{aligned}
\|\mathfrak{F}(X_t, U_t)\|_\infty &= \|X_t + \mathcal{U}(U_t)\|_\infty \\
&\leq \|X_t\|_\infty + \|\mathcal{U}(U_t)\|_\infty \\
&\leq \|X_t\|_\infty + \| < 0, \alpha(U_t), \beta(U_t)_{L,R} \|_\infty \\
&= |X_t| + \max\{|L^{-1}(1)| \cdot \alpha(U_t), |R^{-1}(0)| \beta(U_t)\} \\
&\leq |X_t| + |L^{-1}(1)| \cdot \alpha(U_t) + |R^{-1}(0)| \cdot \beta(U_t)
\end{aligned}$$

In view of this

$$\begin{aligned}E\big(\|\mathfrak{F}(X_t,U_t)\|_\infty\big) &\leq E\big(|X_t|\big)+E\big(|L^{-1}(1)|\cdot\alpha(U_t)\big)+E\big(|R^{-1}(0)|\alpha(U_t)\big)\\ &\leq E\big(|X_t|\big)+|L^{-1}(1)|\cdot E\alpha(U_t)+|R^{-1}(0)|\cdot E\beta(U_t)\\ &<\infty.\end{aligned}$$

Notice that in the previous equation we are taking the expectancy of a (crisp) random variable, since the norm transforms the fuzzy random variable into a crisp one.

The last inequalities are due to the fact that by definition $\mathcal{U}^+(\cdot)(r), \mathcal{U}^-(\cdot)(r)$ are linear for every r and the functions α, and β are also linear, which implies that the expectations $E\alpha(U_t)$ and $E\beta(U_t)$ are finite. Since X_t has finite expectation $E|X_t|$ must be finite. Which implies that $\mathfrak{F}(X_t,U_t)$ is locally bounded.

$$\begin{aligned}\|\mathfrak{F}(X_t,U_t)\|_2 &= \|X_t+\mathcal{U}(U_t))\|_2\\ &\leq \|X_t\|_2+\|\mathcal{U}(U_t)\|_2\\ &= \|X_t\|_2+\int_0^1\big(\max\{|\mathcal{U}(U_t)^+(r)|,|\mathcal{U}(U_t)^-(r)|\}\big)dr\\ &= \|X_t\|_2+\int_0^1\big(\max\{|-L^{-1}(r)\cdot\alpha(U_t)|,|R^{-1}(r)\cdot\beta(U_t)|\}\big)dr\\ &= \|X_t\|_2+\int_0^1\big(\max\{|L^{-1}(r)|\cdot\alpha(U_t),|R^{-1}(r)|\cdot\beta(U_t)\}\big)dr\\ &\leq \|X_t\|_2+\int_0^1\big(|L^{-1}(r)|\cdot\alpha(U_t)+|R^{-1}(r)|\beta(U_t)\big)dr\\ &\leq \|X_t\|_2+\alpha(U_t)\cdot\int_0^1|L^{-1}(r)|dr+\beta(U_t)\int_0^1|R^{-1}(r)|dr\end{aligned}$$

Again we find that

$$E\big(\|\mathfrak{F}(X_t,U_t)\|_2\big)\leq E\big(\|X_t\|_2\big)+E\big(\alpha(U_t)\big)\cdot\int_0^1|L^{-1}(r)|dr+E\big(\beta(U_t)\big)\int_0^1|R^{-1}(r)|dr<\infty.$$

Concluding that $\mathfrak{F}(X_t,U_t)\in\mathcal{L}^2(\mathcal{F}(\mathbb{R}))$.

The variance of $\mathfrak{F}(X_t,U_t)$ can be computed using the covariance

$$\begin{aligned}\operatorname{var}\big(\mathfrak{F}(X_t,U_t)\big) &= \operatorname{cov}\big(\mathfrak{F}(X_t,U_t),\mathfrak{F}(X_t,U_t)\big)\\ &= \frac{1}{2}\int_0^1\Big(\operatorname{cov}\big(\mathfrak{F}(X_t,U_t)^-(r),\mathfrak{F}(X_t,U_t)^-(r)\big)+\operatorname{cov}\big(\mathfrak{F}(X_t,U_t)^+(r),\mathfrak{F}(X_t,U_t)^+(r)\big)\Big)dr\\ &= \frac{1}{2}\int_0^1\Big(\operatorname{var}\big(\mathfrak{F}(X_t,U_t)^-(r)\big)+\operatorname{var}\big(\mathfrak{F}(X_t,U_t)^+(r)\big)\Big)dr\\ &= \frac{1}{2}\int_0^1\Big(\operatorname{var}\big((X_t+\mathcal{U}(U_t))^-(r)\big)+\operatorname{var}\big((X_t+\mathcal{U}(U_t))^+(r)\big)\Big)dr\\ &= \frac{1}{2}\int_0^1\Big(\operatorname{var}\big(X_t+\mathcal{U}(U_t)^-(r)\big)+\operatorname{var}\big(X_t+\mathcal{U}(U_t)^+(r)\big)\Big)dr\\ &= \frac{1}{2}\int_0^1\Big(\operatorname{var}\big(X_t-L^{-1}(r)\cdot\alpha(U_t)\big)+\operatorname{var}\big(X_t+R^{-1}(r)\cdot\beta(U_t)\big)\Big)dr\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{2}\int_0^1 2\mathrm{var}(X_t) + L^{-1}(r)^2 \cdot \mathrm{var}\alpha(U_t) + \\
&\quad - 2L^{-1}(r)\cdot \mathrm{cov}(X_t, \alpha(U_t)) + (R^{-1}(r))^2 \cdot \mathrm{var}\beta(U_t) + 2R^{-1}(r)\mathrm{cov}(X_t, \beta(U_t))\Big) dr \\
&= \mathrm{var}(X_t) - \mathrm{cov}(X_t, \alpha(U_t)) \int_0^1 L^{-1}(r)dr + \mathrm{cov}(X_t, \beta(U_t)) \int_0^1 R^{-1}(r)dr + \\
&\quad + \frac{\mathrm{var}\big(\alpha(U_t)\big)}{2} \cdot \int_0^1 L^{-1}(r)^2 dr + \frac{\mathrm{var}\big(\beta(U_t)\big)}{2} \cdot \int_0^1 R^{-1}(r)^2 dr \\
&< \infty.
\end{aligned}
$$

Therefore, the proof is completed. □

Theorem 3. *Suppose that $(X_t)_t$ and $(U_t)_t$ are stationary and cross-stationary stochastic processes. Then every fuzzification $(\mathcal{X}_t)_t = (\mathfrak{F}(X_t, U_t))_t$ is a fuzzy stationary stochastic process.*

Proof. By Theorem 2, we know that the series $(\mathfrak{F}(X_t, U_t))_t$ has finite expectation and variance.

Let us calculate the expectation. Knowing that by the stationarity of U_t and X_t, their expectations are constant $EX_t = \mu_X$, $EU_t = \mu_I$.

$$
\begin{aligned}
[E\big(\mathfrak{F}(X_t, U_t)\big)]_r &= [E\big(\mathfrak{F}(X_t, U_t)^-(r)\big), E\big(\mathfrak{F}(X_t, U_t)^+(r)\big)] \\
&= [E\big(X_t + \mathcal{U}(U_t)^+(r)\big), E\big(X_t + \mathcal{U}(U_t)^-(r)\big)] \\
&= [E\big(X_t - L^{-1}(r)\cdot \alpha(U_t)\big), E\big(X_t + R^{-1}(r)\cdot \beta(U_t)\big)] \\
&= [E(X_t) - L^{-1}(r)\cdot E\big(\alpha(U_t)\big), E\big(X_t\big) + R^{-1}(r)\cdot E\big(\beta(U_t)\big)] \\
&= [E(X_t) - L^{-1}(r)\cdot \alpha E(U_t), E\big(X_t\big) + R^{-1}(r)\cdot \beta\big(E(U_t)\big)] \\
&= [\mu_X - L^{-1}(r)\cdot \alpha(\mu_I), \mu_X + R^{-1}(r)\cdot \beta(\mu_I)]
\end{aligned}
$$

the last equality is due to the linearity of α and β. We have deduced that the expectation of $\mathfrak{F}(X_t, U_t)$ is a constant fuzzy number.

Let us compute the autocovariance,

$$
\begin{aligned}
\mathrm{cov}\big(\mathfrak{F}(X_t, U_t), \mathfrak{F}(X_{t'}, I_{t'})\big) &= \frac{1}{2}\int_0^1 \Big(\mathrm{cov}\big(\mathfrak{F}(X_t, U_t)^-(r), \mathfrak{F}(X_{t'}, I_{t'})^-(r)\big) \| \\
&\quad + \mathrm{cov}\big(\mathfrak{F}(X_t, I_t)^+(r), \mathfrak{F}(X_{t'}, I_{t'})^+(r)\big)\Big) dr \\
&= \frac{1}{2}\int_0^1 \Big(\mathrm{cov}\big(X_t - L^{-1}(r)\cdot\alpha(U_t), X_{t'} - L^{-1}(r)\cdot\alpha(I_{t'})\big) \\
&\quad + \mathrm{cov}\big(X_t + R^{-1}(r)\cdot\beta(I_t), X_{t'} + R^{-1}(r)\cdot\beta(I_{t'})\big)\Big) dr \\
&= \frac{1}{2}\int_0^1 \Big(\mathrm{cov}(X_t, X_{t'}) + \mathrm{cov}\big(X_t, -L^{-1}(r)\alpha(I_{t'})\big) \\
&\quad + \mathrm{cov}\big(-L^{-1}(r)\alpha(I_t), X_{t'}\big) + \mathrm{cov}\big(-L^{-1}(r)\alpha(I_t), R^{-1}(r)\beta(I_{t'})\big) \\
&\quad \mathrm{cov}(X_t, X_{t'}) + \mathrm{cov}\big(X_t, R^{-1}(r)\beta(I_{t'})\big) \\
&\quad + \mathrm{cov}\big(R^{-1}(r)\beta(I_t), X_{t'}\big) + \mathrm{cov}\big(R^{-1}(r)\beta(I_t), R^{-1}(r)\beta(I_{t'})\big)\Big) dr \\
&= \frac{1}{2}\int_0^1 \Big(\mathrm{cov}(X_{t+h}, X_{t'+h}) + \mathrm{cov}\big(X_{t+h}, -L^{-1}(r)\alpha(I_{t'+h})\big) \qquad (13) \\
&\quad + \mathrm{cov}\big(-L^{-1}(r)\alpha(I_{t+h}), X_{t'+h}\big) + \mathrm{cov}\big(-L^{-1}(r)\alpha(I_{t+h}), R^{-1}(r)\beta(I_{t'+h})\big)
\end{aligned}
$$

$$
\begin{aligned}
&\operatorname{cov}(X_{t+h}, X_{t'+h}) + \operatorname{cov}\big(X_{t+h}, R^{-1}(r)\beta(I_{t'+h})\big) \\
&\quad + \operatorname{cov}\big(R^{-1}(r)\beta(I_{t+h}), X_{t'+h}\big) + \operatorname{cov}\big(R^{-1}(r)\beta(I_{t+h}), R^{-1}(r)\beta(I_{t'+h})\big)\Big)\,dr \\
&= \operatorname{cov}\big(\mathfrak{F}(X_{t+h}, I_{t+h}), \mathfrak{F}(X_{t'+h}, I_{t'+h})\big)
\end{aligned}
$$

We used that due to the stationarity of X_t and U_t, $\operatorname{cov}(X_t, X_{t'}) = \operatorname{cov}(X_{t+h}, X_{t'+h})$, $\operatorname{cov}(I_t, I_{t'}) = \operatorname{cov}(I_{t+h}, I_{t'+h})$, and by the cross-stationarity $\operatorname{cov}(X_t, I_{t'}) = \operatorname{cov}(X_{t+h}, I_{t'+h})$. This combined with the linearity of α and β gives us Equation (13). ☐

4. Fuzzification algorithm

In the cases in which the uncertainty set U is a subset of the real numbers there is a natural way to define a **fuzzification** which is described in the following algorithm:

Algorithm 2 Fuzzification of informed time series $x = (x_t, u_t)_t$.

```
  Input: Informed time series $x = (x_t, u_t)_t$ and maximum spread parameters $\alpha_0$, $\beta_0$
  Output $(X_t)_t$ fuzzified time series
Step 1 Define the functions
    $\alpha : \mathbb{R}^+ \to \mathbb{R}$
    $u \mapsto \frac{u - \min((u_t)_t)}{\max((u_t)_t) - \min((u_t)_t)} \cdot \alpha_0$          (14)
    $\beta : \mathbb{R}^+ \to \mathbb{R}$
    $u \mapsto \frac{u - \min((u_t)_t)}{\max((u_t)_t) - \min((u_t)_t)} \cdot \beta_1$          (15)
Step 2 Define the uncertainty mapping
    $\mathcal{U} : \mathbb{R}^+ \to \mathcal{F}(\mathbb{R})$
    $u \mapsto U =< 0, \alpha(u), \beta_(u) >_{L,R}$
Define the fuzzification $(X_t)_t$ where
    $X_t = \mathfrak{F}(x_t, u_t) = x_t + U_t =< x_t, \alpha(u_t), \beta_(u_t) >_{LR}$
  return $(X_t)_t$
```

Remark 7. Notice that the parameters α_0, β_0, are provided as inputs, as it is specified in the pseudocode, these parameters refer to the maximum spread values in the fuzzification.

There are many ways to select these parameters, one possibility for the value of α_0 is to choose:

$$\alpha_0 = \operatorname{Median}\big((x_t)_t\big) - Q_1\big((x_t)_t\big) \tag{16}$$

Similarly, we can select β_0 as

$$\beta_0 = Q_3\big((x_t)_t\big) - \operatorname{Median}\big((x_t)_t\big) \tag{17}$$

Notice that the idea behind these choices, namely Equation (16) and Equation (17), is to assign the maximum left and right spread in the fuzzification that corresponds to the dispersion of the data (x_t) measured as the distance between the median and the quartiles. Any other quantification of the left and right dispersion will work similarly.

The factor $\frac{u - \min((u_t)_t)}{\max((u_t)_t) - \min((u_t)_t)}$ in Equation (14) and Equation (15) is there to ensure that the maximum (minimum) value of u_t is associated with the maximum (minimum) amount of spread and therefore the maximum (minimum) value of uncertainly in the resulting fuzzy number of the fuzzification.

Example 4. An illustrative example of a fuzzy time series can be constructed using gold price values. Let x_t represent the average price of gold on a given day t. In this example, we quantify the volatility or unreliability of each observation (uncertainty) u_t by calculating the difference between the maximum and minimum values:

$$u_t = (\text{high value of gold within day } t) - (\text{low value of gold within day } t)$$

The application of the fuzzy time series algorithm, as outlined in Algorithm 2, yields the resulting representation presented in Fig. 1.

For the representation displayed in the Fig. 1, we draw at each time t the fuzzy number $\mathfrak{F}(x_t, u_t)$ returned by Algorithm 2 as a blue gradient. Darker values of this gradient represent higher values of the membership function of $\mathfrak{F}(x_t, u_t)$, which in this context can be interpreted as certainty or security of the value x_t at time t.

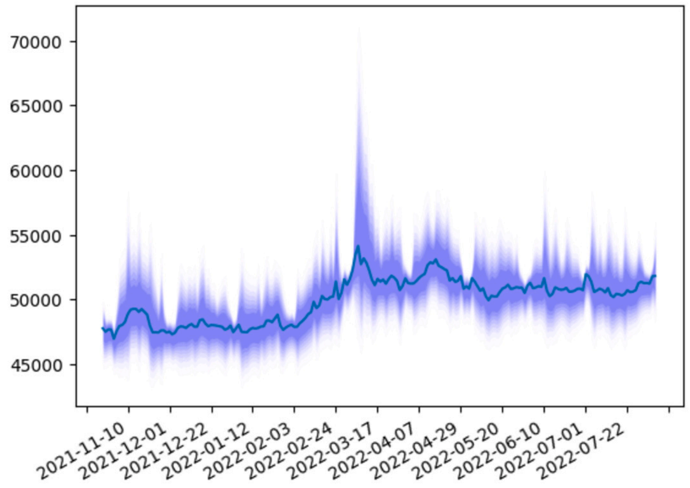


Fig. 1. Prize of gold fuzzified time series. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

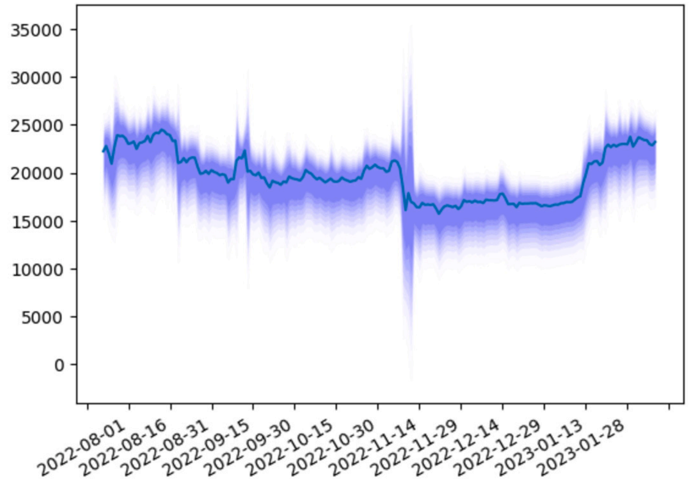


Fig. 2. Prize of bitcoin fuzzified time series.

Notably, a spike in March 2023 is observed, which is likely attributed to the war in Ukraine. Surrounding this high point, a significant degree of uncertainty is evident within the fuzzy time series. The expanded spread can be attributed to the substantial variation observed in the gold price during that particular period, as reflected by the values of u_t.

It is worth noting that the depiction of uncertainty in Fig. 1 effectively highlights the proposed algorithm's potential for visualization purposes, akin to the utilization of candle-bar representations in financial time series analysis.

Example 5. In the spirit of Example 4 we will apply our analysis to study a highly volatile time series as the bitcoin prizes. In this case x_t will represent the close value of bitcoin at a given day. To construct the informed time series we have used

$$u_t = \Big((\text{high value within day } t) - (\text{low value within day } t) \Big)^2$$

We use the square here to emphasize the high volatility measurements.

Using Algorithm 2 with the informed time series $(x_t, u_t)_t$ we obtain the fuzzy-valued time series showed in Fig. 2. For these figures we have use the same technique for representing fuzzy-valued time series thet we described in Example 4.

Notice the sudden increase in uncertainty reflected in Fig. 2 at around November 2022, this corresponds with the bitcoin crash of 2022.

5. Fuzzy stationary time series with structural breaks

Stationarity with structural breaks in fuzzy time series refers to the examination of whether a fuzzy time series retains its stationary properties despite significant shifts in its underlying dynamics. This analysis explores the impact of structural breaks on the stationarity assumption and provides insights into the persistence and stability of patterns within the data. Understanding stationarity with structural breaks in fuzzy time series has practical implications for forecasting, decision-making, and risk assessment, enabling the refinement of models and improving the reliability of fuzzy time series analysis.

Definition 12 *([48]).* A crisp time series (x_t) is called **stationary with structural breaks** if it satisfies

$$x_t = (\eta_t + \alpha)D_t + (1 - D_t)\varepsilon_t \tag{18}$$

where

$$D_t = \begin{cases} 1 & \text{if } t \in A, \\ 0 & \text{otherwise,} \end{cases}$$

for $\alpha \in \mathbb{R}$, η_t, ε_t stationary time series and A an interval (in most applications of the form (T_b, ∞))

These time series were introduced by Perron [48,49]. This type of time series is a very interesting generalization of the notion of stationarity since it includes a series with all sorts of pathologies but that preserve some of the interesting properties of stationary time series but locally.

The following definition is the adaptation of Definition 12 in the context of fuzzy time series.

Definition 13. A fuzzy time series $(\mathcal{X}_t)$ is called **stationary with structural breaks** if it satisfies an equation as Equation (18) but for fuzzy stationary time series.

Theorem 4. *Let us assume (X_t, U_t) is stationary with structural breaks in the sense that*

$$X_t = (\eta_t + \alpha)D_t + (1 - D_t)\varepsilon_t$$

$$U_t = (\eta'_t + \alpha')D_t + (1 - D_t)\varepsilon'_t$$

$$D_t = \begin{cases} 1 & \textit{if } t \in A, \\ 0 & \textit{otherwise.} \end{cases}$$

Suppose also that η_t, η'_t and $\varepsilon_t, \varepsilon'_t$ are cross-stationary, then every fuzzification $\mathfrak{F}(X_t, U_t)$ is also stationary with structural breaks.

Proof. By Proposition 1,

$$\begin{aligned} \mathfrak{F}(X_t, U_t) &= \mathfrak{F}(\eta_t + \alpha)D_t + (1 - D_t)\varepsilon_t, (\eta'_t + \alpha')D_t + (1 - D_t)\varepsilon'_t) \\ &= \mathfrak{F}\big((\eta_t + \alpha)D_t, (\eta'_t + \alpha')D_t\big) + \mathfrak{F}\big((1 - D_t)\varepsilon_t, (1 - D_t)\varepsilon'_t\big) \\ &= \mathfrak{F}\big(\eta_t + \alpha), (\eta'_t + \alpha')\big)D_t + \mathfrak{F}\big(\varepsilon_t, \varepsilon'_t\big)(1 - D_t) \\ &= \big(\mathfrak{F}(\eta_t, \eta'_t) + \mathfrak{F}(\alpha, \alpha')\big)D_t + \mathfrak{F}(\varepsilon_t, \varepsilon'_t)(1 - D_t) \end{aligned} \tag{19}$$

Equation (19) follows easily from the definition of $\mathfrak{F}$. By Theorem 3, since the pairs η_t, η'_t and $\varepsilon_t, \varepsilon'_t$ are stationary and cross-stationary, we deduce that both $\mathfrak{F}(\eta_t, \eta'_t)$ and $\mathfrak{F}(\varepsilon_t, \varepsilon'_t)$ are stationary fuzzy stochastic processes. Since $\mathfrak{F}(\alpha, \alpha')$ is simply a fuzzy number, we deduce that $\mathfrak{F}(X_t, U_t)$ also satisfies the definition of structural breaks but in the sense of fuzzy stochastic processes (Definition 13). ☐

6. Applications and computational examples

The application of fuzzy logic in dynamic time warping is a subject that has been widely explored in the last years ([50], [51], [52], [53]). These works deploy different strategies to apply fuzzy logic in the process of processing time series data through variations of the DTW algorithm and utilize this method in problems such as clustering or forecasting.

In this section, we apply a very simple extension of the DTW algorithm that takes advantage of the methods developed in this work. Algorithm 1 can be easily adapted for informed time series, the trick simply consists of measuring the cost using the distance d_∞ instead of the euclidean distance d.

The pseudocode for this version of DTW is the following:

Algorithm 3 Dynamic time warping distance between informed time series.

```
  Input: $(x_t, i^x_t)^N_{t=0}, y = (y_t, i^y_t)^M_{t=0}$
  Output DTW distance

$\mathcal{X} \leftarrow \mathfrak{F}((x_t, i^x_t)_t)$
$\mathcal{Y} \leftarrow \mathfrak{F}((y_t, i^y_t)_t)$
$D \leftarrow (\infty)_{N \times M}$
$D_{ij} \leftarrow 0$

for $i = 1$ to $N$ do
    for $j = 1$ to $M$ do
        $cost \leftarrow d_\infty(\mathcal{X}_i, \mathcal{Y}_j)$
        $D_{ij} \leftarrow cost + \min\{D_{i-1j}, D_{ij-1}, D_{i-1j-1}\}$
    end for
end for
return $D_{NM}$
```

In the simulated and real datasets in the following sections, we will use the described algorithm to fuzzify informed time series and test if the proposed version of DTW is capable of classifying the time series better that the conventional DTW method with crisp time series.

6.1. *Simulated dataset*

We will start testing Algorithm 2 in conjunction with 3 for the classification of time series obtained using simulations.

To generate the simulated dataset, we used stationary time series with structural breaks composed in the following way:

$$x_t = D_t \cdot v_t + (1 - D_t) \cdot \varepsilon_t \tag{20}$$

where

$$D_t = \begin{cases} 1 & t \in [t_0, t_1] \\ 0 & \text{otherwise} \end{cases}$$

and ε_t is a white noise and v_t is an AR(1) time series of the form

$$v_t = \phi_1 v_{t-1} + \delta_t$$

Where δ_t is also white noise and the parameter ϕ_1 takes different values depending on the class that we want to generate to test the classification algorithm.

The simulated time series (20) has a very characteristic noisy region between t_0 and t_1 and an AR(1) region with parameter ϕ_1 outside of this interval. The value of ϕ_1 of the non-noisy region is what really characterizes (x_t) since negative values of ϕ_1 produce much more oscillatory time series than positive values of ϕ_1.

To apply Algorithm 2 we need *informed time series*, in the case of the described dataset there is a very natural way to transform them into this kind of series. To that end, define the information i_t as

$$i_t = D_t \tag{21}$$

so that it takes the value 1 inside the noise interval and zero outside. The choice (21) ensures that the less informative values are those inside the noisy interval $[t_0, t_1]$ since those are the most irregular or unimportant for the task of classification. Therefore, it is penalized with higher values of i_t.

Notice that Theorem 4 implies that the fuzzifications of this type of series will produce stationary fuzzy time series with structural breaks.

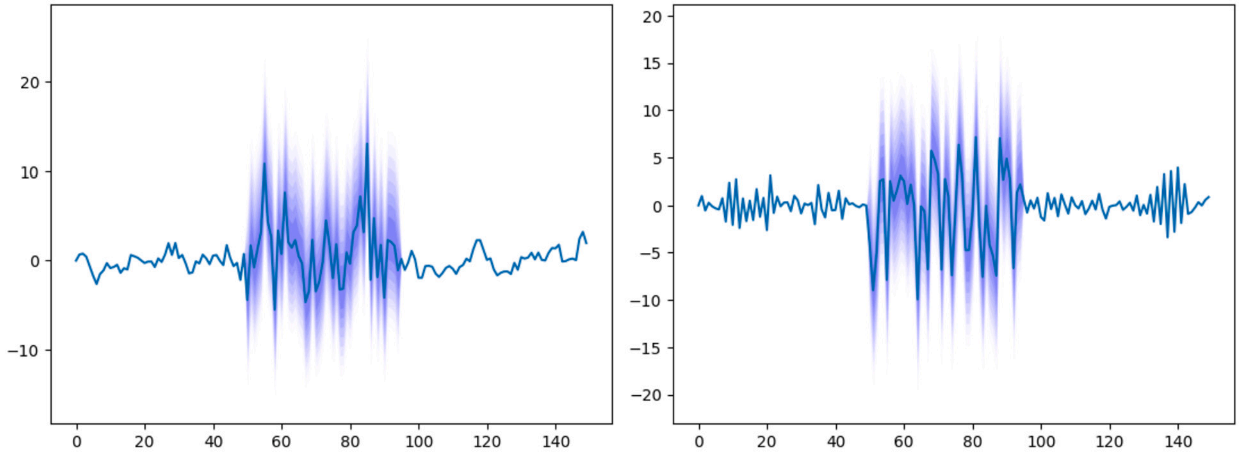
The following figure contains two examples of fuzzified time series simulated using this procedure for positive and negative values of ϕ_1. The fuzzification was obtained using Algorithm 2. See Fig. 3.

As we see by taking $i_t = D_t$ the fuzzification blurs the noisy regions in the time series which are the less relevant values for the classification. As we will see this impacts positively in the classification results using the proposed fuzzy version of the Algorithm 3.

Table 1 shows the results of the simulation experiments comparing the DTW classification 1 precision (measured on the test dataset) in comparison with the fuzzy version of DTW 3 applied on the fuzzification of the time series.

The objective of the classification problem here is to differentiate negative values of the parameter ϕ_1 (what we grouped and tagged as class 1 dataset) from positive values of ϕ_1 (tagged as class 2 dataset) As usual in classification problems, for the algorithm evaluation we used two datasets for each experiment: A training dataset (used for training the model) and a test dataset (used for evaluating the effectiveness of the model).

The fuzzy version of DTW algorithm consistently outperforms the crisp version for the simulated datasets as can be seen in Table 1.



AR(1) with $\phi_1 = 0.8$ and noisy structural break

AR(1) with $\phi_1 = -0.8$ and noisy structural break

Fig. 3. Fuzzification of simulated dataset examples. The left side is a simulation (20) for positive values of the AR(1) parameter ϕ_1. The right side is a similar simulation but for negative values of ϕ_1. Notice that the distinct characteristics of the left and right graphs appear outside the blurred noisy areas.

Table 1
Simulated dataset classification algorithm comparison. Simulated AR with structural breaks (20) fuzzified using Algorithm 2.

Test dataset description	Training dataset description	Precision results DTW	Precision results fuzzy DTW
• 50 instances of AR simulated with $\phi_1 = 0.2$ (Class 1). • 50 instances of AR simulated with $\phi_1 = -0.2$ (Class 2)	• 100 instances of AR simulated with $\phi_1 = 0.2$ (**Class 1**) • 100 instances of AR simulated with $\phi_1 = -0.2$ (**Class 2**)	0.45	**0.55**
• 50 instances of AR simulated with $\phi_1 = 0.5$ (Class 1). • 50 instances of AR simulated with $\phi_1 = -0.5$ (Class 32)	• 100 instances of AR simulated with $\phi_1 = 0.5$ (**Class 1**). • 100 instances of AR simulated with $\phi_1 = -0.5$ (**Class 2**)	0.52	**0.54**
• 50 instances of AR simulated with $\phi_1 = 0.6$ (Class 1). • 50 instances of AR simulated with $\phi_1 = -0.6$ (Class 2)	• 100 instances of AR simulated with $\phi_1 = 0.6$ (**Class 1**). • 100 instances of AR simulated with $\phi_1 = -0.6$ (**Class 2**)	0.47	**0.54**
• 50 instances of AR simulated with $\phi_1 = 0.7$ (Class 1). • 50 instances of AR simulated with $\phi_1 = -0.7$ (Class 2)	• 100 instances of AR simulated with $\phi_1 = 0.7$ (**Class 1**). • 100 instances of AR simulated with $\phi_1 = 0.7$ (**Class 2**)	0.48	**0.54**

6.2. *Real dataset*

Food spectrographs are used in chemometrics to classify food types, a task that has obvious applications in food safety and quality assurance. The objective of this type of analysis is to build classifiers so that food type can be identified from the spectrum alone.

We will focus on the Olive Oil Dataset, which was introduced in [54]. This dataset consists of four classes of time series, each class representing an extra virgin olive oil from alternative countries. The classes are composed in the following way:

- **Class 1**. 10 instances of olive oil spectrographs from Greece.
- **Class 2**. 17 instances of olive oil spectrographs from Italy.
- **Class 3**. 8 instances of olive oil spectrographs from Portugal.
- **Class 4**. 25 instances of olive oil spectrographs from Spain.

To apply our proposed fuzzification algorithm, we need to induce which observation is less reliable or informative (that is, we need to generate *informed time series*. The decided approach here is to penalize regions with high variance. To do this we obtained the structurally different regions using the hidden Markov chains, and associated higher values of i_t to regions with high variance. The result can be visualized in Fig. 4.

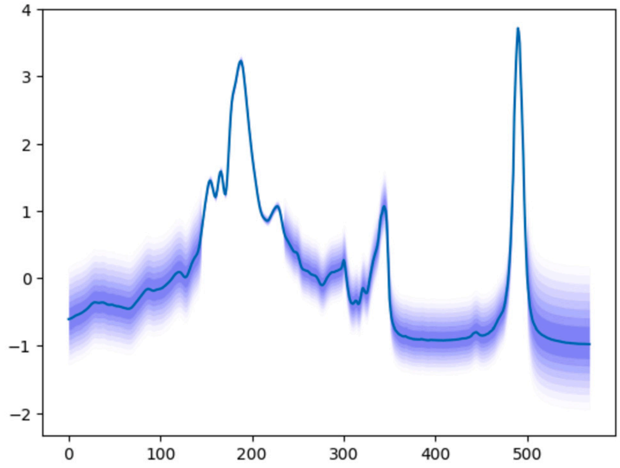


Fig. 4. Fuzzification of olive oil dataset example.

Table 2
Olive oil dataset analysis comparison.

Test dataset description	Train dataset description	Precision results DTW	Precision results fuzzy DTW
30 instances of time series from classes 1 to 4	30 instances of time series from classes 1 to 4	0.83	**0.86**

Following the same idea as with the simulated dataset, we applied the DTW algorithm and the proposed DTW fuzzy version dividing the dataset into train and test datasets. The following Table 2 contains the results of the analysis.

As we see, the proposed DTW algorithm outperforms the classical DTW for the olive oil dataset.

7. Conclusion and future works

Our results contribute to the field of fuzzy time series analysis by introducing a novel method that leverages fuzzy numbers as a foundational component. By incorporating fuzzy numbers into the analysis of time series data, a significant representation of uncertainty and vagueness has been achieved.

The utilization of fuzzy numbers offers a framework for capturing and modeling the inherent uncertainty present in time series data. The proposed algorithm enhances our ability to comprehend and analyze complex temporal patterns, particularly in domains where traditional crisp time series methods may fall short.

The findings presented in this article have demonstrated the efficacy, showcasing its potential for yielding meaningful insights and results. By employing fuzzy numbers, the method enables a robust representation of uncertainty, enabling researchers and practitioners to better understand and interpret time series data in various domains. We emphasize that our algorithm outperforms the classical one.

Furthermore, this article underscores the significance of employing topological tools in the analysis of fuzzy time series. One key result in this direction is Corollary 1 which shows how fuzzifications are additive and continuous operators. This fact gives us an unexpected connection between the topology of crisp and fuzzy time series. The utilization of topological tools in this article highlights the interdisciplinary nature of fuzzy time series analysis, incorporating mathematical concepts from topology to enhance the analysis and interpretation of fuzzy time series data.

Future research endeavors may involve refining the methodology, exploring alternative approaches for fuzzy number representation, or investigating the method's applicability to diverse real-world datasets. In addition, similar studies from the epistemic point of view for fuzzy data can be considered.

A promising avenue for future research in fuzzy time series analysis involves advancing dynamic time warping (DTW) techniques. Expanding DTW to accommodate fuzzy time series data, considering their inherent uncertainty, presents an opportunity to enhance similarity analysis, pattern recognition, classification, and forecasting in practical domains. Developing novel distance measures and alignment strategies tailored to fuzzy time series would enable accurate and effective analysis, contributing to a deeper understanding and application of fuzzy time series analysis in real-world scenarios.

CRediT authorship contribution statement

Hugo J. Bello: Writing – review & editing, Writing – original draft, Investigation, Formal analysis, Conceptualization. **Manuel Ojeda-Hernández:** Writing – review & editing, Writing – original draft, Investigation, Formal analysis, Conceptualization. **Domingo**

López-Rodríguez: Writing – review & editing, Writing – original draft, Investigation, Formal analysis, Conceptualization. **Carlos Bejines:** Writing – review & editing, Writing – original draft, Investigation, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Carlos Bejines reports financial support was provided by University of Malaga. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work has been partially funded by the State Agency of Research (AEI), the Ministerio de Ciencia, Innovación y Universidades (MCIU), the European Social Research Fund (FEDER), the Junta de Andalucía (JA), y la Universidad de Málaga (UMA) through the PhD contract FPU19/01467 (MCIU), the VALID research project (PID2022-140630NB-I00) and the research project PID2021-127870OB-I00.

Data availability

No data was used for the research described in the article.

References

[1] C.-C. Wang, A comparison study between fuzzy time series model and arima model for forecasting Taiwan export, Expert Syst. Appl. 38 (8) (2011) 9296–9304.

[2] R. Hyndman, A.B. Koehler, J.K. Ord, R.D. Snyder, Forecasting with Exponential Smoothing: the State Space Approach, Springer Science & Business Media, 2008.

[3] D. Cheng, F. Yang, S. Xiang, J. Liu, Financial time series forecasting with multi-modality graph neural network, Pattern Recognit. 121 (2022) 108218.

[4] Q. Song, B.S. Chissom, Fuzzy time series and its models, Fuzzy Sets Syst. 54 (3) (1993) 269–277.

[5] Q. Song, B.S. Chissom, Forecasting enrollments with fuzzy time series – part I, Fuzzy Sets Syst. 54 (1) (1993) 1–9.

[6] Q. Song, B.S. Chissom, Forecasting enrollments with fuzzy time series – part II, Fuzzy Sets Syst. 62 (1) (1994) 1–8.

[7] M. Bose, K. Mali, Designing fuzzy time series forecasting models: a survey, Int. J. Approx. Reason. 111 (2019) 78–99.

[8] Y.-M. Wang, Centroid defuzzification and the maximizing set and minimizing set ranking based on alpha level sets, Comput. Ind. Eng. 57 (1) (2009) 228–236.

[9] H. Hellendoorn, C. Thomas, Defuzzification in fuzzy controllers, J. Intell. Fuzzy Syst. 1 (2) (1993) 109–123.

[10] M. Namazov, O. Basturk, Dc motor position control using fuzzy proportional-derivative controllers with different defuzzification methods, Turk. J. Fuzzy Syst. (2010).

[11] S. Naaz, A. Alam, R. Biswas, Effect of different defuzzification methods in a fuzzy based load balancing application, Int. J. Comput. Sci. Issues 8 (5) (2011) 261.

[12] H. Kwakernaak, Fuzzy random variables—ii. Algorithms and examples for the discrete case, Inf. Sci. 17 (3) (1979) 253–278.

[13] M.L. Puri, D.A. Ralescu, L. Zadeh, Fuzzy random variables, in: Readings in Fuzzy Sets for Intelligent Systems, Elsevier, 1993, pp. 265–271.

[14] G. Wang, Y. Zhang, The theory of fuzzy stochastic processes, Fuzzy Sets Syst. 51 (2) (1992) 161–178.

[15] C.-H.L. Lee, A. Liu, W.-S. Chen, Pattern discovery of fuzzy time series for financial prediction, IEEE Trans. Knowl. Data Eng. 18 (5) (2006) 613–625.

[16] R.-C. Tsaur, et al., A fuzzy time series-Markov chain model with an application to forecast the exchange rate between the Taiwan and us dollar, international journal of innovative computing, Inf. Control 8 (7) (2012) 4931–4942.

[17] S.-M. Chen, J.-R. Hwang, Temperature prediction using fuzzy time series, IEEE Trans. Syst. Man Cybern., Part B, Cybern. 30 (2) (2000) 263–275.

[18] H.J. Sadaei, P.C.d.L. e Silva, F.G. Guimaraes, M.H. Lee, Short-term load forecasting by using a combined method of convolutional neural networks and fuzzy time series, Energy 175 (2019) 365–377.

[19] D. Dubois, H. Prade, Using fuzzy sets in flexible querying: why and how?, in: Flexible Query Answering Systems, Springer, 1997, pp. 45–60.

[20] D. Dubois, J. Mengin, H. Prade, Possibilistic Uncertainty and Fuzzy Features in Description Logic. A Preliminary Discussion, Capturing Intelligence, vol. 1, Elsevier, 2006, pp. 101–113.

[21] D. Dubois, E. Hüllermeier, H. Prade, A systematic approach to the assessment of fuzzy association rules, Data Min. Knowl. Discov. 13 (2006) 167–192.

[22] I. Couso, D. Dubois, Statistical reasoning with set-valued information: ontic vs. Epistemic views, Int. J. Approx. Reason. 55 (7) (2014) 1502–1518.

[23] R. Körner, On the variance of fuzzy random variables, Fuzzy Sets Syst. 92 (1) (1997) 83–93.

[24] R. Kruse, K.D. Meyer, Statistics with Vague Data, vol. 6, Springer Science & Business Media, 2012.

[25] J.J. Salamanca, I. Couso, The minimum variance of a random set on a Euclidean space, Fuzzy Sets Syst. 443 (2022) 106–126.

[26] I. Couso, L. Sánchez, Machine learning models, epistemic set-valued data and generalized loss functions: an encompassing approach, Inf. Sci. 358 (2016) 129–150.

[27] I. Molchanov, F. Molinari, Applications of random set theory in econometrics, Annu. Rev. Econ. 6 (1) (2014) 229–251.

[28] M. López-Díaz, M.Á. Gil, P. Grzegorzewski, O. Hryniewicz, J. Lawry, I.S. Molchanov, Applications of random sets in image analysis. How to average a cat and a dog?, in: Soft Methodology and Random Information Systems, Springer, 2004, pp. 8–18.

[29] I. Couso, D. Dubois, On the variability of the concept of variance for fuzzy random variables, IEEE Trans. Fuzzy Syst. 17 (5) (2009) 1070–1080, https://doi.org/10.1109/TFUZZ.2009.2021617.

[30] D. Dubois, H. Prade, Fuzzy numbers: an overview, Read. Fuzzy Sets Inf. Syst. (1993) 112–148.

[31] D. Dubois, H. Prade, Fundamentals of Fuzzy Sets, vol. 7, Springer Science & Business Media, 2012.

[32] D. Dubois, H. Prade, Operations on fuzzy numbers, Int. J. Syst. Sci. 9 (6) (1978) 613–626.

[33] D.J. Dubois, Fuzzy Sets and Systems: Theory and Applications, vol. 144, Academic Press, 1980.

[34] M. Hanss, Applied Fuzzy Arithmetic, Springer, 2005.

[35] P. Diamond, P. Kloeden, Metric spaces of fuzzy sets, Fuzzy Sets Syst. 35 (2) (1990) 241–249.

[36] Y. Feng, L. Hu, H. Shu, The variance and covariance of fuzzy random variables and their applications, Fuzzy Sets Syst. 120 (3) (2001) 487–497.

[37] S. Montes, I. Couso, P. Gil, C. Bertoluzza, Divergence measure between fuzzy sets, Int. J. Approx. Reason. 30 (2) (2002) 91–105.

[38] R. Féron, Ensembles aleatoires flous, 1976.

[39] H. Kwakernaak, Fuzzy random variables—i. Definitions and theorems, Inf. Sci. 15 (1) (1978) 1–29.

[40] E.P. Klement, M. Puri, D. Ralescu, Limit theorems for fuzzy random variables, Proc. R. Soc. Lond. Ser. A, Math. Phys. Sci. 407 (1986) 171–182, (1832).

[41] M.L. Puri, D.A. Ralescu, The concept of normality for fuzzy random variables, Ann. Probab. (1985) 1373–1379.
[42] R. Körner, W. Näther, On the variance of random fuzzy variables, in: Statistical Modeling, Analysis and Management of Fuzzy Data, Springer, 2002, pp. 25–42.
[43] I. Couso, D. Dubois, S. Montes, L. Sánchez, On various definitions of the variance of a fuzzy random variable, in: 5th International Symposium on Imprecise Probability: Theories and Applications (ISIPTA 2007), Sipta, 2007, pp. 135–144.
[44] V. Krätschmer, A unified approach to fuzzy random variables, Fuzzy Sets Syst. 123 (1) (2001) 1–9.
[45] M.A. Lubiano, M.A. Gil, Estimating the expected value of fuzzy random variables in random samplings from finite populations, Stat. Pap. 40 (3) (1999) 277–295.
[46] R.H. Shumway, D.S. Stoffer, D.S. Stoffer, Time Series Analysis and Its Applications, vol. 3, Springer, 2000.
[47] E. Fix, J. Hodges, Discriminatory analysis, nonparametric discrimination, consistency properties, Technical Report 4, United States Air Force, 1951, School . . . .
[48] P. Perron, The great crash, the oil price shock, and the unit root hypothesis, Econometrica J. Econom. Soc. (1989) 1361–1401.
[49] J. Lee, M. Strazicich, Testing the null of stationarity in the presence of a structural break, Appl. Econ. Lett. 8 (6) (2001) 377–382.
[50] A. Safari, R. Hosseini, M. Mazinani, A type-2 fuzzy time series model for pattern similarity analysis: a case study on air quality forecasting, IEEE Intell. Syst. 37 (2) (2021) 92–102.
[51] A. Safari, R. Hosseini, M. Mazinani, Dynamic type-2 fuzzy time warping (dt2ftw): a hybrid model for uncertain time-series prediction, Inf. J. Fuzzy Log. Int. Syst. 21 (4) (2021) 338–348.
[52] P. D'Urso, L. De Giovanni, R. Massari, Trimmed fuzzy clustering of financial time series based on dynamic time warping, Ann. Oper. Res. 299 (2021) 1379–1395.
[53] H. Izakian, W. Pedrycz, I. Jamal, Fuzzy clustering of time series data using dynamic time warping distance, Eng. Appl. Artif. Intell. 39 (2015) 235–244.
[54] H.S. Tapp, M. Defernez, E.K. Kemsley, Ftir spectroscopy and multivariate analysis can distinguish the geographic origin of extra virgin olive oils, J. Agric. Food Chem. 51 (21) (2003) 6110–6115.