

Monitoring of segmented processes using p-box aggregated data and fuzzy control charts*

by

Olgierd Hryniewicz* and Katarzyna Kaczmarek-Majer

Systems Research Institute, Polish Academy of Sciences,
Newelska 6, 01-447 Warszawa, Poland

* Corresponding author: Olgierd.Hryniewicz@ibspan.waw.pl

Abstract: The monitoring of inhomogeneous and non-stationary processes composed of segments and subsegments is considered. The structure of this segmentation is typical for medical data, describing voice characteristics of Bipolar Disorder (BD) psychiatric patients, calculated from their recorded smartphone calls. Data from subsegments are described by different probability distributions and are represented by histograms. Then, data from subsegments belonging to the same segment are aggregated using probability boxes (p-boxes) methodology and a simple probabilistic method. Finally, the mean value of each of the aggregated segments is described by a fuzzy triangular number. Therefore, the stream of consecutive segments is represented by the stream of fuzzy numbers. Several control charts for such fuzzy data are proposed. Their statistical properties are evaluated using simulated synthetic data. The simulation model is related to the real-life data obtained from the monitoring of BD patients. The results of simulations demonstrate the applicability of the proposed procedure for monitoring of BD patients.

Keywords: monitoring of segmented processes, aggregation of segmented processes, fuzzy control charts, p-boxes, monitoring of Bipolar Disorder patients

1. Introduction

In recent years, we have witnessed a proliferation of data-collecting devices, such as sensors, smartphones, network bots, and others. The data collected by these devices are typically represented as streams of numerical values that require preprocessing for further analysis. Such preprocessing becomes necessary

*Submitted: November 2025; Accepted: December 2025.

when applying analytical tools directly to raw data is computationally demanding. One effective method of preprocessing is to aggregate process data into consecutive segments, thereby creating a more granular structure that is often essential for meaningful interpretation.

Real-life data streams are usually inhomogeneous and non-stationary. This implies that data collected during different observation periods may follow different probability distributions, with parameters that change over time. For instance, continuously measured medical data may exhibit varying characteristics depending on the time of day or the patient's condition. In such cases, data collected at different times of the day can be viewed as subsegments, while data collected across multiple days form larger segments. This segmentation is crucial for medical practitioners, who use short-term data to make immediate decisions, while longer-term data are used for overall health evaluations.

The motivation for this paper stems from a real-life case, involving the monitoring of voice characteristics in patients diagnosed with Bipolar Disorder - a serious mental illness (see Section 2 for more details). In analyzing such data, psychiatrists aim to assess a patient's mental health status. Observational data are derived from recorded smartphone calls, and the daily aggregation of these calls is used to evaluate the patient's condition for that day. This approach, originally proposed by Hryniewicz and Kaczmarek-Majer (2021), involves treating the data as a process composed of aggregated daily segments.

When monitoring a complex segmented process, it is typically necessary to represent the data using summary statistics such as averages or medians. Hryniewicz and Kaczmarek-Majer (2021, 2023) described these aggregated segments using possibility distributions (i.e., fuzzy numbers). They then derived representative values from these fuzzy numbers for the purpose of use in control charts. However, from a practical standpoint, this approach is limited by the fact that possibility distributions and their representative values lack a probabilistic interpretation that is easily understood by practitioners. Consequently, their effectiveness - even in theory - can only be validated through simulation studies, as demonstrated in Hryniewicz and Kaczmarek-Majer (2023). To overcome this limitation, we propose in this paper a new alternative aggregation method, grounded in the probabilistic concept of p-boxes.

While the aggregation of complex segmented processes was briefly discussed in a previous conference paper by Hryniewicz and Kaczmarek-Majer (2024), the focus of this paper is not on the aggregation method itself. Instead, the main objective is to propose effective methods for monitoring such processes using fuzzy control charts. We introduce and compare four types of control charts: one based on probabilistically aggregated data and three novel charts, founded on p-box-based aggregation. The key contribution of this work lies in the fuzzy representation of the aggregated complex process. The statistical

performance of the proposed fuzzy control charts is evaluated through extensive simulations across various data-generation scenarios. The results confirm the practical applicability of the proposed methodology.

The remainder of the paper is organized as follows. Section 2 presents the practical motivation behind this research and a brief review of the literature related to the problems considered. Section 3 introduces a mathematical model for the considered class of complex segmented processes. Section 4 describes two proposed aggregation methods: one based on a probabilistic approach and the other using a p-box-based approach. In Section 5, we introduce a monitoring procedure using fuzzy control charts, focused on evaluating the stability of the process mean. Section 6 reports the results of simulation experiments assessing the statistical properties of the proposed fuzzy control charts. Finally, Section 7 presents the conclusions.

2. Practical motivations of the research and related work

As mentioned in the Introduction, the motivation for this paper stems from a real-life case, involving the monitoring of voice characteristics in patients diagnosed with Bipolar Disorder (BD). In the first subsection of this section, we present important background information on this serious psychiatric illness and describe the real-life data observed for monitored BD patients. The mathematical model of data generation proposed in this paper is designed to describe data of this type.

The subject of this paper relates to several areas of data analysis, including granular computing and statistical process monitoring. Therefore, the second subsection presents a brief overview of the literature related to the problems considered in this work.

2.1. Practical motivations

Bipolar disorder (BD) is a mental illness, affecting more than 2% of the world's population, see Grande et al. (2006). BD is characterized by episodic fluctuations (shifts) between mood phases, ranging from depression, through mixed or euthymic states, to manic episodes. It is a chronic and recurrent disease that has a serious impact on psychosocial functioning, cognition, and quality of life, see Catala-Lopez et al. (2013). For example, the suicide rate among psychiatric patients suffering from BD is the highest from among the patients of all mental disorders, see Chen and Dilsaver (1996). Therefore, early detection of the onset of illness episodes - so-called affective states - is crucial. Moreover, early detection and timely treatment significantly reduce symptom severity, improve patient well-being, and lower treatment costs.

Rapid growth and widespread adoption of smartphone technology enable real-time patient monitoring in naturalistic settings through self-monitoring as well as automatically collected objective data, such as speech and behavioral activity (see, e.g., Faurholt-Jepsen et al., 2014). Consequently, smartphone applications that monitor patients' behavior, social interactions, voice signals, movement, or location changes can be used to detect health status. Moreover, objective data, collected via smartphones are correlated, for instance, with scores from depression (e.g., HDRS) and mania (e.g., YMRS) scales (Grünerbl et al., 2015) and may therefore serve as state markers for monitoring the phases of illness in patients with BD. Recent research (Faurholt-Jepsen et al., 2014; Grünerbl et al., 2015) confirms that smartphones are becoming increasingly effective tools for assessing affective states and detecting phase changes in BD patients. However, detecting changes in a patient's state using smartphone-based monitoring remains challenging, as the patient's condition is inferred from indirectly observed data (e.g., voice characteristics) and must ultimately be confirmed by psychiatrists during clinical visits.

Hryniewicz, Kaczmarek-Majer and Opara (2019) investigated the monitoring of health states in patients with BD using behavioral characteristics, such as average daily number of phone calls, average call duration, or average number of SMS messages. These characteristics were used as input attributes for a classifier, predicting the patient's current health state, constructed using a semi-supervised learning approach. However, when the proportion of properly labeled data was small - as in their concrete observational study - the resulting classifiers were unreliable. Consequently, an alternative approach, based on the analysis of voice characteristics, was proposed to better discriminate between different health states of BD patients.

In this approach, the study participants received a dedicated mobile application, called *BDMon*, which collected a wide range of features about patient's voice during phone calls, e.g., time-domain descriptors (zero crossing rate, amplitude statistics, signal energy), and spectral features (distribution of energy, mel-cepstral coefficients, fundamental frequency and its harmonics). The voice signal of each phone call was divided into short, 10-20 ms, frames, next, with the use of an adopted version of the *openSMILE* library, see Eyben et al. (2013), the acoustic descriptors were extracted from each frame. The observations of the acoustic parameter, describing zero crossing rate (the zero crossing rate, *zcr*, is the rate, at which a signal changes its sign during the frame; it denotes the number of times the signal changes value, from positive to negative and vice versa, divided by the total length of the frame) gave the main motivation for the research described in this paper.

An example of real-life data describing zero-crossing rate observations calculated for recorded smartphone calls of BD patients was analyzed by Hryniewicz

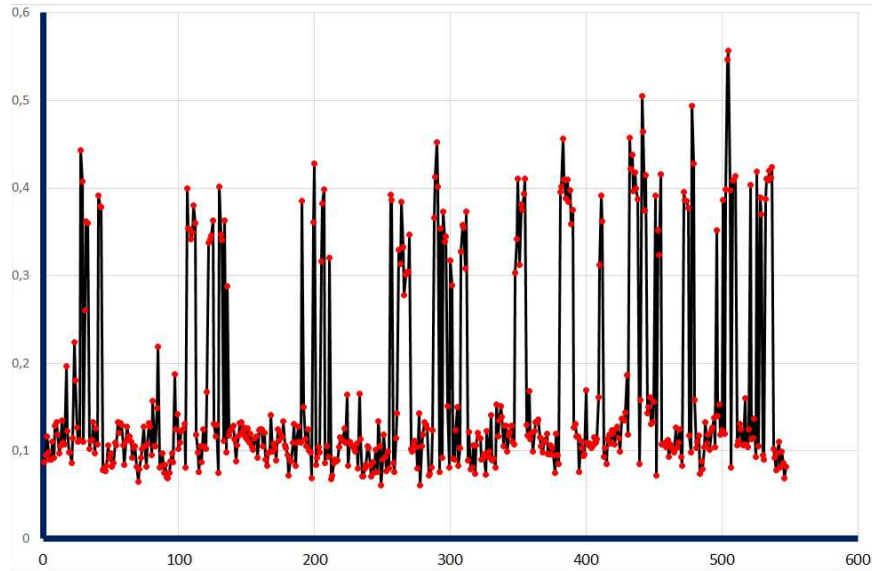


Figure 1. BD calls data - average values of the voice *zcr* characteristic

and Kaczmarek-Majer (2021). An example of such process data is presented in Fig. 1.

The process, shown in Fig. 1, appears highly inhomogeneous, even though physicians considered the patient's mental state to be stable. Moreover, histograms of voice characteristic measurements recorded for different smartphone calls exhibit significantly different probability distributions. Table 1 presents an example of real-life data collected from five recorded calls of a single BD patient during a single day.

Table 1. Measurement data (histograms) of the *zcr* voice characteristic

call-1	1260	794	114	38	16	18	16	8
call-2	525	558	117	30	20	16	2	0
call-3	25	390	1021	1308	964	748	420	48
call-4	7	141	448	412	282	232	145	12
call-5	872	602	85	26	24	11	1	0

Analysis of the observations, presented in Fig. 1 and Table 1, indicates that processes describing voice characteristics of BD patients cannot be adequately modeled using classical time-series analysis methods. This motivated

Hryniewicz and Kaczmarek-Majer (2021) to aggregate subsegment data (individual calls) within each segment (day) and to treat the process as a sequence of aggregated observations. For this purpose, a possibilistic aggregation approach was proposed, in which histograms representing subsegment data were aggregated into possibility distributions. Characteristics of these distributions were subsequently used for process monitoring, see Hryniewicz and Kaczmarek-Majer (2023). Although the results demonstrated the effectiveness of this approach using classical SPC methods (e.g., Shewhart and CUSUM charts), it was criticized by researchers favoring traditional statistical interpretations. This criticism provided additional motivation for the new approach, proposed in this paper.

2.2. Related work

Classical methods for time-series analysis are well documented in the literature and are primarily designed for homogeneous data. However, analyzing complex data requires more advanced techniques. These methods generally fall into two categories: nonlinear and non-stationary time-series analysis, Priestley (1988), and machine learning methods for data streams. The former are often mathematically rigorous, but difficult to apply in practice (see, e.g., Kuznetsov and Mohri, 2015). The latter have gained significant attention in recent years, particularly in the context of Big Data; relevant references include Gama (2010) and Sayed-Mouchaweh (2019). See also the recent book by Rutkowski, Jaworski and Duda (2020).

Fuzzy set theory has been applied to time-series analysis, too, primarily for short series or real-time prediction using Fuzzy Evolving Systems (FES), see Angelov and Zhou (2006), Lughofer (2011), and Geng and Zheng (2020). However, these approaches have rarely been used for process monitoring, due to their limited explainability, which is particularly critical in medical applications.

The real-life example motivating this research involves a naturally segmented process (days as segments, phone calls as subsegments). The methods proposed here may also be applicable to other segmented processes, particularly in monitoring scenarios, where change-point detection and dimensionality reduction are essential. Extensive literature on time-series segmentation exists, with early algorithms proposed over thirty years ago (Chu, 1995; Liu, Lin and Wanh, 2008) and more recent advances reported in Wang, Zhu and Jia (2020), Silva et al. (2023), He et al. (2024), Wang et al. (2024) and Wang and Wang (2026).

In previous work, we modeled complex processes using possibility distributions and monitored their characteristics with control charts. However, these characteristics often fail to meet the statistical assumptions of classical control charts. Although bootstrap control charts were proposed as an alternative (Hryniewicz and Kaczmarek-Majer, 2024b) they do not fully capture data im-

precision. Therefore, in this paper, we propose fuzzy control charts as a generalization of those introduced in Grzegorzewski and Hryniewicz (2000), along with other fuzzy control chart variants discussed later.

Early fuzzy control charts were developed in the 1990s for linguistically described quality data (Wang and Raz, 1990; Kanagawa, Tamaki and Ohta, 1993). Since then, numerous fuzzy control charts for attribute (i.e. discrete) and variable (real valued) data have been proposed, including fuzzy versions of Shewhart, CUSUM, EWMA, and bootstrap charts (see Grzegorzewski, 1997; Grzegorzewski and Hryniewicz, 2000; Cheng, 2005; Gülbay and Kahraman, 2007; Şenturk and Erginel, 2009; Shu and Wu, 2011; Kahraman and Bolturk, 2016; Wang and Hryniewicz, 2013, 2015; Erginel and Şenturk, 2016; Hesamian, Akbari and Ranjbar, 2019; Hryniewicz and Kaczmarek-Majer, 2024b; Razali et al., 2025). Applications of fuzzy control charts in medical contexts have also been reported, notably by Aslam et al. (2024, 2025).

3. Mathematical model of data generation

Let us assume that we observe a stream of numerical data, consisting of a large number of observations. This stream is divided into segments, and each segment is further subdivided into subsegments. We assume that this segmentation structure may either be known in advance or determined unequivocally using independent statistical or machine learning methods. In a general model of such a process we may assume that data belonging to the same subsegment are described by a certain marginal distribution and may be interdependent with this dependency characterized by a certain copula. We may also assume that the averages of subsegments belonging to the same segment are also interdependent. Moreover, we may assume that consecutive segments are also interdependent. As a matter of fact, the real-life data, representing voice characteristics of BD patients, may be described by this very general model. The identification of this model is extremely difficult, and for the purpose of data analysis we shall assume some simplifications of this model.

The first simplification pertains to the assumption that data observed within a subsegment follow a multivariate distribution with identical marginal distributions and a simplified dependence structure. In this model, the dependence between two consecutive interdependent observations is characterized by a specific bivariate copula $C(u, v)$, following the standard copula notation, see Nelsen (2006). For instance, if observations within a subsegment follow a normal distribution and the copula $C(u, v)$ is Gaussian, then the subsegment can be modeled using a simplified multivariate normal distribution. Additional simplifications have been introduced to model dependencies between subsegments and between segments. Dependencies among subsegments within the same segment are also

modeled using the aforementioned copula-based approach. For dependencies between consecutive segments, we additionally employ a classical first-order autoregressive model, $AR(p)$.

Assume now that the observed process consists of n segments, consisting of $n_i, i = 1, \dots, n$ subsegments. Moreover, we assume that each of those subsegments consists of $n_{i,j}, i = 1, \dots, n, j = 1, \dots, n_i$ observations. Then, we will assume that each of these observations is represented by the additive model

$$X_{i,j,k} = Y_{i,j,k} + Z_{i,j} + W_i, i = 1, \dots, n, j = 1, \dots, n_i, k = 1, \dots, n_{i,j}, \quad (1)$$

where $Y_{i,j,k}$ is the random component responsible for the interdependence within a subsegment, $Z_{i,j}$ is a random component responsible for the subsegment interdependence within the i th segment, and W_i is a random component responsible for the interdependence between n observed segments. Note that the model (1) describes three sources, mentioned above, of possible autocorrelations: within a subsegment (random variable X), between averages of subsegments within a segment (random variable Z), and between averages of segments (random variable W).

In the simulation experiments, described in Section 6, we assume that the marginal distributions of random variables $Y_{i,j,k}$, $Z_{i,j}$ and W_i are known, and their dependence structures are described by the respective copulas. Moreover, we assume that the numbers of subsegments n_i and the lengths of subsegments $n_{i,j}$ may vary randomly according to given distributions.

For the purposes of data aggregation let us make another important simplification. We will assume that within a given subsegment, each of observed data points is classified into one of K classes, defined by the set of class limits $(c_0, c_1, \dots, c_K)$, and that for all considered subsegments these class limits are the same. Because the aggregation method is the same for all considered n segments, then for the reasons of clarity we will omit the index i representing segments, and denote the number of subsegments in the concrete analyzed segment by M , instead of n_i . Thus, instead of the original observations, we observe discrete data, whose numbers are $m_{j,k}, j = 1, \dots, M, k = 1, \dots, K$, where $m_j = m_{j,1} + \dots + m_{j,K}$ is the total number of observations in the j -th subsegment, index k denoting classes. Hence, the observed segment data are described by M multinomial distributions with parameters estimated from the formula $\hat{p}_{j,k} = m_{j,k}/m_j, j = 1, \dots, M, k = 1, \dots, K$.

The multinomial data, described above, can be considered as grouped data, generated by the continuous probability distributions that describe subsegment observations. If the functional form of these distributions is known, it is possible to estimate the parameters of the cumulative distribution functions (CDFs) $F_j(x)$, describing the subsegment data. However, for complex and inhomogeneous data streams, the functional forms of these distributions are rarely known.

In such cases, we can estimate the CDF values at known class limits using the following formulas:

$$\hat{F}_j(c_0) = 0$$

and

$$\hat{F}_j(c_k) = \sum_{i=1}^k p_{j,i}, j = 1, \dots, M, k = 1, \dots, K. \quad (2)$$

In Table 1, we have presented an example of real-life data collected from five recorded calls of a Bipolar Disorder (BD) patient during a single day. For this type of data, we can compute the empirical probability density functions (epdf) $\hat{p}_{j,k}, j = 1, \dots, M, k = 1, \dots, K$, based on the respective histograms. This allows us to estimate the values of an unknown probability distribution only at discrete points, corresponding to the class limits.

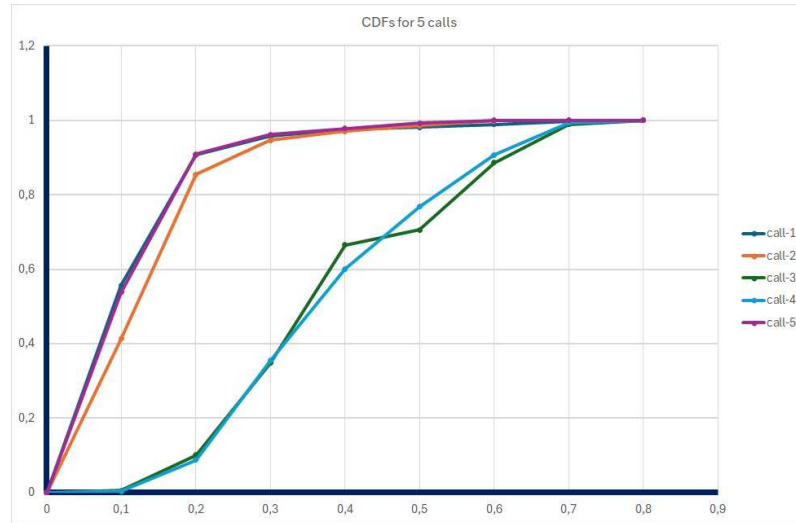


Figure 2. CDF's of the recorded *zcr* voice characteristic

In practice, however, we often approximate this distribution using, for example, a piecewise linear function. This approximation assumes that unobserved data within each class are uniformly distributed. The respective cumulative probability functions, calculated under this assumption for the data in Table 1, are shown in Fig. 2.

Looking at Table 1 and Fig. 2, we can observe that the recorded calls vary in length, and the probability distributions of two calls differ significantly from

those of the others. This confirms our claim that this observed real-life process is quite complex and cannot be easily described by a simple mathematical model.

4. Aggregation of the considered segmented processes

4.1. The setting

In many practical applications, the segmented stream of data is interpretable at the level of individual segments. Therefore, there is a need to aggregate all data from subsegments belonging to the same segment and to find its appropriate representation. Hryniewicz and Kaczmarek-Majer (2021) considered two approaches to this aggregation problem. First, we proposed a simple probabilistic method that leads to the construction of a probability distribution, representing a given segment. Second, they proposed a possibilistic method that describes all segment data using a possibility distribution, defined by the convex and normalized upper envelope of all histograms, belonging to the considered segment. For such possibility distributions, we computed certain characteristic values and used these values for process monitoring. In this section, we extend our previous analysis and introduce a new method, dedicated to the use of p-boxes (probability boxes). The proposed methods will be illustrated using an artificial dataset, shown in Table 2.

Table 2. Multinomial data describing 5 subsegments

	Class-1	Class-2	Class-3	Class-4	Class-5	Sum
ssg-1	200	20	40	10	10	280
ssg-2	100	50	100	20	30	300
ssg-3	50	80	120	130	150	430
ssg-4	30	30	90	20	120	290
ssg-5	20	20	50	20	290	400
Sum	400	200	400	200	500	1700

These data describe 5 subsegments of varying lengths, described by probability distributions of different shapes. Moreover, we assume that these data represent continuous data grouped into 5 classes, with class limits at 0.0, 0.2, 0.4, 0.6, 0.8, 1.0, respectively. These segments will be aggregated using the two methods considered in this paper.

4.2. Probabilistic aggregation of segmented processes

The statistical analysis of grouped data is a special case of the analysis of interval-valued data. This subject has attracted the attention of statisticians for

nearly one hundred years. Statisticians have sought to find good estimates for population characteristics, such as the average, variance, skewness, and kurtosis. For example, the well-known Sheppard's corrections are widely used when class midpoints are employed in the estimation of data characteristics. These corrections are optimal when the underlying continuous data are normally distributed, but they should not be used for some other types of probability distributions (e.g., the uniform distribution). A comprehensive review of the basic methods for the statistical analysis of grouped data can be found in the paper by Heitjan (1989). In general, effective statistical methods for dealing with grouped data have been proposed when the type of the probability distribution that describes the underlying continuous data is known (see, e.g., the book by Kulldorff, 1961).

Hryniewicz and Kaczmarek-Majer (2021) used the same interpretation of grouped data and proposed combining subsegment empirical probability density functions (epdfs) by calculating their weighted sum, with the weights equal to their respective subsegment sizes divided by the total number of observations in the whole segment. Simple arithmetic shows that this approach is equivalent to constructing a combined histogram $(m_1^0, \dots, m_K^0)$ for each segment, where $m_k^0 = m_{1,k} + \dots + m_{M,k}$ for $k = 1, \dots, K$. The total number of segment observations is equal to $N = m_1^0 + \dots + m_K^0$. Thus, the aggregated histogram, obtained in this way, follows a multinomial distribution $MN(p_1^0, \dots, p_K^0; K)$, described by the estimated class probabilities p_k^0 for $k = 1, \dots, K$, where $p_k^0 = m_k^0/N$ for $k = 1, \dots, K$, and the total number of segment observations N .

For the purposes of process monitoring, which will be discussed in the next section, the results of data aggregation should be summarized using statistics such as the average or standard deviation. In practically all textbooks dedicated to applied statistics, when the type of the underlying continuous data distribution is unknown, it is assumed that unobserved data within each class are uniformly distributed. This approach is called the "equiprobability model of intervals" (see Ferson et al., 2007). A list of important references dedicated to this model is provided in the same reference. Thus, most practitioners assume that uncertain (e.g., grouped) data are represented by the midpoints of their respective intervals. For grouped data, the midpoints are calculated using the formula $x_l^m = (c_{l-1} + c_l)/2$ for $l = 1, \dots, K$. The first two moments of the data distribution are calculated using the following formulas:

$$E_{mp} = \sum_{l=1}^K x_l^m p_l^0, \quad (3)$$

$$E_{mp}^{(2)} = \sum_{l=1}^K (x_l^m)^2 p_l^0. \quad (4)$$

Hence, the distribution's average is equal to E_{mp} , the variance V_{mp} is equal to $E_{mp}^{(2)} - (E_{mp})^2$, and the standard deviation σ_{mp} is equal to $\sqrt{V_{mp}}$. For the exemplary data from Table 2, where aggregated multinomial data are presented in the bottom row, we have $E_{mp} = 0.5271$, $V_{mp} = 0.0772$, and $\sigma_{mp} = 0.2778$.

When accounting for data uncertainty, related to the effect of grouping, the lower and upper limits of the data average can be calculated. Let us introduce the concept of grouped data configuration. This configuration is defined by the set of points, located in all K classes, where the actual (non-observed) data values are located. Now, let us define two boundary configurations: minitive (where all actual data values are located at the lower class limits) and maxitive (where all actual data values are located at the upper class limits). The lower limit of the expected value (average) corresponds to the minitive configuration, and is calculated as:

$$E_{lp} = \sum_{l=1}^K c_{l-1} p_l^0. \quad (5)$$

On the other hand, the upper limit of the expected value (average) is related to the maxitive configuration and is calculated from the formula:

$$E_{up} = \sum_{l=1}^K c_l p_l^0. \quad (6)$$

Note that by shifting data points, we only shift the location of the data distribution, not its shape, which is described by statistical characteristics such as standard deviation, skewness, and kurtosis. When all considered classes have the same width, say h , we have $E_{up} - E_{lp} = h$, and the interval (E_{lp}, E_{up}) is symmetrically centered around E_{mp} . However, when the class widths are different, this property does not hold. In our exemplary data, where the widths of all classes are equal to 0.2, we have $E_{lp} = 0.4271$ and $E_{up} = 0.6271$.

Regarding calculation of the data variance, taking into account data uncertainty, the situation becomes much more complicated. Intuitively, we may assume that the maximum value of the variance is associated with a configuration, in which actual data values for certain classes are located at the lower limits and for others at the upper limits. The optimal (maximal) variance is obtained by considering different configurations for each grouped observation. Formally, we define the data configuration as the set $(z_1, \dots, z_K)$, where $z_k \in \{c_{k-1}, c_k\}$ for $k = 1, \dots, K$. Denote by $v(z_1, \dots, z_K; p_1, \dots, p_K)$ the function used to calculate the data variance from grouped data. Then, the optimal data configuration is expressed as:

$$(z_1^*, \dots, z_K^*) = \arg \max_{z_1, \dots, z_K} v(z_1, \dots, z_K; p_1, \dots, p_K). \quad (7)$$

Ferson et al. (2002) shows that calculating the maximal variance for interval-valued data is an NP-hard problem. In the case of probabilistically aggregated data, where the total number of observations is N , the number of possible data configurations is 2^N , making the calculation of the maximal variance usually infeasible (note that in our example, $N = 1700$), even with the approximations proposed in various papers.

Let us consider two lower approximations for the maximal variance. First, suppose that for every subsegment, there exists one out of 2^K possible data configurations describing all the observations in this subsegment. Since there are M subsegments, the total number of possible configurations equals 2^{MK} . When M and K are small, the calculation of the maximal variance, using, for example, the algorithms proposed for the analysis of the so-called credal sets (see Montes and Miranda, 2018), may be feasible.

For larger values of M and K , we propose a weaker approximation, assuming that the same data configuration applies to all observations within the same segment. In this case, the number of possible configurations in the search for the optimal solution is 2^K , and it is possible to find the optimal configuration for moderate values of K through complete search. Note that by using this assumption, we are seeking the optimal configuration for the probabilistically aggregated data. For our exemplary data, such an optimal configuration is 0.0, 0.2, 0.4, 0.8, 1.0 (i.e., the lower class limits for the first three classes and the upper class limits for the last two classes). For this configuration, we have $V_{max} = 0.1337$ and $\sigma_{max} = 0.3656$. The data average value for this configuration is 0.5082. Note that the variability of the data, calculated for the exemplary grouped observations, is significantly greater than the variability calculated using data midpoints, and the respective average is not the same as E_{mp} .

4.3. Data aggregation of segmented processes using p-boxes

The need to address imprecisely known information, such as imprecise measurement data, has been recognized for many years. The first approach, which is purely probabilistic, was promoted by statisticians, who proposed solving the problem within the framework of Bayesian statistics. In the second approach, mainly (but not exclusively) advocated by specialists in computer science, new mathematical models of uncertainty, alternative to classical probability theory, were proposed. Destercke, Dubois and Chojnacki (2008) discuss the mutual relationships between different models of uncertainty in their two papers. According to them, dealing with the most general models presents significant challenges. Therefore, less general and simpler models, such as possibilities (see Dubois and Prade, 1988) or p-boxes, introduced by Ferson et al. (2003), are commonly used in practice.

The concept of p-boxes was introduced by Ferson et al. (2003). Ferson et al. (2003, 2007) list several practical problems that can be addressed using this concept, including issues related to non-stationarity (non-constant distributions) and interval-censoring. Given the data model, which is considered in this paper and the representation of measured data through histograms, we can assume that p-boxes are an appropriate model for analyzing this type of data.

Ferson et al. (2003) defined a p-box as follows: Suppose $\bar{F}$ and $\underline{F}$ are non-decreasing functions from the real line $\mathcal{R}$ to $[0, 1]$, with $[\underline{F} \leq \bar{F}]$ for all $x \in \mathcal{R}$. Let $[\underline{F}, \bar{F}]$ denote the set of all non-decreasing functions F from $\mathcal{R}$ to $[0, 1]$ such that $\underline{F}(x) \leq F(x) \leq \bar{F}(x)$ for all x . When the functions $\underline{F}$ and $\bar{F}$ circumscribe an imprecisely known probability distribution, we refer to $[\bar{F}, \underline{F}]$ - defined by this pair of functions - as a ‘probability box’ or a ‘p-box’ for that distribution. This implies that if $[\bar{F}, \underline{F}]$ is a p-box for a random variable X , whose distribution F is unknown but constrained within the p-box, then $\underline{F}(x)$ represents a lower bound on $F(x)$, which is the (imprecisely known) probability that X is less than x . Similarly, $\bar{F}(x)$ is an upper bound on the same probability.

The original concept of p-boxes was later generalized by other authors and related to various models of uncertainty (see Destercke and Dubois, 2006, 2009; Destercke, Dubois and Chojnacki, 2008; and Troffaes and Destercke, 2011). For instance, Destercke, Dubois and Chojnacki (2008) presented an equivalent definition of a p-box using the general concept of credal sets. They also generalized the classical concept of p-boxes by considering a broader space X (the classical p-boxes are defined on a subset of the real line). This generalization enabled the analysis of the relationships between p-boxes and other representations of imprecise information, such as random sets, probability intervals, and possibilities.

Now, let us consider the aggregation of grouped data from n subsegments that belong to the same segment. Then, the estimated cdf’s for grouped data, calculated according to (2) may be denoted as $(\hat{F}_j(c_0), \dots, \hat{F}_j(c_K)), j = 1, \dots, M$. Using these values we can estimate the upper and lower bounds of the p-box that describes the data belonging to the considered segment

$$\bar{\hat{F}}(c_k) = \max_j \hat{F}_j(c_k), k = 0, \dots, K \quad (8)$$

$$\underline{\hat{F}}(c_k) = \min_j \hat{F}_j(c_k), k = 0, \dots, K. \quad (9)$$

To illustrate this problem let us consider the values of the cdf’s calculated for the artificial multinomial data presented in Table 2. First five rows of Table 3 represent the cdf’s values, calculated at the class limits, for five considered segments. The data displayed in the last three rows represent the aggregated data set obtained using a probabilistic model of aggregation, considered in Hryniewicz

and Kaczmarek-Majer (2021), upper p-box bound $\overline{F}$ and lower p-box bound $\underline{F}$, respectively.

Table 3. Cdf values of data describing 5 subsegments

	Class-1	Class-2	Class-3	Class-4	Class-5
ssg-1	0.5	0.75	0.875	0.95	1.0
ssg-2	0.1	0.35	0.75	0.9	1.0
ssg-3	0.1	0.35	0.65	0.875	1.0
ssg-4	0.05	0.15	0.8	0.9	1.0
ssg-5	0.02	0.08	0.18	0.42	1.0
Prob.aggr.	0.165	0.341	0.594	0.765	1.0
$\overline{F}$	0.5	0.75	0.875	0.95	1.0
$\underline{F}$	0.02	0.08	0.18	0.42	1.0

As we have already mentioned, p-boxes are related to credal sets. Below, we represent the p-box described by the data presented in Table 3 using the concept of credal sets. The respective credal set is now described by Table 4.

Table 4. Credal set representation of the p-box for segment data

Focal.elm	prob.mass	low.limit	up.limit	L.exp	U.exp
Felm-1	0.02	0.	0.2	0.	0.004
Felm-2	0.06	0.	0.4	0.	0.024
Felm-3	0.1	0.	0.6	0.	0.06
Felm-4	0.24	0.	0.8	0.	0.192
Felm-5	0.08	0.	1.	0.	0.08
Felm-6	0.25	0.2	1.	0.05	0.25
Felm-7	0.125	0.4	1.	0.05	0.125
Felm-8	0.075	0.6	1.	0.045	0.075
Felm-9	0.05	0.8	1.	0.04	0.05

The graphical representation of this p-box in terms of a credal set is given, following Montes (2021), in Fig. 3.

The rectangles between the dotted lines represent the focal elements of this credal set with respective probabilities equal to the heights of these rectangles.

When we look for real-life applications of p-boxes, we usually need to calculate expectations of a random variable X , characterized by the p-box $(\overline{F}(x), \underline{F}(x))$ with respect to a certain real-valued function $h(x)$. In general, this is a very difficult task, but it can be solved relatively easily using some approximations. Utkin and Destercke (2009) approximate the distribution function $F(x)$ by the set of their N values $F(x_i), i = 1, \dots, N$. Using this approximation, they find

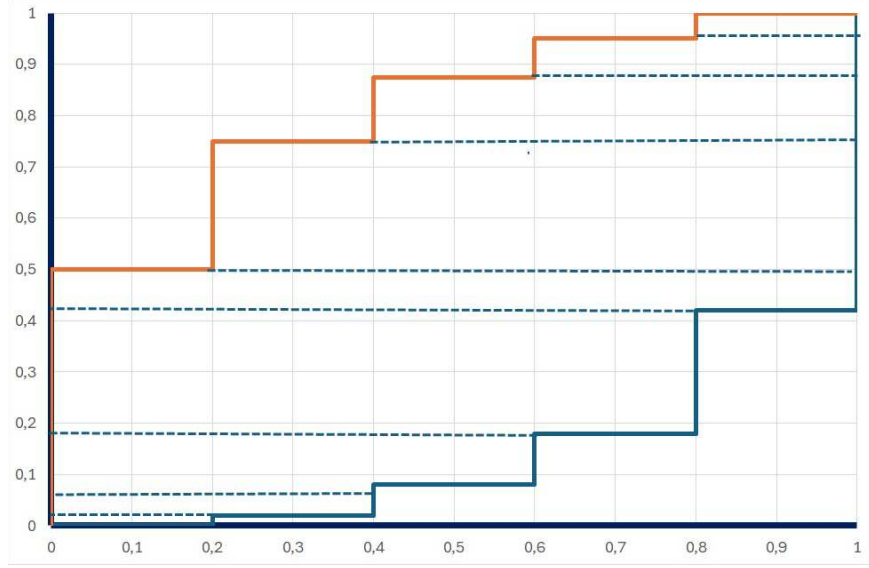


Figure 3. Representation of P-box using credal sets

the approximate upper (lower) expectation of the random variable X by solving two linear programming problems (one for maximization and one for minimization). A simpler solution can be found if $h(x)$ is a monotone non-decreasing (non-increasing) function. In this case, the upper and lower expectations are completely determined by $\bar{F}(x)$ and $\underline{F}(x)$ functions. However, the functional form of these functions is usually unknown, and even in this special case, we have to use their approximations.

When $h(x) = x$, i.e. when we want to calculate the classical expected value of X , and our data are described by $\bar{F}(x)$ and $\underline{F}(x)$ functions, the lower and upper expectations are calculated from the following formulas (see Ferson et al., 2007, for a more general case of interval-valued data),

$$\underline{E} = \sum_{i=1}^m c_{i-1} * (\bar{F}(c_i) - \bar{F}(c_{i-1})). \quad (10)$$

$$\bar{E} = \sum_{i=1}^m c_i * (\underline{F}(c_i) - \underline{F}(c_{i-1})). \quad (11)$$

The calculation of the lower and upper variances of the p-box is not so simple, but the respective algorithm can also be found in Ferson et al. (2007).

5. Monitoring of segmented processes using their representations by fuzzy triangular numbers

Control charts are simple tools for the monitoring of process stability using representative values (such as average, median, variance, etc.) that have probabilistic interpretation. In the case of the probabilistic aggregation, for the simplest representation of the aggregated data we can take the average E_{mp} , calculated according to (3), and use this average for monitoring consecutive data segments using any control chart. However, when we want to take into account uncertainty resulting from data grouping, we can also calculate the lower possible value of the average E_{lp} , calculated according to (5) and the upper possible value of the average E_{up} , calculated according to (6). Then, we can use the approach proposed in Hryniewicz and Kaczmarek-Majer (2024a) and represent the aggregated segment average by the *fuzzy triangular number* (E_{lp}, E_{mp}, E_{up}) .

5.1. Representation of uncertain data described by p-boxes for the purpose of process monitoring

If we use the concept of p-boxes for aggregation purposes, we can employ a similar approach. When we want to take into account uncertainty resulting not only from data grouping, but also resulting from the variety of probability distributions, representing aggregated subsegments we can calculate the lower possible value of the average $\underline{E}$ and the upper possible value of the average $\bar{E}$, calculated according to (10) and (11), respectively. These averages represent, respectively, the lowermost and the uppermost values of the segment average.

The p-boxes, considered in this paper, are rather focused on the description of borderline extremes, and not on their impact on an average process level. Therefore, we need a variable that can provide additional information about a certain ‘average’ of the process level. For this purpose, we can use the probabilistic aggregation method and take E_{mp} as the most plausible value of the segment average. Therefore, we can use the fuzzy triangular number $(\underline{E}, E_{mp}, \bar{E})$ as a more realistic representation of the uncertain segment average.

Hryniewicz and Kaczmarek-Majer (2024a) proposed an even narrower representation of the aggregated segment average. They propose a different method for the construction of approximate p-box expectations for grouped data. They assume that instead of discrete-valued cdfs we should consider the following piecewise-linear cdfs

$$\widehat{F}^*(x) = \widehat{F}(c_{l-1}) + \frac{x - \widehat{F}(c_{l-1})}{\widehat{F}(c_l) - \widehat{F}(c_{l-1})}, l = 1, \dots, K, \quad (12)$$

$$\hat{\underline{F}}^*(x) = \hat{\underline{F}}^*(c_{l-1}) + \frac{x - \hat{\underline{F}}(c_{l-1})}{\hat{\underline{F}}(c_l) - \hat{\underline{F}}(c_{l-1})}, l = 0, \dots, K, \quad (13)$$

where functions $\bar{\hat{F}}(x)$ and $\hat{\underline{F}}(x)$ are given by (8) - (9), respectively. Note, that this proposal is based on the assumption that the random variables represented, respectively, by the cdfs $\bar{\hat{F}}(x)$ and $\hat{\underline{F}}(x)$, are uniformly distributed between the consecutive class limits $c_0, c_1, \dots, c_K$.

Thus, the expected values of random variables distributed according to (12) - (13) can be calculated as follows

$$\underline{E}^* = \sum_{i=1}^K \frac{c_{i-1} + c_i}{2} * (\bar{\hat{F}}(c_i) - \bar{\hat{F}}(c_{i-1})) \quad (14)$$

$$\bar{E}^* = \sum_{i=1}^K \frac{c_{i-1} + c_i}{2} * (\hat{\underline{F}}(c_i) - \hat{\underline{F}}(c_{i-1})). \quad (15)$$

Using this approach we can represent the uncertain aggregated segment average by the fuzzy triangular number $(\underline{E}^*, E_{mp}, \bar{E}^*)$.

In the description of the p-box using a triangular fuzzy number one can consider possibility to replace E_{mp} with another value representing the p-box average, such as, e.g., the average calculated directly from the equivalent credal set (see Table 4 and Fig. 3). However, in this paper we will not consider this possibility.

5.2. Design of a control chart for fuzzy data

In Statistical Process Control (SPC), control charts are used to generate alarm signals that the probability distribution describing process observations has been changed. In process monitoring, we usually monitor either central tendency (the so-called ‘process level’) or variability. An alarm signal is generated when the value of a monitored statistic (e.g., an average), plotted on a control chart, falls beyond a certain control limit. Consider, for example, a classical Shewhart control chart applied for the case when alarm signals should be generated when the process level has been increased. In this case, we compare the value of the observed statistic, say z , with the one-sided upper control limit defined as

$$UCL = \bar{z} + k\sigma(\bar{z}), \quad (16)$$

where $\bar{z}$ is the average of n_S values $z_1, z_2, \dots, z_{n_S}$ observed while the monitored process is in a stable state, $\sigma(\bar{z})$ is the sample standard deviation of $\bar{z}$ and k is a certain coefficient, usually equal to 3 (when $k = 3$ and observed values are independent and normally distributed, then the probability of a false alarm is very low, equal to 0.0027).

In contrast to classical monitoring procedures, our data, described in Section 5.1, are fuzzy. Therefore, for monitoring of data of this type we may use *fuzzy control charts*, described, e.g., in the papers mentioned in Section 2.2. When the observed data are uncertain and describe the aggregated segments, we consider in this paper three types of their representation. When we use the probabilistic aggregation the data are represented by triples (E_{lp}, E_{mp}, E_{up}) . When for data aggregation we use the concept of p-boxes, in this paper we propose to use two representations: $(\underline{E}, E_{mp}, \bar{E})$, which for computations uses class limits, and $(\underline{E}^*, E_{mp}, \bar{E}^*)$ which uses class mid-points.

Let $\tilde{z} = (z_L, z_C, z_U)$ be the generic description of an uncertain datum described by a triangular fuzzy number. Denote by $\bar{z}_L, \bar{z}_C, \bar{z}_U$ the sample averages of (z_L, z_C, z_U) and by $\sigma(\bar{z}_L), \sigma(\bar{z}_C), \sigma(\bar{z}_U)$ the standard deviations of these averages. Then, instead using one crisp control limit UCL we have three control limits. We propose to interpret these control limits as a fuzzy control limit $\tilde{Z} = (Z_L, Z_C, Z_U)$, where $Z_L = \bar{z}_L + k\sigma(\bar{z})_L$, $Z_C = \bar{z}_C + k\sigma(\bar{z})_C$, $Z_U = \bar{z}_U + k\sigma(\bar{z})_U$, respectively, and k is a certain design parameter of the chart. In the same way we can generalize one-sided lower control limits and two-sided control limits to their fuzzy counterparts.

When observed imprecise process data are described by fuzzy triangular numbers, the question arises about a method that should be used for the generation of alarm signals. Hryniewicz and Kaczmarek-Majer (2024a) proposed to generate alarm signals using the generalization of rules proposed by Grzegorzewski and Hryniewicz (2000), who proposed to calculate a crisp control limit, say C_U , and to compare fuzzy observations $\tilde{z}$ with that limit using the necessity measure of the strength of the relation $\tilde{z} > C_U$ (see: Dubois and Prade, 1989). The alarm signal is generated when the value of this measure, denoted by α is not smaller than a certain critical value $\alpha_{crt} \in (0, 1]$.

Given the above described circumstances, in the present paper we propose the use of a more general method as a mechanism of generation of alarms. In this suggested method, instead of a crisp control limit C_U , we propose to use the fuzzy control limit $\tilde{Z}$. Then, we propose to calculate the necessity measure of the strength of the relation $\tilde{z} > \tilde{Z}$, denoted by α , and the possibility measure of the strength of this relation, denoted by β . Because $\tilde{z}$ and $\tilde{Z}$ are both fuzzy triangular numbers, then the values of these measures can be calculated from simple formula,

$$\alpha = Nec(\tilde{z} > \tilde{Z}) = \begin{cases} 0 & , & z_U < Z_L \\ \frac{\Delta z}{Z_C - z_L + \Delta z} & , & z_L < Z_C < z_U \\ 1 & , & z_L \geq Z_U \end{cases} \quad (17)$$

$$\beta = Pos(\tilde{z} > \tilde{Z}) = \begin{cases} 0 & , & z_U < Z_L \\ \frac{z_U - Z_L}{z_U - Z_L + \Delta z} & , & (Z_L < z_U) \cap (z_C < Z_C) \\ 1 & , & z_C \geq Z_C \end{cases} \quad (18)$$

where $\Delta z = z_C - Z_C$. The alarm signal is generated when the value of the necessity measure α is not smaller than a certain critical value $\alpha_{crt} \in (0, 1]$ (a stronger requirement) or the possibility measure β is not smaller than a certain critical value $\beta_{crt} \in (0, 1]$ (a weaker requirement).

6. Statistical properties of the proposed control charts - simulation experiment

6.1. Introductory remarks

In SPC, the statistical properties of control charts used for process monitoring are described in many textbooks, such as Montgomery (2011). The most commonly used characteristic for this purpose is the Average Run Length (ARL). The ARL is defined as the expected (average) number of process observations between consecutive alarms. When the monitored process remains in a desirable stable state, these alarms are referred to as false alarms, and the corresponding value of ARL is usually denoted as ARL_0 . However, the main purpose of a chart is to trigger alarms when the process goes out of control, e.g., when its average value (sometimes called the process level) increases or decreases by a certain amount, typically expressed as a multiple of the standard deviation of the statistic used for monitoring. Let σ_m denote the standard deviation of the monitored statistic (e.g., the standard deviation of the sample average) calculated from all observations collected in the initial (training) Phase I. Assume that the shift in the average value of the process observation (i.e., the shift in the process level) is equal to $\delta \cdot \sigma_m$, where σ_m is the standard deviation of the average calculated for the monitored process characteristic. The average waiting time until an alarm is triggered by a shift of this magnitude is denoted as $ARL(\delta)$.

When the monitored process is described by a known probability distribution, the values of $ARL(\delta)$ can be computed relatively easily using closed formulas. However, when the distribution type is known (e.g., normal or exponential) but its parameters are estimated using data collected in Phase I, closed

formulas for calculating $ARL(\delta)$ are available only in certain simple cases, such as when the data are normally distributed. In all other cases, specially designed numerical methods have been developed to calculate the values of $ARL(\delta)$.

When monitored processes are complex and described, for example, by multiple different probability distributions, calculating $ARL(\delta)$ values, even using numerical methods, is often not feasible. In the case of the segmented processes, considered in this paper, the situation is even more challenging, as the number of subsegments in individual segments and the number of observations in each subsegment may vary randomly. This occurs, for instance, when monitoring Bipolar Disorder patients based on their voice characteristics. In such cases, a simulation experiment is the only feasible approach for evaluating the proposed monitoring procedures. In these experiments, synthetic data, that are similar to those observed in real-life experiments, are generated for various data-generation scenarios. These data are used for the qualitative evaluation of the efficiency of the proposed monitoring methods. It is important to emphasize, however, that the results of such experiments cannot be used for precise evaluation of the proposed methods, serving only as a ‘proof of concept’.

In this paper, we assume that the number of subsegments in an aggregated segment and the lengths of these subsegments vary randomly. This scenario is exactly what occurs when monitoring voice data, where the number of recorded calls (subsegments) and their duration (length of a subsegment) vary randomly.

6.2. Description of simulation experiments

In real-life complex processes the number of subsegments belonging to the same segment and their lengths, i.e. numbers of observations, may vary randomly. Consider, for example, the monitoring of Bipolar Disorder patients whose voice characteristics are calculated from their recorded smartphone calls. The number of calls in each day (i.e., the number of subsegments) and the number of the observed data points (proportional to the length of a call) are described by respective random variables. In our simulation experiments we generate these random variables using a predefined normal distribution. However, in real-life experiments the distributions of these parameters may be estimated from the available data. The observations representing subsegments of the considered complex process may be described by different probability distributions (see Table 1). In our experiments, we have considered the following five data scenarios

1. Subsegments of varying length described by the same symmetric (normal) distribution.
2. Subsegments of varying length described by symmetric (normal) distributions with randomly varying parameters.

3. Subsegments of varying length described by the same asymmetric (Weibull) distribution.
4. Subsegments of varying length described by asymmetric (Weibull) distributions with randomly varying parameters.
5. Subsegments of varying length described by a mixture of symmetric (normal) and asymmetric (Weibull) distributions with varying parameters.

These scenarios were designed with real-world situations in mind, such as those observed during the monitoring of Bipolar Disorder patients. In one of our experiments that is described in this paper we assumed the normal $N(3, 10)$ distribution for Scenario 1. For Scenario 2 we also assumed the normal distribution, but with the expected value randomly chosen from the interval $[8, 12]$ and the standard deviation randomly chosen from the interval $[2.5, 3.5]$. In Scenario 3 and Scenario 4, instead of a normal distribution we assumed the asymmetric Weibull distribution. In Scenario 3 the value of its scale parameter was set to 10, and the value of its shape parameter was set 3. In Scenario 4 both these parameters were chosen randomly: the scale parameter from the interval $[8, 12]$, and the shape parameter from the interval $[2, 4]$. Finally, in Scenario 5 the probability distribution describing subsegment data was chosen with probability $p = 0.5$ according to the method used in Scenario 2, and with probability $p = 0.5$ according to the method used in Scenario 4.

For each scenario we generated data sets of the length of 100 elements each, and grouped them into segments. For each of these segments, subsegments were aggregated using the probabilistic method. Then, for the obtained aggregated probability distributions, we estimated (using Sheppard's correction for grouped data) the values of their average, standard deviation, coefficient of variation, skewness and kurtosis. The values of these statistical characteristics, averaged over altogether 1000 segments are presented in Table 5.

Table 5. Statistical characteristics of the segment aggregated data for the scenarios assumed

	Scen1	Scen2	Scen3	Scen4	Scen5
Average	10.01	10.07	8.93	8.93	9.461
StnDev	3.05	3.26	3.37	3.56	3.428
CoVar	0.305	0.324	0.377	0.399	0.362
Skewness	0.002	0.006	0.167	0.377	0.182
Kurtosis	-0.025	0.013	-0.252	0.278	0.085

The data presented in Table 5 show that the averages for all scenarios considered in this experiment are similar, but the remaining parameters (standard deviation, coefficient of variation, skewness and kurtosis) may be significantly different, depending on the applied scenario. This suggests the use of different

monitoring procedures when the observed probability distributions have different shapes. In all our experiments we have considered four types of monitoring (charting) models

1. PRB - monitored are fuzzy characteristics (E_{lp}, E_{mp}, E_{up}) of the probabilistically aggregated data,
2. PBX - monitored are fuzzy characteristics $(\underline{E}, E_{mp}, \bar{E})$ of the aggregated data described by p-boxes,
3. PBXE - version of PBX with different definitions of fuzzy control limits,
4. PBXM - monitored are fuzzy characteristics $(\underline{E}^*, E_{mp}, \bar{E}^*)$ of the aggregated data described by p-boxes and mid-points of histogram classes.

The fuzzy control limits $\tilde{Z} = (Z_L, Z_C, Z_U)$, where $Z_L = \bar{z}_L + k\sigma(\bar{z})_L$, $Z_C = \bar{z}_C + k\sigma(\bar{z})_C$, $Z_U = \bar{z}_U + k\sigma(\bar{z})_U$, were obtained using the method presented in Section 5.2 for the fuzzy characteristics describing the monitoring models listed above.

In the case of aggregated segment data described by histograms or by complex structures like p-boxes the evaluation of $ARL(\delta)$ is much more difficult. First of all, let us notice that a shift in a process level changes not only the location of the considered probability distribution, but also its shape. Thus, shifts in the location parameters of the probability distributions describing underlying crisp subsegment data result in different, and depending on the applied method of aggregation, shifts of the characteristics describing the imprecise aggregated data.

In our experiments, shifted data were obtained by adding respective shift values to the fuzzy values of non-shifted data

1. PRB - $(E_{lp} + \delta\sigma(\bar{E}_{mp}), E_{mp} + \delta\sigma(\bar{E}_{mp}), E_{up} + \delta\sigma(\bar{E}_{mp}))$,
2. PBX - $(\underline{E} + \delta\sigma(\bar{E}), E_{mp} + \delta\sigma(\bar{E}_{mp}), \bar{E} + \delta\sigma(\bar{E}))$,
3. PBXE - $(\underline{E} + \delta\sigma(\bar{E}_{mp}), E_{mp} + \delta\sigma(\bar{E}_{mp}), \bar{E} + \delta\sigma(\bar{E}_{mp}))$,
4. PBXM - $(\underline{E}^* + \delta\sigma(\bar{E}^*), E_{mp} + \delta\sigma(\bar{E}_{mp}), \bar{E}^* + \delta\sigma(\bar{E}^*))$.

An important parameter in the simulation experiments for evaluating monitoring procedures is the number of segments in Phase I, i.e., the volume of training data. In our experiments, we considered several values of this parameter, but in this paper, we present results when this parameter is set to 25 (which is considered the minimal length for Phase I in SPC theory). It is worth noting, however, that in many practical cases (e.g., when aggregated segments represent consecutive days of the monitored process), bigger values of this parameter cannot be used, as the length of Phase I cannot be too big.

6.3. Sketch of the simulation algorithm

The operation of our simulation algorithm is divided into M_{sym} (usually 500) runs. Data generated during the first part of each run are considered as the training data, and are used for the design of a control chart. Therefore, in our experiment we have M_{sym} control charts. Then, for each of these charts we generate a sequence of the aggregated (fuzzy) representations of consecutive process segments. The generation of this sequence is terminated when a control chart produces an alarm signal. Thus, the length of this sequence is equal to the run length (RL) of the chart. For each chart this procedure is repeated n_{run} (usually 50) times, and the average value of the run lengths (ARL) is established. Finally, for the M_{sym} values of ARL's some statistical characteristics (average, median, standard deviation, etc.) are calculated. These values are used for the evaluation of performance of different types of control charts applied for data generated according to the different scenarios, defined in Section 6.2.

Algorithm 1 Control Chart Simulation

- 1: **Input:** Parameters describing data generation and control chart operation
 - 2: **for** run = 1 to 500 **do**
 - 3: **Design Control Chart**
 - 4: Generate subsegment training data (25 segments)
 - 5: Aggregate segment data
 - 6: Represent aggregated segment data as triangular fuzzy numbers
 - 7: Calculate fuzzy control chart parameters (fuzzy requirements)
 - 8: **Test Designed Control Chart**
 - 9: Generate subsegment test data
 - 10: Aggregate segment test data
 - 11: Represent aggregated segment data as triangular fuzzy numbers
 - 12: Compare fuzzy test data with fuzzy requirements
 - 13: **Establish ARL (Average Run Length)**
 - 14: Calculate ARL for the simulation run
 - 15: **end for**
 - 16: **Output:** Calculate the average chart characteristics
-

Note that each item of this sketch, given as Algorithm 1, represents one or more specialized algorithms such as, e.g., generation of random numbers, computation of p-boxes, or computations of possibility and necessity measures. Additional information about these simulation models can be found in Section 6.4.3.

6.4. Results of simulation experiments

6.4.1. The setting

The data generation model, described in Sections 6.2 and 6.3, involves many parameters that might influence statistical characteristics of monitored processes. Therefore, taking into account many feasible combinations of these parameters will result in performing extremely huge, and difficult for interpretation, simulation experiments. Thus, we have decided to split our experiments in two parts. The goal of the first part of our experiment is to answer the question as to the basic differences between the control charts, considered in this paper. In this part of our experiment we have assumed that the process data are *uncorrelated*. This assumption means that the complexity of the considered process is related only to the diversity of probability distributions that describe sub-segment data. In the second part of our simulation experiment we will try to demonstrate how possible autocorrelation of observed data may influence the operating characteristics of the proposed fuzzy control charts.

In all our experiments we used one-sided control charts with decision rules based on the possibilistic approach. The chosen critical values β_{crt} represent different practical requirements regarding the values of ARL's. For example, for correlated data (Section 6.4.3) we set $\beta_{crt} = 0.95$. For this particular critical value the rate of false alarms is higher than it is usually assumed in SPC literature, but it is fully acceptable in the context of monitoring of BD patients, where very short time to reveal a necessary alarm is required, and frequent false alarms are not costly.

6.4.2. Results for uncorrelated data

At the beginning of our experiments we tried to answer a question that is usually asked by practitioners, namely: whether it is possible to use for our simulated processes a classical Shewhart control chart (with $k = 3$, i.e., a classical '3-sigma' control limit), applied for the average value of the observed probability distribution. In other words, is it advisable to use only the value of E_{mp} that is used in all fuzzy data models considered in this paper as the most plausible value? Note that it is possible to evaluate the properties of this classical chart by taking either the critical possibility value $\beta = 1$ or critical necessity value $\alpha = 0$. The simulated values of ARL0 and ARL(1) for five considered scenarios are presented in Table 6. Additionally, in this table we present the values of the respective medians, MRL0 and MRL(1), of the run length.

It is interesting to note that in the cases when subsegments are described by the same distributions (Scenario 1 and Scenario 3) the behavior of a Shewhart control chart is similar. For the remaining three cases, the false alarm rate is

Table 6. The values of ARL for a classical Shewhart control chart

	Scen1	Scen2	Scen3	Scen4	Scen5
ARL0	759	1226	710	1038	1041
ARL(1)	52	57	53	57	52
MRL0	259	419	243	334	325
MRL(1)	24	25	26	26	26

very low (large values of ARL0 and MRL0). However, in all considered cases the values of ARL(1) and MRL(1) are too high (i.e., the waiting time for the actually needed alarm is too long). Therefore, the use of a classical Shewhart control chart for monitoring complex segmented process should not be recommended.

In the first group of our main experiments we investigated the impact of critical values of possibility (β) and necessity (α) on the performance of control charts for four considered monitoring models. We calculated not only average run lengths (ARLs), but also median run lengths (MRLs). We calculated these values for two cases: processes without a shift (ARL0, MRL0) and processes with a shift of one σ (ARL1, MRL1). In all considered cases we calculated the coefficient of *length reduction* CLR, defined as $ARL0/ARL1$ ($MRL0/MRL1$). Greater values of CLR show better discrimination between shifted and non-shifted process levels.

In Table 7 we present the simulation results for a certain realization of Scenario 5 (mixture of symmetric and asymmetric distributions). We display the values of ARL0, ARL1, and the respective coefficient of length reduction CLRA.

Table 7. Performance of monitoring models with respect to ARL and critical values of possibility β , $k = 3$

	β :	0.97	0.96	0.95	0.94	0.93
PRB	ARL0	595.2	525.6	458.6	402.9	354.8
	ARL1	38.6	34.5	30.8	27.2	24.1
	CLRA	15.4	15.2	14.9	14.8	14.7
PBX	ARL0	586.9	470.9	376.0	301.8	233.4
	ARL1	27.4	21.7	17.7	14.1	11.4
	CLRA	21.4	21.7	21.3	21.5	20.5
PBXE	ARL0	486.9	383.3	295.8	230.0	177.6
	ARL1	24.7	19.5	15.6	12.7	10.3
	CLRA	19.7	19.6	18.9	18.2	17.2
PBXM	ARL0	699.7	644.0	590.9	537.6	490.7
	ARL1	36.4	31.9	28.5	25.5	22.8
	CLRA	19.2	20.2	20.7	21.0	21.5

Table 8. Performance of monitoring models with respect to MRL and critical values of possibility β , $k = 3$

	β :	0.97	0.96	0.95	0.94	0.93
PRB	MRL0	181	155.5	133	115	100
	MRL1	17	15	14	12	11
	CLRM	10.6	10.4	9.5	9.6	9.1
PBX	MRL0	140	105	84	65	49
	MRL1	14	11	9	8	6
	CLRM	10.1	9.5	9.3	8.1	8.2
PBXE	MRL0	129	99	75	57	43
	MRL1	12	10	8	7	6
	CLRM	10.8	9.9	9.4	8.1	7.2
PBXM	MRL0	199	174.5	156.5	137	123.5
	MRL1	17	15	14	12	11
	CLRM	11.7	11.6	11.2	11.4	11.2

The probability distribution of the ARL is highly skewed (for the given constant probability of an alarm this is the geometric distribution). Therefore, some practitioners prefer to use the median of the run length (MRL). This approach is seldom used by theoreticians, as the theoretical probability distribution of the median is more difficult for calculation than the theoretical distribution of the average. However, in simulation experiments, as those described in this paper, we also use MRL characteristics for the evaluation of the proposed control charts. In Table 8 we present the analysis of the performance of monitoring models where the median values of the run length are considered.

From Tables 7 and 8 we can see that by decreasing the critical values of possibility β we also decrease the values of the respective ARL's (MRL's). Moreover, from Table 7 we see that even for the values of β close to 1 the values of ARL's are much higher than those suggested by specialist of quality control. For such large values the probability of observing a false alarm is very low, but on the other hand, the average waiting time for a needed alarm is too long. Therefore, for having practically acceptable values of the run lengths we may either decrease the value of the design parameter k or keep the recommended value of this parameter equal to 3, and use the critical possibility requirement β smaller than one.

Another way to having practically acceptable values of the run lengths is to take a lower value of k , e.g., equal to 2 (another value recommended by specialists in quality control for setting the so called control chart warning limits) and apply the necessity critical values α greater than 0. The effect of this approach for tuning a control chart is presented in Tables 9 and 10. We can

Table 9. Performance of monitoring models with respect to ARL and critical values of necessity α , $k = 2$

	α :	0.05	0.06	0.07	0.08	0.09	0.1
PRB	ARL0	128.1	151.4	178.6	215.6	255.5	302.5
	ARL1	10.3	11.3	12.5	13.9	15.3	16.9
	CLRA	12.5	13.4	14.3	15.5	16.7	17.9
PBX	ARL0	268.0	368.3	491.9	652.6	865.0	1104.8
	ARL1	16.8	21.2	27.0	33.8	42.2	54.2
	CLRA	16.0	17.3	18.2	18.3	20.5	20.4
PBXE	ARL0	279.5	378.9	518.7	676.8	872.6	1113.4
	ARL1	17.7	22.2	27.6	34.9	43.2	54.6
	CLRA	15.8	17.0	18.8	19.4	20.2	20.4
PBXM	ARL0	99.7	114.9	131.6	148.3	170.1	191.9
	ARL1	10.0	11.0	12.3	13.6	15.2	16.7
	CLRA	10.0	10.4	10.7	10.9	11.2	11.5

see that by increasing the critical value of necessity α we may observe the same final effect of the control chart tuning, i.e. the design of control chart with the recommended value of the ARL0.

A similar pattern, as that presented in Tables 7 through 10, has been observed for the remaining four considered data scenarios.

One of the goals of our simulation experiment is the comparison of different charts for different data scenarios. As we can see from Tables 7 through 10, the effectiveness of control charts, measured by the coefficient of length reduction CLR (CLRM) depends upon the value of ARL0 (MRL0). Therefore, for the purpose of comparison, we have chosen only those charts with the values of the ARL0 belonging to the interval (350.0, 390.0), i.e., the values recommended by specialists in quality control. We have compared charts designed according to four models (PRB, PBX, PBXE, PBXM) applied for the five considered data generation scenarios and two models for making decisions: ($k=3$ and the possibility critical value β) and ($k=2$ and the necessity critical value α). The required values of the ARL0 have been obtained by tuning the values of β and α , respectively. The obtained values of CLR for control charts that fulfill the requirement for the value of ARL0 are displayed in Table 11.

From Table 11 we can see that the overall discrimination rate, described by the value of CLR, is similar for all considered cases. However, certain interesting patterns can be recognized. For example, the values of CLR for the charts based on the possibility critical value β have, on average, better discrimination rate (18.2) than the charts based on the necessity critical value α

Table 10. Performance of monitoring models with respect to MRL and critical values of necessity α , $k = 2$

	α :	0.05	0.06	0.07	0.08	0.09	0.1
PRB	MRL0	48	55	64	75	86	103
	MRL1	6	7	8	8	9	10
	CLRM	8	7.9	8	9.4	9.6	10.3
PBX	MRL0	101	134	176.5	243	337	472
	MRL1	11	13	16	20	25	32
	CLRM	9.2	10.3	11.0	12.2	13.5	14.8
PBXE	MRL0	100	129	177	251	335	464.5
	MRL1	11	13	16	21	25	32
	CLRM	9.1	9.9	11.1	12.0	13.4	14.5
PBXM	MRL0	46	52	60	67	77	87.5
	ARLM	6	7	8	9	9	10
	CLRM	7.7	7.4	7.5	7.4	8.6	8.8

Table 11. Values of the coefficient of run reduction (CLR) for different scenarios

	Possibility β and $k=3$					Necessity α and $k=2$				
	PRB	PBX	PBXE	PBXM	Avg	PRB	PBX	PBXE	PBXM	Avg
Sc-1	15.9	14.2	14.9	12.8	14.5	12.1	11.9	12.4	11.6	12.0
Sc-2	22.8	21.2	25.0	27.9	24.2	19.0	12.4	17.1	14.1	17.1
Sc-3	13.1	14.6	12.6	14.8	13.8	14.6	14.1	16.6	18.2	15.9
Sc-4	16.4	18.5	17.8	24.6	19.3	15.4	15.1	16.0	14.8	15.3
Sc-5	14.7	21.3	19.6	20.5	15.7	19.0	17.3	17.0	12.4	15.6
Avg;	16.6	18.0	18.0	20.1	18.2	15.4	15.3	15.8	14.2	15.2

(15.2). More interesting results of analysis are obtained if we divide our scenarios into three clusters: symmetric distributions (Scenario 1 and Scenario 2), asymmetric distributions (Scenario 3 and Scenario 4), and the mixture of symmetric and asymmetric distributions. For the symmetric distributions the preferred (the best average CLR, and the best average ranking) control chart is the PRB chart - based on the probabilistic aggregation. On the other hand, for the asymmetric distributions the preferred (using the same criteria) control chart is the PBXM chart - based on the aggregation with the usage of p-boxes. In the third case, i.e. the mixture of symmetric and asymmetric distributions, the preferred control chart is the PBX chart - based on the aggregation with the usage of p-boxes with wider p-box limits. The proposed preference structure is easily verifiable using simple statistical methods, and thus easy to implement in practice.

Another problem that has to be solved in practice is the choice of the critical values: α or β . This can be done using a simulation experiment. Because we do not know the mathematical model of the monitored process, we cannot generate sequences of test data. Instead, we can apply bootstrapping and generate test data sequences using the method described in Hryniewicz and Kaczmarek-Majer (2024b). By repeating this simulation experiment for different values of α or β , we can find such values of these critical values that for the evaluated control charts the value of ARL0 is practically acceptable.

6.4.3. Results for correlated data

In our simulation experiments, we considered three sources of data autocorrelation: within subsegments, between subsegment averages within a single segment, and between segment averages (see Equation 1 and its description). All these autocorrelation patterns have been observed in real-world data describing the voice characteristics of patients with Bipolar Disorder. We evaluated the impact of these dependencies on the statistical performance of PBX control charts used for monitoring complex processes through extensive simulations. Specifically, we analyzed two data generation models (Scenario 2 and Scenario 5) across seven levels of dependency strength (measured using Pearson's correlation coefficient, ρ). In each simulation, we generated 500 different control charts. For each chart, we simulated 50 monitoring processes and estimated the Average Run Length (ARL), resulting in a sample of 500 ARL values per experiment.

These samples were used to calculate overall statistical characteristics (three types of averages and the median) and to conduct comparative analyses using appropriate statistical tests. Due to substantial process inhomogeneity, the samples were highly skewed and contained many outliers. Consequently, in addition to conventional means, we computed robust estimates, including winsorized and

Table 12. Impact of dependence of subsegment data

ρ	Scenario-2				Scenario-5			
	ARL	ARLW	ARLT	MRL	ARL	ARLW	ARLT	MRL
-0.9	181.1	121.0	90.7	53.7	114.2	94.2	73.9	44.0
-0.5	178.6	125.0	103.1	59.6	129.2	91.5	78.0	44.5
-0.1	165.0	134.9	106.7	59.4	129.3	100.7	83.7	52.4
0.0	177.2	129.1	98.0	55.3	126.7	98.8	83.9	50.6
0.1	173.0	135.1	110.6	62.1	122.1	75.4	89.9	47.4
0.5	177.4	124.2	98.8	54.9	117.6	95.9	83.0	48.0
0.9	186.2	140.9	113.7	57.9	142.1	109.8	94.8	65.2

trimmed averages. For example, in trimmed samples, we discarded the lowest and highest 5% of values. We also calculated medians, which are more suitable for comparing the central tendency in skewed distributions.

To assess the effect of dependency strength on the performance of PBX control charts, we applied ANOVA to compare overall ARL averages. While ANOVA is a common omnibus test for comparing sample means, its applicability is limited in our context, due to non-normality and the presence of outliers. Therefore, we also employed the non-parametric Kruskal-Wallis test, which is more robust under these conditions.

We begin by examining the potential impact of subsegment-level autocorrelation on ARL values. In our model, dependencies between consecutive observations within a subsegment are defined using bivariate copulas with identical, fixed marginal distributions, unaffected by ρ . Subsegment data are presented as histograms, whose shapes reflect these marginal distributions. Thus, we argue that neither the nature nor the strength of dependence within subsegments influences the aggregated data or the control chart performance. To verify this claim, we simulated ARL values under both Scenario 2 and Scenario 5 for seven levels of ρ , with results presented in Table 12. The first four columns in Table 12 display results for Scenario 2 (normally distributed subsegments with random parameters), while the next four columns correspond to Scenario 5 (subsegments drawn from either symmetric normal or asymmetric Weibull distributions). We report the total ARL averages (ARL), winsorized averages (ARLW), trimmed averages (ARLT), and medians (MRL).

For Scenario 2, the estimated averages and medians appear consistent across ρ values, confirmed by ANOVA ($p = 0.99$) and Kruskal-Wallis ($p = 0.3$) tests. For Scenario 5, ANOVA fails to reject the equality of means ($p = 0.76$), but the Kruskal-Wallis test does reject distributional similarity ($p = 8.65 \times 10^{-5}$). Post-hoc analyses (Nemenyi, Dunn, and Wilcoxon-Mann-Whitney tests) indicate that the sample with strong positive dependence ($\rho = 0.9$) differs significantly. This

Table 13. Impact of dependence of subsegment averages

ρ	Scenario-2				Scenario-5			
	ARL	ARLW	ARLT	MRL	ARL	ARLW	ARLT	MRL
-0.9	142.9	103.2	85.8	52.0	132.4	100.1	81.9	51.6
-0.5	141.2	114.6	98.6	51.4	138.9	110.8	89.2	50.1
-0.1	136.8	111.0	91.7	58.2	140.8	106.2	88.9	55.9
0.0	140.3	118.8	98.3	58.5	136.6	104.0	88.5	55.7
0.1	152.6	118.0	101.0	63.9	159.8	121.9	102.3	57.0
0.5	194.7	144.4	118.7	69.8	169.7	131.7	110.8	64.9
0.9	297.5	243.0	198.0	106.4	238.5	197.2	163.9	89.0

may be due to the effect of fixed class limits used to form histograms, which effectively winsorizes the data - an effect particularly influential with skewed distributions like the Weibull. This supposition is confirmed by the statistical analysis of fuzzy control limits $\tilde{Z} = (Z_L, Z_C, Z_U)$. The results of experiments show that the lower value Z_L does not depend upon ρ . However, ρ has an impact on the value of Z_C , and especially on the value of Z_U . It means that in segment data aggregation the value of ρ exerts impact on the lower bound of a p-box. This lower bound represents those data, which are shifted to the right, i.e., the data whose average value is mostly affected by winsorization. Removing this outlying group restores distributional similarity (Kruskal-Wallis $p = 0.18$). It is worth noting that low p -values in the Kruskal-Wallis test could also result from differing variances rather than from shifts in central tendency.

Next, we assessed the effect of dependence between subsegment averages on chart performance, summarized in Table 13. For both scenarios, the presence of moderate ($\rho = 0.5$) and strong ($\rho = 0.9$) dependencies visibly impacts the data. This is supported by both ANOVA and Kruskal-Wallis tests. When excluding the most dependent cases (two bottom rows of Table 13), similarity among the remaining datasets is confirmed for Scenario 2 (ANOVA $p = 0.93$, Kruskal-Wallis $p = 0.4$) and Scenario 5 (ANOVA $p = 0.1$, Kruskal-Wallis $p = 0.33$). Again, fixed class limits and resulting winsorization may explain the observed discrepancies.

We then examined the influence of inter-segment average dependence. It is well-established in Statistical Process Control (SPC) literature that such autocorrelation can significantly affect chart performance (see, e.g., Montgomery, 2011). Since the 1970s, two main approaches have emerged: adjusting control limits for the original data, and using residuals (differences between observed and predicted values) for monitoring. These approaches typically model data as time series, e.g., autoregressive processes. Schmid (1995) provided the first rigorous analysis of a Shewhart chart under such conditions. A broader overview

Table 14. Impact of the interdependence (Gaussian copula) of segment averages

ρ	Scenario-2				Scenario-5			
	ARL	ARLW	ARLT	MRL	ARL	ARLW	ARLT	MRL
-0.9	727.9	715.2	534.0	99.5	534.1	493.0	232.7	90.3
-0.5	349.0	308.7	241.8	109.0	301.4	256.9	209.1	105.2
-0.1	332.2	273.9	227.5	119.4	267.0	223.6	183.6	105.9
0.0	328.5	283.2	227.2	126.2	263.0	216.4	186.4	115.9
0.1	320.3	276.3	225.8	128.4	267.3	232.7	179.5	87.5
0.5	329.1	286.8	229.3	109.1	271.6	215.4	172.9	95.2
0.9	327.7	254.5	185.5	72.1	274.9	212.0	158.4	70.1

was later presented by Knoth and Schmid (2004). In our work, we limit analysis to the Shewhart chart under a first-order autoregressive (AR(1)) process.

Alternatively, we model subsegment dependencies using a copula-based approach, following Hryniewicz (2012). This method, rare in SPC outside financial applications, shows that Shewhart chart behavior depends not only on the strength of dependence (e.g., Kendall's τ) but also on its type (copula form).

We begin with Gaussian copula-based dependence (bivariate normal with normal marginals). Table 14 presents the respective simulation results. Across both scenarios, ARL averages remain stable except when $\rho = -0.9$ (strong negative correlation). Except for this case, the remaining datasets support average equivalence (Scenario 2: $p = 0.99$, Scenario 5: $p = 0.89$). However, analysis suggests that for both strong positive and strong negative correlations the median run lengths (MRLs) are significantly smaller than in the remaining cases. If we exclude these extreme situations, the remaining five datasets show median similarity (Scenario 2: $p = 0.6$), though Scenario 5 fails to confirm this even at low correlation ($p = 0.03$).

With AR(1) segment dependence (Table 15), results differ. While average ARL values increase with $|\rho|$, as described in SPC literature, the medians deviate markedly: they remain consistent for weak/moderate positive dependence but drop for strong correlations. This is statistically supported and represents an unexpected finding. Notably, this effect does not arise when monitoring data with fixed normal distributions.

Finally, we acknowledge that simulation outcomes may be sensitive to seed values. For complex processes, 500 chart sample sizes, suitable for homogeneous data, may be insufficient. However, larger samples can also lead to detecting statistically significant yet practically irrelevant differences due to high test sensitivity.

Table 15. Impact of the interdependence (AR(1) process) of segment averages

ρ	Scenario-2				Scenario-5			
	ARL	ARLW	ARLT	MRL	ARL	ARLW	ARLT	MRL
-0.9	731.7	712.4	538.8	103.2	729.6	724.0	538.5	119.8
-0.5	361.4	393.5	235.0	113.5	445.0	299.6	293.9	122.0
-0.1	314.0	331.7	211.2	117.6	374.3	260.1	266.2	131.7
0.0	299.8	267.3	210.2	105.3	311.3	259.0	221.6	119.5
0.1	296.2	279.0	206.1	115.4	333.5	253.6	159.5	128.1
0.5	323.6	304.0	203.0	96.5	363.6	259.2	228.5	106.0
0.9	348.3	341.1	190.0	64.1	406.6	270.2	246.6	83.8

7. Conclusions

In this paper, we describe the analysis of a segmented, non-stationary, and complex process using the concept of p-boxes. P-boxes are employed to provide an imprecise description of aggregated segment data. The two expected values (lower and upper) of the p-box describing a data segment, together with the expected value of its probabilistically aggregated distribution, are used to represent the segment's average as a triangular fuzzy number. It is worth noting that, in constructing these fuzzy numbers, medians could be used instead of expected values; however, this option is not considered in the present paper.

For the purpose of process monitoring, these triangular fuzzy numbers are compared with fuzzy control limits on a fuzzy Shewhart control chart. An alarm signal is generated when the possibility (or necessity) that the observed triangular fuzzy number, representing the aggregated data segment, exceeds the corresponding fuzzy control limit is sufficiently high. Based on this principle, four charting schemes are proposed. Extensive simulation experiments reveal that three of these schemes may be preferable under different, but well-defined, data scenarios. The proposed approach is not computationally demanding; therefore, after introducing further simplifications - such as the use of crisp control limits - it can be applied to the monitoring of continuous data streams measured at discrete time points.

We have demonstrated that complex segmented processes can be represented as streams of fuzzy numbers constructed using a fully objective and probabilistically interpretable method. These streams of aggregated fuzzy data may also be monitored using other types of control charts, such as those described in the references cited in Section 2.2, which were proposed for monitoring fuzzy data. From a practical point of view, this constitutes the most important result of this paper.

References

- ANGELOV, P. AND ZHOU, X. (2006) Evolving fuzzy systems from data streams in real time. In: *Proc. of 2006 IEEE Symp. on Evolving Fuzzy Systems*, IEEE, 29–35.
- ASLAM, M. U., XU, S., NOOR-UL-AMIN, M., HUSSAIN, S. AND WAQAS, M. (2024) Fuzzy Control Charts for Individual Observations to Analyze Variability in Health Monitoring Processes. *Applied Soft Computing*, **169**, 111961.
- ASLAM, M.U, XU, S.-H., RASHEED, Z., NOOR-UL-AMIN, M., HUSSAIN, S. AND WAQAS, M. (2025) Improved fuzzy control charts for monitoring defined health ranges using trapezoidal fuzzy numbers. *Expert Systems With Applications* **278**, 127310.
- CARMONA-POYATO, A., FERNANDEZ-GARCIA, N.L., MADRID-CUEVAS, F.J., MUNOZ-SALINAS, R. AND ROMERO-RAMIREZ, F.J. (2024) Optimal on-line time series segmentation. *Knowledge and Information Systems*. **66**, 2417–2438.
- CATALA-LOPEZ, F., GENOVA-MALERAS, R., VIETA, E. AND TABARES-SEISDEDOS, R. (2013) The increasing burden of mental and neurological disorders. *Eur. Neuropsychopharmacology* **23**, 1337–1339.
- CHENG, C.-B. (2005) Fuzzy process control: construction of control charts with fuzzy numbers. *Fuzzy Sets and Systems*, **154**, 287–303.
- CHEN, Y.W. AND DILSAVER, S.C. (1996) Lifetime rates of suicide attempts among subjects with bipolar and unipolar disorders relative to subjects with other Axis I disorders. *Biol Psychiatry*. **39**, 896–899 (1996)
- CHU, C.S. (1995) Time series segmentation: A sliding window approach. *Information Science*, **85**, 147–173.
- SILVA, V.D.C., ZARPELAO, B.B., MEDVET, E. AND BARBON, JR S. (2023) Explainable Time Series Tree: An Explainable Top-Down Time Series Segmentation Framework. *IEEE Access*, **11**, 120845–120865.
- DESTERCKE, S. AND DUBOIS, D. (2006) A unified view of some representations of imprecise probabilities. In: J. Lawry, E. Miranda, A. Bugarin, S. Li, M.A. Gil, P. Grzegorzewski, O. Hryniewicz (eds.), *Soft Methods for Integrated Uncertainty Modelling*. Springer, *Advances in Soft Computing* **37**, 249–257.
- DESTERCKE, S. AND DUBOIS, D. (2009) The role of generalised p-boxes in imprecise probability models. In: T. Augustin, F.P.A. Coolen, S. Moral, M.C.M. Troffaes (eds.), *ISIPTA'09: Proc. of the Sixth Intern. Symposium on Imprecise Probability: Theories and Applications*, Durham, 179–188.
- DESTERCKE, S., DUBOIS, D. AND CHOJNACKI, E. (2008) Unifying practical uncertainty propagations: I generalized p-boxes, II clouds. *Int. J. of Approx. Reas.* **49**, 649–677.

- DUBOIS, D. AND PRADE, H. (1988) *Possibility Theory - An approach to Computerized Processing of Uncertainty*, Plenum Press, New York.
- ERGINEL, N. AND ŞENTURK, S. (2016) Fuzzy EWMA and fuzzy CUSUM control charts. In: C. Kahraman and O. Kabak, (eds.) *Fuzzy Statistical Decision-Making: Theory and Applications*, Springer, 281–295.
- EYBEN, F., WENINGER, F., GROSS, F. AND SCHULLER, B. (2013) Recent developments in opensmile, the Munich open-source Multimedia Feature Extractor. *Proc. of the 21st ACM Int. Conf. on Multimedia*, 835–838
- FERSON, F., GINZBURG, L., KREINOVICH, V., LONGPRÉ, L. AND AVILES, M. (2002) Computing variance for interval data is NP-hard. *ACM SIGACT News*, **33**(2), 108–118.
- FERSON, F., GINZBURG, L., KREINOVICH, V., MYERS, D. AND SENTZ, K. (2003) Constructing probability boxes and Dempster-Shafer structures. *Technical report SAND 2002-4015*, Sandia National Laboratories.
- FERSON, F., KREINOVICH, V., HAJAGOS, J., OBERKAMPF, W. AND GINZBURG, L. (2007) Experimental Uncertainty Estimation and Statistics for Data Having Interval Uncertainty. *Technical report SAND 2007-0939*, Sandia National Laboratories.
- FAURHOLT-JEPSEN, M., FROST, M., VINBERG, M., CHRISTENSEN, E.M., BARDRAM, J.E. AND KESSING, L.V. (2014) Smartphone data as objective measures of bipolar disorder symptoms. *Psychiatry Research*. **217**, 124–127.
- GAMA, J. (2010) *Knowledge Discovery from Data Streams. Data Mining and Knowledge Discovery*. Chapman and Hall, London.
- GENG, D. AND ZENG, X.-J. (2020) Learning data streams on line - An evolving fuzzy system approach with self-learning/adaptive thresholds. *Information Sciences*, **507**, 172–184.
- GRANDE, I., BERK, M., BIRMAHER, B. AND VIETA, E. (2016) Bipolar disorder. *The Lancet* **387** (10027), 1561–1572.
- GRÜNERBL, A., MUAREMI, A., OSMANI, V., BAHLE, G., OEHLER, S., TRÖSTER, G., MAYORA, O., HARING, C. AND LUKOWICZ, P. (2015) Smartphone-based recognition of states and state changes in bipolar disorder patients. *IEEE Journal of Biomedical and Health Informatics*. **19**(1), 140–148.
- GRZEGORZEWSKI, P. (1997) Control charts for fuzzy data. In: *Proceedings of the 5th European Congress on Intelligent Techniques and Soft Computing EUFIT'97*, Aachen, 1326–1330.
- GRZEGORZEWSKI, P. AND HRYNIEWICZ, O. (2000) Soft methods in statistical quality control. *Control & Cybernetics* **29**(1), 1–22.
- GÜLBAY, M., KAHRAMAN, C. AND RUAN, D. (2004) α -Cuts fuzzy control charts for linguistic data. *Intl. Journ. of Intelligent Systems*, **19**, 1173–1196.

- GÜLBAY, M. AND KAHRAMAN, C. (2007) An alternative approach to fuzzy control charts: Direct fuzzy approach. *Information Sciences*, **177**, 1463–1480.
- HEITJAN, D.F. (1989) Inference from Grouped Continuous Data: A Review. *Statistical Science* **4**(2), 164–183.
- HE, Z., ZHANG, S., HU, J. AND DAI, F. (2024) An adaptive time series segmentation algorithm based on visibility graph and particle swarm optimization. *Physica A*, **636**, 129–563.
- HESAMIAN, G., AKBARI, M.G. AND RANJBAR, E. (2019) Exponentially weighted moving average control chart based on normal fuzzy random variables. *Int. J. Fuzzy Syst.* **21**(4), 1187–1195.
- HRYNIEWICZ, O. (2012) On the Robustness of the Shewhart Control Chart to Different Types of Dependencies in Data. In: H.-J. Lenz, W. Schmid and P.-Th. Wilrich (eds.), *Frontiers in Statistical Quality Control* **10**, Springer, 19–33.
- HRYNIEWICZ, O. AND KACZMAREK-MAJER, K. (2021) Possibilistic aggregation of inhomogeneous streams of data. *Proc. of FUZ-IEEE Conference*, doi: 10.1109/FUZZ45933.2021.9494583.
- HRYNIEWICZ, O. AND KACZMAREK-MAJER, K. (2023) Monitoring of Possibilistic Aggregated Complex Time Series. In: L. A. Garcia-Escudero et al. (eds.), *Building Bridges between Soft and Statistical Methodologies for Data Science*, Springer, *Advances in Intelligent Systems and Computing* **1433**, 208–215.
- HRYNIEWICZ, O. AND KACZMAREK-MAJER, K. (2024a) Aggregation of Inhomogeneous Non-stationary Time Series using P-boxes. In: J. Ansari, S. Fuchs, W. Trutschnig, M. A. Lubiano, M. A. Gil, P. Grzegorzewski and O. Hryniewicz (eds.), *Combining, Modelling and Analyzing Imprecision, Randomness and Dependence. Advances in Intelligent Systems and Computing Springer*, **1458**, 184–191.
- HRYNIEWICZ, O. AND KACZMAREK-MAJER, K. (2024b) Monitoring Complex Segmented Streams of Data Using Bootstrap Control Charts. In: S. Knoth, Y. Okhrin and P. Otto (eds.), *Advanced Statistical Methods in Process Monitoring, Finance, and Environmental Science: Essays in Honour of Wolfgang Schmid*, Springer, 105–125.
- HRYNIEWICZ, O., KACZMAREK-MAJER, K. AND OPARA, K.R. (2019) Monitoring two-state processes using indirectly observed data. *Proc. of XIIIth Inter. Workshop on Intell. Stat. Qual. Contr.*, Hong Kong, City Uni. of Hong Kong, 59–78.
- KAHARAMAN, M.G. AND BOLTURK, E. (2016) Fuzzy Shewhart control charts. In: C. Kahraman, O. Kabak (eds.) *Fuzzy Statistical Decision-Making: Theory and Applications*, Springer, 263–280.
- KANAGAWA, A., TAMAKI, F. AND OHTA, H. (1993) Control charts for process

- average and variability based on linguistic data. *Intl. Journ. of Prod. Res.*, **31**, 913–922.
- KNOTH, S. AND SCHMID, W. (2004) Control charts for time series. A review. In: H. J. Lenz, P.-Th. Wilrich (eds.), *Frontiers in Statistical Quality Control* **7**, Springer, 210–236.
- KULLDORFF, G. (1961) *Contribution to the Theory of Estimation from Grouped and Partially Grouped Samples*. Almqvist and Wicksell, Stockholm.
- KUZNETSOV, V. AND MOHRI, M. (2015) Learning theory and algorithms for forecasting non-stationary time series. In: C. Cortes, D.D. Lee, M. Sugiyama and R. Garnet (eds), *Proc. of NIPS'15*, 1, Cambridge, MA: MIT, 541–549.
- LIU, X., LIN, Z. AND WANH, H. (2008) Novel Online Methods for Time Series Segmentation. *IEEE Trans. on Knowledge and Data Engin.*, **20**, 1616–1626.
- LUGHOFFER, E. (2011) *Evolving Fuzzy Systems - Methodologies, Advanced Concepts and Applications*. Springer, Heidelberg.
- MONTES, I. (2021) On the Comonotone Natural Extension of Marginal p-Boxes. *Proceedings of Machine Learning Research* **147**, 210–218.
- MONTES, I. AND MIRANDA, E. (2018) Extreme Points of the Core of Possibility Measures and Maxitive p-Boxes. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* **31**, 1017–1051.
- MONTGOMERY, D.C. (2011) *Introduction To Statistical Quality Control* (6th Ed.). Wiley, New York.
- NELSEN, R. B. (2006) *An Introduction to Copulas*. Springer Verlag, New York.
- PRIESTLEY, M. B. (1988) *Non-Linear and Non-Stationary Time Series Analysis*. Academic Press, New York.
- RAZALI, H.M., LAZIM, A., AB GHANI, A.T., ZAHARUDIN, Z.A. AND AFTHANORHAN, A. (2025) Exponentially Weighted Moving Average Charts Based on Interval Type-2 Fuzzy Numbers: Analyses of Quality Control and Performance. *Int. Journ. of Fuzzy Systems*, **27**, 657–679.
- RUTKOWSKI, L., JAWORSKI, M. AND DUDA, P. (2020) *Stream Data Mining: Algorithms and Their Probabilistic Properties*. Springer, Cham.
- SABEGH, M.H.Z., MIRZAZADEH, A., SALEHIAN, S. AND WEBER, G.-W. (2014) A Literature Review on the Fuzzy Control Chart; Classifications & Analysis. *Int. Journ. of Supply and Oper. Managm.*, **1**, 167–189.
- SAYED-MOUCHAWEH, M. (Ed.) (2019) *Learning from Data Streams in Evolving Environments. Methods and Applications*, Springer, Cham.
- SCHMID, W. (1995) On the run length of a Shewhart chart for correlated data. *Statistical Papers* **36**, 111–130.
- ŞENTURK, S. AND ERGINEL, N. (2009) Development of $\tilde{\bar{X}} - \tilde{R}$ and $\tilde{\bar{X}} - \tilde{S}$ control charts using α -cuts. *Information Sciences* **179**, 1542–1551.

- ŞENTURK, S., ERGINEL, N., KAYA, İ. AND KAHRAMAN, C. (2014) Fuzzy exponentially weighted moving average control chart for univariate data with a real case application. *Applied Soft Computing* **22**, 1–10.
- SHU, M.-H. AND WU, H.-C. (2011) Fuzzy $\bar{X}$ and $\bar{R}$ control charts: Fuzzy dominance approach. *Computers & Industrial Engineering* **61**, 676–685.
- TROFFAES, M. AND DESTERCKE, S. (2011) Probability boxes on totally pre-ordered spaces for multivariate modelling. *Int. J. of Approx. Reas.* **52**, 767–791.
- UTKIN, L. AND DESTERCKE, S. (2009) Computing expectations with continuous p-boxes: Univariate case. *Int. J. of Approx. Reas.* **50**, 778–798.
- WANG, D. AND HRYNIEWICZ, O. (2013) The design of a CUSUM control chart for LR-fuzzy data. *Proceedings of the 2013 Joint IFSA World Congress, NAFIPS Annual Meeting*, Edmonton, Canada, July 2013, 175–180.
- WANG, D. AND HRYNIEWICZ, O. (2015) A fuzzy nonparametric Shewhart chart based on bootstrap approach. *Int. J. of Appl. Math. and Comp. Science* **25**, 389–401.
- WANG, L., ZHU, H. AND JIA, G. (2020) Adaptive G-G clustering for fuzzy segmentation of multivariate time series. *Stoch. Env. Res. and Risk AS.* **34**, 1353–1367.
- WANG, C., LI, X., ZHOU, T. AND CAI, Z. (2024) Unsupervised Time Series Segmentation: A Survey on Recent Advances. *Computers, Materials & Continua* **80**, 2657–2672.
- WANG, J.H. AND RAZ, T. (1990) On the construction of control charts using linguistic variables. *Intl. Journ. of Prod. Res.* **28**, 477–487.
- WANG, L. AND WANG, R. (2026) Parameter adaptive segmentation algorithm based on temporal-weighted clustering for time series tensor. *Information Sciences* **728**, 122795.