applied sciences

MDPI

Article

Analysis and Application of Selected Control Charts Suitable for Smart Manufacturing Processes

Tereza Smajdorová and Darja Noskievičová *

Department of Quality Management, Faculty of Materials Science and Technology, VŠB-Technical University of Ostrava, 17. listopadu 2172/15, 708 00 Ostrava, Czech Republic; tereza.smajdorova@vsb.cz

* Correspondence: darja.noskievicova@vsb.cz

Abstract: Nonparametric control charts (NPCC) have shown great potential for monitoring processes in conditions of smart manufacturing with complex structures, various monitored characteristics and the need to process big data. Practical applications of NPCCs are very rare. The main reasons for this situation are a deficiency in software support and a lack of simple but complete instructions for their application. The introduction of such manual, which is based on the authors' own simulations of performance of wide spectrum of NPCCs in conditions of different violations of data prerequisites, leading to recommendations for the selection of the most effective NPCC in various practical situations, is the main goal of this paper. Compared to other similar studies, this approach covers a wider range of control charts, and it was applied to a wider spectrum of data assumption violations. As an integral part of these analyses, an examination of various control chart performance indicators such as ARL, MRL, x_5 and x_{95} was performed using simulations to select the best of them. The designed methodology was verified using real data.

Keywords: nonparametric control charts; control chart performance simulations; control chart performance indicators; data assumption violations; smart manufacturing

check for updates

Citation: Smajdorová, T.; Noskievičová, D. Analysis and Application of Selected Control Charts Suitable for Smart Manufacturing Processes. *Appl. Sci.* **2022**, *12*, 5410. https://doi.org/10.3390/app12115410

Received: 2 April 2022
Accepted: 23 May 2022
Published: 26 May 2022

1. Introduction

Statistical process monitoring methods have an irreplaceable position in smart process control.

Various control charts were developed for monitoring processes. Classical parametric control charts (CC) assume the presence of data normality, constant mean and variance and data independence [1–3]. However, such prerequisites for data are quite often not met in practical conditions, which is a frequent situation in smart manufacturing processes with complex structure, various monitored characteristics and the need to process big data. In conditions of big data, which have already been enabled by massive deployment of various sensors, large amounts of processed data with various structures and high frequency of collection have led to the fact that statistical properties of data such as data non-normality of various types or data dependence are more typical of modern manufacturing processes than they were of previous manufacturing systems.

However, improper application of parametric control charts, including Shewhart's control charts, in such situations can lead to wrong conclusions about a process's statistical stability. Thus, more dependable nonparametric control charts were designed for these situations] [4,5]. Nonparametric control charts (NPCC) have shown great potential for monitoring processes in conditions of smart manufacturing due to their ability to overcome the shortcomings of classical parametric control charts [6–9], as they were developed for situations when the distribution of the analyzed process is arguable or unknown.

The main benefits of NPCCs are as follows: it is not necessary to presume a particular distribution for the analyzed process, they are greatly efficient in detecting changes when data are not normally distributed, they are more robust and resistant to outliers, and it is not required to assess the variance when constructing a control chart for the location parameter [10,11]. However, practical utilization of NPCC is very rare, predominantly because of a deficiency in software support and a lack of simple but complete instructions for their application. The introduction of such a manual is the main output of this paper.

The heart of this developed methodology can be found in the phase of selection of the most effective NPCC. The main goal of the research described in this paper is the selection of the most effective NPCC in relation to various considered deviations from data preconditions and the determination of the most robust NPCC. The phase mentioned above is based on the authors' simulations and comparison of the performance of 6 various univariable NPCCs not only for monitoring location but also for monitoring process variability: the Shewhart sign control chart (SSCC), the nonparametric exponentially weighted moving average signed-rank control chart (NP-EWMA), the nonparametric sign cumulative sum chart based on the Mann–Whitney statistic (NP-CUSUM), the nonparametric progressive mean control chart (NP-PM), the nonparametric control chart based on the Mood statistic for dispersion (NP-MOOD) and the nonparametric control chart based on the median absolute deviation (NP-MAD).

In contrast to the study described in this paper, other similar studies compare fewer control charts, and very often they compare NPCCs to some classical Shewhart control chart. For instance, in [12], the author makes a comparison between an NPCC based on a two-sample rank-sum test with the classical Shewhart S chart. In [13], an NPCC based on Mann–Whitney statistics is compared to the classical Shewhart X-bar chart. Oprime [14] designed an NPCC with simultaneous monitoring of location and scale and compared it with classical Shewhart control charts. In [15], Shewhart S charts and NP-MAD charts are compared. In [16], authors put the stress on the performance of three control charts—median, bootstrap and Hodges-Lehmann control charts [17,18] in contrast to a Shewhart X-bar control chart with estimated control limits. The most various control charts are analyzed by the authors of [19]: the empirical likelihood-based control chart, classical Shewhart X-bar chart, classical EWMA chart and bootstrap chart.

Another goal of this research was to analyze the performance of 6 selected NPCCs to cover various deviations from the data assumptions as much as possible. For this reason data from 7 distributions, i.e. normal distribution, distributions with greater and lower kurtosis as compared to normal distribution, asymmetrical, two bimodal distributions and autocorrelated data were incorporated into these simulations. Other authors when simulating NPCC performance used fewer distributions, distributions with different parameters or quite different distributions, such as the β distribution [20]; normal and exponential distributions [14]; normal, Laplace and uniform distributions [12]; normal, Laplace and Gamma distributions [13]; normal, Student, Gamma, two symmetrical bimodal and one asymmetrical bimodal distributions [19]; and normal, uniform, Student, double exponential and Cauchy distributions [16].

To assess control chart performance, an ARL (average run length) indicator is applied most frequently [12,14,19,21]. However, some criticism of this indicator can be found in the literature [10,22], based on the fact that "*the run length is a positive integer valued random variable, so the ARL loses much of its attractiveness as a typical summary if the distributions is skewed*" ([10], p. 304). Some authors recommend and use MRL (median run length) or quantiles x_{10} and x_{90} [13,16]. As the quality of the performance indicators has a crucial impact on precise evaluation of NPCC performance and so on meeting the main goal of this research, the authors incorporated into the research an additional partial goal—to analyze ARL, MRL, x_5 and x_{95} to identify the most precise indicators for simulations.

To reach all goals and realize the planned output, the authors set several hypotheses that were verified during the research:

Hypothesis 1. *Some NPCCs are more suitable (they have better performance) for the particular data assumption violations than others*.

Hypothesis 2. *Some NPCCs are robust (distribution-free) against most data precondition violations*.

Hypothesis 3. *Performance indicators such as MRL or other quantiles of the run-length distribution are better for the evaluation of control chart performance*.

Hypothesis 4. *Nonparametric control charts are resistant against outliers (which cannot be explained and for this reason cannot be removed from the data or repaired)*.

The paper consists of the following sections: Section 2 focuses on the theoretical framework and the solved topic. Section 3 describes the used methods, instruments and analyses. Section 4 is devoted to the results obtained from the previous analyses. The designed instructions for the practical use of NPCCs are applied to real data in Section 5. Finally, in Section 6, discussion of the results can be found.

2. Theoretical Background

In the technical literature, numerous examples of the utilization of parametric statistical process control when data come from the normal distribution can be found. Unfortunately, in practice, there are many processes where normality cannot be met. The results of classical parametric statistical process control may not be correct in these cases (see [23–25]). These sources provide sufficient justification for the development and practical application of control charts in cases of noncompliance with data normality or other data preconditions for the use of parametric control charts. NPCCs and distribution-free control charts have been designed to achieve this aim. The term nonparametric is not entirely accurate, as this would mean that no parameter needs to be involved. This is not true because at least the assumption of a continuous distribution is required. The notion of distribution-free seems to be more correct because it captures the essence of these control charts It means they are not dependent on a specific probability distribution, mostly the normal distribution. However, these terms are taken as synonyms, and the term nonparametric is used more frequently [10].

In the same source, a very valuable overview of the literature on NPCCs can be found. The authors divided NPCCs into four classes: Shewhart-type charts, CUSUM-type charts, EWMA-type charts and other methods.

Shewhart-type control charts are practically applicable because of their relative simplicity. The first nonparametric Shewhart procedure for monitoring the location parameter of a continuous process when the value of the parameter for the statistically stable process is not specified was described in [26]. Other Shewhart-type methods are discussed in [27,28].

CUSUM-type control charts are more suitable for the detection of small changes in processes. In 1975 Reynolds [29] introduced the nonparametric CUSUM based on sequential signed ranks of observations. Other contributions to nonparametric CUSUM-type charts can be found in [30–32].

Various EWMA-type charts, also suitable for detecting small changes, are described for instance in [33–38].

In addition to the previous three classes of nonparametric methods, the fourth class of NPCC defined in [10] covers various special cases that were introduced for instance in [15,21,39–41]. Adekeye and Azubuike in [15] compared the classical Shewhart X-bar control chart with the MAD (median absolute deviation) control chart. In [21], a new NPCC MIN is presented. Interesting special nonparametric methods are discussed in [39]. Murakami and Matsuki in [40] suggested an NPCC working with the Mood statistic.

In the following subsections there are described in more detail some NPCCs that were selected for analyses described in this paper.

2.1. SSCC Control Chart

The Shewhart sign control chart (SSCC) is one of the simplest NPCCs [42]. Its principle is as follows: Let θ_0 be a target value; each measurement x_{ij} from some unknown continuous distribution is transformed according to Formula (1):

$$sign(x) = x_{ij} - \theta_0 = \begin{cases} 1, & if\ x_{ij} > \theta_0 \\ 0, & if\ x_{ij} = \theta_0 \\ -1, & if\ x_{ij} < \theta_0 \end{cases} \tag{1}$$

The statistic SN_i recorded into the control chart is computed using Formula (2)

$$SN_i = \sum_{j=1}^{n} (x_{ij} - \theta_0) \tag{2}$$

and the control limits of the NPCC are set using Formulas (3)–(5):

$$UCL = c \tag{3}$$

$$CL = 0 \tag{4}$$

$$LCL = -c \tag{5}$$

where

$$c = 2t - n \tag{6}$$

where n is a sample size, and t is constant that can be found in [42,43].

2.2. NP-CUSUM

This NPCC is based on the Mann–Whitney statistic [44,45]. It is supposed that the quality characteristic X has a mean μ. For every measurement is computed the deviation $y_i = x_i - \mu$, and then it is transformed into values I_i using Formula (7):

$$I_i = \begin{cases} 1\ if\ \ y_i > 0 \\ 0\ if\ \ y_i \leq 0 \end{cases}\ for\ i = 1, 2, \ldots, n. \tag{7}$$

Let MW_j be the total number of positive deviations for every subgroup.

The standardized Mann–Whitney statistic SMW_j is computed using Formula (8):

$$SMW_j = \frac{MW_j - E_0(MW_j)}{\sqrt{var_0(MW_j)}} \tag{8}$$

In the case of a two-sided NP-CUSUM, statistics $S_j{}^+$ and $S_j{}^-$, computed according to Formulas (9) and (10), are recorded into the control chart:

$$S_j{}^+ = \max\{0; S^+{}_{j-1} + SMW_j - K\} \tag{9}$$

$$S_j{}^- = \min\{0; S^-{}_{j-1} + SMW_j + K\} \tag{10}$$

$$S\_0 = 0 \tag{11}$$

$$K = \Delta/2 \tag{12}$$

where Δ is the size of the shift that should be detected as soon as possible. The control limits are set as follows:

$$UCL = H,\ LCL = -H \tag{13}$$

For parameters K and H:

$$H = h \cdot \sigma, K = k \cdot \sigma \tag{14}$$

where σ is the standard deviation of the statistic used for the creation of NP-CUSUM, and k and h are standardized forms of parameters K and H, respectively.

In general, the values of parameters K and H should be set in a such way that the combination of k and h will lead to the required value of ARL(0) [44]. For instance, when the required ALR(0) = 370, the combination of k = 0.5 and h = 4.77 is quite good [22].

2.3. Nonparametric Control Chart Based on Mood Statistic

NP-Mood was derived for monitoring dispersion. The principle of this NPCC is well described in [40].

Suppose that we have a reference sample X from a statistically stable process with m values and an arbitrary test sample Y of n values. $R_1 < \cdots < R_n$ are the combined-sample ranks of the X value arranged in increasing order of magnitude. The statistics recorded into this NPCC are computed using Formula (15):

$$M_{m,n} = \sum_{i=1}^{m} \left(R_i - \frac{N+1}{2} \right)^2, \text{ where } N = m + n \quad (15)$$

The control limits are then computed as follows:

$$UCL = E(M\_(m,n)) + c\sqrt{}(var(M\_(m,n))) \quad (16)$$

$$LCL = E(M\_(m,n)) - c\sqrt{}(var(M\_(m,n))) \quad (17)$$

where $E(M\_(m,n))$ is the mean, and $\sqrt{}(var(M\_(m,n)))$ is the variance of the statistic $(M\_(m,n))$; it is recommended to compute them using Formulas (18) and (19) [40], p. 758.

$$E(M_{m,n}) = \frac{m(N^2-1)}{12} \quad (18)$$

$$var(M_{m,n}) = \frac{mn(N+1)(N^2-4)}{180} \quad (19)$$

The value of c is set to correspond to the required value of ARL(0). For instance, for ARL(0) = 370.4, c = 2.782177 [40], p. 761.

3. Methodological Framework

In this section, the types of performed analyses and used methods will be described.

3.1. Selection of Nonparametric Control Charts

This study covers a wide range of control charts for location and for dispersion to represent NPCCs of various types. The criteria for the selection of control charts were simplicity of practical use, simplicity of the creation of SW support and possible former knowledge of the principles of similar parametric versions of the chosen control charts.

Based on these criteria, the following control charts were chosen: the Shewhart sign control chart (SSCC) [42], the nonparametric exponentially weighted moving average signed-rank control chart (NP-EWMA) [20,46,47], the nonparametric sign cumulative sum chart based on the Mann–Whitney statistic (NP-CUSUM) [32,44,45], the nonparametric progressive mean control chart (NP-PM) [48,49], the nonparametric control chart based on the Mood statistic for dispersion (NP-MOOD) [40] and the nonparametric control chart based on the median absolute deviation (NP-MAD) [15,50].

3.2. Nonparametric Control Chart SW Support

To be able to create analyzed NPCCs and realize our own simulations for judgment of the control charts, a performance program in MS Excel was produced.

The Excel sheet consists of several tables. One table is used for the input values. The sheet also contains basic information about the subgroups' size and number, the value of the mean, the standard deviation and other statistics necessary for the ensuing computations.

In the biggest table, computations of the statistics necessary for construction of the selected NPCC are realized. The program also enables calculation of the run length (RL) values and the control chart performance indicators ARL, MRL, x_5, and x_{95}. In the lower part of the sheet, the resulting NPCC is constructed. All computations and diagrams are made automatically after inserting the input values [51].

3.3. Selection of Control Chart Performance Indicators

Before simulations, attention had to be paid to the selection of the most precise control chart performance indicators that are most stable for different probability distributions and so the most convenient for appraising and comparing the performance of various NPCCs [52–54]. Indicators for statistically stable processes such as average run length ARL(0), median run length MRL(0), 5% quantile x_5 and indicators for statistically unstable processes such as ARL(δ), MRL(δ) and 95% quantile x_{95} were analyzed. For a statistically stable process, the values of the performance indicators must be as high as possible [12]. Calculations of these indicators are based on RL(0), which represents the quantity of points, plotted in the control chart, that are located inside the control limits between two points lying outside of these limits. For a statistically unstable process, the values of the performance indicators must be as small as possible. To compute these indicators, it is necessary to set RL(δ) (δ is a standardized value of the shift that is supposed to be detected by the control chart as soon as possible), which is a quantity of points plotted in the control chart starting from the point when the shift in the process occurred until the shift is announced by the point outside of the control limit.

To assess the quality of these control chart performance characteristics, their values were sorted in descending order. The course of the resulting pointed lines expresses the rate of stability of the analyzed performance indicators for various simulated deviations from the data prerequisites represented by 7 probability distributions. The more tardily the line goes down, the more stable the performance indicator can be considered. Based on this conclusion, it can be said that for statistically stable processes, the 5% quantile x_5 is the most stable indicator for various deviations from the data prerequisites. Regarding ARL(0) and MRL(0), the latter is the more stable indicator. Accordingly, the following performance evaluation of analyzed NPCCs for statistically stable processes was made using only x_5 and MRL(0) performance indicators [51]. As an example of this analysis, there are plots of performance indicators in relation to various violations of data assumptions represented by seven different probability distributions (see Table 1) for the SSCC control that can be found in Appendix A. Seven points on every curve in Figures A1–A6 (Appendix A) correspond to the seven probability distributions, and three points in a vertical direction represent three analyzed performance indicators. Such plots were constructed for all the rest of the analyzed NPCCs.

Table 1. Summary of various simulated violations of data prerequisites and their connection to the utilized probability distributions. Source: own research.

Unmet Data Assumption	Distribution/Data Properties	Distribution Parameters
None	Normal distribution	N (0,1)
Data normality	Pearson distribution/skewed distribution	χ^2_3
	Uniform distribution/less kurtosis	$R\left(-\sqrt{3};\sqrt{3}\right)$
	Student's distribution/greater kurtosis	t_3
Constant variance	Mixed distribution/different variances	50% N(0,1) + 50% N(0,4) MIX 2
Constant mean	Mixed distribution/different means	50% N(0,1) + 50% N(2,1) MIX 1
Data independence	Autocorrelated data	AR (1): $x_i - 0.5x_i + a_i$

Based on a similar analysis of statistically unstable processes, only x_{95} and MRL(δ) were utilized for evaluation of NPCC performance [51].

3.4. Design of Simulations of Statistically Stable Process

The statistically stable process corresponds to an in-control process which is influenced only by random causes of variability. Table 1 contains the probability distributions used for simulating different data prerequisites violations.

The quantity of subgroups m was equal to 300, 100 and 20, and the sizes of the subgroup n were equal to 10 and 5. The simulations were replicated 10,000 times. The values of n and m were set according to the results of simulations of statistically stable processes [52] (see also [53]) so that the ARL was about 370. To assess the performance of the analyzed NPCC, the indicators MRL(0) and x_5 were applied. This simulation methodology was utilized for all NPCCs analyzed in this study.

Let us have a sample of 100 values randomly generated from the applied distribution. This sample is divided into m subgroups with the size n. From these values, control limits for the constructed NPCC are calculated. These limits are drawn into the control chart. Afterwards, n values representing the particular subgroup are randomly generated from the applied distribution, and the control statistic is computed from these n values and entered into the control chart. Then the RL(0) values are determined, and the values of the performance indicators x_5 and MRL (0) are calculated from them.

The resulting graphs for all analyzed NPCCs for location and all simulated deviations from the data assumptions can be found in Appendix B (Figures A7 and A8). The outputs of the evaluation of NPCCs for monitoring variability are in Appendix C.

3.5. Design of Simulations of Statistically Unstable Process

The statistically unstable process corresponds to an out-of-control process, i.e., a process which is affected by both random causes and assignable causes of variability. The quantity of subgroups m was set as 300, 100 and 20, and the size of the subgroup n was equal to 10 and 5. The simulations were again replicated 10,000 times for each combination of n and m and for each particular violation of the data prerequisites.

The experiment was performed similarly to the simulations for the statistically stable process (see Section 3.4), but deviations of various sizes were inserted into the data sets additionally. First, isolated shifts of 1.5σ, 2σ and 3σ were inserted during the 30th, 50th and 99th repetitions, respectively. After that, simulations of persistent deviations of 1.5σ, 2σ and 3σ were performed but only for 20 $\times$ 5 combination of m and n and selected distributions. The deviation appeared before signaling by a point beyond the limit. Subsequently, "the process was intervened," and the following subgroup was without any deviation. Afterward, the deviation returned.

To assess the performance of an NPCC during a statistically unstable process, only the MRL(δ) and x_{95} indicators were applied (see Section 3.3).

4. Results

The previous analyses led to the results discussed in the following subsections.

4.1. Summary of Simulations of Statistically Stable Process

Based on the previous analyses (see Section 3.4 and Appendix B), it can be determined which NPCCs for location monitoring are the most effective for a given data prerequisite violation. The results are based on the prerequisite that for a statistically stable process, the values of the chart performance indicators should be as high as possible. The conclusions are summarized in Table 2.

Table 2. Summary of analysis of NPCCs for location monitoring. Source: own research.

SSCC	NP-CUSUM
• Distribution with bigger kurtosis than for normal distribution • Distribution with smaller kurtosis than for normal distribution • Asymmetric distribution • Nonconstant variance • Autocorrelated data (for larger samples)	• Distribution with bigger kurtosis than for normal distribution • Distribution with smaller kurtosis than for normal distribution • Nonconstant variance (for larger samples and $n = 5$)

As can be seen from the figures in Appendix B, the SSCC and NP-CUSUM chart have the best values of the used performance indicators in relation to the particular violations of the data presumptions as summarized in the previous table. But as the most universal (distribution-free, robust) NPCC, i.e., the CC that has very good performance indicator values over all or nearly all simulated violations of the data prerequisites, the control chart SSCC was selected [55].

Regarding the analysis of the performance of NPCCs for variability monitoring (see Appendix C), the following conclusions were made (see Table 3).

Table 3. Summary of analysis of NPCCs for variability monitoring. Source: own research.

MAD	Mood
• Distribution with smaller kurtosis than for normal distribution • Distribution with bigger kurtosis than for normal distribution • Nonconstant variance (for larger samples)	• Distribution with bigger kurtosis than for normal distribution • Asymmetrical distribution • Nonconstant variance (for smaller samples) • Nonconstant mean • Autocorrelated data

4.2. Summary of Simulations of Statistically Unstable Process

Regarding persistent deviations, analysis of the MRL(δ) and x_{95} performance indicators revealed that the persistent deviation is rapidly identified by NPCCs. The analyses showed that the NP-EWMA chart is the most powerful NPCC. Other NPCCs with acceptable x_{95} results are the NP-CUSUM chart and the SSCC [55].

However, for a statistically unstable process with an isolated deviation, the performance of NPCCs is quite poor, particularly for small process changes. It can be concluded that with increasing subgroup size, the performance indicators improve, as Das validated in [47].

This can be explained by the fact that NPCCs perceive isolated change as a random deviation, against which they are robust [51].

The best results for isolated deviations were demonstrated by the NP-EWMA chart, the SSCC and the NP-CUSUM chart.

5. Suggestion of Methodology for NPCC Practical Utilization

Based on previous simulations, obtained outcomes and general principles of statistical process control, a methodology for practical utilization of NPCCs was created. The steps of this methodology are depicted using the following scheme in Figure 1.

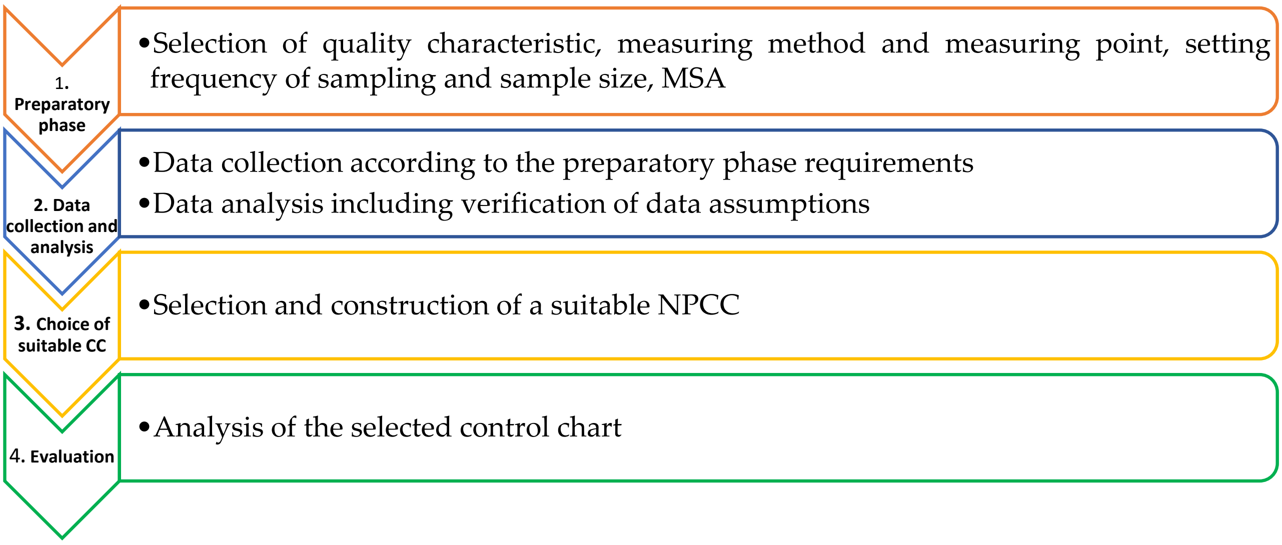

Figure 1. Scheme of the designed methodology. Source: own research.

Phases 1 and 4 contain similar activities, as they are well-known for the standard methodology of statistical process control applications. For that reason, only Phases 2 and 3, which represent the innovative core of the designed methodology, will be described in more detail. The collection of data and their analysis are very significant activities in the frame of statistical monitoring because especially verification of data assumptions forms the base for the objective decision of whether to apply the classic Shewhart CC or some NPCC. The effective verification of data assumptions calls for combining multiple statistical tests with graphical methods.

Significant attention must be paid to the analysis of potential outliers or suspicious values. When causes of them cannot be identified, these values must not be removed from the data set and must remain part of the following data analysis.

The phase of the selection of a suitable CC is wholly based on the previous results of performed simulations, and it creates the heart of this methodology design. When all data prerequisites are met, the classic Shewhart CC can be used, but when some of the data prerequisites are violated, applying the classic Shewhart CC could result in incorrect conclusions. Thus, the utilization of a suitable NPCC is a reasonable alternative in these cases. Phase 3 is clearly described using flowcharts (see Figures 2 and 3).

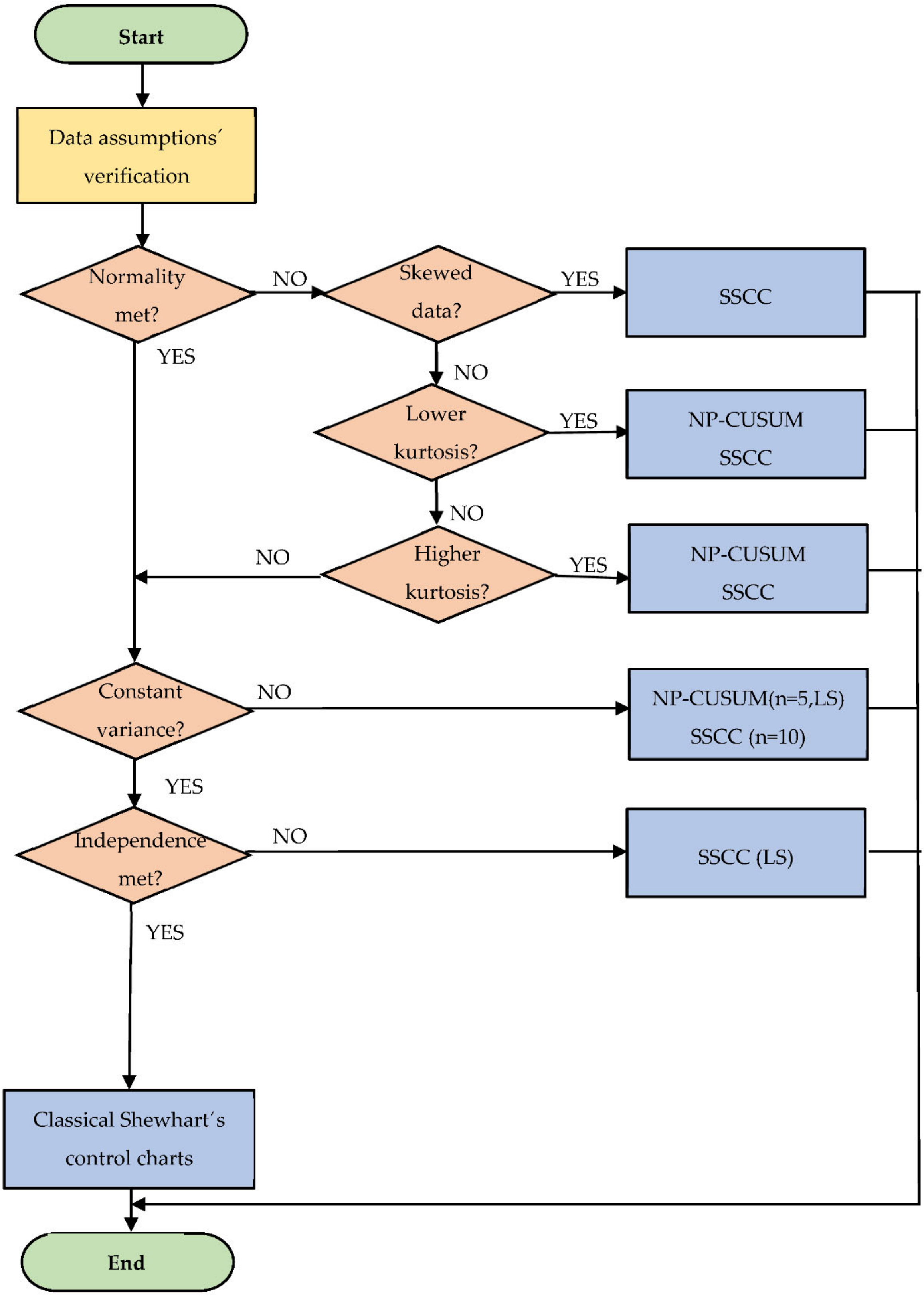

Figure 2. Flowchart for practical utilization of NPCCs for location monitoring. Source: Adapted from [55]. Legend: LS—large samples.

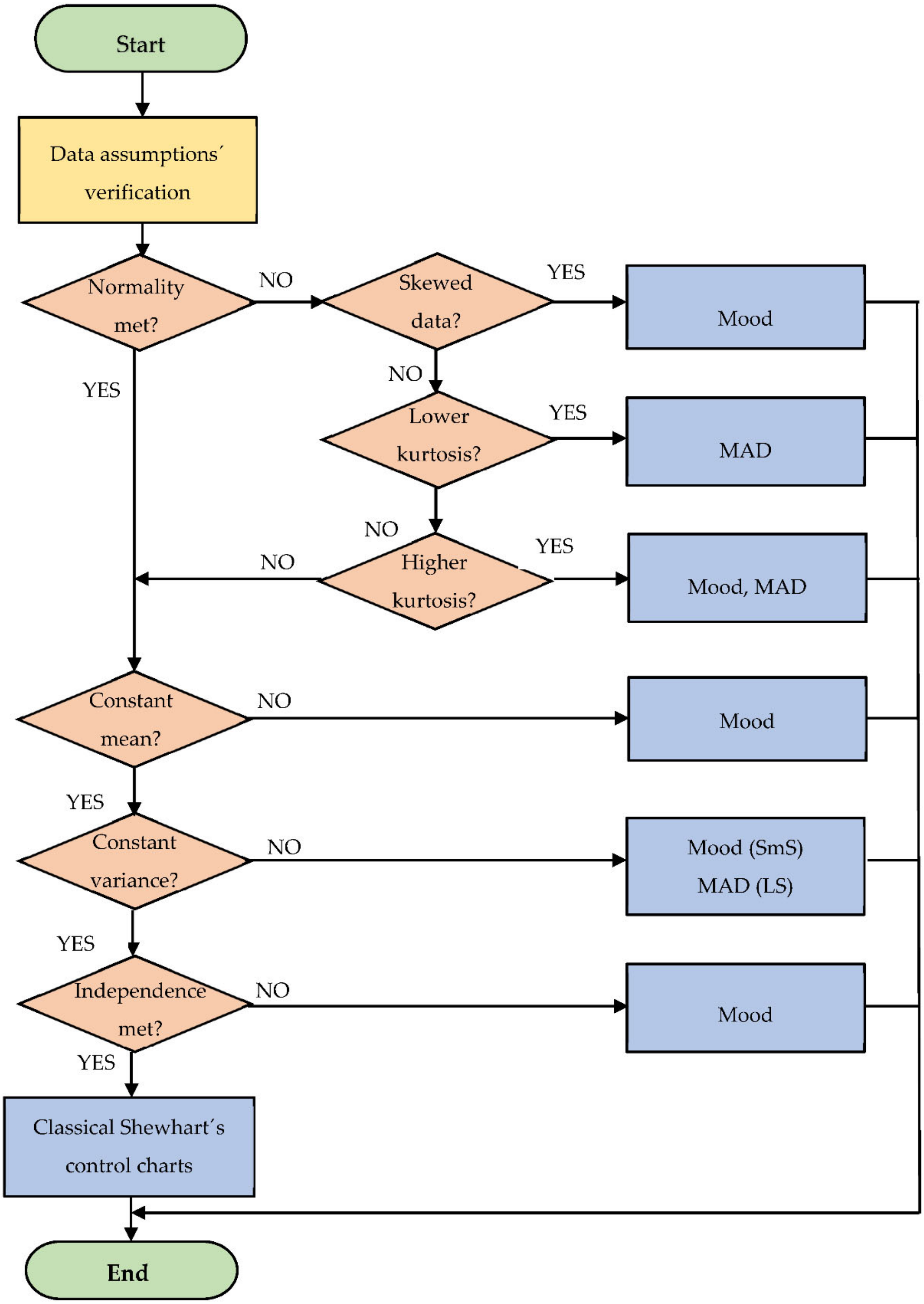

Figure 3. Flowchart for utilization of NPCCs for variability monitoring. Source: Own research. Legend: LS—large samples, SmS—small samples.

6. Practical Application

Verification of the designed methodology was performed using real data from a supplier for the automotive industry. The organization specializes in coating interior components with natural and synthetic leather. The weight of the adhesive that is automatically applied by a robot was set as a controlled characteristic. The weight of the adhesive after drying is calculated as the difference in the weight of the plastic part before application of the adhesive and the weight of the part after application of the adhesive and its drying. The weight of each piece is determined. The measured values were divided into 30 subgroups with a range of 5 units.

Subsequently, a data analysis was performed to determine whether the basic data assumptions had been met. The Anderson–Darling test was used to verify normality. Its results showed that the data are not normally distributed. Based on Figure 4, it can be concluded that the data are slightly skewed and contain some potential outliers or suspicious values.

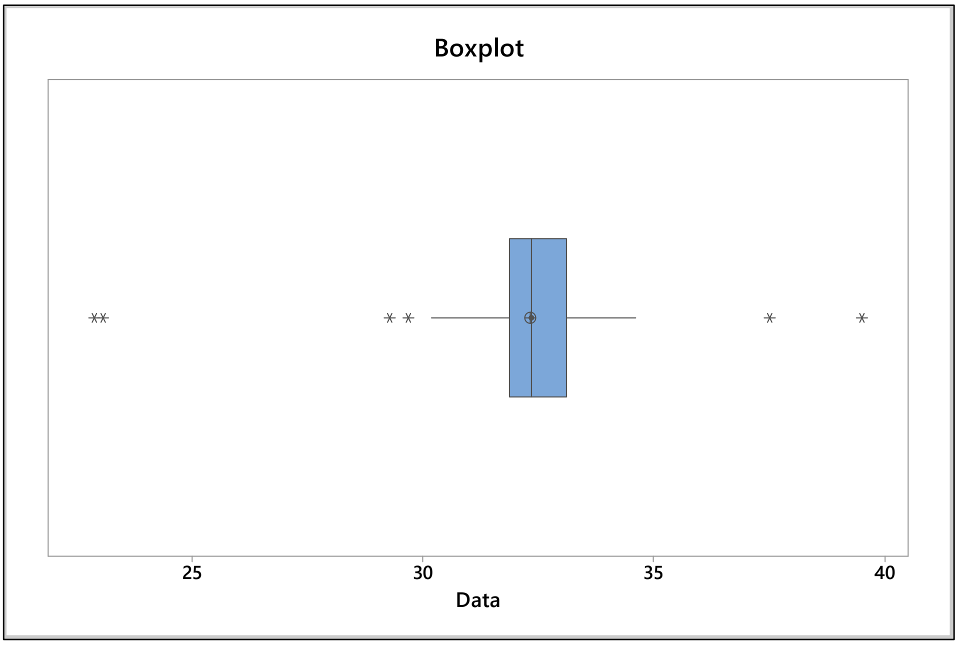

Figure 4. Box plot. Source: own research using SW Minitab.

Verification of data independence was performed using the Box–Pierce test. The resulting P-value of the test was 0.676, so it is greater than 0.05, and we can say that the data are independent. However, the most strongly violated data prerequisite seems to be kurtosis, as is clear from Figure 5 and from the value of the coefficient of kurtosis, which is equal to 17, 88.

The potential outliers and suspicious values were analyzed, but the reasons for their occurrence were not identified. For this reason, it was inappropriate to remove these values from the data set. However, using classical control charts, several points should be outside control limits UCL or LCL (see Figure 6), leading to useless searching for the nonexistent causes and time losses for the operators or forcing them to make unnecessary process adjustments (this is called tampering).

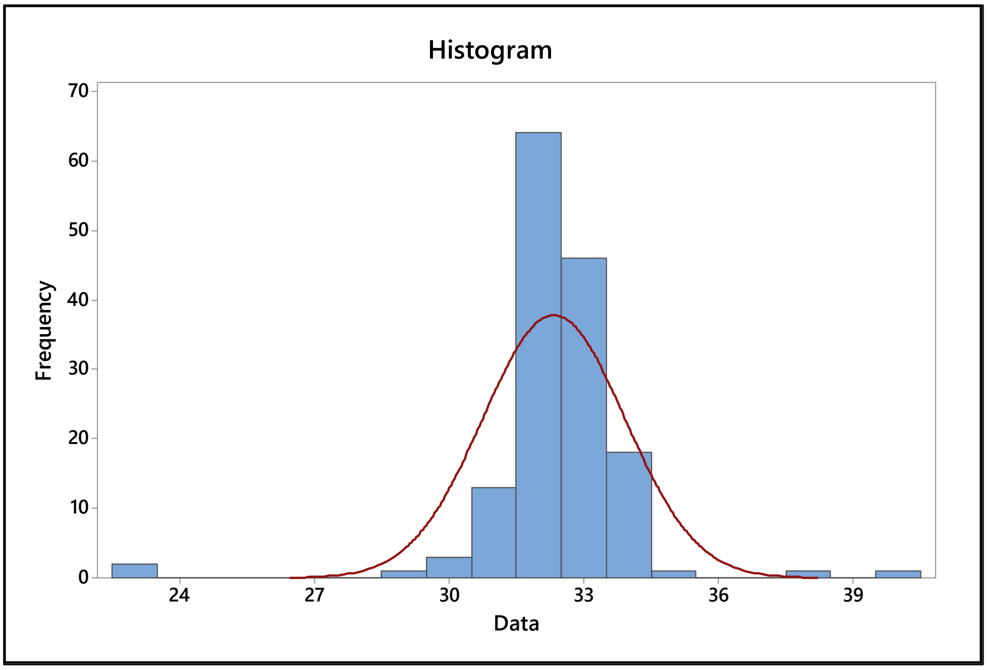

Figure 5. Histogram. Source: own research using SW Minitab.

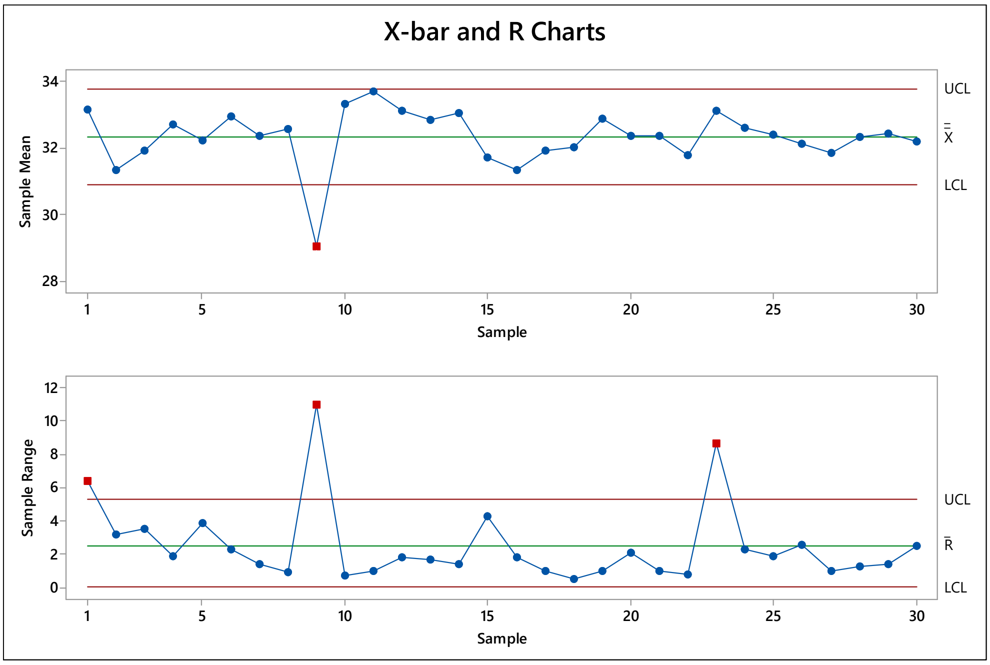

Figure 6. Parametric Shewhart control charts for mean and range. Source: own research using SW Minitab.

For these reasons, a suitable NPCC was selected using the suggested algorithm depicted in Figures 2 and 3, i.e., an NPCC based on the Mood statistic (for monitoring process variability) and an NP-CUSUM chart (for monitoring process position) were applied.

The values of the calculated statistics necessary for the construction of the selected control charts can be found in Table A1 (Appendix D). Figure 7 shows the NPCC based on the Mood statistic. As can be seen, all points are located inside the control limits. In terms of variability, the process is statistically stable.

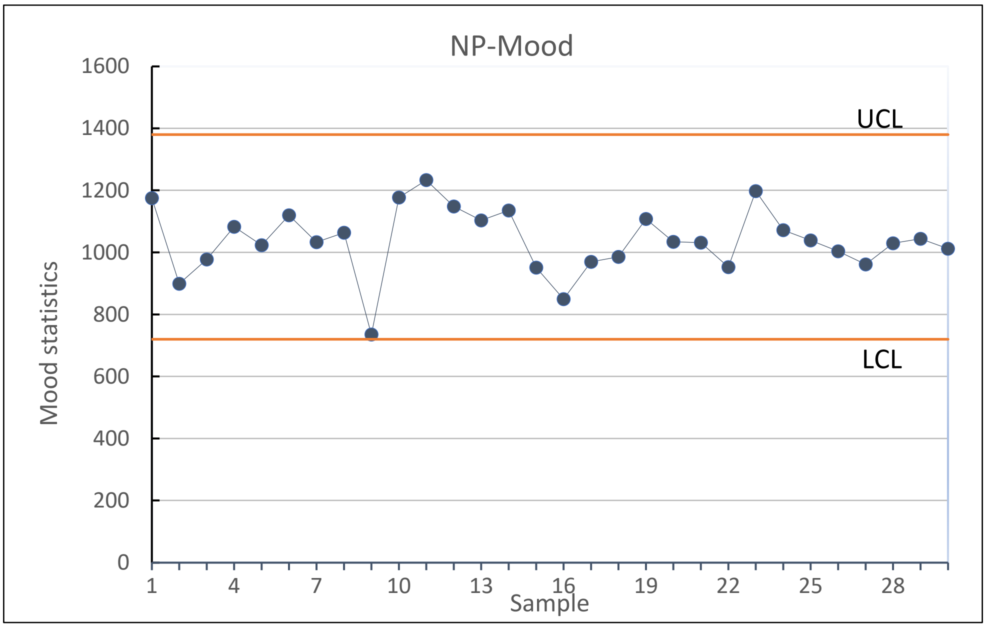

Figure 7. Nonparametric control chart based on the Mood statistic. Source: own research.

Regarding the evaluation of the process's statistical stability in terms of location, an NP-CUSUM chart was constructed. The values of the required statistics are given again in Table A1 in Appendix D. The resulting CC is shown in Figure 8. This control chart does not contain any points beyond the control limits, so it can be said that the process is statistically stable.

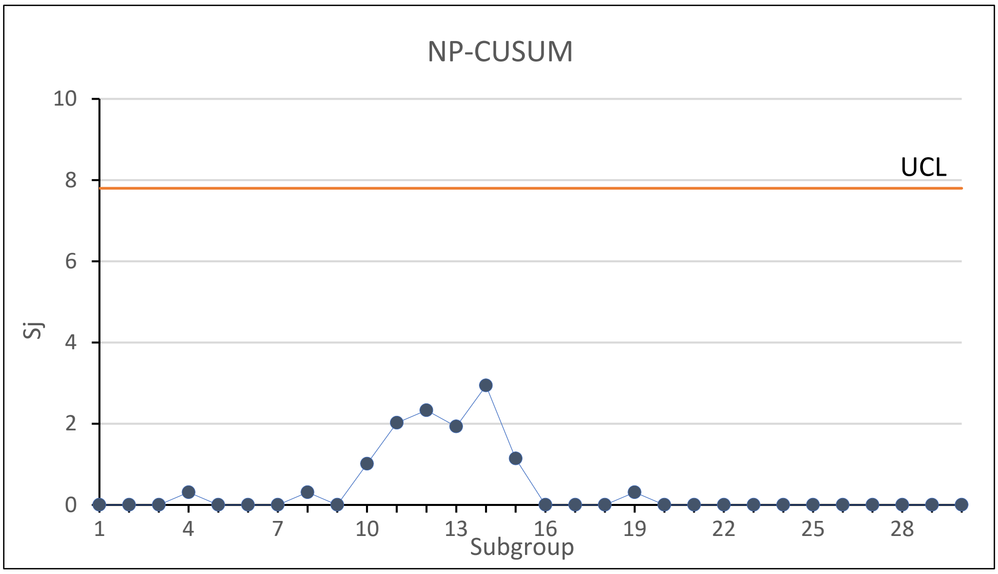

Figure 8. Nonparametric CUSUM control chart. Source: own research.

The results of this case study support Hypothesis 4 that NPCCs are resistant against potential outliers, which is suitable in situations when it is not possible to find causes of their occurrence.

7. Discussion

7.1. Theoretical Implications

An inevitable part of our research was studying the theoretical sources and investigation of practical experiences of workers in factories with SPCs and especially NPCCs. Research of the theoretical sources has shown that during the last 20 years, a large amount of research, concerning this topic, has been conducted. Many NPCC variations, some of them quite simple but many of them very complex, were developed and analyzed. The work of researchers devoting time to revealing the benefits of these methods and their ability overcome some practical problems with SPC applications was also very useful. These analyses have shown that NPCCs could be very suitable even in a new situation in manufacturing that asks for more and more collecting and processing of huge amounts of various data. However, a survey performed in manufacturing organizations revealed a large gap between theory and practice. The lack of software support and nonexistence of simple but complete instructions for the practical application of these methods were identified as the main reasons for this situation. For these reasons, we decided to propose such a methodology based predominantly on quite simple variations of the NPCC—nonparametric variants of the known parametric CC. There is at least some probability that people in companies are aware of their principles or have some practical experience with them (Shewhart, CUSUM and EWMA types of control charts). To make the methodology practically applicable, we had to create an SW application, as the situation relating to NPCCs on the statistical programs market is very poor (only Mathematica offers the option to create one of the types of nonparametric control charts—an NPCC based on Mann–Whitney statistics).

The survey also revealed another obstacle to the penetration of these methods into practice—a lack of professional courses and training focusing on NPCCs.

In the introductory part of this paper, the main research goals and four research hypotheses were defined. Table 4 summarizes the results of verification of the formulated hypotheses.

7.2. Practical Implications

The research has shown that the designed methodology is applicable in practice. It is based on quite simple NPCCs that are similar to parametric CC counterparts known in practice. The defined rules are clear and simple, and the research narrowed the field of options from a huge number of NPCCs to four charts. It should also simplify requirements for training and programming.

In addition, it has been shown that in practice, the widely used MS Excel program allows for the creation of an SW support, using both classic formulas and functions and more advanced programming using Visual Basic for Applications (VBA).

The confirmation of Hypothesis 4 yields an effective solution for situations that occur quite frequently during SPC applications: points outside of control limits whose real cause is impossible to identify. It could be good instrument against frequent disappointment in relation to SPC applications.

At present, the proposed methodology for the application of NPCCs can already support statistical control and monitoring with a high rate of predictability of behavior of modern manufacturing processes in a big data environment, as has been discussed in the introductory section of this paper. In the future, it can be incorporated into computer-based algorithms for self-control of cyber–physical systems (CPS), which are the fundamental technology platform of smart factories.

7.3. Limitations and Future Research Directions

One of the limiting factors of this analysis is the selected probability distributions. Even if the distributions used in our study covered all possible deviations from the data assumptions, future research could focus on verifying the results using the same distributions but with other parameters and other types of distributions as well. Another possibility for future research is to extend the study to other nonparametric control charts.

Table 4. Summary of verification of formulated hypotheses.

Number	Formulation of Hypothesis	Methods of Verification and Results
1.	Some NPCCs are more suitable (they have better performance) for particular data assumption violations than others.	Comparative analysis of NPCC performance using the best performance indicators. Appendices B and C Result: Hypothesis cannot be rejected: Analysis showed that some NPCCs are better for a particular data prerequisite violation then others. For instance, for monitoring the process location for data with greater kurtosis, the SSCC and NP-CUSUM chart have very good performance as compared to the rest of the analyzed NPCCs; For asymmetric distribution, the SSCC is the best NPCC. For monitoring the process dispersion for data with greater kurtosis, NP-Mood is the best, and for the data with smaller kurtosis, NP-MAD is the best.
2.	Some NPCCs are robust (distribution-free) against most data precondition violations.	Comparative analysis of NPCC performance using the best performance indicators. Appendices B and C Result: Hypothesis cannot be rejected: Analysis showed that some NPCCs are highly robust. For monitoring the process location, SSCC is very robust. For monitoring the process dispersion, NP-Mood is very robust.
3.	Performance indicators MRL or other quantiles of the run-length distribution are better for the evaluation of control chart performance.	Stability analysis of performance indicators Appendix A Result: Hypothesis cannot be rejected: MRL and other quantiles are better performance indicators as compared to ARL.
4.	Nonparametric control charts are resistant against outliers whose occurrence cannot be explained and for this reason removed from the data or repaired.	Comparative analysis of NPCC performance for statistically unstable process when isolated shifts have occurred. Example Result: Hypothesis cannot be rejected: An example practically proved validity of hypothesis.

Another limitation of this study was the nonexistent SW support for nonparametric control charts leading to the necessity of creating our own SW. Hand in hand with the extensions defined above, the SW that was created for the experiments described in this paper should be extended and improved to cover the abovementioned ideas.

8. Conclusions

Nonparametric statistical process control methods was well-described in various professional journals, but their practical applications have been infrequent. According to the results of the authors' simulations, it can be said that various NPCCs are differently effective for miscellaneous violations of data prerequisites. The issues of nonparametric statistical process monitoring methods are very extensive and present a huge number of new research opportunities, including improvement of the SW support. Nonparametric statistical process monitoring methods could become a permanent part of the teaching of statistical methods at universities, as well as in the frame of expert training.

Author Contributions: Conceptualization, D.N. and T.S.; methodology, D.N. and T.S.; software, T.S.; validation, D.N. and T.S.; formal analysis, D.N.; investigation, D.N. and T.S.; resources, T.S.; data curation, T.S.; writing—original draft preparation, T.S.; writing—review and editing, D.N.; visualization, T.S.; supervision, D.N.; project administration, D.N.; funding acquisition, D.N. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the European Regional Development Fund in A Research Platform focused on Industry 4.0 and Robotics in the Ostrava Agglomeration project, CZ.02.1.01/0.0/0.0/17_049/0008425 within the Operational Program Research, Development and Education.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

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

Appendix A

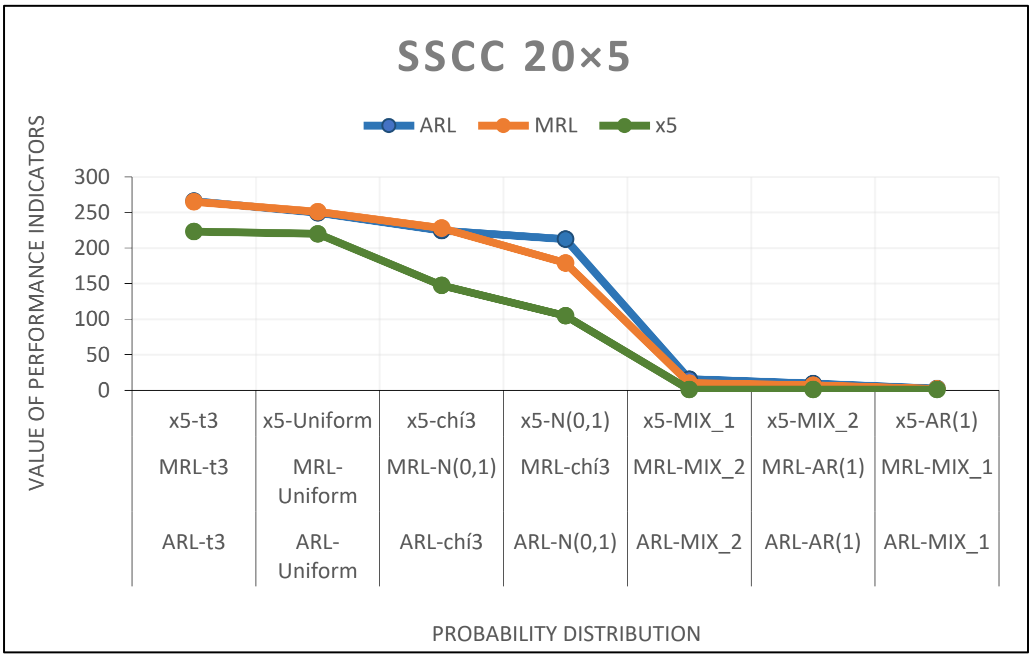

Figure A1. Values of performance indicators ARL(0), MRL(0) and x_5 for 7 used distributions, number of subgroups m = 20 and subgroup size n = 5 (control chart SSCC). Source: [51].

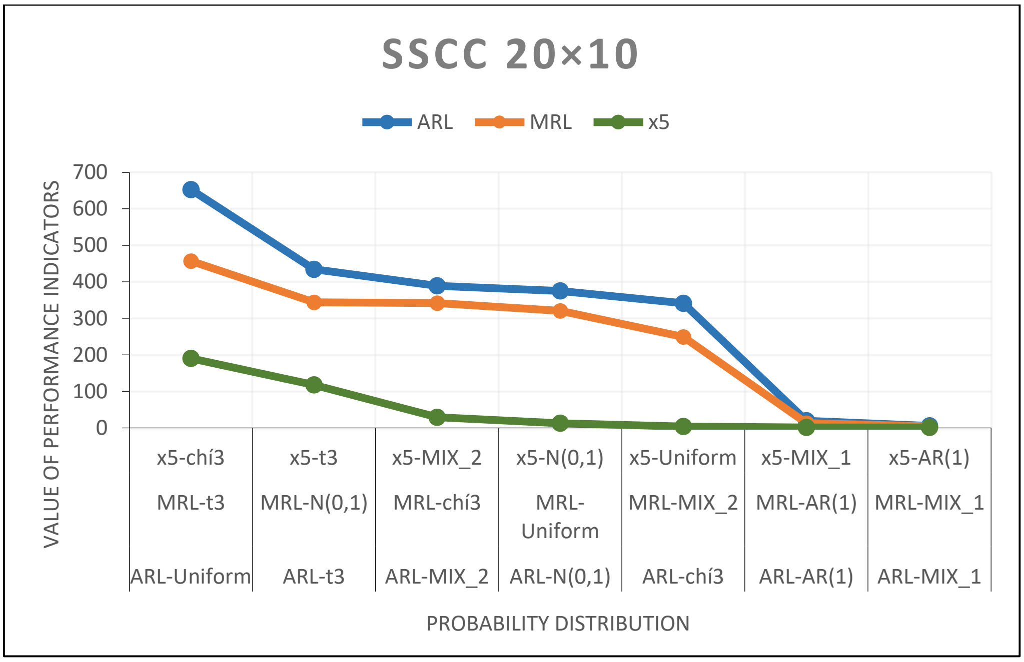

Figure A2. Values of performance indicators ARL(0), MRL(0) a x_5 for various 7 used distributions, number of subgroups m = 20 and subgroup size n = 10 (control chart SSCC). Source [51].

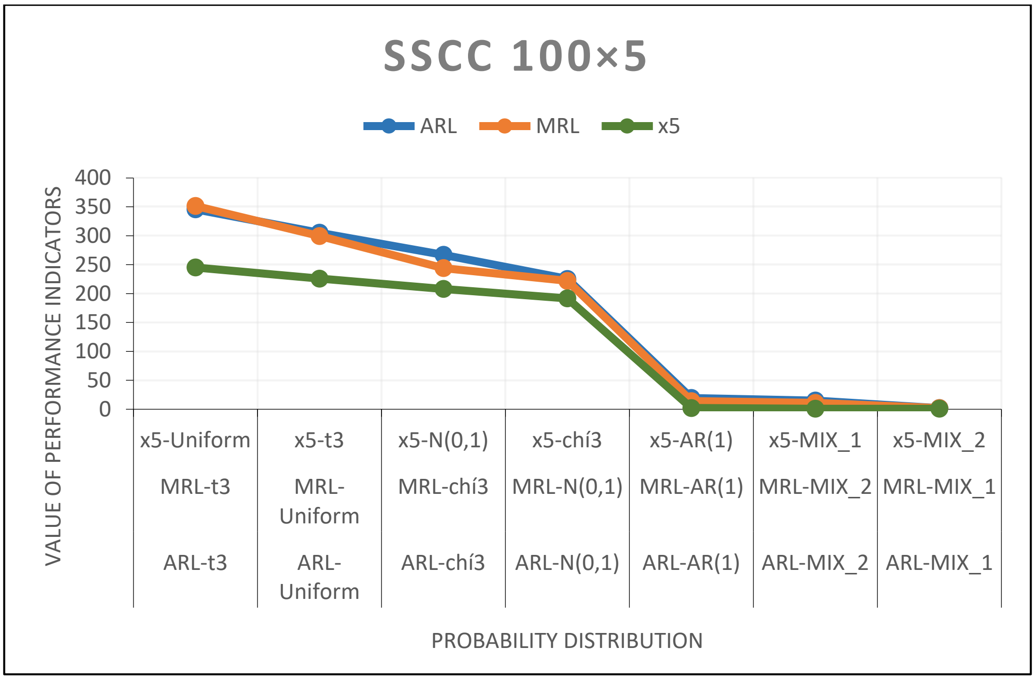

Figure A3. Values of performance indicators ARL(0), MRL(0) and x_5 for 7 used distributions, number of subgroups m = 100 and subgroup size n = 5 (control chart SSCC). Source [51].

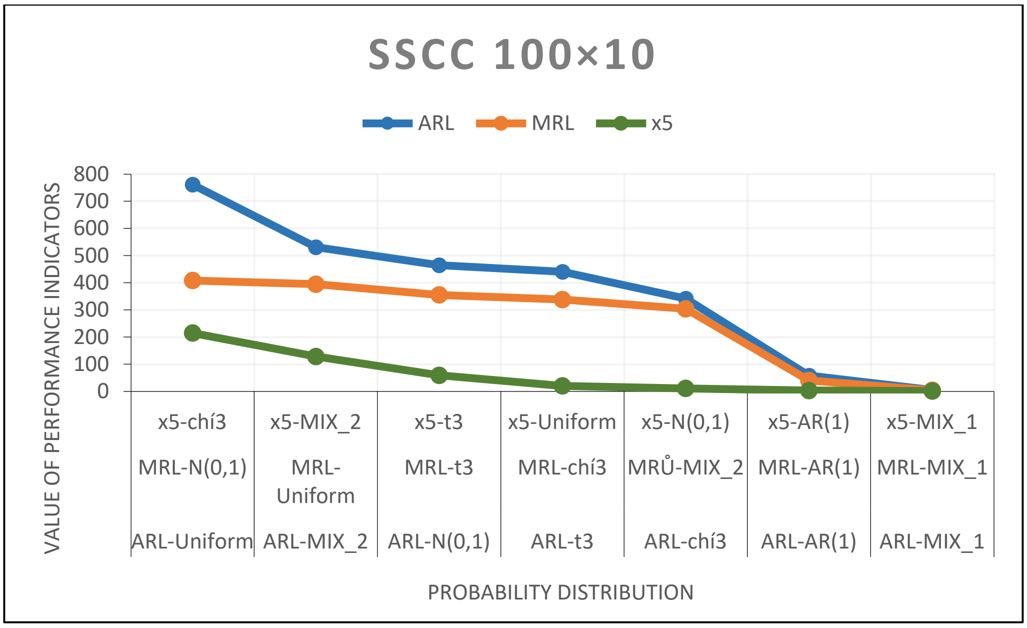

Figure A4. Values of performance indicators ARL(0), MRL(0) and x_5 for 7 used distributions, number of subgroups m = 100 and subgroup size n = 10 (control chart SSCC). Source [51].

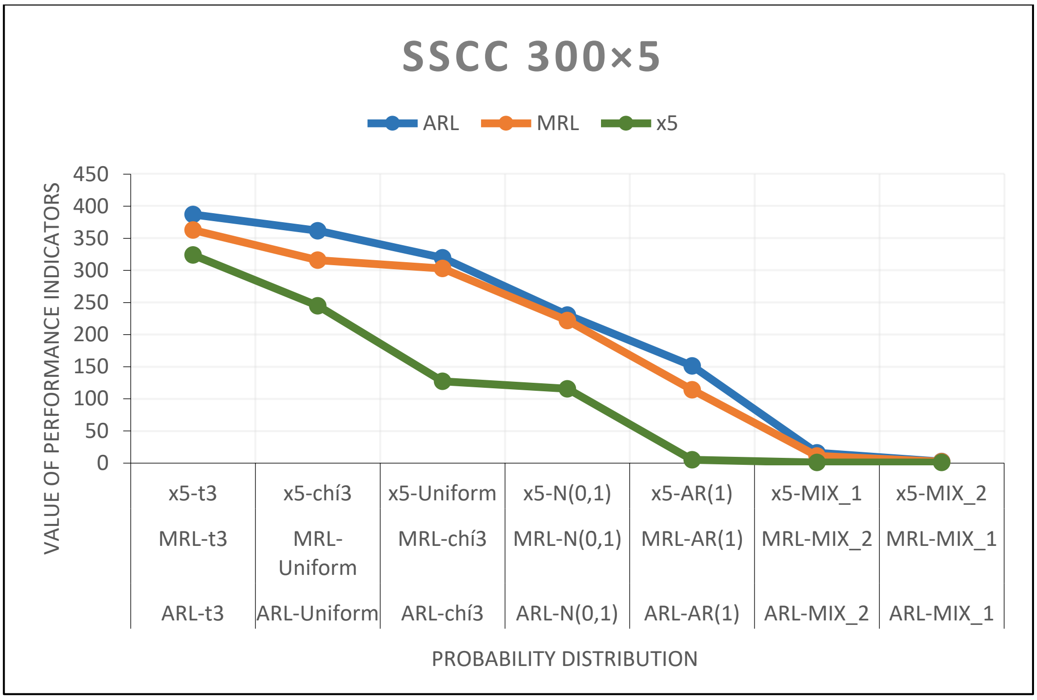

Figure A5. Values of performance indicators ARL(0), MRL(0) and x_5 for 7 used distributions, number of subgroups m = 300 and subgroup size n = 5 (control chart SSCC). Source: [51].

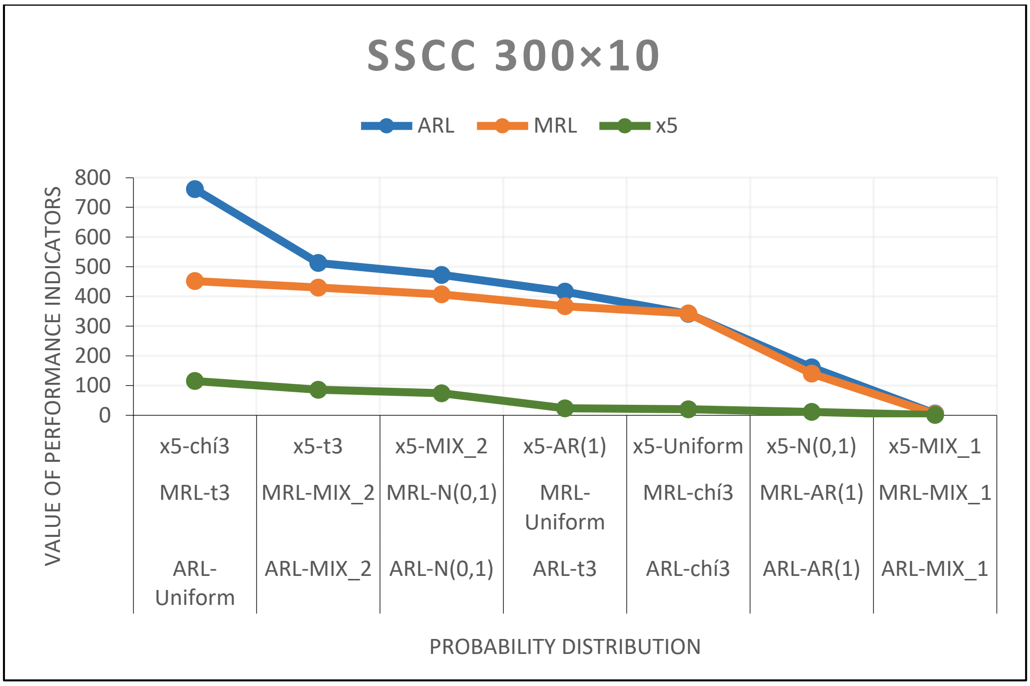

Figure A6. Values of performance indicators ARL(0), MRL(0) and x_5 for 7 used distributions, number of subgroups m = 300 and subgroup size n = 10 (control chart SSCC). Source: [51].

Appendix B

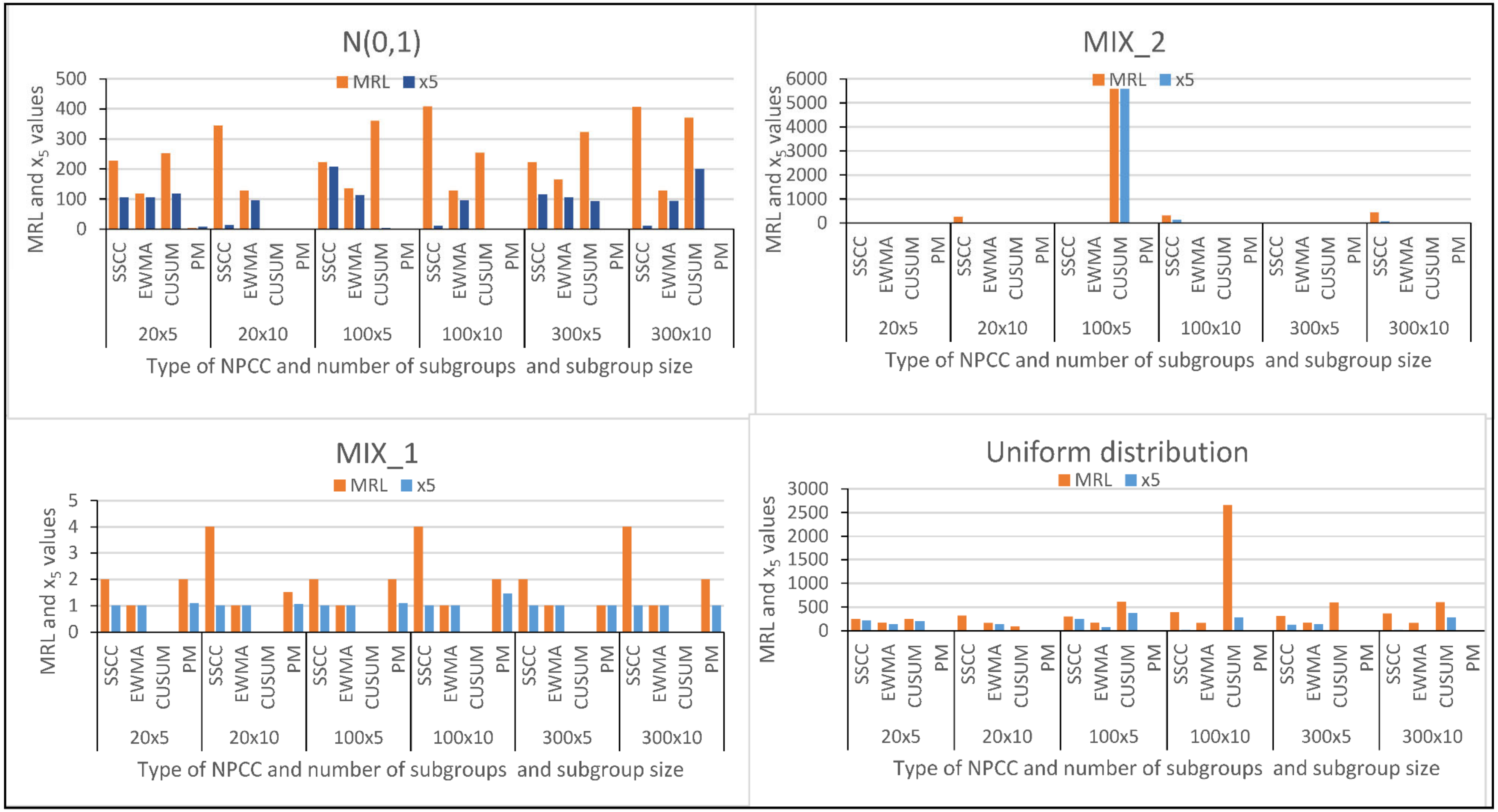

Figure A7. Values of MRL (0) and x_5 for various NPCCs for monitoring process location, different subgroup number and subgroup size—normal, two bimodal and uniform distributions. Source: [51].

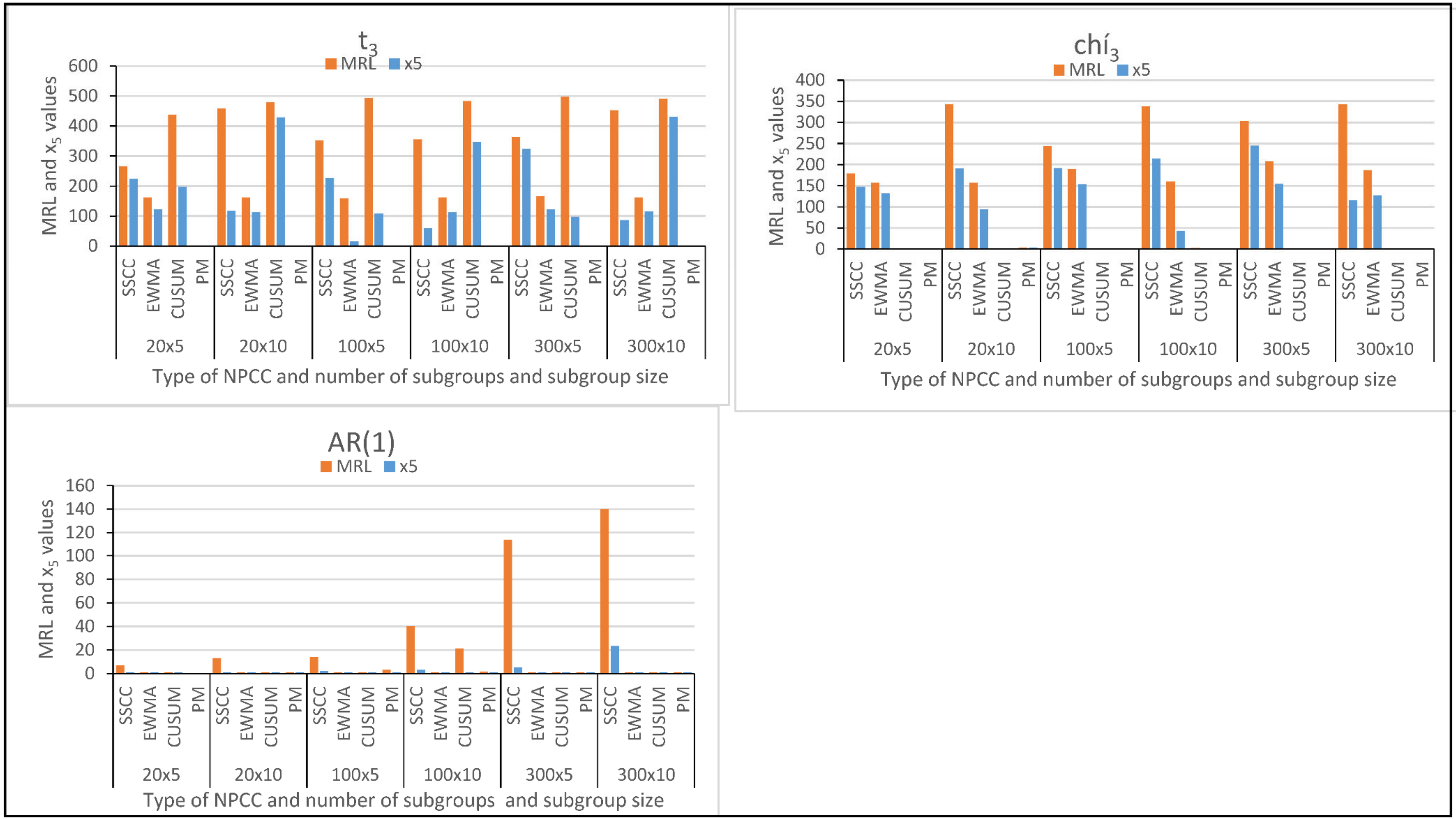

Figure A8. Values of MRL (0) and x_5 for NPCCs for monitoring process location—Student, chi-squared distributions and autocorrelated data. Source: [51].

Appendix C

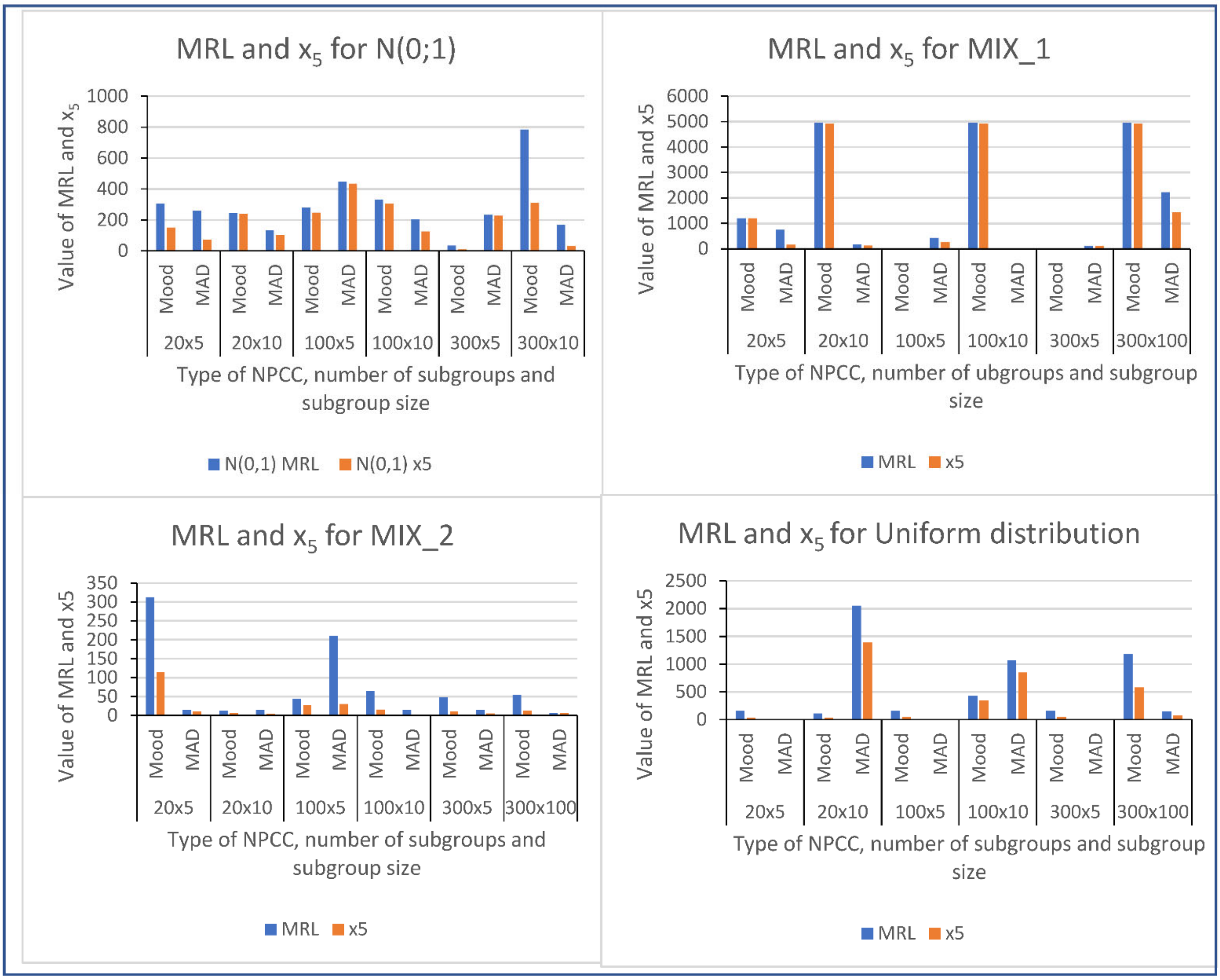

Figure A9. Values of MRL (0) and x_5 for NPCCs for monitoring process dispersion—normal, 2 bimodal and uniform distributions [51].

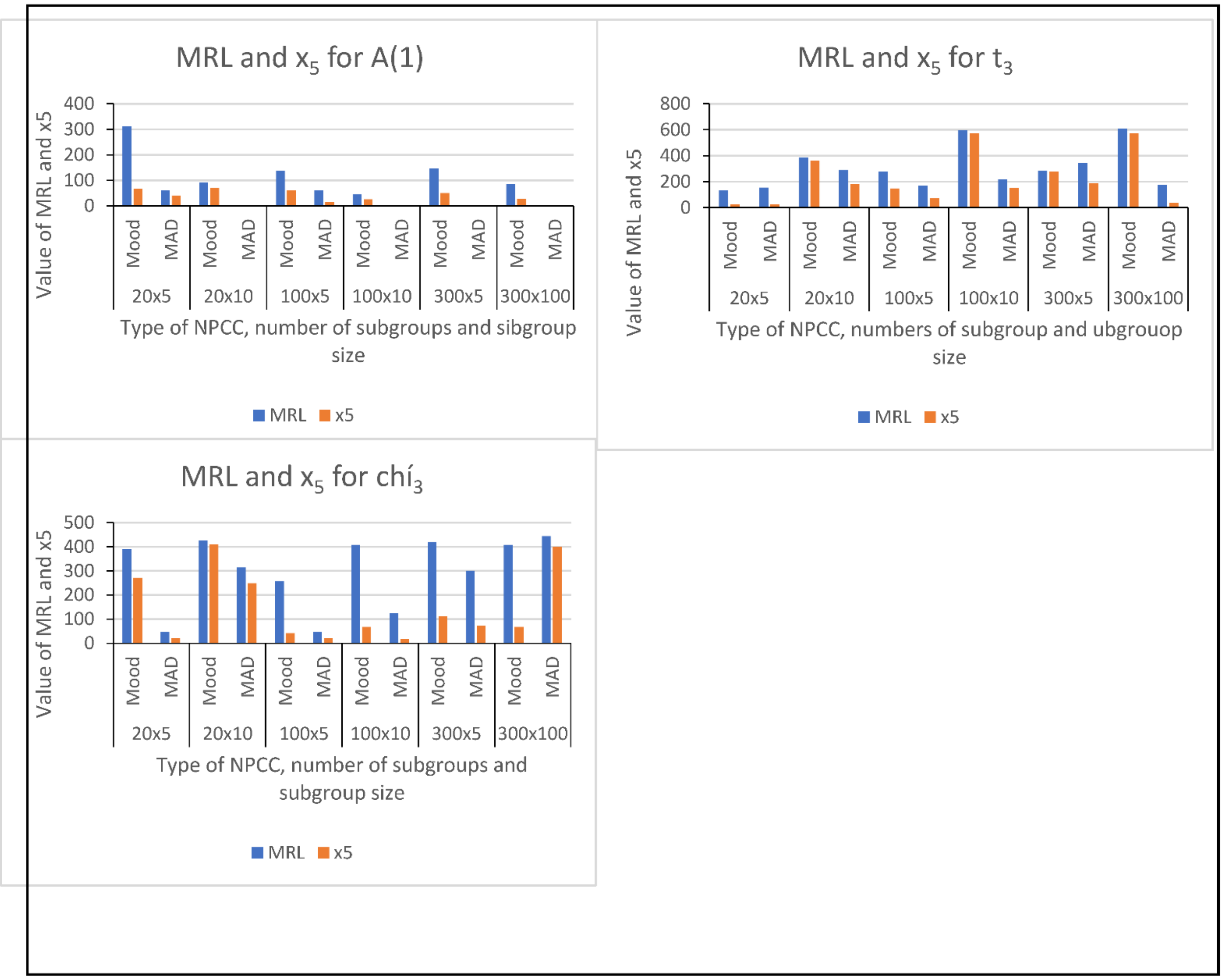

Figure A10. Values of MRL(0) and x_5 for NPCCs for monitoring process dispersion—dependent data, Student, chi-squared distributions [51].

Appendix D

Table A1. Statistics for construction of NPCC based on Mood statistic $M_{30,5}$ and for NP-CUSUM. Source: own research.

Subgroup j	$\bar{x}_j$	$M_{30,5}$ (Mood Statistics)	MW (NP-CUSUM)	SMW (NP-CUSUM)	S_i (NP-CUSUM)
1	33.16	1174.66	3	0.11	0
2	31.36	899.02	1	−1.30	0
3	31.92	977.42	2	−0.60	0
4	32.70	1082.47	4	0.81	0.31
5	32.24	1023.00	3	0.11	0
6	32.94	1119.67	3	0.11	0
7	32.36	1032.80	3	0.11	0

Table A1. *Cont.*

Subgroup j	$\bar{x}_j$	$M_{30,5}$ (Mood Statistics)	MW (NP-CUSUM)	SMW (NP-CUSUM)	S_i (NP-CUSUM)
8	32.58	1063.31	4	0.81	0.31
9	29.06	735.09	3	0.11	0
10	33.34	1176.87	5	1.51	1.01
11	33.70	1233.01	5	1.51	2.02
12	33.14	1148.03	4	0.81	2.33
13	32.84	1103.16	3	0.11	1.93
14	33.06	1135.09	5	1.51	2.94
15	31.72	950.92	1	−1.30	1.14
16	31.36	894.30	1	−1.30	0
17	31.92	969.34	1	−1.30	0
18	32.04	985.78	1	−1.30	0
19	32.88	1107.92	4	0.81	0.31
20	32.36	1034.10	1	−1.30	0
21	32.36	1031.58	3	0.11	0
22	31.80	952.66	0	−2.00	0
23	33.14	1197.67	1	−1.30	0
24	32.62	1071.81	3	0.11	0
25	32.38	1038.84	1	−1.30	0
26	32.14	1003.35	1	−1.30	0
27	31.86	961.09	1	−1.30	0
28	32.34	1029.27	3	0.11	0
29	32.44	1043.82	3	0.11	0
30	32.20	1011.78	2	−0.60	0

References

1. Jarošová, E.; Noskievičová, D. *Pokročilejší metody statistické regulace procesů (Advanced Methods of Statistical Process Control)*, 1st ed.; Grada Publishing: Praha, Czech Republic, 2015; pp. 139–144.
2. Qiu, P. *Introduction to Statistical Process Control*, 1st ed.; Chapman and Hall/CRC: Boca Raton, FL, USA, 2014; pp. 73–254.
3. Shewhart, W.A. *Statistical Methods from the View-Points of Quality Control*, 1st ed.; Dover Publication: New York, NY, USA, 1986.
4. Bakir, S.T. Classification Of Distribution-Free Quality Control Charts. In Proceedings of the Annual Meeting of the American Statistical Association, Atlanta, GA, USA, 5–9 August 2001.
5. Hollander, M.; Wolfe, D.A.; Chicken, E. *Nonparametric Statistical Methods*, 3rd ed.; John Wiley & Sons: Hoboken, NJ, USA, 2014; pp. 744–759.
6. Dempír, J.; Dohnal, L. Některé robustní postupy určení střední hodnoty a rozptýlení souboru výsledků a jejich použití (Some Robust Methods for Determination of Mean and Variance and Their Application). *Klin. Biochem. Metab.* **2005**, *13*, 139–144.
7. Woodall, W.H.; Montgomery, D.C. Research Issues and Ideas in Statistical Process Control. *J. Qual. Technol.* **1999**, *3*, 376–386. [CrossRef]
8. Figueiredo, F.; Gomes, M.I. Monitoring Industrial Processes with Robust Control Charts. *REVSTAT Stat. J.* **2009**, *7*, 151–170.
9. Koutras, M.V.; Triantafyllou, I.S. A General Class of Nonparametric Control Charts. *Qual. Reliab. Eng. Int.* **2018**, *34*, 427–435. [CrossRef]
10. Chakraborti, S.; Van Der Laan, P.; Bakir, S.T. Nonparametric Control Charts: An Overview and Some Results. *J. Qual. Technol.* **2001**, *33*, 304–315. [CrossRef]
11. Coelho, M.; Chakraborti, S.; Graham, M.A. A Comparison of Phase I Control Charts. *SAJIE* **2015**, *26*, 178–190. [CrossRef]
12. Das, N. A Note on the Efficiency of Nonparametric Control Chart for Monitoring Process Variability. *Econ. Qual. Control* **2008**, *23*, 85–93. [CrossRef]
13. Chakraborti, S.; Van de Wiel, M.A. A Nonparametric Control Chart Based on the Mann-Whitney Statistic. In *IMS Collection*; Institute of Mathematical Statistics: Beachwood, OH, USA, 2008; pp. 156–172. [CrossRef]

14. Oprime, P.C.; de Toledo, J.C.; Gonzáles, M.O.A.; Chakraborti, S. Methods for Determining the Control Limits of Nonparametric Charts for Monitoring Location and Scale. *Gest. Prod.* **2016**, *23*, 146–164. [CrossRef]
15. Adekeye, K.S.; Azubuike, P.I. Derivation of the Limits for Control Chart Using the Median Absolute Deviation for Monitoring Non-Normal Process. *Math. Stat. J.* **2012**, *8*, 37–41. [CrossRef]
16. Laha, A.K.; Gupta, D.; Choubey, A. Performance of Nonparametric Control Charts. IIMA Working Papers; Indian Institute of Management Ahmedabad, Research and Publication Department. 2005. Available online: http://hdl.handle.net/11718/1918 (accessed on 1 April 2022).
17. Pappanastos, E.A.; Adams, B.M. Alternative Designs of the Hodges-Lehmann Control Chart. *J. Qual. Technol.* **1996**, *28*, 223–231. [CrossRef]
18. Alloway, J.A.; Raghavachari, M. Control Chart Based on the Hodges-Lehmann Estimator. *J. Qual. Technol.* **1991**, *23*, 336–347. [CrossRef]
19. Aparisi, F.; Lillo, S.; Bartolucci, F. An Empirical Likelihood Based Control Chart for the Process Mean. Technical Report. 2005. Available online: https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.622.2767&rep=rep1&type=pdf (accessed on 1 April 2022).
20. Yang, S.-F.; Lin, J.S.; Cheng, S.W. A New Nonparametric EWMA Sign Control Chart. *Expert Syst. Appl.* **2011**, *38*, 6239–6243. [CrossRef]
21. Albers, W.; Kallenberg, W.C.M. Normal Control Charts with Nonparametric Safeguard. *Stat. Neerl.* **2009**, *63*, 63–81. [CrossRef]
22. Montgomery, D.C. *Introduction to Statistical Quality Control*, 8th ed.; John Wiley & Sons: Hoboken, NJ, USA, 2019; p. 768.
23. Tukey, J.W. A Survey of Sampling from Contaminated Distributions. In *Contributions to Probability and Statistics*; Oklin, I., Ed.; Standford University Press: Stanford, CA, USA, 1960.
24. Jacobs, D.C. Statistical Process Control: Watch for Nonnormal Distribution. *Chem. Eng. Progr.* **1990**, *86*, 19–27.
25. Woodall, W.H. Controversies and Contradictions in Statistical Process Control (with discussion). *J. Qual. Technol.* **2000**, *32*, 341–378. [CrossRef]
26. Park, C.; Reynolds, M.R. Nonparametric Procedures for Monitoring a Location Parameter Based on Linear Placement Statistics. *Seq. Anal.* **1987**, *6*, 303–323. [CrossRef]
27. Gibbons, J.D.; Chakraborti, S. *Nonparametric Statistical Inference*, 5th ed.; Chapman & Hall: Boca Raton, FL, USA, 2011; pp. 168–211.
28. Willemain, T.R.; Runger, G.C. Designing Control Charts Using an Empirical Reference Distribution. *J. Qual. Technol.* **1996**, *28*, 31–38. [CrossRef]
29. Reynolds, M.R. A Sequential Signed-Rank Test for Symmetry. *Ann. Stat.* **1975**, *3*, 382–400. [CrossRef]
30. Bakir, S.T.; Reynolds, M.R. A Nonparametric Procedure for Process Control Based on Within-Group Ranking. *Technometrics* **1979**, *21*, 175–183. [CrossRef]
31. Yashchin, E. Analysis of Cusum and Other Markov-Type Control Scheme by Using Empirical Distribution. *Technometrics* **1992**, *34*, 54–63. [CrossRef]
32. Abid, M.; Nazir, H.Z.; Riaz, M.; Lin, Z. Investigating the Impact of Ranked Set Sampling in Nonparametric CUSUM Control Charts. *Qual. Reliab. Eng. Int.* **2017**, *33*, 203–214. [CrossRef]
33. Hackl, P.; Ledolter, J.A. Control Chart Based on Ranks. *J. Qual. Technol.* **1991**, *23*, 117–124. [CrossRef]
34. Amin, R.; Searcy, A.J. A Nonparametric Exponentially Weighted Moving Average Control Scheme. *Commun. Stat. Simul. Comput.* **1991**, *20*, 1049–1072. [CrossRef]
35. Yue, J.; Liu, L. Multivariate Nonparametric Control Chart with Variable Sampling Interval. *Appl. Math. Model.* **2017**, *52*, 603–612. [CrossRef]
36. Haq, A. A New Nonparametric EWMA Control Chart for Monitoring Process Variability. *Qual. Reliab. Eng. Int.* **2017**, *33*, 1499–1512. [CrossRef]
37. Abbas, Z.; Nazir, H.Z.; Abid, M.; Akhtar, N.; Riaz, M. Enhanced Nonparametric Control Charts under Simple and Ranked Set Sampling Schemes. *Trans. Inst. Meas. Control.* **2020**, *42*, 2744–2759. [CrossRef]
38. Raza, M.A.; Nawaz, T.; Aslam, M.; Bhatti, S.H.; Sherwani, R.A.K. A New Nonparametric Double Exponentially Weighted Moving Average Control Chart. *Qual. Reliab. Eng. Int.* **2020**, *36*, 68–87. [CrossRef]
39. Boone, J.M.; Chakraborti, S. Two Simple Shewhart-Type Multivariate Nonparametric Control Charts. *Appl. Stoch. Models Bus. Ind.* **2012**, *28*, 130–140. [CrossRef]
40. Murakami, H.; Matsuki, T. A Nonparametric Control Chart Based on the Mood Statistic for Dispersion. *Int. J. Adv. Manuf. Technol.* **2010**, *49*, 757–763. [CrossRef]
41. Zhang, C.; Chen, N.; Zou, C. Robust Multivariate Control Chart Based on Goodness-of-Fit Test. *J. Qual. Technol.* **2016**, *48*, 139–161. [CrossRef]
42. Bakir, S.T.; Prater, T.; Kiser, S. A Simple Nonparametric Quality Control Chart for Monitoring Students' GPAs. *SOP Trans. Stat. Anal.* **2015**, *1*, 8–16. [CrossRef]
43. Graham, M.A. Theory and Applications of Univariate Distribution-Free Shewhart, CUSUM and EWMA Control Charts. M.Sc. Thesis, University of Pretoria, Faculty of Natural & Agricultural Sciences, Pretoria, South Africa, 3 September 2011.
44. Yang, S.-F.; Cheng, S.W. A New Non-Parametric CUSUM Mean Chart. *Qual. Reliab. Eng. Int.* **2011**, *27*, 867–875. [CrossRef]
45. Wang, D.; Zhang, L.; Xiong, Q. A Non-Parametric CUSUM Control Chart Based on the Mann—Whitney Statistic. *Commun. Stat.–Theory Methods* **2013**, *46*, 4713–4725. [CrossRef]

46. Graham, M.A.; Chakraborti, S.; Human, S.W. A Nonparametric Exponentially Weighted Moving Average Signed-Rank Chart for Monitoring Location. *Comput. Stat. Data Anal.* **2011**, *55*, 2490–2503. [CrossRef]
47. Lu, S.L. An Extended Nonparametric Exponentially Weighted Moving Average Sign Control Chart. *Qual. Reliab. Eng. Int.* **2014**, *31*, 3–13. [CrossRef]
48. Abbas, N.; Zafar, R.F.; Riaz, M.; Hussain, Z. Progressive Mean Control Chart for Monitoring Process Location Parameter. *Qual. Reliab. Eng. Int.* **2013**, *29*, 357–367. [CrossRef]
49. Abbasi, S.; Miller, A.; Riaz, M. Nonparametric Progressive Mean Control Chart for Monitoring Process Target. *Qual. Reliab. Eng. Int.* **2013**, *29*, 1069–1080. [CrossRef]
50. Abu-Shawiesh, M.O.A. A Simple Robust Control Chart Based on MAD. *J. Math. Stat.* **2008**, *4*, 102–107. [CrossRef]
51. Smajdorová, T. Využití Neparametrických Metod Ve Statistické Regulaci Procesů (Application of Non-Parametric Methods in Statistical Process Control). Ph.D. Thesis, VSB—Technical University of Ostrava, Ostrava, Czech Republic, 24 June 2019.
52. Chen, G. The Mean and Standard Deviation of the Run Length Distribution of X¯ Charts when Control Limits Are Estimated. *Stat. Sin.* **1997**, *7*, 789–798.
53. Chakraborti, S. Run Length Distribution and Percentiles: The Shewhart Chart with Unknown Parameters. *Qual. Eng.* **2007**, *19*, 119–127. [CrossRef]
54. Low, C.K. The Revised M-of-K Runs Rules Based On Median Run Length. M.Sc. Thesis, Universiti Sains Malaysia, Penang, Malaysia, May 2011.
55. Smajdorová, T.; Noskievičová, D. Methodology for the Application of Nonparametric Control Charts into Practice. *Emerg. Sci. J.* **2020**, *4*, 272–282. [CrossRef]