WILEY

Statistical testing for sufficient control chart performances during monitoring of grouped processes

Kevin Nikolai Kostyszyn ⬤ | **Tobias Claus Brandstätter** | **Thomas Vollmer** |
Robert Schmitt

Department for Production Quality,
Fraunhofer Institute for Production
Technology IPT, Aachen, Germany

Correspondence
Kevin Nikolai Kostyszyn, Department for
Production Quality, Fraunhofer Institute
for Production Technology IPT, Aachen,
Germany.
Email: kevin.kostyszyn@ipt.fraunhofer.de

Abstract

With ISO 7870-8, a standardized application of charting techniques for short
runs and small mixed batches was presented in 2017. Similar to various scientific
approaches, it requires that sample values from grouped processes follow nearly
identical distributions. In practice, however, there tend to be differences between
distribution parameters. Moreover, equal parameters do not ensure that distributions are properly aligned to the center line and control limits of the chart. These
facts can lead to undesired control chart performances which can be expressed
by average run lengths (ARL) during in-control and out-of-control conditions. In
this work, a statistical test for sufficient control chart performances during monitoring of grouped processes based on preliminary samples is proposed. Control
chart performances are defined as sufficient when they deviate within acceptable
ranges from usual performances during single process monitoring in mass production. The ARL resulting from estimated distributions and planned production
sequences is used as test statistic and calculated via the Markov chain approach.
Exemplary tests are executed for scenarios with individuals and cumulated sum
(CUSUM) charts. A simulative determination of error rates resulting from the
ARL-based testing demonstrates its effectiveness in testing for sufficient control
chart performances compared to an indirect testing with Levene's test and a one-
way analysis of variance (ANOVA).

KEYWORDS
average run length, control chart performances, grouped processes monitoring, ISO 7870-8,
short run SPC

1 | INTRODUCTION

The growing demand for customized products results in an increasing proportion of small batch production runs. Hence,
the implementation of suitable quality control techniques such as statistical process control (SPC) is gaining importance.
SPC techniques for small batch productions are often associated with the term *short run SPC*. Short runs are characterized
by situations in which products are produced in small quantities. Castillo et al[1] differentiate between two cases: short runs
in nonrepetitive and short runs in repetitive production processes.

Nonrepetitive processes refer to production with completely different machines and setups. Overviews of suitable SPC techniques such as control charts with modified control limits, self-starting control charts, and change-point models are provided by Capizzi and Masarotto,[2] Elam,[3] and Marques et al.[4]

Repetitive processes are characterized by the fact that the same machine is used to produce different products and quality characteristics without requiring major setup changes. Assuming a similar behavior of repetitive processes, sample values can be plotted within a joint control chart. For this, new developed chart types are presented in the scientific literature such as Q-charts proposed by Quesenberry[5,6] and enhanced by Castillo and Montgomery[7] or the t-chart proposed by Gu et al.[8] These have some drawbacks that include difficulty in interpretation of charting statistics and worse performance than the control chart types, which perform well in large batch and mass production such as Shewhart, cumulative sum (CUSUM), and exponentially weighted moving average (EWMA) charts. Control chart performances can be expressed by the average run length (ARL), which is the average number of samples required to receive an alarm. Typically, high control chart performances mean a large ARL during in-control conditions and small ARLs during out-of-control conditions. In order to ensure industrial acceptance, the application of proven chart types is preferable.

Charting techniques for short runs and small mixed batches based on Shewhart-styled control charts are proposed under ISO 7870-8,[9] which was published in 2017 and is partly based on the British standard BS5702-3.[10,11] Their application is explained in three steps. In the first two steps, repetitive processes are identified and grouped. This includes the identification of process influencing characteristics and the determination if differences in characteristics cause processes to behave significantly different. In the third step, the monitoring of grouped processes begins. Plotted samples are measurements of quality characteristics or process parameters, in the following briefly denoted as characteristics. Thus, groups of processes can also be denoted as groups of characteristics. Samples are required to have a size of one and to be approximately normally distributed. In case of different aim values or process spreads, sample values are transformed to deviations from their aim values and normalized through division by their standard deviations.

Alternative grouping approaches to ISO 7870-8 can be found in the scientific literature. These can be based on coding systems such as those proposed by Al-Salti and Statham[12] and Zhu et al[13] or on clustering algorithms as introduced by Wiederhold.[14] Shewhart-styled or similar charting techniques for grouped characteristics are also presented by Bothe,[15] Evans and Hubele,[16] Kimbler and Sudduth,[17] Koh et al,[18,19] Lin et al,[20] Koons and Luner,[21] and Wheeler.[22]

A main problem with these charts is the requirement that samples from different processes follow nearly identical distributions. Through this, it shall be indirectly ensured that resulting control chart performances are similar to usual performances during single process monitoring in large batch or mass production. Thus, it is widely proposed to statistically test distributions from grouped processes for equal means and variances based on preliminary samples.[12–14,16,23] In practice, however, parameters tend to be different. Even in case of equal parameters, distributions are not necessarily aligned to the center line and control limits of the control chart. The reason is that the transformation of samples as well as the construction of the control chart rely on predefined aim values, experience, or distribution parameters, which are estimated based on strongly limited sample values. Possible consequences are undesired control chart performances. Quesenberry[24] shows that estimated control limits can lead to drastic decreases of control chart performances and thus may cause high financial losses, for example, due to machine downtime. Hence, before monitoring a group of processes in a joint control chart, it is desirable to statistically evaluate the control chart performances that may be achieved.

Since there exists no suitable evaluation approach, a new statistical test for sufficient control chart performances based on preliminary samples is proposed as follows. Testing results serve as decision base for or against a monitoring of a formed group of processes in a joint control chart. It can be combined with several existing grouping approaches and control chart types. In the following, the testing is presented as an intermediate step of the procedure according to ISO 7870-8. The standard itself proposes to support the process grouping with suitable statistical analyses besides simulations, experiments, or expert knowledge. Sufficient control chart performances are defined as acceptable deviations from usual performances during single process monitoring in mass production for in-control conditions and relevant process shifts. The ARL resulting from estimated distributions and the planned production sequence is taken as test statistic and calculated via the Markov chain approach. The testing method is applied but not limited to exemplary scenarios with Shewhart's individuals and CUSUM charts plotting samples of size one. Its effectiveness in testing for sufficient control chart performances is demonstrated through simulative determinations of error rates (type I and type II) and compared to the indirect testing with Levene's test[25] and the one-way analysis of variance (ANOVA).[26]
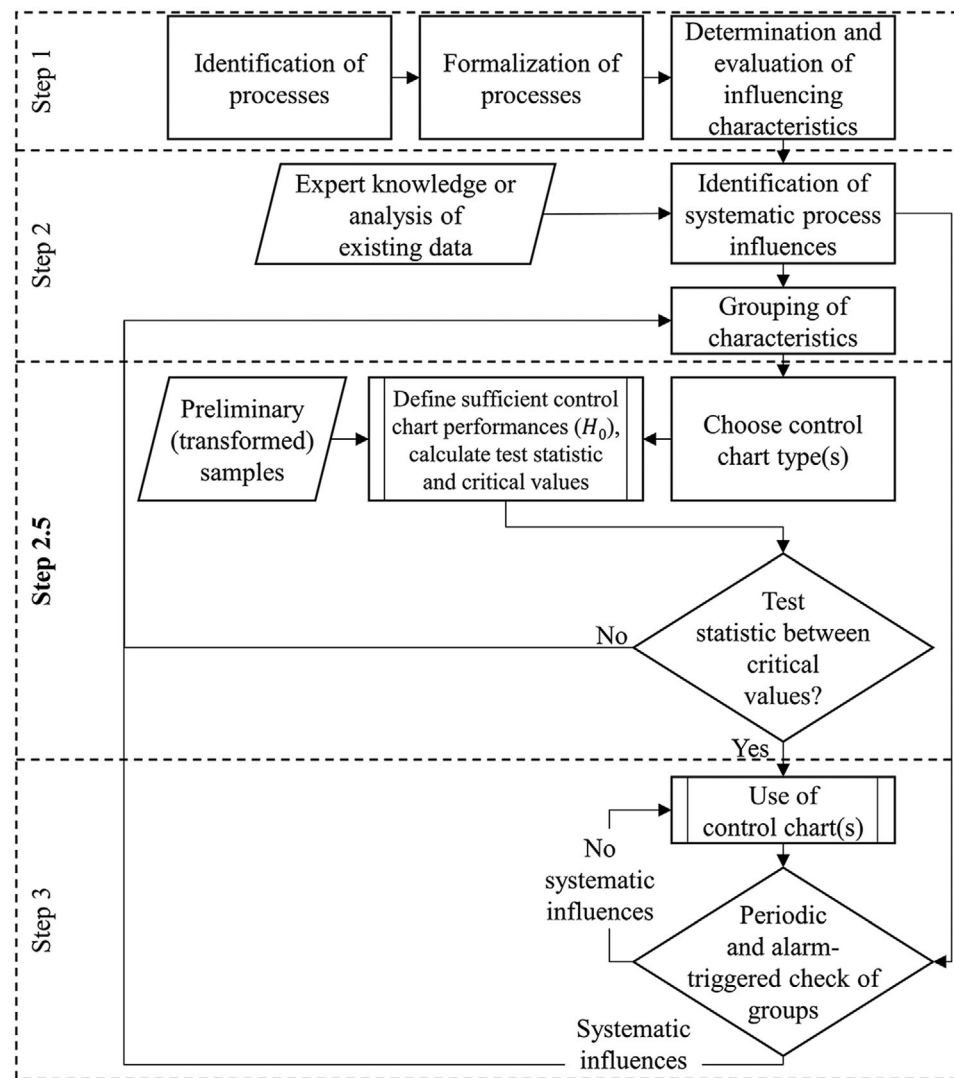
FIGURE 1 Procedure according to ISO 7870-89 with new intermediate step 2.5

2 ⎸ ARL-BASED TESTING FOR SUFFICIENT CONTROL CHART PERFORMANCES

The integration of the developed testing method into the procedure of ISO 7870-8 as an intermediate step 2.5 is depicted in Figure 1. After repetitive processes are identified and grouped, the desired control chart types are selected as they determine achievable control chart performances. For example, an individuals chart performs well in quickly detecting large process shifts. A CUSUM chart is preferable if small process shifts have to be detected early.[26] ISO 7870-8 recommends the use of individuals, moving mean, and moving range charts.

After selection, conditions for sufficient control chart performances are formulated as null hypothesis H_0 and the test statistic as well as critical values are calculated. The estimated ARL is used as test statistic and results from estimated distributions and the planned production sequence. As according to the ISO standard, it is assumed that samples from the same process follow an identical normal distribution. However, normal distributions from different processes may be nonidentical. For the estimation of parameters, at least two preliminary samples per process during in-control conditions are required.

The statistical test is two-tailed, leading to two critical values. At the end of step 2.5, the determined test statistic is compared to critical values. If it falls between the critical values, it is expected that the combination of grouped characteristics and the control chart types will lead to sufficient control chart performances. Otherwise, this expectation must be rejected and the validity of grouping conditions is recommended to be revised.
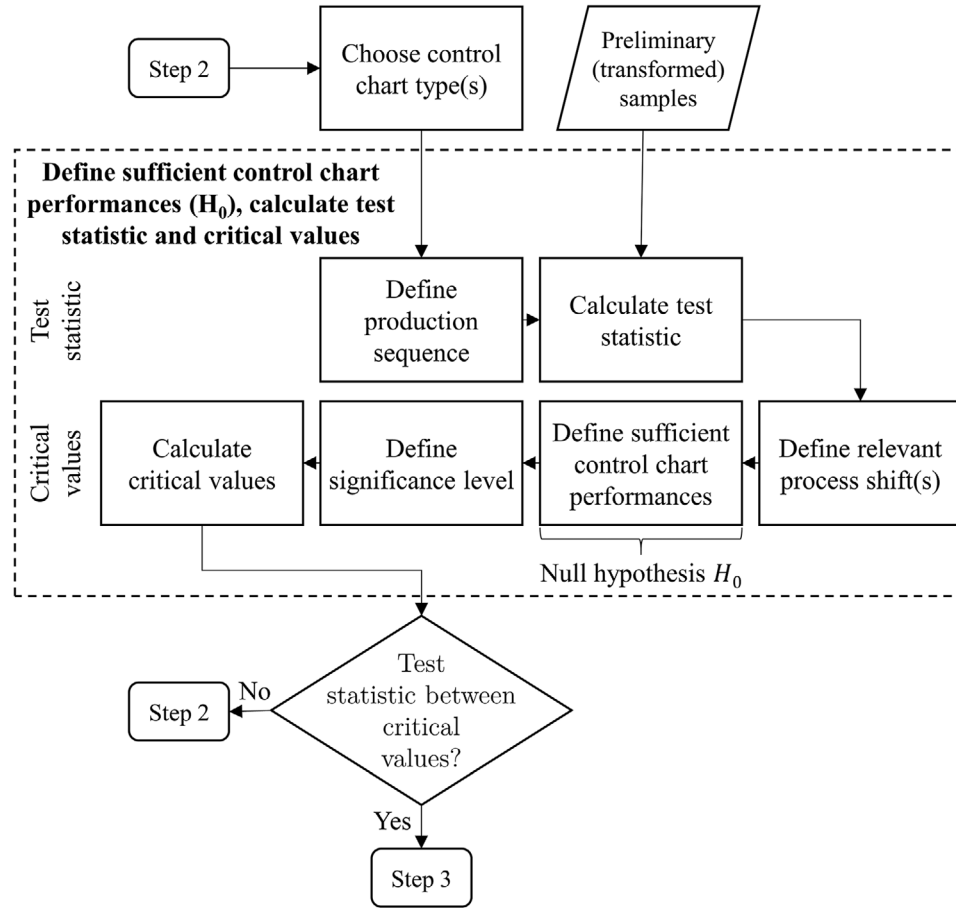
FIGURE 2 Subprocedure *Define sufficient control chart performances (H0), calculate test statistic and critical values* of step 2.5

The subprocedure *Define sufficient control chart performances (H_0), calculate test statistic and critical values* is depicted in detail in Figure 2. It can be divided into two different tasks. The first task is to calculate the test statistic. In the second task, critical values are calculated which also includes the formulation of the null hypothesis H_0. All single substeps are described below.

2.1 | Define production sequence

For the developed testing method, we consider that samples from different processes tend to follow different normal distributions. Thus, the planned production sequence determines the sequence of distributions from which plotted sample values are drawn. Considering that the sequence has an impact on the ARL-based test statistic, it must be previously defined. It is assumed that production sequences are infinitely often repeated.

A sequence can be simply described with a vector

$$q = (q_1, q_2, \dots, q_V). \tag{1}$$

Each of the V vector elements refers to one of m grouped characteristics. For example, considering $m = 2$ grouped characteristics, an example production sequence can be defined with $q = (1, 1, 1, 2)$. This means that the first characteristic is produced and measured three times before the second one is produced and measured. Afterwards, the sequence starts again from its beginning. In this way, different proportions of different produced characteristics can be considered.

2.2 | Calculate test statistic

The test statistic ARL_{est} is the estimated ARL, which results from the chosen control chart type, the grouped distributions with estimated parameters, and the previously defined production sequence. At the beginning, the estimated mean $\bar{x}_i$ and standard deviation s_i for each characteristic i are calculated. It is

$$\bar{x}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} x_{i,j}, \qquad (2)$$

$$s_i = \sqrt{\frac{1}{n_i - 1} \sum_{j=1}^{n_i} \left(x_{i,j} - \bar{x}_i\right)^2}, \qquad (3)$$

where $n_i \geq 2$ is the number of preliminary obtained samples and $x_{i,j}$ is the jth preliminary sample of characteristic i.

2.2.1 | Calculation of ARL_{est} for single process monitoring ($m = 1, q = (1)$) via the Markov chain approach

With the Markov chain approach, the selected control chart is interpreted to determine a discrete state space. With every new plotted sample, the control chart can change from its current state to another state. For example, the individuals chart is characterized by two different states. The first state refers to a sample lying between the control limits and the second state is achieved when one of both control limits is exceeded by a sample. More complex state spaces result from Shewhart charts applying further sensitizing pattern tests such as proposed by ISO 7870-2.[27] Champ and Woodall[28] describe a procedure of how to derive all available states for different combinations of applied rules. Brook and Evans[29] and Woodall[30] describe state spaces for one and two-sided CUSUM charts based on artificial discretization.

Assuming the existence of $h + 1$ different states, probabilities $p_{a,b}$ referring to the transition from state a to state b can be summarized in a transformation matrix T:

$$T = \begin{pmatrix} Q & R \\ 0^T & 1 \end{pmatrix} = \begin{pmatrix} p_{1,1} & p_{1,2} & \cdots & p_{1,h} & p_{1,h+1} \\ p_{2,1} & p_{2,2} & \cdots & p_{2,h} & p_{2,h+1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ p_{h,1} & p_{h,2} & \cdots & p_{h,h} & p_{h,h+1} \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix}. \qquad (4)$$

For the calculation of ARL_{est}, each probability is determined by the estimated distribution of the sample values. The $(h + 1)$-th state represents the out-of-control condition of the chart. Taking the example of the individuals chart, this would be the second state. It is called absorbing state as the state cannot be left anymore. Thus, the lower right probability of T is set to $p_{h+1,h+1} = 1$. Brook and Evans[29] show that the ARL of the control chart can be calculated based on the submatrix Q. It is the first element of the vector

$$M = \sum_{w=1}^{\infty} \left(Q^{w-1} R w\right) 1$$

$$= \sum_{w=1}^{\infty} \left(Q^{w-1} (1 - Q) w\right) 1 \qquad (5)$$

$$= (I - Q)^{-1} 1.$$

The matrix I represents the identity matrix and 1 is a column vector with elements all equal to one. The index w of the infinitive sum represents the run length, which typically starts from one. $(Q^{w-1} R)1$ is a vector in which the first element describes the probability of changing from the initial state to the absorbing state after w plotted samples. Within the infinitive sum, each of these vectors is weighted with the run length w. Thus, the first element of M is the test statistic
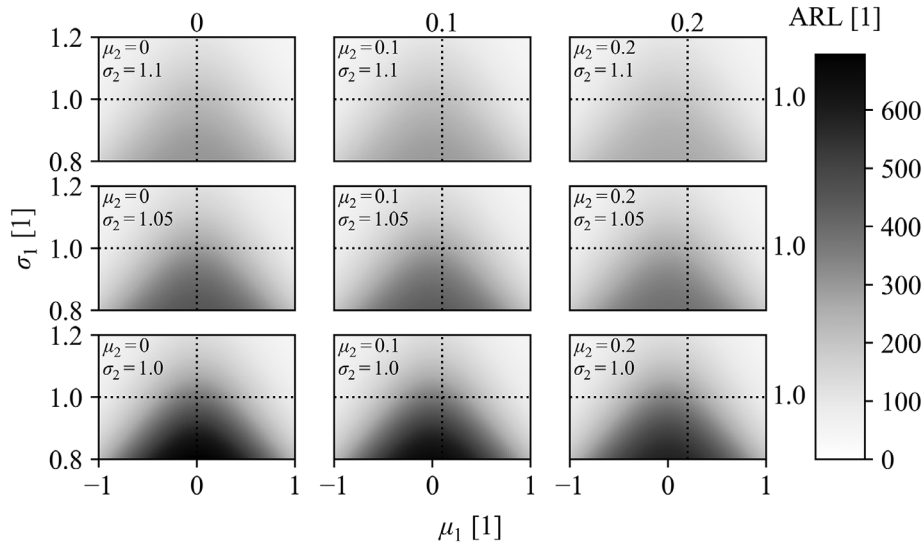
FIGURE 3 ARLs resulting from two grouped processes monitored in an individuals chart with control limits $CL = \pm 3$ and a production sequence $q = (1, 2)$

ARL_{est}. A simple example can be given for an individuals chart with samples drawn from a standard normal distribution. With only one state next to the absorbing state, it is $Q \approx 99.73\%$. This leads to $M = ARL_{est} \approx 370.37$.

2.2.2 | Calculation of ARL_{est} for monitoring of grouped processes $(m \geq 2, q = (q_1, q_2, \ldots, q_V), V \geq 2)$ via the Markov chain approach

Let Q_v with $v = 1, 2, \ldots, V$ describe the transformation submatrix of characteristic q(v) within the production sequence. Each submatrix is derived based on the discrete state space given by the control chart and on the estimated sample distributions of grouped characteristics. Let further $Q_{c,d}$ be a product of matrices defined with

$$Q_{c,d} = \begin{cases} I & c > d \\ Q_c \cdot Q_{c+1} \cdot \ldots \cdot Q_d & 1 \leq c \leq d \end{cases}. \tag{6}$$

The vector M can now be formulated in a similar way as shown by Brook and Evans.[29] The only difference is that Q matrices can change for every new plotted sample as determined by the infinitive production sequence q and the estimated sample distributions of characteristics:

$$M = \sum_{v=1}^{V} (m_v) \mathbf{1} = \sum_{v=1}^{V} \sum_{w=0}^{\infty} \left(Q_{1,V}^w Q_{1,v-1} (I - Q_v)(wV + v) \right) \mathbf{1}$$

$$= (I - Q_1 Q_2 \ldots Q_V)^{-1} (I + Q_1 + Q_1 Q_2 + \cdots + Q_1 Q_2 \ldots Q_{V-1}) \mathbf{1}. \tag{7}$$

The transformation of the formula to the simplified expression is shown in detail in the last section (Appendix).

Example ARLs resulting from known distribution parameters of two grouped characteristics monitored in an individuals and a CUSUM chart are visualized in Figures 3 and 4. Each diagram shows a heatmap consisting of 100×100 ARLs generated for different normal distributions $N(\mu_1, \sigma_1^2)$ and $N(\mu_2, \sigma_2^2)$. For both cases, the production sequence is set to $q = (1, 2)$, meaning that both characteristics are produced and measured in an alternating sequence.

The control limits of the individuals chart are set to $CL = \pm 3$. The monitored statistic are individual sample values x_w, which are directly drawn from both distributions according to the production sequence. For the CUSUM chart, the monitored statistics are

$$C_w^+ = \max \left(0, \ x_w - k + C_{w-1}^+ \right),$$

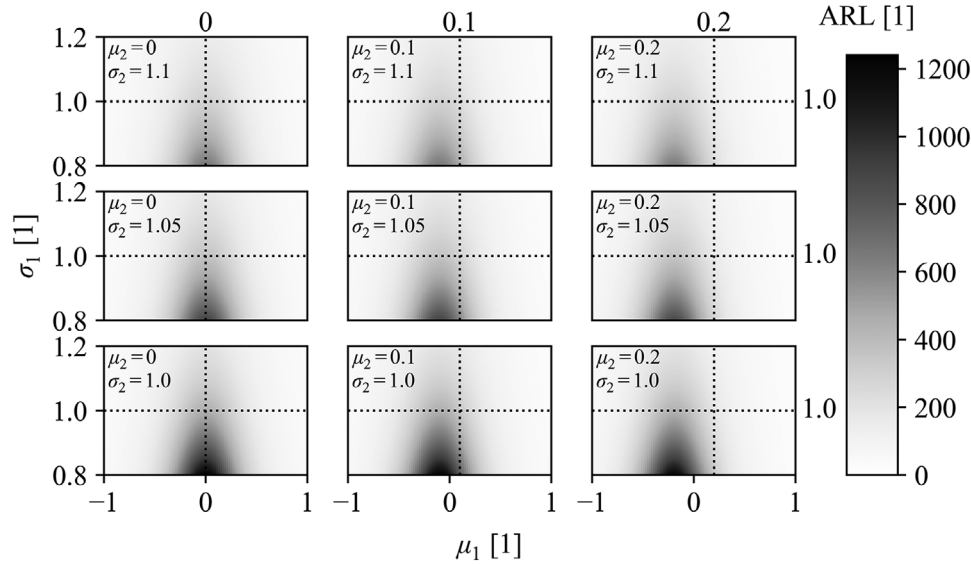$$C_w^- = \min \left(0, \ x_w + k + C_{w-1}^- \right). \tag{8}$$

FIGURE 4 ARLs resulting from two grouped process monitored in a double-sided CUSUM chart with reference value $k = 0.5$, limits $h = 5$, discretization parameter $t = 15$, and a production sequence $q = (1, 2)$

An out-of-control condition is met when $C_w^+ \geq h$ or $C_w^- \leq -h$. For the values in Figure 4, it is $h = 5$ and the reference value k is set to 0.5, which is typically chosen when process shifts of ± 1 are desired to be quickly detected.[26] As a CUSUM chart does not naturally provide discrete zones that can be used for the determination of a discrete state space, an artificial discretization is required. For this, the discretization parameter proposed by Woodall[30] is set to $t = 15$.

In each lower left diagram of both figures, the ideal ARL resulting from standard normal distributions $N(0, 1)$ is located at the intersection of the dotted lines. For the individuals chart, it is $ARL_{ideal,0} = 370.37$ and for the CUSUM chart it is $ARL_{ideal,0} = 452.41$. It is obvious that small changes of parameters lead to small deviations from these values. The same is visible for shifted distributions. The intersections in the lower middle and lower right diagrams show ARLs after ideal process shifts meaning that original standard normal distributions have shifted by a positive value of 0.1 and 0.2 to $N(0.1, 1)$ and $N(0.2, 1)$. For the individuals chart, it is $ARL_{ideal, 0.1} = 352.93$ and $ARL_{ideal,0.2} = 308.43$ and for the CUSUM chart it is $ARL_{ideal, 0.1} = 340.55$ and $ARL_{ideal,0.2} = 188.31$. Also in these cases, small differences of distribution parameters partly lead to similar ARLs.

Considering that small deviations from ideal ARLs for in-control and out-of-control conditions can be acceptable from industrial perspective, conditions for sufficient control chart performances are formulated as null hypothesis H_0 within the following two substeps.

2.3 | Define relevant process shifts

As grouped distributions with unequal parameters lead to deviations from ideal ARLs, the user can define relevant process shifts. Corresponding control chart performances are desired to be similar to ideal ARLs. In this approach, process shifts are understood as absolute. After a process shift δ, distribution means change to

$$\mu_{i,\delta} = \mu_i + \delta. \tag{9}$$

All relevant shifts can be summarized in a shift vector d with

$$d = (\delta_1, \delta_2, \dots, \delta_D). \tag{10}$$

If process shifts are expected to be random, the shift vector can be filled with values covering a wide range of possible shifts, for example, $d = (-3, -2.5, -2, -1.5, -1, 0, 1, 2, 3)$.

2.4 | Define acceptable deviations from ideal ARLs

In this substep, the understanding of sufficient control chart performances is formulated as null hypothesis H_0 of the statistical test. At first, a maximum acceptable deviation $r \in (0, 1)$ from ideal ARLs during in-control conditions and out-of-control conditions is defined. Considered out-of-control conditions are determined by the relevant shifts in d. The null hypothesis H_0 is defined as follows:

$$
\begin{aligned}
& ARL_0 \geq (1 - r) \, ARL_{ideal,0} \\
& \wedge \, ARL_0 \leq (1 + r) \, ARL_{ideal,0} \\
& \wedge \, ARL_{\delta_1} \leq (1 + r) \, ARL_{ideal,\delta_1} \\
& \wedge \, ARL_{\delta_2} \leq (1 + r) \, ARL_{ideal,\delta_2} \\
& \vdots \\
& \wedge \, ARL_{\delta_D} \leq (1 + r) \, ARL_{ideal, \delta_D}.
\end{aligned}
\tag{11}
$$

It implies that ARLs during out-of-control conditions are desired to be smaller than ideal ARLs while accepting a small range of greater ARLs. The ARL during in-control conditions is limited in both directions. While it is typically desired to be as large as possible, the upper limit is set in order to consider the continuous development of the ARL with increasing or decreasing process shifts. A very large ARL during in-control conditions would also lead to large ARLs for process shifts that are not defined as relevant but are between the zero shift and the nearest relevant process shift in positive or negative direction. Thus, with the upper limit, very large ARLs for small process shifts in positive and negative directions can be prevented. The alternative hypothesis H_1 applies in any situation in which at least one of the inequations is not true.

2.5 | Define significance level

As typical for statistical hypothesis tests, a significance level α must be defined in order to determine the rejection area of the statistic distribution, which results from a true null hypothesis. In the following, a conventional value of $\alpha = 5\%$ is chosen.

2.6 | Calculate critical values

Since the developed testing is two-tailed, two critical values c_1 and c_2 must be determined:

$$
\begin{aligned}
& P(ARL_{est} \leq c_1) = \frac{\alpha}{2}, \\
& P(ARL_{est} \geq c_2) = 1 - \frac{\alpha}{2}.
\end{aligned}
\tag{12}
$$

Thus, H_0 is not rejected as long as the following condition is met:

$$
c_1 < ARL_{est} < c_2.
\tag{13}
$$

Both critical values c_1 and c_2 are derived through simulations written in Python. Within each simulation loop, random sample distribution parameters for each characteristic are generated until a group of distributions is found, which fulfils the null hypothesis. For the calculation of ARLs, the selected control chart type and production sequence q are considered. For each characteristic i, n_i random preliminary samples are drawn from its representing distribution. The samples are then used to calculate estimated distribution parameters and the test statistic ARL_{est}.

After multiple simulation loops, calculated statistics are sorted to an ascending order. c_1 equals the value that is greater than $\frac{\alpha}{2}$ of all statistic values. c_2 is greater than a percentage share of $1 - \frac{\alpha}{2}$. For the following application scenarios, each simulation runs through 50 000 loops. Past experimental calculations have shown that resulting critical values mostly vary by less than 1%.

TABLE 1 Preliminary sample values from three grouped drilling processes (grouped x-positions of drill holes)

Group members	Measured x-positions of drill holes $x_{i,j}$ [mm]	Exp. mean μ_{exp} [mm]	Exp. std. dev. σ_{exp} [mm]	Transformed values ($x_{i,j}$ − μ_{exp})/ σ_{exp} [mm]	Est. mean $\bar{x}_i$ [mm]	Est. std. dev. s_i [mm]
1	5.28, 5.59, 4.91	5.00	0.20	1.40, 2.95, −0.45	1.30	1.70
2	2.92, 3.22, 2.68	3.00	0.30	−0.27, 0.73, −1.07	−0.20	0.90
3	7.40, 7.77, 7.45	7.50	0.40	−0.25, 0.67, −0.13	0.10	0.50
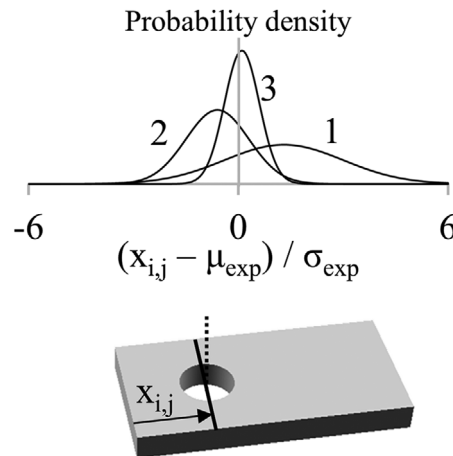
FIGURE 5 Estimated normal distributions of measured and transformed x-positions from three grouped drilling processes (see Table 1)

3 | APPLICATION SCENARIOS AND TESTING POWER

For an exemplary application of the ARL-based testing method, drilling processes for three different products are grouped. Measured x-positions of drilled holes are planned to be plotted within a joint control chart. Three preliminary sample values from each of three processes are listed in Table 1 (see also Figure 5).

x-positions of each process are assumed to be normally distributed. Since there are different expected means and standard deviations, all measured positions are transformed. Estimated means and standard deviations are derived based on Equations (2) and (3). With the application of the testing method, we investigate if sufficient control chart performances can be expected for the use of individuals or CUSUM charts. All testing parameters and results are given in Table 2. Although only three preliminary sample values per process are available, further testing results are generated for

TABLE 2 Testing parameters and results for different numbers of preliminary samples ($r = 10\%, \alpha = 5\%$)

Scenario	Control chart Type	Numbers of preliminary samples n_i	Chart-specific param.	Prod. seq. p_s	Test statistic ARL_{est}	Null hypothesis H_0	Critical values c_1	c_2	H_0 rejected?
1		(3, 3, 3)				$ARL_0 \geq 333.26$	16.14	8567405759105.84	No
2	Individuals	(5, 5, 5)	$CLs = \pm 3$	(1, 2, 3)	16.15	$^\wedge ARL_0 \leq 407.44$	24.47	21192943.25	Yes
3		(10, 10, 10)				$^\wedge ARL_{-2} \leq 6.93$	42.15	89283.80	Yes
						$^\wedge ARL_2 \leq 6.93$			
4		(3, 3, 3)	$k = 0.5,$			$ARL_0 \geq 407.17$	15.66	525214324176.66	No
5	CUSUM	(5, 5, 5)	$h = 5,$	(1, 2, 3)	38.42	$^\wedge ARL_0 \leq 497.65$	24.73	6554765.24	No
6		(10, 10, 10)	$t = 15$			$^\wedge ARL_{-1} \leq 11.43$	43.63	67843.61	Yes
						$^\wedge ARL_1 \leq 11.43$			

TABLE 3 Rates of error type I while testing for sufficient control chart performances with the test statistic ARL_{est} ($\alpha = 5\%$) and a combination of Levene's test and the one-way ANOVA ($\alpha_{Eq} = 2.53\%$)

						Error [%]	
Scenario	μ_{min} [mm]	μ_{max} [mm]	σ_{min} [mm]	σ_{max} [mm]	Fulfill H_0 ?	$err_{I,ARL}$	$err_{I,Eq}$
1						5.07	90.06
2						4.93	98.08
3						5.11	99.69
4	−3	3	0	2	Yes	5.08	82.89
5						5.29	94.66
6						4.94	98.71

varying numbers of preliminary samples resulting in six different scenarios. For every scenario, it is assumed that estimated distribution parameters are the same as in Table 1. The specific example from Table 1 is considered with scenarios 1 and 4.

As indicated by the production sequence, we assume that different products are produced and measured one by one. Measured ARLs are calculated based on estimated parameters given in Table 2. Individuals charts are typically applied if only large shifts shall be quickly detected. Thus, relevant shifts are set to $d = (-2, 2)$ as indicated by the null hypothesis. CUSUM charts offer high performances for smaller shifts. For them, relevant shifts are set to $d = (-1, 1)$. As numbers of preliminary samples vary among different scenarios, corresponding critical values are derived. Through comparison of test statistics with critical values, it can be concluded that the null hypotheses are only rejected in case of larger numbers of preliminary samples. For scenarios 1, 4, and 5, considered charts are expected to provide sufficient control chart performances.

Considering that sufficient control chart performances are a crucial decision criterion for the monitoring of grouped processes, the ARL-based testing method seems to deliver a useful decision base. For a quantitative assessment of the testing effectiveness, error rates of type I and type II ($err_{I,ARL}$, $err_{II,ARL}$) are determined based on simulations hereafter. In order to demonstrate that the direct testing for sufficient control chart performances based on the test statistic ARL_{est} is more effective than the indirect approach requiring equal distribution parameters, the null hypotheses shown in Table 2 are also tested with a combination of Levene's test and the one-way ANOVA. Resulting error rates are denoted as $err_{I,Eq}$ and $err_{II,Eq}$. With consideration of error cumulation, the significance level for each of both tests is set to $\alpha_{Eq} = 1 - \sqrt{1 - \alpha}$. Each simulation result presented in the following is based on a parameter scenario selected from Table 2. Within 50 000 simulation loops, normal distributions $N(\mu_i, \sigma_i^2)$ with random parameters $\mu_i \in [\mu_{min}, \mu_{max}]$ and $\sigma_i \in [\sigma_{min}, \sigma_{max}]$ are generated. Considering the scenario-specific control chart type and production sequence, it is determined whether generated distributions lead to sufficient control chart performances or not.

For the determination of error type I rates, only distributions leading to sufficient control chart performances are required. Therefore, in each loop, new distributions are randomly generated until a suitable set is found. Chosen intervals for random parameter generation are expected to cover nearly all suitable combinations of distribution parameters. The determination of error type II rates requires the generation of distributions leading to insufficient control chart performances. For random parameter generation, we choose once narrow intervals and once wide intervals.

After drawing n_i random sample values from corresponding distributions $N(\mu_i, \sigma_i^2)$, the ARL-based testing method as well as the combination of Levene's test and the one-way ANOVA are applied. Errors of both types are detected through comparison of sample-based testing results with the known control chart performances based on known distributions $N(\mu_i, \sigma_i^2)$. In the following both Tables 3 and 4, resulting error rates of type I and II are listed. Error rates of type II are also converted to powers expressing the probability of detecting real process shifts.

In both tables, the ARL-based testing mostly outperforms the combination of Levene's test and the one-way ANOVA when testing for sufficient control chart performances. Error rates of type I are always smaller with differences of up to more than 90%. Differences increase with increasing numbers of preliminary samples. As expected, $err_{I,ARL}$ is always close to 5% as determined by the significance level α. The powers of the new statistical test are mostly higher except for the CUSUM chart in scenarios 5 and 6 for distributions with $\mu_i \in [-0.5, 0.5]$ and $\sigma_i \in [0.5, 1.5]$. In all other cases, powers are higher up to more than 60%.

TABLE 4 Rates of error type II while testing for sufficient control chart performances with the test statistic ARL_{est} ($\alpha = 5\%$) and a combination of the Levene's test and the one-way ANOVA ($\alpha_{Eq} = 2.53\%$)

Scenario	μ_{min} [mm]	μ_{max} [mm]	σ_{min} [mm]	σ_{max} [mm]	Fulfill H_0 ?	Error [%] $err_{II,ARL}$	$err_{II,Eq}$	Power [%] $1 - err_{II,ARL}$	$1 - err_{II,Eq}$
1						82.47	91.90	17.53	8.10
2						78.90	83.80	21.10	16.20
3	−0.5	0.5	0.5	1.5	No	70.43	70.92	29.57	29.08
4						91.55	91.80	8.45	8.20
5						88.96	83.69	11.04	16.31
6						82.44	71.03	17.56	28.97
1						18.43	82.27	81.57	17.73
2						10.33	65.21	89.67	34.79
3	−1.5	1.5	0.3	3.3	No	4.79	35.17	95.21	64.83
4						27.55	83.03	72.45	16.97
5						15.67	65.81	84.33	34.19
6						8.91	36.67	91.09	63.33

4 | CONCLUSION

The presented method allows to test for sufficient control chart performances before deciding for a monitoring of grouped processes within a joint control chart. Sufficient control chart performances are defined based on acceptable deviations from usual performances during single processes monitoring for in-control and selected out-of-control conditions. For the calculation of the ARL-based test statistic, a Markov chain approach was presented, which considers both the selected control chart type as well as the planned production sequence. For industrial application, the method can be integrated as an intermediate step into the procedure proposed by ISO 7870-8. Besides Shewhart-styled control charts, the application can be extended to other types such as CUSUM charts.

The application of the testing method was demonstrated for a group of drilling processes. Through simulations of exemplary scenarios, it was shown that an ARL-based testing for sufficient control chart performances can be more effective than the indirect approach requiring equal distribution parameters.

In future research, the method applicability will be extended to nonnormal distributions such as folded normal or Rayleigh distributions. Also, the consideration of random production sequences shall be enabled. For the determination of critical values, an alternative approach for random generation of distributions is required in order to reduce calculation time.

APPENDIX

In the following, the transformation of Equation (7) is demonstrated in detail. It is

$$m_v = \sum_{w=0}^{\infty} \left(Q_{1,V}^w Q_{1,v-1} (I - Q_v)(wV + v) \right)$$

$$= \left(Q_{1,v-1} - Q_{1,v} \right) v + Q_{1,V} \sum_{w=0}^{\infty} Q_{1,V}^w \left(Q_{1,v-1} - Q_{1,v} \right) (Vw + v + 1)$$

$$= \left(I - Q_{1,V} \right)^{-1} \left(\left(Q_{1,v-1} - Q_{1,v} \right) v + \sum_{w=0}^{\infty} \left(Q_{1,V}^w Q_{1,V} \left(Q_{1,v-1} - Q_{1,v} \right) \right) \right)$$

$$= \left(I - Q_{1,V}\right)^{-1} \left(\left(Q_{1,v-1} - Q_{1,v}\right) v + \left(I - Q_{1,V}\right)^{-1} Q_{1,V} \left(Q_{1,v-1} - Q_{1,v}\right) V\right)$$

$$= \left(I - Q_{1,V}\right)^{-2} \left(\left(I - Q_{1,V}\right)\left(Q_{1,v-1} - Q_{1,V}\right) v + Q_{1,V}\left(Q_{1,v-1} - Q_{1,v}\right) V\right), \tag{14}$$

where the infinite sum

$$\sum_{w=0}^{\infty} Q_{1,V}^{w} = \left(I - Q_{1,K}\right)^{-1} \tag{15}$$

is known as the Neumann series.[31] Now, putting (14) into (7), it is

$$M = \left(I - Q_{1,V}\right)^{-2} \sum_{v=1}^{V} \left(\left(I - Q_{1,V}\right)\left(Q_{1,v-1} - Q_{1,v}\right) v + Q_{1,V}\left(Q_{1,v-1} - Q_{1,v}\right) V\right) 1$$

$$= \left(I - Q_{1,V}\right)^{-2} \left(\left(I - Q_{1,V}\right)\left(I + Q_{1,1} + Q_{1,2} + \cdots + Q_{1,V-1} - Q_{1,V}V\right) + Q_{1,V}\left(I - Q_{1,V}\right) V\right) 1$$

$$= \left(I - Q_{1,V}\right)^{-1} \left(I + Q_{1,1} + Q_{1,2} + \cdots + Q_{1,V-1}\right) 1$$

$$= \left(I - Q_1 Q_2 \dots Q_V\right)^{-1} \left(I + Q_1 + Q_1 Q_2 + \cdots + Q_1 Q_2 \dots Q_{V-1}\right) 1. \tag{16}$$

ORCID
Kevin Nikolai Kostyszyn https://orcid.org/0000-0002-7477-7811

REFERENCES
1. Castillo ED, Grayson JM, Montgomery DC, Runger GC. A review of statistical process control techniques for short run manufacturing systems. *Commun Stat—Theor Meth*. 1996;25(11):2723-2737.
2. Capizzi G, Masarotto G. An enhanced control chart for start-up processes and short runs. *Qual Technol Quant Manage*. 2012;9(2):189-202.
3. Elam ME. Control charts for short production runs. In: Balakrishnan N, ed. *Wiley StatsRef: Statistics Reference Online*. Hoboken, New Jersey: Wiley; 2014.
4. Marques PA, Cardeira CB, Paranhos P, Ribeiro S, Gouveia H. Selection of the most suitable statistical process control approach for short production runs: a decision-model. *Int J Inform Educ Technol*. 2015;5(4):303-310.
5. Quesenberry CP. SPC Q charts for start-up processes and short or long runs. *J Qual Technol*. 1991;23(3):213-224.
6. Quesenberry CP. On properties of Q charts for variables. *J Qual Technol*. 1995;27(3):184-203.
7. Castillo ED, Montgomery DC. Short-run statistical process control: q-chart enhancements and alternative methods. *Qual Reliab Eng Int*. 1994;10(2):87-97.
8. Gu K, Jia X, You H, Zhang S. A t-chart for monitoring multi-variety and small batch production run. *Qual Reliab Eng Int*. 2014;30(2):287-299.
9. International Standard. *ISO 7870-8:2017: Control Charts: Part 8: Charting Techniques for Short Runs and Small Mixed Batches* (ISO 7870-8:2017). Beuth Verlag: Berlin; 2017.
10. British Standard. *BS 5702-3:2008: Guide to Statistical Process Control (SPC) Charts for Variables: Part 3: Charting Techniques for Short Runs and Small Mixed Batches*. Beuth Verlag: Berlin; 2008.
11. Schultz W. Die crux mit der streuung: sPC für kleine Losgrößen und Einzelfertigung. *QZ Qualität und Zuverlässigkeit*. 2018;63(10):44-46.
12. Al-Salti M, Statham A. The application of group technology concept for implementing SPC in small batch manufacture. *Int J Qual Reliab Manage*. 1994;11(4):64-76.
13. Zhu YD, Wong YS, Lee KS. Framework of a computer-aided short-run SPC planning system. *Int J Adv Manufact Technol*. 2007;34(3-4):362-377.
14. Wiederhold M. *Clustering of Similar Features for the Application of Statistical Process Control in Small Batch and Job production*. Aachen: Apprimus Wissenschaftsverlag; 2017.
15. Bothe DR. SPC for short production runs. *Proceedings of the IEEE National Aerospace and Electronics Conference*. Piscataway, New Jersey: IEEE; 1989:1960-1963.
16. Evans ME, Hubele NF. *A Case Study of Family Formation for Statistical Process Control in Small-Batch Manufacturing*. Dearborn, Michigan: Society of Manufacturing Engineers; 1993.

17. Kimbler DL, Sudduth BA. Using statistical process control with mixed parts in FSM. In: Gakkai SSJ, ed. *Proceedings of the 1992 Japan-USA Symposium on Flexible Automation*. New York, New York: ASME; 1992:441-445.

18. Koh CK, Chin JF, Kamaruddin S. Modified SPC for short run test and measurement process in multi-stations. *IOP Conf Ser: Mater Sci Eng*. 2018;328:012009.

19. Koh CK, Chin JF, Kamaruddin S. Modified short-run statistical process control for test and measurement process. *Int J Adv Manufact Technol*. 2019;100(5-8):1531-1548.

20. Lin S-Y, Lai Y-J, Chang SI. Shprt-run statistical process control: multicriteria part family formation. *Qual Reliab Eng Int*. 1997;13(1):9-24.

21. Koons GF, Luner JJ. SPC in low-volume manufacturing: a case study. *J Qual Technol*. 1991;23(4):287-295.

22. Wheeler DJ. *Short RunSPC*. Knoxville, Tennessee: SPC Press; 1991.

23. Jiang X, Zhao K, Tang L, Wang W. Research on multi-type & small batch oriented process quality control system under network environment. *2008 Asia Simulation Conference—7th International Conference on System Simulation and Scientific Computing*. Piscataway, New Jersey: IEEE; 2008:1050-1056.

24. Quesenberry CP. The effect of sample size on estimated limits for $\bar{X}$ and X control charts. *J Qual Technol*. 1993;25(4):237-247.

25. Levene H. Robust tests for equality of variances. In: Olkin I, ed. *Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling*. Redwood City, California: Stanford University Press; 1960:278-292.

26. Montgomery DC. *Introduction to Statistical Quality Control*. 7th ed. Hoboken, New Jersey: John Wiley & Sons; 2013.

27. International Standard. *ISO 7870-2:2013: Control Charts: Part 2: Shewhart Control Charts* (ISO 7870-2:2013). Berlin: Beuth Verlag; 2013.

28. Champ CW, Woodall WH. Exact results for Shewhart control charts with supplementary runs rules. *Technometrics*. 1987;29(4):393-399.

29. Brook D, Evans DA. An approach to the probability distribution of Cusum run length. *Biometrika*. 1972;59(3):539-549.

30. Woodall WH. On the Markov chain approach to the two-sided CUSUM procedure. *Technometrics*. 1984;26(1):41-46.

31. Werner D. *Funktionalanalysis*. 8th ed. Berlin, Heidelberg: Springer Spektrum; 2018.

AUTHOR BIOGRAPHIES

Kevin Nikolai Kostyszyn received his MSc degrees in automotive engineering and transport from RWTH Aachen University, Germany, and in mechanical engineering from Tsinghua University, Beijing, China. He has worked as research fellow at Fraunhofer Institute for Production Technology IPT and is currently finishing his doctoral studies at RWTH Aachen University with the research focus on statistical process control in short run production.

Tobias Claus Brandstätter obtained his MSc in mechanical engineering from RWTH Aachen University, Germany. At present, he is a research fellow at Fraunhofer Institute for Production Technology IPT. His main areas of research address the support of employees in manufacturing.

Thomas Vollmer obtained his doctorate from RWTH Aachen University, Germany. Since 2018, he has been head of Department for Production Quality at Fraunhofer Institute for Production Technology IPT. He previously worked in the domains of resource efficiency, production optimization, and networked production IT as research fellow and group leader.

Robert Schmitt has been a professor at RWTH Aachen University, Germany, since 2004. He is head of Metrology and Quality Management Department at the Laboratory for Machine Tools and Production Engineering (WZL) and member of the board of directors at the Fraunhofer Institute for Production Technology IPT where he is also head of Production Metrology and Production Quality Department.