

Precautionary Statistical Criteria in the Monitoring Quality of Technological Process

Eugenij Volodarsky^{1(✉)}, Zygmunt Warsza^{2(✉)}, Larysa A. Kosheva³,
and Adam Idzkowski⁴

¹ Department of Automation of Experimental Studies,
National Technical University of Ukraine “KPI”, Kiev, Ukraine
vet-1@ukr.net

² Industrial Research Institute of Automation and Measurement (PIAP),
Warsaw, Poland
zlw@op.pl

³ Department of Biocybernetics and Aerospace Medicine,
National Aviation University of Ukraine, Kiev, Ukraine
l.kosh@ukr.net

⁴ Faculty of Electrical Engineering,
Bialystok University of Technology, Bialystok, Poland
a.idzkowski@pb.edu.pl

Abstract. The quality of products depends on a stability of production process. In practice to identify reasons of the process degradation, the Shewhart control chart is typically used. At the construction of control X-charts, it is supposed that the dispersion of sample means in subgroups of data measured in process is caused by the influence of random factors and the limited sample size. In such cases it is an improbable event to obtain the output sample values, which are outside the interval $\pm 3\sigma$. Its appearance indicates the presence of systematic influence and it is the need to adjust the controlled parameters of technological process. Based on practical experience in the ISO 7870-2: 2013 standard it is recommended to pay attention to “... any unusual structure of data points, which may indicate about a manifestation of special (non-random) reasons”. In the numerical example presented in this work the analysis of a structure of points on the control chart showed the presence of non-random values, although if the sample mean values were within interval $\pm 3\sigma$. The indicator of existence of non-randomness was the probability that the minimum number of consecutive selective averages, which got to a certain area did not exceed 0.003. As a result of executed analysis the criteria are established and the algorithm is developed. It helped to identify the dysfunction of technological process at an early stage.

Keywords: Quality control · Technological process · Shewhart control chart

1 Introduction

A quality of production depends on the stability and absence of changes in a technological process. Therefore, it is important to identify timely the reasons of dysfunction of technological process and the signals corresponding to them. If the process is carried out in normal conditions then the scattering of parameters, which characterize the properties and quality of a product, depends only on the influence of random variables. The possible dispersion of parameters (their standard deviations) are usually standardized in the normalized conditions.

A stability control of results is based on a series of control procedures. In this case a number of observations and the intervals between them have to be established on the basis of relations between the rates of changes of measured statistical characteristics under the influence of different random variables. This should be realized in such a way that the influence of deviations in carrying out the technological process could be neglected.

As proven practice the control charts are widely used for the statistical control of stability and quality of processes. This method was developed by Shewhart [1]. The main idea of control charts is to divide observations in subgroups, in which variations due to random causes are only permitted. The differences between these subgroups may not only be caused by random specific causes and this must be identified by control charts [2]. For this purpose, a reference value is established. The deviations from this value are detected as observations.

Depending on a control algorithm, the warning and action signals are selected and the control limits are calculated. A warning signal is an event which testifies a confidence level greater than (0.95 ... 0.99) about the withdrawal of process from the statistically controlled conditions and about the requirement of technological process correction. An action signal is an event, which indicates to withdraw the process from the statistically controlled conditions with a confidence level of 0.997. The monitored results of a process are presented on the control charts after recording each current observation on them.

The practice of using charts for the control of process has shown that acceptable results are obtained when a number of elements in the subgroups is not more than 4 to 5. The number of elements has to be identical in the assumption of their normal distribution. Under this condition, the coefficients for calculating the control limits are derived. Since the control limits are used as empirical criteria for decision-making, it is allowed to ignore small deviations from normality.

If the volumes of experimental data are larger ($n > 10$) then the standard deviation (SD) adequately displays a scattering of results. It characterizes a stability of the controlled process. When the samples of small volume are considered then a sample range R_n (the absolute difference between the highest and lowest values of subgroup sample) gives better estimate of the scattering of results than the standard deviation and it is calculated more quickly [1]. In addition, for the evaluation of process stability it allows to have only two observations in each subgroup. That should often be enough due to a dynamics of process or an economic feasibility.

The sample range R_n and standard deviation σ is statistically connected. For normal distribution it is [3]

$$M\left(\frac{R_n}{\sigma}\right) = \alpha_n, \quad (1)$$

where: α_n – tabulated value depending on a number of elements n in sample, R_n - range of n -element sample.

Due to these statistics the mean value of R_n can be identified as

$$M(R_n) = \alpha_n \sigma. \quad (2)$$

In addition, as it can be seen from the expression (1), α_n is an unbiased estimate, and consequently, also, an unbiased estimate $M(R_n)$ is important and it may be taken as the centre of possible scattering of inspection results.

The range of possible values of ratio R_n/σ at a fixed value of n due to the influence of random variables and the limited sample size is also tabulated. Thus, there is a relationship

$$\text{var}\left(\frac{R_n}{\sigma}\right) = \beta_n, \quad (3)$$

which allows, for a given σ , to set the possible values of scattering amplitude R_n with respect to $M(R_n)$ in the form

$$\text{var}(R_n) = \beta_n \sigma. \quad (4)$$

The most common charts in the monitoring of process stability are:

- average value (\bar{X} -chart) and range (\bar{R} -chart) or sample standard deviation s ,
- individual measured values (X -chart) and moving range (R -chart).

2 Average Value \bar{X} Control Chart and Range \bar{R} Control Chart

The control chart of average values \bar{X} is used to demonstrate what the average value of process is and what its stability is. Moreover, it allows identifying variations between subgroups that cannot be explained only by the influence of random variables and their relation to the total variation of the mean. If this type of control has to be reliable, the samples should be stable for a time period between repeated measurements.

Range control chart \bar{R} identifies any undesirable variation within a subgroup and it is an indicator of variability of a controlled process. If \bar{R} – chart shows that the variation within a subgroup are not changed then it informs about the uniformity of process. It is necessary to analyze \bar{R} – chart prior to the analysis of \bar{X} – chart.

Since the average value \bar{X} control chart and the range \bar{R} control chart (sample or standard deviations) reflect the state of process through the spread (variability from unit to unit) and through the centre of location (the average value of process), they are always used inseparably. Thus, to ensure the stability of process is necessary, firstly, to monitor the changes of σ in time which can be caused by an influence of random variables and a dysfunctional process.

By using the expressions (2) and (3) one can determine the absolute value of the variation quantile span R_n as

$$M(R_n) \pm k(P)\text{var}(R_n), \tag{5}$$

where: $k(P)$ – coefficient depending on a value of confidence interval.

If the range is positive number then for the left quantile the condition should be satisfied

$$[M(R_n) - k(P)\text{var}(R_n)] > 0. \tag{6}$$

These relations form the basis for the construction of Shewhart control charts. To make a decision on the stability of technological process, there are introduced precautionary warning limits $k(P) = 2$ and action limits $k(P) = 3$. It corresponds to the probability of decision $P = 95\%$ and $R = 99.7\%$.

When a range control chart \bar{R} is constructed a tabulated value α_n is used as a centre line which in the standard [4] is indicated as d_2 . For example: when $n = 2$ in accordance with [3] $d_2 = 1.128$ and $M(R_n) = 1.128\sigma$ is taken as a centre line CL. The action limits in the \bar{R} control chart must be separated from the centre line by $\pm 3\text{var}(R_n)$. For example, the upper action limit is

$$UCL_a(n) = d_2\sigma + 3d_3\sigma = D_2\sigma, \tag{7}$$

where: d_3 corresponds to the value where β_n , taken from the same table, for $n = 3$, $d_3 = 0.853$.

Thus, calculated values $D_2 = (d_2 + 3d_3)$ are provided in the table [4]. Similarly, one can obtain for the lower action limit

$$LCL_a = d_2\sigma - 3d_3\sigma, \tag{8}$$

or

$$LCL_a = D_1\sigma, \tag{9}$$

where $D_1 = d_2 - 3d_3$.

After analyzing the relation (6) one can conclude that for $k(P) = 3$ it will be executed if $n \geq 7$ only. For $n < 7$ as the lower action limit of Shewhart chart is taken zero line.

To calculate the upper and lower warning limits, one can use the expressions (5) and (6). For example, if $k(P) = 2$

$$UCL_w = D_2(2)\sigma \tag{10}$$

where: $D_2(2) = d_2 + 2d_3$,

$$LCL_w = D_1(2)\sigma \tag{11}$$

where: $D_1(2) = d_2 - 2d_3$.

In this case Eq. (6) will only be carried out when $n \geq 4$. In other cases the numerical value of the lower warning limit is absent – it is replaced by 0.

The calculated values for the centre line $CL = d_2\sigma$, upper and lower action limits $LCL_a = D_1\sigma$, as well as upper and lower warning limits and $LCL_w = D_1(2)$, respectively, are used to build the Shewhart charts.

In Fig. 1 an example of construction \bar{R} – charts is presented. The ordinate axis represents the value of magnitude and on the horizontal axis are a number of observations of subgroup.

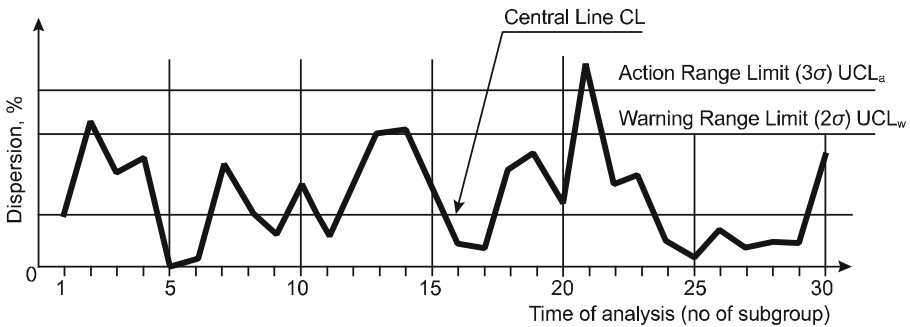


Fig. 1. Example of construction – Shewhart chart

The values of coefficients used to determine the control limits of Shewhart charts are presented in Table 1 [4]. The coefficients for calculating the warning limits are derived according to formulas

$$D_1(2) = d_2 - 2d_3, \tag{12}$$

$$D_2(2) = d_2 + 2d_3. \tag{13}$$

Assessing the bias stability of technological process should be carried out on \bar{X} – chart using a standard sample with a SD value of μ . As mentioned before, a requirement is the invariance of its characteristics in time. A monitoring is carried out with a standard sample several times ($n \geq 2$) and average \bar{x}_i is calculated for each subgroup.

In this case as the centre line in the construction of the \bar{X} – chart is used $CL = \mu$, in relation to which changes \bar{x}_i are considered in i -subgroups. The warning and action limits of are defined as

Table 1. The values of coefficients to determine the limit values of Shewhart charts

The coefficients to calculate the median line and the limits of efficiency			Coefficients for warning ranges		
Number of observations in subgroup	Coefficient for the centre line	Coefficient for the action range			
n	d_2	D_2	d_3	$D_1(2)$	$D_2(2)$
2	1.128	3.686	0.853	–	2.834
3	1.693	4.358	0.888	–	3.469
4	2.059	4.698	0.880	0.299	3.819
5	2.326	4.918	0.864	0.598	4.054

$$UCL = \mu + k(P) \frac{\sigma}{\sqrt{n}}, \tag{14}$$

$$LCL = \mu - k(P) \frac{\sigma}{\sqrt{n}}, \tag{15}$$

where: σ – standard deviation (normalized value) of estimated results in the control due to the influence of random variables.

3 Individuals X Control Chart and Moving Range R Control Chart

In some cases for practical or economic reasons there is no possibility to carry out the multiple observations and $n = 1$. In such situation, the moving range R_i may be used to monitor the stability of process. The μ value of standard sample does not need to be known accurately. However, the values of the samples must be stable. So-called initial “reference” measurement of performance of standard sample is realized with the result y_0 . Then, there is the difference between the result of the first measurements y_1 and y_0 , i.e. it is calculated the first implementation of bias (offset) results $\hat{\delta}_i = y_1 - y_0$. Subsequently, the offset is determined as the difference values obtained in the current and previous times. Thus, the moving range of i -th control chart estimated as

$$R_i = |\hat{\delta}_i - \hat{\delta}_{i-1}|. \tag{16}$$

A fragment of the moving range R -chart with the results of calculations [5] using the standard object with $\mu = 10.29$ is presented in Table 2.

Since the centre line CL at moving range R chart is the zero line then the previously received formula to calculate the action and warning limits should be modified. The starting point is known and the SD value of the process $\sigma_I = 0.06645$ is used. Then the action limits and warning limits are defined as

Table 2. Data to build the charts of moving range R

No of subgroup i	Result of analysis y_i	Estimate of bias $\hat{\delta}_i$	Moving range R_i
1	10.30	0.01	0.01
2	10.29	0.00	0.01
3	10.28	-0.01	0.02
4	10.30	0.01	0.01
5	10.29	0.00	0.00
6	10.29	0.00	0.09

$$\begin{aligned}
 UCL_a &= +3\sigma_I = 0.1994, & LCL_a &= -3\sigma_I = -0.1994, \\
 UCL_w &= +2\sigma_I = 0.1329, & LCL_w &= -2\sigma_I = -0.1329.
 \end{aligned}$$

In this case unlike the \bar{R} - charts, lower warning and action limits cannot be equal to zero because deviation from μ value can be positive or negative.

The centre line $M(\hat{\delta}) = 0$ is taken as the zero line to create the X -chart. Symmetrically in relation to it, the upper and lower warning limits are determined with factor $k(P) = 2$ and upper and lower action limits with factor $k(P) = 3$. Further actions and solutions taken for X -chart are similar as for R -chart. The Shewhart chart for the example from Table 2 is presented in Fig. 2 a,b.

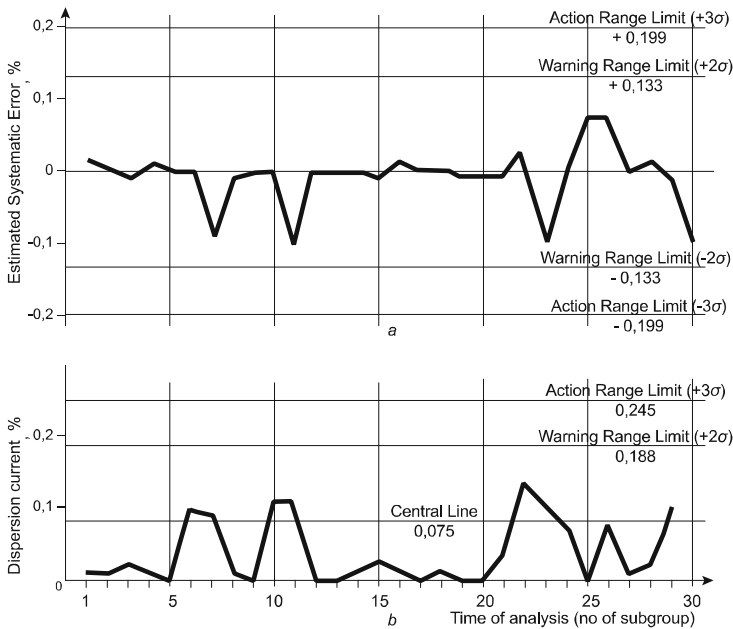


Fig. 2. Shewhart control charts for: (a) assessment of displacement $\hat{\delta}$ stability (b) current dispersion of results

4 Display of Process Instability

As it is presented in Fig. 2, there are some periods of time when the bias of process and the changes of range are low. There are other periods where results indicate an increased instability. This situation requires the identification of reasons that caused an increase of process instabilities over a certain period although the current values of the parameters are within the required limits.

When creating a control chart it is assumed that during monitoring the change in the X value (the dispersion of sample means as an estimate of bias) results from the influence of random factors and the limited sample size. In this case, the value outside $\pm 3\sigma$ is a low probable event. Its appearance indicates the presence of systematic influence, which leads to a process dysfunction and a change of its control. Based on practical experience in [4] is recommended to pay attention to "... any unusual structure of data points, which may indicate about a manifestation of the special (non-random) reasons". Such an event corresponds to the probability of 0.003 [4]. The observed situation can be attributed as "critical" which constitutes a violation of the conditions of process.

Figure 3 presents the probability of getting the results displayed in control zones A, B and C. It characterizes the relationship between σ and the number of observations n .

An indication of the influence of random variables is chaotic incidence of sub-groups results in all areas A, B and C of the Shewhart chart. Tracking the emergence of a systematic trend in a distribution of points corresponding to the mean values of samples (bias estimate) can be used as evidence of the trend of mean value of controlled process parameter.

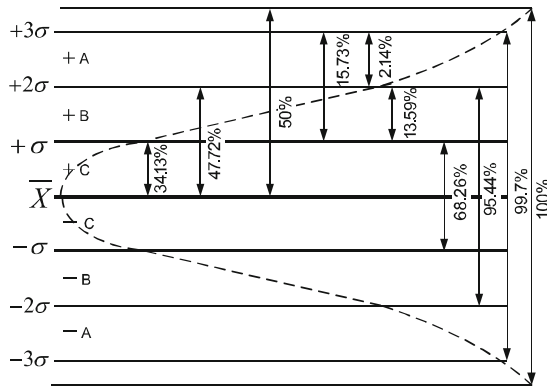


Fig. 3. The probability of getting the results into control zones A, B, C

The trends on the control chart, which emerged under the influence of special causes, can lead to a breakdown of process. They will be called as series criteria. In this case, the manifestation of a systematic influence on the background of random

dispersion is to find a sequence of a certain number of points on the control chart in one of zones or in area covering several zones. This approach allows to justify theoretically that the incidence of trend reports at an early stage about the possibility of violation of a correct process. In this way a sequence of control points on the chart located within the warning or action limit, the probability of which is less than 0.003, can be regarded as a manifestation of joint influence of random and systematic reasons. This property can be used as the basis for the creation of preventive warning criteria. It would allow to adjust the progress of process without waiting for situation when it could be disordered that the results would be outside the control limits.

Checkpoints should be sufficiently distant from each other in time and space that the effect of autocorrelation cannot be included [7]. In accordance with the multiplication theorem for independent events the probability of getting a normally distributed random variable in area A, B, C is equal to the product of individual probabilities

$$P(A_1, A_2, \dots, A_s) = P\left(\prod_{i=1}^s P(A_i)\right). \quad (17)$$

During analysis the criteria were established. According to them it can be found the probability of several independent, consecutive values in a particular field of control chart. It may indicate a trend of dysfunctional process. For example if we assume that the average value (bias estimate) of consecutive subgroups are independent random variables then the probability of getting the result for any subgroup above (or below) the centre line in each of zones A, B, C is 0.4986 approximately 0.5 (Fig. 3). Probability, that two consecutive sample values are e.g. above the centre line, is equal to $0.5 \cdot 0.5 = 0.25$.

It is necessary to find out what the minimum number of successive results, arranged in a row on one side of the center line of the control chart, corresponds to the probability of 0.003 (0.0027). It is the probability that a single sample value is not within the control limits $\pm 3\sigma$. It turns that this condition can be fulfilled by a sequence of nine points, i.e. probability that a series of nine control chart points will be on one side of central line is 0.00195. If this criterion is satisfied then a change in average value of the overall process can be considered.

Similar reasoning can be made for the possible sequences of points, which are located in corresponding zones of the control chart with a certain probability. One can define a set of preventive criteria when a dysfunctional process begins but results still do not go outside the warning limits or action limits. An example would be a situation when finding 15 consecutive results in a completely “safe” zones $\pm C$ cannot cause concern. However, the emergence of the next 16-th result in this zone corresponds to the probability 0.0023, which exceeds the value of 3σ . Therefore, this criterion is also “critical” as shown in Fig. 4.

With the use of similar analysis one can define a number of simple criteria shown in Table 3, which demonstrate that the mean value of controlled process is biased.

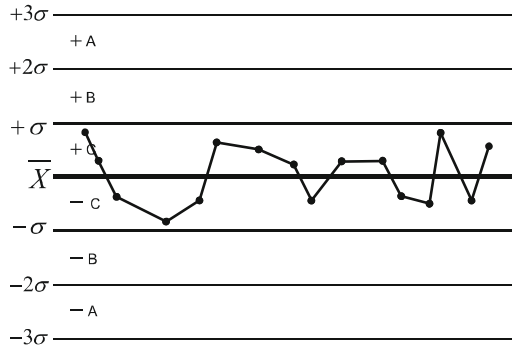


Fig. 4. The probability of warning signals and symptoms of the action

Table 3. Examples criteria that indicate the shift of mean value of controlled process variable

Situation	Number of consecutive points	Location of points	Probability of an elementary event in %
1	2	Only in zone plus A or minus A	2.14
2	3	Only in zone plus B or minus B	13.59
3	6	Only in zone plus C or minus C	34.13
	6	Permanent increase or decrease in points of any zone ^a	
4	7	Only in two zones[+A; +B] or [-A; -B]	15.73
5	8	Series arrangement only in zones [+A; +B] and [-A; -B], passing zones +C and -C	15.46
6	9	In any zones [+A; +B; +C] or [-A; -B; -C]	49.86
7	16	Only in zone + C or -C	68.26

^a Sample values are independent variables, which have the same distribution. When comparing the current value with the previous ones we will have two outcomes: more than “+” or less “-”. Under the influence of the random variable, signs will have the same frequency. Wald-Wolfowitz run test [6] establishes a relationship between the number of identical characters in a series and probability magnitude of the systematic influences.

5 Conclusions

In evaluating the stability of process it is possible an very early detection of trends leading to dysfunction of the process.

The considered approach and the resulting criteria for identifying trends can be treated as the basis of sequential analysis and the identification of the “critical” situation for the controlled process, depending on a location of area and a number of points in their sequence on the control chart.

The established laws for the control of individual sections of the X chart allow to introduce corrective actions rapidly, without waiting for the actual dysfunction of process parameters.

This area of research is promising and requires further detailed development and formalization of the results in the form of adaptive decision-making algorithms.

References

1. Wheeler, D.J., Chambers, D.S.: Understanding Statistical Process Control. Addison-Wesley Publishing Company (2010)
2. Nishina, K., Kuzuya, K., Ishi, N.: Reconsideration of Control Charts in Japan. *Front. Stat. Qual. Control* **8**, 136–150 (2005)
3. Gatti, P.L.: Probability Theory and Mathematical Statistics for Engineers. Taylor & Francis (2004)
4. ISO 7870-1, -2 ... -6: 2014. Shewhart control charts – part 1-6
5. ISO 5725-6: 1994 (reviewed in 2012) Accuracy (trueness and precision) of measurement methods and results. Part 6: Use in practice of accuracy values
6. Alhakim, A., Hooper, W.: A non-parametric test for several independent samples. *J. Nonparametric Stat.* **20**(3), 253–261 (2008)
7. Warsza, Z.L.: Evaluation of the type A uncertainty in measurements with autocorrelated observations. *J. Phys. Conf. Ser.* **459**(1), Article no 012035 (2013)