TECHNICAL TRANSACTIONS | **CZASOPISMO TECHNICZNE**
MECHANICS | MECHANIKA

4-M/2016

JACEK PIETRASZEK*, MARIUSZ KRAWCZYK**, ANDRZEJ SOBCZYK**,
EWA SKRZYPCZAK-PIETRASZEK***

THE DOMINANT FACTOR DETECTION IN THE SHAININ'S APPROACH

WYKRYWANIE DOMINUJĄCEGO CZYNNIKA ZA POMOCĄ PODEJŚCIA SHAININA

Abstract

Shainin's approach is a specific sequential heuristic aimed at finding and ranking the most important factors which impact the investigated process. The sequential aspect of the approach is simultaneously its strongest and weakest side, because just after detection of the most important factor, the further analysis is stopped without any additional cost. However, such a detection may take place at the end of the whole sequence. This paper tries to answer the question if the dominant factor may be hidden by interactions with other factors.

Keywords: *Shainin's approach*, *Red X*, *design of experiment*

Streszczenie

Podejście Shainina jest specyficzną heurystyką sekwencyjną nakierowaną na wykrywanie i rangowanie najważniejszych czynników wpływających na badany proces. Sekwencyjność podejścia jest jednocześnie jego najmocniejszą i najsłabszą stroną, gdyż po wykryciu najważniejszego czynnika cała dalsza analiza jest przerywana bez ponoszenia dodatkowych kosztów, ale wykrycie tego czynnika może nastąpić dopiero przy końcu całej sekwencji. Artykuł poszukuje odpowiedzi na pytanie, czy istnieje ryzyko maskowania istnienia czynnika dominującego wskutek działania interakcji.

Słowa kluczowe: *podejście Shainina*, *Red X*, *planowanie doświadczeń*

DOI: 10.4467/2353737XCT.16.238.5987

---

* D.Sc. Ph.D. Eng. Jacek Pietraszek, Assoc Prof., M.Sc. Eng. Mariusz Krawczyk, Institute of Applied Informatics, Faculty of Mechanical Engineering, Cracow University of Technology.

** D.Sc. Ph.D. Eng. Andrzej Sobczyk, Assoc Prof., Institute of Machine Design, Faculty of Mechanical Engineering, Cracow University of Technology.

*** Ph.D. Ewa Skrzypczak-Pietraszek, Chair and Department of Pharmaceutical Botany, Jagiellonian University Collegium Medicum.

1. Introduction

Even after a statistical process control (SPC) tuning, an industrial process is affected by instabilities induced by many known and unknown factors. Therefore, the process must be continuously tuned up to produce goods which meet consumer requirements. Many of those factors which control such processes and their proper settings are crucial for the final success. Some factors may be affected by instability or a drift which becomes a source of the whole process instability. A typical SPC reacts if the deviations of important parameters go over the trigger limit, however it is *a posteriori* operation and some scrap is unavoidable. Typically, the process is more sensitive to a few factors while the rest is less important. The selection of these factors is one of the most important phases in SPC [1].

In the 1930s and 1940s, the standard approach to make this selection was to use the Yates's factorial design [2] and to construct the Pareto ranking. In 1948 Placket and Burman proposed a more effective approach [3]: screening designs based on Hadamard matrices. All these approaches have the same requirement: all tests have to be conducted before the analysis may be processed. It means that the costs are constant and the engineering heuristic knowledge is not applied *a priori* to reduce the cost.

In the 1950s and 1960s, the aerospace engineer D. Shainin changed this requirement through his "Red X" approach. Its most spectacular application was related to the reliability assessment of the Lunar Expedition Module (NASA Apollo missions) produced by Grumman where Shainin was responsible for the quality [4].

The ideas which influenced the core concept of the Shainin's approach date back to the results of 19th century Italian sociologist W. Pareto, later in the 1950s enhanced by J. Juran [5]. He created the following statement [6]: "vital few and trivial many". This statement has been mistakenly named Pareto principle, however the particular form has been expressed by Juran. In the 1990s Juran tried to explain this mistake [7], but it was too late. The name "Pareto principle" appeared to be unchangeable.

2. Methods

A clear detection of the most dominant factor is shown in two of Shainin's heuristics phases: the"Component Search" and "Variable Search" developed in 1956 and 1973, respectively. The mathematical structure of both is the same, however in the"Component Search" swapping is applied to physical components, while in the "Variable Search" – to the settings of selected factors.

The analysis starts from the selection of "Green Y", the quantitative variable of interest which should be stabilized. Then two objects are selected, the best product (BoB – Best of Bests) and the worst product (WoW – Worst of Worsts) from a production lot. Next, both objects are twice disassembled and reassembled in the "Component Search" or the process is reset and tuned in the "Variable Search". After each rebuild/resetting, the "Green Y" is measured. The aim of this procedure is to determine the noise originated from the disassembly/reassembly process or adjustment devices. At this time, two subsets of three measurements are obtained.

The first analysis starts from two subsets $\mathbf{M}_{\mathrm{BoB}}$ and $\mathbf{M}_{\mathrm{WoW}}$, each containing three measurements. The noise analysis is provided based on median and ranges, as opposed to the classic mean and standard deviations. Shainin argued that such statistics are more robust to instability from outliers. The median and ranges are calculated for BoB:

$$\begin{aligned} m_{\mathrm{BoB}} &= \operatorname{median}(\mathbf{M}_{\mathrm{BoB}}) \\ d_{\mathrm{BoB}} &= \max(\mathbf{M}_{\mathrm{BoB}}) - \min(\mathbf{M}_{\mathrm{BoB}}) \end{aligned} \tag{1}$$

and for WoW:

$$\begin{aligned} m_{\mathrm{WoW}} &= \operatorname{median}(\mathbf{M}_{\mathrm{WoW}}) \\ d_{\mathrm{WoW}} &= \max(\mathbf{M}_{\mathrm{WoW}}) - \min(\mathbf{M}_{\mathrm{WoW}}) \end{aligned} \tag{2}$$

Next, the difference between the medians and average range are calculated:

$$\begin{aligned} D &= \left| m_{\mathrm{BoB}} - m_{\mathrm{WoW}} \right| \\ d &= \frac{\left| d_{\mathrm{BoB}} - d_{\mathrm{WoW}} \right|}{2} \end{aligned} \tag{3}$$

The median interval D is treated as the measure of a significance difference between BoB and WoW, while the average range d is treated as a measure of noise factors from the assembly process or adjustment devices. The ratio between these two variables i.e. D divided by d is the key value which decides about the further analysis. The critical value of the ratio is set by Shainin at 1.25. This is the first of "magic numbers" placed in this approach. The value was argued as median/range critical level equivalent for typical t Student test for equality of two means.

If the ratio is less than 1.25, it means that the noise from the assembly process/adjustment devices is too loud and it should be reduced first i.e. the assembly process/adjustment devices should be repeatable earlier than the main analysis starts. It should be noted that such a noise reduction may require additional expenses like e.g. repairing of machines in machining industry or significant changes of raw materials. It may be very difficult in e.g. biochemistry and biotechnology processes [8–10], where noise factors origin directly from raw materials and from the natural or semi-natural environment.

If the ratio is greater than 1.25, it means that the noise from the assembly process/adjustment devices is small enough and this noise will not mask the effects from controlled factors. If these conditions are satisfied, then the control lines may be calculated based on values from Eqs. 1–3, first for BoB:

$$\begin{aligned} \mathrm{UCL}_{\mathrm{BoB}} &= m_{\mathrm{BoB}} + 2.776 \cdot \frac{d}{1.81} \\ \mathrm{LCL}_{\mathrm{BoB}} &= m_{\mathrm{BoB}} - 2.776 \cdot \frac{d}{1.81} \end{aligned} \tag{4}$$

and next for WoW:

$$\begin{aligned} \mathrm{UCL}_{\mathrm{WoW}} &= m_{\mathrm{WoW}} + 2.776 \cdot \frac{d}{1.81} \\ \mathrm{LCL}_{\mathrm{WoW}} &= m_{\mathrm{WoW}} - 2.776 \cdot \frac{d}{1.81} \end{aligned} \tag{5}$$

The “magic” values, 2.776 and 1.81, are described by Bhote [11] as taken from *t* Student distribution, however the authors could not find any precise reference to any statistical source.

Subsequently, a specific control card is created based on mentioned control lines and two triplets associated with BoB and WoW. The next points on this card are taken from measurement made at the sequential swapping of components. After each swap, a decision needs to be made about the component/factor status: not-important, important (Pink X), the most important (Red X). The status is read from the mutual position of measurements on the control card (Fig. 1).

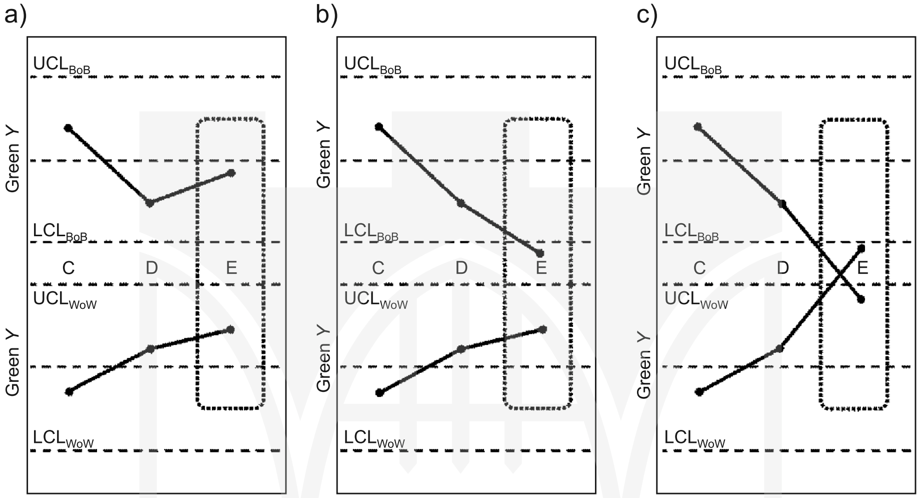

Fig. 1. Possible behavior of the response: a) not-important component/factor, b) important component/factor (Pink X), c) the most important component/factor (Red X)

If both measures are located inside respective control lines (Fig. 1a), it means the swapped component/factor is not important. If at least one measure is located outside its control lines but relative position to the second measure will not be replaced (Fig. 1b), it means the swapped component/factor is important (Pink X), but it does not explain all variability – other important factors are expected. If both measures are located outside their control lines and their relative position is swapped (Fig. 1c), it means the swapped component/factor is the most important (Red X).

The aim of the research was to check when Red X is detected and whether the interaction may mask the existence of Red X.

3. Results

The first check was conducted based on a fixed-effect model with an interaction for two factors A, R and additive term ε (Eq. 6):

$$\text{Green } Y = \mu + A + R + AR + \varepsilon \tag{6}$$

Factor *A* represents potential Red X while *R* describes effects of other components. The additive term ε describes the noise from the disassembly/assembly process and the additive μ describes the mean response. Two test objects, BoB and WoW are modelled at maximum contrasts i.e. at opposite effects:

$$\begin{aligned} \text{Green } Y_{\text{BoB}}(+1,+1) &= \mu + a + r + ar + \varepsilon_{\text{BoB}} \\ \text{Green } Y_{\text{WoW}}(-1,-1) &= \mu + (-a) + (-r) + ar + \varepsilon_{\text{WoW}} \end{aligned} \tag{7}$$

The responses after a swap of *A* component are described by Eq. 8:

$$\begin{aligned} \text{Green } Y_{\text{BoB}|A}(-1,+1) &= \mu - a + r - ar + \varepsilon_{\text{BoB}} \\ \text{Green } Y_{\text{WoW}|A}(+1,-1) &= \mu + a - r - ar + \varepsilon_{\text{WoW}} \end{aligned} \tag{8}$$

The Red X detection is made when – after a swap of *A* component – the response of BoB is worse than the response of WoW:

$$\text{Green } Y_{\text{BoB}|A} < \text{Green } Y_{\text{WoW}|A} \tag{9}$$

After substituting respective terms from Eq.8, the condition is transformed into:

$$a > r + \frac{\varepsilon_{\text{BoB}} - \varepsilon_{\text{WoW}}}{2} \tag{10}$$

The form of the expression shows that the interaction effect *AR* was completely eliminated. It means that the interaction cannot mask Red X detection. If the value of noise realizations are relatively small, the fact of Red X detection means that such factor explains more than 50% of the total variability.

4. Conclusions

The fixed-effect model with linear effects and an interaction was created to describe the process of detection of the most important factor in Shainin's Red X approach. The theoretical analysis reveals that the interaction does not interfere with the detection process, which means that the interaction existence or non-existence is not important for the detection procedure. Additionally, it was concluded that Red X factor explains more than 50% of the total variability if noise factors are relatively small i.e. *D*/*d* is far from the limit 1.25.

References

[1] Montgomery D.C., *Introduction to Statistical Quality Control*, John Wiley & Sons, Hoboken, 1997.

[2] Yates F., *Complex experiments*, J. Roy. Statist. Soc., vol. Suppl.2, 1935, 181-247.

[3] Plackett R.L., Burman, J.P., *The design of optimum multifactorial experiments*, Biometrika, vol. 33, 1946, 305-325.

[4] Shainin D., *Reliability: Managing a reliability program*, "Apollo lunar module engine exhaust products", Science, vol. 166, 1969, 733-738.

[5] Juran J.M., *Quality Control Handbook*, McGraw-Hill, New York 1951.
[6] Juran J.M., *Universals in Management Planning and Controlling*, The Management Review, vol. 43 (11), 1954, 748-761.
[7] *The Non-Pareto Principle; Mea Culpa*, [in:] Juran J.M. (Ed.), Juran on Quality by Design, Free Press, New York 1992, 68-71.
[8] Pietraszek J., Skrzypczak-Pietraszek E., *The Optimization of the Technological Process with the Fuzzy Regression*, Terotechnology, vol. 874, 2014, 151-155.
[9] Skrzypczak-Pietraszek E., Pietraszek J., *Seasonal Changes of Flavonoid Content in Melittis melissophyllum L. (Lamiaceae)*, Chemistry & Biodiversity, vol. 11 (4), 2014, 562-570.
[10] Skrzypczak-Pietraszek E., Slota J., Pietraszek J., *The influence of Lphenylalanine, methyl jasmonate and sucrose concentration on the accumulation of phenolic acids in Exacum affine Balf. f. ex Regel shoot culture*, Acta Biochimica Polonica, vol. 61 (1), 2014, 47-53.
[11] Bhote K.R., *World Class Quality: using design of experiment to make it happen*, AMACOM, New York 1991.