MECHANICAL ENGINEERING | RESEARCH ARTICLE

Process adjustment by a Bayesian approach

Daniel Duret1 and Maurice Pillet2*

Abstract: In a production or measure situation, operators are required to make corrections to a process using the measurement of a sample. In both cases, it is always difficult to suggest a correction from a deviation. The correction is the result of two different deviations: one in set-up and the second in production. The latter is considered as noise. The objective of this paper is to propose an original approach to calculate the best correction using a Bayesian approach. A correction formula is given with three assumptions as regards adjusting the distribution: uniform, triangular and normal distribution. This paper gives a graphical interpretation of these different assumptions and a discussion of the results. Based on these results, the paper proposes a practical rule for calculating the most likely maladjustment in the case of a normal distribution. This practical rule gives the best adjustment using a simple relation (Adjustment = K × sample mean) where K depends on the sample size, the ratio between the maladjustment and the short-term variability and a Type I risk of large maladjustment.

Subjects: Control Engineering; Engineering Productivity; Technology

Keywords: process adjustment; capability; statistical process control; estimation

1. Introduction

Statistical process control has always taken an interest in the detection of out-of-control situations. Since Shewhart (1931), who first proposed the eponymous control chart, many contributions have been made to detect a process that is deemed as out of control. Several types of control charts have improved the Average Run Length (ARL) such as the EWMA chart proposed by Lucas and Saccucci (1990) and the CUSUM chart (Hawkins & Olwell, 1998; Page, 1962). More recently, several authors have contributed to the improvement of the concept and suggested new advances on control charts, either to reduce the size of samples (Torng, Lee, & Liao, 2009), to improve the ARL (Nezhad & Niaki, 2010) or to adapt them for specific situations such as monitoring the mean of a measurable quality characteristic under 100% inspection (Wu, Wang, & Jiang, 2010). Process adjustment is critical for small run, so some research turned to the improvement of control cards for small series (Pillet, 1996).
The adjustment of the process after detecting a situation that is out of control was also studied in parallel to the work on the detection. The statistical approach to the tools setting was first addressed by Grubbs (1954). Grubbs proposed a sequence of adjustments using successive approaches. More recently, Castillo (1998), Pan (2002), and Trietsch (1998) have proposed a detailed analysis and an extension of the Grubbs rule for the adjustment of machines. Recent contributions have focused on the adjustment of the tools using the in situ measurement (Kibe, Okada, & Mitsui, 2007) and on the interest of the use of neural networks (Liao, Xie, Zhou, & Xia, 2007; Lin & Lin, 2009).

It is not easy to find the best adjustment when the value of the deviation from a given target has been estimated. This distance is the result of maladjustment and the process spread. Yet this problem arises more frequently with the use of machines in autonomy using automatic monitoring. The objective of this paper is to propose an approach for estimating the best correction using a Bayesian approach.

2. The adjustment problem
The adjustment problem can be applied to different situations such as:

- Monitoring a production parameter,
- Controlling a gauge tool,
- Monitoring the bias on a measuring device.

Production is considered as a random process described by an $X$ variable, depending on a $\theta$ adjustment parameter.

To simplify the writing of mathematical expressions, the parameter $\theta$ represents the deviation from the target (which corresponds to a change of origin for the parameter). $\theta$ is unknown, the result of this study will attempt to estimate it.

The variable $X$ is distributed according to a $\mathcal{L}$ distribution depending on the parameter $\theta$ (Figure 1). It represents the random deviation of production for a given setting, $\theta$:

- Mean: $\mu = \theta$ (we assume that the distribution is symmetric because $X$ is often a sample mean).
- Variance: $\sigma^2/n$ (if $n$ is very small, the previous symmetry assumption can be questioned).
- Distribution function: $Y = f(X)$.
- A particular individual estimate of $X$ will be written $x_i$.

If we take a sample of $n$ individual values for the variables $X_1, \ldots, X_i, \ldots, X_n$, considered a priori as totally independent, then we can determine the adjustment is to analyse the event $A$, a posteriori.

$$A = \bigcap_{i=1}^{n} \{X_i = x_i\}$$

(1)

Adjuster aggregates the information of the sample by retaining only the average (or median) of the sample. This simpler method of treatment leads to a loss of information. We will replace (1) by:

![Figure 1. Graphic representation of the adjustment problem.](image)
The information obtained using this sample informs us, a posteriori, of the probable values of $\theta$.

Before analysing this sample, we had ideas about a priori the parameter $\theta$, which were:

- $\theta$ belongs to a domain defined as $D_\theta$ (corresponding to the knowledge of the setting: set point, tracking drifts, expertise, etc.).
- $g(\theta)$ being the density of probability on the interval.

For example, the a priori distribution $g(\theta)$ is uniform:

- $D_\theta = \pm \delta$ then, $g(\theta) = 1/(2 \cdot \delta)$ (an adjustment greater than $\delta$ is not possible).

After reviewing the event $A$, a new a posteriori density can be calculated for the parameter $\theta$, density conditioned by knowledge of the event $A$ (Delsart & Vaneecloo, 2011):

$$g(\theta/A) = \frac{\int f(\bar{x}/\theta) \cdot g(\theta) \cdot d\theta}{\int_{D_\theta} f(\bar{x}/\theta) \cdot g(\theta) \cdot d\theta}$$  

Assuming that all the individual variables $X$ are distributed according to the same law $\mathcal{L}$, it means taking a sample over a relatively short period.

3. $\theta$ parameter estimation

There are several possibilities for estimating the parameter $\theta$. For example:

$$\hat{\theta} = E(\theta/A)$$  

We propose to retain the most likely value of $\theta$ on the definition of the domain $D_\theta$ (it may differ from the mean in the case of non-symmetrical distribution).

This is equivalent to finding $g(\theta/A)$ maximum; that is to say, finding the maximum of the numerator for the expression (3). It is then necessary to know the probability law of the parameter $\theta$ and the domain boundaries of $D_\theta$. This paper examines three assumptions for $g(\theta)$: uniform, triangular and normal distribution (Figure 2).

4. Assuming a uniform distribution for $g(\theta)$

4.1. Best adjustment

In this case, where $g(\theta)$ is a constant, it may be omitted in order to look for the most likely value of $\theta$. This is equivalent to maximizing the function:

$$L(\theta) = f(\bar{x}/\theta)$$  

In the case where the estimated parameter is the mean, the value of the parameter $\theta$ for which this function (5) reaches its maximum, is called the “maximum likelihood estimator”.

![Figure 2. Different assumptions to estimate the maximum likelihood of maladjustment.](image-url)
Based on the assumption that the variable $X = \bar{x}$ is normally distributed with a mean $\theta$ and a variance $\sigma^2/n$:

$$L(\theta) = f(\bar{x} / \theta) = \frac{\sqrt{n}}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{\bar{x} - \theta}{\sigma} \right)^2}$$  \hspace{1cm} (6)

To find the value of $\theta$ that maximizes this function, it is easier to find the maximum of the logarithm (the logarithm function is monotone):

$$l_n(L(\theta)) = l_n(\sqrt{n}) - l_n(\sigma \sqrt{2\pi}) - \frac{n}{2} \cdot \left( \frac{\bar{x} - \theta}{\sigma} \right)^2$$  \hspace{1cm} (7)

The first derivative gives:

$$\frac{d(l_n(L(\theta)))}{d\theta} = \frac{n}{\sigma^2} \cdot (\bar{x} - \theta)$$  \hspace{1cm} (8)

As the second derivative is negative, it is a maximum for:

$$\theta_M = \bar{x}$$  \hspace{1cm} (9)

In this particular case, with the assumption of uniform distribution for the parameter $\theta$, the best adjustment corresponds to the maximum likelihood estimator.

4.2. Graphic illustration with a uniform distribution for $g(\theta)$

Let us assume that the distribution of $X = \bar{x}_{ex}$ is symmetrical in relation to $\theta$. In addition we will use the following variable change:

$$X = \theta + X_p$$  \hspace{1cm} (10)

With $X_p$ as the random short-term distribution, we will retain a normal marginal distribution:

$$Y_p = \frac{\sqrt{n}}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{Y_p}{\sqrt{\sigma}} \right)^2}$$  \hspace{1cm} (11)

and for the maladjustment $\theta$, an uniform distribution

$$g(\theta) = \frac{1}{2\delta} \text{ with } \theta \in [-\delta; +\delta]$$

$$g(\theta) = 0 \text{ with } \theta \notin [-\delta; +\delta]$$  \hspace{1cm} (12)

The distribution function for the pair $(\theta, X_p)$ is given by:

$$f(\theta, X_p) = \frac{\sqrt{n}}{2 \sqrt{2\pi \cdot \delta \cdot \sigma}} \cdot e^{-\frac{1}{2} \left( \frac{\bar{x} - \theta}{\sqrt{\sigma}} \right)^2} |\theta| \leq \delta$$  \hspace{1cm} (13)

**Example**

$$\sigma = 2 \mu m, \delta = 8 \mu m, n = 5$$

For a particular instance of $X$, $X = \bar{x}_{ex}$, the information for the pair $(\theta, X_p)$ is not detailed. Therefore, several values of $\theta$ can correspond to the equation:

$$\bar{x}_{ex} = \theta + X_p$$  \hspace{1cm} (14)

Figure 3 illustrates all the possible solutions for $\bar{x} = 3 \mu m$.  

In the case of a uniform distribution setting, the best setting \( \theta = \bar{x}_{\text{ex}} \) corresponds to the maximum likelihood estimator.

5. Assuming a triangular distribution for \( g(\theta) \)

5.1. Best adjustment

In this assumption, \( g(\theta) \) is not a constant, so the equation to maximize is changed and becomes:

\[
L(\theta) = f(\bar{x}/\theta) \cdot g(\theta)
\]

The likelihood of a strong adjustment decreases proportionally to \( \theta \):

\[
g(\theta) = \frac{1}{\frac{\theta}{\delta} + 1} \quad \text{for} \quad \theta \in [0;\delta]
\]

\[
g(\theta) = \frac{1}{\frac{-\theta}{\delta} + 1} \quad \text{for} \quad \theta \in [-\delta;0]
\]

Otherwise \( g(\theta) = 0 \)

With the same assumptions as previously stated, (the variable \( X = \bar{x}_{\text{ex}} \) is normally distributed with a mean \( \theta \) and a variance \( \sigma^2/n \)) we obtain on the half domain (0; \( \delta \)):

\[
L(\theta) = f(\bar{x}/\theta) \cdot g(\theta) = \frac{\sqrt{n}}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{\bar{x} - \theta}{\sigma} \right)^2} \cdot \frac{1}{\delta} \left( -\frac{\theta}{\delta} + 1 \right)
\]

By taking the logarithm (the logarithm function is monotone):

\[
l_n(L(\theta)) = l_n(\sqrt{n}) - l_n(\sigma \sqrt{2\pi}) - \frac{n}{2} \left( \frac{\bar{x} - \theta}{\sigma} \right)^2 - l_n(\delta) + l_n\left( -\frac{\theta}{\delta} + 1 \right)
\]

The first derivative gives:

\[
\frac{d(l_n(L(\theta)))}{d\theta} = \frac{n}{\sigma^2} \cdot (\bar{x} - \theta) - \frac{1}{\delta - \theta}
\]

By canceling this derivative, we have a quadratic equation where the physical solution is:

\[
\theta_m = \frac{(\delta + \bar{x})}{2} - \sqrt{\left( \frac{\delta - \bar{x}}{2} \right)^2 + \frac{\sigma^2}{n}}
\]
5.2. Graphic illustration with a triangular distribution \( g(\theta) \)

We will take the same example as above and the same experimental values. Only the distribution of the pair \((\theta, X_p)\) is modified, and the density near the point \((0,0)\) is at its maximum.

With the same sample mean of 3 µm, we have the picture given in Figure 4.

Figure 5 gives a view from the plane containing \( \theta \) and \( f(\theta, X_p) \) allows us to estimate the maximum likelihood.

In the previous example, the correction suggested is practically minimal compared to the sample mean for small values of the mean. However, when the experimental sample mean is large (close to \( \delta \)), the difference between the average sample and the suggested correction is bigger (Figure 6).

For different values of the sample means (from 0 to 8 µm) with \( \sigma = 4 \) µm and \( n = 5 \), Figure 7 gives the density of the probability of the maladjustment. The maximum likelihood corresponds to the peak of the curve.

5.3. Generalization setting in the triangular case

Equation (20) gives the likelihood of a maladjustment for the case where \( \bar{x} > 0 \). To reflect a negative maladjustment, we propose the following modification:

\[
\theta_M = \begin{cases} 
\frac{\delta + \bar{x}}{2} - \sqrt{\left(\frac{\delta - \bar{x}}{2}\right)^2 + \frac{\sigma^2}{n}} & \text{for } \bar{x} \geq 0 \\
\frac{-\delta + \bar{x}}{2} + \sqrt{\left(\frac{\delta + \bar{x}}{2}\right)^2 + \frac{\sigma^2}{n}} & \text{for } \bar{x} < 0 
\end{cases}
\]  

(21)

Figure 4. Maximum likelihood when \( g(\theta) \) is triangular.

Figure 5. Maximum likelihood and sample mean.
Equation (20) was found by canceling the derivative. However, when \( \bar{x} \) tends towards 0, the limits where \( \bar{x} \to 0^+ \) and \( \bar{x} \to 0^- \) are not identical and are different from 0. In this case, the maximum likelihood will be forced to 0. Rule (21) must be slightly modified:

\[
\begin{align*}
\text{For } \bar{x} \geq 0 &\Rightarrow \theta_M = \frac{\delta + \bar{x}}{2} - \sqrt{\left(\frac{\delta - \bar{x}}{2}\right)^2 + \frac{\sigma^2}{n}} \\
\text{If } \theta_M < 0 &\text{ then } \theta_M = 0 \\
\text{For } \bar{x} < 0 &\Rightarrow \theta_M = \frac{-\delta + \bar{x}}{2} + \sqrt{\left(\frac{-\delta + \bar{x}}{2}\right)^2 + \frac{\sigma^2}{n}} \\
\text{If } \theta_M > 0 &\text{ then } \theta_M = 0
\end{align*}
\]  
(22)

6. Assuming a normal distribution for \( g(\theta) \)

With \( g(\theta) \) corresponding to a normal distribution \( \mathcal{N}(0, \sigma_\theta) \), reflecting a probability for a strong extremely low maladjustment, the calculation of the maximum likelihood for \( \theta \) gives (Duret, 2012):

\[
g(\theta) = \frac{1}{\sigma_\theta \sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{\theta}{\sigma_\theta} \right)^2}
\]  
(23)

Keeping the same assumptions (the variable \( X = \bar{x}_{ex} \) is normally distributed with a mean \( \theta \) and a variance \( \sigma^2/n \)), then the \( L(\theta) \) is:

\[
\begin{align*}
L(\theta) &= f(\bar{x}/\theta) \cdot g(\theta) = \frac{\sqrt{n}}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{\bar{x}}{\sigma} \right)^2} \cdot \frac{1}{\sigma_\theta \sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{\theta}{\sigma_\theta} \right)^2} \\
L(\theta) &= \frac{\sqrt{n}}{\sigma \sigma_\theta \sqrt{2\pi}} e^{-\frac{1}{2} \left( \bar{x} - 2\bar{x}\theta + \frac{\theta^2}{\sigma^2/n} \right)}
\end{align*}
\]  
(24)
By taking the logarithm:

\[ l_n(L(\theta)) = l_n\left(\sqrt{n}\right) - \frac{n \cdot \sigma \sqrt{2\pi}}{2}\sigma^2 \cdot \left(\bar{x} - 2\bar{x} \sigma + (1 + \frac{\sigma^2}{n\sigma^2})\vartheta\right) \]  

(25)

The first derivative gives:

\[ \frac{d(l_n(L(\theta)))}{d\vartheta} = \frac{n}{\sigma^2} \cdot \left(\bar{x} - (1 + \frac{\sigma^2}{n\sigma^2})\vartheta\right) \]  

(26)

By canceling this derivative, the solution is:

\[ \vartheta_M = \frac{1}{1 + \frac{\sigma^2}{n\sigma^2}} \cdot \bar{x} \]  

(27)

7. Discussion

The choice of the a priori distribution is relatively important for the calculation of the best adjustment, and the difference increases when the ratio \( \sigma / \delta \) increases. For example, it is obvious that if the short-term variability of the production process (or measurement) is low compared to the variations of the parameter \( \theta \) (\( \sigma \ll \delta \)), the difference between a uniform distribution and a triangular one is not very significant. Here, correcting the value of the sample average would be a good idea. On the other hand, if the setting is easily compared to the process spread, the correcting of the average is excessive. Using the previous example, the process spread is doubled (\( \sigma = 4 \mu m \) instead of \( \sigma = 2 \mu m \), \( \delta = 8 \mu m \), \( n = 3 \)). The corrections are given in Figure 8.

In addition to the choice of distribution law, the choice of the uncertainty involved setting strongly in the calculation. Several approaches deserve further developments:

1. Knowledge of the probability distribution can be obtained from an observation of historical data.

2. Determining the prior distribution can be enriched with the knowledge of the sample average at the sampling time.

The second point allows refining the calculation of the best adjustment. In the case of a normal distribution, Equation (27) was demonstrated. By taking \( \bar{X} = Z_a \sigma / \bar{x} \) depending on the Type I risk \( \alpha \).
We suppose that the knowing of the sample mean gives us information on the maladjustment distribution. The Type I risk $\alpha$ is the risk of very large maladjustment.

Equation (27) becomes:

$$\theta_M = \frac{1}{1 + \frac{r^2\sigma^2}{\hat{x}^2}} \hat{x}$$

(28)

Posing: $r = \hat{x}/\sigma$ (Pillet & Pairel, 2011) this means that the offset is evaluated a posteriori in short-term sigma:

$$\theta_M = \frac{n.r}{n.r + Z_u^2} \hat{x} = K\hat{x}$$

(29)

Figure 9 indicates the proportion $K$ of the most likely adjustment with the ratio $r$ for $Z_u = 1$.

8. Validation by Monte Carlo simulation

To validate the various assumptions used in this paper, it is necessary to compare the result in terms of variability in the same initial maladjustment situations. To achieve this validation, we conducted a Monte Carlo simulation with the following conditions:

• Simulation of a small run of 10 workpieces ($n = 1$).
• An initial maladjustment following a uniform distribution $[-9; 9]$.
• Common causes are $\mathcal{N}(0, 1)$.
• 100,000 run simulations.

Assumptions tested are the following:

• Shewhart benchmark if $\hat{x} > \text{Control Limit} \theta_M = \hat{x}$ else $\theta_M = 0$.
• Uniform distribution $\theta_M = \hat{x}$.
• Triangular distribution (21).
• Gaussian distribution (27).
• Enriched Gaussian distribution (28).
• Enriched Gaussian distribution (29).

Figure 9. Corrections for different sample sizes and $r$ with $Z_u = 1$. 

![Graph showing corrections for different sample sizes and $r$ with $Z_u = 1$.]
Results of the simulation are global variation (standard deviation) and the average of a million measurements. Less the standard deviation is large, better is the assumption. More close to 0 is the average, better is the assumption (see Table 1).

Best result is for Equation (29) with $z_α = 3$, worst result is for the uniform assumption and only the Gaussian assumptions improve the results vs. the Shewhart benchmark.

### Table 1. Monte Carlo results

|                     | Shewhart benchmark | Uniform Equation (9) | Triangle Equation (21) | Gauss Equation (27) | Gauss Equation (29) $z_α = 1$ | Gauss Equation (29) $z_α = 3$ |
|---------------------|--------------------|----------------------|------------------------|----------------------|-----------------------------|-----------------------------|
| Mean                | 0.29               | 1.36                 | 0.01                   | 0.01                 | 0.00                        | 0.01                        |
| Standard deviation  | 1.63               | 1.69                 | 1.75                   | 1.55                 | 1.56                        | 1.49                        |

9. Conclusion

In this paper, we proposed a method of calculation of the most probable maladjustment using three hypothesis of a priori distribution of the maladjustment. The rules we proposed are heavily dependent on the arbitrary choice:

- The domain of definition $D_θ = \pm δ$ for the parameter $θ$.
- The probability density $g(θ)$ for the parameter $θ$.

Between an assumed normal distribution for $g(θ)$ (Duret & Pillet, 2011; Pillet & Pairel, 2011) and an assumption of uniform distribution, the choice of a triangular distribution might seem a good compromise, but simulation shows the superiority of the Gaussian hypothesis with $z_α = 3$.

Future development must be done in order to propose a method to identify the good probability density and the parameter $K$. 

### References

- Castillo, E. D. (1998). A note on two process adjustment models. Quality and Reliability Engineering International, 14, 23–28. http://dx.doi.org/10.1002/(ISSN)1099-1638
- Delsart, V., & Vaneecloo, N. (2011). Estimation, test, échantillonnage [Estimation, testing, sampling]. Paris: Septentrion.
- Duret, D. (2012). Estimation d’une correction de réglage d’une erreur de justesse [Estimation of adjustment for a bias error]. Controles Essais Mesures, 38, 97–100.
- Grubbs, F. E. (1954). An optimum procedure for setting machines or adjusting processes. Industrial Quality Control; Reprinted in Journal of Quality Technology, 1983, 15, 186–189.
- Hawkins, D. M., & Olwell, D. H. (1998). Cumulative sum charts and charting for quality improvement. New York, NY: Springer. http://dx.doi.org/10.1007/978-1-4612-1686-5
- Kibe, Y., Okada, Y., & Mitsui, K. (2007). Machining accuracy for shearing process of thin-sheet metals—Development of initial tool position adjustment system. International Journal of Machine Tools and Manufacture, 47, 1728–1737. http://dx.doi.org/10.1016/j.ijmachtools.2006.12.006
- Liao, X. P., Xie, H. M., Zhou, Y. J., & Xia, W. (2007). Adaptive adjustment of plastic injection processes based on neural network. Journal of Materials Processing Technology, 187–188, 676–679. http://dx.doi.org/10.1016/j.jmatprotec.2006.11.068
- Lin, H. D., & Lin, W. T. (2009). Automated process adjustments of chip cutting operations using neural network and statistical
approaches. Expert Systems with Applications, 36, 4338–4345. http://dx.doi.org/10.1016/j.eswa.2008.05.035
Lucas, J. M., & Saccucci, M. S. (1990). Exponentially weighted moving average control schemes: Properties and enhancements. Technometrics, 32, 1–12. http://dx.doi.org/10.1080/00401706.1990.10484583
Nezhad, F. M. S., & Niaki, S. T. A. (2010). A new monitoring design for uni-variate statistical quality control charts. Information Sciences, 180, 1051–1059.
Page, E. S. (1962). Cumulative sum schemes using gauging. Technometrics, 4, 97–109. http://dx.doi.org/10.1080/00401706.1962.10489990
Pan, R. (2002, August). Statistical process adjustment methods for quality control in short-run manufacturing (Thesis in Industrial Engineering). Pennsylvania State University, University Park.
Pillet, M. (1996). A specific SPC chart for small-batch control. Quality Engineering, 8, 581–586. http://dx.doi.org/10.1080/08982119608904668
Pillet, M., & Pairel, E. (2011). Determination of an adjusting rule in the case of multi-criteria inertial piloting. International Journal of Metrology and Quality Engineering, 2, 51–59. http://dx.doi.org/10.1051/ijmqe/2011009
Shewhart, W. A. (1931). Economic control of quality of manufactured products (501 p.). New York, NY: Van Nostrand. (1980 by American Society for Quality).
Tong, C. T., Lee, P. H., & Liao, N. Y. (2009). An economic-statistical design of double sampling X control chart. International Journal of Production Economics, 120, 495–500. http://dx.doi.org/10.1016/j.ijpe.2009.03.013
Trietsch, D. (1998). The harmonic rule for process setup adjustment with quadratic loss. Journal of Quality Technology, 30, 75–84.
Wu, Z., Wang, Z., & Jiang, W. (2010). A generalized conforming run length control chart for monitoring the mean of a variable computers & industrial engineering, 59, 185–192.