Optimum Process Targets for Multiple Quality Characteristics Using Regression Analysis

Jirarat Teeravaraprug
Department of Industrial Engineering
Thammasat University, Rangsit Campus
Khlong Lung, Pathum Thani 12121

Abstract

The selection of optimum process target has become one of the focused research areas to increase productivity and improve product quality. Although the quality engineering literature contains a vast collection of work related to this issue, a couple of questions still remain unanswered. First, a quality loss function approach using conventional quality loss functions, such as step-loss and Taguchi loss functions, has been extensively used to determine the optimum process target mainly due to mathematical convenience. When historical data concerning a customer loss associated with product performance are available, a quality loss function using a well-established statistical method, such as regression analysis, may be a more practical alternative procedure. Second, many researchers have carried out their studies based on a single quality characteristic. From the customer’s viewpoint, however, products are often judged based on more than one characteristic. To address these questions, this paper first develops a multivariate empirical loss function based on historical data associated with product performance and its associated customer loss and then proposes an optimization scheme for the most economical process target.

Keywords: Process target, Regression analysis, Multiple quality characteristics, Empirical loss function, Optimization.

1. Introduction

One of the most important decision-making problems encountered in a wide variety of industrial processes is the determination of process target (mean). Selecting the optimum process target is critically important since it affects a process defective rate, material cost, scrap or rework cost, and the loss to the customer due to a deviation of a product performance from the customer-identified target value. Furthermore, the process target should be reset frequently and promptly due to unpredictable random variation in many manufacturing processes.

2. Related Research

Techniques to determine the optimum process target have been discussed and developed for more than forty years. The initial work on this issue probably began with Springer [1] who considered the problem of determining the optimum process target with specified upper and lower specification limits under the assumption of constant net income functions. This problem is often referred to as the ‘filling problem’ or ‘canning problem.’ There are some situations in which the minimum content is often dictated by legislation. In such a case, underfilled cans whose product performance falls below the legal minimum often need to be reprocessed. Along this line, Bettes [2] modeled the process target setting with a fixed lower specification limit and arbitrary upper specification limit when underfilled and overfilled cans are reprocessed at a fixed cost. In some situations, however, the cans that do not meet the minimum content requirement may be sold at a reduced price. Hunter and Kartha [3] presented a model to determine the optimum process target with the assumption that the cans...
meeting the minimum content requirement are sold in a regular market at a fixed price, while the underfilled cans are sold at a reduced price in a secondary market. Nelson [4-5] determined approximate solutions to the Hunter and Kartha model [3] and developed a nomograph for the Springer model [1]. The Hunter and Kartha model [3] was later modified by Bisgaard et al. [6] who assumed that underfilled cans are sold at a price proportional to their content, and Carlsson [7] who included a more general income function. In addition, Arcelus and Banerjee [8] extended the work of Bisgaard et al. [6] under the situation where there is a linear shift in process target.

Golhar [9] developed a model for the optimum process target with the assumptions that overfilled cans can be sold in the regular market while underfilled ones can be reprocessed. Golhar and Pollock [10] modified this model by treating both the upper specification limit and the process target as control variables, and Golhar [11] developed a computer program to solve the Golhar and Pollock model [10]. Arcelus and Rahim [12] presented a model for the most profitable process target where both variable and attribute quality characteristics of a product are considered simultaneously, while Boucher and Jafari [13] addressed the same problem by extending the line of research under the context of a sampling plan. Schmidt and Pfeifer [14] extended the models of Golhar [9] and Golhar and Pollock [10] by considering a limited process capacity. Al-Sultan [15] developed an algorithm to find the optimal machine setting when two machines are connected in series, and Das [16] presented a non-iterative numerical method for solving the Hunter and Kartha model [3]. Usher et al. [17] considered the process target problem under the situation where a demand for a product has not exactly met the capacity of a filling operation. Liu and Taghavachari [18] considered a general problem of the determination of both the optimal process target and the upper specification limit when a filling amount follows an arbitrary continuous distribution, and showed that the optimal upper specification limit can be presented by a very simple formula regardless of the shape of the distribution. Pulak and Al-Sultan [19] developed a set of FORTRAN-based computer codes, and Pollock and Golhar [20] reconsidered the process target problem under the environment of capacitated production and fixed demand. Teeravaraprug and Cho [21] used Taguchi's loss function to determine the optimal process target for multiple quality characteristics and Teeravaraprug [22] showed an optimization model for determining an optimal process mean when considering two-market products. Further, Hong and Elsayed [23] studied the effects of measurement errors on process target, and Pfeifer [24] showed the use of an electronic spreadsheet program as a solution method.

3. Observations and Research Motivations

The analysis in this paper differs from the previous studies in two ways. First, most studies in the literature deal with the problem of how to determine the optimum process target using the conventional quality loss functions such as step-loss and Taguchi loss functions. Some applications of these loss functions can be found in Refs. [25-30]. Even though these loss functions dominate the research as a quality evaluation mechanism mainly due to mathematical convenience, they may not be the best representation of the quality loss for a product. A close look at these loss functions reveals a shortcoming from a practical point of view. That is, these loss functions inherently assume that there is little or no information about the functional relationship between product performance and its associated quality loss. Second, most studies deal with a single quality characteristic to determine the optimum process target. From the customer's point of view, however, products are often judged by more than one quality characteristic. Hence, to meet or surpass the customer needs and satisfaction, the determination of the optimum process target for key quality characteristics is one of the most important tasks in early design stage.

To incorporate these ideas, this paper gives an attempt to exhibit a quality loss function, called a multivariate empirical loss function, using historical data associated with product performance and its customer loss, to deal with multiple quality characteristics. This paper then develops an optimization scheme to determine the optimum process target. This optimization scheme is demonstrated by numerical examples.
4. Multivariate Empirical Loss Function

Since the choice of path to quality enhancement depends heavily upon the type of quality loss function used, the selection of a proper quality loss function to relate key quality characteristics of a product to its product performance is critically important. Although there are several multivariate quality loss functions (see Refs 31-34), the exact form of quality loss is rarely known. If empirical data concerning quality loss associated with product performance are available, which is often the case in many industries, the empirical loss function may ensure a better estimate of quality loss. To develop such a quality loss function, the concept of regression analysis, which is one of the most widely used tools in engineering and many other fields for investigating cause and effect relationships, can be employed.

Note that each quality characteristic \( y_i \) has its own customer-identified target \( \tau_i \), and quality losses incurred due to the deviation from the target value in positive and negative directions may not be equal. Hence, \( 2^n \) different sets \( (\Omega_1, \Omega_2, \ldots, \Omega_n) \) form the specification region where \( n \) quality characteristics are present. More detailed descriptions on the sets forming the specification region are discussed in Appendix A.

Let \( \hat{L}(y_1, y_2, \ldots, y_n) \) denote an empirical estimate of quality loss associated with \( n \) quality characteristics, where \( y_i \) is the \( i \)th quality characteristic for \( i = 1, 2, \ldots, n \). A first-order estimated multivariate empirical loss function is then given by Eq. (1):

\[
\hat{L}(y_1, y_2, \ldots, y_n) = \beta_0 + \beta_{y_1} y_1 + \ldots + \beta_{y_n} y_n + \ldots + \beta_{y_{1,2}} y_{1,2} + \ldots + \beta_{y_{1,n}} y_{1,n} + \ldots + \beta_{y_{2,n}} y_{2,n} + \ldots + \beta_{y_{n,1}} y_{n,1} + \ldots + \beta_{y_{n,n}} y_{n,n} + \cdots
\]

where \( \beta_{y_i} \)'s are empirical estimates of unknown parameters that are determined by using the least squares method associated with the \( i \)th quality characteristic in the set of \( \Omega_j \). A numerical example is shown in a later section to show how to develop Eq. (1).

Similarly, a second-order estimated multivariate empirical loss function is given by:

\[
i(y_1, y_2, \ldots, y_n) = \hat{\alpha}_0 + \hat{\alpha}_{y_1} y_1 + \hat{\alpha}_{y_2} y_2 + \ldots + \hat{\alpha}_{y_{1,2}} y_{1,2} + \ldots + \hat{\alpha}_{y_{1,n}} y_{1,n} + \ldots + \hat{\alpha}_{y_{2,n}} y_{2,n} + \ldots + \hat{\alpha}_{y_{n,1}} y_{n,1} + \ldots + \hat{\alpha}_{y_{n,n}} y_{n,n} + \cdots
\]

where \( \hat{\alpha}_0 \) is an empirical estimate of loss intercept, \( \hat{\alpha}_{y_{ij}} \)'s are empirical estimates of unknown parameters of the first degree associated with the \( i \)th quality characteristic, and \( \hat{\alpha}_{y_{ij}} \)'s are empirical estimates of unknown parameters of the second degree associated with the \( k \)th and \( l \)th quality characteristics.

The first- and second-order models, with the consideration of two quality characteristics are presented in Appendix B. It should be noticed that the shapes of the empirical loss functions defined in Eqs. (1-2) are not required to be symmetrical about the target, whereas the step-loss and Taguchi loss functions are symmetrical about the target. This means that empirical losses due to the equal deviation from the target value in either positive or negative directions may not lead to equal quality losses, which is the general case in many industrial problems. Although this paper discusses two types of the estimated empirical loss functions, other higher-order estimated models may be generated by regression analysis. Hence, the empirical loss function is more flexible in terms of a functional form and evaluates customer perception of product quality in a more effective manner. In practice, a functional form (usually, a first- or second-order) is first assumed to obtain an estimated multivariate empirical loss function, and then it needs to be checked using a lack-of-fit test, if the assumed estimated function is appropriate.

5. Analysis of Costs

When designing an optimum process target, three types of costs are generally considered in the early design stage. First, a rejection cost is incurred by a manufacturer when a product
performance fails to fall within the specification region of interest, and hence corrective actions on the rejected products, such as repairing, scrapping, or returning the products to the manufacturer, need to be taken. Second, an inspection cost is also incurred by the manufacturer when inspections are performed on products. Finally, a quality loss is incurred by the customer when the performance deviates from the customer-identified target within the specification region. The term ‘specification region’ refers to the jointly-intersected set by the individual sets of specifications for each quality characteristic. That is, when a product performance falls within the specification region, customers are willing to accept the product.

Denoting the expected total rejection cost, expected quality loss, and inspection cost by $E[TC_R(y_1, y_2, \ldots, y_n)]$, $E[L(y_1, y_2, \ldots, y_n)]$, and $C_I$, respectively, the expected total cost becomes:

$$E[TC(y_1, y_2, \ldots, y_n)] = E[TC_R(y_1, y_2, \ldots, y_n)] + E[L(y_1, y_2, \ldots, y_n)] + C_I$$

(3)

In many industrial settings, a rejection cost may vary under different rejection regions (see Ref. 35). For example, consider the case of two quality characteristics as shown in Figure 1. Note that region $I$ represents the rejection region where only one quality characteristic falls within the specification region formed by two quality characteristics, and region $II$ represents the rejection region where both quality characteristics fail to fall within the specification region. From a practical point of view, the rejection cost when a product performance falls within region $I$ is lower than the one when a product performance falls within region $II$. It is further noted that the number of different types of rejection regions is equal to the number of quality characteristics, and their rejection costs are different. Extending this idea to the case of $n$ quality characteristics, let $C^q_R$ and $P(R^q_n)$ be the unit rejection cost associated with region $q$ and the probability that a product performance falls within region $q$, respectively, where $q = 1, 2, \ldots, n$. Appendix C analyzes the set of probabilities that a product performance falls in each rejection region where $n$ quality characteristics are present. The expected total rejection cost is then obtained by:

$$E[TC_R(y_1, y_2, \ldots, y_n)] = \sum C^q_R P(R^q_n)$$

(4)

Considering the second part of the right-hand-side of Eq. (3), the expected quality loss can be expressed as:

$$E[L(y_1, y_2, \ldots, y_n)] = \int_{LSL_1}^{USL_1} \int_{LSL_2}^{USL_2} \cdots \int_{LSL_n}^{USL_n} L(y_1, y_2, \ldots, y_n) f(y_1, y_2, \ldots, y_n) dy_1 dy_2 \cdots dy_n$$

(5)

Similarly, the expected quality loss with two quality characteristics is obtained by:

$$E[L(y_1, y_2)] = \int_{LSL_1}^{USL_1} \int_{LSL_2}^{USL_2} L(y_1, y_2) f(y_1, y_2) dy_1 dy_2$$

(6)

6. The Optimization Models

The objective is to minimize expected total cost with the constraint that a process target is within the specification region of interest. Using the mathematical results obtained in the previous sections and applying a multivariate normal distribution for the quality characteristics of interest, Table 1 shows a general optimization model. The objective function comprises rejection costs, quality loss and inspection cost, and the constraints are employed to ensure that the process distribution follows a multivariate normal distribution and the resulting process

Figure. 1 Rejection regions for a two-quality characteristic case
target is in the specification region of interest. Note that the matrix \( V \) is a variance-covariance matrix of the quality characteristics and \( |V| \) is the determinant of the matrix \( V \). As a special case, Table 2 presents the optimization model for the optimum process target with the consideration of two quality characteristics.

7. Numerical Examples

Two examples are presented to demonstrate how the proposed models can be applied to real-world problems. We show how to determine the optimum process target by showing the procedures of developing the first-and second-order estimated multivariate loss.

7.1 Application of the first-order multivariate empirical loss function

To construct a prototype of an infrared laser imaging media, layers of material were coated in a laboratory. These included heat absorbing dyes \( (y_1) \) and thermally activated initiator \( (y_2) \). The objective was to find the optimum setting of those quality characteristics. The ideal dimensions of absorbing dyes and thermally activated initiator are 7.5 and 12.5, respectively. If the ideal dimensions (i.e., customer-identified target) are not met, a loss would be incurred by the customer and the losses associated with the product performance based on the deviations from the target were collected as shown in Table 3 and Figure 2. Moreover, a manufacturer provides the specifications of those quality characteristics which are \((LSL_1=5, USL_1=10)\) and \((LSL_2=10, USL_2=15)\). Note that the specification region is the jointly-intersected region formed by two sets of specifications. If only one quality characteristic fails within the specification region, the unit rejection cost of 20 (i.e., \( C_R^l = 20 \)) would be incurred. Similarly, if both quality characteristics fail to fall within the specification region, the higher unit rejection cost of 25 (i.e., \( C_R^h = 25 \)) would be incurred. The inspection cost of 0.75 (i.e., \( C_I = 0.75 \)) is also incurred by the manufacturer. Furthermore, two quality characteristics (i.e., \( y_1 \) and \( y_2 \)) are statistically independent and follow the bivariate normal distribution with \( \sigma_1^2 = 1 \) and \( \sigma_2^2 = 4 \).

Using regression analysis to examine the data set shown in Table 3, the first-order estimated empirical loss functions are appropriate.

| Table 1. Optimization model for an n-quality characteristic case. |

Minimize

\[
E[TC(y_1, y_2, \ldots, y_n)] = \sum_{q=1}^{n} C_q^2 P(R_q^t) + \int_{LSL_1}^{USL_1} \int_{LSL_2}^{USL_2} \cdots \int_{LSL_n}^{USL_n} \hat{L}(y_1, y_2, \ldots, y_n) \cdot f(y_1, y_2, \ldots, y_n) \, dy_1 \, dy_2 \, \cdots \, dy_n + C_I
\]

subject to

\[
f(y_1, y_2, \ldots, y_n) = \frac{1}{(2\pi)^{n/2}|V|^{n/2}} \exp \left[ -\frac{1}{2} (y - \mu)^T V^{-1} (y - \mu) \right]
\]

\[
\hat{L}(y_1, y_2, \ldots, y_n) = \begin{cases} 
\text{use Eq. (1), when the first-order empirical loss function is appropriate} \\
\text{use Eq. (2), when the second-order empirical loss function is appropriate}
\end{cases}
\]

\[
P(R_q^t) = \Pr[\text{product performance falls within region } q]
\]

\[
LSL_i \leq \mu_i \leq USL_i, \quad i = 1, 2, \ldots, n
\]
Table 2. Optimization model for a two-quality characteristic case.

Minimize

\[
E[TC] = C_1 \left[ \int LSL_i \int USL_2 \int f(y_1, y_2) \, dy_1 \, dy_2 + \int LSL_2 \int USL_i \int f(y_1, y_2) \, dy_1 \, dy_2 + \int USL_2 \int USL_i \int f(y_1, y_2) \, dy_1 \, dy_2 \right]
\]

\[
+ \int USL_2 \int LSL_2 \int f(y_1, y_2) \, dy_1 \, dy_2 \right]\]

\[
+ \int USL_2 \int LSL_i \int f(y_1, y_2) \, dy_1 \, dy_2 + \int LSL_2 \int USL_i \int f(y_1, y_2) \, dy_1 \, dy_2 \right]
\]

\[
+ \int LSL_i \int USL_i \int LSL_2 \int USL_i \int LSL_2 \int USL_i \int \hat{L}(y_1, y_2) \cdot f(y_1, y_2) \, dy_1 \, dy_2 + C_i
\]

subject to

\[
f(y_1, y_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp \left[ -\frac{1}{2(1-\rho^2)} \left( \frac{y_1 - \mu_1}{\sigma_1} \right)^2 - 2\rho \left( \frac{y_1 - \mu_1}{\sigma_1} \right) \left( \frac{y_2 - \mu_2}{\sigma_2} \right) + \left( \frac{y_2 - \mu_2}{\sigma_2} \right)^2 \right]
\]

\[
\hat{L}(y_1, y_2) = \begin{cases} 
\text{use Eq. (1), when the first-order empirical loss function is appropriate} \\
\text{use Eq. (2), when the second-order empirical loss function is appropriate} 
\end{cases}
\]

\[
LSL_i \leq \mu_1 \leq USL_1
\]

\[
LSL_2 \leq \mu_2 \leq USL_2
\]
Table 3. Empirical data of quality losses for the first numerical example.

<table>
<thead>
<tr>
<th>$y_2$</th>
<th>$y_1$</th>
<th>5</th>
<th>5.5</th>
<th>6</th>
<th>6.5</th>
<th>7</th>
<th>7.5</th>
<th>8</th>
<th>8.5</th>
<th>9</th>
<th>9.5</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td></td>
<td>16.8</td>
<td>15.8</td>
<td>14.8</td>
<td>13.7</td>
<td>12.8</td>
<td>11.8</td>
<td>16</td>
<td>17.4</td>
<td>19.4</td>
<td>20.5</td>
<td>22</td>
</tr>
<tr>
<td>10.5</td>
<td></td>
<td>15.1</td>
<td>13.5</td>
<td>12.7</td>
<td>12.2</td>
<td>11.2</td>
<td>9.6</td>
<td>14.3</td>
<td>15.1</td>
<td>16.5</td>
<td>18.3</td>
<td>19.5</td>
</tr>
<tr>
<td>11</td>
<td></td>
<td>14.3</td>
<td>12.3</td>
<td>11.3</td>
<td>9.8</td>
<td>9.1</td>
<td>7.9</td>
<td>11.4</td>
<td>12.8</td>
<td>14.3</td>
<td>15.5</td>
<td>17.7</td>
</tr>
<tr>
<td>11.5</td>
<td></td>
<td>11.5</td>
<td>10.5</td>
<td>10.9</td>
<td>8.4</td>
<td>6.7</td>
<td>6.2</td>
<td>8.7</td>
<td>9.6</td>
<td>11.5</td>
<td>13.9</td>
<td>14.5</td>
</tr>
<tr>
<td>12</td>
<td></td>
<td>9.2</td>
<td>8.7</td>
<td>6.7</td>
<td>5.9</td>
<td>4.9</td>
<td>4.1</td>
<td>6.5</td>
<td>7.9</td>
<td>9.4</td>
<td>10.5</td>
<td>12</td>
</tr>
<tr>
<td>12.5</td>
<td></td>
<td>6.9</td>
<td>6.8</td>
<td>4.7</td>
<td>4.3</td>
<td>3</td>
<td>1.2</td>
<td>3.1</td>
<td>4.3</td>
<td>5.2</td>
<td>6</td>
<td>7.1</td>
</tr>
<tr>
<td>13</td>
<td></td>
<td>11.2</td>
<td>9.3</td>
<td>7.7</td>
<td>6.1</td>
<td>5.2</td>
<td>3.5</td>
<td>5.8</td>
<td>8.1</td>
<td>9.3</td>
<td>11.1</td>
<td>13.2</td>
</tr>
<tr>
<td>13.5</td>
<td></td>
<td>12.7</td>
<td>11.2</td>
<td>9.3</td>
<td>8.3</td>
<td>6.1</td>
<td>5.2</td>
<td>7.1</td>
<td>9.4</td>
<td>11.1</td>
<td>12.6</td>
<td>14.7</td>
</tr>
<tr>
<td>14</td>
<td></td>
<td>13.3</td>
<td>12.1</td>
<td>10.9</td>
<td>9.6</td>
<td>7.8</td>
<td>6.7</td>
<td>8.9</td>
<td>10.7</td>
<td>12.3</td>
<td>14.3</td>
<td>16.7</td>
</tr>
<tr>
<td>14.5</td>
<td></td>
<td>15.2</td>
<td>13.4</td>
<td>12.2</td>
<td>11.1</td>
<td>9.1</td>
<td>7.8</td>
<td>10.3</td>
<td>12.2</td>
<td>13.7</td>
<td>16.1</td>
<td>18.3</td>
</tr>
<tr>
<td>15</td>
<td></td>
<td>16.9</td>
<td>15.2</td>
<td>13.7</td>
<td>12.4</td>
<td>10.7</td>
<td>9.5</td>
<td>11.7</td>
<td>13.8</td>
<td>15.2</td>
<td>17.5</td>
<td>20.1</td>
</tr>
</tbody>
</table>

Figure 2. Graphical presentation of the data set illustrated in the first numerical example.
Table 4. ANOVA tables for the first-order empirical loss functions.

<table>
<thead>
<tr>
<th>SOURCE</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>2</td>
<td>966.5212</td>
<td>483.2606</td>
<td>576.84</td>
<td>0.0001</td>
</tr>
<tr>
<td>Error</td>
<td>33</td>
<td>27.6463</td>
<td>0.8378</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>35</td>
<td>994.1675</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

R-Square: 0.9722, C.V.: 7.7953, Root MSE: 0.9153

| VARIABLE | DF | ESTIMATE | STANDARD ERROR | T FOR H0: PARAMETER = 0 | PROB>|T| |
|----------|----|----------|----------------|--------------------------|------|
| Intercept| 1  | 41.3964  | 2.5507         | 16.23                    | 0.0001 |
| y1       | 1  | 3.2219   | 0.1786         | 18.03                    | 0.0001 |
| y2       | 1  | -5.1419  | 0.1786         | -28.78                   | 0.0001 |

SOURCE | DF | SS       | MS      | F       | p   |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>2</td>
<td>525.6940</td>
<td>262.8470</td>
<td>1136.29</td>
<td>0.0001</td>
</tr>
<tr>
<td>Error</td>
<td>33</td>
<td>7.6335</td>
<td>0.2313</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>35</td>
<td>533.3275</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

R-Square: 0.9857, C.V.: 5.0143, Root MSE: 0.4810

| VARIABLE | DF | ESTIMATE | STANDARD ERROR | T FOR H0: PARAMETER = 0 | PROB>|T| |
|----------|----|----------|----------------|--------------------------|------|
| Intercept| 1  | 67.1821  | 1.2108         | 55.49                    | 0.0001 |
| y1       | 1  | -2.2048  | 0.0939         | -23.49                   | 0.0001 |
| y2       | 1  | -3.8943  | 0.0939         | -41.48                   | 0.0001 |

SOURCE | DF | SS       | MS      | F       | p   |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>2</td>
<td>488.6155</td>
<td>244.3077</td>
<td>898.59</td>
<td>0.0001</td>
</tr>
<tr>
<td>Error</td>
<td>33</td>
<td>8.9720</td>
<td>0.2719</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>35</td>
<td>497.5875</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

R-Square: 0.9820, C.V.: 5.6831, Root MSE: 0.5214
where the unknown parameters were estimated using the least-squares method. The respective Analysis of Variance (ANOVA) tables and tests for individual parameters based on the above models are given in Table 4. The results in the ANOVA tables show sufficient evidence that the first-order models are appropriate since the $R^2$ value of each estimated model is 0.972, 0.986, 0.982, and 0.959, respectively. The parameter estimates described in Table 4 show that all parameters are statistically significant. Although not shown here, the assumptions for linear regression are supported by residual and normality plots.

Once the loss function is obtained, an optimization model is developed using Table 2. Solving the optimization model by using Excel Solver®, the optimum process target is obtained by $(\mu_1^* = 7.40, \mu_2^* = 12.82)$, with $E^*[TC_k(y_1, y_2)] = 4.57$, and $E^*[TC(y_1, y_2)] = 9.31$. Note that if the process target is set at the customer-identified target

\begin{align*}
L(y_1, y_2) &= \begin{cases}
41.3964 + 3.2219y_1 - 5.1419y_2, & 7.5 \leq y_1 \leq 10 \text{ and } 10 \leq y_2 \leq 12.5 \\
67.1821 - 2.2048y_1 - 3.8943y_2, & 5 \leq y_1 \leq 7.5 \text{ and } 10 \leq y_2 \leq 12.5 \\
-18.0274 - 2.8238y_1 + 3.2619y_2, & 7.5 \leq y_1 \leq 10 \text{ and } 12.5 < y_2 \leq 15 \\
-73.2833 + 3.6324y_1 + 3.7733y_2, & 7.5 < y_1 \leq 10 \text{ and } 12.5 < y_2 \leq 15
\end{cases}
\end{align*}

(7)

...
7.2 Application of the second-order multivariate empirical loss function

The second example considers machining operations for precision-machined alloy foundry products. The product considered here is a valve seat insert, and the machining operations for this product consist of removing metal chips to produce a smooth surface and to assure proper fit within the specified specification. The key quality characteristics for the valve seat insert are outside diameter \( (y_1) \) and width \( (y_2) \). The specified specifications are given by \( (LSL_1 = 7, USL_1 = 10) \) and \( (LSL_2 = 18, USL_2 = 21) \). Further, \( \tau_1 = 8, \tau_2 = 20, C^l_1 = 30, C^u_1 = 50, \) and \( C_2 = 2 \). The losses due to the deviations from the customer-identified target are shown in Table 5 and the graphical presentation of the data set is shown in Figure 3. The quality characteristics \( (y_1, y_2) \) follow a bivariate normal distribution with \( \sigma_1^2 = 0.49 \) and \( \sigma_2^2 = 0.25 \). The problem is where the process target should be set in order to minimize the expected total cost.

Using Figure 3, the following second-order empirical loss function is appropriate.

\[
\hat{L}(y_1, y_2) = 1150.3658 - 67.9622y_1 - 87.8513y_2 \\
+ 3.0034y_1^2 + 1.9973y_2^2 + 0.9953y_1y_2
\]

(8)

where the unknown parameters were estimated using the least-squares method. The ANOVA table and tests for individual parameters for the above model are given in Table 6, and the second-order model is appropriate since the \( R^2 \) value is 0.99.

Once this empirical loss function is obtained, an optimization model is developed using Table 2. Solving the optimization model by using Excel Solver\textsuperscript{a}, the optimum process target is obtained by \( (\mu_1 = 8.33, \mu_2 = 19.72) \), with \( E^*[TC_R(y_1, y_2)] = 1.28 \), \( E^*[\hat{L}(y_1, y_2)] = 2.06 \), and \( E^*[TC(y_1, y_2)] = 5.34 \). Note that if the process target is set at the customer-identified target \( (\mu_1 = 8, \mu_2 = 20) \), then \( E^*[TC_R(y_1, y_2)] = 3.03 \), \( E^*[\hat{L}(y_1, y_2)] = 1.42 \), and \( E^*[TC(y_1, y_2)] = 19.60 \).

For this particular example, some magnitude of savings as high as 17.21% \( =((6.45-5.34)/6.45)*100 \) would be realized by implementing the optimum process target. This example also indicates that the customer-identified target \( (\tau_1 = 8 \text{ and } \tau_2 = 20) \) is not the most economical process target. The graphical presentation of this example associated with the customer-identified target and optimum process target is depicted in Figure 4.

8. Conclusion

Quality engineers are often faced with the problem of how to determine the most economical process target. Although this issue has been extensively studied in the research community, there is room for improvement. This paper attempts to incorporate the customer’s overall perception of product quality using regression analysis when multiple quality characteristics are considered. This paper first develops a multivariate empirical loss function and then presents optimization models for the most economical process target by taking into account a customer loss based on empirically gathered data and the costs incurred by a manufacturer such as rejection costs and inspection cost. The particular numerical examples reveal that a significant amount of savings would be realized by implementing the optimum process target, indicating that the customer-identified target may not be the most cost-effective setting to pursue. It should be noted that the models shown in Tables 1 and 2 are based on multivariate and bivariate normal distributions. In the case of quality characteristics falling in other distributions, these two models cannot be used directly but they can be applied by changing the distribution functions \( f(y_1, y_2, \ldots, y_n) \) and \( f(y_1, y_2) \) to an appropriate function.
Table 5. Empirical data of quality losses for the second numerical example.

<table>
<thead>
<tr>
<th>( y_1 )</th>
<th>18</th>
<th>18.2</th>
<th>18.4</th>
<th>18.6</th>
<th>18.8</th>
<th>19</th>
<th>19.2</th>
<th>19.4</th>
<th>19.6</th>
<th>19.8</th>
<th>20</th>
<th>20.2</th>
<th>20.4</th>
<th>20.6</th>
<th>20.8</th>
<th>21</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>12.98</td>
<td>11.25</td>
<td>9.70</td>
<td>8.30</td>
<td>7.00</td>
<td>6.00</td>
<td>5.00</td>
<td>4.30</td>
<td>3.75</td>
<td>3.30</td>
<td>3.02</td>
<td>2.90</td>
<td>3.00</td>
<td>3.20</td>
<td>3.50</td>
<td>4.00</td>
</tr>
<tr>
<td>7.2</td>
<td>11.50</td>
<td>9.68</td>
<td>8.37</td>
<td>7.00</td>
<td>5.80</td>
<td>4.65</td>
<td>4.00</td>
<td>3.20</td>
<td>2.60</td>
<td>2.20</td>
<td>1.95</td>
<td>1.80</td>
<td>2.00</td>
<td>2.14</td>
<td>2.60</td>
<td>3.10</td>
</tr>
<tr>
<td>7.4</td>
<td>10.30</td>
<td>8.60</td>
<td>7.20</td>
<td>5.85</td>
<td>4.72</td>
<td>3.70</td>
<td>3.00</td>
<td>2.30</td>
<td>1.70</td>
<td>1.25</td>
<td>1.08</td>
<td>1.00</td>
<td>1.20</td>
<td>1.50</td>
<td>1.90</td>
<td>2.50</td>
</tr>
<tr>
<td>7.6</td>
<td>9.30</td>
<td>7.71</td>
<td>6.30</td>
<td>5.00</td>
<td>3.76</td>
<td>2.90</td>
<td>2.00</td>
<td>1.50</td>
<td>0.96</td>
<td>0.67</td>
<td>0.50</td>
<td>0.50</td>
<td>0.70</td>
<td>1.00</td>
<td>1.44</td>
<td>2.00</td>
</tr>
<tr>
<td>7.8</td>
<td>8.50</td>
<td>7.00</td>
<td>5.50</td>
<td>4.30</td>
<td>3.25</td>
<td>2.32</td>
<td>1.50</td>
<td>1.00</td>
<td>0.50</td>
<td>0.25</td>
<td>0.20</td>
<td>0.20</td>
<td>0.40</td>
<td>0.80</td>
<td>1.25</td>
<td>1.95</td>
</tr>
<tr>
<td>8</td>
<td>8.05</td>
<td>6.50</td>
<td>5.10</td>
<td>3.90</td>
<td>2.80</td>
<td>2.00</td>
<td>1.30</td>
<td>0.80</td>
<td>0.40</td>
<td>0.07</td>
<td>0.01</td>
<td>0.08</td>
<td>0.32</td>
<td>0.70</td>
<td>1.28</td>
<td>2.00</td>
</tr>
<tr>
<td>8.2</td>
<td>7.74</td>
<td>6.30</td>
<td>5.00</td>
<td>3.80</td>
<td>2.76</td>
<td>1.92</td>
<td>1.25</td>
<td>0.70</td>
<td>0.40</td>
<td>0.15</td>
<td>0.12</td>
<td>0.24</td>
<td>0.52</td>
<td>1.00</td>
<td>1.60</td>
<td>2.30</td>
</tr>
<tr>
<td>8.4</td>
<td>7.70</td>
<td>6.30</td>
<td>5.00</td>
<td>3.85</td>
<td>3.00</td>
<td>2.00</td>
<td>1.45</td>
<td>0.99</td>
<td>0.64</td>
<td>0.50</td>
<td>0.50</td>
<td>0.70</td>
<td>1.00</td>
<td>1.50</td>
<td>2.10</td>
<td>2.90</td>
</tr>
<tr>
<td>8.6</td>
<td>7.91</td>
<td>6.50</td>
<td>5.26</td>
<td>4.20</td>
<td>3.10</td>
<td>2.50</td>
<td>1.88</td>
<td>1.50</td>
<td>1.20</td>
<td>1.05</td>
<td>1.07</td>
<td>1.30</td>
<td>1.70</td>
<td>2.10</td>
<td>2.80</td>
<td>3.70</td>
</tr>
<tr>
<td>8.8</td>
<td>8.29</td>
<td>7.01</td>
<td>5.80</td>
<td>4.60</td>
<td>3.85</td>
<td>3.20</td>
<td>2.54</td>
<td>2.20</td>
<td>1.85</td>
<td>1.90</td>
<td>2.20</td>
<td>2.50</td>
<td>3.10</td>
<td>3.90</td>
<td>4.80</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>9.02</td>
<td>7.69</td>
<td>6.58</td>
<td>5.50</td>
<td>4.60</td>
<td>4.10</td>
<td>3.50</td>
<td>3.20</td>
<td>2.95</td>
<td>2.90</td>
<td>3.00</td>
<td>3.30</td>
<td>3.70</td>
<td>4.30</td>
<td>5.00</td>
<td>6.00</td>
</tr>
<tr>
<td>9.2</td>
<td>10.00</td>
<td>8.70</td>
<td>7.50</td>
<td>6.60</td>
<td>5.80</td>
<td>5.20</td>
<td>4.70</td>
<td>4.30</td>
<td>4.20</td>
<td>4.20</td>
<td>4.30</td>
<td>4.70</td>
<td>5.12</td>
<td>5.70</td>
<td>6.60</td>
<td>7.50</td>
</tr>
<tr>
<td>9.4</td>
<td>11.05</td>
<td>9.90</td>
<td>8.79</td>
<td>7.80</td>
<td>7.00</td>
<td>6.40</td>
<td>6.00</td>
<td>5.80</td>
<td>5.65</td>
<td>5.70</td>
<td>5.90</td>
<td>6.30</td>
<td>6.70</td>
<td>7.50</td>
<td>8.30</td>
<td>9.30</td>
</tr>
<tr>
<td>9.6</td>
<td>12.50</td>
<td>11.31</td>
<td>10.23</td>
<td>9.40</td>
<td>8.50</td>
<td>8.00</td>
<td>7.70</td>
<td>7.50</td>
<td>7.40</td>
<td>7.50</td>
<td>7.70</td>
<td>8.10</td>
<td>8.75</td>
<td>9.40</td>
<td>10.24</td>
<td>11.30</td>
</tr>
<tr>
<td>9.8</td>
<td>14.16</td>
<td>13.00</td>
<td>12.00</td>
<td>11.20</td>
<td>10.50</td>
<td>10.00</td>
<td>9.60</td>
<td>9.40</td>
<td>9.30</td>
<td>9.50</td>
<td>9.80</td>
<td>10.20</td>
<td>10.74</td>
<td>11.50</td>
<td>12.50</td>
<td>13.50</td>
</tr>
<tr>
<td>10</td>
<td>16.00</td>
<td>15.00</td>
<td>13.99</td>
<td>13.10</td>
<td>12.50</td>
<td>12.00</td>
<td>11.70</td>
<td>11.52</td>
<td>11.50</td>
<td>11.70</td>
<td>12.00</td>
<td>12.50</td>
<td>13.20</td>
<td>14.00</td>
<td>15.00</td>
<td>16.00</td>
</tr>
</tbody>
</table>

Figure 3. Graphical presentation of the data set illustrated in the second numerical example.

Figure 4. Customer-identified target and optimum process target.
Table 6. ANOVA table for the second-order empirical loss function.

<table>
<thead>
<tr>
<th>SOURCE</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>5</td>
<td>2948.7441</td>
<td>589.7488</td>
<td>803.94</td>
<td>0.0001</td>
</tr>
<tr>
<td>Error</td>
<td>250</td>
<td>183.3929</td>
<td>0.7336</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>255</td>
<td>3132.1370</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>3132.1370</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

R-Square: 0.9414  
C.V.: 13.7426  
Root MSE: 0.8565

| VARIABLE | DF | ESTIMATE | STANDARD ERROR | T FOR H0: PARAMETER = 0 | PROB>|T| |
|----------|----|----------|----------------|--------------------------|-------|
| Intercept| 1  | 880.7056 | 29.2948        | 30.06                    | 0.0001|
| $y_1$    | 1  | -54.9711 | 1.7208         | -31.94                   | 0.0001|
| $y_2$    | 1  | -65.1682 | 2.8143         | -23.16                   | 0.0001|
| $y_1^2$  | 1  | -2.2345  | 0.0708         | 31.55                    | 0.0001|
| $y_2^2$  | 1  | -1.4101  | 0.0708         | 19.91                    | 0.0001|
| $y_1y_2$ | 1  | 0.9823   | 0.0630         | 15.60                    | 0.0001|

**Appendix A**

When there are $n$ quality characteristics, $n$ customer-identified target values exist and hence $2^n$ different sets form the specification region. For example, when $n=1$, two different sets around the target value, such as $\Omega_1 = \{(y_1) : LSL \leq y_1 \leq \tau_1\}$ and $\Omega_2 = \{(y_1) : \tau_1 \leq y_1 \leq USL\}$, form the specification region. When $n=2$, four different sets, which are $\Omega_1 = \{(y_1, y_2) : LSL_1 \leq y_1 \leq \tau_1, LSL_2 \leq y_2 \leq \tau_2\}$, $\Omega_2 = \{(y_1, y_2) : \tau_1 \leq y_1 \leq USL_1, \tau_2 \leq y_2 \leq USL_2\}$, $\Omega_3 = \{(y_1, y_2) : \tau_1 \leq y_1 \leq USL_1, LSL_2 \leq y_2 \leq \tau_2\}$ and $\Omega_4 = \{(y_1, y_2) : \tau_1 \leq y_1 \leq USL_1 \text{ and } \tau_2 \leq y_2 \leq USL_2\}$, form the specification region. Similarly, for $n$ quality characteristics, $2^n$ different sets are realized as follows:

$\Omega_1 = \{(y_1, y_2, ..., y_n) : LSL_1 \leq y_1 \leq \tau_1, ..., LSL_n \leq y_n \leq \tau_n\}$

$\Omega_2 = \{(y_1, y_2, ..., y_n) : LSL_1 \leq y_1 \leq \tau_1, ..., LSL_n \leq y_n \leq \tau_n, \tau_n < y_n \leq USL_n\}$

$\Omega_3 = \{(y_1, y_2, ..., y_n) : \tau_1 \leq y_1 \leq USL_1, LSL_2 \leq y_2 \leq \tau_2, ..., \tau_n \leq y_n \leq USL_n\}$

$\Omega_4 = \{(y_1, y_2, ..., y_n) : \tau_1 \leq y_1 \leq USL_1 \text{ and } \tau_2 \leq y_2 \leq USL_2\}$

**Appendix B**

If there are two quality characteristics of interest and the first-order estimated empirical loss function is appropriate, then $\hat{L}(y_1, y_2)$ is given by:
\[ L(y_1, y_2) = \begin{cases} \hat{\beta}_{10} + \hat{\beta}_{11}y_1 + \hat{\beta}_{12}y_2, & LSL_1 \leq y_1 \leq \tau_1, \text{ and } LSL_2 \leq y_2 \leq \tau_2, \\ \hat{\beta}_{20} + \hat{\beta}_{21}y_1 + \hat{\beta}_{22}y_2, & LSL_1 \leq y_1 \leq \tau_1, \text{ and } LSL_2 < y_2 \leq USL_2, \\ \hat{\beta}_{30} + \hat{\beta}_{31}y_1 + \hat{\beta}_{32}y_2, & \tau_1 < y_1 \leq USL_1, \text{ and } LSL_2 \leq y_2 \leq \tau_2, \\ \hat{\beta}_{40} + \hat{\beta}_{41}y_1 + \hat{\beta}_{42}y_2, & \tau_1 < y_1 \leq USL_1, \text{ and } \tau_2 < y_2 \leq USL_2, \end{cases} \]

Similarly, if the second-order estimated empirical loss function is appropriate,
\[ L(\hat{y}_1, \hat{y}_2) \] is defined as:
\[ \tilde{L}(\hat{y}_1, \hat{y}_2) = \tilde{\alpha}_0 + \tilde{\alpha}_1y_1 + \tilde{\alpha}_2y_2 + \tilde{\alpha}_11y_1^2 + \tilde{\alpha}_22y_2^2, \]

Note that these models should be verified using lack-of-fit tests and supported by ANOVA tables.

**Appendix C**

Considering \( n \) quality characteristics, the probability that a product performance falls within region \( I \) can be shown as:
\[ P(R^i) = \int \int \cdots \int f(y_1, y_2, \ldots, y_n) dy_1 dy_2 \ldots dy_n \]
\[ + \int \int \cdots \int f(y_1, y_2, \ldots, y_n) dy_1 dy_2 \ldots dy_n \]
\[ + \cdots + \int \int \cdots \int f(y_1, y_2, \ldots, y_n) dy_1 dy_2 \ldots dy_n \]
\[ + \int \int \cdots \int f(y_1, y_2, \ldots, y_n) dy_1 dy_2 \ldots dy_n \]
\[ + \cdots + \int \int \cdots \int f(y_1, y_2, \ldots, y_n) dy_1 dy_2 \ldots dy_n \]

The probability that a product performance falls within region \( II \) can be obtained as:
\[ P(R^i) = \int \int \cdots \int f(y_1, y_2, \ldots, y_n) dy_1 dy_2 \ldots dy_n \]
\[ + \int \int \cdots \int f(y_1, y_2, \ldots, y_n) dy_1 dy_2 \ldots dy_n \]
\[ + \cdots + \int \int \cdots \int f(y_1, y_2, \ldots, y_n) dy_1 dy_2 \ldots dy_n \]

The probability that a product performance falls within region \( III \) can be shown as:
\[ P(R^i) = \int \int \cdots \int f(y_1, y_2, \ldots, y_n) dy_1 dy_2 \ldots dy_n \]
\[ + \int \int \cdots \int f(y_1, y_2, \ldots, y_n) dy_1 dy_2 \ldots dy_n \]
\[ + \cdots + \int \int \cdots \int f(y_1, y_2, \ldots, y_n) dy_1 dy_2 \ldots dy_n \]

Considering two quality characteristics as shown in Figure 1, rejection regions \( I \) and \( II \) represent the ones where a product performance fails to meet one of the specification sets and both specification sets, respectively. From a practical point of view, the rejection cost when a
The expected rejection cost associated with region 1, denoted by $E_1^r [C_R]$, can be shown as:

$$E_1^r [C_R] = C_R^l \cdot P(R_1^l),$$

where $C_R^l$ and $P(R_1^l)$ represent the unit rejection cost associated with region 1 and the probability that a product performance falls within region 1, respectively. That is:

$$E_1^r [C_R] = C_R^l \int_{USL_1}^{LSL_1} \int_{USL_1}^{LSL_1} f(y_1, y_2) dy_1 dy_2 + \int_{USL_1}^{LSL_1} \int_{USL_1}^{LSL_1} f(y_1, y_2) dy_1 dy_2,$$

where $f(y_1, y_2)$ is the joint probability density function of $y_1$ and $y_2$. Similarly, the expected rejection cost associated with region II, denoted by $E_2^r [C_R]$, can be obtained as:

$$E_2^r [C_R] = C_R^l \cdot P(R_2^l),$$

where $C_R^l$ and $P(R_2^l)$ represent the unit rejection cost associated with region II and the probability that the product performance falls within region II, respectively. Hence, the expected total rejection cost, denoted by $E^r [TC_R]$, is:

$$E^r [TC_R] = E_1^r [C_R] + E_2^r [C_R] = C_R^l \left( \int_{USL_1}^{LSL_1} \int_{USL_1}^{LSL_1} f(y_1, y_2) dy_1 dy_2 + \int_{USL_1}^{LSL_1} \int_{USL_1}^{LSL_1} f(y_1, y_2) dy_1 dy_2 \right) + C_R^l \left( \int_{USL_1}^{LSL_1} \int_{USL_1}^{LSL_1} f(y_1, y_2) dy_1 dy_2 + \int_{USL_1}^{LSL_1} \int_{USL_1}^{LSL_1} f(y_1, y_2) dy_1 dy_2 \right).$$

**References**


Production Research, 38(10), 2309-2325 (2000).


