Multiresponse Optimization of Powder Metals via Probabilistic Loss Functions

Abstract

Quadratic loss functions have been used extensively within the context of quality engineering and experimental design for process and product optimization and robust design. In general, this approach determines optimal parameter settings based on minimizing the sum of individual or mean loss of the associated response(s) of interest in a defined response surface. While the method is neat and handy, it totally neglects the effect of deviations on the desirable value of loss function. This paper utilizes variance and probability distribution of loss functions for developing an in depth optimization scheme that balances mean and variance of loss in a Pareto optimal manner. Since losses are usually defined in financial terms, this model then further improved to handle the user determined risk levels so that financial losses are being restricted within a certain region of interest. Application of the model is illustrated on a multiresponse optimization problem from powder metallurgy industry.

Word count = 151

Key Words: Loss functions; multiresponse optimization; experimental design; powder metallurgy; quality engineering; risk constrained

1. Introduction

Finding the best operating settings for industrial and various other processes is a challenging task for quality engineers, as the robustness of the process or product depends on different design parameters (inputs) that produce various responses (outputs) of interest. In many cases, these responses of interest conflict with each other and cannot be optimized simultaneously. Best balance between these competing responses can only be achieved by modeling the problem at hand as a multiresponse optimization problem.

In order to solve this dilemma, several optimization techniques have been developed with the help of designed experiments and response surface methodology. Desirability functions, first introduced by Harrington (1965) and enhanced by Derringer and Suich (1980), Wu (2004), and, Aksezer (2008) is one of the most widely used multiresponse optimization method both by researchers and modeling software developers. First, each predicted response is plotted to a desirability scale from 0 to 1(0 denoting an undesirable and 1 denoting a completely desirable value) based on the response's target value on the specification band and how much it deviates from this target, and then all weighted responses are incorporated to an overall desirability function with maximization objective. Taguchi's signal to noise perspective is also studied by numerous authors (Box 1988, Maghsoodloo 1990, Sii *et al.* 2001, Antony *et al.* 2006, and Xue *et al.* 2008) in order to generate cost effective optimization models. Kraus *et al.* (2000) which aims to maximize the probability of being within the specification limits subject to prediction equations on mean and variance of the response that is obtained from experimental model and Koksoy and Doganaksoy's (2003) dual response formulations which is based on mean and standard deviation, one being the objective constrained on the other, are the examples of mathematical programming approaches proposed within the literature.

Quadratic loss functions were introduced in the early 80's by Taguchi and Wu (1979) and more recently have been advocated by many researchers and practitioners (Hunter 1985, Ross 1988, Byrne and Taguchi 1987, Phadke 1989, Spiring 1993, Fowlkes and Creveling 1995, Benneyan and Aksezer 2006). Loss function establishes a financial measure of the user dissatisfaction with a product's performance as it deviates from target. The loss, $L(y)$, represents the total loss ("loss to society") of a response Y, such as a critical product dimension, taking on a specific value y. The goal is then to determine the design or operating parameters that minimize the total loss or expected loss, which might include both immediate and less tangible costs of failure, warranty, rework, replacement, customer ill-will, environmental impact, product perception, and others. While Taguchi used loss functions as a single response optimization tool, since then many approaches are developed for utilizing these functions in multiresponse problems. These studies focus on two different kinds of loss modeling; loss at a certain point on the response design surface or mean loss on overall response surface (Pignatiello and Ramberg 1991, Artiles-Leon 1996, Kapur and Cho 1996, Ames *et al.* 1997, Spiring and Yeung 1998).

The purpose of this paper is to expand the mean loss modeling to include the variance and probabilistic nature of loss functions. Next section briefly revisits the variance and probability distribution (PDF) of loss functions and then proposes a new optimization scheme that transforms the problem into a multiresponse optimization problem. Lastly, use of the model is illustrated on an optimal powder metal production process. Results gathered are discussed and compared by means of effectiveness and applicability.

2. Loss Functions

There are three most common quadratic loss functions in the literature for applications in which the objectives are to achieve values (1) as close to a target as possible, (2) as small as possible, or (3) as large as possible. These 3 general cases ("nominal-the-best", "smaller-the-better", and "larger-the-better") are illustrated in Figure 1 and defined mathematically as:

Nominal-the-best (NTB):

$$L_N(Y) = k(Y-T)^2 \tag{1}$$

Smaller-the-better (STB):

$$L_S(Y) = k(Y)^2 \tag{2}$$

Larger-the-better (LTB):

$$L_L(Y) = k\left(\frac{1}{Y}\right)^2 \tag{3}$$

where *L(Y)* is the total loss due to deviation from target, *Y* is the random variable of interest, *T* is the target value for the product's response in the nominal-the-best case, and *k* is a constant sometimes referred to as the quality loss coefficient. In each case, increasing losses are incurred as the measurement *Y* deviates by greater amounts from its desired target, irrespective of specifications.

Figure 1

Long term (life cycle) loss optimization procedure only involves the minimization of expected loss over the response surface. However, in some applications along with a measure of location, it also is of interest to measure the variability of the loss acquired and the distribution shape of the loss incurred from the process. These properties can be helpful in understanding the behavior of the process in details. For instance; a product involving multiple quality characteristics can achieve a small expected loss on a given response surface. But in the long run, the deviations from the expected loss are also important for the assessment and management of operational risks. In order to calculate these effects, practitioners have to go one step further from the classical optimization technique of minimization of loss per unit or expected loss and should include other measures such as variance and distribution of loss into the optimization model.

2.1. Expectation and Variance of a Loss Function

Assume that a response Y with mean $E(Y) = \mu$ and variance $V(Y) = \sigma^2$ is considered with k^{th} non-central moment about the origin $E(Y^k) = \mu_k^{'}$ and k^{th} central moment about the mean μ_k. Expectation and variance term of each case then can be derived as following. For a NTB response; $L_N(Y) = k(Y-T)^2$ and since $V(Y) = E(Y^2) - \mu^2$ with $E(Y^2) = \sigma^2 + \mu^2$, then

$$E[L_N(Y)] = E\left[k(Y-T)^2\right] = kE\left[Y^2 - 2YT + T^2\right]$$

$$= k\left[E(Y^2) - E(2TY) + E(T^2)\right] = k\left[(\sigma^2 + \mu^2) - 2TE(Y) + T^2\right] ,$$

$$E[L_N(Y)] = k\left[\sigma^2 + (\mu^2 - 2T\mu + T^2)\right] = k\left[\sigma^2 + (\mu - T)^2\right]$$

and

$$V[L_N(Y)] = E\left[(L_N(Y))^2\right] - E^2[L_N(Y)]$$

$$= E\left[\left(k(Y-T)^2\right)^2\right] - E\left[k\left(\sigma^2 + (\mu - T)^2\right)\right]^2 ,$$

$$= k^2 E\left(Y^4 - 4Y^3T + 6Y^2T^2 - 4YT^3 + T^4\right) - k^2\left(\sigma^4 + 2\sigma^2(\mu - T)^2 + (\mu - T)^4\right),$$

which can further be simplified into,

$$V[L_N(Y)] = k^2\left[\mu_4' + 4T\left(T\sigma^2 + \mu(\sigma^2 + \mu^2) - \mu_3'\right) - \left(\sigma^2 + \mu^2\right)^2\right].$$

In a similar fashion, for a STB response; given $L_S(Y) = kY^2$ then

$$E[L_S(Y)] = E\left[kY^2\right] = k\left[E\left(Y^2\right)\right] = k\left[\sigma^2 + \mu^2\right]$$

and

$$V[L_S(Y)] = E\left[L_S(Y)^2\right] - E[L_S(Y)]^2 = E\left(k^2Y^4\right) - \left[k\left(\sigma^2 + \mu^2\right)\right]^2,$$

$$= k^2\left[E\left(Y^4\right) - \left(\sigma^2 + \mu^2\right)^2\right] = k^2\left[\mu_4' - \left(\sigma^2 + \mu^2\right)^2\right],$$

where $\mu_k' = E\left(Y^k\right) = \int_y y^k f(y)dy$ is the k^{th} (here the 4^{th}) moment about the origin.

Since in general $E(1/X) \neq 1/E(X)$, LTB does not have an exact expression for the mean and variance loss. However, a series approximation can be used by manipulating the formulation.

$$E[L_L(Y)] = kE\left[\frac{1}{Y^2}\right] = kE\left[\frac{1}{(Y - \mu + \mu)^2} \cdot \frac{\mu^2}{\mu^2}\right] = \frac{k}{\mu^2} E\left[\left(1 + \left(\frac{Y - \mu}{\mu}\right)\right)^{-2}\right]$$

Using the Taylor expansion, this becomes

$$E[L_L(Y)] = \frac{k}{\mu^2} E\left[1 - 2\left(\frac{Y - \mu}{\mu}\right) + 3\left(\frac{Y - \mu}{\mu}\right)^2 - 4\left(\frac{Y - \mu}{\mu}\right)^3 + 5\left(\frac{Y - \mu}{\mu}\right)^4 - \ldots\right]$$

$$= \frac{k}{\mu^2}\left[1 + \frac{3\mu_2}{\mu^2} - \frac{4\mu_3}{\mu^3} + \frac{5\mu_4}{\mu^4} - \frac{6\mu_5}{\mu^5} + \frac{7\mu_6}{\mu^6} - \ldots\right], \text{ since } E(Y - \mu) = 0,$$

where $\mu_k = E(Y - \mu)^k$, the k^{th} central moment about the mean. The mean loss therefore can be approximated using only the first four central moments as

$$E[L_L(Y)] \approx \frac{k}{\mu^2}\left[1+\frac{3\sigma^2}{\mu^2}-\frac{4\mu_3}{\mu^3}+\frac{5\mu_4}{\mu^4}\right]=\frac{k}{\mu^2}\left[20+\frac{45\sigma^2}{\mu^2}-\frac{24\mu_3^{'}}{\mu^3}+\frac{5\mu_4^{'}}{\mu^4}\right].$$

Using a similar series approximation approach, the variance of $L_L(Y)$ can be shown as

$$V[L_L(Y)] = E\left[(L_L(Y))^2\right] - E^2[L_L(Y)] = E\left(\frac{k}{Y^2}\right)^2 - \left[E\left(\frac{k}{Y^2}\right)\right]^2$$

$$= \frac{k^2}{\mu^4} E\left(1+\left(\frac{Y-\mu}{\mu}\right)\right)^{-4} - \left[\frac{k}{\mu^2} E\left(1+\left(\frac{Y-\mu}{\mu}\right)\right)^{-2}\right]^2,$$

and by using only the first five terms in each expansion, the approximation becomes

$$V[L_L(Y)] \approx \frac{k^2}{\mu^{12}}\left[\mu_4^{'}\left(-165\mu^4 - 450\mu^2\sigma^2 - 25\mu_4^{'}\right) + \mu_3^{'}\left(800\mu^5 + 2160\mu^3\sigma^2 - 576\mu_3^{'}\mu^2 + 240\mu_4^{'}\mu\right) - \mu^4\sigma^2\left(274\mu^4 + 1520\mu^2 + 2025\sigma^2\right)\right]$$

in terms of central and noncentral moments, respectively. Note that, if the distribution of Y is not very skewed or kurtotic (e.g., normal) so that $\mu_3 \approx 0$ and $\mu >> \mu_4$, then the higher central moments are negligible. Then the mean and variance loss approximates to

$$E[L_L(Y)] \approx \frac{k}{\mu^2}\left[1+\frac{3\sigma^2}{\mu^2}\right] \text{ and, } V[L_L(Y)] \approx \frac{k^2\sigma^2}{\mu^8}\left[4\mu^2 - 9\sigma^2\right].$$

2.2. Probability Density of a Loss Function

Let y be a continuous random variable of a function *f(y)* with known distribution over a sample space and $z = h(y)$ is a strictly increasing or a strictly decreasing function of x over the same sample space. Then the PDF of a transformed continuous random variable z can be written as $p(z) = f[y(z)]\left|\frac{dy}{dz}\right|$, where $\left|\frac{dy}{dz}\right|$ is the *Jacobian* of the transformation and $f[y(z)]$ is the functional relationship substitution between random variables z in terms of y.

The transformation of a discrete random variable can be accomplished in a similar fashion and will be identical to the form showed above except for the Jacobian term that is used for the mapping of a continuous function. For nominal-the-best case, the functional relationship between L and Y and the Jacobian term is found as

$$L(y) = \ell = k(y-T)^2 \Rightarrow y = \mp\sqrt{\frac{\ell}{k}} + T \quad \text{and} \quad \left|\frac{dy}{d\ell}\right| = \frac{1}{2\sqrt{\ell k}} \quad .$$

The probability density function for the nominal-the-best loss case then becomes

$$p(\ell) = \frac{1}{2\sqrt{k\ell}}\left[f\left(\sqrt{\frac{\ell}{k}} + T\right) + f\left(-\sqrt{\frac{\ell}{k}} + T\right)\right].$$

Following a similar notion for smaller-the-better case,

$$L(y) = \ell = ky^2 \Rightarrow y = \mp\sqrt{\frac{\ell}{k}} \quad \text{and} \quad \left|\frac{dy}{d\ell}\right| = \frac{1}{2\sqrt{\ell k}},$$

respectively, and the probability density function for smaller-the-better loss case becomes

$$p(\ell) = \frac{1}{2\sqrt{k\ell}}\left[f\left(\sqrt{\frac{\ell}{k}}\right) + f\left(-\sqrt{\frac{\ell}{k}}\right)\right].$$

Similarly, for larger-the-better case, the functional relationship between L and Y and the Jacobinan term are

$$L(y) = \ell = k\frac{1}{y^2} \Rightarrow y = \mp\sqrt{\frac{k}{\ell}} \quad \text{and} \quad \left|\frac{dy}{d\ell}\right| = \frac{\sqrt{k}}{2\sqrt{\ell^3}},$$

respectively, and the probability density function for larger-the-better loss case becomes

$$p(\ell) = \frac{\sqrt{k}}{2\sqrt{\ell^3}}\left[f\left(\sqrt{\frac{k}{\ell}}\right) + f\left(-\sqrt{\frac{k}{\ell}}\right)\right].$$

Interestingly, expressions for the variance include the first four moments of *Y* and thus are related to not only the mean and variance of the response *Y* but also its skewness and kurtosis. These may then be simplified to common cases where *Y* follows a normal distribution, as well as various other response distributions (lognormal, Weibull, exponential, and uniform) where higher ordered moments are readily available. Application of normal distribution is especially important since many problems can be normalized through proper transformation. Also, normal distribution has a skewness of 0 and kurtosis of 3 which leads to simplified versions of variance and PDF expressions of the associated loss as illustrated in Table 1. Lognormal, Weibull, and exponential distributions may also be appropriate in quality and reliability applications or for strictly non-negative responses, especially when the mean is close to zero.

Table 1

3. Loss Modeling

The classical optimization approach minimizes the expected loss subject to a set of constraints on the process parameters such as the relationships among responses and process factors (prediction equations), distribution of the quality characteristic in evaluation, specific costs incurred from quality control, inspection, maintenance, labor, tools etc., tolerance and specification limits on responses, and process capability indices. The general form of this kind of model for *m* responses of interest along with *n* model variables can be given as the following:

$$
\begin{aligned}
&\textit{Minimize} \qquad \sum_{i=1}^{m} E_L\left(y_i, T_i\right) \\
&\textit{subject to} \\
&\qquad \mu_m = \beta_0 + \sum_{j=1}^{n} \beta_j x_j \\
&\qquad \sigma_m = \alpha_0 + \sum_{j=1}^{n} \alpha_k x_k
\end{aligned}
\tag{4}
$$

The availability of data necessary to identify the functional relationship between these parameters and responses may not readily be available. For example; the costs associated from each parameter setting, necessary for the realization of loss coefficient, cannot be easily calculated in many production systems. In reality, this coefficient is calculated according to familiarity with the process, meaning that the quality engineer can make the judgment of economic losses due to off target response, or if the cost at the specification limits are known (this is the cost incurred when the response is totally undesirable), then it can be calculated for any point on the response surface. However, there are kinds of processes for which the practitioner cannot calculate this coefficient at all. Then, instead of true costs, estimates of the comparative importance of different characteristics are generally evaluated. These relative importance factors are used to weight the individual losses from each response and integrate them into an expected overall loss on the whole response surface, such as applied in desirability function methodology.

A similar approach to the one given above is the minimization of variance of loss. This is especially important since higher variations in loss would signal inconsistencies in the product or process. The objective function (5) involves the variance of the associated loss function from Table 1 and is the sum of all variances from individual response y_j in multiresponse problems. The constraints are once again the prediction equations on mean and standard deviation of the associated response. If there is a correlation between the quality characteristics then the interaction terms and covariances should be included to the formulation. Model can be given as the following:

$$\textit{Minimize} \qquad \sum_{i=1}^{m} V_L\left(y_i, T_i\right)$$

$$\textit{subject to}$$

$$\mu_m = \beta_0 + \sum_{j=1}^{n} \beta_j x_j \qquad (5)$$

$$\sigma_m = \alpha_0 + \sum_{j=1}^{n} \alpha_j x_j$$

$$\mu_m^{'} = \delta_0 + \sum_{j=1}^{n} \delta_j x_j$$

The question to be investigated then becomes what exactly is the difference between minimizing the expected loss and the variance of loss. While in some problems both lead to the same product settings, there are cases where certain tradeoffs must be made between these two.

Sequential minimization of expected loss and variance of loss clearly illustrates this tradeoff and proposes a dual response system that minimizes these simultaneously. The problem becomes a multiple objective model, which can be solved by ε-constraint approach that reduces the feasible region by introducing one of the objectives as a constraint at its threshold level and optimizing the other on this reduced region. The problem formulation then becomes:

$$\textit{Minimize} \qquad \sum_{i=1}^{m} E_L(y_i, T_i)$$

subject to

$$\sum_{i=1}^{m} V_L(y_i, T_i) \le \varepsilon$$

$$\mu_m = \beta_0 + \sum_{j=1}^{n} \beta_j x_j \tag{6}$$

$$\sigma_m = \alpha_0 + \sum_{j=1}^{n} \alpha_j x_j$$

$$\mu'_m = \delta_0 + \sum_{j=1}^{n} \delta_j x_j$$

ε : upper bound on variance

This model utilizes expected loss as the primary objective and the associated mean loss expression is used in the objective function, while the variance of loss becomes secondary objective and treated as a constraint bounded by ε, which is the desired lower bound constant for variance of loss while minimizing the expected loss. It is almost always impossible to predetermine ε, instead the analyst goes through several different values of ε, which provide the feasible tradeoff frontier between mean and variance loss. Any point on this frontier will be Pareto optimum and can be chosen as optimal solution, based on the decision criteria.

It is also practical to present a probabilistic constraint to the model that involves the distribution of the loss function. This transforms the proposed multiple criteria optimization model into a minimization of total expected loss subject to minimal valued variance of loss and a risk constraint in which the probability of the total loss, given the upper bound for sum of the total loss, is limited to a certain value *p* within the design surface. Under the assumption of independency between responses, we can calculate the joint probability function as the simple product of all individual PDFs. However, when a dependency exists among responses, product of PDFs leads to incorrect results. Instead, the PDFs of individual responses have to be convoluted on each other and be used in the model. Because of the complexity of PDFs, there is no closed form to this convolution. Only when the number of competing responses is assumed to be large enough and independent, we know from the central limit theorem that the resulting overall distribution should approach to normal distribution.

$$\text{Minimize} \qquad \sum_{i=1}^{m} E_L(y_i, T_i)$$

subject to

$$\begin{aligned}
&\sum_{i=1}^{m} V_L(y_i, T_i) \leq \varepsilon \\
&F\left[f(l_1) \otimes f(l_2) \otimes \ldots \otimes f(l_m)\right] \geq p \\
&\mu_m = \beta_0 + \sum_{j=1}^{n} \beta_j x_j \\
&\sigma_m = \alpha_0 + \sum_{j=1}^{n} \alpha_j x_j \\
&\mu_m' = \delta_0 + \sum_{j=1}^{n} \delta_j x_j
\end{aligned} \tag{7}$$

ε: upper bound on variance

p: probabilistic risk level $(0 < p < 1)$

4. Numerical Example

To illustrate the application of the models given above, we will use an example from powder metal production. Today, spherical metal powders are used in many applications ranging from aerospace to medical implants for paint and varnish material production, as catalyst in chemical processes, for thermal diffusion, and as antifriction & antiwear components in forms of aluminum, titanium, zinc, copper etc. They are desirable in these applications because of the flow characteristics of the powder and the resulting packing efficiencies for increased part densities (Upadhyaya 1998, Minagawa *et al.* 2005). In order to benefit from these characteristics, metal powders must be perfectly spherical and clean. While the sphericity of a certain particle is measured by its desirable diameter, the cleanliness is measured by the smoothness of the surface. This smoothness can only be achieved by manufacturing the particle surface free of satellites. Satellites are the smaller particles that are sintered to the powder itself (around the surface).

There exist several manufacturing methods in the literature for powder metallurgy. Solid state reduction, atomization and centrifugal disintegration are the most prominent methods used in practice. Problem at hand is based on manufacturing via centrifugal disintegration method in which the metal to be powdered is formed into a rod that is introduced into a chamber through a rapidly rotating spindle. Opposite to the spindle tip there is an electrode from which an arc is established, heating the metal rod. As the tip material fuses, the rapid rod rotation throws off tiny melt droplets, which solidify before hitting the chamber walls. A circulating gas sweeps particles from the chamber. Based on these production goals, the design has two response characteristics:

- Mean Diameter (micron, μm)
- Satellites (ratio, %)

Note that the chosen powder characteristics are usually based on compromise, since many of the factors are in direct conflict with each other. Obtaining larger spherical particles will increase irregularity on the surface since contact surface gets larger. The first response is an NTB type of response where the target is to produce particles at a certain size and the second is an STB type of response seeking to minimize the satellites on the particle. Since these particles are very small, the measure to be minimized will be the percentage of the ratio of the largest satellite diameter on the surface of the particle to the powder diameter.

Typical process parameters that affect the output characteristics of powder metal production with centrifuging are the rotational speed (x_1) of the electrode which is typically between 15000 – 17000 rotations per minute, amperage (x_2) supplied to generate a plasma arc is being between 925 – 1025 amps and gas pressure (x_3) necessary to flow the particles out of the chamber is measured to be between 90 – 110 psi. These process settings are used in a 2-level full factorial design with two replications as illustrated on Table 2 in standard order with coded terms. Higher level designs or inclusion of center points are not generally advisable to be used in multiresponse optimization schemes. One of the reasons for not considering this type of designs is the high ordered (e.g. quadratic) nature of the resulting prediction equations. Such nonlinear response equations are not always guaranteed to be globally optimized by common algorithms used in statistical software packages. These designs also require more runs and complicating data collection process even further when used in conjunction with replications.

The process constraints for this kind of problem are the specification levels for the diameter. The process has a lower specification level of 90μm and an upper specification limit of 110μm and the target value is the midpoint of this specification band which is 100μm.

Table 2

Prediction Equations for Diameter:

$$\hat{\mu} = 96 - 6.56x_1 - 2.81x_2 - 13x_3 - 2.56x_1x_2 + 2.56x_2x_3$$

$$\hat{\sigma} = 5.74 - 0.62x_1 - 0.27x_2 - 2.21x_3$$

Prediction Equations for Smoothness:

$$\hat{\mu} = 18.88 - 3.13x_1 - 1.38x_2 - 3.25x_3 - 0.25x_1x_3 + 0.75x_2x_3$$

$$\hat{\sigma} = 2.12 + 0.35x_1 + 0.53x_3 + 0.18x_1x_3$$

The process is found to be normally distributed after a least squares fit of each response and an *R-sqr* > 0.98, allowing the use of the simplified moment terms in the general form of the loss equations, which only require the mean and standard deviation. As noted by Pignatiello (1993), loss function based optimization methodologies can only be used with multiply replicated experimental design, in which the experiment itself can be costly because of the multiple runs. However, multiple runs are crucial to identify the necessary prediction equations for the standard deviation, the third and the fourth moment terms.

Solving this problem separately both for minimization of expected loss (4) and variance of loss (5), yields to the following minimum achievable values. Note that these are optimized individually without being constraint on each other.

$$Min(E_L) = \$356.25$$
$$Min(V_L) = \$9447$$

Although these values imply some economical terms, this study assumes fictitious dollar units since there exist no reasonable loss coefficient *k* available at hand. We simply ignored this constant and assumed it as 1 for all responses. This means that deviations from each response will cause same amount of financial impact.

These two responses of interest are then combined by the proposed hybrid model. This model provides a general idea of how the two objectives interact with each other within the objective space and whether they could be optimized simultaneously. Solving the problem simultaneously by minimizing the expected loss, subject to constraint on different feasible values of the variance loss leads to the tradeoff frontier shown in Figure 2. Analyses of the frontier allow us to understand the tradeoffs of the process and to find a Pareto optimum point, which best fits the process or product needs. The steep slope of the curve shows that by sacrificing some from the mean loss, the process gains much from the minimization of the variability of the loss.

As previously discussed, the application of the last model involves the PDF convolution for the responses. This convolution should be included in the model as the probabilistic constraint that balances the probability of total loss being less than a certain value. Once again, selecting the total loss risk level requires familiarity with the process at hand, but by also running the model for different values of total loss, the analyst can find the feasible value that leads to the desired solution. For the given example, this last model can be constructed as following:

$$\begin{aligned} &\textit{Minimize} \quad E_L \\ &\textit{subject to} \\ &\qquad V_L \leq \$10000 \\ &\qquad P(\textit{Total Loss} \leq \$500) \geq 0.7 \\ &\qquad -1 \leq x_i \leq 1, \quad i = 1 \ldots n \end{aligned} \tag{8}$$

Solution of this final and most insightful model yields to a result of minimum expected loss of, E_L = \$369, coded parameter setting of (x_1, x_2, x_3) = (1, -0.91, -0.23) and actual parameter setting of (17000, 932, 97.7).

Figure 2

5. Conclusions

This paper presents the use of variances and probability distributions of quadratic loss functions in various optimization schemes. Since loss functions are an invaluable tool for multiresponse design optimization problems, inclusion of these characteristics to the model will extend the current expected loss minimization conception to analyze the further aspects of behavior of the product or process in question under the assumption of known response distributions. Previous results indicate that all three types of loss functions frequently exhibit high variance and high skew, which can be important in determining optimal production and operating conditions. In many cases, the loss distributions are significantly asymmetrical and unique in shape proving the importance of using these characteristics in the optimization schemes.

The proposed models provide the opportunity of an in depth look both from practitioners' and theoreticians' point of view by minimizing expected loss along with the variance of loss in a dual manner, producing a flexible and beneficial Pareto optimal solution set that consists all optimal solution pairs. The solution can also be illustrated in a more fashionable way by constructing a collection of all Pareto optimal solutions as a tradeoff frontier that will even highlight the conflicts between responses clearer. Depending on our process characteristics and product needs, most sensitive point on this frontier then can be chosen as the solution to the problem at hand where the feasible area of the solution set can be reduced by the stated probabilistic constraint written in the form of the product of the loss PDFs of the associated response's.

The only troublesome part of the optimization process seems to be the necessity of multiple replications in the planning and conducting phase of the experimental design. These replications are crucial for the estimation of standard deviation and the necessary higher ordered moment terms. More importantly, replicated designs become more robust by identifying the noise present in the design. The models also tend to be statistically significant as a result of higher sample size used in the estimation of response variance. However, replicated designs can be costly and time consuming depending on the data collection process environment so the designer should plan accordingly.

An increase in the number of competing responses usually complicates the marginal rate of substitution mechanism of the Pareto optimal domain, which eventually erodes the computational efficiency of any multiresponse modeling technique. Authors are currently investigating the performance of the proposed models compared to the alternative methods from literature with respect to computational efficiency in reaching global optimum. Extension of the models to include the sensitivity analysis of its primary parameters (input variables that are excluded from the experimental design; e.g. loss coefficient) on the Pareto optimal frontier would also appear to be an important direction for future research.

References

Aksezer, S.C. (2008) 'On the Sensitivity of Desirability Functions for Multiresponse Optimization', *Journal of Industrial and Management Optimization*, Vol. 4, No. 4, pp. 685-696.

Ames, A.A., Mattucci, N., Macdonald, S., Szonyi, G. and Hawkins, D.M. (1997) 'Quality Loss Functions For Optimization Across Multiple Response Surfaces', *Journal of Quality Technology*, Vol. 29, No. 3, pp.339-346.

Antony, J., Anand, R.B., Kumar, M. and Tiwari, M.K. (2006) 'Multiple response optimization using Taguchi methodology and neurofuzzy based model', *Journal of Manufacturing Technology Management*, Vol. 17, No. 7, pp.908-925.

Artiles-Leon, N. (1997) 'A Pragmatic Approach To Multiple Response Problems Using Loss Functions', *Quality Engineering*, Vol. 9, No. 2, pp.213-220.

Benneyan, J. and Aksezer, S.C. (2006) "Probability Distributions and Variances of Quadratic Loss Functions", *Proceedings of the 36th International Conference on Computers and Industrial Engineering*, Taipei, Taiwan, pp.2890-2899.

Byrne, D.M. and Taguchi, S. (1987) 'The Taguchi Approach to Parameter Design', *Quality Progress*, Vol. 20, No. 12, pp.19-26.

Box, G.E.P. (1988) 'Signal-to-Noise Ratios, Performance Criteria, and Transformations', *Technometrics*, Vol. 30, No. 1, pp.1-40.

Derringer, G.C. and Suich, R. (1980) 'Simultaneous Optimization of Several Response Variables', *Journal of Quality Technology*, Vol. 12, No. 4, pp.214-219.

Fowlkes, W.Y. and Creveling, C.M. (1995) *Engineering Methods for Robust Product Design: Using Taguchi Methods in Technology and Product Development*, Addison-Wesley, Massachusetts.

Harrington, E.C. (1965) 'The Desirability Function', *Industrial Quality Control*, Vol. 21, No. 10, pp.494-498.

Hunter, J.S. (1985) 'Statistical Design Applied to Product Design', *Journal of Quality Technology*, Vol. 17, No. 4, pp.210-221.

Kapur, K.C. and Cho, B. (1996)'Economic Design of the Specification Region For Multiple Quality Characteristics', *IIE Transactions*, Vol. 28, No. 3, pp.237-248.

Kraus, P.D., Benneyan, J.C. and Mackertich, N.A. (2000) 'Use of mathematical programming in the analysis of constrained and unconstrained industrial experiments', *Quality Engineering*, Vol. 12, No. 3, pp.395-406.

Koksoy, O. and Doganaksoy, N. (2003) 'Joint optimization of mean and standard deviation using response surface methods', *Journal of Quality Technology*, Vol. 35, No. 3, pp.239-252.

Maghsoodloo, S. (1990) ‘The Exact Relation of Taguchi’s Signal-to-Noise Ratios to His Quality Loss Function’, *Journal of Quality Technology*, Vol. 22, No. 1, pp.57-67.

Minagawa, K., Kakisawa, H., Osawa, Y., Takamori, S. and Halada K. (2005) ‘Production of fine spherical lead-free solder powders by hybrid atomization’, *Science and Technology of Advanced Materials*, Vol. 6, pp.325-329.

Phadke, M.S. (1989) *Quality Engineering Using Robust Design*, Prentice-Hall, Princeton.

Pignatiello, J.J. and Ramberg, J.S. (1991) ‘The Ten Triumphs and Tragedies of Genichi Taguchi’, *Quality Engineering*, Vol. 4, No. 2, pp.211-225.

Pignatiello, J.J. (1993) ‘Strategies For Robust MultiResponse Quality Engineering’, *IIE Transactions*, Vol. 25, No. 3, pp.5-15.

Ross, P.J. (1988) *Taguchi Techniques for Quality Engineering*, McGraw-Hill, New York.

Sii, H.S., Ruxton, T. and Wang, J. (2001) ‘Taguchi concepts and their applications in marine and offshore safety studies’, *Journal of Engineering Design*, Vol. 12, No. 4, pp. 331-358.

Spiring, F.A. (1993) ‘The Reflected Normal Loss Function’, *The Canadian Journal of Statistics*, Vol. 21, No. 3, pp.321-330.

Spiring, F.A. and Yeung, A. (1998) ‘A General Class of Loss Functions with Industrial Applications’, *Journal of Quality Technology*, Vol. 30, No. 2, pp.152-162.

Taguchi, G. and Wu, Y. (1979) *Introduction to Off-Line Quality Control*, Central Japan Quality Control Association (available from the American Supplier Institute), Dearborn.

Upadhyaya, G.S. (1998) *Powder Metallurgy Technology*, Cambridge International Science Publishing.

Wu, F-C. (2004) ‘Optimization of Correlated Multiple Quality Characteristics Using Desirability Function’, *Quality Engineering*, Vol. 17, No. 1, pp.119-126.

Xue, D., Cheing, S.Y. and Gu, P. (2008) ‘Parameter design considering the impact of design changes on downstream processes based upon the Taguchi method’, *Journal of Engineering Design*, Vol. 19, No. 4, pp.299-319.

Table 1. Summary of loss function characteristics

	Nominal-the-Best	***Smaller-the-Better***	***Larger-the-Better***
Loss Function	$k(y-T)^2$	$k(y)^2$	$k\left(1/y\right)^2$
Expected Loss	$k\left[\sigma^2+(\mu-T)^2\right]$	$k\left[\sigma^2+\mu^2\right]$	$\cong \frac{k}{\mu^2}\left[1+\frac{3\sigma^2}{\mu^2}-\frac{4\mu_3}{\mu^3}+\frac{5\mu_4}{\mu^4}\right]$
Variance Loss	$k^2\left[\begin{array}{l}\mu_4^{'}+4T\left(T\sigma^2+\mu(\sigma^2+\mu^2)-\mu_3^{'}\right)\\ -\left(\sigma^2+\mu^2\right)\end{array}\right]$	$k^2\left[\mu_4^{'}-\left(\sigma^2+\mu^2\right)^2\right]$	$\cong \frac{k^2}{\mu^{12}}\left[\begin{array}{l}\mu_4^{'}\left(-165\mu^4-450\mu^2\sigma^2-25\mu_4^{'}\right)+\\ \mu_3^{'}\left(800\mu^5+2160\mu^3\sigma^2-576\mu_3^{'}\mu^2+240\mu_4^{'}\mu\right)\\ -\mu^4\sigma^2\left(274\mu^4+1520\mu^2+2025\sigma^2\right)\end{array}\right]$

Table 2. Data for 2^3-full factorial replicated design

Run#	Speed	Amperage	Pressure	Mean Diameter	Smoothness
1	-	-	-	125	26
2	+	-	-	103	22
3	-	+	-	121	23
4	+	+	-	88	19
5	-	-	+	88	18
6	+	-	+	82	11
7	-	+	+	84	21
8	+	+	+	75	9
9	-	-	-	112	29
10	+	-	-	117	20
11	-	+	-	110	22
12	+	+	-	95	16
13	-	-	+	86	20
14	+	-	+	77	16
15	-	+	+	94	17
16	+	+	+	78	13

	Speed(rpm)	Amperage(amp)	Pressure(lb/in2)
-	15000	925	90
+	17000	1075	110

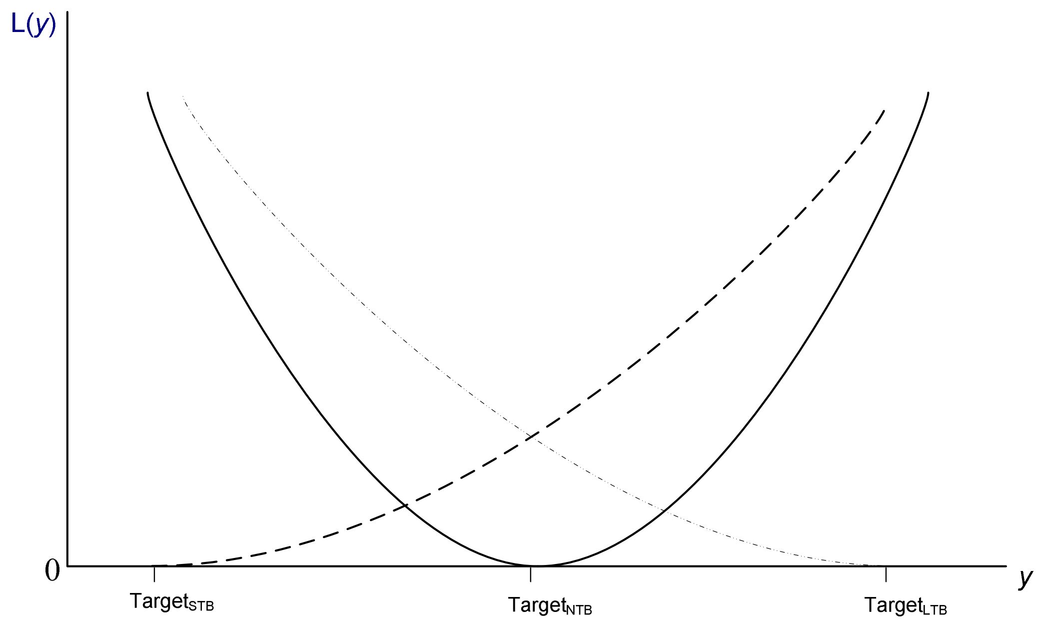

Figure 1. Types of response optimization problems

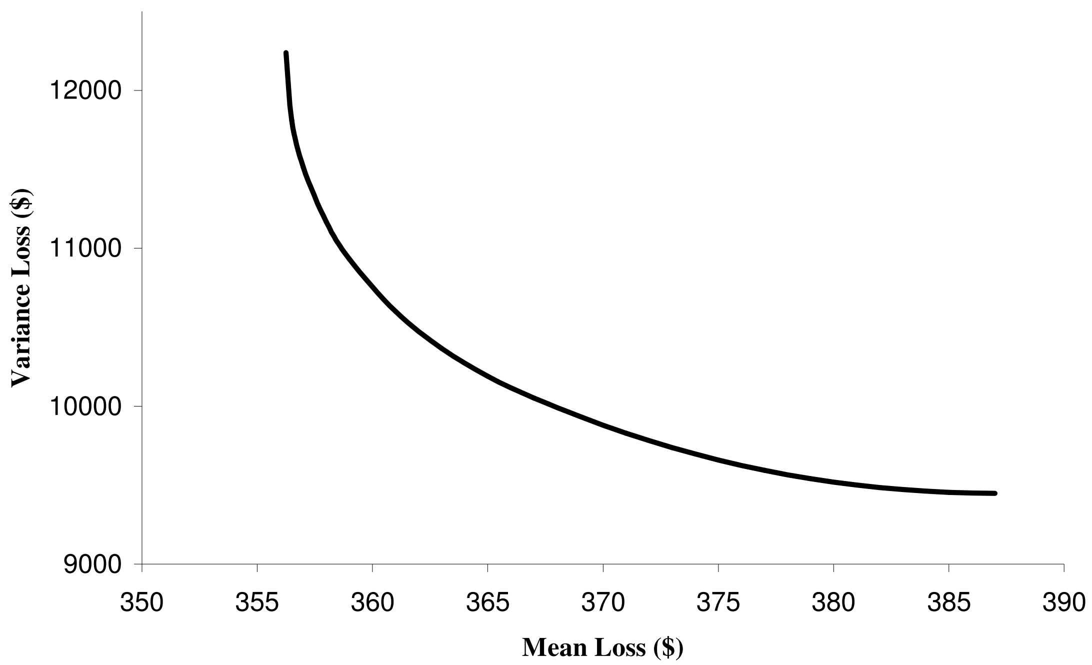

Figure 2. Tradeoff frontier for mean and variance of loss