COST-DRIVEN PARAMETER DESIGN

P.R. Moorhead and C.F.J. Wu

IIQP Research Report
RR-93-09

December 1993

COST-DRIVEN PARAMETER DESIGN

P.R. Moorhead and C.F.J. Wu
Department of Statistics and Actuarial Science
University of Waterloo, Waterloo, Ontario, Canada

December 1993

ABSTRACT

Parameter design is a method for improving quality in a product or manufacturing process by minimizing expected loss due to noise variation. Understanding and development of the methodology has focused primarily on the quadratic loss function, which leads to a two step procedure involving the minimization of a dispersion measure and then adjusting the mean to target. In some practical situations, however, the loss can be far from quadratic as would be the case if it were highly skewed. By building on a theory developed by León and Wu (1992), we develop a modelling and data analysis strategy for parameter design experiments with general loss functions. It turns out that the only additional effort is to model a location measure which is different from the mean for non-quadratic loss. The technique is illustrated on an experiment involving epitaxial layer growth in IC fabrication.

Key words: General loss function, Nominal-the-best parameter design, Additive model, Dispersion measure, Location measure, Adjustment factor.

1 Introduction

The goal of parameter design, as introduced by the Japanese quality expert G. Taguchi, is to improve quality by making the product/process insensitive to variations in the *noise* factors. Because these factors are difficult or expensive to control, Taguchi advocates minimizing their effect by exploiting their interactions with *control* factors, which are easy to manipulate. It is often convenient to measure the quality of a product/process in terms of the closeness of some quantitative characteristic to an intended target. When the target is fixed at a finite (usually nonzero) value, the problem is referred to as *nominal-the-best.* A standard approach in this case is to measure sensitivity to the noise factors with the quadratic loss function, i.e. squared deviation from the target, and then to identify those settings of the control factors which minimize expected loss. Nair (1992), Section 2, contains a comprehensive discussion of these issues.

For parameter design problems of this nature, many improvements to Taguchi's methodology appear in the statistical literature; however, until León and Wu (1992), the loss function was invariably assumed to be quadratic. It is important to recognize that this loss function will not be adequate in all cases. For example, when a windshield deviates from its nominal size the loss is not symmetric. Smaller ones will be scrapped, whereas slightly larger ones can be accommodated by the molding around the frame. Similarly, when a brake rotor is too soft, it must be scrapped. However, if it is too hard, the squealing sound produced may be an annoyance yet the rotor still functions adequately. Asymmetric loss can also be observed in the food packaging industry where under- and over-fill have unequal effects on the producer's reputation. In many situations, quadratic loss over penalizes large deviations from the target. In particular, when the specification limits are violated and the product is scrapped, the loss is usually bounded by the cost of production. The purpose of this paper is to develop a modeling and data analysis strategy for general loss functions.

As previously mentioned, a product/process is usually monitored by some quality characteristic or output y. Following León, Shoemaker and Kacker (1987) (henceforth denoted by LSK) and León and Wu (1992) (henceforth denoted by LW), for given settings of the control factors Θ and the noise factors N, the output y is determined by some transfer function $f(N;\Theta)$. The noise is assumed random and, when the output deviates from an ideal target t, a loss proportional to $(y-t)^2$ is incurred. Utilizing this notation, parameter design can be described formally as a method for ascertaining those settings of the control factors Θ which minimize average loss $E_N[(f(N;\Theta)-t)^2]$ due to variation from the noise factors N. An approach popularized by Taguchi involves a two-step procedure where attention initially focuses on reducing process variability and then turns to the problem of process targeting. To this end, Taguchi divides the control factors into two groups, $\Theta=(a,d)$, where a and d are called respectively the *adjustment* and *nonadjustment* factors. The adjustment factors derive their name from the fact that they can be used to *adjust* the mean output without affecting the variability. Much progress in the understanding and refinement of two-step procedures can be found in LSK, LW and Box (1988). The procedure can be generically described as follows.

1. Find d^* which minimizes some dispersion measure.
2. With d fixed at d^*, find a^* by identifying the setting of a that adjusts the expected output to the target.

Obviously the existence of an adjustment factor is essential (see LSK and LW). In LW, a theoretical basis for generalizing both the loss function and the two-step procedure was introduced. This paper proposes a corresponding data analysis strategy and in particular develops an efficient two-step procedure for the case when y follows an additive model.

In Section 2, the general measures of dispersion and location proposed in LW are reviewed. These measures are then considered for a general class of loss functions. In Section 3, the traditional two-step procedure for quadratic loss is generalized and then adapted to an additive model. Under this model, it is proved that minimizing the dispersion measure is equivalent to minimizing variance. Methods for estimating the location measure are considered. In Section 4, a strategy for data analysis under an additive model is presented and then applied to an industrial experiment in Section 5. Since the loss functions considered in this paper have the flexibility to better portray true costs, the methodology developed in this paper will be called *cost-driven parameter design.*

2 Cost-Driven Loss Functions

Before the work of LW, two-step procedures for nominal-the-best problems were largely limited to quadratic loss under the familiar relation

$$R_t = \mathrm{E}[(Y-t)^2] = \mathrm{Var}[Y] + (\mathrm{E}[Y]-t)^2. \tag{1}$$

As argued in Section 1, quadratic loss ignores the possible asymmetric nature of loss about the target and often over penalizes large deviations. To extend (1) for a general loss function $L(y,t)$, where the output is y and the target is t, let Y be the random variable associated with y, F be the corresponding distribution induced by the noise N and $R_t = \mathrm{E}_F[L(Y,t)]$ be the risk function. LW defined the *dispersion measure* as the minimum risk incurred when t is free to assume any value in a set T,

$$D = \min_{t \in T} R_t = R_{t^*},$$

and the *location measure* as the value of t that minimizes R_t, denoted by t^*. Noting that D and t^* do not depend on t, LW completed the analogue of (1) by calling the excess risk,

$$O_t = R_t - R_{t^*} = R_t - D,$$

the *off-target measure.* Rearranging this relation gives

$$R_t = D + O_t,$$

a fundamental formula for developing two-step procedures. Note that O_t vanishes when $t^* = t$. For quadratic loss, D becomes the variance of Y, t^* the mean of Y, and O_t the bias squared.

LW derive their generalized measures for two asymmetric loss functions $L_1(y,t) = w_t|y-t|$ and $L_2(y,t) = w_t(y-t)^2$, where

$$w_t = \begin{cases} b_1 & \text{if } y \le t, \\ b_2 & \text{if } y > t, \quad b_1, b_2 > 0. \end{cases}$$

Since the expressions for D and O_t are straight forward once t^* is determined, we will only give the expression for t^* in the following. LW showed that, for L_1,

$$t^* = F^{-1}\left(\frac{b_2}{b_1+b_2}\right), \tag{2}$$

which is the $100b_2/(b_1+b_2)$ percentile of Y, and for L_2,

$$t^* = \frac{b_1 \int_{-\infty}^{t^*} y dF(y) + b_2 \int_{t^*}^{\infty} y dF(y)}{b_1 F(t^*) + b_2(1 - F(t^*))}, \tag{3}$$

which is implicitly defined. Since $\frac{\partial^2 R_t}{\partial t^2} > 0$ in either case, t^* uniquely minimizes R_t.

Here we consider another interesting loss function,

$$L_{1,2}(y,t) = \begin{cases} b_1|y-t| & \text{if } y \le t, \\ b_2(y-t)^2 & \text{if } y > t, \quad b_1, b_2 > 0, \end{cases}$$

which combines L_1 and L_2. This intrinsically asymmetric function could be used to represent a process where a product exhibiting significant deviation below the target can be reworked at a cost which is linearly proportional to the amount of deviation. For example, when a brake rotor is too hard (characterized by small readings on the brinell scale), it is possible to anneal the rotor back to the intended target at a cost which is approximately proportional to $|y-t|$. However, excessively soft rotors must be scrapped. For the sake of simplicity, no upper specification limit has been imposed on y; however, the loss function in (5) is capable of incorporating this additional feature. It is proved in the appendix that

$$t^* = \frac{\int_{t^*}^{\infty} y dF(y)}{1 - F(t^*)} - \frac{b_1}{2b_2} \frac{F(t^*)}{1 - F(t^*)} \tag{4}$$

is the location measure. When $L_{1,2}$ is changed to $b_1(y-t)^2$ for $y \leq t$ and $b_2|y-t|$ for $y > t$, a similar result follows.

It turns out that the loss function can be chosen very generally as

$$L(y,t) = \begin{cases} L_I(t-y) & \text{if } y \leq t, \\ L_{II}(y-t) & \text{if } y > t, \end{cases} \tag{5}$$

where L_I and L_{II} are nondecreasing functions of $|t-y|$ and $L_I(0) = L_{II}(0) = 0$, while still admitting a tractable two-step procedure under an additive model. Note that a one-sided derivative of $L(y,t)$ with respect to t exists (for every fixed y) since L_I and L_{II} are monotonic. In general, the smoothing effect of integration is sufficient to guarantee the differentiability of R_t. These remarks are relevant to the proof of Theorem 1 in the following section where we investigate D under the present loss function and assuming an additive model for y.

While it is inappropriate to assume symmetric loss in general, it should also be recognized that manufacturing processes usually demand that upper and lower specification limits about the target not be exceeded. When such violations occur, the product is often scrapped in which case the loss is a constant (i.e. the cost of production) for y outside these limits. The general loss function in (5) is capable of handling this situation.

3 Two-Step Procedure for the Additive Model

The traditional two-step procedure for quadratic loss would suggest that the general dispersion and location measures be modeled as functions of the control factors. Expected loss could then be minimized by implementing the following two-step procedure.

Procedure 1

1. Find d^* which minimizes $D(d)$.

2. With d fixed at d^*, find a^* by identifying the setting of a that adjusts $t^*(a, d^*)$ to the target t.

When the quality characteristic y follows an additive model, it will be demonstrated that *Procedure 1* can be greatly simplified.

Under a location-scale distribution, the additive model for the output is

$$y(a,d) = \mu(a,d) + \sigma(d)\epsilon, \tag{6}$$

where ϵ has a standardized distribution H independent of a and d. It is thus assumed that the distribution of y depends on the control factors only through μ and σ. For the most general loss function given in (5), let $t = \mu + \sigma z$, where z is the standardized target corresponding to the standardized output ϵ. Then,

$$R_t = \int_{-\infty}^{z} L_I[\sigma(z-\epsilon)]dH(\epsilon) + \int_{z}^{\infty} L_{II}[\sigma(\epsilon - z)]dH(\epsilon) = R_z \tag{7}$$

and it follows that z^*, the standardized location measure, is the solution of

$$\frac{\partial R_z}{\partial z} = \int_{-\infty}^{z} L_I'[\sigma(z-\epsilon)]dH(\epsilon) - \int_{z}^{\infty} L_{II}'[\sigma(\epsilon - z)]dH(\epsilon) = 0. \tag{8}$$

This leads us to the following theorem which reveals a fundamental relationship between the dispersion measure associated with L and the standard deviation of Y. Its proof is given in the Appendix.

Theorem 1 *Let σ_Y represent the standard deviation of Y under the additive model in (6) and D_Y the dispersion measure of Y under the general loss function in (5). Then D_Y is a nondecreasing function of σ_Y.*

From Theorem 1 minimizing D_Y is equivalent to minimizing σ_Y. Therefore, we can replace $D(d)$ in Step 1 of *Procedure 1* by the much simpler function $\sigma(d)$. Next, replace $t^*(a,d)$ in Step 2 by $\mu(a,d) + \sigma(d)z^*(d)$. Intuitively we see that the mean is adjusted toward that side of t with lesser cost (i.e. a cost-adjustment to the target). *Procedure 1* can now be restated as follows.

Procedure 2

1. Find d^* that minimizes $\sigma(d)$.

2. With d fixed at d^*, find a^* by identifying the level of a so that $\mu(a, d^*)$ equals the cost-adjusted target $t - \sigma(d^*)z^*(d^*)$, where z^* is a solution to (8).

For the rest of this section we discuss the estimation of $z^*(d^*)$. Two methods are being considered.

Sample Estimation of z^*: Let $\epsilon_{ij} = (y_{ij} - \mu_i)/\sigma_i$ be the j^{th} standardized observation from the i^{th} setting of the control factors and replace it by the corresponding sample estimate $e_{ij} = (y_{ij} - \widehat{\mu}_i)/\widehat{\sigma}_i$, where $\widehat{\mu}_i = \bar{y}_i$ and $\widehat{\sigma}_i = s_i$. A sample estimate of z^* can thus be obtained from (8) upon replacing σ by $\widehat{\sigma}$ and H by $\widehat{H}$—the empirical distribution function of $\{e_{ij}\}$. Since ϵ_{ij} is independent of (a, d), we can pool the experimental data and estimate z^* from a single representative sample.

Expressions for estimating z^* under L_1, L_2 and $L_{1,2}$ are now presented. Let $e_{(k)}$ denote the k^{th} order statistic from the combined sample $\{e_{ij}\}$ of size n. Then for L_1,

$$\widehat{z^*} = e_{(n^*)} + \left(\frac{b_2}{b_1 + b_2}n - n^*\right)\left(e_{(n^*+1)} - e_{(n^*)}\right), \tag{9}$$

where n^* is the inter part of $b_2 n/(b_1 + b_2)$. This estimator is the $100b_2/(b_1 + b_2)$ percentile of e. For L_2,

$$\widehat{z^*} = \frac{\sum_{k=1}^{n} w_{\widehat{z^*}}(k) e_{(k)}}{\sum_{k=1}^{n} w_{\widehat{z^*}}(k)}, \qquad \text{where} \qquad w_z(k) = \begin{cases} b_1 & \text{if } e_{(k)} < z, \\ b_2 & \text{if } e_{(k)} > z. \end{cases} \tag{10}$$

Finally, for $L_{1,2}$ let $e_{(n_{\widehat{z^*}})}$ be the largest order statistic less than or equal to $\widehat{z^*}$. Then,

$$\widehat{z^*} = \frac{\sum_{k=n_{\widehat{z^*}}+1}^{n} e_{(k)}}{n - n_{\widehat{z^*}}} - \frac{b_1}{2b_2\widehat{\sigma}(d^*)}\left(\frac{n_{\widehat{z^*}}}{n - n_{\widehat{z^*}}}\right). \tag{11}$$

The reader will note that the expression in (11) requires an estimate of $\sigma(d^*)$, while (9) and (10) do not.

Although this method of estimation does not require specific assumptions concerning the distribution of the data, caution must be exercised. When the sample size corresponding to the control factor settings is small and particularly if the cost ratio b_1:b_2 is extreme, the range of the data may not include z^*. Obviously sample estimation is inefficient in such instances and an alternate method of estimation is available as follows.

Theoretical Computation of z^*: When the form of $F(y)$ and hence of $H(\epsilon)$ is known or assumed, z^* can be computed directly as the solution of (8).

For example, consider the following loss function,

$$L_{p,q}(y,t) = \begin{cases} b_1(t-y)^p & \text{if } y \le t, \\ b_2(y-t)^q & \text{if } y > t, \end{cases} \tag{12}$$

for integers $p, q \geq 1$. With respect to $L_{p,q}$, it can be shown that, by using the binomial expansion, z^* is the solution of

$$p\sigma^p \sum_{i=0}^{p-1} \binom{p-1}{i}(-1)^i z^{p-i-1} b_1 \int_{-\infty}^{z} \epsilon^i dH + (-1)^q q\sigma^q \sum_{i=0}^{q-1} \binom{q-1}{i}(-1)^i z^{q-i-1} b_2 \int_{z}^{\infty} \epsilon^i dH = 0. \tag{13}$$

The proof is cumbersome but not difficult. When $p = q$ this relation reduces to

$$\sum_{i=0}^{p-1} \binom{p-1}{i}(-1)^i z^{p-i-1} \left(b_1 \int_{-\infty}^{z} \epsilon^i dH + (-1)^p b_2 \int_{z}^{\infty} \epsilon^i dH \right) = 0,$$

which does not depend on σ. Therefore, z^* can be computed analytically (i.e. without data) when $p = q$ are integers; otherwise, the computation of z^* requires an estimate of $\sigma(d^*)$.

To solve for z^* one must evaluate, $\int_{-\infty}^{z} \epsilon^i dH(\epsilon)$ and $\int_{z}^{\infty} \epsilon^i dH(\epsilon)$, the truncated moments of H. When H is standard normal, these moments have nice expressions. For the sake of brevity, we only give the solution to (13) for the three loss functions L_1, L_2 and $L_{1,2}$. For L_1, (13) reduces to $(b_1 + b_2)\Phi(z) - b_2 = 0$, which leads to

$$z^* = \Phi^{-1}\left(\frac{b_2}{b_1 + b_2}\right), \tag{14}$$

the $100b_2/(b_1 + b_2)$ percentile of the standard normal distribution. For L_2, (13) reduces to

$$z^*\Phi(z^*) + \phi(z^*) + \left(\frac{b_2}{b_1 - b_2}\right) z^* = 0, \tag{15}$$

and for $L_{1,2}$, (13) reduces to

$$\left(\frac{b_1}{2b_2\sigma} - z^*\right)\Phi(z^*) - \phi(z^*) + z^* = 0. \tag{16}$$

The three expressions above clearly demonstrate the dependence of z^* on the cost ratio $b_2{:}b_1$. In Table 4 of the appendix, we give selected values of z^*, under the standard normal distribution and as a function of b_2/b_1, for L_1, L_2 and $L_{1,2}$ above. In practice, L_1 and L_2 are preferred since then z^* is independent of σ. Note, however, that this approach is valid only if the skewness exhibited by a loss model can be captured by the coefficients b_1 and b_2 alone.

To summarize, the implication of Theorem 1 has been to greatly simplify the implementation of a two-step procedure under cost-driven loss functions. In fact, the only complexity is that of estimating or computing the standardized location measure z^*. Note, however, that this procedure is completely dependent on the existence of an adjustment factor.

4 Strategy for Data Analysis: Additive Model

The following five-step procedure outlines an analysis strategy for the implementation of *Procedure 2* under an additive model.

Procedure 3

1. Identify the set of control factors d that influence σ and the levels d^* that minimize $\sigma(d)$.
2. Estimate $\sigma(d)$ at d^* and denote it by $\widehat{\sigma}(d^*)$.
3. Identify an adjustment factor a.
4. Estimate $z^*(d^*)$ in (8) and denote it by $\widehat{z^*}(d^*)$. Then adjust a to the level a^* that satisfies

$$\mu(a^*, d^*) = t - \widehat{\sigma}(d^*)\widehat{z^*}(d^*).$$

5. After performing a follow-up experiment, update the adjustment of factor a in step 4.

Under the traditional quadratic loss function in (1), $z^* = 0$ and an estimate of $\sigma(d^*)$ is no longer required in step 4. Thus step 2 can be eliminated and we have the familiar procedure for conducting a parameter design experiment when y is additive.

The usual approach in step 1 is to model $\log(s^2)$ as an additive function of the control factors. Significant factors are thus identified and the derived model can be used to implement the estimation of $\sigma(d^*)$ in step 2. In addition, if replicates of $s^2(d^*)$ are available when the experiment is restricted to the setting d^*, then a second estimate of $\sigma(d^*)$ becomes available. An alternative approach is to model the response y as a function of the control and noise factors and then compute $\widehat{\sigma}^2(d)$ as the variance of $\widehat{y}(d)$ with respect to the noise variation (see Welch, Yu, Kang, and Sacks (1990) and Shoemaker, Tsui and Wu (1991)).

In many cases the adjustment factor required in step 3 is known àpriori from engineering knowledge and design specifications (see LSK for further discussion). When this is not the case, one can search for an adjustment factor empirically by modeling the mean as an additive function of the control factors. Any significant factor for the mean but not for the variance is a potential candidate (see Nair and Pregibon (1986) and Box (1988)).

In step 4, z^* is estimated by one of the two methods discussed in Section 3. If z^* depends on d through $\sigma(d)$, then use the estimate $\widehat{\sigma}(d^*)$ from step 2 in its place. It is emphasized at this point that step 4 merely provides an intermediate estimate of $t - \sigma(d^*)z^*(d^*)$ necessary to target the follow-up experiment. By focusing more resources on a single setting, the follow-up experiment provides a superior estimate of this quantity.

In step 5, a follow-up experiment is performed not only to confirm the conclusions drawn from the experiment, but to estimate $t - \sigma(d^*)z^*(d^*)$ with more precision as well.

5 Cost-Driven Analysis of an Epitaxial-Layer Growth Experiment in IC Fabrication

The following example, from integrated-circuit fabrication, was first analyzed by Kacker and Shoemaker (1986) (henceforth denoted by KS). The manufacturing process begins with the growth of an epitaxial layer of silicon on top of silicon wafers. Reliable performance of future components demands that the layer be uniformly close to the target thickness of 14.5 micrometers. To this end, process specifications called for a layer between 14 and 15 micrometers (see KS for a detailed description of the process). This example will now be analyzed following the strategy outlined in Section 4.

Table 1. Control and Noise Factors and Their Settings

Factor	Experimental settings			
Control Factors				
A Rotation method	Continuous		Oscillating	
B Wafer code	668G4		678D4	
C Deposition temperature	1,210		1,220	
D Deposition time	High		Low	
E Arsenic flow rate	55%		59%	
F HCI etch temperature	1,180		1,215	
G HCI flow rate	10%		14%	
H Nozzle position	2		6	
Noise Factors				
L Location	Top		Bottom	
F_i Facet	1	2	4	6

Step 1: KS considered eight control factors for their potential to reduce variability. The factors, which are listed in Table 1, were varied according to a 2^{8-4}_{IV} fractional factorial design. To minimize the epitaxial-thickness variance, they modeled $\log(s^2)$ as an additive function of these factors. From their analysis it was determined that A and H were significant. The new settings given in Table 2 led to a 60% reduction in the epitaxial-thickness standard deviation.

Table 2. Setting Changes Suggested by the Analysis in Step 1

Factor	Initial setting	New setting
A Rotation method	Oscillating	Continuous
H Nozzle position	4	6

Step 2: Following the model developed in KS, $\log(s^2)$ was regressed on the control factors A and H resulting in the estimated model,

$$\widehat{\log(\sigma^2)} = -1.82 + 0.619X_A - 0.982X_H. \tag{17}$$

where $X_A = \pm 1$ denotes the two settings of A and similarly for X_H. Inserting the new settings for A and H from Table 2 produced the estimate $\widehat{\sigma}(d^*) = 0.181$. Under the reduced model, four replicates of $s^2(d^*)$ were available and their average produced the estimate $\widehat{\sigma}(d^*) = 0.257$, based on a combined sample size of 32.

From these results, it would seem that the $\log(s^2)$ model is significantly underestimating $\sigma^2(d^*)$. This can be explained, at least partially, by the fact that $\exp\{E[\log(s^2)]\}$ is a biased estimate of σ^2. By a Taylor series argument, it can be shown that $\exp\{E[\log(s^2)]\} \approx \sigma^2 - \frac{1}{2}\sigma^2 \text{Var}[\log(s^2)] < \sigma^2$. In any event, since the magnitude of the cost-adjustment to t is directly proportional to σ, conservative estimation of σ would be prudent.

The results above emphasize how difficult it is to estimate dispersion effects from fractional factorial experiments. Typically such experiments do not take enough data to generate reliable estimates. However, since the follow-up experiment can provide a more accurate estimate of $\sigma(d^*)$, this is not a problem of great concern.

Step 3: For this process, deposition time (D) is a scaling factor that was traditionally used by engineers to adjust epitaxial-thickness. We therefore choose D as the adjustment factor. Note that the original data in KS and the response model analysis in Shoemaker, Tsui and Wu (1991) provided empirical evidence supporting this choice.

Step 4: For this example, a cost ratio of $b_1 : b_2 = 1:6$ was chosen because it produced a significant cost-adjustment to the target without pushing it too close to the lower specification limit. Although there was no inherent reason to apply asymmetric loss to epitaxial thickness, some of the physical attributes of the model such as additivity, normality, and nominal-the-best targeting made it ideal for illustrative purposes.

Sample estimates of z^* are obtained from the experimental data and compared with their counterparts under the standard normal distribution.

For L_1, $\widetilde{z^*} = 1.065$ as compared with $z^* = 1.068$.

For L_2, $\widetilde{z^*} = 0.688$ as compared with $z^* = 0.707$.

For $L_{1,2}$, $\widetilde{z^*} = 0.319$ as compared with $z^* = 0.269$, given $\widehat{\sigma}(d^*) = 0.181$.

Under L_1 and L_2, we see that $\widetilde{z^*}$ is slightly less than z^*. One possible explanation is that a small sample provides very conservative information with respect to the tails of its distribution. As a result, $\widetilde{z^*}$ would tend to underestimate z^*.

Given the intended target, $t = 14.5$, and estimates of z^* and σ, deposition time for the follow-up experiment would then be adjusted until the mean thickness of the epitaxial layer is approximately equal to $14.5 - \widehat{\sigma}\widehat{z^*}$.

Step 5: To confirm conclusions drawn from the experiment, KS performed a follow-up experiment at the new settings for A and H. They conducted three independent test runs, each consisting of 70 wafers, and averaged the three sample variances to get the estimate $\widehat{\sigma}(d^*) = 0.239$. Since the original data are not available, estimation of $t - \sigma z^*$ under the assumption of normality is the only recourse. Since L_1 and L_2 are independent of σ, z^* remains unchanged in either case; however, under $L_{1,2}$ the new estimate of σ yields $z^* = 0.404$.

Given that no serious anomalies were observed in the follow-up experiment, an estimate of the targeting factor for mean epitaxial-thickness would be computed. It would then be used, in conjunction with the control factor settings of Table 2, to optimize the process. In

Table 3, the cost-adjusted targets, $t - \widehat{\sigma}\widehat{z}^*$, under each loss function and for each method of estimation are collected together. First, note that the method of estimating z^* makes little difference in this example. Second, estimating $\sigma(d^*)$ from the $\log(s^2)$ model generates conservative estimates of the targeting factors. Finally, the targeting factors from Table 3 are well above the lower specification limit of 14.0. In fact, within the interval $(14.0, 14.5)$, the cost-adjusted targets range from 50% to 20% below $t = 14.5$.

Table 3. Target Adjustments for $b_1 : b_2 = 1:6$

Loss Function	L_1	L_2	$L_{1,2}$
Factorial Experiment	(Pooled Std. Data)		
$\hat{\sigma}(d^*) = 0.181$	14.31	14.37	14.44
$\hat{\sigma}(d^*) = 0.257$	14.23	14.32	14.38
Factorial Experiment	(Std. Normal Dist.)		
$\hat{\sigma}(d^*) = 0.181$	14.31	14.37	14.45
$\hat{\sigma}(d^*) = 0.257$	14.23	14.32	14.39
Follow-up Experiment	(Std. Normal Dist.)		
$\hat{\sigma}(d^*) = 0.239$	14.24	14.33	14.40

Based on results from the follow-up experiment for the asymmetric linear loss function L_1, we see that $t - \sigma z^* \pm \sigma = 14.24 \pm .239$ is contained in the lower half of the tolerance interval $(14.0, 15.0)$. Assuming that any product for which $y < 14.0$ can be reworked whereas $y > 15.0$ leads to scrap, the benefits of employing an asymmetric loss model can be quantified. Approximately 68% of the product produced will have y is in the interval $(14.0, 14.5)$, leaving 16% of the production with $y \leq 14.0$ and 16% with $y \geq t = 14.5$. Operating the process under the settings in Tables 2 and 3 virtually eliminates product waste due to $y > 15.0$ and significantly reduces the probability of y falling in the interval $(14.5, 15.0)$. However, approximately 16% of the product must be reworked due to $y \leq 14.0$.

In contrast, a symmetric linear loss model leaves 2.5% of the product to be reworked and 2.5% for scrap. To eliminate scrap under this model would require a 30% reduction in $\sigma(d^*) \approx 0.239$; in other words, new technology would be required to meet the same conditions achieved by the asymmetric linear loss model. If the cost of reworking 16% of the product is significantly less than the estimated cost of developing and implementing new technology, then a cost-efficient route for quality improvement has been identified.

6 Concluding Remarks

The data analysis strategy proposed in this paper has two major advantages. First, it extends the scope of the nominal-the-best parameter design to include loss functions of a much more general nature. Second, it achieves this generality while retaining the inituitive appeal and much of the mathematical simplicity of the quadratic loss function. However, since this methodology pertains to the additive model, it can not be applied directly to others such as the multiplicative model,

$$y(a,d) = \mu(a,d)\eta(d),$$

where $\mu, \eta > 0$ and η is a multiplicative error with $E[\eta] = 1$, $Var[\eta] = \sigma^2$. In this particular instance a log-transformation of y would achieve approximate additivity and the analysis strategy proposed in the paper would be suitable for $\log(y)$. In general, a variance stabilizing transformation which eliminates dependence of the variance on the mean should be sufficient (see Nair and Pregibon(1986)); however, this problem requires further research. Although it can be difficult to identify the true loss function in a given industrial setting, approximation to the true loss function by L_1, L_2 or $L_{1,2}$ with an appropriate choice of b_1 and b_2 should be quite effective unless the percent of items with large deviations from the nominal value is substantial.

Appendix

A Proof of (4).

Noting that

$$R_t = \mathrm{E}_F[L_{1,2}(Y,t)] = b_1 \int_{-\infty}^{t} (t-y)dF(y) + b_2 \int_{t}^{\infty} (y-t)^2 dF(y),$$

it follows that $\frac{\partial R_t}{\partial t}|_{t=t^*} = 0$ gives

$$b_1 F(t^*) - 2b_2 \int_{t^*}^{\infty} y dF(y) + 2b_2 t^* (\ 1 - F(t^*)\ ) = 0,$$

or equivalently,

$$t^* = \frac{\int_{t^*}^{\infty} y dF(y)}{1-F(t^*)} - \frac{b_1}{2b_2} \frac{F(t^*)}{1-F(t^*)}.$$

B Proof of Theorem 1.

From (5),

$$D = \int_{-\infty}^{z^*} L_I[\sigma(z^*-\epsilon)]dH(\epsilon) \ + \ \int_{z^*}^{\infty} L_{II}[\sigma(\epsilon - z^*)]dH(\epsilon)$$

where z^* is the solution of (8). To prove that D is a nondecreasing function of σ, it is sufficient to show that $\frac{\partial D}{\partial \sigma} \geq 0$. To this end,

$$\begin{aligned} \frac{\partial D}{\partial \sigma} &= \int_{-\infty}^{z^*} L_I'[\sigma(z^*-\epsilon)]\ (z^*-\epsilon)\ dH(\epsilon) + \int_{z^*}^{\infty} L_{II}'[\sigma(\epsilon - z^*)]\ (\epsilon - z^*)\ dH(\epsilon) \\ &+ \ \sigma \frac{\partial z^*}{\partial \sigma} \left(\int_{-\infty}^{z^*} L_I'[\sigma(z^*-\epsilon)]dH(\epsilon) - \int_{z^*}^{\infty} L_{II}'[\sigma(\epsilon - z^*)]dH(\epsilon) \right), \end{aligned}$$

From (8), the term in parentheses above is zero. Since L_I and L_{II} are nondecreasing, i.e. L_I' and L_{II}' are nonnegative, $\frac{\partial D}{\partial \sigma} \geq 0$ as required.

C Table 4.

Tabulated values of z^* under the Standard Normal Distribution†

b_2/b_1	z_1^*	z_2^*	$z_{1,2}^*$	$z_{1,2}^*(b_1/b_2)$	b_2/b_1	z_1^*	z_2^*	$z_{1,2}^*$	$z_{1,2}^*(b_1/b_2)$
1.1	0.006	0.038	0.275	0.182	11.0	1.383	0.937	1.344	-0.905
1.2	0.114	0.073	0.317	0.139	12.0	1.426	0.970	1.382	-0.944
1.3	0.164	0.105	0.356	0.100	13.0	1.465	1.000	1.415	-0.979
1.4	0.210	0.134	0.392	0.064	14.0	1.501	1.028	1.447	-1.011
1.5	0.253	0.162	0.426	0.030	15.0	1.534	1.053	1.475	-1.041
1.6	0.293	0.187	0.457	-0.001	16.0	1.565	1.077	1.502	-1.068
1.7	0.331	0.211	0.487	-0.031	17.0	1.593	1.099	1.527	-1.094
1.8	0.366	0.234	0.514	-0.059	18.0	1.620	1.120	1.550	-1.119
1.9	0.399	0.256	0.540	-0.085	19.0	1.645	1.140	1.573	-1.142
2.0	0.431	0.276	0.565	-0.110	20.0	1.668	1.159	1.593	-1.163
2.1	0.460	0.295	0.588	-0.134	25.0	1.769	1.240	1.683	-1.256
2.2	0.489	0.314	0.611	-0.156	30.0	1.849	1.305	1.754	-1.331
2.3	0.516	0.331	0.632	-0.178	35.0	1.915	1.360	1.814	-1.393
2.4	0.541	0.348	0.652	-0.198	40.0	1.971	1.407	1.865	-1.447
2.5	0.566	0.364	0.672	-0.218	45.0	2.019	1.449	1.909	-1.493
3.0	0.674	0.436	0.758	-0.305	50.0	2.062	1.485	1.948	-1.534
3.5	0.765	0.497	0.831	-0.379	55.0	2.100	1.518	1.983	-1.571
4.0	0.842	0.549	0.893	-0.442	60.0	2.135	1.548	2.015	-1.605
4.5	0.908	0.595	0.947	-0.498	65.0	2.166	1.576	2.044	-1.635
5.0	0.967	0.636	0.996	-0.547	70.0	2.195	1.601	2.071	-1.664
6.0	1.068	0.707	1.078	-0.631	75.0	2.222	1.624	2.096	-1.690
7.0	1.150	0.766	1.147	-0.702	80.0	2.246	1.646	2.119	-1.714
8.0	1.221	0.817	1.206	-0.763	85.0	2.269	1.666	2.140	-1.736
9.0	1.282	0.862	1.258	-0.816	90.0	2.291	1.686	2.160	-1.758
10.0	1.335	0.901	1.304	-0.863	100.0	2.330	1.721	2.197	-1.796

†The notation used in this table identifies z_1^*, z_2^* and $z_{1,2}^*$ as the location measures for L_1, L_2 and $L_{1,2}$ respectively. The following relations:

$$\begin{aligned} z_i^*(b_1/b_2) &= -z_i^*(b_2/b_1), \\ z_{2,1}^*(b_2/b_1) &= -z_{1,2}^*(b_1/b_2), \\ z_{2,1}^*(b_1/b_2) &= -z_{1,2}^*(b_2/b_1), \end{aligned}$$

extend the scope of Table 1. To obtain $z_{i,j}^*$ when σ is not equal to one, simply replace $\frac{b_2}{b_1}$ by $\frac{b_2}{b_1\sigma}$ and take this new cost ratio to Table 4.

Acknowledgements

This research was supported by the Manufacturing Research Corporation of Ontario and the Natural Sciences and Engineering Research Council of Canada. The authors thank M. Hamada and V.N. Nair for their comments.

References

[1] Box , G. E. P. (1988), "Signal-to-Noise Ratios, Performance Criteria, and Transformations," *Technometrics*, 30, 1–40.

[2] Kacker, R. N., and Shoemaker, A. C. (1986), "Robust Design: A Cost-Effective Method for Improving Manufacturing Processes," *AT&T Technical Journal*, 65, 39–50.

[3] León, R. V., Shoemaker, A. C., and Kacker, R.N. (1987), "Performance Measures Independent of Adjustment: An Explanation and Extension of Taguchi's Signal-to-Noise Ratios," *Technometrics*, 29, 253–285.

[4] León, R. V., and Wu, C. F. J. (1992), "A Theory of Performance Measures in Parameter Design," *Statistica Sinica*, 2, 335–358.

[5] Nair, V. N. (editor) (1992), "Taguchi's Parameter Design: A Panel Discussion," *Technometrics*, 34, 127–161.

[6] Nair, V. N., and Pregibon, D. (1986), "A Data Analysis Strategy for Quality Engineering Experiments," *AT&T Technical Journal*, 65, 73–84.

[7] Shoemaker, A. C., Tsui, K. L., and Wu, C. F. J. (1991), "Economical Experimentation Methods for Robust Design," *Technometrics*, 33, 415–427.

[8] Welch, W. J., Yu, T. K., Kang, S. M., and Sacks, J. (1990), "Computer Experiments for Quality Control by Parameter," *Journal of Quality Technology*, 22, 15–22.