TECHNOMETRICS, AUGUST 1987, VOL. 29, NO. 3

Editor's Note: This article and the first three discussions were presented at the *Technometrics* session of the 30th Annual Fall Technical Conference, cosponsored by the Chemical and Process Industries and the Statistics Divisions of the American Society for Quality Control and the Section on Physical and Engineering Sciences of the American Statistical Association. The conference was held in Charlotte, North Carolina, October 23–24, 1986.

Performance Measures Independent of Adjustment

An Explanation and Extension of Taguchi's Signal-to-Noise Ratios

Ramón V. León, Anne C. Shoemaker, and Raghu N. Kacker

Quality Assurance Center
AT&T Bell Laboratories
Holmdel, NJ 07733

Parameter design is a method, popularized by Japanese quality expert G. Taguchi, for designing products and manufacturing processes that are robust to uncontrollable variations. In parameter design, Taguchi's stated objective is to find the settings of product or process design parameters that minimize average quadratic loss—that is, the average squared deviation of the response from its target value. Yet, in practice, to choose the settings of design parameters he maximizes a set of measures called signal-to-noise ratios. In general, he gives no connection between these two optimization problems. In this article, we show that for certain underlying models for the product or process response maximization of the signal-to-noise ratio leads to minimization of average quadratic loss. The signal-to-noise ratios take advantage of the existence of special design parameters called adjustment parameters. When these parameters exist, use of the signal-to-noise ratio allows the parameter design optimization procedure to be conveniently decomposed into two smaller optimization steps, the first being maximization of the signal-to-noise ratio. We show, however, that under different models (or loss functions) other performance measures give convenient two-step procedures, but the signal-to-noise ratios do not.

KEY WORDS: Parameter design; Quality; Robust product design.

1. THE ROLE OF SIGNAL-TO-NOISE RATIOS IN PARAMETER DESIGN

1.1 Parameter Design: Reducing Sensitivity to Variation

When the Ina Tile Company of Japan found that the uneven temperature profile of its kilns was causing unacceptable variation in tile size, it could have attempted to solve the problem with expensive modification of the kilns. Instead, it chose to make an inexpensive change in the settings of the tile design parameters to reduce sensitivity to temperature variation. Using a statistically planned experiment, the company found that increasing the lime content of the clay from 1% to 5% reduced the tile size variation by a factor of 10 (see Taguchi and Wu 1980).

This simple example illustrates the method of parameter design for quality engineering. Parameter design is the operation of choosing settings for the design parameters of a product or manufacturing process to reduce sensitivity to noise. Noise is hard-to-control variability affecting performance. For example, all of the following are considered noise: deviations of raw materials from specifications, changes in the manufacturing or field operation environment such as temperature or humidity, drift of parameter settings over time, and deviation of design parameters from their nominal values because of manufacturing variability.

In parameter design, noise is assumed to be uncontrollable. After parameter design, if the loss caused by noise is still excessive, the engineer may proceed to

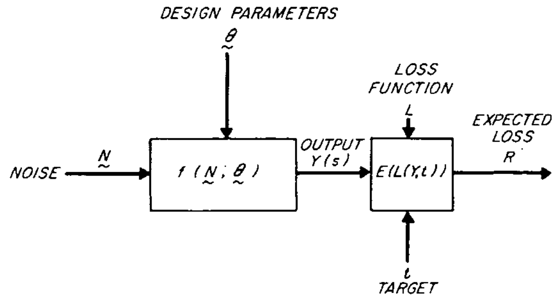

*Figure 1. A Block Diagram Representation of a Simple Parameter-Design Problem. The output Y is determined by the noise **N** through the transfer function f. The transfer function depends on the design parameters θ. Loss is incurred if the output is not equal to the target t.*

control noise with relatively expensive countermeasures, such as the use of higher-grade raw materials or higher-precision manufacturing equipment. (For a more detailed discussion of parameter design and the subsequent design stage, tolerance design, see Taguchi and Wu 1980.)

1.2 A Formalization of a Basic Parameter-Design Problem

Figure 1 shows a block diagram representation of a simple parameter design problem like the tile example. For a given setting of design parameters, $\boldsymbol{\theta}$, noise $\mathbf{N}$ produces a characteristic output Y. The output is determined by the transfer function, $f(\mathbf{N}; \boldsymbol{\theta})$. The noise is assumed to be random; hence the output will be random. A loss is incurred if the output differs from a target t, which represents the ideal output.

The goal of parameter design is to choose the setting of the design parameter $\boldsymbol{\theta}$ that minimizes average loss caused by deviations of the output from target. This average loss is given by

$$R(\boldsymbol{\theta}) = E_{\mathbf{N}} L(Y, t),$$

where L is a loss function.

Taguchi (see Taguchi 1977; Taguchi and Phadke 1984) called this problem the *static* parameter design problem, because the target is fixed. As we show later, other parameter design problems are generalizations of the static problem. He recommended that loss be measured using the quadratic loss function

$$L(y, t) = K(y - t)^2, \tag{1.1}$$

where K is determined using a simple economic argument (see Taguchi and Wu 1980, p. 8). Yet, as we will see in Section 1.3, in practice he optimized different measures called signal-to-noise (SN) ratios.

1.3 Taguchi's Approach to Parameter Design

To solve the problem just described, Taguchi (1977) generally divided the design parameters into two groups, $\boldsymbol{\theta} = (\mathbf{d}, \mathbf{a})$, and used the following two-step procedure (see Phadke 1982):

Procedure 1.
Step 1. Find the setting $\mathbf{d} = \mathbf{d}^*$ that maximizes the SN ratio.
Step 2. Adjust $\mathbf{a}$ to $\mathbf{a}^*$ while $\mathbf{d}$ is fixed at $\mathbf{d}^*$.

The division of design parameters into two groups is motivated by the idea that the parameters $\mathbf{a}$ are fine-tuning adjustments that can be optimized after the main design parameters $\mathbf{d}$ are fixed. Taguchi (see Phadke 1982) claimed that this two-step approach has the advantage that, once a design $(\mathbf{d}^*, \mathbf{a}^*)$ is established, certain changes to product or process requirements can be accommodated by changing only the setting of the adjustment parameter, $\mathbf{a}$. The initial setting of $\mathbf{d} = \mathbf{d}^*$ remains optimal. For instance, in the Ina Tile example mentioned earlier, the clay formulation could be chosen to minimize an SN ratio measurement of tile size variation, and then the tile mold size could be used to adjust the average tile size to target.

Recognizing that the appropriate performance measure depends on the characteristics of the problem, Taguchi (see Taguchi and Phadke 1984) classified parameter design problems into categories and defined a different SN ratio for each category. For example, one of these categories consists of all static parameter design problems.

In general, however, Taguchi (1976, 1977) gave no justification for the use of the SN ratios and no explanation of why the two-step procedure that he recommended will minimize average loss. In this article, we explore this point, investigating the connection between a two-step procedure similar to Procedure 1 and parameter design problems similar to the one stated in Section 1.2.

1.4 Overview

In Section 2, we examine performance measures in the most common parameter design situation, the static problem of Section 1.2. In Section 2.1, we show that for this problem Taguchi's SN ratio and corresponding two-step procedure is valid assuming quadratic loss and a particular multiplicative transfer-function model. We also discuss the danger of using this SN ratio and corresponding two-step procedure when a different type of transfer-function model may hold. Taguchi's recommendation to use a single SN ratio for all static parameter design problems, however, may be motivated by a belief that the transfer-function model for most engineering systems is multiplicative in form. In Section 2.2, we see how a multiplicative model arises from a general situation in which noise is manufacturing variation, deterioration, or drift in design-parameter values.

In Section 3, we discuss block-diagram representations of more general parameter design problems. These generalizations include "dynamic" problems, in which there is a "signal" input variable that can be used to dynamically control the output. In that section we also generalize the recipe followed in Section 2 to derive a two-step procedure equivalent to Procedure 1. Since the appropriate performance measure to use in Step 1 of this procedure is not always equivalent to an SN ratio, we suggest the term *performance measure independent of adjustment* (PerMIA).

Sections 4 and 5 are devoted to generic examples of dynamic parameter design problems, for which we follow the recipe to obtain performance measures independent of adjustment and corresponding two-step procedures. These examples show that Taguchi's various SN ratios for dynamic problems arise from certain types of transfer-function models under quadratic loss. If different types of transfer functions hold, however, a convenient two-step procedure may still be obtained if a performance measure different from the SN ratio is used. In fact, use of the SN ratio in these situations is incorrect and may not lead to good design-parameter settings. Section 4 focuses on dynamic problems in which input and output variables are continuous. Section 5 focuses on a binary-input–binary-output dynamic problem.

Finally, in Section 6, we discuss how adjustment parameters may arise in practice and the benefits of the two-step procedure made possible by their existence.

2. PERFORMANCE MEASURES FOR A STATIC PARAMETER DESIGN PROBLEM

2.1 Explanation and Extension of Taguchi's SN Ratio

Taguchi's SN ratio for the static parameter design problem of Figure 1 is

$$\text{SN} = 10 \log_{10}(E^2 Y/\text{var } Y), \tag{2.1}$$

where EY and var Y are, respectively, the mean and variance of Y (see Taguchi and Phadke 1984).

We show that if the transfer function in Figure 1 is given by

$$Y = \mu(\mathbf{d}, \mathbf{a})\varepsilon(\mathbf{N}, \mathbf{d}), \tag{2.2}$$

where $EY = \mu(\mathbf{d}, \mathbf{a})$ is a strictly monotone function of each component of $\mathbf{a}$ for each $\mathbf{d}$, then use of SN ratio (2.1) in a two-step procedure leads to minimization of quadratic loss. [Note that $EY = \mu(\mathbf{d}, \mathbf{a})$ implies that $E\varepsilon(\mathbf{N}, \mathbf{d}) = 1$.] Model (2.2) could hold, for example, if the noise affects the output, Y, uniformly over increments of time of distance. For example, if a film is being deposited on a quartz plate, the thickness of the film will tend to vary a certain amount for each micron of film surface deposited. That is, a 10-micron-thick film will have 10 times the standard deviation of a 1-micron-thick film. For another example, see Section 2.2.

Note that model (2.2) essentially says that var Y/E^2Y does not depend on $\mathbf{a}$ (or approximately, for log Y, that $\mathbf{a}$ affects location but not dispersion).

The argument linking the SN ratio (2.1) to quadratic loss is straightforward. We first note that we can find $(\mathbf{d}^*, \mathbf{a}^*)$ such that

$$R(\mathbf{d}^*, \mathbf{a}^*) = \min_{\mathbf{d}\,\mathbf{a}} R(\mathbf{d}, \mathbf{a})$$

by following a general two-step optimization procedure (Procedure 2). Before stating Procedure 2, we will define $P(\mathbf{d})$ by the equation $P(\mathbf{d}) = \min_{\mathbf{a}} R(\mathbf{d}, \mathbf{a})$.

Procedure 2.
Step 1. Find $\mathbf{d}^*$ that minimizes $P(\mathbf{d}) = \min_{\mathbf{a}} R(\mathbf{d}, \mathbf{a})$.
Step 2. Find $\mathbf{a}^*$ that minimizes $R(\mathbf{d}^*, \mathbf{a})$.

To see that this procedure always works [provided that $(\mathbf{d}^*, \mathbf{a}^*)$ exists], let $(\mathbf{d}, \mathbf{a})$ be any arbitrary parameter values and $(\mathbf{d}^*, \mathbf{a}^*)$ be the result of Procedure 2. Then note that

$$\begin{aligned} R(\mathbf{d}, \mathbf{a}) &\geq \min_{\mathbf{a}} R(\mathbf{d}, \mathbf{a}) \\ &\geq \min_{\mathbf{a}} R(\mathbf{d}^*, \mathbf{a}) \text{ (by stage 1)} \\ &= R(\mathbf{d}^*, \mathbf{a}^*) \text{ (by stage 2).} \end{aligned}$$

Even though this procedure is always possible, it is not often useful. If there is a shortcut for calculating $P(\mathbf{d}) = \min_{\mathbf{a}} R(\mathbf{d}, \mathbf{a})$, however, this two-step procedure may be the preferred minimization method.

If we are using quadratic loss and model (2.2) holds, then $P(\mathbf{d})$ is a decreasing function of Taguchi's SN ratio (2.1). This means that in this case Taguchi's two-step procedure, described in Section 1.3, is equivalent to Procedure 2.

To see this, note that under model (2.2) and quadratic loss,

$$R(\mathbf{d}, \mathbf{a}) = \mu^2(\mathbf{d}, \mathbf{a})\sigma^2(\mathbf{d}) + (\mu(\mathbf{d}, \mathbf{a}) - t)^2, \tag{2.3}$$

where $\sigma^2(\mathbf{d}) = \text{var } \varepsilon(\mathbf{N}, \mathbf{d})$. [For convenience, choose $K = 1$ in loss function (1.1).] Setting

$$\frac{\partial R(\mathbf{d}, \mathbf{a})}{\partial \mathbf{a}} = \frac{2\partial\mu(\mathbf{d}, \mathbf{a})}{\partial \mathbf{a}} \{\mu(\mathbf{d}, \mathbf{a})[1 + \sigma^2(\mathbf{d})] - t\}$$

equal to 0, we get

$$\mu(\mathbf{d}, \mathbf{a}^*(\mathbf{d})) = t/(1 + \sigma^2(\mathbf{d})),$$

where $\mathbf{a}^*(\mathbf{d})$ is defined by $R(\mathbf{d}, \mathbf{a}^*(\mathbf{d})) = \min_{\mathbf{a}} R(\mathbf{d}, \mathbf{a})$. Then, substituting into formula (2.3), we get

$$P(\mathbf{d}) = t^2\sigma^2(\mathbf{d})/[1 + \sigma^2(\mathbf{d})].$$

But

$$\begin{aligned} \text{SN} &= 10 \log_{10} (E^2 Y/\text{var } Y) \\ &= 10 \log_{10} [\mu^2(\mathbf{d}, a)/\mu^2(\mathbf{d}, a)\sigma^2(\mathbf{d})] \\ &= -10 \log_{10} \sigma^2(\mathbf{d}), \end{aligned}$$

and $t^2\sigma^2(\mathbf{d})/[1 + \sigma^2(\mathbf{d})]$ is an increasing function of $\sigma^2(\mathbf{d})$.

Hence the two-step procedure (Procedure 2) for solving the parameter design problem is equivalent to the following:

Procedure 3.

Step 1. Find $\mathbf{d}^*$ that maximizes the SN ratio $\text{SN} = 10 \log_{10} (E^2 Y/\text{var } Y)$.

Step 2. Find $\mathbf{a}^*$ such that $\mu(\mathbf{d}^*, \mathbf{a}^*) = t/[1 + \sigma^2(\mathbf{d}^*)]$.

With transfer function (2.2), the use of the SN ratio (2.1) leads to minimization of quadratic loss. (Note that even if the minimization of the expected squared error loss is constrained by the requirement that the mean be equal to t, Step 1 remains the same. What changes is Step 2, which now becomes "adjust to t.")

Although use of the SN ratio (2.1) under transfer function (2.2) is justified by the preceding argument, there may be other measures that are better in practice. In particular, if information about the loss function is not precise, there may be no reason to prefer quadratic loss over, say, quadratic loss on the log scale:

$$L(y, t) = (\log y - \log t)^2.$$

Under this loss function, transfer function (2.2) leads to use of var(log Y) in Step 1 of Procedure 2. To see this, note that

$$\begin{aligned} R(\mathbf{d}, \mathbf{a}) &= E(\log Y - \log t)^2 \\ &= \text{var}(\log Y) + [E(\log Y) - \log t]^2. \end{aligned}$$

Since, from model (2.2),

$$\text{var}(\log Y) = \text{var}(\log \varepsilon(\mathbf{N}, \mathbf{d}))$$

is a function of $\mathbf{d}$ alone, it follows that $\min_{\mathbf{a}} R(\mathbf{d}, \mathbf{a}) = \text{var}(\log Y)$.

The quantity log var(log Y) was suggested by Box (1986) as an alternative to the SN ratio of Taguchi. When there is no practical reason to prefer either of the preceding loss functions, Box's suggestion should probably be followed, because the estimate of log var(log Y) usually has better statistical properties than those of the SN ratio (see Box 1986). (Note, however, that problems may arise with this loss function if t is close to 0.)

We also note that if the transfer function in (2.2) is replaced by

$$Y = \mu(\mathbf{d}, \mathbf{a}) + \varepsilon(\mathbf{N}, \mathbf{d}) \tag{2.4}$$

with $E\varepsilon(\mathbf{N}, \mathbf{d}) = 0$ and $L(y, t) = (y - t)^2$, then var Y should be used instead of the SN ratio (2.1) in the two-step procedure. Note that the SN ratio is not independent of the adjustment parameter $\mathbf{a}$ under model (2.4). This example illustrates that blanket use of the SN ratio in static problems, as Taguchi and Phadke (1984) seem to have advocated, could lead to far from optimal design-parameter settings.

2.2 Example: Static Problem With Variation in Design-Parameter Values

Suppose that the output, Y, of the system is given by

$$Y = A^p g(\mathbf{D}), \qquad -\infty < p < \infty, \tag{2.5}$$

where $(A, \mathbf{D})$ are random design-parameter settings that, because of manufacturing variability or deterioration, are random variables with means equal to the nominal parameter settings $(a, \mathbf{d})$ set by the designer. [Taguchi (1976, 1977) referred to this kind of noise as "inner" noise.] Model (2.5) with $p = 1$ (suggested in this case by Paul Sherry, a quality engineer at AT&T Bell Laboratories) frequently occurs in applications. In these applications parameter a is called a *scale* parameter because it can be used to change the "scale" of the product or process. Examples of scale parameters [and model (2.5)] are mold size in tile fabrication, deposition time in surface film deposition, mask dimension in integrated circuit fabrication, and exposure time in window photolithography.

Engineers frequently observe that variability of the actual scale parameter value (as measured by the standard deviation) is proportional to nominal settings. Hence we assume that

$$\text{var } A = k^2 a^2, \qquad k \geq 0.$$

To see that model (2.5) is a special case of model (2.2), represent A by $A = aZ$, where Z is a random variable with $EZ = 1$ and var $Z = k^2$. Then

$$Y = a^p Z^p g(\mathbf{D}) = \mu(\mathbf{d}, a)\varepsilon(\mathbf{N}, \mathbf{d}),$$

where $\mu(\mathbf{d}, a) = a^p E[Z^p g(\mathbf{D})]$ and $\varepsilon(\mathbf{N}, \mathbf{d}) = Z^p g(\mathbf{D})/E[Z^p g(\mathbf{D})]$. From the result of Section 2.1, we see that under model (2.5) Taguchi's two-stage pro-

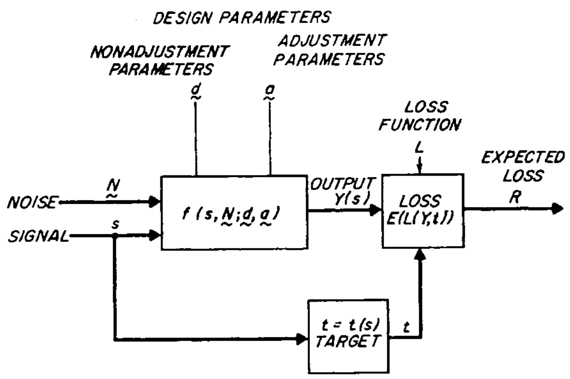

*Figure 2. A Generic Parameter-Design Problem. The output Y(s) is determined by an input signal s and noise **N** through a transfer function f. The transfer function depends on the design parameters d and **a**. Loss is incurred if the output is not equal to the target value t(s).*

cedure leads to the minimization of average quadratic loss.

3. GENERAL MODEL FOR PARAMETER DESIGN PROBLEMS AND A CONVENIENT TWO-STEP SOLUTION

As mentioned in Section 1.2, a convenient way of thinking about parameter design problems is in terms of block diagrams similar to that in Figure 1. Another example of a block-diagram representation is given in Figure 2. In this case, for a given setting of the design parameters **d** and **a**, the output Y is determined by an input signal s and by noise N. Loss is incurred if the output is different from a target that may depend on the input signal. As before, the goal of parameter design is to choose the setting of the design parameters that minimizes expected loss,

$$R(\mathbf{d}, \mathbf{a}) = E_s E_{\mathbf{N}}(L(Y, t(s)) \mid \mathbf{d}, \mathbf{a}), \qquad (3.1)$$

where the distribution of the signal, s, reflects the relative frequency of its different values.

This kind of system is called *dynamic*, because the output and the target depend on a signal that is not fixed by the designer. For example, a measuring instrument such as a bathroom scale is a dynamic system because its output, a weight reading, depends on the input signal, the actual weight of the person standing on the scale. The weight reading is also affected by noise factors such as temperature.

The block diagram in Figure 2 does not include every parameter-design problem. For example, there may not be a signal, as we saw in the static problem of Section 1.2. Moreover, in the control problem of Section 4.2 the signal is determined by the target rather than vice versa. (The block diagram corresponding to this problem is shown in Fig. 3.) The block diagram can be appropriately modified to represent most parameter-design problems, however. For example, Khosrow Dehnad of AT&T Bell Laboratories has suggested that for some problems, such as the design of an amplifier, the function $t(s)$ might be replaced by the class of all linear functions, since to avoid distortion an amplifier's output must be close to a linear function of its input.

The objective of parameter design is to find the values of design parameters (**d**, **a**) that minimize average loss $R(\mathbf{d}, \mathbf{a})$. This can conveniently be done in many parameter design problems by following Procedure 2. As mentioned in Section 2, this two-step procedure is always possible although it may not be useful. Under certain models and loss functions, however, $P(\mathbf{d})$ is equivalent to Taguchi's SN ratio and Procedure 2 is equivalent to Procedure 1. In Section 2 we demonstrated this claim for the static parameter-design problem. In the rest of the article we show other examples in which $P(\mathbf{d})$ is equivalent to the SN ratio given by Taguchi.

Under other models and other loss functions, however, the same two-step procedure is possible, but $P(\mathbf{d})$ is not equivalent to Taguchi's SN ratio. For this reason, when a procedure like Procedure 2 can be conveniently used, we call $P(\mathbf{d})$ (or a monotone function of it) a PerMIA, rather than an SN ratio.

For fixed **d** we let $\mathbf{a}^*(\mathbf{d})$ be the value of **a**, where $R(\mathbf{d}, \mathbf{a})$ is minimized. That is, $P(\mathbf{d}) = \min_{\mathbf{a}} R(\mathbf{d}, \mathbf{a}) = R(\mathbf{d}, \mathbf{a}^*(\mathbf{d}))$.

4. PERFORMANCE MEASURES FOR DYNAMIC PARAMETER DESIGN PROBLEMS

As discussed in Section 3, parameter design problems in which the output and the target depend on a signal are called dynamic. As in all parameter design problems, the objective is to find the settings of design parameters of a product or process so that

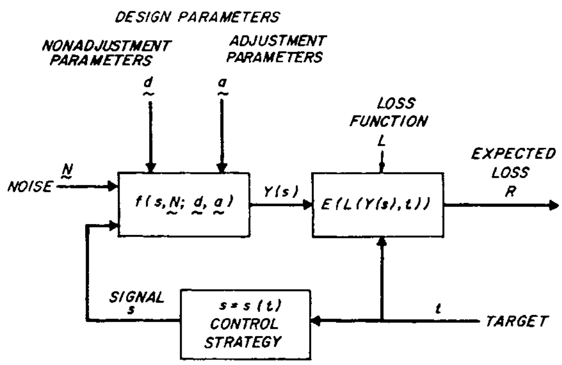

*Figure 3. Parameter Design of a Control Problem. The input signal is determined by the control strategy s(t) and the intended target t. The object of the parameter design is to find the setting of the design parameters (**d**, **a**) that minimizes expected loss R(**d**, **a**). A more general objective of the parameter-design problem would be to find the minimum expected loss not only over the design parameter settings, but also over the permissible control strategies.*

loss caused by deviation of output from target is minimized. In this section, we consider two examples of dynamic parameter design problems. The first example concerns the design of a measuring instrument, such as a spring scale. The performance measure is independent of calibration parameters, allowing comparison of scale designs without calibration. In the second example, we derive a performance measure for a generic control problem in which the objective is to design a product so that the output can be varied over a certain interval by changing the input signal. In Appendix A, we discuss in detail a specific control problem involving the design of a truck-steering mechanism.

Both of these examples are of the type that Taguchi and Phadke (1984) called continuous-continuous dynamic problems because the input and output are continuous variables. Taguchi (1976, 1977) suggested a single SN ratio for all such problems. To see what this SN ratio is, let the general form of the continuous-continuous dynamic problem be given by the formula

$$Y = \alpha + \beta' s + \varepsilon', \qquad (4.1)$$

where s is a continuous signal controlling a continuous output Y, and ε' is some deviation from linearity. Then Taguchi (see Taguchi and Phadke 1984) gave the SN ratio

$$\mathrm{SN} = \log_{10}[\beta'^2/\mathrm{var}(\varepsilon')]. \qquad (4.2)$$

This SN ratio is defined by formulas (4.1) and (4.2) regardless of the actual form of the model for Y.

In the first example considered in this section, we follow Procedure 2 for finding a PerMIA. We show that the SN ratio (4.2) is a PerMIA under a certain special model, and therefore its use leads to minimization of squared-error loss. If a different model holds, however, use of this SN ratio is not appropriate. In the second example, the SN ratio is not independent of adjustment parameters, but a heuristic argument is presented that suggests that the SN ratio may sometimes lead to a solution close to the optimum.

4.1 An Explanation of an SN Ratio: Measuring Instrument Example

Suppose a manufacturer wants to design a measuring instrument and that the dial reading Y of the instrument satisfies

$$Y = \alpha(\mathbf{d}, a_1) + \beta(\mathbf{d}, a_2)(\gamma(\mathbf{d})s + \varepsilon(\mathbf{N}; \mathbf{d})), \qquad (4.3)$$

where $\mathrm{var}[\varepsilon(\mathbf{N}; \mathbf{d})] = \sigma^2(\mathbf{d})$ and $E[\varepsilon(\mathbf{N}; \mathbf{d})] = 0$. Here s is the true value of the measured quantity and Y is the reading on the instrument's dial.

To see why this model might be appropriate for a measuring instrument that can be calibrated, consider the case of a spring scale such as a postal scale. The scale consists of two parts, the sensor, which is the spring, and a dial, which translates the spring's compression into a weight reading. A model for the sensor part of the scale is $Z = \gamma(\mathbf{d})s + \varepsilon(\mathbf{N}; \mathbf{d})$, where Z is the compression (in inches) of the spring, s is the true weight of an object, $\mathbf{d}$ is the design parameters of the spring such as spring size and alloy and $\mathbf{N}$ is noises, such as imperfections in the spring, which affect compression. Since the location and spacing of markings on the dial can be chosen by the designer, a model for the dial reading, Y, is $Y = \alpha(\mathbf{d}, a_1) + \beta(\mathbf{d}, a_2)Z$, where a_1 and a_2 are adjustments chosen by the designer. Substituting for Z gives model (4.3).

For the general measuring instrument described by model (4.3), the desired value for the dial reading Y is equal to the true value s. Suppose loss is measured by $(Y - s)^2$; then the objective of parameter design is to find the setting of $(\mathbf{d}, \mathbf{a})$ that minimizes $E_s E_{\mathbf{N}}[(Y - s)^2 \mid \mathbf{d}, \mathbf{a}]$. In addition, suppose the designer requires that

$$(\mathbf{d}, \mathbf{a}) \in \{(\mathbf{d}, \mathbf{a}): \ E_{\mathbf{N}}(Y \mid \mathbf{d}, \mathbf{a}, s) = s\},$$

that is, the design must give an unbiased estimate of s. Then, to find the optimal setting, $(\mathbf{d}^*, \mathbf{a}^*)$, note that

$$\begin{aligned} &\min_{\mathbf{d}} \min_{\mathbf{a}} E_s E_{\mathbf{N}}[(Y - s)^2 \mid \mathbf{d}, \mathbf{a}] \\ &= \min_{\mathbf{d}} \min_{\mathbf{a}} \{\beta^2(\mathbf{d}, a_2)\, \mathrm{var}_{\mathbf{N}}(\varepsilon(\mathbf{N}; \mathbf{d})) \\ &\quad + E_s[\alpha(\mathbf{d}, a_1) + \beta(\mathbf{d}, a_2)\gamma(\mathbf{d})s - s]^2\} \\ &= \min_{\mathbf{d}}[\mathrm{var}_{\mathbf{N}}(\varepsilon(\mathbf{N}; \mathbf{d}))/\gamma^2(\mathbf{d})]. \end{aligned}$$

The last of the preceding equalities holds because $a_1^*(\mathbf{d})$ and $a_2^*(\mathbf{d})$ must satisfy $\alpha(\mathbf{d}, a_1^*(\mathbf{d})) = 0$ and $\beta(\mathbf{d}, a_2^*(\mathbf{d}))\gamma(\mathbf{d}) = 1$ for each $\mathbf{d}$, under the unbiasedness constraint.

The function $\mathrm{var}_{\mathbf{N}}(\varepsilon(\mathbf{N}; \mathbf{d}))/\gamma^2(\mathbf{d})$ is a PerMIA, because it does not depend on the adjustment parameters, a_1 and a_2. This PerMIA measures the performance of the measuring instrument as it would be after proper calibration.

Hence parameter design can be done using the following two-step procedure.

Procedure 4.

Step 1. Find $\mathbf{d}^*$ that minimizes the PerMIA. $\mathrm{var}_{\mathbf{N}}(\varepsilon(\mathbf{N}; \mathbf{d}))/\gamma^2(\mathbf{d})$.

Step 2. Find a_1^* and a_2^* for which $\alpha(\mathbf{d}^*, a_1^*) = 0$ and $\beta(\mathbf{d}^*, a_2^*) = 1/\gamma(\mathbf{d}^*)$.

As seen by comparing formulas (4.1) and (4.3), $\beta' = \beta(\mathbf{d}, a_2)$ and $\varepsilon' = \beta(\mathbf{d}, a_2)\varepsilon$. So do optimizing the PerMIA is equivalent to optimizing the SN ratio given in formula (4.2), the one suggested by Taguchi for all continuous-continuous parameter design problems. In fact, the derivation of the PerMIA given in this section is motivated by the discussion in chapter 22 of Taguchi (1977).

Let us see what happens, however, if we retain the same loss function but change the model for the measuring instrument as follows:

$$Y = \alpha(\mathbf{d}, a_1) + \beta(\mathbf{d}, a_2)s + \varepsilon(\mathbf{N}; \mathbf{d}),$$

where $E(\varepsilon(\mathbf{N}; \mathbf{d})) = 0$ and $\text{var}(\varepsilon(\mathbf{N}; \mathbf{d})) = \sigma^2(\mathbf{d})$. Then, retracing the previous argument, the PerMIA for the measuring-instrument parameter design can be shown to be $\sigma^2(\mathbf{d})$. But under this model the PerMIA is not equivalent to Taguchi's SN ratio. In fact, under this model the SN ratio depends on the adjustment parameter a_2. Hence using the SN ratio as if it were independent of both adjustment parameters could lead to unfair comparisons of alternative settings of the nonadjustment design parameters **d**.

We emphasize throughout this article that models and loss functions should be carefully stated for each parameter design problem, because the appropriate performance measure may be very dependent on them. The SN ratio formula (4.2) cannot be safely used for all continuous-continuous parameter design problems.

4.2 Performance Measure for a Control System Example

As an illustration of a control system, consider a truck's steering mechanism. The truck driver and the truck make up a control system. The driver chooses a certain steering angle, $s(t)$, to accomplish a turn of radius t. The design of the steering mechanism must allow the driver to find a steering angle for any turning radius over some range. The possible steering angles may also be restricted to some interval. In addition, the turn of radius of the truck resulting from a given steering angle should be consistent and not affected by noise conditions such as road surface, load distribution, and tire air pressure.

The parameter design of a truck's steering mechanism is considered in detail in Appendix A. In this section we present a generic control problem for which we postulate a model and derive a PerMIA. Although the model assumed here is different from the model that is derived for the steering mechanism in Appendix A, the PerMIA's turn out to be the same in both cases.

The block diagram in Figure 3 describes a general control problem. The input signal s is determined by the intended target t and the control strategy $s(t)$. For simplicity we assume that $s(t)$ is the unbiased control strategy given by

$$s(t) = \{s\colon\ E(Y(s)\,|\,\mathbf{d}, \mathbf{a}, s) = t\},$$

where $Y(s)$ is the system output corresponding to signal s. The objective of parameter design is to find the setting of the design parameters, $(\mathbf{d}, \mathbf{a})$, which minimizes expected loss $R(\mathbf{d}, \mathbf{a})$.

A more general objective of the parameter design problem would be to minimize expected loss not only over the settings of the design parameters but also over some set of permissible control strategies. In this article, however, we restrict that set to consist of only one strategy.

For a given setting of the design parameters, $(\mathbf{d}, \mathbf{a})$, suppose the following model describes the relationship between the output Y, the signal s, and noise $\mathbf{N}$:

$$Y = \alpha(\mathbf{d}, \mathbf{a}) + \beta(\mathbf{d})s + \varepsilon(\mathbf{N}, \mathbf{d}), \tag{4.4}$$

where $E(\varepsilon(\mathbf{N}, \mathbf{d})) = 0$ and $\text{var}(\varepsilon(\mathbf{N}, \mathbf{d})) = \sigma^2(\mathbf{d})$. Assuming that loss caused by deviation of Y from t is measured by $(Y - t)^2$, the objective is to find the setting of $(\mathbf{d}, \mathbf{a})$ that minimizes $E_t E_{\mathbf{N}}[(Y - t)^2\,|\,\mathbf{d}, \mathbf{a}]$. For any target t in (t_L, t_H), the unbiased control strategy dictates that $s = s(t)$ will be chosen from (s_L, s_H) so that $E(Y\,|\,\mathbf{d}, \mathbf{a}, s) = t$. This will be possible only if the slope $\beta(\mathbf{d})$ is such that

$$|\beta(\mathbf{d})| \geq |(t_H - t_L)/(s_H - s_L)|.$$

Consequently, the design parameters **d** must satisfy this constraint. Notice that there is no similar constraint for the y-intercept $\alpha(\mathbf{d}, \mathbf{a})$, because, if the slope is steep enough, the adjustment variable **a** can be used to shift $E(Y\,|\,\mathbf{d}, \mathbf{a}, s)$ up or down.

It follows that the two-step decomposition of the parameter design is the following.

Procedure 5.

Step 1. Find $\mathbf{d}^*$ to minimize $\sigma^2(\mathbf{d})$ subject to the constraint

$$\beta^2(\mathbf{d}) \geq [(t_H - t_L)/(s_H - s_L)]^2.$$

Step 2. Choose $\mathbf{a}^*$ so that $\alpha(\mathbf{d}^*, \mathbf{a}^*) + \beta(\mathbf{d}^*)s$, $s \in [s_L, s_H]$ can cover the target interval $[t_L, t_H]$.

The constraint on **d** is frequently important in practice, since values of **d** that tend to make the effects of the noise small—that is, $\sigma^2(\mathbf{d})$ small—may also tend to make the effect of the signal small—that is, β small. As an extreme example, the truck's steering mechanism is least sensitive to noise if it is welded solid, but then the ability to steer is lost completely.

As mentioned previously, Taguchi and Phadke (1984) and Taguchi (1976, 1977) recommended a different performance measure for all continuous-continuous dynamic problems, namely

$$\text{SN} = 10\,\log_{10}[\beta^2(\mathbf{d})/\sigma^2(\mathbf{d})]. \tag{4.5}$$

In the parameter design of a control system such as the design of a truck-steering mechanism, however, the principal objective is to minimize the effect of noise on output while maintaining at least a *minimum* sensitivity to the signal. Maximizing SN in Equation (4.5) may approximately accomplish this goal, because values of $\mathbf{d}$ that make $\beta^2(\mathbf{d})/\sigma^2(\mathbf{d})$ large will tend to make $\sigma^2(\mathbf{d})$ small and make $\beta^2(\mathbf{d})$ large (and hence satisfy our constraint). It seems unnecessary, however, to try to make β much larger than the minimum, especially if that results in a larger value for σ^2.

5. PERFORMANCE MEASURE FOR A BINARY-INPUT BINARY-OUTPUT DYNAMIC PROBLEM

A type of dynamic parameter-design problem that arises in the communications and electronics industries is the binary-input–binary-output problem. This case is identical to the continuous dynamic problem discussed in Section 4, except that both the input and output are binary rather than continuous.

In this section we will derive a performance measure for a simple binary-input–binary-output problem, the design of a binary transmission channel. We show that certain model assumptions lead to a performance measure independent of a channel-adjustment parameter. This derivation reveals conditions under which the SN ratio that Taguchi proposed for binary-input–binary-output problems is independent of the adjustment parameter and leads to minimization of average loss.

5.1 Performance Measure for a Binary Transmission Channel

The purpose of a binary channel is to transmit a 0–1 signal over a wire. The signal must actually be sent over the wire as a voltage, however, so a 0 is converted to a voltage μ_0 and a 1 to a voltage μ_1 ($\mu_0 \le \mu_1$). The voltages actually received are $\mu_0 + \sigma\varepsilon$ and $\mu_1 + \sigma\varepsilon$, because electrical disturbance in the wire adds a random noise ε to the signal. This noise is assumed to be normally distributed with mean 0 and variance (see Pierce 1980, chap. 8; Pierce and Posner 1980, chap. 8). The value of σ is a positive number that reflects the channel's reaction to the noise. The values of μ_0, μ_1, and σ depend on the settings of some design parameters, $\mathbf{d}$. If the received voltage is less than some voltage threshold a, it is assumed that a 0 was sent. If the received voltage is greater than a, it is assumed that a 1 was sent. Note that a is also a design parameter (see Fig. 4).

To describe the problem in the framework of Figure 2, note that the signal s is the 0–1 input and the output Y is the received 0–1 message. Since in this example the purpose is to reproduce the signal, the target t is related to the signal by the identify function $t(s) = s$. The noise $\mathbf{N}$ is the electrical disturbance in the wire that affects the system through the function $\varepsilon = \varepsilon(\mathbf{N})$. The transfer function is as given in Figure 4, or in symbols,

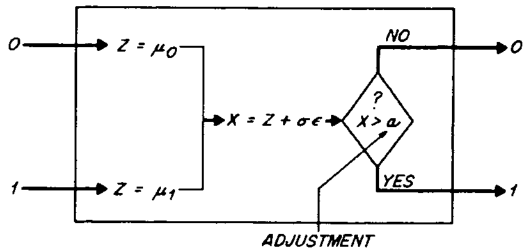

Figure 4. Schematic of Binary Channel. The engineer can send a known stream of 0s and 1s, and observe the resulting output of 0s and 1s. The engineer can change the misclassification probabilities by adjusting a.

$$\begin{aligned} Y &= f(s, \mathbf{N}; \mathbf{d}, a) \\ &= 1 \quad \text{if } \mu_s + \sigma\varepsilon > a \\ &= 0 \quad \text{if } \mu_s + \sigma\varepsilon \le a, \end{aligned}$$

where μ_0, μ_1, and σ are functions of the design parameters $\mathbf{d}$. The letter a stands, as before, for the voltage threshold. The loss function is given by

$$\begin{aligned} L(Y, t) = (Y - t)^2 &= 0 \quad \text{if } Y = t \\ &= 1 \quad \text{if } Y \ne t. \end{aligned}$$

Assuming that 0s and 1s are sent with equal frequency, the expected loss is given by

$$\begin{aligned} R(\mathbf{d}, a) &= E_s E_{\mathbf{N}}(L(y, t) \mid \mathbf{d}, a) \\ &= \Pr(s = 0, Y = 1) + \Pr(s = 1, Y = 0) \\ &= \Pr(Y = 1 \mid s = 0) \Pr(s = 0) \\ &\quad + \Pr(Y = 0 \mid s = 1) \Pr(s = 1) \\ &= \tfrac{1}{2}(p_0 + p_1), \end{aligned}$$

where $p_0 = p_0(\mathbf{d}, a)$ and $p_1 = p_1(\mathbf{d}, a)$ are, respectively, the probabilities of misclassifying a 0 and a 1.

The designers would like to compare several candidate settings of the design parameters, $(\mathbf{d}, a)$, to find the one that minimizes average loss caused by misclassifying the signal. [More correctly, the engineer would like to maximize channel capacity. This criterion from information theory leads to a similar analysis (see Pierce 1980, chap. 8).] Assume that the engineers cannot observe voltages but that they can send a known sequence of 0s and 1s into the channel, and observe the resulting output.

What is a convenient way to find the optimal value of the design parameters $(\mathbf{d}, a)$? One can follow Procedure 2. To see how, first note that the error induced by noise is assumed to have a normal distribution, so

$$p_0 = \Phi[(\mu_0 - a)/\sigma], \qquad p_1 = \Phi[(a - \mu_1)/\sigma],$$

where a is the current voltage threshold and Φ is the standard normal cumulative distribution function. Then note that

$$R(\mathbf{d}, a) = \frac{1}{2}\left\{\Phi\left(\frac{\mu_0 - a}{\sigma}\right) + \Phi\left(\frac{a - \mu_1}{\sigma}\right)\right\}$$

is minimized at $a^*(\mathbf{d}) = [\mu_0(\mathbf{d}) + \mu_1(\mathbf{d})]/2$. That is, for a given value of $\mathbf{d}$ the optimal value of the voltage threshold is halfway between the two average transmission voltages (this is easy to prove).

Using the fact that

$$\Phi[(\mu_0 - a^*(\mathbf{d}))/\sigma] = \Phi[(a^*(\mathbf{d}) - \mu_1)/\sigma],$$

we then see that

$$R(\mathbf{d}, a^*(\mathbf{d})) = \Phi\left(\frac{(\mu_0 + \mu_1)/2 - \mu_1}{\sigma}\right) = \Phi\left(\frac{\mu_0 - \mu_1}{2\sigma}\right) = \Phi\left(\frac{(\mu_0 - a)/\sigma + (a - \mu_1)/\sigma}{2}\right).$$

That is,

$$R(\mathbf{d}, a^*(\mathbf{d})) = \Phi\left(\frac{\Phi^{-1}(p_0) + \Phi^{-1}(p_1)}{2}\right). \quad (5.1)$$

The function $P(\mathbf{d}) = \min_a R(\mathbf{d}, a) = R(\mathbf{d}, a^*(\mathbf{d}))$ is a PerMIA, because it can be used to determine the value of the nonadjustment design parameters $\mathbf{d}$ at which channel performance is optimal while ignoring the value of the adjustment parameter a.

It follows that the engineer can follow the following two-step procedure given to find the optimal value $(\mathbf{d}^*, a^*)$ of $(\mathbf{d}, a)$.

Procedure 6.

Step 1. Find d^* that minimizes

$$\Phi\left(\frac{\Phi^{-1}(p_0(\mathbf{d}, a)) + \Phi^{-1}(p_1(\mathbf{d}, a))}{2}\right),$$

where a is in the arbitrary present position.

Step 2. Find a^* for which $p_0(\mathbf{d}^*, a^*) = p_1(\mathbf{d}^*, a^*)$.

Step 2 follows, because the misclassification probabilities are equal iff a minimizes $R(\mathbf{d}^*, a)$. Step 2 is carried out by empirically adjusting a back and forth until misclassification probabilities are equal.

We note that without Procedure 6, finding the optimal value $\mathbf{d}^*$ would involve doing the previous time-consuming adjustment of a for each candidate value of $\mathbf{d}$.

5.2 Explanation of Taguchi's SN Ratio

Taguchi and Phadke (1984) (following Taguchi 1976, 1977) suggested a formula other than (5.1) for their two-step procedure. We have found that their formula follows from the preceding argument if the error ε caused by noise has a standard logistic distribution rather than a standard normal distribution. Taguchi and Phadke did not use that value as the performance measure, however. Instead, they used a decreasing function of it, which they called the SN ratio. In Appendix B we show that this SN ratio would be a meaningful measure if the binary channel problem were modified so that the received voltages were observable. The reader interested in understanding the relevance of the SN ratio concept of communication engineering to parameter design may want to read Appendix B.

6. ADJUSTMENT PARAMETERS IN PRACTICE

6.1 How Adjustment Parameters Arise in Practice

The problem of identifying adjustment design parameters may seem difficult, but in practice these parameters are often suggested by engineering knowledge and product or process design conventions. In particular, adjustment parameters are frequently those design parameters used to fine-tune performance after the other, more difficult-to-change design parameters have been set.

As illustrated previously, examples of adjustment parameters include the size of the mold used in making tiles, scale markings on a measuring instrument, and the gear ratio in a steering mechanism (App. A). Other examples of adjustment parameters are the length of the pendulum in a clock and the treble, bass, volume, and balance knobs on a stereo receiver.

6.2 The Value of Adjustment Parameters for Product and Process Designers

Adjustment parameters are often specifically designed into a product or process to make the design more flexible. For example, certain changes in design specifications may be accommodated by changing the setting of an adjustment parameter. In the tile example, a change in the desired tile size is easily accomplished by changing the mold size. The clay formulation need not be changed.

In other examples, adjustment parameters may make a product adaptable to a variety of customer use conditions. The carpet-height adjustment of a vacuum cleaner, for example, allows good cleaning on a range of carpet-pile depths.

6.3 Advantages of Measuring Performance Independent of Adjustment Parameters

When adjustment parameters are included in the design of a product or process, measuring that product's or process's performance independent of adjustment parameter settings can be very advantageous. For example, when an adjustment parameter is meant to accommodate future changes in design specifications (such as the desired tile size), use of a PerMIA as performance measure ensures that non-adjustment design parameter values that are optimal for one specification are also optimal for other specifications.

If the adjustment is going to be made by the customer, a PerMIA can be particularly useful to a product designer. It allows the designer to gauge the performance of the product independently of adjustments that the customer might make to suit his particular situation. For the product designer, the PerMIA measures performance *as it would be* after proper adjustment. There is no need to test each alternative product design under every customer-use condition.

Use of a PerMIA also reduces the dimensionality of the parameter design optimization problem. This may be advantageous, because lower-dimensional optimization problems are generally easier to solve.

A PerMIA can sometimes be used to turn a constrained parameter design optimization problem into an unconstrained one, as is the case in the *static* parameter design problem (see the discussion following Procedure 3).

PerMIA's can be used to simplify parameter-design experiments. PerMIA's can be derived *analytically* using minimal engineering knowledge of product or process behavior (as captured by the model). The empirical model identification and fitting is limited to $P(\mathbf{d})$ and $R(\mathbf{d}^*, a)$ rather than to the more complex $R(\mathbf{d}, \mathbf{a})$, which could require cross-terms between the parameters $\mathbf{d}$ and $\mathbf{a}$.

6.4 Empirical Methods for Identifying a PerMIA

Choice of a performance measure independent of adjustment depends on some knowledge of the form of the transfer-function model. In a particular application, engineering knowledge might indicate a multiplicative model like (2.2) rather than an additive one like (2.4). Empirical studies can check or supplement engineering knowledge, however. For the static case, Box (1986) described how the form of the transfer function model can be investigated by estimating the transformation of the output Y required to induce additivity in the transfer function model.

7. SUMMARY

Parameter design is a process for finding product designs that are insensitive to noise such as manufacturing and environmental variability. Operationally, the objective of parameter design is to find the setting of the product's design parameters that minimizes expected loss caused by noise. G. Taguchi pioneered the use of statistically planned experiments for parameter design. His work has led to wide application of this method in Japan and its increasing use by U.S. industry.

In parameter design, Taguchi (see Taguchi 1977; Taguchi and Phadke 1984) used performance measures that he called SN ratios. In general he gave no connection between these performance measures and his stated objective of minimizing loss caused by noise.

In this article we have shown that *if certain models for the product or process response are assumed*, then maximization of the SN ratio leads to minimization of average squared-error loss. The SN ratios take advantage of the existence of special design parameters called adjustment parameters. When these parameters exist, use of the SN ratio allows the parameter design optimization procedure to be conveniently decomposed into two smaller optimization steps.

In these situations, the desire to find a two-step optimization procedure seems to be a major motivation behind Taguchi's use of SN ratios. When different models hold, however, a two-step procedure is possible only if a performance measure different from the SN ratio is used. Hence there are many real problems in which the SN ratio is not independent of adjustment parameters, and its use could lead to far from optimal design parameter settings.

Because adjustment parameters bring definite advantages to product and process designers, we propose a type of performance measure that takes advantage of the existence of adjustment parameters and is more general than Taguchi's SN ratios. We call these measures performance measures independent of adjustment, or PerMIA's.

When an adjustment parameter exists, a PerMIA can be conveniently derived directly from knowledge of the loss function and the general form of the transfer-function model. We have given a procedure for doing this and illustrated its use in several generic parameter design problems. In some of these problems we saw that following the procedure leads directly to the SN ratio that Taguchi recommended.

ACKNOWLEDGMENTS

We gratefully acknowledge the helpful comments of R. Berk, K. Dehnad, R. Easterling, C. Fung, A. B. Godfrey, J. H. Hooper, W. Q. Meeker, V. N. Nair, W. A. Nazaret, P. G. Sherry, C. S. Sherrerd, C. F. Wu, and the referees.

APPENDIX A: DESIGNING A TRUCK-STEERING MECHANISM, A DETAILED EXAMPLE

In this appendix, we show how a PerMIA is derived in a realistic control parameter design problem, the design of a truck-steering mechanism. This example, a modification of one given by Taguchi and Wu (1980), also illustrates how adjustment parameters arise in practice. The truck driver chooses the steering angle s to make a turn of radius t. (Turning angle may seem a more natural measure of the truck's response to a certain steering angle, but turning radius is much easier to measure.) The chosen steering angle is determined by the driver's control strategy. In designing the truck-steering mechanism, the engineer's objective is to minimize the expected loss caused by deviation of the truck's actual turning radius from the driver's intended turning radius. This deviation is caused by noise such as road condition, load position, and tire pressure. [Driving speed is another noise that has a profound effect on how the truck responds to a given steering angle. For simplicity, we assume that speed is constant. See Taguchi and Wu (1980) for a solution that includes speed as a noise factor.]

As mentioned previously, one way to make the steering mechanism insensitive to noise is to weld the steering mechanism so that the truck always goes straight. Of course, this would not do, because, in addition to minimizing sensitivity to noise, the design must allow the driver to make any needed turn using a comfortable steering angle. In particular, since the size of the steering angle is inversely related to the force required to turn the steering wheel, comfortable steering angles are neither too large nor too small—small angles require too much exertion; large angles require too many turns of the steering wheel. Suppose the engineer expects that the driver will need to make turns between t_L meters and t_H meters and force required to turn the steering wheel, comfortable steering angles are neither too large nor too small—small angles require too much exertion; large angles require too many turns of the steering wheel. Suppose the engineer expects that the driver will need to make turns between t_L meters and t_H meters and comfortable steering angles for making these turns are between s_L and s_H degrees.

As before, the block diagram in Figure 3 summarizes the problem from the point of view of parameter design. The truck driver who intends to make a turn of radius t chooses a steering angle s using his control strategy $s(t)$. The truck responds with a turn of radius Y, which depends on the steering angle s, the noise conditions $\mathbf{N}$, and the design of the steering mechanism.

In this example the design of the steering mechanism is determined by three design parameters: hardness of front springs (A), type of steering geometry (B), and gear ratio (G). From the geometry of steering, it can be shown that it is reasonable to assume that the relationship between Y, s, $\mathbf{N}$, A, B, and G is given by the transfer function $Y = (Gs)^{-\delta(\mathbf{N}, A, B)}$, where δ is a positive function. Taking logs and doing some obvious renaming, the transfer function can be rewritten as

$$\log Y = [\alpha(A, B, G) - \beta(A, B) \log s]\varepsilon(\mathbf{N}, A, B), \qquad (A.1)$$

where we force the condition $E_{\mathbf{N}}(\varepsilon(\mathbf{N}, A, B)) = 1$.

Model (A.1) is similar to model (4.4) of Section 4.2, except that the error term is now multiplicative rather than additive. Both models, however, lead to the same PerMIA, as shown in the following.

Recall that the designer's objective is to find the setting of the design parameters that minimizes the expected loss caused by deviation of Y from t. If this loss is $L(Y, t) = (Y - t)^2$, then no PerMIA seems to be available. A PerMIA can be obtained, however, if the loss could be well approximated by $L(Y, t) = \log Y - \log t)^2$. Then expected loss is

$$\begin{aligned} R(A, B, G) &= E_s E_{\mathbf{N}}[(\log Y - \log t)^2 \mid A, B, G, s(t)] \\ &= E_s\{\mathrm{var}_{\mathbf{N}}(\log Y \mid A, B, G, t) \\ &\quad + [E_{\mathbf{N}}(\log Y \mid A, B, G, t) - \log t]^2\}. \end{aligned}$$

Hence, assuming the control strategy

$$s(t) = \{s\colon \ E_{\mathbf{N}}[\log Y \mid A, B, G, t] = \log t\}$$

under model (A.1), we get

$$R(A, B, G) = E_s[(\log t)^2]\sigma^2(A, B),$$

where $\mathrm{var}_{\mathbf{N}}(\varepsilon(\mathbf{N}, A, B)) = \sigma^2(A, B)$. Therefore, the ability of a driver to make a desired turn is measured by $\sigma^2(A, B)$.

It is not enough to find the values of A and B to minimize $\sigma^2(A, B)$, however. Recall that the steering mechanism must allow the driver to make any turn of radius between t_L meters and t_H meters using steering angles between s_L and s_H degrees. From model (A.1), this will be possible on the average if the slope coefficient satisfies the constraint

$$\beta^2(A, B) \geq \left[\frac{\log t_L - \log t_H}{\log s_H - \log s_L}\right]^2. \qquad (A.2)$$

If a design satisfies this constraint, the gear ratio G can be adjusted so that

$$\begin{aligned} E(\log Y \mid A, B, G, s_H) &\leq \log t_L < \log t_H \\ &\leq E(\log Y \mid A, B, G, s_L). \end{aligned}$$

Notice that neither the performance measure $\sigma^2(A, B)$ nor the constraint (4.4) depends on the gear ratio

G. That is, $\sigma^2(A, B)$ is a PerMIA, and its constrained optimization can be done with no dependence on the value of the adjustment parameter G.

Notice that with Model (A.1) it is not even clear how to substitute into Taguchi's SN ratio formula (4.2) for the continuous-continuous dynamic parameter-design problem.

APPENDIX B: TAGUCHI'S SN FOR THE BINARY CHANNEL EXAMPLE

Communication engineers often use the SN ratio as a measure of a transmission channel's performance. This is done because in many types of channels the probability of transmission error is a decreasing function of the SN ratio, or the channel capacity is an increasing function of the SN ratio. The channel capacity is the theoretical maximum bit rate of information that can be transmitted over the channel (see Pierce 1980, chaps. 8 and 9).

Assume that in Section 5 the received voltages $\mu_0 + \sigma\varepsilon$ and $\mu_1 + \sigma\varepsilon$ are observable. Let $\mu = (\mu_0 + \mu_1)/2$, $\alpha_0 = \mu_0 - \mu$, $\alpha_1 = \mu_1 - \mu$, and $\varepsilon' = \sigma\varepsilon$. Then the received voltages X_i ($i = 1, 2, \ldots$) have the form

$$X_{i,k} = \mu + \alpha_k + \varepsilon'_{i,k}, \qquad k = 0, 1$$
$$\alpha_0 + \alpha_1 = 0$$
$$\varepsilon'_i \sim N(0, \sigma^2). \tag{B.1}$$

The variance σ^2 is the average power of the noise and $\sigma_\alpha^2 = (\alpha_0^2 + \alpha_1^2)/2 = \alpha_0^2 = \alpha_1^2$ is the average power of the signal. (Since power is voltage times current and by Ohm's law voltage is proportional to current, it follows that power is proportional to voltage squared.) Hence σ_α^2/σ^2 is the ratio of average signal power to average noise power (SN ratio), and the quantity $10 \log_{10}(\sigma_\alpha^2/\sigma^2)$ is how strong the signal power is relative to the noise power in decibel units.

Now as shown by Pierce and Posner (1980, chap. 8) the transmission error probability after leveling is a decreasing function of the SN ratio. Hence an engineer trying to find the settings of design variables that minimize this probability could pick the settings that maximize the SN ratio or equivalently its decibel value. In other words, the SN ratio is an appropriate performance measure.

The problem is how to estimate the SN ratio on the basis of sending a known sequence of n zeros and n ones through the channel. One can think of model (B.1) as a one-way layout analysis of variance problem. Then we see that the expected mean squared for treatment (MST) is given by $E(\text{MST}) = \sigma^2 + n(2\sigma_\alpha^2)$, and the expected mean squared error (MSE) is given by $E(\text{MSE}) = \sigma^2$. Hence a simple estimate of the SN ratio σ_α^2/σ^2 is the method of moments estimate, given by (MST − MSE)/2nMSE. So the engineer conducting the experiment would try to find the setting of design variables that maximizes

$$10 \log_{10}\{(\text{MST} - \text{MSE})/(n\text{MSE})\}5. \tag{B.2}$$

In the original binary channel problem, however, voltage is not observable. In solving this problem, Taguchi and Phadke (1984) suggested that the estimate (B.2) still be used on the binary output. The binary output is used to calculate MST and MSE, which are then substituted into formula (B.2). After some algebra, this process yields

$$\eta = 10 \log_{10}\left\{\frac{(1 - p_0 - p_1)^2}{p_0(1 - p_0) + p_1(1 - p_1)}\right\}. \tag{B.3}$$

Then they used a formula that is equivalent to formula (5.1) if ϕ is the standard logistic density instead of the standard normal density. Upon substitution of this formula in formula (B.3), they got what they called the SN ratio after leveling:

$$\eta = 10 \log_{10}[(1 - 2q)^2/2q(1 - q)], \tag{B.4}$$

where q is the counterpart of $R(\mathbf{d}, a^*(\mathbf{d}))$ in formula (5.1) for the logistic distribution.

Unfortunately, model (B.1) cannot hold when $X_{i,k}$ ($i = 1, 2, \ldots, n$; $k = 1, 2$) are binary variables. Moreover, since these binary variables are not voltages (or currents), it is not clear in what sense η in formula (B.4) is an SN ratio. But η is a decreasing function of q on $[0, \frac{1}{2}]$, since after some algebra it can be shown that

$$\eta = 10 \log_{10}\left[\frac{8}{1/(q - .5)^2 - 4}\right].$$

Thus the setting of the design parameters that maximizes η also minimizes q. It follows that the approach of Taguchi and Phadke (1984) is equivalent to that given in Section 5.1.

[*Received September 1985. Revised March 1987.*]

REFERENCES

Box, G. E. P. (1986), "Studies in Quality Improvement: Signal-to-Noise Ratios, Performance Criteria and Statistical Analysis: Part I," Technical Report 11, University of Wisconsin, Center for Quality and Productivity Improvement.

Phadke, M. S. (1982), "Quality Engineering Using Design of Experiments," in *Proceedings of the Section on Statistical Education, American Statistical Association*, pp. 11–20.

Pierce, J. R. (1980), *An Introduction to Information Theory: Symbols, Signals, and Noise* (2nd ed.), New York: Dover Publications.
Pierce, J. R., and Posner, E. C. (1980), *An Introduction to Communication Science and Systems*, New York: Plenum Press.
Taguchi, G. (1976), *Experimental Design* (Vol. 1), Tokyo: Maruzen (in Japanese).
——— (1977), *Experimental Design* (Vol. 2), Tokyo: Maruzen (in Japanese).
Taguchi, G., and Phadke, M. S. (1984), "Quality Engineering Through Design Optimization," in *Proceedings of GLOBECOM 84 Meeting*, Piscataway, NJ: IEEE Communication Society, pp. 1106–1113.
Taguchi, G., and Wu, Y. (1980), *Introduction to Off-Line Quality Control*, Nagoya, Japan: Central Japan Quality Control Association (available from the American Supplier Institute, 32100 Detroit Industrial Expressway, Romulus, MI 48174).

TECHNOMETRICS, AUGUST 1987, VOL. 29, NO. 3

Discussion

Thomas J. Boardman

Department of Statistics
Colorado State University
Fort Collins, CO 80523

The contributions to quality improvement made by G. Taguchi have been extensive. His separation of the design process into three stages (system, parameter, and tolerance designs) forces engineers and scientists to consider carefully the impact of the various factors contributing to variation in performance as observed by the customer. As I have heard Tom Barker at the Rochester Institute of Technology say, "We are learning to squeeze and then shift" the process. The concept of narrowing variation in the parameter design stage and then shifting the process in the tolerance design stage is a major contribution of Taguchi.

The article under discussion deals with one particular aspect of the Taguchi parameter design process—the use of SN ratios. Instead of considering one or more dependent variables separately, Taguchi uses certain SN ratios. Statisticians (e.g., Box 1986) and others have criticized Taguchi for the use of these ratios. The authors of this article have attempted to help us understand the use of these SN ratios by showing that under certain assumptions concerning the assumed loss function it is possible to use SN ratios that, when maximized, will also lead to results with minimum squared-error loss. Furthermore, they show that under certain assumptions and models one may be able to accomplish the process in a two-step procedure, first by choosing parameters to narrow the variation (reduce noise) and then by selecting other independent parameters that can be used to adjust the process to further reduce the expected loss.

One real concern that has been expressed by others as well as me is whether it is possible in real situations to find adjustment factors that are truly independent of the other parameters. Then too if the performance measures that we select have the analytical property of being independent of adjustment, do these measures correctly reflect the true state of the situation for the process or are they merely convenient metrics that satisfy the chosen criteria? I caution those implementing Taguchi's procedures to study and understand the relationship between the parameters in **d** and those in the adjustment set **a**. Planning experiments to study the interrelationship between the **d** and **a** set seems appropriate rather than assuming away any possible dependence. Thus the conclusion at the end of Section 6.3 on how the use of PerMIA's can simplify design-stage experimentation should be viewed with caution. Until the improvement team fully understands the implications of ignoring cross-terms between the parameters in the **d** and **a** set, it should proceed with caution. The authors do point out the need for knowledge of the form of the transfer function. The team should select a function not for convenience but because it appears to model the situation adequately.

All of the examples that are included in the article include only one parameter in the **a** set. It would have been helpful to include a more general example with several adjustment parameters. Perhaps they could direct readers to one or more examples in the case studies published by Taguchi's American Supplier Institute.

The major contributions of this article, I believe, are that the authors have (a) helped us to further our understanding of Taguchi's procedures, (b) shown us the relationship between the use of SN ratios and minimizing squared-error loss, and (c) given several specific examples in which PerMIA's work. I greatly appreciate their continued contribution to this important interface between engineering, science, and statistics.

Discussion

George Box and Conrad A. Fung

Center for Quality and Productivity Improvement
University of Wisconsin
610 Walnut Street
Madison, WI 53705

The authors are to be congratulated on this very interesting article. In our discussion, we will try to clarify the case in which the desired PerMIA is a measure of *variation* that is unaffected by changes in location. In this situation Taguchi proposes to use for this measure of variation not σ_y but

$$SN_T = 10 \log_{10}(\mu_y^2/\sigma_y^2).$$

If it happens to be true that the location measure μ_y can be set by changing process "adjustment" factors that do *not* affect SN_T, then it will be a simple matter first to minimize process variation by changing factors that *do* affect SN_T and then to set the process mean on target by using the adjustment factors.

Now SN_T is a function only of σ_y/μ_y, the coefficient of variation of y, which is, as is well known, nearly proportional to $\sigma_{\log y}$. Thus the circumstances in which the procedure discussed previously will work are precisely those in which, after a log transformation of the response y, the desired separation of the measures μ_y and $\sigma_{\log y}$ (or of $\mu_{\log y}$ and $\sigma_{\log y}$) will occur. The assumptions underlying all of this are, therefore, that after a log transformation of the data the standard deviation will be independent of the mean, and that it is then more likely that the design factors will separate into a few that affect variation and some others that affect location without changing variation.

Let us suppose that a desirable simplification of this kind *does* occur after some transformation of the data such as the log transformation. Then it is easy to see that unnecessary complication will occur if we do *not* transform or in fact if we use any other metric. This is because in other metrics the standard deviation will depend on the mean, and factors that affect location will produce dispersion effects simply because of this dependence.

This discussion raises a number of important issues.

1. If we are to look at dispersion as well as location, we must very carefully choose the metric in which we analyze our data. There does not seem to be any particular reason why greatest simplicity will always occur in the log. Wider possibilities should also be considered, including the possibility that no transformation is needed.

2. After we have made an appropriate transformation $Y = f(y)$, which gives maximum separation and simplification in terms of location and dispersion effects, our final analysis for dispersion should be based on log s_Y (see, e.g., Bartlett and Kendall 1946). This is because s_Y itself has a standard deviation, which is proportional to its mean. In the particular case in which the desirable transformation is $Y =$ log y, then log s_Y is approximately a linear function of $SN_T = 10 \log_{10}(\bar{y}^2/s_y^2)$, which is the PerMIA recommended by Taguchi.

3. Although the maximization of the performance measure μ_y/σ_y is essentially equivalent to minimization of $\sigma_{\log y}$, new questions arise as soon as we consider matters of *estimation*. If the log is the transformation in which the standard assumptions of constancy of variance, normality of errors, and so forth most nearly apply, then it would be best to conduct our analysis of dispersion in terms of $s_{\log y}$, which is a sufficient statistic, rather than in terms of SN_T, the use of which can be accompanied by considerable loss of efficiency. The ramifications of these arguments were taken up in more detail by Box (1986) and Box and Ramírez (1986).

ACKNOWLEDGMENT

This discussion was supported by National Science Foundation Grant DMS-8420968.

ADDITIONAL REFERENCES

Bartlett, M. S., and Kendall, D. G. (1946), "The Statistical Analysis of Variance—Heterogeneity and the Logarithmic Transformation," *Journal of the Royal Statistical Society*, Ser. B, 8, 128–150.

Box, G. E. P., and Ramírez, J. G. (1986), "Studies in Quality Improvement: Signal to Noise Ratios, Performance Criteria, and Statistical Analysis: Part II," Report 12, University of Wisconsin, Center for Quality and Productivity Improvement.

TECHNOMETRICS, AUGUST 1987, VOL. 29, NO. 3

Discussion

Robert G. Easterling

Sandia National Laboratories
Albuquerque, NM 87185

1. INTRODUCTION

Ideas and methods advanced by Genichi Taguchi of Japan continue to stimulate much attention among American statisticians and others interested in enhancing the quality of products and processes. Among these ideas is that improving quality requires attention to variability of performance, as well as to average performance. Although some might dispute claims as to the origins of this notion, the fact remains that conventional statistical methodology, the type taught in engineering statistics courses, focuses on means. Unequal variances are only a nuisance that complicates a t test in introductory courses or something to be homogenized away by transformations in later courses. The fact that unequal variances can be conveying vital information about different versions of a process does not receive the attention it should.

Two manifestations of attention to variation in the Taguchi lore are the squared-error loss function as a measure of quality and a panoply of performance measures termed SN ratios that generally combine average performance and performance variability. Heretofore, there has been no explicit connection between these two treatments of variation, but now Ramón León, Anne Shoemaker, and Raghu Kacker have forged a connection. Sometimes, choosing some of the design parameters so as to minimize expected squared-error loss is equivalent to choosing those parameters so as to maximize a particular SN ratio. As will be shown in this discussion, however, such coincidence does not mean complete agreement between the authors' optimum design parameter settings and those arrived at in Taguchi's approach. Statistical matters such as estimation also point to differences between the two approaches, so a second purpose of this discussion will be to go beyond the authors' mathematical analysis and consider statistical issues.

2. THE PROBLEM OF STATIC DESIGN

I will address the problem of static design (Sec. 2 of the article) and try to illuminate the difference between the Taguchi approach and that of León, Shoemaker, and Kacker (LSK). The problem setup is as follows: There are a vector, θ, of design parameters and a vector, N, of noise variables that combine to produce a response, Y. The desired response is a target, t. Two parameters that reflect the effect of the design parameters on performance are $E_N(Y) = \mu(\theta)$ and $\text{var}_N(Y) = \sigma^2(\theta)$. Furthermore, a performance measure of apparent interest is the reciprocal of the coefficient of variation squared, $\text{SN}(\theta) = \mu^2/\sigma^2$. [I will keep the typography simpler by not taking the logarithm of $\text{SN}(\theta)$.] The design problem is to choose θ appropriately.

Obviously, this somewhat vague problem statement doesn't lead to a well-defined solution. There is room to improvise, which Taguchi has done and which LSK wish to formalize. Taguchi's solution (as I understand it) is:

1. Assume that θ can be partitioned: $\theta = (d, a)$, such that $\text{SN}(\theta) = \text{SN}(d)$; that is, SN depends only on a subset, d, of the design parameters.
2. Then, for arbitrary a, find d^* such that $\max_d \text{SN}(d) = \text{SN}(d^*)$.
3. Finally, adjust a to a^* such that—? LSK do not say what Taguchi would do, but I infer that "adjusting to target" means that a^* is selected to satisfy the unbiasedness criterion, $\mu(d^*, a^*) = t$.
4. The vector $\theta^* = (d^*, a^*)$ is the solution—the preferred, if not necessarily optimum by some explicit criterion, settings of the design parameters.

Obviously, some arbitrariness remains: How should θ be partitioned? How did SN get into the problem? Why unbiasedness? LSK seek to bring order out of chaos by the following approach:

1. Introduce the loss function, $L(Y, t) = (Y - t)^2$, and its expectation, $R(\theta) = E_N[L(Y, t)]$. The authors, explicit objective is to find θ^* that minimizes $R(\theta)$. This solution is derived in the following three steps.
2. Separate θ into (d, a). (This separation is arbitrary; any partition will work, but some may be more tractable to work with than others.)
3. For each d, find a^*, which is the value of a that minimizes $R(d, a)$. The result is a function, $a^*(d)$.
4. Find d^*, which is the value of d that minimizes $P(d) = R(d, a^*(d))$. Then the solution, which does in fact minimize $R(d, a)$, is thus $\theta^* = [d^*, a^*(d^*)]$.

The preceding steps 3 and 4 are what LSK actually do to solve the static design problem. What they say is somewhat different:

3′. Find d^*, which is the value of d that minimizes $P(d) = \min_a R(d, a)$.

4′. Find a^*, which minimizes $R(d^*, a)$.

LSK are seeking to emphasize the similarity of their development to Taguchi's: He starts with an SN ratio; they will start with a PerMIA. But actually one cannot start with a PerMIA; one has to derive one. Moreover, the final step consists only of evaluating the function $a^*(d)$ at $d = d^*$, not another minimization. This may be just a quibble, but still I think it helps to illustrate the difference between the two solutions.

2.1 Comparison of the Taguchi and LSK Solutions

Both solutions to the static design problem involve partitioning the design-parameter vector θ into two subsets, (d, a). Taguchi optimizes d first, then a. LSK optimize a first, then d. The partitioning of parameters need not be the same, and the optimization criteria are decidedly not the same, but in some cases, either via insight or serendipity, the LSK-derived PerMIA $P(d)$ turns out to be equivalent to the Taguchi-selected SN(d), in which happy case $d_T^* = d_{LSK}^*$, the subscripts denoting the two solutions. This occasional coincidence is the theme of the article. Although such a coincidence is nice, apparently the only way to know when it will occur is to work out the two solutions.

Even when the two approaches yield the same solution for d^*, however, there is no guarantee that they will yield the same a^*. The following example illustrates this point.

2.2 Example

As a simple example, similar to LSK's in Section 3, consider a situation in which the design parameters d and a are scalars and the response is given by $Y = a\beta_d z_d$, where z_d is the random noise variable, with $E(z_d) = 1$ and $\text{var}(z_d) = \sigma_d^2$. In this case, $\text{SN}(\theta) = \text{SN}(d) = 1/\sigma_d^2$. Thus the Taguchi solution is to choose d^* as that value of d that maximizes SN, which is the value of d that minimizes σ_d^2. The other parameter setting would then be $a_T^* = t/\beta_{d*}$ in order to achieve unbiasedness.

The LSK solution would begin with the expected loss function,

$$R(d, a) = a^2\beta_d^2\sigma_d^2 + (\alpha\beta_d - t)^2.$$

Minimizing this function with respect to a, holding d fixed, leads to $a^*(d) = t/\beta_d(1 + \sigma_d^2)$. Substituting $a^*(d)$ into $R(d, a)$ leads to the PerMIA,

$$P(d) = R(d, a^*(d)) = t^2\sigma_d^2/(1 + \sigma_d^2).$$

Clearly, this is minimized by minimizing σ_d^2, so the Taguchi and LSK solutions for d^* coincide in this problem. The LSK solution for a^*, however, is

$$a_{LSK}^* = a^*(d^*) = t/\beta_{d*}(1 + \sigma_{d*}^2).$$

This value of a^* may not be appreciably different from the Taguchi solution, but still they are unequal.

3. STATISTICS

The entire LSK analysis is based on known transfer functions and parameters. Although such assumptions may illuminate the Taguchi notion of an SN ratio, it still seems appropriate, at least in this journal, to address statistical issues. If those functions and parameters are not known, then what data should be collected (experimental design) and what function of them (estimation) should be used to set the design parameters? In the article, these topics are barely mentioned. Statistics, that noble profession, is summarily described as "Empirical studies [that] can check or supplement engineering knowledge" (p. 262). So let us introduce statistical matters in the context of the preceding example.

Suppose that the form of the transfer function $Y = a\beta_d z_d$ is known, but the parameters β_d and σ_d^2, for a finite set of designs, indexed by $d = 1, 2, \ldots, M$, are not known. Suppose that the "empirical study" consists of setting $a = a_0$, arbitrarily, then for each design, d, obtaining a sample of n observations, $Y_1, Y_2, \ldots, Y_n$. From these data, one way to estimate the unknown parameters is

$$\tilde{\sigma}_d^2 = s_d^2/\bar{Y}_d^2, \qquad \tilde{\beta}_d = \bar{Y}_d/a_0,$$

where $\bar{Y}_d$ and s_d^2 are the sample mean and variance of the n observations. (Alternatively, σ_d^2 might be estimated by the variance of ln Y.) Substituting these estimates into the LSK solution for a^* yields $\tilde{a}_{LSK}^* = t/\tilde{\beta}_d(1 + \tilde{\sigma}_d^2)$, but the Taguchi solution would be $\tilde{a}_T^* = t/\tilde{\beta}_d$. Thus the LSK solution, treated statistically, stochastically shrinks the corresponding Taguchi solution, which itself involves the reciprocal of a mean, toward 0. It is with some trepidation that I open the doors of this and other journals to a flood of Taguchi-inspired ridge estimators, but there it is.

Comparison of Estimators

Even though when the parameters are known the LSK solution for a^* has smaller expected loss than the Taguchi solution, it does not follow that when the parameters are estimated and estimation error is incorporated a_{LSK}^* will have a smaller risk than a_T^*. To investigate this matter, consider the case of a fixed d. (The error incurred by selecting d^* based on estimates of σ_d^2 would be common to both ap-

Table 1. Relative Risks of Using an Estimator of a^* Versus Using a^*_{LSK}

Sample Size	Estimator	γ = logarithmic SD .1	.3	.5
5	LSK	1.20	1.22	1.39
		(.03)	(.03)	(.05)
	Taguchi	1.20	1.38	1.67
		(.03)	(.05)	(.10)
	Log	1.23	1.51	2.26
10	LSK	1.11	1.14	1.19
		(.01)	(.02)	(.05)
	Taguchi	1.13	1.19	1.45
		(.02)	(.02)	(.06)
	Log	1.12	1.37	1.98
	$L(a^*_{LSK})$	.010	.086	.221
	$L(a^*_T)$	.010	.094	.284

NOTE: Standard errors are in parentheses. The last two lines give the expected losses for a^*_{LSK} and a^*_T. SD represents standard deviation.

proaches.) The loss associated with using a particular estimate $\tilde{a}$ is given by

$$L(\tilde{a}) = E_z[(Y - t)^2 \mid \tilde{a}]$$
$$= \tilde{a}\beta^2\sigma^2 + (\tilde{a}\beta - t)^2.$$

In general, one might seek to solve the estimation problem by finding the estimator $\tilde{a}^*$ that minimizes $R = E[L(\tilde{a}^*)]$, where the expectation is with respect to the distribution of $\tilde{a}^*$. The objective here, however, is only to use R to compare the Taguchi and LSK estimators.

Suppose for the sake of simplicity, but without loss of generality, we take $a_0 = 1$, $\beta = 1$, and $t = 1$. Suppose the noise variable z has the property that log $z \sim N(-\gamma^2/2, \gamma^2)$. Then $E(z) = 1$ and $\sigma^2 = \text{var}(z) = e^{\gamma^2} - 1$. Three candidate estimators will be considered:

Taguchi: $\tilde{a}_1 = \tilde{a}^*_T = 1/\bar{Y}$.
LSK: $\tilde{a}_2 = \tilde{a}^*_{LSK} = 1/\bar{Y}(1 + s^2/\bar{Y}^2)$.
log: $\tilde{a}_3 = \exp(-\bar{X})$, where $X_i = \ln Y_i$.

This last estimator is based on the common practice of linearizing a multiplicative model by taking logarithms. Another reason for its inclusion is that $R_3 = E[L(\tilde{a}_3)]$ can be evaluated analytically, but R_1 and R_2 will have to be estimated by simulation. Table 1 gives the results of 100 simulation runs at three values of γ, the standard deviation of log z, and two values of the sample size, $n = 5$ and 10. The values tabulated are the (estimated) relative risks, defined as the ratio of R_i to the expected loss achieved by a^*_{LSK} when the parameters are known. The results in Table 1 indicate that in this problem the LSK estimator outperforms the Taguchi-based estimator and both outperform the somewhat naive log-based estimator, which is distorted by the fact that the logarithmic error has nonzero expectation.

In addition to providing a comparison of the Taguchi-based LSK-based estimators, the preceding analysis also points the way toward how experimental design questions might be addressed. How much additional risk caused by estimation error can be tolerated: 10%? 25%, ... ? With some preliminary information about γ and a specified tolerable risk, the required sample size could be determined. Moreover, in many applications it might be prudent to consider collecting data at different a values in order to check the assumptions of a transfer function that is linear through the origin, multiplicative error, and error independent of a.

4. OTHER COMMENTS

Some of the authors' examples puzzle me—for example, the knobs on a stereo or the height adjustment of a vacuum cleaner. These are adjustments that the user makes, not a subset of the design parameters. The designer has to design the knobs or adjustment mechanisms, and these would have parameters that the designer would have to specify. This seems trivial, however, compared with the problem of designing the stereo circuitry or the vacuum motor. How do we separate the d's and the a's in these problems? In another example, I am similarly not enlightened by finding out that to make different-sized tiles one uses different-sized molds.

Any parameter can play the role of an adjustment parameter a because, as the authors point out, the two-step optimization can, in principle, always be carried out. Some choices just make the analysis more tractable, but one can always get an intermediate function that is independent of a. This hardly constitutes a "general principle for choosing performance measures in parameter design." How then should a be picked to make the analysis most tractable? The answer, it appears, is to let a be those parameters that do not influence the error term. For a completely specified model, the choice is obvious; lacking that, the nature of the problem may suggest the choice, or one may have to fall back on "empirical studies," in which case I think there is much work to be done.

ACKNOWLEDGMENT

The authors are to be commended for providing a mathematical basis for a portion of the Taguchi ensemble of ideas and methods. I appreciate the opportunity to add my commentary and hope that the cause of intelligent use of statistics to improve process and product designs has been advanced.

Discussion

Khosrow Dehnad

Quality Assurance Center
AT&T Bell Laboratories
Holmdel, NJ 07733

1. INTRODUCTION

León, Shoemaker, and Kacker's (LSK) explanation and extension of Taguchi's SN ratio is interesting. I am in agreement with them that the use of Taguchi's SN ratio in parameter design results in the minimization of the expected loss only if certain assumptions about the loss function and the underlying model of the product or process hold. Here I will discuss the following issues concerning the concept of PerMIA, which was introduced by LSK as an extension of SN ratios:

1. need for a priori knowledge of a model
2. need for an operational procedure to derive a PerMIA
3. dependence of a PerMIA on the desired output of the product or process (system).

2. NEED FOR A PRIORI KNOWLEDGE OF A MODEL

A PerMIA by its very definition depends on the loss function and the model that describes the working of the system, so to use a PerMIA we should have some knowledge of these two items. What should be the extent of this knowledge? If the loss function and the model are completely known, then the settings of the design parameters that minimize the expected loss can more easily be determined through standard analytic or computational methods, and there would be no need for experimentation or a PerMIA. On the other hand, if the model and the loss function are merely hypotheses to be validated first, then through the course of validation and by observing the system under various configurations, we probably gather enough information to optimize the system without use of a PerMIA.

The remaining situation is when the loss function and the general form of a model are known but contain unknown parameters. In such cases, the concept of a PerMIA can be used to an advantage, and the examples in LSK are all of this kind. For example, in Section 4.1 the dial reading y of an instrument is supposed to satisfy

$$y = \alpha(d, a_1) + \beta(d, a_2)(\gamma(d)s + \varepsilon(N; d)).$$

The exact expressions for α, β, and ε are not known, however. By the way, in this example, since γ is not uniquely estimable, it would be simpler to absorb it into β and write

$$y = \alpha(d, a_1) + \lambda(d, a_2)[s + \xi(N; d)].$$

In short, if our knowledge of a system is within a certain range, then we can take advantage of a PerMIA; otherwise using such a measure is either unnecessary or inappropriate.

3. NEED FOR AN OPERATIONAL PROCEDURE TO DERIVE A PERMIA

Clearly, in many parameter design problems a PerMIA is not known beforehand. Then the question is how to derive this measure in such cases. As mentioned previously, if we know the model and the loss function completely, then finding a PerMIA is basically an exercise in calculus; that is, we have to manipulate the expression for the expected loss and derive this measure. Otherwise, we use engineering knowledge and the available information to find a PerMIA. In the absence of such knowledge, the task of deriving a PerMIA could be a difficult one even in cases in which the loss function and the existence of such a measure are already known. For illustration, consider a system with design parameters a and d and an output having

$$\text{mean} = \mu(d, a) = ad\, g_1(d, a) + a + d$$

and

$$\text{variance} = a^2d^2g_2^2(d, a) + 1,$$

where g_1 and g_2 are finite and real-valued functions of design parameters a and d. Let the loss due to deviation of the output y from its target y_0 be $L(y) = (y - y_0)^2$. It is easy to check that $(a_2 - y_0)^2$ and $(d - y_0)^2$ are PerMIA's because the expected loss is

$$R(d, a) = [adg_1(d, a) + a + d - y_0]^2 + a^2d^2g_2^2(d, a) + 1 \qquad (1)$$

and

$$\begin{aligned} P_1(a) &= \min_d R(d, a) \\ &= \min_d dG_1(d, a) + 1 + (a - y_0)^2 \\ &= 1 + (a - y_0)^2, \\ P_2(d) &= \min_a R(d, a) \\ &= \min_a aG_2(d, a) + 1 + (d - y_0)^2 \\ &= 1 + (d - y_0)^2. \end{aligned}$$

The exact expressions for G_1 and G_2 can be easily derived from expression (1).

Now suppose we are only given the loss function and the fact that a PerMIA exists. It is unlikely that we could derive such a measure solely based on this information and some data, particularly if g_1 and g_2 were complicated expressions. Note that approximations to the mean and variance will not necessarily be of the form that would reveal the expression for a PerMIA. Therefore, some practical rules should be developed for deriving a PerMIA when it is not known beforehand. Otherwise, the application of such measures will be limited to cases in which a PerMIA can either be derived analytically or is already known, say, based on our assumptions of a model and a loss function.

4. DEPENDENCE ON THE RANGE OF RESPONSE

The existence of a PerMIA is not only a property of the system (i.e., the structural model of the system). It depends on the loss function too. In addition, a PerMIA could also depend on the desired output of the system. It could happen, for example in a power supply, that the resistance of a certain resistor is an adjustment parameter with a corresponding PerMIA if the desired output voltage is less than, say, 130 volts, and otherwise it is not. To illustrate this point, suppose the output of a system has

$$\text{mean} = \mu(d, a) = ad$$

and

$$\text{variance} = (d - 1)^2 + 1,$$

with $0 \leq a \leq 1$ because of technological limitations. If the loss due to deviation of y, the output, from its target value y_0 is $(y - y_0)^2$, then the expected loss is

$$R(d, a) = (ad - y_0)^2 + (d - 1)^2 + 1.$$

For this system $(d - 1)^2$ is PerMIA only if $0 \leq y_0/d \leq 1$. To see this note

$$\begin{aligned} P(d) = \min_a &= (d - 1)^2 + 1 \quad \text{if } 0 \leq y_0/d \leq 1 \\ &= (d - 1)^2 + 1 + [d - y_0]^2 \quad \text{otherwise.} \end{aligned}$$

Therefore, before using $(d - 1)^2$ as a PerMIA we should determine if it is such a measure of performance over the range of response that is of interest to us. Such restrictions could complicate the application of a PerMIA in practice.

5. ADDITIONAL COMMENTS

LSK demonstrate through examples how the use of a PerMIA and the corresponding two-step optimization minimize the expected loss. This indeed should be so by the very definition of a PerMIA. Recall that if $R(d, a)$ is the expected loss from a system with design parameters a and d, then

$$P(d) = \min_a R(d, a) \tag{2}$$

is a PerMIA, and minimizing $P(d)$ is the same as minimizing the expected loss. Moreover, the only property that the adjustment parameter a has is that expression (2) can be evaluated for any given d. This means that any subset of design parameters can serve as an adjustment parameter(s) as long as we can evaluate (2). For example, if $a = (a_1, a_2)$ and $d = (d_1, d_2)$, then (a_1, d_2) can be taken as adjustment parameters provided we can compute

$$P(d_1, a_2) = \min_{a_1, d_2} R(d_1, d_2, a_1, a_2).$$

It is not clear to me that a PerMIA and an adjustment parameter defined in this way should necessarily have all of the properties that LSK mention in Sections 6.2 and 6.3 unless we make this part of their definition. In this sense, a PerMIA might not necessarily have all of the practical properties that are usually associated with Taguchi's SN ratios. For example, in practice, to verify that an SN ratio does not depend on the adjustment parameter, we simply evaluate this ratio for various settings of this parameter and see if we come up with roughly the same numbers. By the way, this is how we discover or confirm that a parameter can be used for adjustment. On the other hand, it is not clear how a PerMIA [e.g., $P(d) = \min_a R(d, a)$] can be evaluated for various levels of the adjustment parameter a.

Finally, two-step optimizations are not unknown in other fields. For example, in source coding there is a two-step procedure for minimizing the degradation of the signal based on a distinction between distortion and uncorrelated additive noise (Jayant and Noll 1984). Still closer to home, in statistics to find the maximum likelihood estimate of the parameters of a normal distribution we first estimate the mean (independent of the variance); next, we adjust the estimate of the variance to maximize the likelihood. In fact, this can be stated in terms of the following separability condition that uses the notion of a PerMIA. This result was also observed and elaborated on by C. F. J. Wu (1986, private communication with A. Shoemaker).

Let $R(d, a)$ be the expected loss for a system. A sufficient condition for the existence of a performance measure independent of the adjustment parameter a is that for a monotone increasing function g the expected loss satisfies the separability condition

$$g[R(d, a)] = R_1(d, a) + R_2(d)$$

such that for any given d

$$\min_a R_1(d, a) = c,$$

where c does not depend on d. In this case R_2 is a PerMIA, and to minimize the expected loss we first determine d so that R_2 is minimized while a is set at an arbitrary level. Next we adjust a to achieve the minimum of R_1. In particular, if

$$P(d) = \min_a R(a, d)$$

is known, we trivially have a PerMIA. Simply define

$$R_1(d, a) = R(d, a) - P(d)$$

$$R_2(d) = P(d),$$

and note that

$$\min_a R_1(d, a) = \min_a R(d, a) - P(d) = 0.$$

V. Nair has pointed out that the process of finding the maximum likelihood estimator of the parameters of a gamma distribution can be formulated in the preceding terms.

6. CONCLUSION

The use of Taguchi's SN ratio in parameter design will result in the minimization of the expected loss if certain assumptions about the model and the loss function hold. The concept of a PerMIA is useful to simplify the optimization procedure in parameter design if we either know such a measure or have enough information to derive it. Otherwise, the problem of determining a PerMIA could be too complicated. Further, there is a need for an operational procedure to verify whether a PerMIA exists and to derive it if it does exist.

ADDITIONAL REFERENCE

Jayant, N. S., and Noll, P. (1984), *Digital Coding of Wave Form*, Englewood Cliffs, NJ: Prentice-Hall.

Discussion

Joseph J. Pignatiello, Jr., and John S. Ramberg

Systems and Industrial Engineering
University of Arizona
Tucson, AZ 85721

Fewer than five years ago, very little literature on the product and process design methods proposed by Genichi Taguchi was available. His methods, surrounded by a mystique that few could unveil, had not been critiqued, because very few people understood them. His approach was expressed in the terms of communications theory rather than those of statistics. Consequently, some of the communications terminology further added to the confusion.

Since that time, the methods of Taguchi have received considerable attention, no doubt because of, the timing of his publications and his obvious success and the success of his many followers. The role of the mystique and the controversy surrounding these methods should also not be discounted.

Kackar (1985) and Hunter (1985) made important contributions to the understanding of Taguchi's methods. Both successfully translated the Taguchi methods into a language that is understandable by professional quality-control statisticians.

Kackar's (1985) introduction to the methods of Taguchi presented the philosophy, along with the requisite terminology of "loss functions," "design and noise matrices," "orthogonal arrays," and "performance statistics" that Taguchi calls "SN ratios." Kackar's major contribution was the removal of much of the mystique that had previously cloaked the Taguchi methods, allowing insight into the philosophy behind them.

The work of Hunter (1985) was also an important contribution, showing the many problems that can arise when one swallows the Taguchi methods hook, line, and sinker. He directed his comments specifically to problems that arise from the misuse of orthogonal-array experimental designs, such as some three-level designs and some two-level designs that are not (fractional) factorial designs.

León, Shoemaker, and Kacker provide yet another contribution to the Taguchi literature. We thank the editors for giving us the opportunity to comment on this article.

The major contribution of the article is the development of a statistical decision framework that justifies the use of Taguchi's SN ratios under certain assumptions. This framework allows one to determine when the SN ratios are *not* the appropriate performance measures to be used. The authors are to be congratulated for their efforts.

This manuscript prompts us to consider two very different aspects of the Taguchi methods: *strategy* and *tactics*.

Taguchi's *strategy*—that is, his problem definition and conceptual framework for designing and manufacturing a quality product or process—is clearly his most important contribution. It provides a road map for practitioners that has been successful in getting them to try designed experiments as illustrated by the extensive collection of case studies reported by the American Supplier Institute (1984).

An integral part of the Taguchi strategy is to adopt a loss function that relates the loss that is incurred by society when the product's ultimate consumer uses a product whose quality characteristic deviates from the target value. We view this as Taguchi's way of quantifying J. Juran's definition of "quality" (i.e., "fitness for use"). Both Taguchi's definition of quality ("the loss imparted to society from the time the product is shipped") and J. Juran's definition differ significantly from P. Crosby's definition of quality ("conformance to specifications").

The Taguchi *tactics* involve the use of the orthogonal-array designs and the SN ratios, among others. León, Shoemaker, and Kacker have focused their attention on certain types of the SN ratio performance measures. Considering alternative tactics is quite important, because one can implement the Taguchi strategy while using tactics that are modifications of Taguchi's or even quite different from those of Taguchi. As León, Shoemaker, and Kacker have pointed out in their article, there are circumstances in which one should *not* use the Taguchi SN ratios.

The product-design problem (and the process-design problem) involves statistical decision-making under uncertainty. A statistical decision problem can be thought of as a triplet (Θ, D, l) and a random variable Y, where $Y \in \chi$. The random variable Y has some distribution function $f_Y(y \mid \theta)$, where all that is known about θ is that $\theta \in \Theta$. The set $D = \{\mathbf{x}\}$ is the decision space. For each combination of θ and $\mathbf{x}$, the loss function l assigns the real number $l(\theta, \mathbf{x})$. A "decision rule," $s(Y)$, is a function of Y for choosing $\mathbf{x}$ from D; that is, after observing the data Y, the decision $\mathbf{x} = s(Y)$ is chosen from D. The problem is to find the value of $\mathbf{x}$ that minimizes the expected loss (or "risk"):

$$\mathbf{E}_\theta\, l(\theta, s(Y)).$$

In the context of a Taguchi product-design situation, the random variable Y represents the quality characteristic of the product. The *noise space* is represented by Θ, whereas θ represents a point in the noise space. The decision space $D = \{\mathbf{x}\}$ is the set of all possible product (or process) designs in which each product design (or process design) is a combination of the values of the k controllable design variables; that is, $\mathbf{x} = (x_1, \ldots, x_k)$. For each product (or process) design $\mathbf{x}$, there is a distribution of the loss, and θ may be a function of $\mathbf{x}$. Thus to select the "best" product (or process) design, say $\mathbf{x}^*$, one needs to find $\mathbf{x}^*$ such that

$$\mathbf{E}_\theta\, l(\theta, \mathbf{x}^*) = \min_{\mathbf{x} \in D} \mathbf{E}_\theta\, l(\theta, x).$$

For a "closer to target is better" situation, Taguchi's decision rule is based on maximizing

$$10 \log_{10}(\mathbf{E}^2[Y(\mathbf{x})]/\mathrm{var}[Y(\mathbf{x})]). \quad (1)$$

León, Shoemaker, and Kacker have pointed out that such a rule is appropriate under certain restrictions; that is, in some situations, Taguchi's decision rule will lead one to the same best product (or process) design as minimizing the expected loss.

To use such a decision rule, however, one would need to know the functional relationship between Y and $\mathbf{x}$. In most real world product (or process)-design situations, the functions and distributions that are involved in minimizing the expected loss are rarely known; that is, one generally does not know a priori the relationship that exists between the mean of the quality characteristic and the controllable variables. It is even more unlikely that the relationship between the variance of the quality characteristic (or the variance of the noise) and the controllable variables is known. Consequently, one resorts to empirical methods such as one involving a systematic application of the design and analysis of experiments.

Since the entire noise space Θ cannot be observed for each possible product design, the *strategy* is to observe the performance of possible product designs over a *sample*, W from Θ. Taguchi recommended the selection of samples from Θ that are orthogonal arrays. Other choices (i.e., other experimental designs for the noise matrix) would simply be another *tactic*.

Similarly, it is usually not possible to test all possible product designs, so a selection is made. Again, Taguchi would select the product designs to be included in the experiment according to an orthogonal array. Once again, better choices could be made.

The Taguchi layout can be displayed as in Figure 1, where $\mathbf{x}_j = (x_{j1}, \ldots, x_{jk})$ denotes the jth product design included in the experiment and x_{jt} denotes the value of the tth controllable design variable in that jth combination. Also in Figure 1, $\boldsymbol{\omega}_i = (\omega_{i1}, \ldots, \omega_{il})$ represents the ith point in the sample from the noise space and y_{ji} represents the observed value of the quality characteristic using product design $\mathbf{x}_j$ under noise conditions $\boldsymbol{\omega}_i$. Each product design $\mathbf{x}_j$ is to be evaluated using a performance measure (which we would *never* call a "PerM"!), $Z(\mathbf{x}_j)$. Presumably, the performance measure that is used should be a "good estimator" of the expected risk (or at least have the property that the performance measure will lead an experimenter to select the same best product design as a "good estimator" would). It is not clear to us that $10 \log_{10} \bar{y}^2(\mathbf{x})/s^2(\mathbf{x})$ is such an estimator of (1).

An alternative to Taguchi's tactic at this point is to fit models relating the performance measure Z to the controllable variables $\mathbf{x}_j$, to explore the response surface, and to determine which product design (possibly one that was not included in the experiment) is best.

We have been successful in using the Taguchi layout (Fig. 1) in several different situations. We will briefly describe one of those situations.

In Pignatiello and Ramberg (1985), we considered a heat-treatment process involving four process-design variables and one noise variable. The combinations of the process variables to be observed in the experiment were selected with a 2^{4-1} fractional-factorial design. The noise variable was observable at two levels, a "low" level and a "high" level. For each combination of the controllable variables in the experiment, the process was observed three times at the "low" level of the noise variable and three times at the "high" level. Thus, for each combination of the process variables in the experiment, the layout for that experiment consisted of three replicates of a two-level, single-factor, full-factorial design.

Not knowing a priori the relationship between the response and the controllable process variables, we considered the relationship between the controllable variables and the (a) mean response, (b) sample variance of the response, and (c) SN ratio. We greatly preferred the joint analyses of (a) and (b) to that of (c). We found that a variable originally treated as noise was very important. Upon reflection, we concluded that it would be more economical to control it, and thus it became a design parameter. We were also fortunate to find that although several of the

					ω_{11}	...	ω_{i1}	...	$\omega_{n_\omega 1}$	
					⋮		⋮		⋮	
					ω_{1s}	...	ω_{is}	...	$\omega_{n_\omega s}$	
					⋮		⋮		⋮	
					ω_{1l}	...	ω_{il}	...	$\omega_{n_\omega l}$	
x_{11}	...	x_{1t}	...	x_{1k}	y_{11}	...	y_{1i}	...	y_{1n_ω}	$Z(x_1)$
⋮		⋮		⋮	⋮		⋮		⋮	⋮
x_{j1}	...	x_{jt}	...	x_{jk}	y_{j1}	...	y_{ji}	...	y_{jn_ω}	$Z(x_j)$
⋮		⋮		⋮	⋮		⋮		⋮	⋮
x_{d1}	...	x_{dt}	...	x_{dk}	y_{d1}	...	y_{di}	...	y_{dn_ω}	$Z(x_d)$

Figure 1. Generic Taguchi Layout.

process variables did influence the process variability, two process variables were not significant in this regard. These two variables were judged significant in their effect on the process mean and thus were adjustment factors. Furthermore, one was appropriate for short-term control, and the other was better for long-term control. On setting the factors affecting process variability at appropriate levels, the latter two were used for adjustment.

León, Shoemaker, and Kacker have addressed an issue related to the choice of a performance measure *when one has a priori knowledge of their process.* They, assume that experimenters already know, for example, which product (or process) variables affect the variability of the noise and which variables are adjustment variables. Furthermore, they assume that an experimenter knows the functional relationship (apart from some constants) between the quality characteristic and these variables including, for example, whether an error term is additive or multiplicative.

Certainly there may be situations (such as the overused Ina tile example) in which an adjustment variable (the mold size) is known to exist before much data have been collected. Perhaps, too, there are situations (such as in Monte Carlo simulations) in which one may indeed know the needed functional relationship. Such extensive prior knowledge is rare in actual manufacturing practice, however, especially for complicated or new products and processes.

The procedure that we discussed involving the heat-treatment process was simple and straightforward. As in most real-world manufacturing environments, we did not know how the controllable process variables were related (if at all) to the quality characteristic. We did not know whether the loss function was quadratic in the original units or quadratic in *log* units. For that matter, we were not aware of any explicit loss function.

Our analysis isolated those process variables that affected the variance of the quality characteristic from those that did not. We did find a process variable that affected the mean of the quality characteristic and did not affect its variance. Our approach was obvious and it made sense to the process engineers: Find the combination of those variance-affecting process variables that minimizes the variance of the quality characteristic and then, with those variables fixed at that combination, use the adjustment variable to bring the process mean on target.

Research is needed on techniques that discriminate between the product (or process) variables that affect the variability only, those that affect the mean only, those that affect both, and those that affect neither. It is not until this important partitioning of the product (or process) variables is made that one can even consider performance measures that are independent of the adjustment variables. Response surfaces need to be estimated to explore the relationships between the process variables and the mean and variance of the quality characteristic. Much work is needed on the shop floor before the choice of performance measures becomes an issue.

We have one negative comment concerning this article. As previously mentioned, Kackar (1985) carefully stated Taguchi's terms in the more meaningful statistical terminology. This article reverts to the original Taguchi terminology, causing unnecessary confusion. Furthermore, the use of the block diagram (León, Shoemaker, and Kacker's Fig. 1) shows a block representing expected loss, whereas a systems engineer would use this block to express *loss* and regard the *expected loss* as one possible performance measure of the system.

ACKNOWLEDGMENT

Pignatiello's work was supported in part by National Science Foundation Grant MEA-8404888.

ADDITIONAL REFERENCES

American Supplier Institute (1984), *Quality Engineering Using Design of Experiments: A Supplier Symposium on Taguchi Methods* (Vols. 1 and 2), Romulus, MI: Center for Taguchi Methods.

Hunter, J. S. (1985), "Statistical Design Applied to Product Design," *Journal of Quality Technology*, 14, 210–221.

Pignatiello, J. J., and Ramberg, J. S. (1985), "Off-Line Quality Control, Parameter Design, and the Taguchi Method, A Discussion," *Journal of Quality Technology*, 14, 198–206.

TECHNOMETRICS, AUGUST 1987, VOL. 29, NO. 3

Discussion

C. F. J. Wu

Center for Quality and Productivity Improvement
University of Wisconsin
610 Walnut Street
Madison, WI 53705

1. INTRODUCTION

I would like to congratulate the authors for their significant contributions and lucid presentation. Assuming the knowledge of a model and a loss function, they develop the concept PerMIA, give several interesting examples, and outline its advantages. They also assume that separation of the adjustment parameters **a** and the nonadjustment parameters **d** can be identified prior to experimentation. I will call the resulting performance measure *model-driven*. As mentioned in their article, the adjustment parameters can in some situations be identified by engineering considerations. Otherwise, it will be determined empirically from the experimental data, in which case the resulting performance measure is *data-driven*.

2. MODEL-DRIVEN PERFORMANCE MEASURES

The authors have made pioneering contributions to this class of measures. Here I will contribute some further results and comment on some aspects of PerMIA.

2.1 An Extension of PerMIA

Several examples of PerMIA are given in the article, but no general method for obtaining it is provided. One such method is stated in the following theorem.

Theorem 1. If $R(d, a) = R_1(d) + R_2(d, a)$ with $\min_a R_2(d, a) = R_2$ independent of d, then $R_1(d)$ is a PerMIA.

The proof is straightforward.

I will show in Section 2.2 that $\min_a R_2(d, a) = R_2$ in Theorem 1 can impose a severe requirement on the adjustment parameters. Noting that the main consequence of a PerMIA is the two-step Procedure 2, we will relax the preceding requirement on $R_2(d, a)$ without affecting the two-step procedure.

Theorem 2. For $R(d, a) = R_1(d) + R_2(d, a)$, define $D^* = \{d^*: R_1(d^*) = \min_d R_1(d)\}$ and let D^c be the complement of D^* in the space of nonadjustment parameters. Assume that, for $d \in D^*$, $\min_a R_2(d, a) = R_2$ is independent of $d \in D^*$, and $R_2(d, a) \geq R_2$ for any $d \in D^c$. Then minimization of $R(d, a)$ over d and a can be done in the following two-step procedure:

1. Find d^* to minimize $R_1(d)$ over d.
2. For d^* in step 1, find a^* to minimize $R_2(d^*, a)$ over a.

The proof is straightforward. Note that $R_1(d)$ in Theorem 2 is not a PerMIA but plays the same role as a PerMIA in the two-step procedure. In this sense, it may be called a *generalized* PerMIA. Examples of generalized PerMIA's will be given in Section 2.2.

2.2 Assumptions on Adjustment

In several examples the authors assume that adjustment can be freely made so that the mean is on target. Take the example of the Section 2.1, in which

$$R(d, a) = (1 + \sigma^2(d))[\mu(d, a) - t/[1 + \sigma^2(d)]]^2 + t^2\sigma^2(d)/[1 + \sigma^2(d)]. \quad (1)$$

An implicit assumption made here is that, for any d, a can be adjusted to a' so that

$$\mu(d, a') = t/[1 + \sigma^2(d)]. \quad (2)$$

It is conceivable that in some practical situations (2) does not hold for all d, in which case $\sigma^2(d)$ will not be a PerMIA. But the two-step procedure in Theorem 2 holds if, for d^* satisfying $\sigma^2(d^*) = \min_d \sigma^2(d)$, there exists a^* with

$$\mu(d^*, a^*) = t/[1 + \sigma^2(d^*)]. \quad (3)$$

This follows from Theorem 2, with $R_2(d, a)$ and $R_1(d)$ being the first and second terms of (1). The requirement (3) on the adjustability of μ is weaker than (2), and $\sigma^2(d)$ is a generalized PerMIA.

Similarly the assumption on the existence of an unbiased control strategy in the example of León, Shoemaker, and Kacker, Section 4.2 needs more careful examination. It is assumed that for any d there exists a control strategy $s(t)$ from (s_L, s_H) such that

$$E(Y \mid d, a, s(t)) = t \quad (4)$$

for any t from (t_L, t_H). Presumably this depends on the particular engineering problem. If (4) does not hold for every d, $\sigma^2(d)$ is not a PerMIA. As a consequence the two-step decomposition in their Section 4.2 does not work if the previously mentioned unbiasedness does not hold for d with small $\sigma^2(d)$. The validity of this two-step procedure can, however, be guaranteed by a much weaker requirement on the control strategy; namely, for the particular pair (d^*, a^*) in the two-step procedure, there exists an $s(t)$ from (S_L, S_H) such that

$$E(Y \mid d^*, a^*, s(t)) = t$$

for any t from (t_L, t_H). This follows from an obvious modification of Theorem 2.

As a last example, consider the unbiasedness constraint in their Section 4.1 that for any d there exist $a_1^*(d)$ and $a_2^*(d)$ such that

$$\alpha(d, a_1^*(d)) = 0, \qquad \beta(d, a_2^*(d))\gamma(d) = 1. \quad (5)$$

The ability to adjust a_1 and a_2 to satisfy (5) depends on the particular engineering problem. If it does not hold for every d, $\text{var}(\varepsilon)/\gamma^2(d)$ is not a PerMIA and the two-step Procedure 4 in Section 4.1 fails. It is possible that a biased design with a smaller β gives a smaller expected loss $E_s E_n(Y - s)^2$ than the best unbiased design. Note that Theorem 2 is not applicable here. It is not clear how the unbiasedness constraint can be relaxed without affecting the validity of the two-step procedure.

Assumptions on adjustability are also made in Step 2 of the two-step Procedure 5 in Section 4.2 and in the choice of the parameters A, B, and G in (A.2) and the following formula. To save space, this will not be further discussed.

Finally, I will point out the difference between two types of constraints. The unbiasedness constraint (5) for the measuring-instrument example can be achieved by adjusting the parameters a_1 and a_2 at the parameter-design stage. The unbiasedness constraint on the control strategy $s(t)$ for the control-system example (Sec. 4.2) and the truck-steering mechanism (App. A) is a problem of system design, which needs to be solved before parameter design. Is the second type of constraint reasonable from the viewpoint of control engineering? Can the authors comment on how it can be achieved in practice?

3. DATA-DRIVEN PERFORMANCE MEASURES

As remarked in my introduction, identification and separation of the adjustment and nonadjustment parameters a and d may not be achievable without the aid of experimental data. The resulting performance measure based on an empirically built model is data-driven. Its implementation depends on the accuracy of the model. Since building an accurate model requires a substantial amount of data, the usefulness of an empirical PerMIA needs a more careful study.

Take, for example, Box's (1986) empirical approach. He started with the following extension of a PerMIA: There exists a monotone transformation on y such that the standard deviation σ_Y of the transformed variable $Y = f(y)$ is independent of its mean μ_Y, and there are two sets of factors (i.e., parameters) x_a and x_d such that the following conditions apply:

Condition 1. σ_Y is a function of x_d only.
Condition 2. μ_Y is a function of x_a and x_d.

A transformation $f: y \rightarrow Y$ is then determined empirically so that an empirical analog of Conditions 1 and 2 is satisfied. A major limitation of this approach is that the assumption on separation (Conditions 1 and 2) is *global*, whereas transformation of data can usually achieve separation only *locally*. This is illustrated by the contour plot of μ_Y and σ_Y in Figure 1. The assumption on separation (Conditions 1 and 2) holds only in a neighborhood of $\sigma_Y = .6$ and $\mu_Y = 20$. As x_a moves out of the small rectangular area, there is no separation. This has an adverse effect on the validity of the two-step procedure. In Figure 1, x_d is first adjusted so that σ_Y drops from 1.0 to .6 (Step 1). For the chosen x_d, x_a is then adjusted from $\mu_Y = 20$ to the target $\mu_Y = 10$. Since the new x_a is outside the region of separation, the value of σ_Y increases from .6 to 1.0 as a result of the change in x_a.

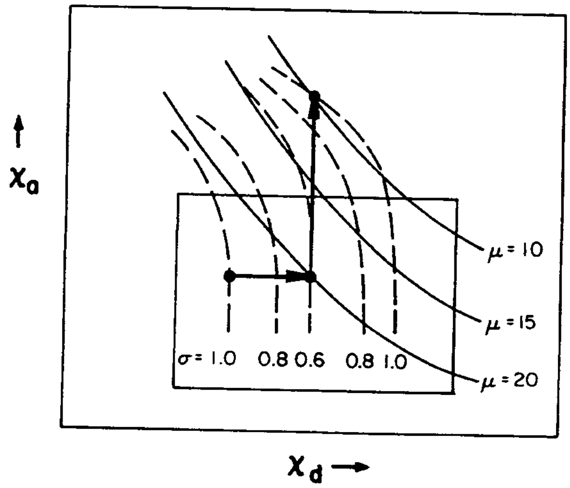

Figure 1. The Contour Plot of μ_Y and σ_Y and Local Separation of x_d and x_a.

To summarize:

1. Data transformation may result in a local separation of x_d and x_a.
2. In this context the proposed two-step procedure is valid only if x_d and x_a are adjusted within the region of separation.

Even if global separation is possible, for fractional designs, empirical determination of x_d and x_a satisfying Conditions 1 and 2 may not be straightforward. Interactions among these factors and their confounding patterns should be carefully handled.

An advertised advantage of PerMIA is the potential saving in sample size of the two-step procedure. For data-driven PerMIA such as those considered by Box (1986), x_d and x_a have to be identified from the data. Since both x_d and x_a are included in the experimentation, there is no saving in the size of the experiment. In the same context, a simple MSE criterion seems to be quite adequate for the stationary target problem. On the other hand, a two-step procedure is possible only after x_d and x_a are separated from the first-round experiment. Additional experiments are needed to carry out the two-step procedure. Is this unnecessarily complicated?

4. CONCLUDING REMARKS

In summary, I think that PerMIA is an important contribution to our understanding of model-driven performance measures. When adjustment parameters can be identified prior to experimentation, the proposed two-step procedure can be effective. On the other hand, if adjustment parameters and PerMIA have to be determined empirically, some difficulties arise (see my Sec. 3). This requires further research before a definitive answer can be given.

I once again thank the authors for their very stimulating article.

ACKNOWLEDGMENT

This research was supported by National Science Foundation Grant DMS-8420968.

Response

Ramón V. León, Anne C. Shoemaker, and Raghu N. Kacker

1. APPROACH FOR DERIVING AND USING PERFORMANCE MEASURES INDEPENDENT OF ADJUSTMENT

Motivated by Robert G. Easterling's comments, we summarize, as a series of steps, the approach we have followed in deriving and using PerMIA's in parameter design problems. These steps will serve as the bases for responding to many of the discussants' comments. In these steps we assume that the object of a parameter design problem is to find the setting of the design parameters that would minimize expected loss. The steps are as follows:

1. Write down as much as you are willing to assume about the model and loss function. (Because Taguchi prefers the squared-error loss function, we usually assume it even though this is not necessary for our approach.)
2. See how far you can go toward minimization of expected loss analytically. As we have shown, in many problems this step consists of deriving a PerMIA, $P(\mathbf{d}) = \min_{\mathbf{a}} R(\mathbf{d}, \mathbf{a})$, which can be done if special design parameters called adjustment parameters are available.

Continue empirically. When a PerMIA is found in Step 2, this consists of the following:

3. Minimize the PerMIA; that is, find $\mathbf{d}^*$ that minimizes $P(\mathbf{d})$.
4. Find $\mathbf{a}^*$ that minimizes $R(\mathbf{d}^*, \mathbf{a})$, which in many problems, but not all, is mathematically equivalent to a procedure that can be interpreted as adjusting the mean.

2. EASTERLING'S COMPARISON OF THE TAGUCHI AND LSK SOLUTION

Judging from Easterling's comments, we did not make clear the connection between the general procedure for using a PerMIA, described in Section 3 of the article, and the mathematical techniques used to find the PerMIA and accomplish the minimizations in particular examples. In light of Easterling's comments, we begin by clarifying the LSK solution. We start with Step 2.

To accomplish Step 2 using a PerMIA, $P(\mathbf{d}) = \min_{\mathbf{a}} R(\mathbf{d}, \mathbf{a})$, we sometimes proceed as follows: First, for a fixed $\mathbf{d}$, we find $\mathbf{a}^*(\mathbf{d})$, the value of $\mathbf{a}$ at which the minimum of $R(\mathbf{d}, \mathbf{a})$ is achieved. Then we evaluate $R(\mathbf{d}, \mathbf{a}^*(\mathbf{d}))$ to obtain $P(\mathbf{d})$. At other times, however, we can find $P(\mathbf{d})$ in a different way. For example, in Section 2 of the article, when the loss function is $L(y, t) = (\log y - \log t)$, we can find $P(\mathbf{d})$ without first finding $\mathbf{a}^*(\mathbf{d})$.

We think that the same confusion also led to Easterling's remark that "the final step consists only of evaluating the function, $\mathbf{a}^*(\mathbf{d})$ at $\mathbf{d} = \mathbf{d}^*$, not another minimization" (p. 268). In a number of problems, evaluating $\mathbf{a}^*(\mathbf{d})$ at $\mathbf{d} = \mathbf{d}^*$ is merely the mathematically most convenient way to carry out the minimization in Step 4. Once this confusion is removed, it will be clear to the reader that we do exactly what we say—that is, Easterling's steps 3′ and 4′.

Easterling also says that the reason we are able to obtain the SN ratios is by "coincidences." Let us examine this point. As we point out in the paragraph following model (2.2), this model is essentially saying that Taguchi's SN ratio, or equivalently the coefficient of variation, depends only on a subset, $\mathbf{d}$, of the design parameters. This is step 1 of Easterling's interpretation of Taguchi's solution; model (2.2) is an explicit statement of what Taguchi is implicitly assuming. It is no coincidence that the LSK procedure leads to the SN ratio under model (2.2).

Easterling also makes the point that Taguchi has an unbiasedness constraint in mind when he "adjusts to target." We agree. But our approach also works if this constraint is added to the minimization of expected loss criterion. For example, under model (2.2), if an unbiasedness constraint is added, $P(\mathbf{d})$ is equal to Taguchi's SN ratio in Step 2. What changes is the result of Step 4, which now fulfills the constraint by adjusting to target.

We chose to do the minimization unconstrained because we want it to be clear that if the object of parameter design is to minimize expected squared error loss, as Taguchi states, then the adjustment cannot be to the target but to something less than the target. We thank Easterling for making us aware that this is connected to the theory of ridge estimators.

Finally, since it bothers Easterling that we have to derive the PerMIA in Step 2, we note that Taguchi (1977) and Taguchi and Phadke (1984) frequently attempted this derivation but with less well-specified "models."

3. HOW MUCH KNOWLEDGE OF THE MODEL AND LOSS FUNCTION IS REALLY ASSUMED?

The comments of several discussants, including Pignatiello and Ramberg, and Dehnad, point to a need for clarification of the degree of knowledge needed to derive a PerMIA in Step 2. For example, examination of models (2.2) and (2.4) in the article reveals that only very general assumptions are being made. The functional forms of the major model components $\mu(\mathbf{d}, a)$ and $\sigma^2(\mathbf{d})$ are completely general. Indeed, these models are no more specialized than those assumed by Taguchi, since the model underlying Taguchi's SN ratio (for static problems) is the same as model (2.2). Box (1986) was more general, but still in the same vein, because he assumed the model

$$Y^\lambda = \mu(\mathbf{d}, \mathbf{a}) + \varepsilon(\mathbf{N}, \mathbf{d})$$

(where $Y^\lambda = \log Y$ when $\lambda = 0$) for some λ. Of course, this family includes model (2.4) and, approximately, model (2.2).

4. COMMENTS ON THE USE OF TRANSFORMATION IN PARAMETER DESIGN

As mentioned previously, the transformation framework referred to by Box and Fung and described in more detail by Box (1986) assumes the following model:

$$Y^\lambda = \mu(\mathbf{d}, \mathbf{a}) + \varepsilon(\mathbf{N}, \mathbf{d}),$$

with Y^λ defined to be log Y when $\lambda = 0$. Once λ has been identified (we briefly describe an empirical way to do this in Sec. 5.2 of this response), the procedure is as follows:

1. Find $\mathbf{d}^*$ that minimizes log var(Y^λ).
2. Use $\mathbf{a}$ to adjust $E(Y)$ [or $E(Y^\lambda)$, Box (1986) does not make it clear which] to the target.

It is important to note that this procedure does not necessarily lead to minimization of the real expected loss function. Instead, this approach leads to minimization of the function $L(Y, t) = (Y^\lambda - t^\lambda)^2$, perhaps subject to an unbiasedness constraint. This approach lets the form of the true model for Y dictate the loss function. This can lead to an unreasonable loss function, for example, when $\lambda = 0$ and t is close to 0.

5. INVESTIGATING THE MODEL AND FINDING AND USING ADJUSTMENT PARAMETERS

Nearly all of the discussants were concerned about the existence, empirical detection, and use of adjustment parameters. We address each of these issues in turn.

5.1 Existence of Adjustment Parameters

As we discussed in Section 6 of the article, candidates for adjustment parameters are often suggested by engineering practice or are deliberately designed into a product or process. In an application of parameter design to improve the process for growing epitaxial layers on silicon wafers (see Kackar and Shoemaker 1986), results verified that the deposition time could be used to adjust the mean layer thickness approximately independently of the thickness variance. In fact, prior to the parameter design study, process engineers had always used deposition time to adjust the process for different target layer thicknesses.

5.2 Empirical Methods for Detecting Adjustment Parameters and Finding PerMIA's

5.2.1 Extending Empirical Methods From SN Ratios to PerMIA's. The same empirical checks used to verify that Taguchi's SN ratio is independent of an adjustment parameter may be applied to check that a candidate PerMIA is independent of a given adjustment parameter. In response to Dehnad's comments on this point, note that if the candidate PerMIA has been derived from a mistaken model, it will indeed be a function of the adjustment parameter, and evaluation of the candidate PerMIA at various adjustment parameter settings will reveal this.

5.2.2 Using the Transformation Framework to Detect Adjustment Parameters. The transformation framework referred to by Box and Fung and described by Box (1986) leads to empirical methods that may be useful for investigating the model and verifying or identifying adjustment parameters. Assume that the model is of the form

$$Y^\lambda = \mu(\mathbf{d}, \mathbf{a}) + \varepsilon(\mathbf{N}, \mathbf{d})$$

(with Y^λ defined to be log Y for $\lambda = 0$) for some λ. Then, for a given value of λ, calculate R^2 statistics for each parameter's effect on estimates of $E(Y^\lambda \mid \mathbf{d}, \mathbf{a})$ and log var$(Y^\lambda \mid \mathbf{d}, \mathbf{a})$. A parameter is a candidate for an adjustment parameter if its R^2 for $E(Y^\lambda \mid \mathbf{d}, \mathbf{a})$ is large and its R^2 for var$(Y^\lambda \mid \mathbf{d}, \mathbf{a})$ is small. This comparison can be repeated for a collection of λ values until an adjustment parameter is identified.

As we saw in Section 4 of this response, however, the transformation approach may not lead to minimization of the real (or even reasonable) expected loss function.

5.3 Advantages of PerMIA's When Empirical Investigation Is Necessary

Wu questions the usefulness of PerMIA's when empirical studies are required to find them. His point is that savings due to simplification of the optimization problem made possible by existence of an adjustment parameter are negated by the effort required to find the PerMIA in the first place.

When an adjustment parameter exists, however, using experiment results to directly minimize expected loss can be incorrect. In parameter design experiments, design parameters are typically varied over wide ranges and the mean response is often off target in many experimental runs. Estimates of expected loss itself calculated from the experiment results do not take into account subsequent tuning of adjustment parameters to bring the mean back to target. The PerMIA, on the other hand, would be independent of this subsequent adjustment parameter shift. Tsui (1987) discusses this point in more detail.

5.4 Danger of Adjusting Out of the Range of the Data

Both Wu and Dehnad point out the danger that adjustment parameters may exist in only a subset of the parameter space. This valuable insight should be kept in mind by experimenters who seek to identify adjustment parameters empirically and use them later to adjust to different targets.

6. NEW RESEARCH DIRECTIONS POINTED OUT BY THE DISCUSSANTS

The discussion has pointed out several valuable areas of research for developing the theory of performance measurement in parameter design. In particular, Wu's discussion of conditions for existence of a PerMIA and his generalized PerMIA could increase the usefulness of these ideas. Both Dehnad and Wu point out interesting parallels between the PerMIA and its associated two-step optimization procedure and other areas of mathematics and science. As mentioned before, Easterling suggests a connection to the theory of ridge estimation.

Although estimation of performance measures was beyond the scope of our article, the discussion points to the need for empirical methods for investigating models, and identifying or verifying adjustment parameters and corresponding PerMIA's. Moreover, studies such as Easterling's comparison of risk associated with different parameter design strategies will be valuable in choosing a useful and powerful procedure.

7. DISAGREEMENTS IN FORMULATION OF PARAMETER DESIGN AND TOLERANCE DESIGN PROBLEMS

Because the area of parameter design is new to the statistical community, the terminology and problem formulation are only now becoming standardized [see Kackar (1985) for a proposed terminology and formulation of parameter design problems]. It seems to us that this has led to some confusion about the difference between parameter design and tolerance design, the stage that follows parameter design in the product-design process.

In describing the "squeeze and shift" process, Boardman states that shifting is part of the tolerance design stage. We feel that shifting is part of parameter design, because the objective is still to reduce output variability *without controlling noise* (see Sec. 1.1 of the article). The tolerance design stage begins after we have gone as far as possible in reducing output variability by finding optimal settings of the design parameters. In the tolerance design stage output variation is reduced further by *controlling noise itself*—for example, by selecting parts that meet more stringent tolerances or by using better process controls. For more discussion of the differences between parameter design and tolerance design, see Kacker (1985).

We also believe that there are some flaws in the decision-theoretic formulation of the "product design problem" ("parameter design problem" in our terminology) given by Pignatiello and Ramberg. For example, the parameter space of decision theory should be identified with the "design parameter space" just as we have done in the article, not with the "noise space" as they suggest. Noise is the source of randomness in the random variable Y.

ACKNOWLEDGMENT

We thank all of the discussants for their stimulating remarks.

ADDITIONAL REFERENCES

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

Tsui, Kwok-Leung (1987), "Analysis of Off-Target Data for Robust Product and Process Design," unpublished manuscript.