

**IMPROVING A CALIBRATION SYSTEM
THROUGH DESIGNED EXPERIMENTS**

Arden Miller and C. F. J. Wu

IIQP Research Report
RR-91-06

August 1991

IMPROVING A CALIBRATION SYSTEM THROUGH DESIGNED EXPERIMENTS

Arden Miller and C. F. J. Wu
Department of Statistics and Actuarial Science
University of Waterloo

August 1991

ABSTRACT

Taguchi (1987) advocates the use of designed experiments to improve calibration systems. In this paper we study the statistical aspects of the problem. Inverse regression is used to estimate the value of some quantity of interest, U , using measurements, Y , of related quantity, W . We give a rigorous justification for Taguchi's signal-to-noise ratio by showing that the expected length of the Fieller intervals for U is a decreasing function of β^2/σ^2 where $Y = \alpha + \beta U + \varepsilon$, $V(\varepsilon) = \sigma^2$. Two methods of modeling this quality characteristic are proposed. Performance measure modeling, which includes Taguchi's signal-to-noise ratio approach as a special case, considers modeling β^2/σ^2 directly as a function of the control factors whereas response function modeling models β and σ separately. These methods are compared both theoretically and using an experiment on drive shaft imbalance.

KEY WORDS: Fieller intervals; Inverse regression; Model building; Performance measure modeling; Response function modeling; Signal-to-noise ratio; Taguchi.

1 Introduction

In industry, the ability to obtain accurate estimates of certain critical quantities is often a crucial part of monitoring a process. For example, the production of a chemical compound may require a complicated process. It is important to have accurate estimates of the composition of the process stream at various stages of the process so that adjustments can be made to maintain a high product yield. A second example is the measurement of residual imbalance in automobile drive shafts. Often, the manufactured drive shafts are not adequately balanced but can be corrected provided the imbalance can be measured accurately. A data set from such a measurement system is analyzed in Section 5.

Measurement systems are a practical application of statistical calibration or inverse regression. Measurements taken on a series of standards are used to calibrate the system. Then measurements taken on samples can be used to estimate the quantity of interest for the sample.

Mathematically, the calibration problem can be defined in the following manner. Let U be the quantity we wish to estimate, W be a related quantity, and Y designate measurements of W . Suppose we have p standards for which the values of U are known with a high degree of precision, and a sample. Given the measured values of W for the standards and for the sample, we wish to estimate the value of U for the sample.

The most common type of calibration assumes that the relationship between U and W is linear and deterministic, i.e., $W = \alpha + \beta U$. Further, it is assumed that the measurements, Y , are normally distributed with mean W . Therefore we have the following relationship between Y and U :

$$Y = \alpha + \beta U + \sigma\epsilon, \tag{1}$$

$$\epsilon \sim N(0, 1).$$

Most of the statistical literature on calibration considers this type of system, where the main interests are focused on obtaining estimates and making inference about U . A more ambitious engineering objective would be the reduction of measurement errors. Recognizing that a measurement system comprises a number of factors that can be adjusted to affect its precision, Taguchi (1987) advocates the use of designed experiments to identify settings of these factors which will reduce measurement errors. Adopting Taguchi's terminology, these factors will be referred to as control factors. We view a measurement system as the whole process involved in obtaining an estimate of U , which may include the sampling procedure and sample preparation as well as the actual measurement process.

Although Taguchi's idea is novel, his proposed signal-to-noise (SN) ratio lacks a firm basis and his statistical analysis can be improved. The purposes of this paper are two-fold: (i) to provide a rigorous justification of and better insight into Taguchi's SN ratio and (ii) to investigate several statistical modeling and analysis methods for achieving the stated objective.

In order to improve a measurement system some measure of the performance of the system is required. Section 2 outlines Taguchi's procedure and demonstrates that his signal-to-noise ratio is equivalent to $\hat{\beta}^2/s^2$. We have chosen the length of the confidence interval obtained for U as a suitable measure. In Section 3, the classical approach to obtaining a confidence interval for U is developed and it is shown that the length of this interval is a decreasing function of $\hat{\beta}^2/s^2$ for the fitted calibration relationship. Section 4 proposes two general methods of investigation; performance measure modeling, which includes Taguchi's procedure as a special case, and response function modeling. Sections 5 and 6 use data from an actual experiment to demonstrate and discuss the procedures. Section 7 outlines a model selection and estimation

procedure for the response function approach.

For convenience we assume the error, ϵ , in (1) and throughout the paper to be normally distributed. The validity of the results in this paper is not sensitive to this assumption since the estimates employed are standard ones in linear model theory which are known to be quite robust against moderate deviations from normality. More serious deviations can be detected by residual plots and other diagnostic methods. In such cases, robust alternatives to the least squares method may be required.

2 Taguchi's SN Ratio Analysis

Taguchi's approach (1987) is motivated by considering the mean square error of the generated estimates. Suppose the true relationship between W and U is known and is deterministic and linear. Also assume the error variance of the measurements of W is known. In terms of the definition of the calibration model (1), this means that α , β and σ^2 are known. The obvious estimator of U given a measured value, y_{obs} , for a sample is

$$u_{est} = \frac{y_{obs} - \alpha}{\beta}.$$

Now let w_t and u_t be the true values of W and U for the sample. Since $y_{obs} = w_t + \epsilon$ and $w_t = \alpha + \beta u_t$, we have

$$u_{est} = u_t + \frac{\epsilon}{\beta}.$$

Therefore,

$$V(u_{est}) = \frac{V(\epsilon)}{\beta^2} = \frac{\sigma^2}{\beta^2}, \quad (2)$$

which is also the MSE since u_{est} is unbiased. Taguchi identifies the goal of an experiment as finding the factor settings which minimize σ^2/β^2 or equivalently maximize $\omega = \beta^2/\sigma^2$. Usually, β and σ are not known and therefore must be estimated. Taguchi

rejects using $\hat{\omega} = \hat{\beta}^2/s^2$ (where $\hat{\beta}$ and s^2 are the classical estimates of β and σ^2) since $\hat{\beta}^2$ is not an unbiased estimator of β^2 . Instead, he defines a signal-to-noise (SN) ratio, $\tilde{\omega}$, as

$$\tilde{\omega} = \frac{\hat{\beta}^2 - s^2/S_{uu}}{s^2} = \frac{\hat{\beta}^2}{s^2} - \frac{1}{S_{uu}}, \quad (3)$$

where S_{uu} is defined in (4). Taguchi's approach is to implement an orthogonal design (usually factorial or fractional factorial) for the control factors. The calibration procedure is performed and $\tilde{\omega}$ is calculated for each run. A standard ANOVA is performed using $10 \log \tilde{\omega}$ as the response with the goal of maximizing ω . For additional work along this line, see Liggett (1991).

Taguchi's derivation ignores the fact that in practice β and σ must be estimated and the estimation error may affect the choice of performance measure. The implications of parameter estimation should be considered in determining control factor settings. A more rigorous derivation is given in the next section.

3 An Alternative Justification of Taguchi's SN ratio based on Fieller Intervals

In this section, the classical method of obtaining a confidence interval for U given an observed value of Y is reviewed. Fieller (1954) used a fiducial argument to develop these intervals. Consequently, confidence intervals of this type are known as Fieller intervals (Williams 1959, Seber 1977).

The measured values y_j of Y and known values u_j , $j = 1, \dots, p$, of U for the standards are used to model the relationship between Y and U . The classical estimates

for α , β , and σ^2 are

$$\begin{aligned}\hat{\beta} &= \frac{S_{yu}}{S_{uu}}, \\ \hat{\alpha} &= \bar{y} - \hat{\beta}\bar{u}, \\ s^2 &= \frac{1}{p-2}(S_{yy} - \hat{\beta}S_{yu}),\end{aligned}$$

where

$$\begin{aligned}S_{yy} &= \sum_1^p (y_j - \bar{y})^2, \\ S_{uu} &= \sum_1^p (u_j - \bar{u})^2, \\ S_{yu} &= \sum_1^p (y_j - \bar{y})(u_j - \bar{u}).\end{aligned}\tag{4}$$

For a specific value of $U = u_o$, the $100(1 - \gamma)\%$ prediction interval for Y is given by

$$\hat{\alpha} + \hat{\beta}u_o \pm ts\sqrt{1 + \frac{1}{p} + \frac{(u_o - \bar{u})^2}{S_{uu}}},\tag{5}$$

where $t = t_{\gamma/2, p-2}$. Suppose that y_o is the measured value of W for a sample which has an unknown value of U . A $100(1 - \gamma)\%$ confidence interval for u_o can be obtained by using the set of values of U for which y_o is in the $100(1 - \gamma)\%$ prediction interval of u . Therefore, the confidence interval will contain all values of u that satisfy:

$$(y_o - \hat{\alpha} - \hat{\beta}u)^2 \leq t^2 s^2 \left(1 + \frac{1}{p} + \frac{(u - \bar{u})^2}{S_{uu}}\right).\tag{6}$$

This can be rewritten as

$$A(u - \bar{u})^2 + B(u - \bar{u}) + C \leq 0,\tag{7}$$

where

$$\begin{aligned}A &= \hat{\beta}^2 - \frac{t^2 s^2}{S_{uu}}, \\ B &= -2\hat{\beta}(y_o - \bar{y}), \\ C &= (y_o - \bar{y})^2 - t^2 s^2 \left(1 + \frac{1}{p}\right).\end{aligned}$$

It is straightforward to characterize the set of values of $u - \bar{u}$ which satisfy (7) using simple calculus. Let

$$\begin{aligned} f(u - \bar{u}) &= A(u - \bar{u})^2 + B(u - \bar{u}) + C, \\ f'(u - \bar{u}) &= 2A(u - \bar{u}) + B, \\ f''(u - \bar{u}) &= 2A. \end{aligned}$$

Consider the following cases:

$A > 0$, $f(u - \bar{u})$ is a quadratic function which opens upward and has a minimum at $u - \bar{u} = -B/2A$. It is straightforward to show that $f(-B/2A) < 0$ if $A > 0$. Therefore, the solution to (7) consists of a finite interval. Note that $A > 0$ implies $\hat{\beta}^2/(s^2/S_{uu}) > t^2$ which is equivalent to the F-test for the hypothesis $H:\beta = 0$ being rejected at the γ level.

$A = 0$, the solution to (7) is a semi-infinite interval.

$A < 0$, the solution to (7) will consist of two semi-infinite intervals if the maximum, $f(-B/2A)$, is greater than zero and the entire real line otherwise.

The first case is the only one of practical interest since it would be unreasonable to use Y to predict U if there is no clear evidence of a relationship between them. Assuming a finite interval, (u_L, u_U) , its length is

$$u_U - u_L = \frac{\sqrt{B^2 - 4AC}}{A}, \quad (8)$$

which can be written as

$$\begin{aligned} u_U - u_L &= 2ts \left[\left(1 + \frac{1}{p} \right) \left(\hat{\beta}^2 - \frac{t^2 s^2}{S_{uu}} \right) + \frac{(y_o - \bar{y})^2}{S_{uu}} \right]^{1/2} \left(\hat{\beta}^2 - \frac{t^2 s^2}{S_{uu}} \right)^{-1} \\ &= 2t \left[\left(1 + \frac{1}{p} \right) \left(\hat{\omega} - \frac{t^2}{S_{uu}} \right) + \frac{(y_o - \bar{y})^2}{s^2} \frac{1}{S_{uu}} \right]^{1/2} \left(\hat{\omega} - \frac{t^2}{S_{uu}} \right)^{-1}. \quad (9) \end{aligned}$$

In this expression, t , p and S_{uu} are controlled by the calibration design so the width of the Fieller interval depends on $\hat{\omega}$ and $(y_o - \bar{y})^2/s^2$. Now, let $u_o = y_o - \hat{\alpha}/\hat{\beta}$ be the classical estimate of U given $Y = y_o$. Therefore

$$y_o - \bar{y} = \hat{\beta}(u_o - \bar{u})$$

and

$$\left(\frac{y_o - \bar{y}}{s}\right)^2 = \hat{\omega}(u_o - \bar{u})^2.$$

Substituting into (9) gives

$$u_U - u_L = 2t \left[\left(1 + \frac{1}{p}\right) \left(\hat{\omega} - \frac{t^2}{S_{uu}}\right) + \hat{\omega} \frac{(u_o - \bar{u})^2}{S_{uu}} \right]^{1/2} \left(\hat{\omega} - \frac{t^2}{S_{uu}}\right)^{-1}, \quad (10)$$

which depends on $\hat{\omega}$, S_{uu} , and $u_o - \bar{u}$. This is equivalent to a result shown by Hoadley(1970) that the width of the Fieller interval depends on the magnitude of the F statistic for testing $H:\beta = 0$:

$$f = S_{uu}\hat{\omega}. \quad (11)$$

The width of the Fieller interval decreases as $\hat{\omega}$ increases for $\hat{\omega} > t^2/S_{uu}$. This can easily be seen by rewriting (10) as

$$u_U - u_L = 2t \left[1 + \frac{1}{p} + \hat{\omega} \left(\hat{\omega} - \frac{t^2}{S_{uu}}\right)^{-1} \frac{(u_o - \bar{u})^2}{S_{uu}} \right]^{1/2} \left(\hat{\omega} - \frac{t^2}{S_{uu}}\right)^{-1/2}. \quad (12)$$

The result is evident since both $\hat{\omega}(\hat{\omega} - t^2/S_{uu})^{-1}$ and $(\hat{\omega} - t^2/S_{uu})^{-1/2}$ are decreasing functions for $\hat{\omega} > t^2/S_{uu}$. It can also be seen from this equation that for a fixed value of $\hat{\omega}$ the size of the Fieller interval decreases as S_{uu} increases. This can be done by increasing the number of standards or by spreading the standards out as far as possible. The second approach is only feasible if the assumption of a linear relation between U and W holds over the extended range. In practical systems, the assumption of linearity is often only valid over a limited region.

As $\hat{\omega}$ is a random variable, the control factors should be set to produce a favorable distribution of $\hat{\omega}$. An obvious choice is to maximize $E(\hat{\omega})$. Since $S_{uu}\hat{\omega}$ has a noncentral F distribution with 1 and $\nu = p - 2$ degrees of freedom and noncentrality parameter $\lambda = S_{uu}\omega$,

$$E(\hat{\omega}) = \frac{\nu}{\nu - 2} \left(\frac{1}{S_{uu}} + \omega \right), \quad (13)$$

$$V(\hat{\omega}) = \frac{2(\nu)^2}{(\nu - 2)^2(\nu - 4)} \left[\left(\frac{1}{S_{uu}} + \omega \right)^2 + \frac{\nu - 2}{S_{uu}} \left(\frac{1}{S_{uu}} + 2\omega \right) \right]. \quad (14)$$

Therefore, maximizing $E(\hat{\omega})$ is equivalent to maximizing ω for the true underlying function.

It should be noted that not all calibration procedures are of the form just described. The drive shaft data which we use to illustrate our procedures in Section 5 is an example of a different type of calibration. Instead of taking measurements on standards, measurements are taken on the sample drive shaft with known amounts of weight attached to it. In this case, the objective is to estimate the amount of weight which should be added to each drive shaft in order to obtain a reading of 0 (balanced drive shaft). It can be shown that ω is also a suitable performance measure for this application. In this case, since we are considering a fixed value of Y , the $100(1 - \gamma)\%$ confidence interval for Y given a specific value of U ,

$$\hat{\alpha} + \hat{\beta}u_o \pm ts \sqrt{\frac{1}{p} + \frac{(u_o - \bar{u})^2}{S_{uu}}},$$

is used to derive the expression for the Fieller interval, instead of the prediction interval (5). Following the same procedure as before, the width of the Fieller interval is

$$u_U - u_L = 2t \left[\frac{1}{p} \left(\hat{\omega} - \frac{t^2}{S_{uu}} \right) + \hat{\omega} \frac{(u_o - \bar{u})^2}{S_{uu}} \right]^{1/2} \left(\hat{\omega} - \frac{t^2}{S_{uu}} \right)^{-1}. \quad (15)$$

Note, in (15), u_o is the classical estimate of U when $Y = 0$. It is clear, that the Fieller interval still decreases as $\hat{\omega}$ increases for $\hat{\omega} > t^2/S_{uu}$.

4 Modeling and Analysis

Having established that a reasonable goal is to determine the control factor settings that maximize $\omega = \beta^2/\sigma^2$, there are two apparent methods of analysis. The first is to calculate some estimate of ω for each run of the experiment. Then this estimate (or some transformation of it) would be used as the response. The philosophy of this approach is to model the performance measure directly as a function of the control factors. This approach will be referred to as *performance measure modeling* (PMM). Taguchi's approach is a specific example of this method where $\tilde{\omega} = \hat{\beta}^2/s^2 - 1/S_{uu}$ is used to estimate ω and then the transformation $g(\tilde{\omega}) = 10 \log_{10} \tilde{\omega}$ is applied.

An alternative approach would be to model the calibration relationship as a function of the control factors. Then the control factor settings which optimize ω could be found for the fitted model. This approach separates the modeling of the calibration relationship and the optimization of the performance measure into two distinct stages. We have assumed the calibration relationship, for a fixed set of control factor levels, can be adequately described by (1). Therefore, we can model this relationship by modeling the parameters (α, β, σ) in (1). This approach will be referred to as *response function modeling* (RFM).

At this point it is useful to consider the structure of the models for the above modeling procedures. In identifying suitable models we recognize that measurement systems are often not stable. By this we mean that the relationship between Y and U is not fixed but may vary with time. For this reason, measurement systems are often recalibrated at regular intervals. It suggests that a good modeling strategy should consider β and σ^2 as random variables instead of constants for a fixed set of control factor settings. These types of models are referred as random coefficient models and are often used in econometrics.

First, consider PMM. The model we propose can be described in the following manner. Let $\hat{\omega}_i = \hat{\beta}_i^2/s_i^2$, $i = 1$ to n , be the value of $\hat{\omega}$ corresponding to the i th experimental run. Assume that the distribution of $\ln(\hat{\omega}_i)$ can adequately be described by the model

$$\ln(\hat{\omega}_i) = \ln(\omega_i) + \sigma_1 \epsilon_i, \quad (16)$$

where the errors are iid $N(0, 1)$. A justification of this is given in the Appendix. It is also shown there that

$$V(\ln \hat{\omega}_i) \approx 2/(\nu - 4) \quad (17)$$

for values of $S_{uu}\omega_i \geq 30$, where ν is the degrees of freedom associated with s_i^2 . Let Θ_ω denote the vector of coefficients for the control factors, X denote the design matrix of the experiment, and \mathbf{x}_i' denote the i th row of X corresponding to the i th control run. We now assume that $\ln(\omega_i)$ can be modeled in the following manner:

$$\ln(\omega_i) = \mathbf{x}_i' \Theta_\omega + \sigma_\omega \kappa_i, \quad (18)$$

where the error terms are iid $N(0, 1)$. Substituting (17) and (18) into (16) gives the following model for $\ln \hat{\omega}_i$:

$$\ln \hat{\omega}_i = \mathbf{x}_i' \Theta_\omega + (\sigma_\omega^2 + 2/(\nu - 4))^{1/2} \epsilon_i, \quad (19)$$

where the errors are iid $N(0, 1)$.

Now consider RFM. In this case we wish to model β and σ^2 separately. To accomplish this we propose the model

$$y_{ij} = \alpha_i + \beta_i u_j + \sigma_i \epsilon_{ij}, \quad (20)$$

$$\beta_i = \mathbf{x}_i' \Theta_\beta + \sigma_\beta \tau_i, \quad (21)$$

$$\ln(\sigma_i^2) = \mathbf{x}_i' \Theta_\sigma + \sigma_\sigma \zeta_i, \quad (22)$$

where ϵ_{ij} , τ_i , and ζ_i are independent $N(0, 1)$ random variables. First consider the modeling of σ^2 . Bartlett and Kendall (1946) have studied the use of $\ln s^2$ to study variance. Adapting this approach to the current situation we have

$$\ln(s_i^2) = \ln(\sigma_i^2) + (2/\nu)^{1/2} \epsilon_i, \quad (23)$$

where ϵ_i is approximately $N(0, 1)$. Substituting (21) into (23) gives

$$\ln(s_i^2) = \mathbf{x}_i' \Theta_\sigma + (\sigma_\sigma^2 + 2/\nu)^{1/2} \varsigma_i, \quad (24)$$

where ς is approximately $N(0, 1)$.

For β , we have the following model

$$\hat{\beta}_i = \beta_i + (\sigma_i^2/S_{uu})^{1/2} \epsilon_i. \quad (25)$$

Substituting in the model for β_i gives

$$\hat{\beta}_i = \mathbf{x}_i' \Theta_\beta + (\sigma_\beta^2 + \sigma_i^2/S_{uu})^{1/2} \epsilon_i, \quad (26)$$

where the error terms are iid $N(0, 1)$.

In Section 7 we give a step-by-step procedure for fitting these models.

5 Analysis of Drive Shaft Data

Taguchi (1987) describes an experiment which investigates a system of measuring the the residual imbalance of automobile drive shafts. Apparently, the manufactured drive shafts are often not adequately balanced which results in noise and vibration. This problem can be corrected using balance weights but this requires an accurate measurement of the amount of imbalance. Therefore, an experiment was undertaken with the goal of finding the control factor settings which produce the most precise

<i>A</i> : Testing Machine	<i>A</i> ₁ : New	<i>A</i> ₂ : Old
<i>B</i> : Master Rotor	<i>B</i> ₁ : #1	<i>B</i> ₂ : #2
	<i>B</i> ₃ : #3	<i>B</i> ₄ : #4
<i>C</i> : Rotations at Handling Time	<i>C</i> ₁ : Current	<i>C</i> ₂ : New
<i>D</i> : Rotations at Measurement	<i>D</i> ₁ : Current	<i>D</i> ₂ : New
<i>E</i> : Signal Sensitivity	<i>E</i> ₁ : 10	<i>E</i> ₂ : 20
	<i>E</i> ₃ : 30	<i>E</i> ₄ : 40
<i>F</i> : Sequence of Correction of Imbalance	<i>F</i> ₁ : Current	<i>F</i> ₂ : Reverse
	<i>F</i> ₃ : New #1	<i>F</i> ₄ : New #2
<i>G</i> : Imbalance Correction Location	<i>G</i> ₁ : Current	<i>G</i> ₂ : New

Table 1: Control Factors for Drive Shaft Data

measurements. The experiment contained eight control factors which are listed in Table 1. Three drive shafts (M'_1, M'_2, M'_3) were tested at each combination of control factors. Four measurements were taken on each drive shaft corresponding to the following conditions:

M_1 : Drive shaft measured as is.

M_2 : Drive shaft measured with 10g weight attached to mass-deficient side.

M_3 : Drive shaft measured with 20g weight attached to mass-deficient side.

M_4 : Drive shaft measured with 30g weight attached to mass-deficient side.

In the original experiment data is collected for both the flange side and the sleeve side of the drive shaft. For the purpose of illustration we will only consider the data for the flange side. Also, to simplify the analysis, the control factors which have four

		Control Factors									Control Factors						
Run	A	B	C	D	E	F	G	Run	A	B	C	D	E	F	G		
1	1	1	1	1	1	1	1	9	2	3	1	1	4	2	1		
2	1	2	1	2	2	2	1	10	2	4	1	2	3	1	1		
3	1	3	2	1	3	3	1	11	2	1	2	1	2	4	1		
4	1	4	2	2	4	4	1	12	2	2	2	2	1	3	1		
5	1	2	1	1	3	4	2	13	2	4	1	1	2	3	2		
6	1	1	1	2	4	3	2	14	2	3	1	2	1	4	2		
7	1	4	2	1	1	2	2	15	2	2	2	1	4	1	2		
8	1	3	2	2	2	1	2	16	2	1	2	2	3	2	2		

Table 2: Design Matrix for Drive Shaft Data

levels (B, E, F) will each be analyzed by considering three linear contrasts. As factor E is quantitative, orthogonal polynomials were used to estimate effects corresponding to the linear, quadratic, and cubic components (E_l, E_q, E_c). Factors B and F are qualitative. The contrasts used for these factors are designated by B₁, B₂, B₃, F₁, F₂, and F₃. B₁ contrasts levels 1 and 2 with levels 3 and 4, B₂ contrasts levels 1 and 3 with levels 2 and 4, and B₃ contrasts levels 1 and 4 with levels 2 and 3. The contrasts for F are designated in the same manner.

It is assumed that a linear relationship exists between the measured quantities and the true imbalance of the drive shafts. This means the response for each control factor combination consists of three simple linear functions corresponding to the three drive shafts. These functions are assumed to have common slopes and error variances but different intercepts. It is always well advised to make sure that the observed data is reasonably consistent with the assumptions. Scatterplots of the data indicate

Run	M'_1				M'_2				M'_3			
	M_1	M_2	M_3	M_4	M_1	M_2	M_3	M_4	M_1	M_2	M_3	M_4
1	-4	6	18	27	-4	6	15	25	-20	-10	2	14
2	-7	10	23	42	-3	15	32	46	-34	-20	4	16
3	-4	9	22	34	-7	6	18	30	-26	-15	0	12
4	-2	10	22	36	-4	8	22	34	-30	-18	-7	10
5	-6	6	16	28	-5	6	16	27	-21	-12	3	13
6	-7	13	32	50	-7	14	31	48	-45	-27	-8	13
7	-13	10	30	52	-8	10	30	50	-37	-18	7	26
8	-19	8	27	48	-14	7	29	52	-42	-25	-2	22
9	-10	4	16	29	-8	6	16	26	-29	-20	-14	-14
10	-14	11	32	51	-18	4	25	46	-44	-26	-11	16
11	-3	2	10	16	-4	2	17	13	-13	-8	-5	7
12	-5	5	16	25	-7	3	12	22	-22	-14	-8	12
13	-4	6	18	30	-8	2	16	28	-23	-15	-10	18
14	-6	16	38	62	-16	6	32	55	-44	-25	-6	21
15	-4	2	7	14	-4	2	6	12	-10	-6	-3	7
16	-5	6	16	30	-7	4	16	27	-25	-14	-6	14

Table 3: Drive Shaft Data (flange side)

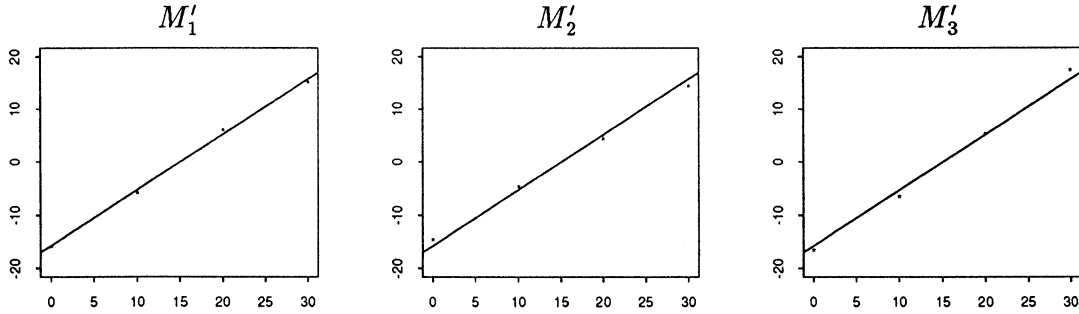


Figure 1: Scatterplots for Run 1 ($y - \bar{y}$ vs. attached weight)

this is the case for most of the runs (Figure 1 contains the scatterplots for Run 1). However, runs 9 and 13 each have one point which appears somewhat unusual. Since the assumptions are reasonably satisfied, we will proceed to illustrate the procedures.

The control factor settings for this experiment are given in Table 2. Table 3 contains the raw data for this experiment and Table 4 contains the estimated values of the parameters for each experimental run.

First consider PMM. Model (19) is valid provided that the values of $S_{uu}\omega_i$ are larger than 30. In the present situation this is reasonable since the smallest value of $S_{uu}\hat{\omega}_i$ occurs for run 9 and is 58.5. Since this model assumes a constant variance for the response, the standard half normal plots should be valid. The half normal plot is given in Figure 2. The plot indicates that effect A is clearly significant and effects Ec and D are marginal.

Suppose the only significant source of variation is due to estimation error. Then using equation (34) from the Appendix, the approximate variance for each $\ln \hat{\omega}_i$ is 0.5. Let Θ_ω denote the vector of true effects for the control factors and $\ln \hat{\omega}$ be the vector of $\ln \hat{\omega}_i$. The covariance matrix for the estimated effects, $\hat{\Theta}_\omega = (X'X)^{-1}X' \ln \hat{\omega}$ will be $0.5(X'X)^{-1}$. For the design matrix used in this example, the effects should be

Run	$\hat{\beta}$	s^2	$\hat{\beta}^2/s^2$	$\ln(\hat{\beta}^2/s^2)$	Run	$\hat{\beta}$	s^2	$\hat{\beta}^2/s^2$	$\ln(\hat{\beta}^2/s^2)$
1	1.050	1.375	0.802	-0.221	9	0.973	24.304	0.039	-3.245
2	1.660	5.075	0.543	-0.611	10	2.080	6.378	0.642	-0.443
3	1.263	0.529	3.017	1.104	11	0.647	10.529	0.040	-3.225
4	1.283	1.417	1.162	0.150	12	1.017	6.667	0.155	-1.863
5	1.117	1.292	0.966	-0.034	13	1.210	17.842	0.082	-2.500
6	1.883	1.417	2.504	0.918	14	2.260	4.842	1.055	0.054
7	2.077	2.992	1.442	0.366	15	0.550	1.875	0.161	-1.824
8	2.183	3.667	1.300	0.262	16	1.180	4.863	0.286	-1.251

Table 4: Estimated Parameters for Drive Shaft Data

independent and have a variance of approximately 0.125. So given the null hypothesis, $H_0: \Theta_\omega = 0$, we can calculate critical points c_1 and c_2 such that,

$$Pr[| \text{any specific estimated effect} | < c_1 | H_0] = 0.95, \quad (27)$$

$$Pr[| \text{all estimated effects} | < c_2 | H_0] = 0.95. \quad (28)$$

For the given situation, $c_1 = 0.69$ and $c_2 = 0.97$. These lines are plotted on the half normal plot. From the plot we see that these lines are consistent with our initial assessment.

Now consider RFM. The analysis for $\ln s^2$ is very similar to that performed for $\ln \hat{\omega}$. For the assumed model (24), the variance of $\ln s_i^2$ is approximately constant for each run. Therefore, the estimated effects for the factors should be approximately iid normal variates. The half normal plot for the estimated effects is given in Figure 3. Also, under the assumption that the only significant source of variation is due to estimation error, we can calculate the individual and simultaneous 95% critical points

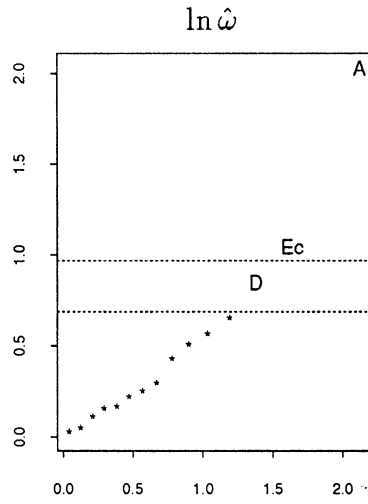


Figure 2: Half Normal Plot for $\ln \hat{\omega}$

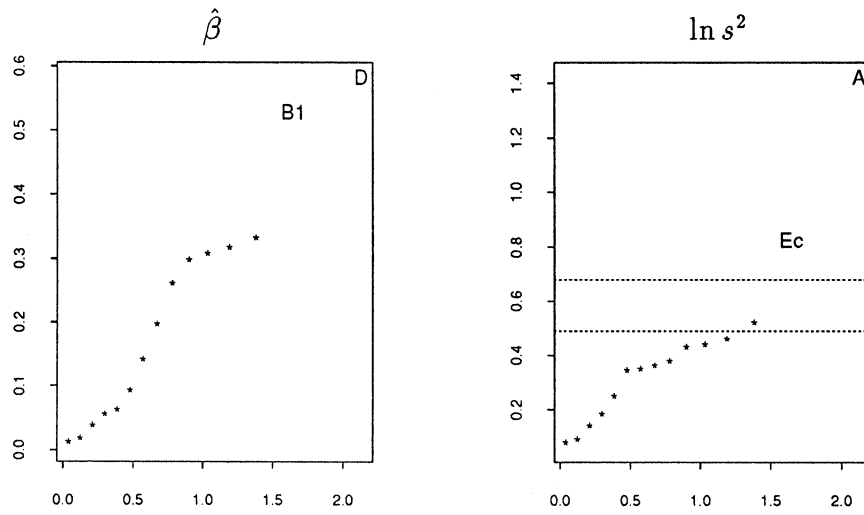


Figure 3: Half Normal Plots for $\hat{\beta}$ and $\ln s^2$

for the hypothesis that the effects are 0. These points are $c_1 = 0.49$ and $c_2 = 0.68$ respectively. Lines corresponding to these points are indicated on the half normal plot. These lines are consistent with the conclusions we would reach just from an inspection of the plot. Therefore it is consistent with the data to assume that estimation error is the only important source of variation. The half normal plot indicates that there is clear evidence that factors A and Ec are significant. There is no evidence that any of the other factors are significant.

Now consider the analysis of $\hat{\beta}$. The model being used is given by (26). In this case, the errors associated with the estimated slopes are not constant. Therefore, in general, weighted least squares (WLS) should be used to estimate the control factor effects. However, for the present example, we have a design matrix which is nearly saturated in that only 1 degree of freedom is available to estimate error. We prefer to estimate an effect corresponding to this degree of freedom and use techniques such as half normal plots to evaluate significance. Therefore, the estimates obtained using WLS will be exactly the same as those obtained using ordinary least squares (OLS). We must be careful in evaluating these estimates since, given the above model, it is not correct to assume that the estimated effects are independent. Suppose we define the matrix D as being diagonal with the i th diagonal element $\sigma_\beta^2 + \sigma_i^2/S_{uu}$. Then the covariance matrix for the estimated effects given the model (26) is $(X'D^{-1}X)^{-1}$. Therefore, we must be careful in using techniques which assume independent estimates, such as half normal plots, in evaluating the importance of effects. A major difficulty is that for a saturated design matrix there are no available degrees of freedom to obtain an estimate of σ_β^2 . This makes it difficult to assess the degree of departure from independence. Note that if the value of σ_β^2 is large with respect to the range of the σ_i^2/S_{uu} terms, the off diagonal elements of $(X'D^{-1}X)^{-1}$ will be relatively small and the half normal plot will give a reasonable representation of the effects. The procedure

we propose is iterative in nature. First, we select a subset of the control factors which we suspect are important and refit the model retaining only these factors. The half normal plot, Figure 3, was used as a guide for this initial selection process.

For this subset of control factors we get refined estimates of the effects as well as an estimate of σ_β^2 by using an iterative reweighted procedure. Details of this procedure are given in Section 7.

The half normal plot would indicate only factors D and B1 are significant if the estimated effects were independent. As this is not the case, the WLS procedure was used to fit a model which contained factors A, G, CD, El and C as well as D and B1. The estimated effects are given in Table 5 under R-1. These estimates are consistent with those obtained for the full model using OLS (first column of Table 5). Finally the model containing just factors D and B1 was considered. The estimated effects for this model are given in Table 5 under R-2. Again, the estimates are consistent with the OLS estimates. The estimated value of σ_β^2 for this model is 0.1251. Since the values of s_i^2/S_{uu} range from 0.0003 to 0.0162, the estimated value of σ_β^2 is large enough to support treating the estimated effects as if they were independent. Therefore, the half normal plot is a valid form of analysis.

6 Discussion of the Drive Shaft Experiment

In analyzing the data it is important to consider the objective of the experiment. Box (1988, 1966) identifies two types of feedback, *empirical* and *scientific*, which can occur as the result of an experiment. Empirical feedback uses some definite rule to determine the response taken as the result of the experimental outcome. For example, suppose we consider the experiment simply as a mechanism to identify recommended control factor settings. The recommended settings are those the analysis predicts will yield

Factors	Model		
	Full	R-1	R-2
Intercept	1.403	1.399	1.406
A	0.324	0.330	-
B1	-0.529	-0.524	-0.532
B2	0.056	-	-
B3	0.048	-	-
C	0.255	0.262	-
D	-0.583	-0.580	-0.574
E1	-0.291	-0.320	-
Eq	-0.031	-	-
Ec	-0.085	-	-
F1	0.133	-	-
F2	0.005	-	-
F3	-0.012	-	-
G	-0.312	-0.342	-
CD	-0.301	-0.298	-
e	-0.189	-	-
$\hat{\sigma}_\beta^2$	-	0.0117	0.1251

Table 5: Estimated Effects for β

the optimal output. Questions of statistical significance are not important since the recommended settings for “nonsignificant effects” are arbitrary choices which should have little impact on the system. For scientific feedback the goal of the experiment is to enhance the investigator’s knowledge of the system. This may result in the investigator using the knowledge gained from the experiment (as well as knowledge from other sources) to determine a future course of investigation. For this type of feedback, questions of significance are relevant. We will consider both types of feedback for comparing PMM and RFM.

Consider PMM and RFM with respect to scientific feedback. First consider performance measure modeling. Factor A is the only clearly significant effect. The half normal plot for $\ln \hat{\omega}$ indicates that effect Ec and possibly effect D are large enough to warrant further investigation although not clearly significant. Now consider response function modeling. Figure 3 contains the half normal plots for $\hat{\beta}$ and $\ln s^2$. The analysis for $\hat{\beta}$ indicates that effects D and B1 are significant. From the half normal plot for $\ln s^2$ it would appear that effects A and Ec are significant. It seems somewhat unusual that the cubic orthogonal contrast for factor E has been identified as being important. The cubic contrast for E is partially aliased with a number of two factor interactions as follows:

$$E(Ec) = \mu_{Ec} + (\mu_{A \times F1} + \mu_{B1 \times G} + \mu_{B3 \times D1})/\sqrt{5} - 2(\mu_{A \times D} + \mu_{C \times F3} + \mu_{G \times F2})/\sqrt{5}.$$

It may well be that one of these interactions is causing the observed effect. In particular, since A is clearly significant the interactions involving A should be considered as strong possibilities.

It is instructive to consider the results for factors A, B1 and D more carefully. Factor A is significant in the analysis of $\ln s^2$, and $\ln \hat{\omega}$ but not $\hat{\beta}$. So there is evidence that A affects σ as well as the performance measure, ω , but no clear evidence that

RFM	$\hat{\beta}$	D B1
	$\ln s^2$	A Ec
PMM	$\ln \hat{\omega}$	A (Ec) (D)

Table 6: Significant Effects

it effects β . We would conclude that the new testing machine gives more precise readings and this is mainly the result of a smaller measurement variance. Factors B1 and D are a different matter. There is evidence that B1 and D affect β but no clear evidence that either affects σ or ω . This means that in both cases we cannot reject either the possibility that the factor has no effect on σ or the possibility it has an effect on σ large enough to offset its effect on β . It may be that a person familiar with the measurement system can determine which of these possibilities is more plausible. Notice that for the analysis using $\ln \hat{\omega}$, B1 is overlooked entirely and D is considered marginally significant. Table 6 summarizes the effects which were found to be significant by the different analyses. The marginal effects are in parentheses.

Now consider the question of empirical feedback. In particular, consider Taguchi's approach of using the experiment as simply a method of determining factor levels. The PMM approach is very straightforward to apply in this situation. Factors A, B, F, and G are all qualitative and therefore the recommended settings are simply the best marginal settings. For factors C and D, the recommended settings depend on the calculated effect of the CD interaction as well as the main effects. Factor E is the only quantitative factor. The fitted cubic model is used to determine the recommended setting for this factor over the range studied in the experiment.

For RFM the strategy is to calculate the effects for $\hat{\beta}$ and s^2 separately. The effects are then combined to determine the preferred level for each factor. The results

	PMM	RFM
Factor	$\ln(\hat{\beta}^2/s^2)$	$\hat{\beta} \ln s^2$
A	1	1
B	3	3
C	1	1
D	2	2
E	33.4	34.0
F	1	1
G	2	2

Table 7: Recommended Control Factor Settings for Empirical Feedback

of both analyses are given in Table 7. The recommended settings for all qualitative factors are the same. The recommended settings for E are only slightly different (34.0 vs. 33.4). In this example the type of empirical feedback we are considering is insensitive to the modeling procedure.

7 General Modeling Procedure for RFM

The general procedure we propose for response function modeling is as follows.

1. Estimation of Regression Coefficients

Estimates of β and σ^2 are obtained for each run in the experiment. In this paper we only consider the standard least square estimators. However, in certain situations it may be valuable to use estimators which are more robust to outliers.

2. Model Identification for σ^2

The general procedure is to obtain maximum likelihood estimates using (24) as the model. This is a straightforward process provided the design matrix is not saturated. In this case the MLE's for Θ_σ are the usual least squares estimates obtained by solving

$$X' \ln s^2 - X'X\Theta_\sigma = 0,$$

where $\ln s^2$ is the vector of $\ln s_i^2$ terms. The MLE for σ_σ^2 is obtained by solving

$$\hat{\sigma}_\sigma^2 = (1/\nu)(\ln s^2 - X\hat{\Theta}_\sigma)'(\ln s^2 - X\hat{\Theta}_\sigma) - (2/\nu).$$

The estimate of σ_σ^2 is constrained to be non-negative. So if the solution is negative, $\hat{\sigma}_\sigma^2$ is taken to be 0. Now, we can calculate an estimated variance for each $\ln s_i^2$ as $\hat{\sigma}_\sigma^2 + 2/\nu$. Therefore, an estimated covariance matrix for $\hat{\Theta}_\sigma$ is $(\hat{\sigma}_\sigma^2 + 2/\nu)(X'X)^{-1}$. In most cases, the design matrix used will be orthogonal and scaled so that the estimated effects have equal variance. This will result in an estimated covariance matrix of the form $(\hat{\sigma}_\sigma^2 + 2/\nu)kI$. Therefore, it is relatively easy to calculate critical values c_1 and c_2 such that

$$Pr[| \text{any specific estimated effect} | < c_1 | H_o] = 0.95, \quad (29)$$

$$Pr[| \text{all estimated effects} | < c_2 | H_o] = 0.95. \quad (30)$$

We recommend that a half normal plot of the effects with lines at c_1 and c_2 be produced. This is an effective visual tool in assessing the importance of the factors.

The situation is only slightly more complicated when the design matrix is saturated. In this case we cannot obtain a direct estimate of σ_σ^2 . However, given the design matrix is orthogonal and properly scaled, the estimated effects will be independent and have equal variance. Therefore, a half normal plot can

be used to assess which effects are significant. An estimate of σ_σ^2 can then be obtained when only this subset of effects is included in the model. Note, that the estimates of Θ_σ are not affected. This estimate of σ_σ^2 can now be used for subsequent inference.

3. Model Identification for β

First, consider the case where the design matrix is not saturated. The coefficients for the model given in (26) can be estimated by using an iteratively reweighted procedure. This procedure uses the equations generated by differentiating the log likelihood function for (26) by Θ_β and σ_β^2 . In these equations each σ_i^2 is replaced by the corresponding s_i^2 . The procedure starts by using the OLS estimates of Θ_β to obtain an estimate $\hat{\sigma}_\beta^2$ of σ_β^2 from solving

$$-\sum_i (\sigma_\beta^2 + s_i^2/S_{uu})^{-1} + \sum_i (\hat{\beta}_i - \mathbf{x}_i' \hat{\Theta}_\beta)^2 (\sigma_\beta^2 + s_i^2/S_{uu})^{-2} = 0. \quad (31)$$

Then new estimates for Θ_β are obtained by solving

$$X'D^{-1}\hat{\beta} - X'D^{-1}X\Theta_\beta = 0,$$

where $\hat{\beta}$ is the vector of $\hat{\beta}_i$ s and D is a diagonal matrix with i th diagonal element equal to $\hat{\sigma}_\beta^2 + s_i^2/S_{uu}$. These two steps are repeated until the estimates for Θ_β and σ_β^2 converge. If the estimated value of σ_β^2 is large with respect to the s_i^2/S_{uu} terms, then the off diagonal elements of the estimated covariance matrix, $(X'D^{-1}X)^{-1}$, will be small with respect to the diagonal elements and we can treat the estimated effects as being approximately independent. In this case, we recommend using a half normal plot as well as approximate critical values for testing hypotheses of the type given in (29) and (30) to evaluate the significance of the control factor effects. If this is not the case, a more rigorous procedure is required. Since the covariance between estimated effects cannot be

neglected, some kind of search procedure is required to identify the subsets of control factors which produce good final models. The search procedure can be simplified by using the initial analysis to divide the control factors into three groups:

- (a) Those factors which clearly should be retained in the model.
- (b) Those factors which are marginal in that it is not clear whether they should be retained in the model or not.
- (c) Those factors which can safely be eliminated from the model.

Only models which contain the factors in (a) and a subset of the the factors in (b) need be considered. Usually, the number of factors in (b) is small enough so that an exhaustive search can be done. A number of criteria, including Akaike's information criteria (AIC) (Miller 1990), are available which can be used to assess the suitability of the models.

A more difficult situation arises when the design matrix is saturated. We recommend the following procedure. First, produce a half normal plot of the estimated effects obtained using OLS. This plot can be used to divide the control factors into three sets as was done previously. The model containing all factors from sets (a) and (b) should then be fitted and an estimate of σ_{β}^2 obtained using (31). If the estimated effects for this model are quite close to those for the original model and the covariance matrix estimated by $(X'D^{-1}X)^{-1}$ has relatively small off diagonal elements, then it is reasonable to use the original half normal plot to evaluate the relative importance of effects. Otherwise, as before, an exhaustive search of models which contain the factors in (a) and subsets of those in (b) can be done using AIC or related methods to evaluate the relative suitability of the models.

In general the procedures may identify several models as suitable. Practical knowledge of the system and subsequent experimentation should be used to determine a final model.

Acknowledgements: This research was supported by the Natural Sciences and Engineering Research Council of Canada, General Motors of Canada Limited, and the Manufacturing Research Corporation of Ontario. We also wish to acknowledge Jock MacKay for his helpful suggestions.

Appendix: Model Justification for ω

In considering an estimator for $\omega = \beta^2/\sigma^2$, it makes sense to consider functions of $\hat{\beta}$ and s since $(\hat{\beta}, s)$ is minimal sufficient for (β, σ) . The most obvious choice would simply be $\hat{\omega} = \hat{\beta}^2/s^2$. Taguchi rejects this estimator on the grounds that $\hat{\beta}^2$ is not unbiased for β^2 . This is not a valid argument as there is no compelling reason to require an unbiased estimator. Further, although Taguchi's alternative, $\tilde{\omega}$, can be constructed from the quotient of an unbiased estimator of β^2 and an unbiased estimator of σ^2 , $\tilde{\omega}$ is not unbiased for ω . It is straightforward to derive, from (13) and (3), that

$$E(\tilde{\omega}) = \frac{\nu}{\nu - 2}\omega + \frac{2}{\nu - 2}\frac{1}{S_{uu}}.$$

In most practical situations, there will be very little difference between $\tilde{\omega}$ and $\hat{\omega}$. From (11) we have

$$1/S_{uu} = \frac{1}{f}\hat{\omega}.$$

Therefore,

$$\tilde{\omega} = \hat{\omega} \left(1 - \frac{1}{f}\right) \approx \hat{\omega}$$

for large f . Note, that the F-statistic, f , should have a large value for the measurement system to be useful (see Section 3).

Now, consider possible transformations of the estimator. Major reasons for transforming the response variable are to stabilize the error variance, to obtain a parsimonious model, and to make the error terms approximately normal.

Stabilization of error variance can be assessed from theoretical considerations. First, note from (13) and (14) that $E(\hat{\omega})$ is linear in ω and $V(\hat{\omega})$ is quadratic in ω . This would suggest that some type of logarithmic transformation may be useful. Consider transformations of the form $\ln(\hat{\omega} + c)$, where c is constant with respect to ω . The delta method can be used to approximate the variance of this transformation:

$$\begin{aligned} V[\ln(\hat{\omega} + c)] &\approx \frac{V(\hat{\omega})}{[E(\hat{\omega}) + c]^2} \\ &= k_1 \frac{(S_{uu}^{-1} + \omega)^2 + (\nu - 2)(S_{uu}^{-2} + 2S_{uu}^{-1}\omega)}{(S_{uu}^{-1} + \omega + c^*)^2} \\ &= k_1 \left[1 + \frac{(\nu - 2)(S_{uu}^{-2} + 2S_{uu}^{-1}\omega) - 2c^*(S_{uu}^{-1} + \omega) - c^{*2}}{(S_{uu}^{-1} + \omega + c^*)^2} \right], \quad (32) \end{aligned}$$

where $k_1 = 2/(\nu - 4)$ and $c = c^*\nu/(\nu - 2)$. Clearly, if we let $c^* = (\nu - 2)S_{uu}^{-1}$, then the coefficient of ω becomes 0 in the numerator of the second term of (32).

Therefore, we obtain

$$\begin{aligned} V[\ln(\hat{\omega} + \nu/S_{uu})] &\approx k_1 \left[1 - \frac{(\nu - 1)(\nu - 2)S_{uu}^{-2}}{(\omega + (\nu - 1)S_{uu}^{-1})^2} \right] \\ &\approx k_1 \left[1 - (\nu - 1)(\nu - 2)S_{uu}^{-2}\omega^{-2} + O(\omega^{-3}) \right]. \quad (33) \end{aligned}$$

Compare this to the variance of $\ln \hat{\omega}$ which can be approximated from (32) by taking $c^* = 0$:

$$V[\ln \hat{\omega}] \approx k_1 \left[1 + 2(\nu - 2)S_{uu}^{-1}\omega^{-1} + O(\omega^{-2}) \right]. \quad (34)$$

This indicates that the transformation $\ln(\hat{\omega} + \nu/S_{uu}^2)$ should stabilize the error variance at lower values of ω than $\ln \hat{\omega}$. For sufficiently large values of $S_{uu}\omega$ either transformation will effectively stabilize the error variance.

Note that Taguchi's SN ratio (3) is equivalent to taking $c^* = (2 - \nu)\nu^{-1}S_{uu}^{-1}$ in (32). From (32),

$$V[\ln \tilde{\omega}] \approx k_1 \left[1 + 2(\nu - 2)(\nu + 1)\nu^{-1}S_{uu}^{-1}\omega^{-1} + O(\omega^{-2}) \right]. \quad (35)$$

The ω^{-1} term is slightly larger in this expression than that for $V(\ln \hat{\omega})$. This is another reason for not favoring the use of Taguchi's SN ratio.

Figure 4 compares the variance of $\ln \hat{\omega}$, $\ln \tilde{\omega} = \ln(\hat{\omega} - 1/S_{uu})$, and $\ln(\hat{\omega} + \nu/S_{uu})$ as a function of $S_{uu}\omega$ when $\nu = 8$. All three variances converge to 0.5. As a rough guide, it appears that all three transformations are adequate for values of $S_{uu}\omega$ above 30. However, there is a distinct advantage in using $\ln(\hat{\omega} + \nu/S_{uu})$ for values of $S_{uu}\omega$ between 10 and 30.

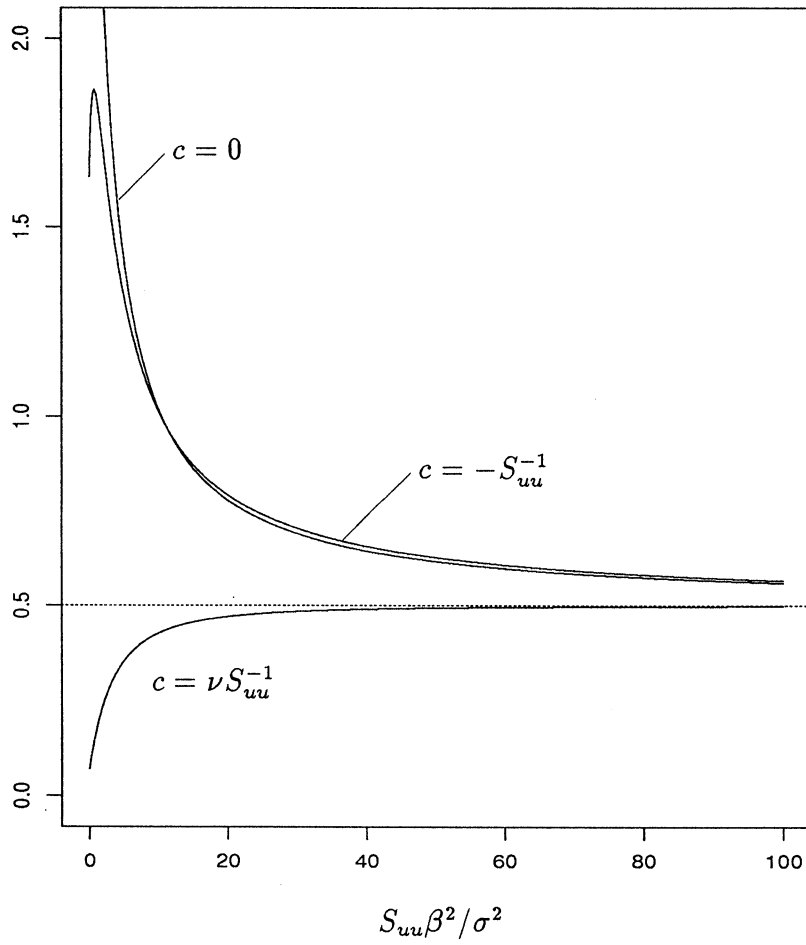


Figure 4: $V[\ln(\hat{\beta}^2/s^2 + c)]$ against $S_{uu}\beta^2/\sigma^2$ for $\nu = 8$

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. (1966), "A Simple System of Evolutionary Operation Subject to Empirical Feedback," *Technometrics*, **8**, 19-26.
- (1988), "Signal-to-Noise Ratios, Performance Criteria, and Transformations," *Technometrics*, **30**, 1-17.
- Fieller, E. C. (1954), "Some Problems in Interval Estimation," *Journal of the Royal Statistical Society, Ser. B*, **16**, 175-185.
- Hoadley, B. (1970), "A Bayesian Look at Inverse Linear Regression," *Journal American Statistical Association*, **65**, 356-369.
- Liggett, W. (1991), "Project Planning for Sequential Parameter-Design Experiments," *Preprint*.
- Miller, A. J. (1990), *Subset Selection in Regression*, Chapman and Hall, London.
- Seber, G. A. F. (1977), *Linear Regression Analysis*, Wiley, New York.
- Taguchi, G. (1987), *System of Experimental Design*, Unipub/Kraus International Publications, White Plains, New York.
- Williams, E. J. (1959), *Regression Analysis*, Wiley, New York.