Metrologia 26, 107–113 (1989)

metrologia

An Extended Error Model for Comparison Calibration

C. Croarkin

Statistical Engineering Division, National Institute of Standards and Technology*, Gaithersburg, MD 20899, USA

Received: September 10, 1988 and in revised form November 18, 1988

Abstract

The usual error model for calibration experiments is extended to situations where there are both short-term and long-term random errors of measurement. Such error models are useful where short-term errors are related to instrumentation, and long-term errors are related to operating procedures, environmental factors or changes in the artifacts themselves. The concept of a check standard is advanced for estimating variability and maintaining statistical control of the measurement process.

Introduction

Comparison calibration relates a characteristic of an artifact or instrument to the defined unit for the quantity of interest. A reference standard, whose value has been independently established, is the basis for assigning a value to the unknown artifact. For calibrations at the highest accuracy levels, very precise comparators with linear responses over a small on-scale range are used to quantify small differences between artifacts of the same nominal value. We describe an error model and analysis where two unknowns are compared with two reference standards according to a specific design.

Calibration Model

In the simplest case, an unknown X with value X^*, yet to be determined, is assumed to be related to a reference standard R with known value R^* by

$$X^* = \Delta + R^*$$

where Δ is small but not necessarily negligible.

Given a measurement x on the unknown and a measurement r on the reference standard, the responses are assumed to be of the form

$$x = \eta + X^* + e_x$$

and (1)

$$r = \eta + R^* + e_r$$

where η is instrumental offset and e_x and e_r are independent random errors which come from a distribution with mean zero and standard deviation σ.

The value of Δ is estimated[1] by the difference $\boldsymbol{\Delta}$ where

$$\boldsymbol{\Delta} = x - r \tag{2}$$

and the value assigned to the unknown artifact is based on the known value of the reference standard, R^*, called the restraint, according to

$$\boldsymbol{X}^* = \boldsymbol{\Delta} + R^* . \tag{3}$$

The standard deviation of this estimate, σ_X, depends on the error structure for $\boldsymbol{X}^*$ which is of the form

$$\boldsymbol{X}^* = X^* + e_x - e_r \tag{4}$$

so that

$$\sigma_X = \sqrt{2}\sigma . \tag{5}$$

Calibration Designs

A more complicated case involves the calibration of several unknowns, such as a weight set of various denominations or a group of voltage cells in a temperature-controlled enclosure, relative to a single reference standard or group of standards. Any difference measurements which compare unknowns and reference standards with one another and each other are candidates for the calibration procedure.

* Formerly, the U.S. National Bureau of Standards

[1] Boldface type is used to denote a least-squares estimate from the data such as $\boldsymbol{\Delta}$

A calibration design is a subset of all candidate measurements which admits a least-squares solution for the unknowns. The design is constructed to be parsimonious so as, on one hand, to minimize the number of measurements and, on the other hand, to give estimates with reasonably high precision. We recognize that precision depends on the number of measurements, and Grabe [1] has shown how precision depends on the construction of the design. As we show in this paper, precision can also be limited by other factors.

In the earliest references to designs by Hayford and Benoit [2, 3], the term "weighing design" is used to describe a sequence of measurements for calibrating a weight set. In papers published in the 1960s and 1970s, Bose and Cameron [4, 5] and Chakravarti and Suryanarayana [6] extend the theory and application of designs; Cameron and Eicke [7] solve a problem peculiar to electrical circuits; and Cameron and Hailes [8] discuss the situation where there is drift in the measurement process. Recent publications [9–12] show that designs now enjoy general acceptance in the calibration laboratory and are routinely used for the calibration of mechanical and electrical units of measurement, as well as for mass measurements.

Expanded Calibration Model

Throughout this development, the one constant assumption has been that random errors of measurement are independent and come from a single error distribution (such as the normal distribution). With more precise measurement systems, we are now able to identify situations where these assumptions are called into question and a more realistic model is needed. We find that random errors of measurement for a single design, which takes at most a few hours' time, are not of the same magnitude as errors which afflict the measurement process over the course of several designs or days[2]. Thus, we are forced to admit two error distributions, one that arises in the short-term and one that arises in the long term.

It is convenient to think in terms of short-term instrumental variations and long-term artifact changes caused by environmental conditions and the like. The latter are assumed to vary randomly from design to design and to be constant for a single design. The model in (1) is expanded to include both types of errors so that

$$\begin{aligned} x &= \eta + \{X^* + \delta_X\} + e_x \\ r &= \eta + \{R^* + \delta_R\} + e_r \end{aligned} \tag{6}$$

where e_x and e_r are short-term errors of (1), and δ_X and δ_R, which represent long-term changes associated with X and R, come from a distribution with mean zero and standard deviation σ_b.

The error structure of the estimate, X^*, given by

$$X^* = X^* + \delta_X - \delta_R + e_x - e_r \tag{7}$$

now contains both types of error terms, and the standard deviation σ_X becomes

$$\sigma_X = (2\sigma_b^2 + 2\sigma^2)^{1/2}.$$

Application to Designs

Standard deviations associated with solutions to a design depend upon the error structures of the model. We illustrate with an example where two unknown artifacts X_1 and X_2 with unknown values X_1^* and X_2^* are calibrated relative to two reference standards R_1 and R_2 with values R_1^* and R_2^*. All items have the same nominal value. A design consisting of the six comparisons $d_1, \ldots, d_6$ that can be made among the four items, two at a time, can be represented as:

Obs	R_1	R_2	X_1	X_2
d_1	1	−1		
d_2	1		−1	
d_3	1			−1
d_4		1	−1	
d_5		1		−1
d_6			1	−1

The model that follows from this design is:

$$\begin{aligned}
d_1 &= \{R_1^* + \delta_{R_1}\} - \{R_2^* + \delta_{R_2}\} + \varepsilon_1 \\
d_2 &= \{R_1^* + \delta_{R_1}\} - \{X_1^* + \delta_{X_1}\} + \varepsilon_2 \\
d_3 &= \{R_1^* + \delta_{R_1}\} - \{X_2^* + \delta_{X_2}\} + \varepsilon_3 \\
d_4 &= \{R_2^* + \delta_{R_2}\} - \{X_1^* + \delta_{X_1}\} + \varepsilon_4 \\
d_5 &= \{R_2^* + \delta_{R_2}\} - \{X_2^* + \delta_{X_2}\} + \varepsilon_5 \\
d_6 &= \{X_1^* + \delta_{X_1}\} - \{X_2^* + \delta_{X_2}\} + \varepsilon_6
\end{aligned} \tag{8}$$

[2] The statistical term for this phenomenon is components of error with the errors sometimes referred to as within-time and between-time random errors

The terms $\varepsilon_1, \ldots, \varepsilon_6$ represent random errors of measurement and the terms $\delta_{R_1}, \delta_{R_2}, \delta_{X_1}$, and δ_{X_2} represent random changes in the artifacts. It is assumed that the ε terms come from a distribution with mean zero and standard deviation σ_w and that the δ terms come from a distribution with mean zero and standard deviation σ_b. All random errors are assumed to be mutually independent.

The solution to the design depends on the restraint. If the restraint is taken to be the average of the reference standards or

$$R^* = \tfrac{1}{2}(R_1^* + R_2^*),$$

then least-squares estimates (see, for example, Cameron et al. [13]) are as follows:

$$\begin{aligned}
R_1^* &= \tfrac{1}{8}(\ 2d_1 + d_2 + d_3 - d_4 - d_5) + R^* \\
R_2^* &= \tfrac{1}{8}(-2d_1 - d_2 - d_3 + d_4 + d_5) + R^* \\
X_1^* &= \tfrac{1}{8}(\ -3d_2 - d_3 - 3d_4 - d_5 + 2d_6) + R^* \\
X_2^* &= \tfrac{1}{8}(\ -d_2 - 3d_3 - d_4 - 3d_5 - 2d_6) + R^*
\end{aligned} \tag{9}$$

We rewrite the solutions in terms of model (8) and collect error terms to obtain

$$\begin{aligned}
R_1^* &= R_1^* + \tfrac{1}{8}(\ 4\delta_{R_1} - 4\delta_{R_2} + 2\varepsilon_1 + \varepsilon_2 + \varepsilon_3 - \varepsilon_4 - \varepsilon_5) \\
R_2^* &= R_2^* + \tfrac{1}{8}(-4\delta_{R_1} + 4\delta_{R_2} - 2\varepsilon_1 - \varepsilon_2 - \varepsilon_3 + \varepsilon_4 + \varepsilon_5) \\
X_1^* &= X_1^* + \tfrac{1}{8}(-4\delta_{R_1} - 4\delta_{R_2} + 8\delta_{X_1} - 3\varepsilon_2 - \varepsilon_3 - 3\varepsilon_4 - \varepsilon_5 + 2\varepsilon_6) \\
X_2^* &= X_2^* + \tfrac{1}{8}(-4\delta_{R_1} - 4\delta_{R_2} + 8\delta_{X_2} - \varepsilon_2 - 3\varepsilon_3 - \varepsilon_4 - 3\varepsilon_5 - 2\varepsilon_6)
\end{aligned} \tag{10}$$

Associated standard deviations are found from (10) as follows[3]:

$$\sigma_{R_1} = \sigma_{R_2} = (\tfrac{1}{2}\sigma_b^2 + \tfrac{1}{8}\sigma_w^2)^{1/2}$$

and (11)

$$\sigma_{X_1} = \sigma_{X_2} = (\tfrac{3}{2}\sigma_b^2 + \tfrac{3}{8}\sigma_w^2)^{1/2}.$$

The structure of (11) indicates how precision depends on the relationship between the components of error. For all four estimates, the contribution to the total variance from σ_b^2 is four times larger than the contribution from σ_w^2; thus, the size of σ_b relative to σ_w determines to what extent precision is affected by the number of design points.

Check Standard

The quantity σ_b can only be estimated from many designs involving the same artifact. Because calibrations are usually performed on a one-time basis, the prerequisite data for this analysis does not usually exist on the unknown itself. Thus, we designate a check standard for this purpose, and values of the check standard from many designs provide the basis for estimating σ_b.

For designs involving two reference standards, we create a check standard based on the difference between the two reference standards. For the design of (8), this difference

$$C = R_1^* - R_2^* \tag{12}$$

which is independent of the restraint, has an error structure of the form

$$C = R_1^* - R_2^* + \tfrac{1}{8}(8\delta_{R_1} - 8\delta_{R_2} + 4\varepsilon_1 + 2\varepsilon_2 + 2\varepsilon_3 - 2\varepsilon_4 - 2\varepsilon_5). \tag{13}$$

with associated standard deviation

$$\sigma_C = (2\sigma_b^2 + \tfrac{1}{2}\sigma_w^2)^{1/2}. \tag{14}$$

Hence

$$\sigma_b = (\tfrac{1}{2}\sigma_C^2 - \tfrac{1}{4}\sigma_w^2)^{1/2} \tag{15}$$

and (11) can be reduced to

$$\sigma_{R_1} = \sigma_{R_2} = \tfrac{1}{2}\sigma_C$$

and (16)

$$\sigma_{X_1} = \sigma_{X_2} = \frac{\sqrt{3}}{2}\sigma_C.$$

Estimates of Standard Deviations from the Data

Given n designs with check standard values $C_1, \ldots, C_n$, the quantity σ_C is estimated with (n − 1) degrees of freedom by

$$s_C = \left(\frac{1}{n-1}\sum_{i=1}^{n}(C_i - \bar{C})^2\right)^{1/2} \tag{17}$$

where $\bar{C}$ is the average of the check standard values[4].

The standard deviation, σ_w, is estimated from a single design with (m − k + 1) degrees of freedom where m is the number of comparisons in the design; k is the number of artifacts; and the additional degree of freedom comes from the known value of the restraint. For the design given by (8), the standard deviation σ_w is estimated with three degrees of freedom by

$$s_w = \left(\frac{1}{3}\sum_{i=1}^{6}(d_i - \hat{d}_i)^2\right)^{1/2} \tag{18}$$

[3] These equations are valid where R^* is known without random error: see the section headed, "A Matrix Approach", for the case where R^* is subject to random error

[4] This method of estimating the standard deviation assumes that the check standard is not drifting over time

where d_i is the predicted value for each difference measurement from the design; i.e.,

$$\begin{aligned} d_1 &= R_1^* - R_2^* \\ d_2 &= R_1^* \qquad - X_1^* \\ d_3 &= R_1^* \qquad\qquad - X_2^* \\ d_4 &= \qquad R_2^* - X_1^* \\ d_5 &= \qquad R_2^* \qquad - X_2^* \\ d_6 &= \qquad\qquad X_1^* - X_2^* \end{aligned}$$

We can improve the estimate of σ_w by pooling the standard deviations $s_{w_1}, \ldots, s_{w_n}$ from the n designs. The pooled value s_p, which has 3 n degrees of freedom, is computed as

$$s_p = \left(\frac{1}{n}\sum_{i=1}^{n} s_{w_i}^2\right)^{1/2}. \tag{19}$$

For the purpose of making statements of precision or uncertainty the population standard deviations σ_w, σ_b and σ_C are replaced by their respective estimates in the appropriate equations.

Process Control

Two aspects of statistical process control are relevant in the calibration process. Short-term control for measurements constituting a single design depends on σ_w, and long-term control for calibrations over time depends on σ_b via check standard measurements. The latter depends upon reliable estimates from historical data for the mean, $\bar{C}$, and the standard deviation, s_C. For any new calibration, the check standard value, C, is tested for agreement with past data by a t statistic where

$$t = \frac{|C - \bar{C}|}{s_C}.$$

The process is judged to be in control if

$$t \le t_{\alpha/2}(\nu)$$

where $t_{\alpha/2}(\nu)$ is the upper $\alpha/2$ percentage point of Student's t distribution [14] with ν degrees of freedom. Otherwise, the calibration is discarded.

Short-term control for each design is exercised by comparing the standard deviation from the design, s_w, with a pooled value s_p from historical data. An F statistic is computed as

$$F = s_w^2 / s_p^2.$$

Short-term precision is regarded as being in control if

$$F \le F_\alpha(\nu_1, \nu_2)$$

where $F_\alpha(\nu_1, \nu_2)$ is the upper α percentage point of Snedecor's F distribution [15] with ν_1 degrees of freedom in s_w and ν_2 degrees of freedom in s_p. Failure to meet this condition is taken as an indication that precision has deteriorated, and the current calibration results are discarded.

Case Study From Mass Calibration

The National Institute of Standards and Technology (NIST) maintains about thirty check standards for mass calibrations. These check standards, which cover a variety of designs, load levels, and balances, constitute the data base for constructing uncertainties associated with mass calibrations and for implementing statistical control of the calibration process.

The data base, which covers the last twenty years of calibration history at NIST, is reviewed on an annual basis to update uncertainty statements and to expose any trends or anomalies in the process. Standard deviations from the designs, s_w, are pooled by balance. Standard deviations for each check standard, s_C, are estimated by (17).

Analysis confirms that the long-term component of error, s_b, is negligible for the mass-calibration process except at the critical kilogram level. The majority of mass calibrations at NIST start at the kilogram level using the design of (8) with the restraint as the average of two reference kilograms and a check standard C as defined by (12). Standard deviations for this process are shown in the table below.

Standard Deviations at the Kilogram Level

Source	Notation	Eq.	Std. dev.
Kg balance	s_p	(19)	0.0316 mg
Check standard	s_C	(17)	0.0277 mg
Long-term change	s_b	(15)	0.0116 mg
Unknowns	s_{X_1}, s_{X_2}	(16)	0.0240 mg

Weights other than kilograms are related to the NIST unit of mass via a hierarchy of designs where the restraint for each design is taken from the solution to the previous design. For example, at the kilogram level, the unknown X_2 is a group of weights totaling a kilogram; the group constitutes the starting restraint for the next design in the series. Thus, any random error that influences the value assigned to X_2 is propagated to all other weights.

Application to Other Designs

The standard deviation associated with a measurement must be defined on a design-by-design basis. A matrix approach is outlined in the next section; also see Croarkin [16, 17] for specific formulations for a design involving two reference standards and three unknowns and a design involving four reference standards and four unknowns.

The problem of definition can sometimes be avoided by judicious choice of a check standard. If one chooses a check standard with the same error structure as the artifacts being calibrated, then the standard deviation for the check standard also applies to the calibrated artifacts. For example, if we make all ten comparisons among five artifacts of the same nominal value, where one artifact is a designated check standard, then the check standard will have the same error structure as the unknowns.

A Matrix Approach

A matrix approach is outlined for estimating components of variance for any measurement design where there are both short-term random errors of measurement and long-term random changes in the artifacts. We also allow for the situation where the restraint has been estimated from a previous experiment, and the random errors associated with that measurement process are taken into account.

Given m difference measurements among k artifacts, where some artifacts are regarded as reference standards and some are regarded as test items or unknowns, the model for the measurement process

$$D = A[X^* + \delta] + \varepsilon \tag{20}$$

is shown in terms of matrix elements. The elements and their respective dimensions are defined as follows:

- $\underset{(m \times 1)}{D}$ a matrix of difference measurements
- $\underset{(m \times k)}{A}$ a matrix of zeroes and ones such that a plus or minus one in the j^{th} position indicates that the j^{th} artifact is involved in the i^{th} comparison and a zero indicates the converse
- $\underset{(k+1)}{X^*}$ a matrix of unknown values for the k artifacts
- $\underset{(k \times 1)}{\delta}$ a matrix of random errors with zero mean and standard deviation σ_b
- $\underset{(m \times 1)}{\varepsilon}$ a matrix of random errors with zero mean and standard deviation σ_w

Because the matrix A has rank $(k-1)$, a solution for an unknown X^*, as shown by Zelen [18], is achieved by imposing upon the system a restraint, or known value for a linear combination of the artifacts. Let the scalar R^* be the restraint, and let $\mathscr{L}_R$ be a vector of zeroes and ones such that a one in the j^{th} position indicates that the j^{th} artifact is in the restraint and a zero indicates the converse.

For example, the vector [5]

$$\underset{(1 \times k)}{\mathscr{L}'_R} = (1 \quad 1 \quad 0 \ldots 0)$$

indicates that the *restraint* R^* is the *summation* for the first two artifacts.

Then a solution can be found from an augmented matrix B where

$$\underset{(k+2)\times(k+2)}{B} = \begin{pmatrix} \underset{(k \times k)}{A'A} & \underset{(k \times 1)}{\mathscr{L}_R} & \underset{(k \times 1)}{A'D} \\ \underset{(1 \times k)}{\mathscr{L}'_R} & \underset{(1 \times 1)}{0} & \underset{(1 \times 1)}{R^*} \\ \underset{(1 \times k)}{0} & \underset{(1 \times 1)}{0} & \underset{(1 \times 1)}{-I} \end{pmatrix}$$

has an inverse of the form

$$\underset{(k+2)\times(k+2)}{B^{-1}} = \begin{pmatrix} \underset{(k \times k)}{Q} & \underset{(k \times 1)}{h} & \underset{(k \times 1)}{X^*} \\ \underset{(1 \times k)}{h'} & \underset{(1 \times 1)}{0} & \underset{(1 \times 1)}{\bullet} \\ \underset{(1 \times k)}{\bullet} & \underset{(1 \times 1)}{\bullet} & \underset{(1 \times 1)}{\bullet} \end{pmatrix}$$

and Q is the covariance matrix; X^* is the vector of estimates for the unknowns; and other entries (•) are irrelevant for this application.

The deviations from the fit are given by the vector ζ where

$$\underset{(1 \times m)}{\zeta'} = [D - AX^*]'$$

and the standard deviation for the design σ_w is estimated by

$$s_w = \left(\frac{\zeta'\zeta}{m-k+1}\right)^{1/2}$$

with $m-k+1$ degrees of freedom.

It is now assumed that a check standard C is tracked for many applications of the same design over time. The estimated value of C for any particular design is given by

$$C = \mathscr{L}'_C[X^*]$$

where, for example,

$$\underset{(1 \times k)}{\mathscr{L}'_C} = (1 \quad -1 \quad 0 \ldots 0)$$

indicates that the check standard is the computed difference between the first and second artifacts.

[5] The mark (′) indicates the transpose of a matrix

The standard deviation σ_b can be estimated from the relationship

$$\sigma_b^2 = \frac{\sigma_C^2 - \mathscr{L}_C'[Q]\,\mathscr{L}_C\,\sigma_w^2}{\mathscr{L}_C'[Q\,A'\,A]\,\mathscr{L}_C}$$

where σ_C and σ_w should be estimated from the data of several designs[6].

Now consider a single unknown X_J whose estimated value is

$$X_J^* = [\mathscr{L}_{X_J}]'[X^*] \tag{21}$$

where, for example,

$$\underset{(1\times k)}{\mathscr{L}_{X_J}'} = (0 \quad 1 \quad 0 \ldots 0)$$

signifies that X_J refers to the second artifact in the design. Then the appropriate standard deviation for X_J^* is given by

$$\sigma_{X_J} = ([\mathscr{L}_{X_J}]'[Q\,A'\,A][\mathscr{L}_{X_J}]\,\sigma_b^2 + [\mathscr{L}_{X_J}]'[Q]\,[\mathscr{L}_{X_J}]\,\sigma_w^2)^{1/2} \tag{22}$$

and the standard deviation associated with any linear combination of the unknowns is computed in a similar fashion. At this stage we assume that R^* is known without random error. Eq. (25) is appropriate if this assumption is not valid.

Mass calibration is a special case because values are assigned to sets of weights covering several denominations of mass. All values are related to a starting restraint, such as a kilogram reference standard, by a series of interrelated designs. The first series includes as an unknown, a single weight or a summation of weights, which becomes the restraint for the following series and so on throughout the entire weight set. Thus, we must account for imprecision associated with restraints after the first series.

Let $\mathscr{L}_\Sigma$ be a $(k \times 1)$ vector that defines the unknown whose value will be used as the restraint in the next series; this out-going restraint has value

$$X_\Sigma = [\mathscr{L}_\Sigma]'[X^*]\,. \tag{23}$$

The standard deviation associated with this restraint is computed as

$$\sigma_{X_\Sigma} = ([\mathscr{L}_\Sigma]'[Q\,A'\,A]\,[\mathscr{L}_\Sigma]\,\sigma_b^2 + [\mathscr{L}_\Sigma]'[Q]\,[\mathscr{L}_\Sigma]\,\sigma_w^2)^{1/2}\,. \tag{24}$$

To account for weights of various denominations, let W be a vector of nominal values for the k weights of the second series so that

$$\underset{(1\times k)}{W'} = (W_1, \ldots, W_k)\,.$$

Now we redefine the design matrix A and the restraint vector $\mathscr{L}_R$ for the next series and let

$$R^* = X_\Sigma$$

and

$$\sigma_R = \sigma_{X_\Sigma}\,.$$

The matrix B and its inverse B^{-1} follow accordingly. The $\mathscr{L}_{X_J}$ vectors are also redefined for the weights in the series so that estimates can be computed according to (21). Then the appropriate standard deviation for the j[th] weight, X_J, is given by (25) which follows:

$$\sigma_{X_J} = \left[[\mathscr{L}_{X_J}]'[Q\,A'\,A]\,[\mathscr{L}_{X_J}]\,\sigma_b^2 + [\mathscr{L}_{X_J}]'[Q]\,[\mathscr{L}_{X_J}]\,\sigma_w^2 + \left(\frac{[\mathscr{L}_{X_J}]'\,W}{[\mathscr{L}_R]'\,W}\right)^2 \sigma_R^2\right]^{1/2} \tag{25}$$

Standard deviations for the check standard for this series and other combinations of weights are computed similarly. It is noted that the process standard deviations, σ_w and σ_b, depend on the balance and the denominations of weights calibrated in the series; thus, they should be estimated separately for each series.

The process is extended to the next series by redefining the vector $\mathscr{L}_\Sigma$ so that it identifies the outgoing restraint whose value is given by (23). Then the standard deviation for this restraint is given by (26) which follows:

$$\sigma_{X_\Sigma} = \left[[\mathscr{L}_\Sigma]'[Q\,A'\,A]\,[\mathscr{L}_\Sigma]\,\sigma_b^2 + [\mathscr{L}_\Sigma]'[Q]\,[\mathscr{L}_\Sigma]\,\sigma_w^2 + \left(\frac{[\mathscr{L}_\Sigma]'\,W}{[\mathscr{L}_R]'\,W}\right)^2 \sigma_R^2\right]^{1/2} \tag{26}$$

The standard deviations given by (22) and (24) are appropriate for values estimated in the initial series of weighings where the starting restraint is a known value. For values assigned by subsequent series of weighings, the imprecision of the estimated restraint contributes a component to the total standard deviation. Thus, (25) and (26) are appropriate.

Concluding Remarks

The proposed error model is especially enlightening where short-term errors are related to instrumentation. Then long-term errors are the result of operating procedures or environmental changes which affect the artifacts over time but are reasonably constant in the short-term so as not to affect the standard deviation from the design. Thus, there is motivation for isolating the long-term component in order to ascer-

[6] See the discussion under "Check Standard" and Eqs. (16) and (18)

tain whether precision can be improved given current instrumentation.

Other models may prove more useful or descriptive for other situations. For example, for mass calibrations which deal with weights of the same nominal mass, it is reasonable to assume that random changes in the weights can be characterized by a single error distribution. However, for weights which are not of the same nominal mass, we would allow for errors proportional to mass or, perhaps, to surface area.

Finally, the analysis of the design for four artifacts demonstrates that improved precision cannot always be attained by increasing the number of measurements in the design. The relative magnitudes of σ_w and σ_b and their contribution to the total variance must be understood before one can improve precision.

Acknowledgments. The author wishes to thank Dr. Richard Davis of the National Institute of Standards and Technology, who is responsible for maintaining the NIST unit of mass, for his interest and advice on the subject at hand.

References

1. M. Grabe: Metrologia **14**, 143–146 (1978)
2. J. A. Hayford: *On the least-square adjustment of weighings*, U.S. Coast and Geodetic Survey Report for 1892, Appendix 10
3. J.-R. Benoît: *L'étalonnage des séries de poids*, Trav. Mém. BIPM vol. **13**, 1907
4. R. C. Bose, J. M. Cameron: J. Res. Natl. Bur. Stand., Sect. **B: 79**, 323–332 (1965)
5. R. C. Bose, J. M. Cameron: J. Res. Natl. Bur. Stand., Sect. **B: 71**, 149–160 (1967)
6. I. M. Chakravarti, K. V. Suryanarayana: J. Combin. Theory **13**, 426–431 (1964)
7. J. M. Cameron, W. G. Eicke: *Designs for the Surveillance of the Volt Maintained by a Small Group of Saturated Standard Cells*, Natl. Bur. Stand. (U.S.) Tech. Note 430 (1967)
8. J. M. Cameron, G. E. Hailes: *Designs for the Calibration of Small Groups of Standards in the Presence of Drift*, Natl. Bur. Stand. (U.S.) Tech. Note 844 (1974)
9. C. P. Reeve: *The Calibration of Indexing Tables by Subdivision*, Natl. Bur. Stand. (U.S.) report NBSIR 75-750 (1975)
10. J. M. Cameron, M. C. Croarkin, R. C. Raybold: *Designs for the Calibration of Standards of Mass*, Natl. Bur. Stand (U.S.) Tech. Note 952 (1977)
11. C. P. Reeve: *The Calibration of a Roundness Standard*, Natl. Bur. Stand. (U.S.) report NBSIR 79-1758 (1979)
12. C. P. Reeve: *The Calibration of Angle Blocks by Intercomparison*, Natl. Bur. Stand. (U.S.) report NBSIR 80-1967 (1980)
13. J. M. Cameron, M. C. Croarkin, R. C. Raybold: Ref. 10 (above), pp. 10–11
14. E. S. Pearson, H. O. Hartley (eds.): *Biometrika Tables for Statisticians, Vol. I* (Cambridge University Press, Cambridge 1956), p. 138
15. Ibid., pp. 157–163
16. M. C. Croarkin: *Measurement Assurance Programs, Part II: Development and Implementation*, Natl. Bur. Stand. (U.S.) Spec. Publ. 676-II (1985), pp. 54–58
17. Ibid., pp. 66–75
18. M. Zelen: *Linear Estimation and Related Topics*, in: *Survey of Numerical Analysis*, J. Todd (ed.) (McGraw Hill, New York 1962), pp. 563–565