Drift Eliminating Designs for Non-Simultaneous Comparison Calibrations

Ted Doiron

National Institute of Standards and Technology, Gaithersburg, MD 20899-0001

The effects of drift on calibrations carried out by comparison have been studied at the National Institute of Standards and Technology for many years, and a number of strategies have been introduced to combat these effects. One strategy, the use of comparison designs which are inherently immune to linear drift, was developed specifically for mass comparison measurements. These techniques, developed for simultaneous comparisons, are extended to the case of non-simultaneous comparisons, such as gage block calibrations, where each artifact is measured separately, and the comparison is made mathematically from the individual measurements.

Key words: calibration design; comparison; drift eliminating; least squares analysis; metrology.

1. Introduction

The sources of variation in measurements are numerous. Some of the sources are truly random noise, $1/f$ noise in electronic circuits for example. Usually the "noise" of a measurement is actually due to uncontrolled systematic effects such as instability of the mechanical setup or variations in the conditions or procedures of the test. Many of these variations are random in the sense that they are describable by a normal distribution. Like true noise in the measurement system, the effects can be reduced by making additional measurements.

Another source of serious problems, which is not random in nature is drift in the instrument readings. This effect cannot be minimized by additional measurement because it is not generally pseudo-random, but a nearly monotonic shift in the readings. In dimensional work the most important cause of drift is thermal changes in the equipment

techniques to address this problem of instrumental drift.

A simple example of the techniques for eliminating the effects of drift by looking at two different ways of comparing two gage blocks, one standard (A) and one unknown (B).

Scheme 1: A B A B

Scheme 2: A B B A

Now let us suppose we make the measurements regularly spaced in time, 1 time unit apart, and there is an instrumental drift of Δ. The actual readings (m_i) from scheme 1 are:

$$m_1 = A \tag{1a}$$
$$m_2 = B + \Delta \tag{1b}$$
$$m_3 = A + 2\Delta \tag{1c}$$

Solving for B in terms of A we get:

$$B = A - \frac{1}{2}(m_1 + m_3 - m_2 - m_4) - \Delta \qquad (2)$$

which depends on the drift rate Δ.

Now look at scheme 2. Under the identical conditions the readings are:

$$m_1 = A \qquad (3a)$$
$$m_2 = B + \Delta \qquad (3b)$$
$$m_3 = B + 2\Delta \qquad (3c)$$
$$m_4 = A + 3\Delta \qquad (3d)$$

Here we see that if we add the second and third readings and subtract the first and fourth readings we find that the Δ drops out:

$$B = A - \frac{1}{2}(m_1 + m_4 - m_2 - m_3) \qquad (4)$$

Thus if the drift rate is constant — a fair approximation for most measurements if the time is properly restricted — the analysis both eliminates the drift and supplies a numerical approximation of the drift rate.

The calibration of a small number of "unknown" objects relative to one or two reference standards involves determining differences among the group of objects. Instrumental drift, due most often to temperature effects, can bias both the values assigned to the objects and the estimate of the effect of random errors. This paper presents schedules for sequential measurements of differences that eliminate the bias from drift and at the same time gives estimates of the magnitude of drift.

Previous works have [1,2] discussed schemes which eliminate the effects of drift for simultaneous comparisons of objects. For these types of measurements the difference between two objects is determined at one instant of time. Examples of these types of measurements are comparisons of masses with a double pan balance, comparison of standard voltage cells, and thermometers which are all placed in the same thermalizing environment. Many comparisons, especially those in dimensional metrology, cannot be done simultaneously. For example, using a gage block comparator, the standard, control (check standard) and test blocks are moved one at a time under the measurement stylus. For these comparisons each measurement is made at a different time. Schemes which assume simultaneous measurements will, in fact, eliminate

the drift from the analysis of the test objects but will produce a measurement variance which is drift dependent and an erroneous value for the drift, Δ.

Calibration designs involve differences between measured items so that unless one or more of them are standards for which values are known, one cannot assign values for the remaining "unknown" items. Algebraically, one has a system of equations that is not of full rank and needs the value for one item or the sum of several items as the restraint to lead to a unique solution. The least squares method used in solving these equations has been presented [3] and refined [4] in the literature and will not be repeated in detail here. The analyses of particular measurement designs presented later in this paper conform to the method and notation presented in detail by Hughes [3].

The schemes used as examples in this paper are those currently used at NIST for gage block comparisons. In our calibrations a control (check standard) is always used to generate data for our measurement assurance plan [5]. It is not necessary, however, to use a control in every design and the schemes can be used with any of the objects as the standard and the rest as unknowns. Schemes of various numbers of unknowns and measurements are presented in the Appendix.

2. Calibration Designs

The term calibration design has been applied to experiments where only differences between nominally equal objects or groups of objects can be measured. Perhaps the simplest such experiment consists in measuring the differences between the two objects of the $n(n-1)$ distinct pairings that can be formed from n objects. If the order is unimportant, X compared to Y is the negative of Y compared to X, and there are only $n(n-1)/2$ distinct pairings. Of course only one measurement per unknown is needed to determine the unknown, but many more measurements are generally taken for statistical reasons. Ordinarily the order in which these measurements are made is of no consequence. However, when the response of the comparator is time dependent, attention to the order is important if one wishes to minimize the effect of time.

When this effect can be adequately represented by a linear drift, it is possible to balance out the effect by proper ordering of the observations. The drift can be represented by the series, ... -3, -2, -1, 0, 1, 2, 3, ... if there are an odd number of

comparisons and by ... $-5/2$, $-3/2$, $-1/2$, $1/2$, $3/2$, $5/2$, ... if there are an even number of comparisons.

As an example let us take $n = 3$. If we make 12 measurements to obtain all possible $n(n-1)=6$ comparisons we get a scheme like that below, denoting the three objects by A, B, C.

Observation	Measurement	
m_1	$A - 6\Delta$	
m_2	$B - 5\Delta$	
m_3	$C - 4\Delta$	
m_4	$A - 3\Delta$	
m_5	$B - 2\Delta$	
m_6	$C - \Delta$	(5)
m_7	$A + \Delta$	
m_8	$C + 2\Delta$	
m_9	$B + 3\Delta$	
m_{10}	$A + 4\Delta$	
m_{11}	$C + 5\Delta$	
m_{12}	$B + 6\Delta$	

If we analyze these measurements by pairs, in analogy to the weighing designs of Cameron we see that:

$$
\begin{array}{llllcccc}
 & & & & A & B & C & \Delta \\
d_1 = m_1 - m_2 & = A - B - \Delta & +1 & -1 & & -1 \\
d_2 = m_3 - m_4 & = C - A - \Delta & -1 & & +1 & -1 \\
d_3 = m_5 - m_6 & = B - C - \Delta & & +1 & -1 & -1 & (6)\\
d_4 = m_7 - m_8 & = A - C - \Delta & +1 & & -1 & -1 \\
d_5 = m_9 - m_{10} & = B - A - \Delta & -1 & +1 & & -1 \\
d_6 = m_{11} - m_{12} & = C - B - \Delta & & -1 & +1 & -1 \\
\end{array}
$$

The notation used here, the plus and minus signs, indicate the items entering into the difference measurement. Thus, d_2 is a measurement of the difference between object C and object A.

The difference between the above design and that of a design for simultaneous comparisons in Ref. [2] is that the drift column is constant. It is simple to see by inspection that the drift is balanced out since each object has two ($+$) and two ($-$) measurements and the drift column is constant. By extension, the effects of linear drift is eliminated in all complete block measurement schemes (those for which all objects are measured in all possible $n(n-1)$ combinations).

Although all schemes in which each object has equal numbers of ($+$) and ($-$) measurements is drift eliminating, there are practical criteria which must be met for the scheme to work. First, the actual drift must be linear. For dimensional measurements the instrument drift is usually due to changes in temperature. The usefulness of drift eliminating designs depends on the stability of the thermal environment and the accuracy required in

the calibration. In the NIST gage block laboratory the environment is stable enough that the drift is linear at the 5 nm level over periods of 5 to 10 min. Our comparison plans are chosen so that the measurements can be made in this period.

Secondly, each measurement must be made in about the same amount of time so that the measurements are made at fairly regular intervals. In a completely automated system this is simple, but with human operators there is a natural tendency to make measurements simpler and quicker if the opportunity presents itself. For example, if the scheme has a single block measured two or more times in succession it is tempting to measure the object without removing it from the apparatus, placing it in its normal resting position, and returning it to the apparatus for the next measurement.

Finally, the measurements of each block are spread as evenly as possible across the design. Suppose in the scheme above where each block is measured four times block A is measured as the first measurement of d_1, d_2, d_3, and d_4. There is a tendency to leave block A near the measuring point rather than its normal resting position because it is used so often in the first part of the scheme. This allows block A to have a different thermal handling than the other blocks which can result in a thermal drift which is not the same as the other blocks.

3. Restraints

Because all of the measurements made in a calibration are relative comparisons, at least one value must be known to solve the system of equations. In the design of the last section, for example, if one has a single standard and two unknowns, the standard can be assigned to any one of the letters. (The same would be true of two standards and one unknown.) If there are two standards and one unknown, the choice of which pair of letters to assign for the standards can be important if all of the possible comparisons are not made. For full block designs (all possible comparisons are made) the uncertainty of the result does not depend on which letter is assigned to the standards or unknowns. For incomplete block designs the uncertainty of the results can depend on which letter the standard and unknowns are assigned. In these cases the customer blocks are assigned to minimize their variance and allow the larger variance for the measurement of the extra master (control).

This asymmetry occurs because every possible comparison between the four items has not been measured. For 4 objects there are 12 possible inter-comparisons. If an eight measurement scheme is used all three unknowns cannot be compared directly to the standard the same number of times. For example, two unknowns can be compared directly with the standard twice, but the other unknown will have no direct comparisons. This indirect comparison to the standard results in a slightly larger variance for the block compared indirectly. Complete block plans, which compare each block to every other block equal number of times, has no such asymmetry, and thus removes any restriction on the measurement position of the control.

4. Example: 4 block, 12 comparison, Single Restraint Design for NIST Gage Block Calibration [4]

The gage block comparison scheme put into operation in 1989 consists of two standards blocks, denoted S and C, and two customer blocks to be calibrated, denoted Y and Z. In order to decrease the random error of the comparison process a new scheme was devised, consisting of all 12 possible comparisons between the four blocks. Because of continuing problems making penetration corrections, the scheme was designed to use either the S or C block as the restraint and the difference $(S - C)$ as the control parameter. The S blocks are all steel, and are used as the restraint for all steel customer blocks. The C blocks are chrome carbide, and are used as the restraint for chrome and tungsten carbide blocks. The difference $(S - C)$ is independent of the choice of restraint.

We chose a complete block scheme which assures that the design is drift eliminating and the blocks can be assigned to the letters of the design arbitrarily. We chose $(S - C)$ as the first comparison. Since there are a large number of ways to arrange the 12 measurements for a complete block design, we added two restrictions as a guide to choose a "best" design.

1. It was decided to avoid schemes which measured the same block two or more times consecutively. There are many possible schemes where one or more blocks are measured twice consecutively. There is a great temptation to not remove and replace the blocks under these conditions. The analysis assumes that each measurement is made with the same motion and that the

measurements are evenly spaced in time. Consecutive repetition threatens both these assumptions.

2. We decided that schemes in which the six measurements of each block were spread out as evenly as possible in time would be less likely to be affected by small non-linearities in the drift. For example, for some schemes measurements of one block are completed by the 8th comparison, leaving the final 1/3 of the comparisons with no sampling of that block.

The new scheme is as follows:

		S	C	Y	Z	Δ	
$d_1 = S - C$		1	-1	0	0	-1	
$d_2 = Z - S$		-1	0	0	1	-1	
$d_3 = Y - Z$		0	0	1	-1	-1	
$d_4 = C - S$		-1	1	0	0	-1	
$d_5 = C - Y$		0	1	-1	0	-1	
$D = d_6 = Z - C$	$X =$	0	-1	0	1	1	(7)
$d_7 = S - Y$		1	0	-1	0	-1	
$d_8 = C - Z$		0	1	0	-1	-1	
$d_9 = S - Z$		1	0	0	-1	-1	
$d_{10} = Y - C$		0	-1	1	0	-1	
$d_{11} = Y - S$		-1	0	1	0	-1	
$d_{12} = Z - Y$		0	0	-1	1	-1	

When the S block is the restraint, with value L, the matrix equation to solve is:

$$\begin{vmatrix} X'X & a \\ a' & 0 \end{vmatrix} \begin{vmatrix} A \\ \lambda \end{vmatrix} = \begin{vmatrix} X'Y \\ L \end{vmatrix} \tag{8}$$

where $[A'] = [S\ C\ Y\ Z\ \Delta]$ is the vector to be estimated and $[a'] = [1\ 0\ 0\ 0\ 0]$ is the restraint vector.

$$\begin{vmatrix} X'X & a \\ a' & 0 \end{vmatrix} = \begin{vmatrix} 6 & -2 & -2 & -2 & 0 & 1 \\ -2 & 6 & -2 & -2 & 0 & 0 \\ -2 & -2 & 6 & -2 & 0 & 0 \\ -2 & -2 & -2 & 6 & 0 & 0 \\ 0 & 0 & 0 & 0 & 12 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{vmatrix} \tag{9}$$

$$\begin{vmatrix} B & h \\ h' & 0 \end{vmatrix} = \begin{vmatrix} X'X & a \\ a' & 0 \end{vmatrix}^{-1} = (1/24) \begin{vmatrix} 0 & 0 & 0 & 0 & 0 & 24 \\ 0 & 6 & 3 & 3 & 0 & 24 \\ 0 & 3 & 6 & 3 & 0 & 24 \\ 0 & 3 & 3 & 6 & 0 & 24 \\ 0 & 0 & 0 & 0 & 2 & 0 \\ 24 & 24 & 24 & 24 & 0 & 0 \end{vmatrix} \tag{10}$$

The estimate for A can be written as

$$<A> = \frac{|BX'\ h|\ |D|}{|L|} = 1/24 \begin{vmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 24 \\ -6 & 3 & 0 & 6 & 3 & -3 & -3 & 3 & -3 & -3 & 3 & -3 & 24 \\ -3 & 3 & 3 & 3 & 3 & 0 & -6 & 0 & -3 & 3 & 6 & -3 & 24 \\ -3 & 6 & -3 & 3 & 0 & 3 & -3 & -3 & -6 & 0 & 3 & 3 & 24 \\ -2 & -2 & -2 & -2 & -2 & -2 & -2 & -2 & -2 & -2 & -2 & -2 & 0 \end{vmatrix} \quad (11)$$

This leads to estimated values for S, C, Y, Z, and Δ which can be written as a function of the comparison measurements:

$$<S> = L \quad (12a)$$
$$<C> = (1/8)(-2d_1 + d_2 + 2d_4 + d_5 - d_6 - d_7 + d_8 - d_9 - d_{10} + d_{11}) + L \quad (12b)$$
$$<Y> = (1/8)(-d_1 + d_2 + d_3 + d_4 - d_5 - 2d_7 - d_9 + d_{10} + 2d_{11} - d_{12}) + L \quad (12c)$$
$$<Z> = (1/8)(-d_1 + 2d_2 - d_3 + d_4 + d_6 - d_7 - d_8 - 2d_9 + d_{11} + d_{12}) + L \quad (12d)$$
$$<\Delta> = (-1/12)(d_1 + d_2 + d_3 + d_4 + d_5 + d_6 + d_7 + d_8 + d_9 + d_{10} + d_{11} + d_{12}) + L \quad (12e)$$

The deviations of the measured values from the estimated values, $\epsilon_1, \epsilon_2, \ldots, \epsilon_{12}$ can be determined from the equations above, or can be calculated directly using matrix methods. For example,

$$\epsilon_1 = (d_1)_{observed} - (<S> - <C>). \quad (13)$$

These deviations provide the information needed to obtain a value, s_w, which is the experimental estimate value for the short term process standard deviation, or within standard deviation σ_w.

$$s^2 = \sum_{i=1}^{n} \frac{\delta_i^2}{(n-m+1)} \quad (14)$$

The number of degrees of freedom, $(n-m+1)$, results from taking the number of observations $(n=12)$ less the number of unknowns $(m=5; S, C, Y, Z, \Delta)$, and then adding one for the restraint. Because of the complete block structure (all 12 possible combinations measured) the standard deviations of the predicted values for the three unknowns are the same:

$$s_C = s_Y = s_Z = (1/2)s . \quad (15)$$

The standard deviation of the predicted drift is

$$S_\delta = \sqrt{\frac{1}{12}s} \cdot \quad (16)$$

5. Process Control

5.1 F-Test

Continued monitoring of the measurement process is required to assure that predictions based on the accepted values for process parameters are still valid. For gage block calibration at NIST, the process is monitored for precision by comparison of the observed standard deviation, s_w, to the average of previous values, σ_w. For this purpose the value of s_w is recorded for every calibration done, and this data set is periodically analyzed to provide an updated value of the accepted process σ_w for each gage block size.

The comparison is made using the F distribution, which governs the comparison of variances. The ratio of the variances s^2 (derived from the model fit to each calibration) and σ_w^2 derived from the history is compared to the critical value $F(8, \infty, \alpha)$, which is the upper α probability point of the F distribution for degrees of freedom 8 and ∞. For calibrations at NIST, α is chosen as 0.01 to give $F(8, \infty, .01) = 2.5$. The condition to be checked is:

$$F = \frac{s^2_{obs}}{\sigma_w^2} < 2.5 . \quad (17)$$

If this condition is violated the calibration fails, and is repeated. If the calibration fails more than once the test blocks are re-inspected and the instrument checked and recalibrated. All calibrations, pass or fail, are entered into the history file.

5.2 t-Test

At NIST a control measurement is made with each calibration by using two known master blocks in each calibration. One of the master blocks is steel and the other chrome carbide. When a customer block is steel the steel master is used as the restraint, and when a customer block is carbide, the carbide master is used as the restraint. The use of a control measurement for calibrations is necessary in order to provide assurance of the continuing accuracy of the measurements. The F-test, while providing some process control, only attempts to control the repeatibility of the process, not the accuracy. The use of a control is also the easiest method for finding the long term variability of the measurement process [5].

While the use of a control in each calibration is not absolutely necessary, the practice is highly recommended. There are systems which use intermittent tests, e.g., measures a control set once a week [4]. This is a good strategy for automated systems because the chances of block to block operator errors is reduced greatly. For manual measurements the process variability, and of course the occurence of operator error is much higher.

The check for systematic error is given by comparison of the observed value of the difference between the standard and control blocks. If S is the standard it becomes the restraint, and if A is used as the control $(S - A)$ is the control parameter for the calibration. This observed control is recorded for every calibration, and is used to periodically update the accepted, or average value, of the control. The process control step involves the comparison of the observed value of the control to the accepted (historical) value. The comparison is made using the Student t-distribution.

The control test demands that the observed difference between the control and its accepted value be less than 3 times the accepted long term standard deviation, σ_t, of the calibration process. This value of the t-distribution implies that a good calibration will not be rejected with a confidence level of 99.7%. The condition to be checked is:

$$ t = \frac{\left| (<S> - <C> - (S-C)_{acc} \right|}{\sigma_t} < 3 . \qquad (18) $$

The value of σ_t is obtained directly from the sequence of values of $(<S> - <C>)$ arising in regular calibrations. The recorded $(<S> - <C>)$ as the restraint, and $<S> - <C>$ is again used as the control.

If both the precision (F-test) and accuracy (t-test) criteria are satisfied, the process is regarded as being "in control" and values for the unknowns, Y and Z, and their associated uncertainties are regarded as valid. Failure of either criterion is an "out-of-control" signal and the measurements are repeated.

The value for drift, (Δ), serves as an indicator or possible trouble if it changes markedly from its usual range of values. However, because any linear drift is balanced out, a change in the value does not of itself invalidate the result.

6. Conclusion

The choice of the order of comparisons is an important facet of calibrations, in particular if chosen properly the comparison scheme can be made immune to linear drifts in the measurement equipment. The idea of making a measurement scheme robust is a powerful one. What is needed to implement the idea is an understanding of the sources of variability in the measurement system. While such a study is sometimes difficult and time consuming because of the lack of reference material concerning many fields of metrology, the NIST experience has been that such efforts are rewarded with measurement procedures which, for about the same amount of effort, produce higher accuracy.

7. Appendix A. Selection of Other Drift Eliminating Designs

The following designs can be used with or without a control block. The standard block is denoted S, and the unknown blocks A, B, C, etc. If a check standard block is used it can be assigned to any of the unknown block positions. The name of the design is simply the number of blocks in the design and the total number of comparisons.

3 - 6 Design (One master block, 2 unknowns 4 measurements each)		3 - 9 Design (One master block, 2 unknowns 6 measurements each)	
y_1 =	$S - A$	y_1 =	$S - A$
y_2 =	$B - S$	y_2 =	$B - A$
y_3 =	$A - B$	y_3 =	$S - B$
y_4 =	$A - S$	y_4 =	$A - S$
y_5 =	$B - A$	y_5 =	$B - S$
y_6 =	$S - B$	y_6 =	$A - B$
		y_7 =	$A - S$
		y_8 =	$B - A$
		y_9 =	$S - B$

4 - 8 Design
(One master block,
3 unknowns
4 measurements each)

$$y_1 = S - A$$
$$y_2 = B - C$$
$$y_3 = C - S$$
$$y_4 = A - B$$
$$y_5 = A - S$$
$$y_6 = C - B$$
$$y_7 = S - C$$
$$y_8 = B - A$$

4 - 12 Design
(One master block,
3 unknowns
6 measurements each)

$$y_1 = S - A$$
$$y_2 = C - S$$
$$y_3 = B - C$$
$$y_4 = A - S$$
$$y_5 = A - B$$
$$y_6 = C - A$$
$$y_7 = S - B$$
$$y_8 = A - C$$
$$y_9 = S - C$$
$$y_{10} = B - A$$
$$y_{11} = B - S$$
$$y_{12} = C - B$$

9 - 18 Design
(One master block,
7 unknowns
4 measurements each)

$$y_1 = S - A$$
$$y_2 = H - F$$
$$y_3 = A - B$$
$$y_4 = D - C$$
$$y_5 = E - G$$
$$y_6 = C - A$$
$$y_7 = B - F$$
$$y_8 = G - H$$
$$y_9 = D - S$$
$$y_{10} = C - E$$
$$y_{11} = H - S$$
$$y_{12} = G - D$$
$$y_{13} = C - S$$
$$y_{14} = A - C$$
$$y_{15} = F - D$$
$$y_{16} = S - H$$
$$y_{17} = E - B$$
$$y_{18} = F - G$$

10 - 20 Design
(One master block,
9 unknowns
4 measurements each)

$$y_1 = S - A$$
$$y_2 = F - G$$
$$y_3 = I - C$$
$$y_4 = D - E$$
$$y_5 = A - H$$
$$y_6 = B - C$$
$$y_7 = G - H$$
$$y_8 = I - S$$
$$y_9 = E - F$$
$$y_{10} = H - I$$
$$y_{11} = D - F$$
$$y_{12} = A - B$$
$$y_{13} = C - I$$
$$y_{14} = H - E$$
$$y_{15} = B - G$$
$$y_{16} = S - D$$
$$y_{17} = F - B$$
$$y_{18} = C - D$$
$$y_{19} = G - S$$
$$y_{20} = E - A$$

5 - 10 Design
(One master block,
4 unknowns
4 measurements each)

$$y_1 = S - A$$
$$y_2 = D - C$$
$$y_3 = S - B$$
$$y_4 = D - A$$
$$y_5 = C - B$$
$$y_6 = A - C$$
$$y_7 = B - S$$
$$y_8 = B - D$$
$$y_9 = C - S$$
$$y_{10} = A - D$$

6 - 12 Design
(One master block,
5 unknowns
4 measurements each)

$$y_1 = S - A$$
$$y_2 = D - C$$
$$y_3 = E - B$$
$$y_4 = E - D$$
$$y_5 = C - A$$
$$y_6 = B - C$$
$$y_7 = S - E$$
$$y_8 = A - D$$
$$y_9 = A - B$$
$$y_{10} = D - S$$
$$y_{11} = B - E$$
$$y_{12} = C - S$$

11 - 22 Design
(One master block,
10 unknowns
4 measurements each)

$$y_1 = S - A$$
$$y_2 = D - E$$
$$y_3 = G - I$$
$$y_4 = C - H$$
$$y_5 = A - B$$
$$y_6 = I - J$$
$$y_7 = H - F$$
$$y_8 = D - S$$
$$y_9 = B - C$$
$$y_{10} = S - E$$
$$y_{11} = A - G$$
$$y_{12} = F - B$$
$$y_{13} = E - F$$
$$y_{14} = J - A$$
$$y_{15} = C - D$$
$$y_{16} = H - J$$
$$y_{17} = F - G$$
$$y_{18} = I - S$$
$$y_{19} = B - H$$
$$y_{20} = G - D$$
$$y_{21} = J - C$$
$$y_{22} = E - I$$

7 - 14 Design
(One master block,
4 unknowns
6 measurements each)

$$y_1 = S - A$$
$$y_2 = E - C$$
$$y_3 = B - D$$
$$y_4 = A - F$$
$$y_5 = S - E$$
$$y_6 = D - B$$
$$y_7 = A - C$$
$$y_8 = B - F$$
$$y_9 = D - E$$
$$y_{10} = F - S$$
$$y_{11} = E - A$$
$$y_{12} = C - B$$
$$y_{13} = C - S$$
$$y_{14} = F - D$$

8 - 16 Design
(One master block,
7 unknowns
4 measurements each)

$$y_1 = S - A$$
$$y_2 = E - G$$
$$y_3 = F - C$$
$$y_4 = D - S$$
$$y_5 = B - E$$
$$y_6 = G - F$$
$$y_7 = C - B$$
$$y_8 = E - A$$
$$y_9 = F - D$$
$$y_{10} = C - S$$
$$y_{11} = A - G$$
$$y_{12} = D - B$$
$$y_{13} = C - S$$
$$y_{14} = G - C$$
$$y_{15} = B - D$$
$$y_{16} = A - F$$

The choice of the order of comparisons is an important facet of calibrations, in particular if chosen properly the comparison scheme can be made immune to linear drifts in the measurement equipment. The idea of making a measurement scheme

robust is a powerful one. What is needed to implement the idea is an understanding of the sources of variability in the measurement system. While such a study is sometimes difficult and time consuming because of the lack of reference material concerning many fields of metrology, the NIST experience has been that such efforts are rewarded with measurement procedures which, for about the same amount of effort, produce higher accuracy.

Acknowledgment

I would like to acknowledge Mr. Chien-ming Wu of the Industrial Technology Research Institute and Ms. Carroll Croarkin of NIST for reading the draft paper and suggesting many useful changes and corrections.

8. References

[1] J. M. Cameron, M. C. Croarkin, and R. C. Raybold, Designs for the Calibration of Standards of Mass, Natl. Bur. Stand. (U.S.) Technical Note 952, 1977.

[2] J. M. Cameron and G. E Hailes, Designs for the Calibration of Small Groups of Standards in the Presence of Drift, Natl. Bur. Stand. (U.S.) Technical Note 844, 1974.

[3] C. G. Hughes, III and H. A. Musk, A Least Squares Method for Analysis of Pair Comparison Measurements, Metrologia 8, 109-113 (1972).

[4] C. Croarkin, An Extended Error Model for Comparison Calibration, Metrologia 26, 107-113 (1989).

[5] C. Croarkin, Measurement Assurance Programs, Part II: Development and Implementation, Natl. Bur. Stand. (U.S.) Special Publication 676-II, 1984.-

[6] Further examples are presented in detail in a larger version of this work to be published as a NIST Technical Note.

[7] M. Priel, Automated Gage Block Calibration, Proceedings of the Measurement Science Conference, Anaheim CA, 1989.

About the author: Ted Doiron is a physicist in the Precision Engineering Division at NIST. The National Institute of Standards and Technology is an agency of the Technology Administration, U.S. Department of Commerce.