JOURNAL OF AIRCRAFT
Vol. 44, No. 2, March–April 2007

Advancements in Aircraft Model Force and Attitude Instrumentation by Integrating Statistical Methods

Peter A. Parker* and Tom D. Finley†
NASA Langley Research Center, Hampton, Virginia 23681

DOI: 10.2514/1.23060

Applying statistical methods in conjunction with instrumentation expertise has resulted in a dramatic reduction in calibration time and expense, while simultaneously improving the calibration quality. In this paper, we illustrate the application of response surface methodology and statistical quality control to two quintessential instruments used in aeronautical wind-tunnel experiments, namely, the force balance and the triaxial accelerometer. We emphasize the benefits that have been achieved by integrating the statistical design with the mechanical calibration system. For both instruments, we discuss the development of an experimental design that accommodates physics-based constraints and highlight an innovative calibration apparatus. As a result of reduced calibration time, the frequency of calibration can be increased, which enables the monitoring of instrument stability over time. Throughout the calibration process, we emphasize efficient allocation of experimental resources to achieve the calibration requirements.

Nomenclature

A_x	=	azimuth angle
b	=	accelerometer bias
k	=	number of experimental factors
g	=	gravitational component
$\boldsymbol{g}$	=	vector of gravitational components
R	=	roll angle
S	=	accelerometer sensitivity
$\mathbf{T}$	=	matrix of temperature compensated sensitivities
v	=	voltage response
$\boldsymbol{v}$	=	vector of voltage responses
x	=	explanatory variable
Y	=	yaw angle
y	=	response variable
α	=	pitch angle, false detection probability
β	=	model parameter
$\boldsymbol{\beta}$	=	matrix of model parameters
ϵ	=	random error
$\boldsymbol{\epsilon}$	=	vector of random errors
σ^2_{est}	=	estimated variance
Ω	=	coning angle

Subscripts

i, j	=	factor indicesx
x, y, z	=	axes

Superscript

t	=	temperature compensated

Presented as Paper 7603 at the U.S. Air Force T&E Days Conference, Nashville, TN, 6–8 December 2005; received 8 February 2006; revision received 28 August 2006; accepted for publication 10 October 2006.

*Research Scientist, Aeronautics Systems Engineering Branch. Senior Member AIAA.

†Senior Research Scientist (retired), Aeronautics Systems Engineering Branch.

I Introduction

IN EXPERIMENTAL aeronautics, research questions are answered by estimating mathematical models that relate aerodynamic performance parameters to simulated flight conditions through wind-tunnel experiments. Clearly, the utility of these mathematical relationships is a function of the experimental design, the ability to set the simulated flight conditions, and the quality of the instrumentation. Until recently, instrumentation research has focused on improved sensor technology and data acquisition systems. However, the basic calibration approach has remained largely unchanged. In this paper, we consider advancements in calibration methodology through the application of statistical methods to force balances and attitude measurement systems based on triaxial accelerometers.

Calibration is the most critical phase in the production of a measurement system. During this phase, the performance characteristics, or quality, of the instrument is assessed. Oftentimes, calibration activities are costly and time consuming requiring a significant investment of resources. Unfortunately, even with the recognition of its criticality and allocation of resources, the quality and stability of the instrument may not be adequately characterized.

An instrument calibration is essentially an experiment where factors (or independent variables) are set and responses, usually in the form of electrical signals, are measured. The objectives of the calibration are to build a prediction-oriented mathematical model and to assess its quality. By viewing this process as an experiment, we propose a disciplined approach that employs a set of statistical techniques, known as response surface methodology (RSM), to efficiently design, execute, and analyze the calibration. Montgomery and Myers [1,2] present these methods in detail and discuss many successful applications to industrial and scientific process and product characterization and optimization. For an exposition on the application of RSM to wind-tunnel experimentation, see [3–5].

After the instrument is placed into wind-tunnel service, periodic recalibration is required. However, for many custom-designed instruments, the calibration interval is dictated by cost and time constraints, rather than an objective assessment of instrument performance. Because of infrequent calibration, a rigorous analysis of calibration stability is usually not possible. Replicated calibration has been proposed as a means of assessing long-term instrument performance [6]. However, in general, systematic methods to distinguish between subtle changes in performance and to diagnose their cause have received little attention.

As a result of more efficient calibration methods, there is an opportunity to increase the frequency of calibration. Replicated calibrations of an instrument over time provide valuable information

regarding its operational condition, precision, and stability. Applying a statistical quality control (SQC) method known as profile monitoring [7] provides a framework to analyze the replicated calibrations in a manner that is easily interpretable and well suited for diagnostic evaluation. Moreover, characterizing the stability of an instrument over time is an important component in the determination of adequate calibration intervals. For a general introduction to SQC, see [8]. For the application of statistical process monitoring to wind-tunnel experiments, see [9,10].

For force and attitude calibration, we discuss the development of an experimental design that accommodates physics-based constraints and highlight its integration with an innovative apparatus that reduces calibration time and expense. From our experience, the most significant improvements in calibration efficiency and quality have been achieved by focusing on integration, rather than a piecewise improvement approach. In addition, we illustrate two methods of statistical process monitoring that provide objective tools to monitor instrument performance over time, detect instability, and diagnose potential sources of instability. Comprehensive references are provided for the statistical and calibration methods.

II. Force-Balance Calibration

A force balance is a multiple-axis load cell that provides simultaneous measurement of six components of aerodynamic force and moment (normal, axial, and side force, and rolling, pitching, and yawing moments) exerted on a wind-tunnel test article. Note that the term force will be used subsequently to denote both forces and moments. The basic operating principle of a force balance is to transduce the applied forces into six electrical signals. Ideally, each signal would respond only to its respective component of force and be insensitive to others. This is not entirely possible, even though balance designs are structurally and electrically optimized to reduce these undesirable interaction effects. As a result, a calibration experiment is performed to generate a mathematical model of the force-response relationship. Even though the basic design of force balances has been largely unaltered since the 1940s [11,12], they remain as the state-of-the-art force measurement instrument used in wind-tunnel experiments.

A. Experimental Design

In this section, we highlight distinctive features of an experimental design based on response surface methodology. In calibration applications, the experimental design is commonly referred to as the load schedule or simply the calibration set points. A formally designed experiment considers the selection of the design points, prediction characteristics, execution protocol, analysis approaches supported by the design and protocol, and mathematical modeling focused on adequacy and parsimony. A detailed exposition of these topics is provided in [13].

In force-balance calibration, we model the relationship between the six force components and the six electrical signals, estimating the parameters of a six-dimensional response surface model for each force-balance channel. Traditionally, the model is based on a second-order Taylor's series approximation to the true underlying force-response relationship given by

$$y = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k} \beta_{ii} x_i^2 + \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} \beta_{ij} x_i x_j + \epsilon \qquad (1)$$

where y is the response (electrical signal), the βs are the unknown regression (or calibration) coefficients, the xs are the independent explanatory variables (applied forces), ϵ is a random error assumed to be independently and identically distributed from a normal distribution with a mean of zero and constant variance, and k is the number of variables. For a force balance, there are six explanatory variables ($k = 6$), resulting in 28 terms in a complete second-order model for each response.

There are three fundamental quality-assurance principles employed in RSM, namely, randomization, blocking, and replication. Randomization of the execution order defends against systematic errors and supports the assumed statistical independence of the random error term, ϵ, in Eq. (1).

Blocking entails organizing an experiment into relatively short blocks of time within which the randomization of point ordering is performed. Typical block boundaries are defined by operators, work shifts, or days. Although randomization defends against systematic within-block variations, the effects of between-block systematic variations can be removed in the analysis of the data, if the experiment is properly blocked. Removing this component of the unexplained variance allows for more precise estimation of model parameters.

Replication is performed throughout the experiment by randomly allocating genuine replicates of specific design points. In contrast to replication, repeated observations are obtained while holding the variables constant rather than resetting their levels between successive data points. Replication provides unbiased estimates of experimental error, known as pure error, which is the component attributable to ordinary chance variations in the data. Estimates of pure error enable objective assessments of the quality of fit, or adequacy, of the mathematical model. In addition, by replication we can increase the precision of the model parameter estimates.

When designing an experiment, we first specify the design space, which refers to the extreme levels of the force components that can be safely applied during the calibration and are considered to be representative of the wind-tunnel operational envelope. A design point refers to the setting of a specific combination of forces and moments at which a single observation of the responses is obtained. A collection of design points is specified to provide high estimation and prediction efficiency [2], resulting in adequate data volume rather than excessive.

An additional set of points is chosen to test the prediction capability of the calibration model, referred to as confirmation points. We use the term prediction in the sense that in the wind-tunnel experiment, the force-balance signals (responses) are measured and the calibration model is used to predict the aerodynamic forces (explanatory variables) through an inverse mathematical model. To select the confirmation points, we randomly choose locations within the design space. If specific knowledge exists about likely combinations of forces during the wind-tunnel experiment, then those points should be included in the confirmation set.

To illustrate the RSM approach, consider an experimental design that has been implemented at NASA Langley Research Center. (For other proposed force-balance experimental designs, see [14–16].) This design features a total of 64 design points consisting of 55 unique combinations of the six force components and 10 replicated center points. A center point denotes an observation without an external calibration force applied. The design is divided into 2 blocks, with the first containing 32 factorial combinations of the six force components at approximately 41% of full scale, based on a spherical design space. The second block contains 22 design points with single- and two-component combinations with the forces at their full-scale level. The combined design in the two blocks has a compositelike structure and was inspired by the central composite design [2]. The blocking boundaries were determined by 8-hour work shifts. Finally, an additional block of 20 confirmation points is performed. Within each block, the ordering of the design points is completely randomized.

Once the calibration is complete, the statistical analysis methods employed enhance the experience of the balance engineer to objectively estimate the model coefficients. For example, the number of model terms is minimized by eliminating those that are too small to resolve with a sufficiently high level of confidence, resulting in a parsimonious model. Also, the total unexplained variance is partitioned into the pure-error and lack-of-fit components to determine the adequacy of the model and the potential for improvement with higher order terms. Once a model has been built, graphical diagnostic techniques are employed to check the statistical assumptions and to assess the quality of the model fit. For more details on these analysis methods, see [2].

Employing RSM design and analysis techniques provides new insights into the calibration results and empowers the balance engineer to objectively analyze and report the results in a manner that is considerably more defendable than previous approaches. Theoretical and experimental results have shown that an RSM approach makes it possible to perform force-balance calibration with an order of magnitude fewer design points than historical methods [13,15].

B. Single-Vector System

To execute the experimental design, a precision mechanical system is required to set the force components (independent variables). The ability to set randomly ordered multicomponent force combinations is difficult for traditional manual systems [16]. Alternatively, studies have shown that automated systems can significantly reduce the calibration time; however, there is a deterioration in accuracy compared to manual systems [17].

Therefore, a new calibration system has been implemented at NASA Langley Research Center, known as the Single-Vector System (SVS), to overcome the weaknesses of previous systems and to be fully integrated with the requirements of RSM, particularly randomly ordered multicomponent force application. A photograph of the SVS is provided in Fig. 1, and a complete description including a comparison to a manual calibration system and an uncertainty analysis is available in [15,18].

The SVS allows for single-vector calibration, meaning that single, calibrated deadweight loads are applied in the gravitational direction generating six component force combinations relative to the balance coordinate system. By using a single force vector, load application inaccuracies caused by the conventional requirement to generate multiple force vectors are fundamentally reduced. The SVS features significantly fewer components than other systems, and therefore fewer sources of systematic error [18]. Overall, the SVS improves on the trusted aspects of manual calibration systems, requires minimal time to operate, and increases calibration accuracy.

To generate a desired combination of the three forces, the force balance is manipulated to a prescribed orientation using a nonmetric positioning system. Its attitude relative to the Earth's gravitational vector is precisely measured on the metric end using a triaxial accelerometer system that provides the components of the gravitational vector projected onto the balance coordinate system. Combining the measured gravitational components and the magnitude of the deadweight enables the determination of the three force components. The three moments are a function of the force components and the position of force application. This position is set using a multiple degree-of-freedom mechanical system that uses a novel system of bearings and knife-edge rocker guides. The use of a single calibration load dramatically reduces the setup time for randomized multi-axis load combinations. Note that the application of all six components of force and moment with a single force vector poses a physics-based constraint [15] that is accommodated with a numerical algorithm during the experimental design generation.

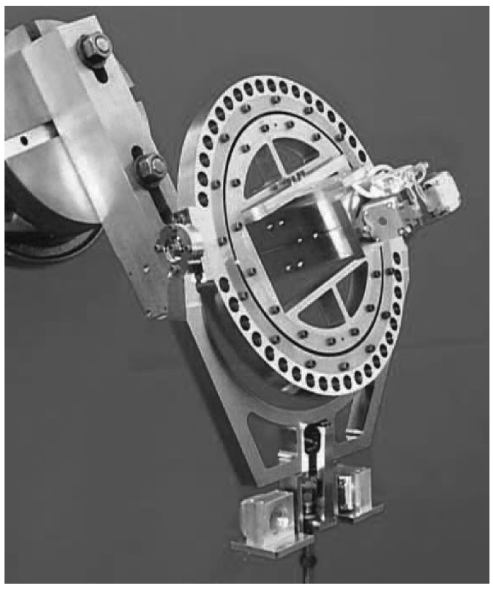

Fig. 1 Single-Vector System.

The integration of RSM with the SVS has resulted in an order of magnitude reduction in calibration time and cost, while simultaneously increasing quality. Because of the increase in calibration efficiency, the interval between successive calibrations can be significantly shortened.

C. Statistical Process Monitoring

Statistical quality control is a disciplined approach to quantify and monitor process variability. Although measurement standards laboratories have historically employed these methods [19,20], they have not been routinely applied to wind-tunnel instrumentation, primarily due to infrequent calibration. In this section, we introduce terminology and review some of the basic concepts of statistical process monitoring.

When analyzing replicated calibrations over time, we expect some level of variability in the model coefficients, regardless of the instrument design, calibration system, or condition of the instrument. Statistical process monitoring methods enable us to quantify the level of expected variability, known as common cause variation, and distinguish it from a departure from this condition, known as special cause variation. In the case of force balances, special causes might include mishandling, structural damage, or electrical malfunction. If the variability of the calibration over time is only due to common cause variation then we say that it is in control. Alternatively, when we detect special cause variation, we refer to the process as exhibiting behavior that is out of control.

Statistical process monitoring is divided into two phases, denoted by phases I and II. In phase I, we analyze historical data and estimate the level of common cause variation. In this initial phase, it is challenging to detect the existence of any special causes, because we have not quantified the level of common cause variation. Throughout this paper, we focus on phase I methods, which is consistent with the current state of implementation to the wind-tunnel instruments discussed.

At the completion of phase I, we have an estimate of common cause variation and are prepared to monitor the calibration stability over time in an online mode. Online monitoring, known as phase II, analyzes each new calibration as it is performed and makes an assessment of its consistency with previous calibrations. Phase II methods are designed to rapidly detect departures from common cause variation, known as signaling. These methods provide an objective decision tool to monitor the health and performance of the instrument throughout its life cycle.

For a calibration application, we propose monitoring the estimated model coefficients to assess instrument stability. Recall that the model coefficients are the deliverable product from the calibration, rather than the raw data, and therefore monitoring them directly is intuitive. The relationship between the independent variables and the responses, captured by the model coefficients, is generically referred to as a profile. In the methods that follow, we discuss techniques that identify changes or shifts among multiple profiles.

Change-point methods are designed to rapidly detect small departures from the in-control state of the calibration process. A change point denotes a location in time when there is a change in the calibration profile, which may require an investigation into the possible causes of the change or it may warrant recalibration. If the change is extreme, then it will be readily detected, not requiring the use of sophisticated statistical methods.

If a change point is detected, then diagnostic tools are used to determine the location(s) of the change point(s) and the nature and severity of the shift. For example, a shift in a slope coefficient or variance term affects the bias and precision of the predicted forces and moments. Alternatively, small shifts in the intercept are less important due to an offset correction procedure employed during

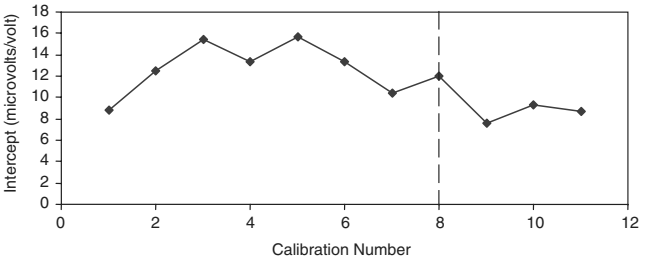

Fig. 2 Estimated intercept values for the axial force signal plotted for sequential calibrations over a period of 16 months. A change point was detected after calibration no. 8, denoted by a dashed line.

wind-tunnel operations, commonly known as adjusting for the wind-off zero. However, a large shift in the intercept may indicate structural or electrical damage. Therefore, it is not only necessary to detect a shift, but also to attribute the shift to a specific model parameter for diagnostic purposes.

To illustrate, consider a change-point method to investigate replicated calibrations provided in [21] that uses a binary segmented regression technique for testing the constancy of the regression parameters in a linear profile data set. (Note that this method is designed to detect between calibration change points, rather than within a calibration. Standard RSM analysis diagnostics can be used to detect within calibration anomalies.) Our example data set consists of 11 calibrations collected over 16 months, which has been traditionally considered a reasonable calibration interval. We illustrate the method using the axial force component. This particular change-point method is limited to simple linear relationships. Therefore, to isolate the simple linear relationship of the axial force prime sensitivity and account for the influence of the other explanatory variables, a partial regression approach was employed (see [22] for details on partial regression). Finally, a nominal false alarm probability of $\alpha = 0.04$ was used. When we obtain an indication of a change point, there is a risk of false detection. The value of α sets this risk to an acceptable level.

Applying this method, a change point was detected after calibration no. 8. Decomposing the test statistic into pieces attributable to the intercept, slope, and variance, indicates that this out-of-control situation is primarily attributable to a shift in the intercept. Note that we simultaneously considered potential shifts in the intercept, slope, and variance and attributed the signal to the intercept.

Splitting the data set into two subsets at calibration no. 8 allows us to recursively apply the method to search for additional change points, of which none were found. A correction to the overall false alarm rate was applied, because with each additional test there is a risk of false detection. A time-based plot of the intercept estimates from the 11 calibrations is provided in Fig. 2 with a vertical dashed line denoting the location of the change point. The change in the intercept is on the order of 5 μV/V, which is approximately 0.5% of the maximum response of the axial force channel. Therefore, this small shift would not have a significant impact on the performance of the force balance, and furthermore it would not warrant recalibration. Because there is no evidence of a change point in the sensitivity or the variance, plots of these parameters are not supplied. In the application of a change-point method to calibration data, a result of not detecting a change point is equally as informative as detection.

In summary, the change-point method gave evidence of a shift in this set of replicated calibrations. Our ability to attribute it to the intercept rather than the sensitivity constant (slope) and interpret its magnitude demonstrates the diagnostic benefits of this method. Based on this particular data set, we have some initial evidence to suggest that 16 months may be a reasonable calibration interval. A more comprehensive assessment of calibration stability requires a representative sample of calibrations mingled with wind-tunnel service.

We underscore that our ability to apply statistical monitoring techniques is a direct result of employing more efficient calibration methods that allow for frequent calibrations. Integrating efficient calibration methods with statistical monitoring tools provides a complete picture of balance performance and stability, rather than focusing our attention solely on the results of an individual calibration.

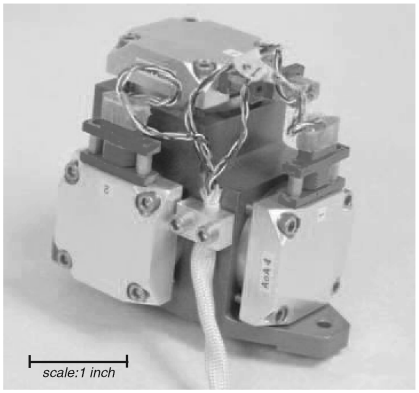

Fig. 3 Three quartz flexure devices orthogonally mounted in a rigid housing to form a triaxial accelerometer system.

III. Calibration of Inertial Attitude Measurement Systems

In this section, we focus on the calibration of inertial-based attitude measurement instruments, commonly known as angle measurement systems. The quartz flexure accelerometer, routinely used in wind-tunnel experimentation, is known colloquially as a Q-flex[‡], a name derived from its internal construction. It is used to measure static acceleration with respect to the Earth's gravitational field and produces an electrical current, converted to a voltage with a precision resistor, whose sign and magnitude are proportional to its orientation. More specifically, the current is proportional to the sine of the angle between a plane that is normal to the Earth's gravitational vector and the internal axis of the accelerometer [23,24]. The change in current with respect to the change in acceleration is called the sensitivity. When the Q-flex is in a level orientation it produces a near-zero current, although not exactly zero due to an internal electrical imbalance, referred to as the bias. The Q-flex also contains an internal temperature sensor that is used for temperature compensation.

Combining three single-axis Q-flex devices forms a triaxial accelerometer system that is used to measure pitch and roll attitude [25], for alignment of components, and detection of relative movement in mechanical joints. As previously discussed, the triaxial system is also a critical component of the SVS. The triaxial system incorporates three mutually orthogonal accelerometers mounted in a rigid housing (a photograph is shown in Fig. 3). Note that the alignment in the housing is more accurately described as near orthogonal, due to fabrication imperfections and the internal misalignment of the sensitive axis of the Q-flex with respect to its external case. Simultaneous measurements of the projection of the gravitational vector onto a three-axis Cartesian coordinate system are provided, thereby enabling the measurement of pitch and roll angles through trigonometric relationships. Note that in any orientation, two angles can be determined, which in some cases are defined by pitch and yaw.

The highest angular sensitivity and thereby the most accurate angle measurement is achieved when the accelerometer's sensitive axis is oriented perpendicular to the gravitational vector. By employing three accelerometers, we ensure that regardless of the orientation of the triaxial system, at least two of the individual devices will be less than 45 deg away from their sensitive attitude.

[‡]Q-flex® is a registered trademark of Honeywell International, Inc. and is used in this report for identification only. It does not constitute an official endorsement, either expressed or implied, by the National Aeronautics and Space Administration.

A. Experimental Design

The current mathematical model used to characterize a triaxial accelerometer system is based on the trigonometric relationship between the electrical signal (response) and the angular orientation of the device given by $\alpha = \sin^{-1}(v - b/S)$ where α is the angle, v is the electrical signal, b is the bias, and S is the sensitivity. In a triaxial system, additional model parameters are required to compensate for the cross-axis sensitivity that results from the near-orthogonal orientation of the accelerometers.

These additional parameters are the coning angle Ω, azimuth angle A_z, pitch angle α, and roll angle R. For the individual Q-flex that is sensitive to level relative to the x axis, the coning angle is defined by the angle from the x axis to the sensitive axis of the accelerometer and is usually less than 1 deg [26]. The azimuth angle is the angle formed by the projection of the sensitive axis onto the y–z plane. Similar definitions exist for the other two accelerometers that are sensitive to level relative to the y and z axes.

In a triaxial system, the model between the response of each accelerometer (v_x, v_y, v_z) and the angular orientation is given by

$$\begin{aligned} v_x &= b_x + S_x[\cos\Omega_x \sin\alpha - \sin\Omega_x \cos\alpha \sin(R + A_x)] \\ v_y &= b_y - S_y[\cos\Omega_y \sin R \cos\alpha - \sin\Omega_y(\sin A_y \sin\alpha \\ &\quad - \cos A_y \cos R \cos\alpha)] \\ v_z &= b_z - S_z[\cos\Omega_z \cos R \cos\alpha - \sin\Omega_z(\cos A_z \sin\alpha \\ &\quad - \sin A_z \sin R \cos\alpha)] \end{aligned} \tag{2}$$

where the subscripts x, y, and z denote the accelerometer that is aligned with the indicated axis. Note that the calibration model also includes a random error term (not explicitly included), which is always present in experimental data. To estimate the calibration coefficients from Eq. (2), an approximate solution can be used [25]; however, an iterative nonlinear regression based on the Gauss–Newton algorithm is recommended [26,27]. In addition, to apply the calibration coefficients in practice during data reduction, a nonlinear iterative solution is required.

We propose the use of a linear model, rather than the nonlinear model in Eq. (2). Instead of modeling the relationship between the response and the angle, we propose to model the relationship between the response and the projection of the gravitational vector onto the Cartesian coordinate system of the triaxial accelerometer, denoted by g_x, g_y, and g_z, which are subsequently referred to as g vectors. Our proposed calibration model is given by

$$\begin{aligned} v_x &= \beta_{0_x} + \beta_{1_x} g_x + \beta_{2_x} g_y + \beta_{3_x} g_z + \epsilon_x \\ v_y &= \beta_{0_y} + \beta_{1_y} g_x + \beta_{2_y} g_y + \beta_{3_y} g_z + \epsilon_y \\ v_z &= \beta_{0_z} + \beta_{1_z} g_x + \beta_{2_z} g_y + \beta_{3_z} g_z + \epsilon_z \end{aligned} \tag{3}$$

where, for the x-axis accelerometer, β_{0_x} is the bias, β_{1_x} is the sensitivity, β_{2_x} is the coefficient for the cross-axis sensitivity of g_y, β_{3_x} is the coefficient for the cross-axis sensitivity of g_z, and ϵ_x is an independently and identically distributed random error. Comparing this model to Eq. (2), the bias term b_x is equivalent to β_{0_x}, and the sensitivity S_x is equivalent to β_{1_x}. Differing from the nonlinear model, the misalignment terms β_{2_x} and β_{3_x} are in vector components rather than polar coordinates. The other model terms are similarly defined for the y- and z-axis accelerometers denoted by their respective subscripts. Note that the terms corresponding to the sensitivity constant for the y and z axes are β_{2_y} and β_{3_z}, respectively.

We can express the proposed linear model in Eq. (3) in a more compact matrix notation given by

$$\boldsymbol{v}' = \boldsymbol{\beta}_0' + \boldsymbol{g}'\boldsymbol{\beta} + \boldsymbol{\epsilon}' \tag{4}$$

where

$$\boldsymbol{v} = \begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix}, \quad \boldsymbol{\beta}_0 = \begin{bmatrix} \beta_{0_x} \\ \beta_{0_y} \\ \beta_{0_z} \end{bmatrix}, \quad \boldsymbol{g} = \begin{bmatrix} g_x \\ g_y \\ g_z \end{bmatrix},$$

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_{1_x} & \beta_{1_y} & \beta_{1_z} \\ \beta_{2_x} & \beta_{2_y} & \beta_{2_z} \\ \beta_{3_x} & \beta_{3_y} & \beta_{3_z} \end{bmatrix}, \quad \boldsymbol{\epsilon} = \begin{bmatrix} \epsilon_x \\ \epsilon_y \\ \epsilon_z \end{bmatrix}$$

To estimate the calibration coefficients, a multiple linear regression is performed separately for each response, thereby providing a noniterative estimation procedure. We denote the estimated vector of bias terms as $\hat{\boldsymbol{\beta}}_0$, and the matrix of estimated coefficients, without the intercept (bias) term, as $\hat{\boldsymbol{\beta}}$. As a consequence of using a linear model, a direct, noniterative solution is implemented in the data reduction to transform the measured signals into the estimated gravitational components. Substituting the estimated coefficients into Eq. (4) and solving with straightforward matrix algebra results in the estimated gravitational components $\hat{\boldsymbol{g}}$, given by

$$\hat{\boldsymbol{g}}' = (\boldsymbol{v} - \hat{\boldsymbol{\beta}}_0)'\hat{\boldsymbol{\beta}}^{-1} \tag{5}$$

As previously discussed, both the bias and the sensitivity need to be compensated for temperature related effects, which we denote with a superscript t. Incorporating these temperature compensated terms into Eq. (5), it can be shown that the estimated, temperature compensated, gravitational components are computed by

$$\hat{\boldsymbol{g}}^{t\prime} = (\boldsymbol{v} - \hat{\boldsymbol{\beta}}_0^t)'\hat{\mathbf{T}}_s\hat{\boldsymbol{\beta}}^{-1}$$

where

$$\hat{\mathbf{T}}_s = \begin{bmatrix} \dfrac{\hat{\beta}_{1_x}}{\hat{\beta}^t_{1_x}} & 0 & 0 \\ 0 & \dfrac{\hat{\beta}_{2_y}}{\hat{\beta}^t_{2_y}} & 0 \\ 0 & 0 & \dfrac{\hat{\beta}_{3_z}}{\hat{\beta}^t_{3_z}} \end{bmatrix}$$

and $\hat{\boldsymbol{\beta}}_0^t$ is the vector of the bias terms compensated for the observed temperature sensor output (TSO), $\hat{\beta}_{1_x}$ is the sensitivity at the calibration reference TSO, and $\hat{\beta}^t_{1_x}$ is the sensitivity compensated for the observed TSO of the x-axis accelerometer. The compact matrix notation and noniterative form of the estimation and data reduction equations can be easily implemented in a software package that performs matrix manipulations or in a spreadsheet.

Once we estimate the gravitational components from the signals, the equations to compute angles are given by

$$\alpha = \sin^{-1}(\hat{g}_x) \qquad R = \tan^{-1}\left(\frac{\hat{g}_y}{\hat{g}_z}\right)$$

and as the pitch angle approaches 90 deg, we compute yaw rather than roll given by

$$Y = \tan^{-1}\left(\frac{\hat{g}_y}{\hat{g}_x}\right)$$

The estimation of the three gravitational components is an advantage of the linear model because it allows for a check of system stability by computing the root-sum square (RSS) of the g vectors, which by definition should equal 1. In the nonlinear approach, the pitch and roll angles are estimated and then converted to the gravitational components using trigonometric relationships, which by definition, always equal 1 and eliminate the possibility of this internal check. Moreover, the gravitational components are uniquely determined in any orientation, thereby eliminating the ambiguity and path dependence of an angular solution over wide angle ranges.

Table 1 A six-point experimental design to enable the estimation of the four calibration coefficients in the proposed linear model. Levels of g_x, g_y, and g_z represent the projection of the gravitational vector on each axis

Design point	g_x	g_y	g_z
1	−1	0	0
2	1	0	0
3	0	−1	0
4	0	1	0
5	0	0	−1
6	0	0	1

As a point of clarification, the proposed linear solution in Eq. (3) is not an approximation of the nonlinear model. Rather it is a reparameterized version of the same model with equivalent degrees of freedom to characterize the triaxial accelerometer system.

An experimental design to support the estimation of the coefficients in Eq. (3) was developed based on the gravitational vector components (explanatory variables). In Eq. (3), there are four calibration coefficients to be estimated for each response. A proposed experimental design is shown in Table 1 that requires the setting of each component independently to its extreme values, resulting in a six-point design. These points are oftentimes referred to as the cardinal points and represent a minimal set of points (with 2 additional degrees of freedom) to determine the four calibration coefficients. In addition, a set of confirmation points consisting of randomly allocated g vectors, not found in the experimental design, is recommended in practice.

A simple geometric interpretation of the design is instructive in understanding the design space. Consider a three-dimensional coordinate system where the axes represent the three components of the gravitational vector (g_x, g_y, g_z). Each of the design points lie on an axis of the coordinate system at a distance of one unit away from the origin. Furthermore, consider a spherical surface of radius one. The design points represent the intersection of the coordinate axes with the sphere. The notion of a spherical surface is particularly relevant in this application, because it represents the domain of all possible design points. Because of the physics-based constraint on the RSS of the gravitational components, every possible design point must lie on the surface of the sphere. Note that a point at the center of the sphere (at the origin of the coordinate system) is infeasible.

B. Accelerometer Calibration and Evaluation Cube

A new mechanical calibration system has been developed that provides an improved and simplified method of calibrating a single or multi-axis accelerometer system. Previous calibration systems rely on precise angular positioning using a mechanically complex sequence of rotary tables mounted in a known (near-level) orientation relative to the gravitational coordinate system. Although these systems represent the state of the art in accelerometer calibration, they possess certain weaknesses. For example, the systems are costly and are permanently mounted in a laboratory, making them incompatible for in situ calibration. In addition, they require a complex calibration procedure to partially compensate for the near levelness of the mounting surface and the orthogonal misalignment among the sequence of rotary tables.

The accelerometer calibration and evaluation cube (ACE cube) has been proposed, shown in Fig. 4, which addresses many of the weaknesses in existing systems. For example, the production cost of the ACE cube is an order of magnitude less expensive compared to the previous calibration apparatus. In addition, the system is fundamentally simpler and requires fewer mechanical components, thereby reducing the sources of uncertainty. Moreover, it is relatively portable allowing for in situ evaluation and calibration.

The primary components of the system are shown in Fig. 4 and include a rigid cuboidal frame where the triaxial accelerometer system is mounted, a set of three positioning balls on each of the six faces of the cuboidal frame (a total of 18 balls), a fence that allows for a repeatable orientation of the cuboidal frame in a plane normal to the gravitational vector, and a flat, stable, near-level surface table.

The design points do not need to be set to their nominal values as long as they are known and incorporated into the analysis. As a result, the fabrication cost of the system is reduced by eliminating precision machining requirements except in the mounting area for the triaxial system in order to repeatedly align the triaxial system with the cuboidal frame. By eliminating the need to achieve exact set points, the fabrication tolerances on the faces of the cube can be relaxed. After the cube is fabricated, a series of experiments are performed to define the orientation of the cube faces using multiple precalibrated triaxial systems. The use of multiple triaxial systems reduces the transferred bias from each device through simple averaging.

To perform the calibration, the triaxial accelerometer is mounted into the cube and the fasteners are installed in a repeatable manner. A clean, near-level granite block is used as the surface table. Note that the stability of the surface table is of primary interest, rather than its levelness. To ensure that the calibration is robust to the table levelness, the experimental design in Table 1 is replicated to include two observations of each design point obtained by rotating the cube 180 deg in a plane normal to the gravitational vector. Position numbers on the upper right corners of the cube guide the operator in setting the orientation for each design point. Position numbers 7–12 indicate 180 deg rotations of numbers 1–6 (an example of position numbers 1 and 7 can be seen in Fig. 4). The fence ensures a 180 deg alignment between these rotations. Three replicates of the 12-point design are performed in a completely randomized order within a single block.

After the calibration is complete, the data undergo a postprocessing step to average the 180 deg opposing positions,

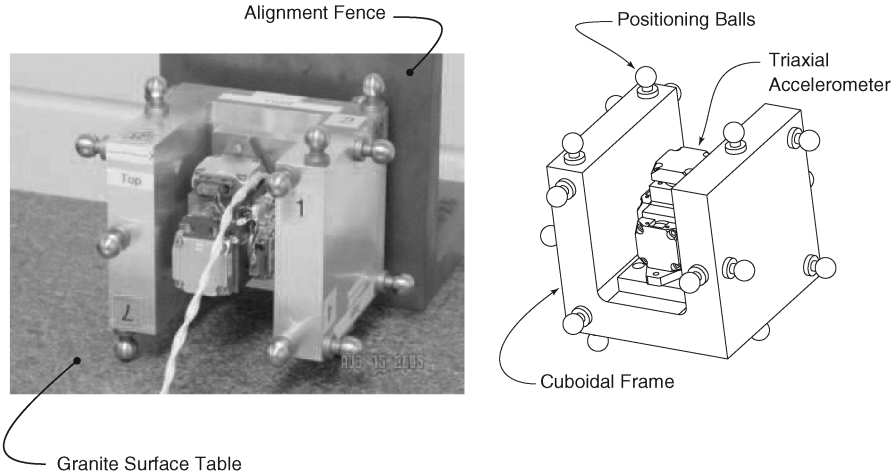

Fig. 4 Accelerometer calibration and evaluation cube, shown resting on a granite surface table with a triaxial accelerometer system mounted for calibration (left panel); primary components are specified (right panel).

thereby eliminating the effect of the near-level condition of the surface table from the estimation of the calibration coefficients. Subsequent analysis is performed on these six averaged pseudo-observations. Note that if the surface table were set perfectly level, then the six design points in Table 1 would be sufficient. We prefer to relax the table leveling requirement and employ the 12-point approach, because it is well suited for in situ calibration in nonlaboratory conditions. In addition, by computing half the difference between opposing positions, we can obtain six independent estimates of table levelness in two axes. A comparison of these estimates forms a convenient approach to detect within calibration operator errors.

The ability of the ACE cube to perform in situ calibrations has two primary advantages. First, in situ calibrations allow users to check the health and performance of their instrumentation connected to their own data acquisition system. This provides an increased level of confidence in the transfer of the coefficients from the calibration to the wind-tunnel laboratory. Second, in situ calibrations provide calibration coefficients that are automatically compensated for the local gravitational field. Current methods require a mathematical correction to the laboratory-determined coefficients based on a measurement of the local gravitational constant. The automatic compensation feature is particularly important when an accelerometer system is used a significant geographic distance from the calibration laboratory.

C. Statistical Process Monitoring

By exploiting the increased accessibility and ease of calibration provided by the ACE cube, more frequent verifications and calibrations of angle measurement systems can be performed. In this section, we demonstrate another phase I method and illustrate it with replicated triaxial accelerometer calibrations performed over a period of 10 weeks.

The most frequently used statistical monitoring tool is the control chart. The basic idea of a control chart is to plot a performance statistic as a function of time on a graph that contains a centerline and upper and lower limits. The limits are statistically derived and represent the expected range of common cause variation. A patternless random variation of the statistic about the centerline, based on the process mean, and within the limits indicates a system that is in control. There are many types of control charts, with the simplest version known as a Shewhart chart for normally distributed data [8]. The phase I statistical monitoring method we present here is based on using multiple Shewhart-type charts [28].

Three replicated calibrations, once per week, over a period of 10 weeks were performed using the ACE cube. Within the three replicated calibrations performed on the same day, the triaxial system was not mechanically removed from the ACE cube. However, in between the sets of three calibrations, the system was removed and remounted to include the variability due to the mounting interface. Throughout the 10 week period, the time of day when the calibration was performed was varied to include warm-up related variability and the effect of normal temperature fluctuations in the calibration laboratory.

Therefore, on a single day a calibration consisted of three replicates of the 12-point design resulting in a total of 18 pseudo-observations. A set of calibration coefficients were estimated for each day, resulting in 10 sets (one set per week). Recall that for each accelerometer there are four calibration coefficients corresponding to the bias, primary sensitivity, and two cross-axis sensitivities. Because of the temperature variation over the 10 weeks, the sensitivity has been temperature compensated. In addition, an estimate of the residual variance was obtained for each calibration. A group of five Shewhart-type charts were constructed for each of the three accelerometers. As an example, consider the x-axis accelerometer shown in Fig. 5.

On each chart, there is an upper and a lower limit that are denoted by dashed lines and a centerline that is denoted by a dash-dotted line. These limits were computed as a function of the variance of the regression parameter estimates, and the overall false alarm rate ($\alpha = .05$) was achieved by adjusting the limits to compensate for the multiple charts and the number of samples.

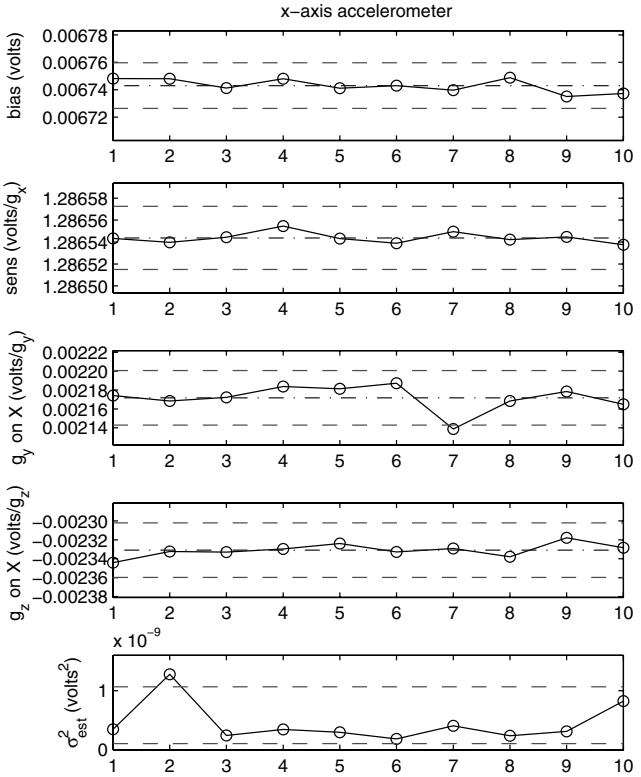

Fig. 5 Shewhart-type control charts for the estimated calibration coefficients and residual variance for replicated calibrations over a period of 10 weeks for the x-axis accelerometer. Centerlines are indicated by dash-dotted lines and limits are dashed lines. Signals occur at calibration no. 7 for the cross-axis sensitivity of g_y and calibration no. 2 for the residual variance. Each statistic is based on a sample size of 18 observations. The overall false alarm rate is $\alpha = 0.05$.

Reviewing Fig. 5, we see that the variance chart signals at the upper limit for calibration no. 2. This signal was investigated and no assignable cause could be determined. Therefore, the calibration was not excluded from subsequent analysis. We recommend that the variance chart be reviewed first since a signal on this chart could indicate inadequacy of the model rather than instability in the coefficients [21]. Another signal occurs on the cross-axis sensitivity of g_y at the lower limit for calibration no. 7. This point was also investigated, and no cause could be found, therefore it was retained. Note that in the phase I analysis, we are estimating common cause variation so that effective phase II methods can be employed. If an assignable cause is found, then that observation should be removed because it could deteriorate the monitoring performance in phase II.

In regard to the lower limit on the variance control chart, it may not be intuitive as to why we would want to signal if the variance decreases. A decrease in the variance would be considered an improvement in the calibration quality, if the decrease was attributable to an intentional change, or improvement, in the calibration process. If the variance decreases, and no such change has been initiated, then this could signal a mistake in the data reduction or experimental protocol. In general, we apply control charting techniques to detect a change in the expected level of variation, rather than a change in a direction that we would consider favorable.

These Shewhart-type control charts provide a simple graphical representation of the variation in the calibration coefficients over time in a retrospective analysis. The limits on the charts also provide valuable information about the expected level of variation in each calibration coefficient. Furthermore, they provide an objective technique to distinguish between common cause and special cause variation, signaling a potential disruption in the calibration process or instrument performance.

IV. Conclusions

We have presented the fundamental concepts of response surface methodology and statistical quality control and demonstrated their application to the calibration of wind-tunnel instrumentation. The integration of the experimental design with the mechanical calibration system has enabled significant reductions in calibration time and expense, while simultaneously producing richer information regarding the performance and stability of an instrument. Overall, we conclude that these statistical methods combined with subject-matter expertise provide new and deeper insights that can lead to the advancement of instrumentation systems.

Acknowledgments

The authors would like to thank Greg Jones of Modern Machine and Tool Company, Inc. for his assistance in the execution of the calibration designs discussed. In addition, we would like to thank the editor and the reviewers for their helpful suggestions that significantly improved this article.

References

[1] Montgomery, D. C., *Design and Analysis of Experiments*, 6th ed., Wiley, New York, 2004.

[2] Myers, R. H., and Montgomery, D. C., *Response Surface Methodology: Process and Product Optimization Using Designed Experiments*, 2nd ed., Wiley, New York, 2002.

[3] DeLoach, R., "Tailoring Wind Tunnel Data Volume Requirements Through the Formal Design of Experiments," AIAA Paper 98-2884, June 1998.

[4] DeLoach, R., "Improved Quality in Aerospace Testing Through the Modern Design of Experiments," AIAA Paper 2000-0825, Jan. 2000.

[5] DeLoach, R., Hill, J. S., and Tomek, W. G., "Practical Applications of Response Surface Methods in the National Transonic Facility," AIAA Paper 2001-0167, Jan. 2001.

[6] Belter, D. L., and Holler, D. J., "Assessing Long-Term Instrumentation Performance and Uncertainty from Mulitple Calibrations," AIAA Paper 2000-2202, June 2000.

[7] Woodall, W. H., Spitzner, D. J., Montgomery, D. C., and Gupta, S., "Using Control Charts to Monitor Process and Product Quality Profiles," *Journal of Quality Technology*, Vol. 36, No. 3, July 2004, pp. 309–320.

[8] Montgomery, D. C., *Introduction to Statistical Quality Control*, 5th ed., Wiley, New York, 2004.

[9] Hemsch, M. J., "Development and Status of Data Quality Assurance Program at NASA Langley Research Center—Toward National Standards," AIAA Paper 96-2214, June 1996.

[10] Hemsch, M., Grubb, J., Krieger, W., and Cler, D., "Langley Wind Tunnel Data Quality Assurance—Check Standard Results," AIAA Paper 2000-2201, June 2000.

[11] Ferris, A. T., "Strain Gauge Balance Calibration and Data Reduction at NASA Langley Research Center," NASA Paper DR-2, Oct. 1996.

[12] Hansen, R. M., "Evaluation and Calibration of Wire-Strain-Gage Wind-Tunnel Balances Under Load," NACA Langley Aeronautical Laboratory, Hampton, VA, 1956.

[13] Parker, P. A., and DeLoach, R., "Response Surface Methods for Force Balance Calibration Modeling," *IEEE 19th International Congress on Instrumentation in Aerospace Simulation Facilities*, IEEE, Reston, VA, Aug. 2001.

[14] Dahiya, R. C., "Statistical Design for an Efficient Calibration of a Six Component Strain Gage Balance," Old Dominion University Research Foundation, work performed under LaRC Contract No. NAS1-15648, Aug. 1980.

[15] Parker, P. A., Morton, M., Draper, N., and Line, W., "A Single-Vector Force Calibration Method Featuring the Modern Design of Experiments," AIAA Paper 2001-0170, Jan. 2001.

[16] Simpson, J., Landman, D., Rhew, R., Giroux, R., and Zeisset, M., "Calibrating Large Capacity Aerodynamic Force Balance Instrumentation Using Response Surface Methods," AIAA Paper 2005-7601, Dec. 2005.

[17] Parker, P. A., and Rhew, R. D., "A Study of Automatic Balance Calibration System Capabilities," *Second International Symposium on Strain Gauge Balances*, May 1999 (unpublished).

[18] Parker, P. A., and Liu, T., "Uncertainty Analysis of the Single-Vector Force Balance Calibration System," AIAA Paper 2002-2792, 2002.

[19] Pontius, P. E., and Cameron, J. M., "Realistic Uncertainties and the Mass Measurement Process," *Precision Measurement and Calibration: Statistical Concepts and Procedures*, NBS SP 300, U.S. GPO, Washington, D.C., Feb. 1969, Vol. 1.

[20] Schumacher, R. B. F., "Statistical Control in a Standards Laboratory," *Measurements and Data*, Vol. 3, No. 15, 1969, pp. 58–64.

[21] Mahmoud, M. A., Parker, P. A., Woodall, W. H., and Hawkins, D. H., "A Change Point Method for Linear Profile Data," *Quality and Reliability Engineering International* (to be published).

[22] Myers, R. H., *Classical and Modern Regression with Applications*, 2nd ed., PWS-Kent Publishing Co., Boston, MA, 1990.

[23] Finley, T. D., "Technique for Calibrating Angular Measurement Devices when Calibration Standards are Unavailable," NASA TM 104148, Hampton, VA, Aug. 1991.

[24] Finley, T. D., and Tcheng, P., "Model Attitude Measurements at NASA Langley Research Center," AIAA Paper 92-0763, Jan. 1992.

[25] Finley, T. D., "A Digital Pitch and Roll Monitor," *Proceedings of the 35th International Instrumentation Symposium*, Proceedings (A91-19651 06-35), Instrument Society of America, Research Triangle Park, NC, May 1989, Vol. 35, pp. 233–250.

[26] Tripp, J. S., and Tcheng, P., "Uncertainty Analysis of Instrument Calibration and Application," NASA TP-1999-209545, 1999.

[27] Marshall, R., and Landman, D., "An Improved Method for Determining Pitch and Roll Angles Using Accelerometers," AIAA Paper 2000-2384, June 2000.

[28] Kim, K., Mahmoud, M. A., and Woodall, W. H., "On the Monitoring of Linear Profiles," *Journal of Quality Technology*, Vol. 35, No. 3, July 2003, pp. 317–328.