AIAA-2001-0170

A Single-Vector Force Calibration Method Featuring the Modern Design of Experiments (Invited)

P.A. Parker, M. Morton, N. Draper, W. Line
NASA Langley Research Center
Hampton, Virginia

39th AIAA Aerospace Sciences Meeting & Exhibit
8-11 January 2001
Reno, Nevada

AIAA-2001-0170

A SINGLE-VECTOR FORCE CALIBRATION METHOD FEATURING THE MODERN DESIGN OF EXPERIMENTS

P.A. Parker*, M. Morton†, N. Draper‡, W. Line§
NASA Langley Research Center
Hampton, Virginia 23681

Abstract

This paper proposes a new concept in force balance calibration. An overview of the state-of-the-art in force balance calibration is provided with emphasis on both the load application system and the experimental design philosophy. Limitations of current systems are detailed in the areas of data quality and productivity. A unique calibration loading system integrated with formal experimental design techniques has been developed and designated as the Single-Vector Balance Calibration System (SVS). This new concept addresses the limitations of current systems. The development of a quadratic and cubic calibration design is presented. Results from experimental testing are compared and contrasted with conventional calibration systems. Analyses of data are provided that demonstrate the feasibility of this concept and provide new insights into balance calibration.

Introduction

Direct force and moment measurement of aerodynamic loads is fundamental to wind tunnel testing at NASA Langley Research Center (LaRC). An instrument known as a force balance provides these measurements. Force balances provide high-precision measurement of forces and moments in six degrees of freedom. The balance is mounted internally in the scaled wind tunnel model, and the components measured by the force balance consist of normal force (*NF*), axial force (*AF*), pitching moment (*PM*), rolling moment (*RM*), yawing moment (*YM*), and side force (*SF*), see Figure 1.

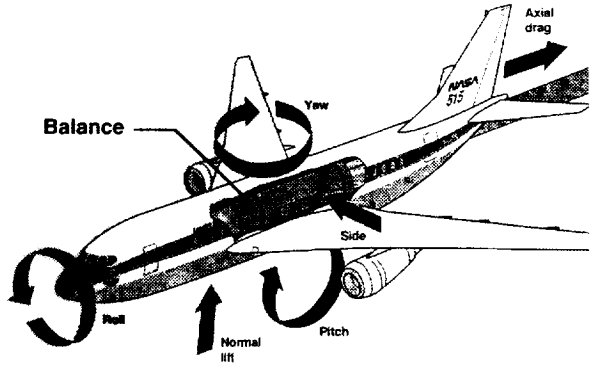


Figure 1. Measured aerodynamic forces and moments.

Force Balance Background

The force balance is a complex structural spring element designed to deflect a specified amount under applied load. This deflection of the balance under load results in a change in strain within the flexural elements. Differential strain is measured with six Wheatstone bridges, each consisting of four foil resistive strain gages. Each bridge is designed to primarily respond to the application of one of the six components of load. The electrical response of these bridges is proportional to the applied load on the balance. This electrically measured strain, as a function of load, forms the basic concept of force balance measurements that has been generally used since the 1940s.

The balance flexural elements are optimized such that the magnitude of the strain response is approximately the same for the individual application of each component of load. The magnitude of the applied load in each component is not equivalent, and therefore the flexural elements do not have the same deflection in all axes. Typical balance load ratios of lift to drag are sixteen to one. These structural requirements dictate that the force balance must be carefully designed in

* Research Scientist, Model Systems Branch, Member AIAA
† Research Statistician, R.J. Reynolds
‡ Emeritus Professor, University of Wisconsin, Madison, WI
§ President, DOES Institute

order to achieve an accurate measurement of axial force (drag) in the presence of a large normal force (lift). These high ratios of component load also create undesired cross-effects in the balance responses.

Ideally, each balance response signal would respond to its respective component of load, and it would have no response to other components of load. This is not entirely possible. The response of a particular balance channel to the application of other components of load is referred to as an interaction effect. Balance designs are optimized to minimize these interaction effects. Ultimately, a mathematical model must correct for these unwanted interaction effects. A calibration experiment provides the data to derive an adequate mathematical model.

Force balances are the state-of-the-art instrument used for high-precision aerodynamic force measurements in wind tunnel testing. Over the past 60 years, there have been many improvements in the areas of balance structural design and analysis, fabrication techniques, strain sensor technology, data acquisition systems and automated calibration loading systems. Relatively little has changed in the area of force balance calibration methodology. The majority of developmental work in the area of calibration has been focused on automation of load application in order to minimize the calibration duration, and thereby provide an increase in calibration frequency. Regardless of the calibration loading mechanism, automated or manual, the experimental approach to the balance calibration has been the same and is detailed in the following section.

Balance Calibration Methodology

The balance calibration is the most critical phase in the overall production of a high quality force transducer. The goal of calibration is to derive a mathematical model that is used during wind tunnel testing of scaled aircraft models to estimate the aerodynamic loading. Furthermore, the accuracy of this model is also determined during the calibration experiment.

In general, force balance calibration consists of setting the independent variables and measuring the response of the dependent variables. The applied loads are the independent variables, and the electrical responses of the balance are the dependent variables. Current mathematical models are based on a polynomial equation where the balance response is a function of the independent variables.

For $k = 2$ design variables, a general polynomial can be expressed by[1]:

$$
\begin{aligned}
f(x,\beta) &= \beta_0 && \text{(y – intercept)}\\
&+ \beta_1 x_1 + \beta_2 x_2 && \text{(linear terms)}\\
&+ \beta_{12} x_1 x_2 && \text{(interaction terms)}\\
&+ \beta_{11} x_1^2 + \beta_{22} x_2^2 && \text{(pure quadratic terms)}\\
&+ \beta_{111} x_1^3 + \beta_{222} x_2^3 + \beta_{112} x_1^2 x_2 + \beta_{122} x_1 x_2^2 && \text{(cubic terms)}\\
&+ \beta_{1111} x_1^4 + \ldots && \text{(quartic terms)}\\
&+ etc.,
\end{aligned}
$$

where k is the number of independent variables, x_i is the i^{th} independent variable, and β represents the coefficients in the mathematical model. The subscript notation of the coefficients is chosen so that each β coefficient can be easily identified with its corresponding x term(s). This model can be thought of as a Taylor's series approximation to a general function. Typically, the higher the degree of the approximating polynomial, the more closely the Taylor series expansion will approximate the true mathematical function. In the current LaRC calibration approach, a degree of two is used, and therefore a second-order model is generated. For balance calibration, k is equal to six, and the quadratic mathematical model for each response contains a total of 28 coefficients that consist of the intercept, 6 linear terms, 15 two-way interactions, and 6 pure quadratic terms.

The magnitude of loads set in the calibration experiment defines the inference space. The error in the prediction of the mathematical model is evaluated within this space. The inference space for LaRC calibrations is based on the plus and minus full-scale design loads of the balance. Current force balance calibration schedules increment one independent variable at a time. Each independent variable is incremented throughout its full-scale range. During this incrementing of the primary variable, all other variables are zero, or are held at a constant magnitude. This approach is referred to as one-factor-at-a-time (OFAT) experimentation.

All of the required set points of independent variables are grouped into a load schedule, or calibration design. Ordering of the points within the design is based on the efficiency of the load application system and the data analysis algorithm. For optimum design execution, efficiency is gained by the application of one load at a

time in incremental levels. Data analysis algorithms used to determine the balance calibration model often rely on a specific grouping of variable combinations. Single variable loading is used to calculate the main linear and quadratic effects. Constant auxiliary loads in combination with single incremented loads are performed in order to assess two-way interactions.

Calibration Systems

In order to set the independent variables of applied load, a system of hardware is required. The NASA LaRC manual calibration consists of a system of levels, free-hanging precision weights, bell cranks, and other mechanical components. After each loading the balance is re-leveled prior to taking data to assure that the applied loads are orthogonal with the balance coordinate system. This type of calibration system is referred to as a repositioning system. There are a total of 81 load sequences performed to calculate the calibration mathematical model coefficients. Each load sequence consists of four ascending increments and four descending increments providing a total of 729 data points. Also, four sequences of a multi-component proof-loading are performed to assess the quality of the math model.[2]

Manual dead weight balance calibration stands have been in use at LaRC since the 1940's. Figure 2 is a photograph of one of these systems. They have undergone continuous improvement in load application, leveling, and data acquisition. These manual systems are the "standard" that other balance calibration systems have been compared against. The simple, accurate methodologies produce high confidence results. The conventional method is generally accurate, but the process is quite complex and labor-intensive, requiring three to four weeks to complete each full calibration. To ensure accuracy, gravity-based loading is utilized, however this often causes difficulty when applying loads in three simultaneous, orthogonal axes. A complex system of levers, bell-cranks, cables, and optical alignment devices must be used, introducing increased sources of systematic error, and significantly increasing the time and labor intensity required to complete the calibration.

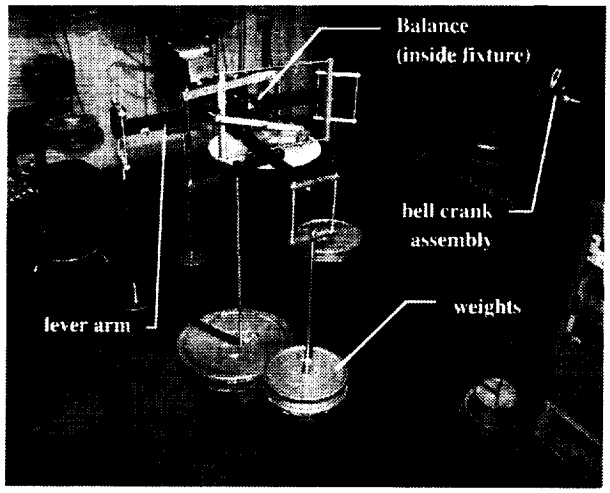


Figure 2. NASA LaRC manual dead-weight calibration system (during multi-component loading).

In recent years, automated balance calibration systems have been developed.[3,4] In general, all automatic balance calibration systems are designed to simulate the manual calibration process. Unfortunately, the automation of this tedious manual process results in a complex mechanical system. Utilizing these automated systems, combined with an abbreviated manual calibration, can reduce the balance calibration time to approximately two days; however, these new automated systems have significant disadvantages.

The mechanical complexity of an automated system makes it quite expensive. As compared to manual systems, this complexity also tends to deteriorate the load application quality due to the added degrees of freedom that factor into the overall accuracy. Since overall system accuracy is based on the elemental accuracy of multiple high precision load cells and position sensors, any complex load introduces multiple sources of systematic error. Furthermore, the calibration system accuracy is difficult to experimentally verify. Automatic systems can theoretically compute their system accuracy[5] or they can infer system accuracy by comparing their calibrations of test balances with calibrations performed using a traditional manual loading system.

A study was performed by LaRC to evaluate these automated systems.[6] One conclusion of this study was that a mixture of an abbreviated manual calibration and an automated calibration would be required in standard practice. The manual calibration, in its simplicity, inherently provides more confidence in the data collected.

Limitations of Current Force Balance Calibration

There are a number of disadvantages to the current calibration methodology and load application systems. In regard to methodology, the OFAT approach has been widely accepted because of its inherent simplicity and intuitive appeal to the balance engineer. Unfortunately, this approach provides no mechanism to defend against systematic errors, common to all calibration systems. These uncontrolled and undetected systematic errors are contradictory to the foundational premise of the OFAT methodology, that all other variables are held constant. The inability to defend against these errors, determine their magnitude and remove their influence on the model is a significant limitation. Also, the OFAT approach does not provide genuine replicates within the calibration, and therefore the repeatability of the measurement environment cannot be adequately separated from the quality of the math model. These data quality issues have been generally overlooked in the field of force balance calibration. Formal experimental design offers techniques to address these data quality issues as well as provide the balance engineer with new insights into force balance calibration. A development of these concepts as applied to balance calibration is provided in subsequent sections.

Load application systems, both manual and automated, also have disadvantages. The manual systems, although generally considered accurate, are slow and tedious and provide many opportunities for systematic error. Automated systems that greatly reduce calibration time include additional sources of systematic error due to their mechanical complexity, and their expense makes them prohibitive for wide spread use. Both of these hardware systems were designed around the OFAT calibration design requirement to set independent variables one at a time and to obtain maximum efficiency of data collection.

A new comprehensive view to force balance calibration was required to address these limitations. This new approach should simultaneously address the data quality and productivity shortcomings of the current calibration systems.

Single-Vector Calibration System

An innovative approach to balance calibration has been proposed. This new calibration system integrates a unique load application system with formal experimental design methodology. The Single Vector Balance Calibration System (SVS) enables the complete calibration of a six component force balance with a single force vector. A primary advantage to this load application system is that it improves on the "trusted" aspects of current manual calibration system, and therefore provides high confidence in the results. Applying formal experimental design methodology to the balance calibration experiment has many features that enable higher quality mathematical models. It also provides new insight into calibration data regarding the adequacy of the math model, and the repeatability of the measurement environment. Moreover, formal experiment design techniques provide an objective means for the balance engineer to evaluate and report the results of a calibration experiment. The SVS improves data quality, while simultaneously improving productivity and reducing cost. A description of the load application system will now presented followed by a description of the application of formal experimental design methodology.

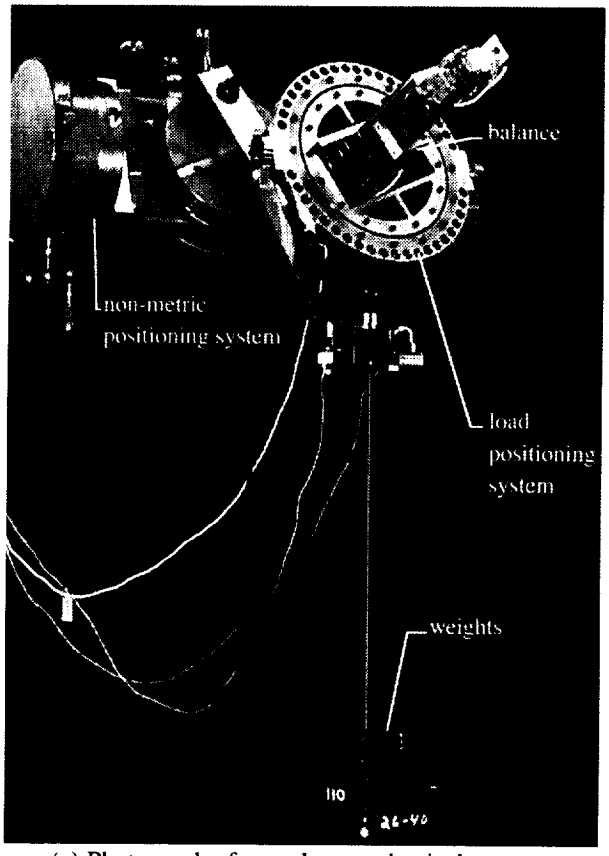
Single-Vector Load Application System

The objectives of the SVS are to provide a calibration system that enables the efficient execution of a formal experimental design, be relatively inexpensive to manufacture, require minimal time to operate, and provide a high level of accuracy in the setting of the independent variables. Photographs of the assembled system are provided in Figure 3.

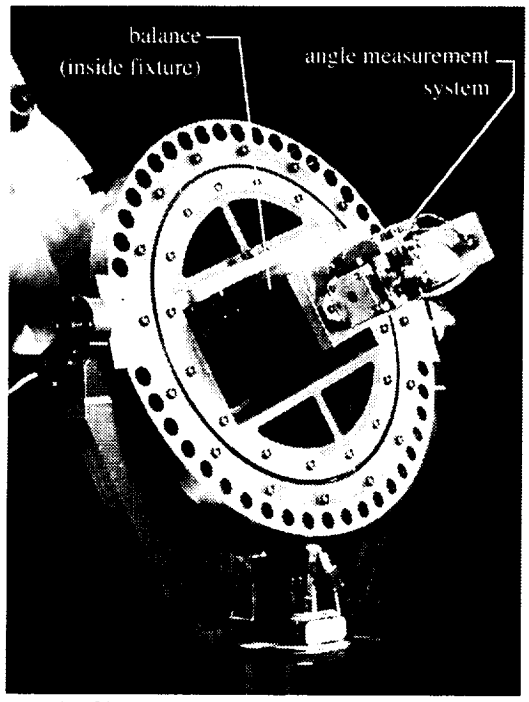
The individual hardware components of this system are critical to setting the independent variables rapidly and with high accuracy. The primary components include a non-metric positioning system, a multiple degree of freedom load positioning system, a three-axis orthogonal angular measurement system[7], and calibrated weights. These components can be seen in Figure 3. All of the system components including the data acquisition system are designed to meet or exceed the current requirements for the LaRC manual calibration system. The system features significantly fewer components then the LaRC manual system and therefore fewer sources of systematic error.

The system allows for single vector calibration, meaning that single, calibrated dead-weight loads are applied in the gravitational direction generating six component combinations of load relative to the coordinate system of the balance. By utilizing this single force vector, load application inaccuracies caused by the conventional requirement to generate multiple force vectors are fundamentally reduced. The angular manipulation of the balance, combined with the load point positioning system, allows the uni-

directional load to be used to produce three force vectors (normal force, axial force, side force) and three moment vectors (pitching moment, rolling moment, yawing moment), with respect to the balance moment center. The non-metric positioning system rotates the force balance being calibrated about three axes. The rotation angle of the balance coordinate system relative to the gravitational force vector is measured using the three-axis angle measurement system. A multiple degree of freedom load positioning system utilizes a novel system of bearings and knife-edge rocker guides to maintain the load orientation, regardless of the angular orientation of the balance. As a result, the use of a single calibration load reduces the set-up time for the multi-axis load combinations required to execute a formal experimental design.



(a) Photograph of complete mechanical system.
Figure 3. Single-Vector Balance Calibration System.



(b) Close-up of load positioning system.
Figure 3. concluded.

The application of all six components of force and moment on a balance with a single force vector poses a physics-based constraint that warrants a brief discussion in light of its impact on the development of the calibration design. Any system of forces and moments, however complex, may be reduced to an equivalent force-couple system acting at a given point. However, for a system of forces and moments to be reduced to a single resultant force vector acting at a point, the system must be able to be reduced to a resultant force vector and a resultant moment vector that are mutually perpendicular.[8] Therefore, an infinite number of combinations of the six components can be loaded with a single force vector, but any arbitrary combination of forces and moments cannot. Figure 4 provides the LaRC force balance coordinate system used for the following derivation of the governing equations.

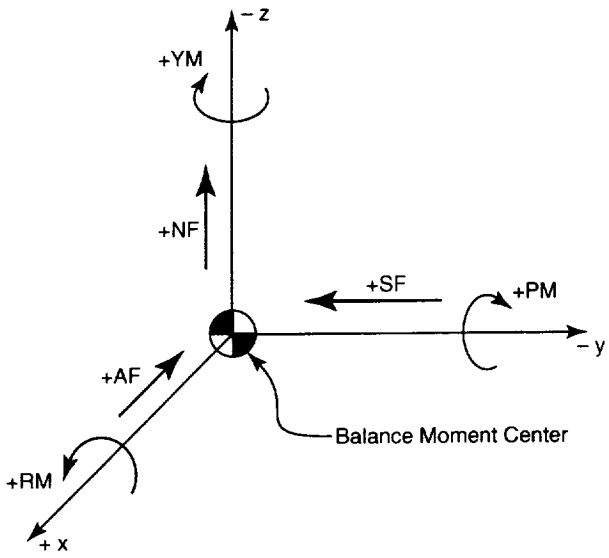


Figure 4. LaRC force balance coordinate system.

The perpendicular relationship can be expressed as the dot product of the resultant force vector and the resultant moment vector. By, setting the cosine of the angle between the two vectors equal to zero, the following relationship can be derived.

$$\frac{-(RM)(AF)+(PM)(SF)-(YM)(NF)}{\left(\sqrt{RM^2+PM^2+YM^2}\right)\left(\sqrt{AF^2+SF^2+NF^2}\right)}=0 \quad (1)$$

where, NF, AF, PM, RM, YM, SF are the component loads in actual units. Therefore for non-zero magnitudes of the resultant force and moment vector, Equation 1 can be simplified as follows.

$$-(RM)(AF)+(PM)(SF)-(YM)(NF)=0 \quad (2)$$

This governing equation constrains the relative direction and magnitude of each of the six components of load applied to a balance with a single force vector. It also introduces a co-linearity between the three products of Equation 2. The implications of this co-linearity are discussed later in this paper.

Furthermore, if the system of forces and moments can be reduced to a single force vector, then there are an infinite number of points in which that resultant force vector can be loaded to generate the desired moments. The following equation governs the location of the point in which the forces can be loaded to generate the desired moments.

$$\begin{bmatrix} 0 & -NF & -SF \\ NF & 0 & -AF \\ SF & AF & 0 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix}=\begin{bmatrix} RM \\ PM \\ YM \end{bmatrix} \quad (3)$$

where, x, y, and z are the linear distances from the balance moment center.

The three balance force components are a function of the applied load and the angular orientation of the balance in three-dimensional space. To generate a desired combination of the three forces, the balance is manipulated to a prescribed angular orientation using the non-metric positioning system, and the orientation of the balance is precisely measured on the metric end using the angle measurement system. This angle measurement system provides the components of the gravitational vector projected onto the three-axes of the balance coordinate system. The magnitude of the applied load is known based on the use of calibrated dead-weights. Combining the measured gravitational components on the balance axes and the known dead-weight enables the determination of the three force components.

The three balance moment components are a function of the three force vectors and the position of the point of load application in three-dimensional space relative to the balance moment center. The balance moment center is a defined location in the balance coordinate system that serves as a reference point in which the moment components are described. The point of load application is set using the multiple degree-of-freedom load positioning system. This system defines the x, y, and z distances from the balance moment center to the point of load application. Due to the degrees of freedom in the load positioning system, the point of load application is independent of the angular orientation of the balance. Stated another way, when the balance is manipulated in three-dimensional space, the point of load application remains constant.

The SVS achieves the objectives of rapid and accurate setting of the independent variables. Even though this load application system would greatly enhance the execution of the current OFAT design, it is particularly well suited to meet the requirements for the execution of a formal experimental design.

Formal Experimental Design Applied to Balance Calibration

LaRC has been conducting research in a "modern design of experiments" (MDOE) approach to force balance calibration. Force balance calibration is an experiment in which independent variables (applied loads) are set and dependent variables (balance electrical responses) are measured. A six dimensional response surface that models the electrical responses as a function of applied load is desired. Formal experimental design techniques provide an integrated view to the force balance calibration process. This scientific approach applies to all three major aspects of an experiment; the design of the experiment, the execution of the experiment, and the statistical analyses of the data.

Formal experimental design has been applied to wind tunnel testing at NASA LaRC since January 1997.[9] The application of formal experiment design to balance calibration was first considered in 1980 at NASA LaRC.[10] A classical quadratic experimental design was conducted in 1980, and the results appeared promising. However, due to the complex load application requirements of an MDOE design, this effort was considered not to be feasible for general application, and the idea remained dormant for nearly two decades. Then, the single-vector calibration method was conceived to provide an efficient means to execute a formal experimental design. To satisfy the requirements of this unique load application system, a custom MDOE calibration design was developed.

An MDOE approach deviates from the current trend of collecting massive data volume, by specifying ample data to meet requirements quantified in the design without prescribing volumes of data far in excess of this minimum. The goal is to efficiently achieve the primary objective of the calibration experiment; namely the determination of an accurate mathematical model to estimate the aerodynamic loads from measured balance responses.

Modern Design of Experiments Fundamentals

The three fundamental quality-assurance principles of MDOE are randomization, blocking, and replication.[11] Randomization of point ordering ensures that a given balance load is just as likely to be applied early as late. If sample means are stable, the point ordering does not matter. However, if some systematic variation (instrumentation drift, temperature effects, operator fatigue, etc) causes earlier measurements to be biased low and later measurements to be biased high then randomization converts such unseen systematic errors to an additional component of simple random error. Random error is easy to detect and also easy to correct by replication and other means. Randomization of point ordering also increases the statistical independence of each data point in the design. This statistical independence is often assumed to exist in the current methods of balance calibration, but systematic variation can cause measurement errors to be correlated and therefore not independent of each other, as required for standard precision interval computations and other common variance estimates to be valid. Even relatively mild correlation can corrupt variance estimates substantially, introducing significant errors into estimates of "95% confidence intervals" and other such quality metrics.

Blocking entails organizing the design into relatively short blocks of time within which the randomization of point ordering ensures stable sample means and statistical independence of measurements. While randomization defends against systematic within-block variation, substantial between-block systematic variation is also possible. For example, calibrations spanning days or weeks might involve different operators, who each use slightly different techniques, or possess somewhat different skill levels. By blocking the design, it is possible in the analysis to remove these between-block components of what would otherwise be unexplained variance.

Replication causes random errors to cancel. This includes otherwise undetectable systematic variation that is converted to random error by randomizing the point order of the loading schedule. Replication also facilitates unbiased estimates of what is called "pure error" - the error component due to ordinary chance variations in the data. These pure-error estimates are critical to evaluating the quality of the calibration model by permitting the fit of the model to the data to be assessed objectively.

Pure-error estimates make it possible to separate unexplained variance into lack-of-fit and pure-error components. The basis of investigating higher order math models is that the lack-of-fit component is significant in current quadratic models, meaning large compared to the pure-error component. If the lack-of-fit is significant, then higher order models can be used to improve balance load estimation quality. If the lack-of-fit is not significant, then adding additional terms to the math model is not justified.

The volume of data required in an MDOE calibration design depends on three primary parameters: 1) the repeatability of the measurement environment, 2) the precision requirement, and 3) the inference error risk. The repeatability of the measurement environment is a function of how repeatable the independent variables can be set, and the balance responses can be measured. The precision requirement is commonly thought of as the required balance accuracy, or the desired quality of the prediction from the math model. Inference error risk refers to the probability that the application of the derived mathematical model used to predict a response would produce a result outside of the precision requirement. Combining these parameters with the form of the mathematical model and the constraints involved in the execution of the design provide the experiment designer with the necessary information to define the data volume required for an adequate design.

Prior to the execution of the design, the quality of the design is evaluated. The goal of this evaluation is to verify the adequacy of the design to estimate all of the required coefficients in the proposed math model. This phase does not require experimental data; rather it is an evaluation of the design itself. There are a number of measures that provide insight into the design that include the distribution of unit standard error, and the computation of variance inflation factors that are a measure of co-linearity between the terms in the model.

The analysis of the data obtained from an MDOE design is quite different from current balance data processing. This analysis of the data involves statistical tools, used in tandem, with the experience of the balance engineer to objectively determine the calibration coefficients. The number of terms in the model is minimized by eliminating those with coefficients that are too small to resolve with a sufficiently high level of confidence. Minimizing the number of terms in the model lowers the average variance, because each coefficient carries some uncertainty.

Also, the total unexplained variance is partitioned into the pure-error and lack-of-fit components. It is generally agreed that we cannot fit the data with a model better than we can repeat the experimental data. An analysis of unexplained variance provides a method to make objective judgements about the adequacy of the model and the potential for improving the model with higher order terms.

This is merely a sampling of the full "toolbox" of statistical analysis and diagnostic techniques that can be applied to the data obtained from an MDOE design. It is important to note that the application of these techniques is tied directly to the manner in which the experiment is performed. With such procedures as randomization and blocking, these techniques produce high-quality results that defend against the kinds of systematic variations that are common in long-duration experiments such as a balance calibration. Many of these analysis techniques are demonstrated later in the analyses of data section in this paper.

Development of the Single-Vector Calibration Design

The generation of the experimental design for the single vector system proved challenging due to the unique physics-based constraint imposed on the set point combinations. An orthogonally blocked third order experiment design was required for six independent variables (*NF, AF, PM, RM, YM, SF*) that could be used to determine force balance calibration mathematical models for each response. The experiment design contains four blocks of thirty-one points each. It features sufficient degrees of freedom to partition the total residual mean square error into pure-error and lack-of-fit error components. All points within the design satisfy the physics-based loading constraint. Furthermore, the third order portion of the design can be performed sequentially after the second order portion. This allows the second order model to be determined and then a decision made about the necessity of higher order terms.

A literature search was performed to determine the state-of-the-art in third order designs. There was a modestly large amount of literature describing third order designs. However, the references in the literature for six variable designs are a little more modest. In order to satisfy the physics-based constraint, to make the force and moment vectors orthogonal to one another, there were no references found.

The approach taken was to generate the design in two different parts, treating the moment and force vectors separately. Focussing on the three moment vector variables a central composite design was generated in the moment vector variables. For each point in the moment vector design, there are two dimensions in the force vector variables that are orthogonal to the moment vector (that is, there is a two-dimensional space orthogonal to every vector in three dimensions). Then a central composite design was generated in the two orthogonal dimensions for each point in the moment vector variable design. This process was then repeated in the force vector variables.

Following the procedure described above generates 112 points in the quadratic design. For symmetry, the same process is carried out interchanging *RM,PM,YM* and *AF,SF,NF* to generate an additional 112 points for a total of 224 points in the cubic design. It is desirable to execute the calibration experiment sequentially. That is, a quadratic design can be executed first followed by the remaining points of a cubic design if a lack-of-fit test indicates that a third order model is necessary. An approach that can be taken in generating a sequential design is to assign the factorial portion of the design to the quadratic portion of the model and the axial points to the cubic portion. The points in the first four blocks contain the quadratic portion of the design and the remaining four blocks can be added to allow estimation of the cubic terms.

The design above was generated subject to the physics-based constraint in Equation 2. This is a constraint in the cross-product terms, so only two of the three cross-product terms in the constraint equation can be estimated at one time. One regression degree of freedom is lost due to this constraint in a quadratic model. This constraint equation can be multiplied by any of the six variables. For example, multiply it by AF produces the equation below.

$$-(RM)(AF)(AF) + (PM)(SF)(AF) - (YM)(NF)(AF) = 0 \quad (4)$$

Only two of the three third order terms can be estimated at one time. That happens similarly for the other five variables, and a total of six of the third order parameters cannot be estimated.

The loss of the single degree of freedom in the quadratic design and six additional degrees of freedom in the cubic design only poses a problem if all of these coefficients are significant in the model. In other words, for the quadratic model, if at least one of the three two-way interaction terms in Equation 2 is equal to zero, then all of the significant model coefficients can be estimated. In the experimental testing to date, this has been the case. For the case of all three interactions being significant to the model, a generalized solution approach was required. Two different approaches have been derived. First, an additional hardware application sub-system was devised that allows for the setting of specific combinations of independent variables that deviate from the constraint imposed by Equation 2. This is not considered the optimum approach since it deteriorates the simplicity of the SVS. Alternatively, an iterative approach to estimating the model coefficients based on all possible permutations of selecting two of the three cross-product interactions at one time can also be employed. The model coefficients can be estimated based on setting one of the three interactions to zero. Using this approach provides two estimates of each of three interactions in the quadratic model. The balance engineer can then decide which combination of these interactions provides the most useful model. Using one of these two approaches will enable the estimation of all significant quadratic and cubic interactions without the limitation caused by Equation 2.

The quadratic design for the balance used in the experimental testing is provided in Table I. The design is listed in standard order, but was executed based on a randomization of the points within each of the four orthogonal blocks.

A common visualization scheme used in the review of balance calibration designs is based on plotting all possible two variable combinations. For a six component balance, this requires 15 plots. The LaRC OFAT 729-point design and the MDOE 124-point design are plotted using this technique in Figure 5 and 6, respectively. These plots contain the values of the actual set points of the independent variables. It is obvious from these plots that the MDOE 124-point is better distributed and more symmetric than the OFAT 729-point design.

There are two significant aspects of the MDOE 124-point design that can not be seen using this plotting method. First, a comparison of the point ordering cannot be seen. The MDOE 124-point design randomly samples these variable combinations within a block as previously described. The OFAT 729-point design methodically moves along the series of points on these plots. Second, the MDOE 124-point is actually distributed in six dimensional space which makes visualization difficult. This six dimensional distribution can be seen in Table I, which consists of many multiple component combinations in contrast to the OFAT 729-point design, which consists primarily of one and two component combinations.

(a) Blocks one and two of the design.

Table I. MDOE 124-point quadratic design for Balance NTF-107.

Standard		Coded Variables (percent of full-scale load)						Actual Units (pounds or inch-pounds)					
Order	Block	NF	AF	PM	RM	YM	SF	NF	AF	PM	RM	YM	SF
1	1	-0.22	-0.87	-0.71	-0.71	-0.71	-0.43	-35	-44	-177	-71	-88	-35
2	1	0.22	-0.87	-0.71	-0.71	0.71	-0.43	35	-44	-177	-71	88	-35
3	1	0.33	-0.94	0.71	-0.71	-0.71	-0.09	52	-47	177	-71	-88	-7
4	1	-0.33	-0.94	0.71	-0.71	0.71	-0.09	-52	-47	177	-71	88	-7
5	1	-0.33	-0.94	-0.71	0.71	-0.71	-0.09	-52	-47	-177	71	-88	-7
6	1	0.33	-0.94	-0.71	0.71	0.71	-0.09	52	-47	-177	71	88	-7
7	1	0.22	-0.87	0.71	0.71	-0.71	-0.43	35	-44	177	71	-88	-35
8	1	-0.22	-0.87	0.71	0.71	0.71	-0.43	-35	-44	177	71	88	-35
9	1	-0.45	0.00	0.00	-1.00	0.00	-0.89	-72	0	0	-100	0	-72
10	1	-0.45	0.00	0.00	1.00	0.00	-0.89	-72	0	0	100	0	-72
11	1	-0.30	-0.95	-1.00	0.00	0.00	0.00	-48	-48	-250	0	0	0
12	1	-0.30	-0.95	1.00	0.00	0.00	0.00	-48	-48	250	0	0	0
13	1	0.00	-0.85	0.00	0.00	-1.00	-0.53	0	-42	0	0	-125	-42
14	1	0.00	-0.85	0.00	0.00	1.00	-0.53	0	-42	0	0	125	-42
15	1	-0.71	-0.71	-0.38	-0.91	-0.16	-0.71	-113	-35	-96	-91	-20	-57
16	1	0.71	-0.71	-0.38	-0.91	0.16	-0.71	113	-35	-96	-91	20	-57
17	1	-0.71	-0.71	-0.27	-0.84	0.48	0.71	-113	-35	-67	-84	60	57
18	1	0.71	-0.71	-0.27	-0.84	-0.48	0.71	113	-35	-67	-84	-60	57
19	1	-0.71	0.71	-0.27	-0.84	-0.48	-0.71	-113	35	-67	-84	-60	-57
20	1	0.71	0.71	-0.27	-0.84	0.48	-0.71	113	35	-67	-84	60	-57
21	1	-0.71	0.71	-0.38	-0.91	0.16	0.71	-113	35	-96	-91	20	57
22	1	0.71	0.71	-0.38	-0.91	-0.16	0.71	113	35	-96	-91	-20	57
23	1	0.00	-1.00	-0.45	0.00	-0.89	0.00	0	-50	-112	0	-112	0
24	1	0.00	1.00	-0.45	0.00	-0.89	0.00	0	50	-112	0	-112	0
25	1	0.00	0.00	0.00	-0.78	-0.62	-1.00	0	0	0	-78	-78	-80
26	1	0.00	0.00	0.00	-0.78	-0.62	1.00	0	0	0	-78	-78	80
27	1	-1.00	0.00	-0.37	-0.93	0.00	0.00	-160	0	-93	-93	0	0
28	1	1.00	0.00	-0.37	-0.93	0.00	0.00	160	0	-93	-93	0	0
29	1	0.00	0.00	0.00	0.00	0.00	0.00	0	0	0	0	0	0
30	1	0.00	0.00	0.00	0.00	0.00	0.00	0	0	0	0	0	0
31	1	0.00	0.00	0.00	0.00	0.00	0.00	0	0	0	0	0	0
32	2	-0.33	0.94	-0.71	-0.71	-0.71	-0.09	-52	47	-177	-71	-88	-7
33	2	0.33	0.94	-0.71	-0.71	0.71	-0.09	52	47	-177	-71	88	-7
34	2	0.22	0.87	0.71	-0.71	-0.71	-0.43	35	44	177	-71	-88	-35
35	2	-0.22	0.87	0.71	-0.71	0.71	-0.43	-35	44	177	-71	88	-35
36	2	-0.22	0.87	-0.71	0.71	-0.71	-0.43	-35	44	-177	71	-88	-35
37	2	0.22	0.87	-0.71	0.71	0.71	-0.43	35	44	-177	71	88	-35
38	2	0.33	0.94	0.71	0.71	-0.71	-0.09	52	47	177	71	-88	-7
39	2	-0.33	0.94	0.71	0.71	0.71	-0.09	-52	47	177	71	88	-7
40	2	0.45	0.00	0.00	-1.00	0.00	-0.89	72	0	0	-100	0	-72
41	2	0.45	0.00	0.00	1.00	0.00	-0.89	72	0	0	100	0	-72
42	2	-0.30	0.95	-1.00	0.00	0.00	0.00	-48	48	-250	0	0	0
43	2	-0.30	0.95	1.00	0.00	0.00	0.00	-48	48	250	0	0	0
44	2	0.00	0.85	0.00	0.00	-1.00	-0.53	0	42	0	0	-125	-42
45	2	0.00	0.85	0.00	0.00	1.00	-0.53	0	42	0	0	125	-42
46	2	-0.71	-0.71	-0.27	0.84	-0.48	-0.71	-113	-35	-67	84	-60	-57
47	2	0.71	-0.71	-0.27	0.84	0.48	-0.71	113	-35	-67	84	60	-57
48	2	-0.71	-0.71	-0.38	0.91	0.16	0.71	-113	-35	-96	91	20	57
49	2	0.71	-0.71	-0.38	0.91	-0.16	0.71	113	-35	-96	91	-20	57
50	2	-0.71	0.71	-0.38	0.91	-0.16	-0.71	-113	35	-96	91	-20	-57
51	2	0.71	0.71	-0.38	0.91	0.16	-0.71	113	35	-96	91	20	-57
52	2	-0.71	0.71	-0.27	0.84	0.48	0.71	-113	35	-67	84	60	57
53	2	0.71	0.71	-0.27	0.84	-0.48	0.71	113	35	-67	84	-60	57
54	2	0.00	-1.00	-0.45	0.00	0.89	0.00	0	-50	-112	0	112	0
55	2	0.00	1.00	-0.45	0.00	0.89	0.00	0	50	-112	0	112	0
56	2	0.00	0.00	0.00	0.78	-0.62	-1.00	0	0	0	78	-78	-80
57	2	0.00	0.00	0.00	0.78	-0.62	1.00	0	0	0	78	-78	80
58	2	-1.00	0.00	-0.37	0.93	0.00	0.00	-160	0	-93	93	0	0
59	2	1.00	0.00	-0.37	0.93	0.00	0.00	160	0	-93	93	0	0
60	2	0.00	0.00	0.00	0.00	0.00	0.00	0	0	0	0	0	0
61	2	0.00	0.00	0.00	0.00	0.00	0.00	0	0	0	0	0	0
62	2	0.00	0.00	0.00	0.00	0.00	0.00	0	0	0	0	0	0

(b) Blocks three and four of the design.

Table 1. concluded.

Standard Order	Block	Coded Variables (percent of full-scale load) NF	AF	PM	RM	YM	SF	Actual Units (pounds or inch-pounds) NF	AF	PM	RM	YM	SF
63	3	0.33	-0.94	-0.71	-0.71	-0.71	0.09	52	-47	-177	-71	-88	7
64	3	-0.33	-0.94	-0.71	-0.71	0.71	0.09	-52	-47	-177	-71	88	7
65	3	-0.22	-0.87	0.71	-0.71	-0.71	0.43	-35	-44	177	-71	-88	35
66	3	0.22	-0.87	0.71	-0.71	0.71	0.43	35	-44	177	-71	88	35
67	3	0.22	-0.87	-0.71	0.71	-0.71	0.43	35	-44	-177	71	-88	35
68	3	-0.22	-0.87	-0.71	0.71	0.71	0.43	-35	-44	-177	71	88	35
69	3	-0.33	-0.94	0.71	0.71	-0.71	0.09	-52	-47	177	71	-88	7
70	3	0.33	-0.94	0.71	0.71	0.71	0.09	52	-47	177	71	88	7
71	3	-0.45	0.00	0.00	-1.00	0.00	0.89	-72	0	0	-100	0	72
72	3	-0.45	0.00	0.00	1.00	0.00	0.89	-72	0	0	100	0	72
73	3	0.30	-0.95	-1.00	0.00	0.00	0.00	48	-48	-250	0	0	0
74	3	0.30	-0.95	1.00	0.00	0.00	0.00	48	-48	250	0	0	0
75	3	0.00	-0.85	0.00	0.00	-1.00	0.53	0	-42	0	0	-125	42
76	3	0.00	-0.85	0.00	0.00	1.00	0.53	0	-42	0	0	125	42
77	3	-0.71	-0.71	0.27	-0.84	0.48	-0.71	-113	-35	67	-84	60	-57
78	3	0.71	-0.71	0.27	-0.84	-0.48	-0.71	113	-35	67	-84	-60	-57
79	3	-0.71	-0.71	0.38	-0.91	-0.16	0.71	-113	-35	96	-91	-20	57
80	3	0.71	-0.71	0.38	-0.91	0.16	0.71	113	-35	96	-91	20	57
81	3	-0.71	0.71	0.38	-0.91	0.16	-0.71	-113	35	96	-91	20	-57
82	3	0.71	0.71	0.38	-0.91	-0.16	-0.71	113	35	96	-91	-20	-57
83	3	-0.71	0.71	0.27	-0.84	-0.48	0.71	-113	35	67	-84	-60	57
84	3	0.71	0.71	0.27	-0.84	0.48	0.71	113	35	67	-84	60	57
85	3	0.00	-1.00	0.45	0.00	-0.89	0.00	0	-50	112	0	-112	0
86	3	0.00	1.00	0.45	0.00	-0.89	0.00	0	50	112	0	-112	0
87	3	0.00	0.00	0.00	-0.78	0.62	-1.00	0	0	0	-78	78	-80
88	3	0.00	0.00	0.00	-0.78	0.62	1.00	0	0	0	-78	78	80
89	3	-1.00	0.00	0.37	-0.93	0.00	0.00	-160	0	93	-93	0	0
90	3	1.00	0.00	0.37	-0.93	0.00	0.00	160	0	93	-93	0	0
91	3	0.00	0.00	0.00	0.00	0.00	0.00	0	0	0	0	0	0
92	3	0.00	0.00	0.00	0.00	0.00	0.00	0	0	0	0	0	0
93	3	0.00	0.00	0.00	0.00	0.00	0.00	0	0	0	0	0	0
94	4	0.22	0.87	-0.71	-0.71	-0.71	0.43	35	44	-177	-71	-88	35
95	4	-0.22	0.87	-0.71	-0.71	0.71	0.43	-35	44	-177	-71	88	35
96	4	-0.33	0.94	0.71	-0.71	-0.71	0.09	-52	47	177	-71	-88	7
97	4	0.33	0.94	0.71	-0.71	0.71	0.09	52	47	177	-71	88	7
98	4	0.33	0.94	-0.71	0.71	-0.71	0.09	52	47	-177	71	-88	7
99	4	-0.33	0.94	-0.71	0.71	0.71	0.09	-52	47	-177	71	88	7
100	4	-0.22	0.87	0.71	0.71	-0.71	0.43	-35	44	177	71	-88	35
101	4	0.22	0.87	0.71	0.71	0.71	0.43	35	44	177	71	88	35
102	4	0.45	0.00	0.00	-1.00	0.00	0.89	72	0	0	-100	0	72
103	4	0.45	0.00	0.00	1.00	0.00	0.89	72	0	0	100	0	72
104	4	0.30	0.95	-1.00	0.00	0.00	0.00	48	48	-250	0	0	0
105	4	0.30	0.95	1.00	0.00	0.00	0.00	48	48	250	0	0	0
106	4	0.00	0.85	0.00	0.00	-1.00	0.53	0	42	0	0	-125	42
107	4	0.00	0.85	0.00	0.00	1.00	0.53	0	42	0	0	125	42
108	4	-0.71	-0.71	0.38	0.91	0.16	-0.71	-113	-35	96	91	20	-57
109	4	0.71	-0.71	0.38	0.91	-0.16	-0.71	113	-35	96	91	-20	-57
110	4	-0.71	-0.71	0.27	0.84	-0.48	0.71	-113	-35	67	84	-60	57
111	4	0.71	-0.71	0.27	0.84	0.48	0.71	113	-35	67	84	60	57
112	4	-0.71	0.71	0.27	0.84	0.48	-0.71	-113	35	67	84	60	-57
113	4	0.71	0.71	0.27	0.84	-0.48	-0.71	113	35	67	84	-60	-57
114	4	-0.71	0.71	0.38	0.91	-0.16	0.71	-113	35	96	91	-20	57
115	4	0.71	0.71	0.38	0.91	0.16	0.71	113	35	96	91	20	57
116	4	0.00	-1.00	0.45	0.00	0.89	0.00	0	-50	112	0	112	0
117	4	0.00	1.00	0.45	0.00	0.89	0.00	0	50	112	0	112	0
118	4	0.00	0.00	0.00	0.78	0.62	-1.00	0	0	0	78	78	-80
119	4	0.00	0.00	0.00	0.78	0.62	1.00	0	0	0	78	78	80
120	4	-1.00	0.00	0.37	0.93	0.00	0.00	-160	0	93	93	0	0
121	4	1.00	0.00	0.37	0.93	0.00	0.00	160	0	93	93	0	0
122	4	0.00	0.00	0.00	0.00	0.00	0.00	0	0	0	0	0	0
123	4	0.00	0.00	0.00	0.00	0.00	0.00	0	0	0	0	0	0
124	4	0.00	0.00	0.00	0.00	0.00	0.00	0	0	0	0	0	0

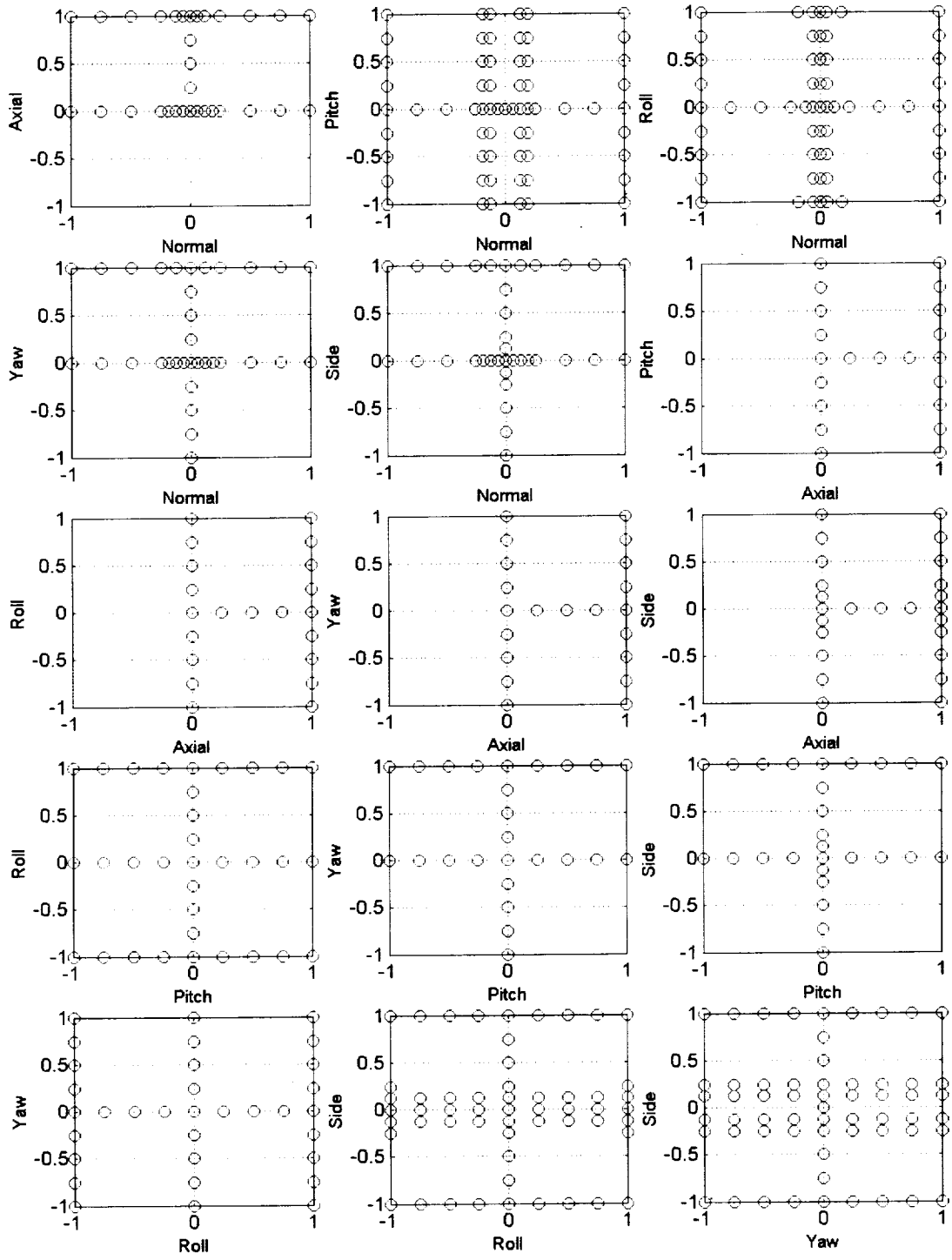


Figure 5. OFAT 729-point design for Balance NTF-107.
Component 1 (percent of full-scale load) versus Component 2 (percent of full-scale load).

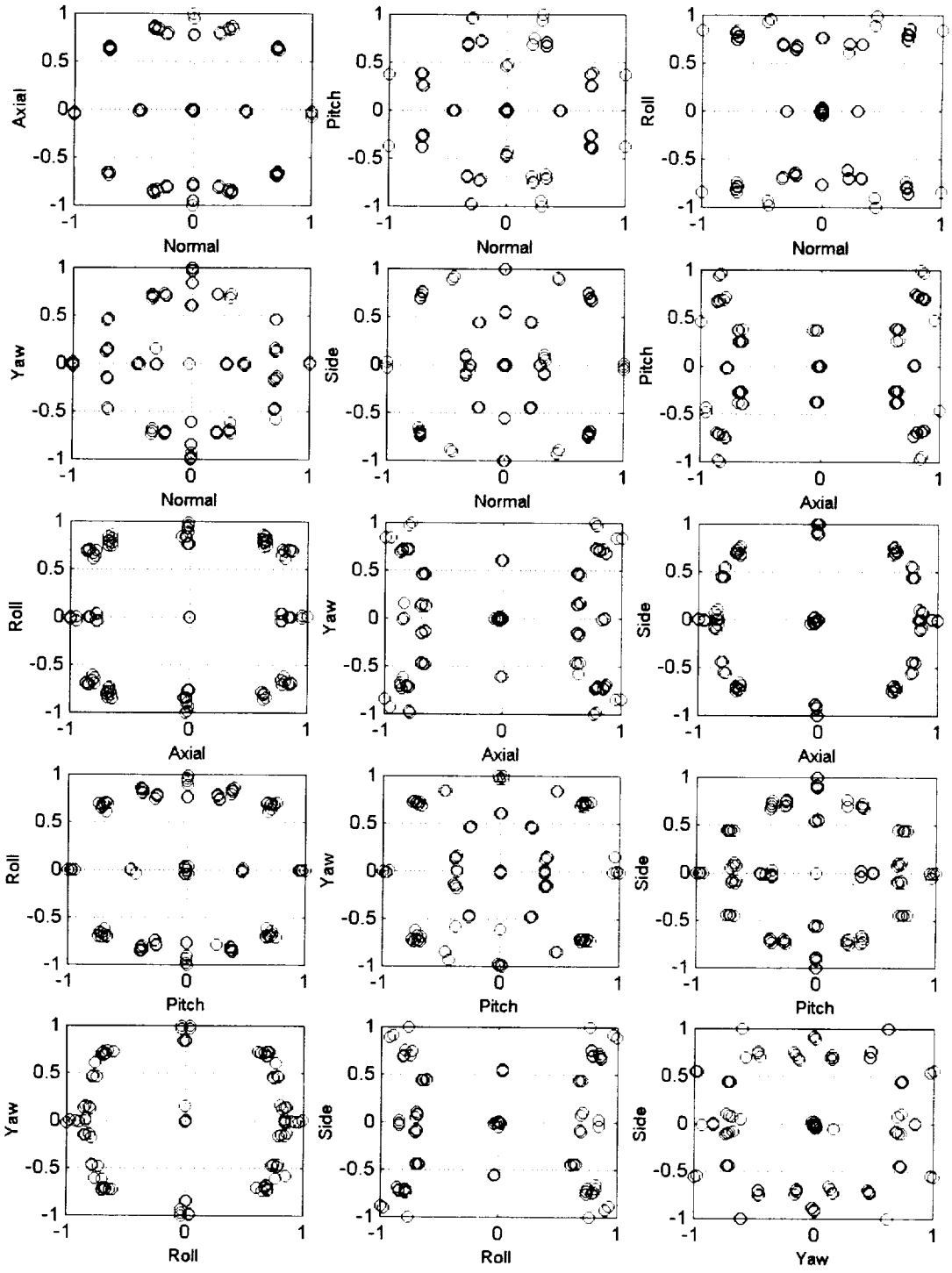


Figure 6. MDOE 124-point quadratic design for Balance NTF-107.
Component 1 (percent of full-scale load) versus Component 2 (percent of full-scale load).

A statistical summary of the quadratic design is provided in Table II.

Table II. Summary of the 124-point quadratic design.

Total Degrees of Freedom	123
Regression Degrees of Freedom	26
Pure Error Degrees of Freedom	11
Lack of Fit Degrees of Freedom	86
Average Leverage	0.217
Average Prediction Variance	1.217
Maximum Prediction Variance	1.264

In the event that the loss in regression degrees of freedom results in an inadequate model, as previously discussed, an augmented design was created. This augmented design adds twelve additional points to the original design. In order to set these variable combinations during the experiment, an additional load application system was devised that allows for the application of a co-linear resultant force and moment vector. This load application system utilizes a novel system of permanent magnets arranged in a repelling orientation to generate a non-contact force vector that is perpendicular to the gravitational force vector.

A summary of the augmented design for the quadratic model is provided in Table III.

Table III. Summary of the 136-point quadratic augmented design.

Total Degrees of Freedom	135
Regression Degrees of Freedom	27
Pure Error Degrees of Freedom	11
Lack of Fit Degrees of Freedom	97
Average Leverage	0.206
Average Prediction Variance	1.206
Maximum Prediction Variance	1.384

Experimental Testing

A sequence of experimental tests was performed to compare the SVS to the LaRC manual calibration system. The goals of these tests were to verify the performance of the SVS and to compare the mathematical model derived from the MDOE based design to the OFAT design. The preparation of the test balance and LaRC standard calibration is described followed by a presentation of the experimental testing of the SVS. This is followed by analyses of the data from these two methodologies.

Experimental Apparatus

The balance selected for the experimental testing was the NTF-107. The NTF-107 is a state-of-the-art six component balance of monolithic construction. The full-scale load range of the balance is provided in Table IV.

Table IV. Test balance full-scale load range.

Component	**Full-scale design load**
Normal Force (pounds)	160
Axial Force (pounds)	50
Pitching Moment (inch-pounds)	250
Rolling Moment (inch-pounds)	100
Yawing Moment (inch-pounds)	125
Side Force (pounds)	80

The balance was refurbished prior to testing to ensure that its performance would not be questionable during the series of experiments performed. This included a refurbishment of the mechanical interfaces and the installation of new instrumentation. Balance stability and repeatability were considered vital to the comparison, since the testing took place over an eight-month period. This particular balance was chosen due to its availability and load range, which were compatible with the testing schedule and load application system design.

A LaRC manual calibration was performed on the NTF-107 in March 2000. This was a full calibration that included all 729 data points in the OFAT calibration design as well as proof load sequences. The calibration required three weeks to perform. This manual calibration was considered the baseline for comparison to the SVS. Figure 2 is a photograph of the experimental setup for the LaRC manual calibration.

The NTF-107 was then calibrated using the SVS in November 2000. Three calibrations were performed and two different calibration designs were evaluated. The results presented in this paper are from the MDOE 124-point quadratic design that has been previously presented. The points within each of the four blocks were randomized to ensure statistical independence. Each of the four orthogonal blocks required approximately four hours to complete. Prior to the execution of the design, a tare sequence consisting of 20 data points was performed. Using a math model derived from this tare sequence eliminates the need to reposition the balance at each data point. The SVS

calibration required a total of three days to complete. The SVS experimental apparatus are shown in Figure 3.

Analyses of the data

The use of formal experimental design methodology enables the application of numerous analyses and diagnostics that have not been utilized for the review of balance calibration data. Many of these analysis techniques are valid only because of the way in which the experiment was conducted. This makes it difficult to directly compare the results from the OFAT calibration to the MDOE calibration. Also, these new analyses are probably unfamiliar to those within the balance community. Therefore, a mixture of conventional calibration reporting methods and the MDOE methods will be presented. Conventional OFAT calibration results include a comparison of the math model coefficients, plots of the residual errors versus data point in the design, and a two-sigma estimate of the residual errors.

There are many sophisticated methods of analyzing the calibration data from the MDOE design. It is not within the scope of this paper to present all of the possible analyses that are available. Rather, selections have been made that will highlight the new types of insight that can be provided to the balance engineer. These examples include the determination of the minimum number of model coefficients and an analysis of the unexplained variance.

Conventional OFAT experiment analyses

This section provides the results from the SVS, using conventional reporting methods. The coefficients that are generated by the mathematical model are not merely numerical values; they have a physical meaning. Experienced balance engineers not only review the mathematical model from the viewpoint of curve fitting, but also from the perspective of the physics behind the balance response. This knowledge of the balance construction and physical performance is invaluable in reviewing the math model.

LaRC force balances are of similar physical construction and have typical interaction patterns that are expected. Model coefficients are typically expressed as a percent of full-scale effect. This enables them to be reviewed based on their relative magnitude to each other. Figure 7 illustrates all of the coefficients from the quadratic model for each balance response. There are 26 coefficients per component displayed on each plot. The main effect, or sensitivity coefficient, is provided in Table V and the intercepts are not shown. It can be seen that the coefficients derived from the MDOE calibration match closely with those from the OFAT calibration.

Table V reveals that the coefficients from the LaRC manual calibration are consistently higher than the coefficients from SVS calibration. This difference was investigated by performing a simple load sequence using the LaRC manual calibration hardware with the balance connected to the SVS data acquisition system. Then, the same load sequence was performed using the SVS. The sensitivity coefficients derived from these load sequences agreed within 0.01% of full-scale response. This agreement, on the order of the resolution, eliminates the load apparatus as the source of the difference in the sensitivity coefficients. At this time, it is suspected that instrumentation differences between the data acquisition systems that were used for the two calibrations are the source of the discrepancy.

At the resolution of Figure 7, it is difficult to discern the minor differences important to the level of precision obtained from force balance measurements. Therefore, Figure 8 depicts the difference between the coefficients. The largest differences on these plots involve the PM term on the YM response, and the NF term on the SF response, and are less than 0.20% of the full-scale response. This difference is most likely due to different level reference surfaces used during the LaRC manual calibration. During the LaRC manual calibration, five different reference surfaces are used to define the balance coordinate system. In the SVS, a single reference surface is used, and the definition of the orthogonal axes is contained within the angle measurement device. The magnitude of these two terms on Figure 7 reveals that the SVS coefficients are smaller. A coefficient of zero would be theoretically predicted based on the balance design, and this would tend to point to the smaller coefficient as being correct. This qualitative analysis cannot be confirmed unless a detailed mechanical inspection is performed.

Table V. Sensitivity Coefficients, units are pounds or inch-pounds per (microvolt/volt).

Design	Normal	Axial	Pitch	Roll	Yaw	Side
OFAT 729-point	1.00206E-01	4.76009E-02	1.47139E-01	9.49683E-02	1.43170E-01	8.74642E-02
MDOE 124-point	1.00028E-01	4.75262E-02	1.46951E-01	9.48932E-02	1.42941E-01	8.73010E-02

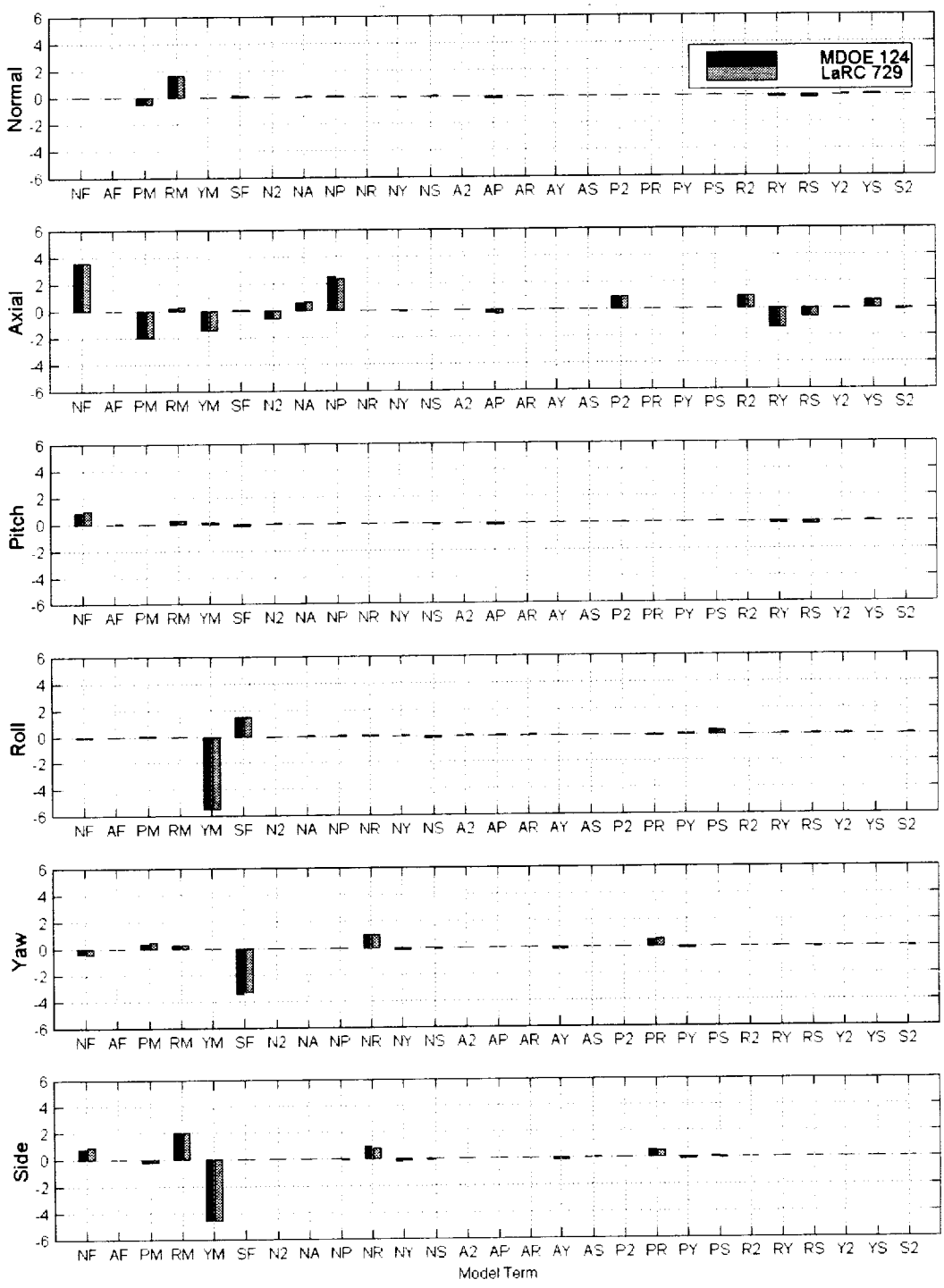


Figure 7. Comparison of calibration model coefficients.
Effect of model term (percent of full-scale response) versus Model term.

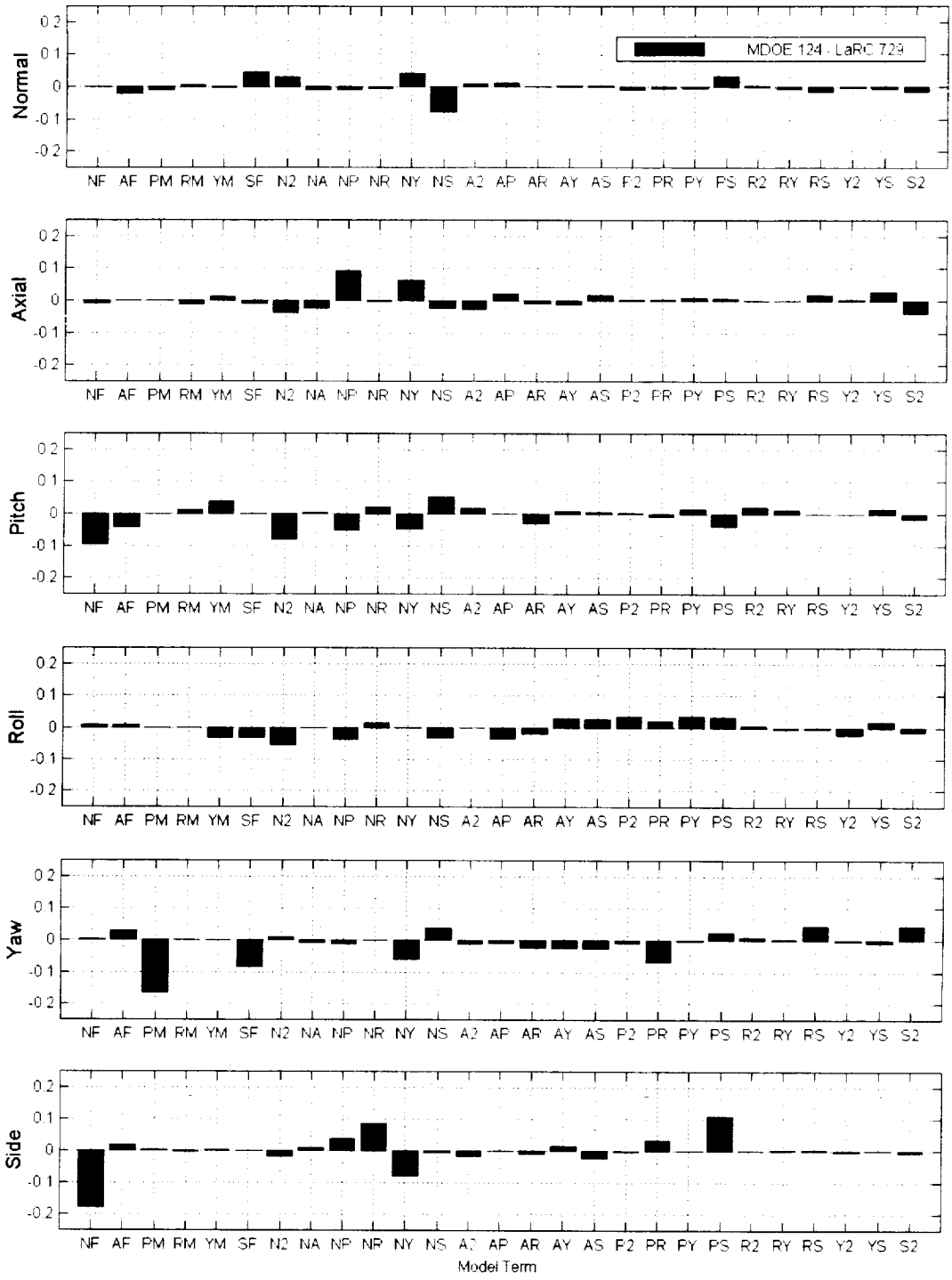


Figure 8. Difference between model coefficients.
Delta effect of model term (percent of full-scale response) versus Model term.

A common plotting scheme that is used to review the quality of the fit of the math model is provided in Figures 9 and 10 for the OFAT 729-point and the MDOE 124-point designs. These plots illustrate the total residual error expressed as a percent of full-scale load for each balance response. It is important to remember that the MDOE 124-point design involves numerous multiple component combinations as compared to the OFAT 729-point design. Typically, the residual errors become considerably larger when the math model derived from the OFAT 729-point design is applied to multiple component combinations. One of the factors that cause this situation is the difficulty in applying orthogonal force vectors with the LaRC manual system, as previously discussed.

Another factor is that the model coefficients are not derived from data that contains multiple component combinations. The OFAT design contains primarily one and two component combinations. We know that the "real" balance calibration model is not quadratic, it is approximated as a quadratic polynomial. Significant higher order terms are believed to exist, and their effects must be absorbed into the available lower order terms in the model. This is not necessarily desirable, but it is required to obtain the best possible model of a specified degree. A technique is available to determine which higher order terms are absorbed into which lower order terms. It involves the computation of the alias matrix[1] and provides excellent insight into this phenomena. It is not within the scope of this paper to provide the results from the alias matrix, simply to point out that a method exists to investigate the design for alias effects. The calibration design must provide combinations of the independent variables that include these higher order terms for their effects to be absorbed into the lower order terms. The MDOE design includes these combinations, and therefore the fit of the model on these multiple component combinations is better. It is also important to note that multiple component combinations are more commonly encountered in the wind tunnel use of the force balance, and this is another benefit of providing them in the calibration design.

A comparison of the two-sigma values for the total residual errors from the two calibration systems is provided in Table VI. This is typically referred to as the balance accuracy, but the potential errors in this calculation due to the lack of statistical independence of the data points within the LaRC OFAT design have been discussed. From this table, the high quality of this particular test balance is apparent since all of the values are less than or equal to 0.10% of full-scale load. These values also compare well between the two calibration systems. To put the resolution of these values in perspective, the weights used during the calibration experiments are accurate to within 0.01% of full-scale magnitude. This means that simply using different physical weights during a single calibration could cause differences on the order of 0.01% of full-scale magnitude.

MDOE experiment analyses

As previously stated, the smallest model with the fewest parameters is desired since each coefficient carries its own uncertainty. The objectives for the determination of the model coefficients are to provide a mathematical model that has insignificant lack-of-fit and meets the precision requirement. An analysis of variance (ANOVA) was performed to achieve these objectives. One aspect of the ANOVA is a systematic method to determine the statistical significance of each model coefficient (parameter). Table VII contains a portion of the data from the ANOVA of the normal force response. In this table, all possible quadratic model terms are provided except the cross product of axial times roll. Due to the physics-based constraint previously discussed, only twenty-six terms can be modeled simultaneously. It was determined that axial times roll had the least significant effect on the normal response from the three cross products in Equation 2, and therefore it was eliminated first. This set of data is used as an example of how a determination is made on which terms to retain in the model. The same procedure applies to the other five response variables.

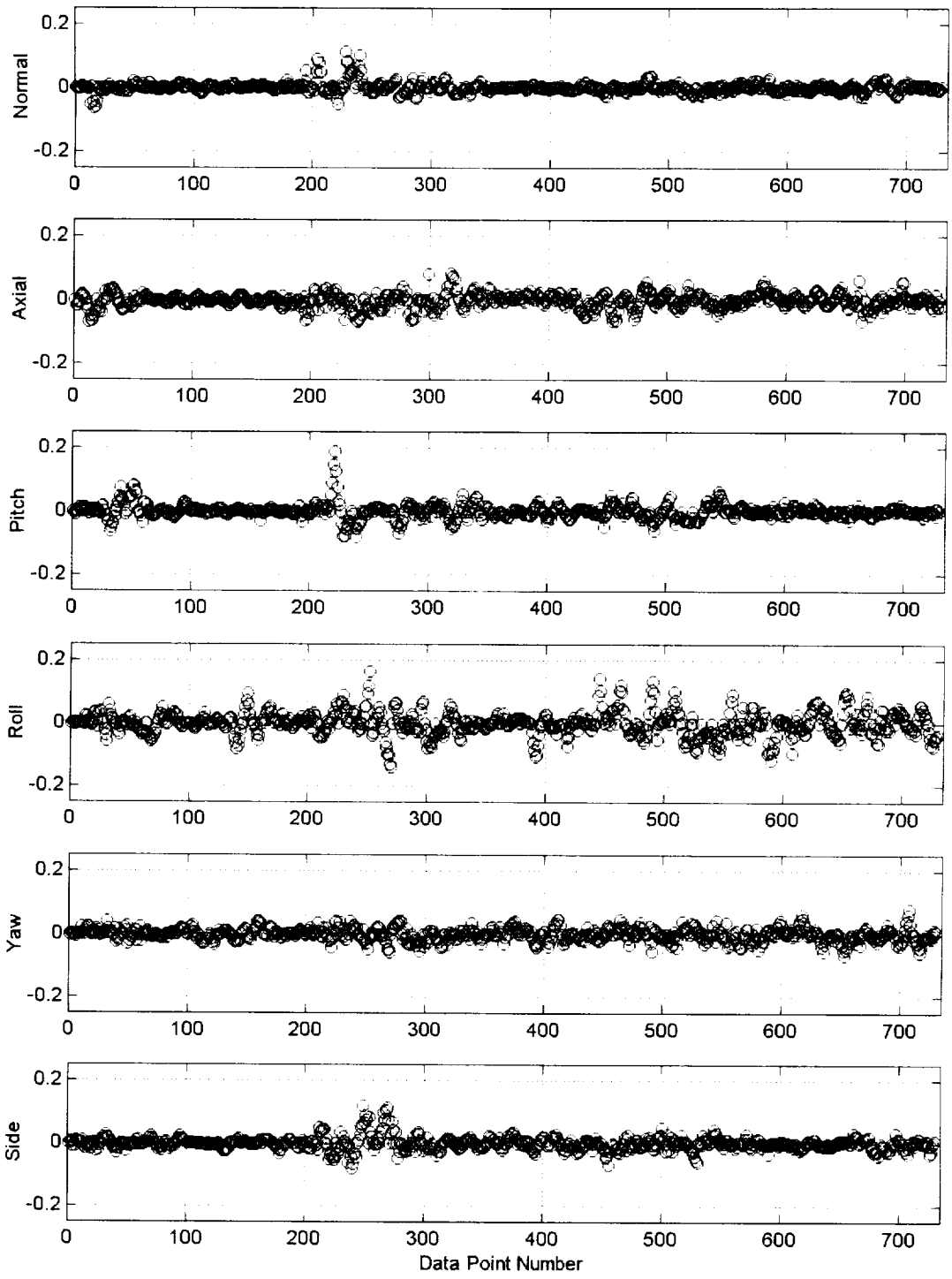


Figure 9. Residual errors from OFAT 729-point calibration experiment.
Total residual (percent of full-scale load) versus Data point.

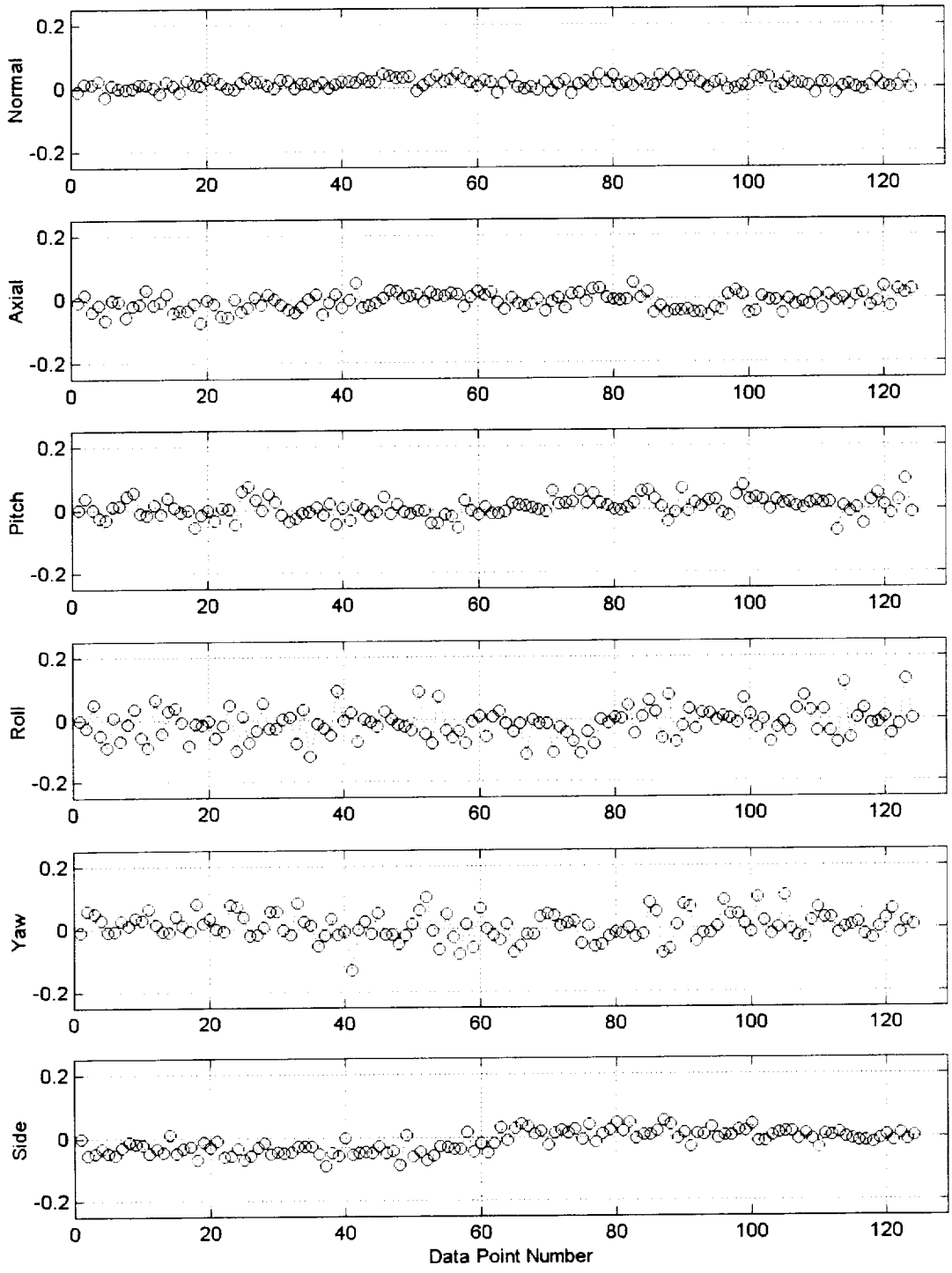


Figure 10. Residual errors from MDOE 124-point calibration experiment.
Total residual (percent of full-scale load) versus Data point.

Table VI. Comparison of two-sigma accuracies.

	Residual Error expressed as two times the standard deviation (95% confidence) (percent of full-scale load)					
Design	**Normal**	**Axial**	**Pitch**	**Roll**	**Yaw**	**Side**
LaRC 729-point	0.03%	0.05%	0.05%	0.08%	0.04%	0.04%
MDOE 124-point	0.03%	0.05%	0.06%	0.10%	0.09%	0.07%

Table VII. Initial model for the normal force response.

Source	DF	Mean Square	F-value	Prob > F
Block	3	1.846E+05		
Model	26	3.272E+06	53,903,167	< 0.0001
N	1	8.326E+07	1,371,566,591	< 0.0001
A	1	3.013E-01	4.96	0.0283
P	1	1.934E+03	31,867	< 0.0001
R	1	2.993E+04	493,064	< 0.0001
Y	1	1.002E+01	165.08	< 0.0001
S	1	1.504E+02	2478.21	< 0.0001
N2	1	5.204E-01	8.57	0.0043
A2	1	4.337E-02	0.71	0.4001
P2	1	1.647E-01	2.71	0.1028
R2	1	6.019E-01	9.92	0.0022
Y2	1	1.581E+00	26.04	< 0.0001
S2	1	4.163E-01	6.86	0.0103
NA	1	3.841E+00	63.27	< 0.0001
NP	1	1.485E+00	24.46	< 0.0001
NR	1	1.682E-02	0.28	0.5998
NY	1	9.687E-04	0.02	0.8997
NS	1	3.558E-02	0.59	0.4459
AP	1	1.632E+02	2687.90	< 0.0001
AY	1	1.476E-02	0.24	0.6231
AS	1	1.250E-03	0.02	0.8862
PR	1	2.173E-01	3.58	0.0616
PY	1	8.436E-04	0.01	0.9064
PS	1	1.845E-02	0.30	0.5827
RY	1	4.322E+01	712.01	< 0.0001
RS	1	3.435E+02	5658.28	< 0.0001
YS	1	4.248E+00	69.97	< 0.0001
Residual	94	6.070E-02		
Lack of Fit	86	6.123E-02	1.11	0.4785
Pure Error	8	5.502E-02		
Total	123			

In the left most column is the source of variance. The sources of variance in this table include the explained variance (model terms and block effects) and the unexplained variance (lack-of-fit and pure-error). The second column contains the number of degrees of freedom (DF) used to estimate the variance from each source. In the third column the variance, or mean square error (MSE), of each model coefficient is provided in units of microvolts per volt quotient squared. The fourth column contains the F-value, which is equal to the ratio, of the variance of each coefficient divided by the residual variance. The right-most column contains the probability that an F-value this large, with the associated degrees of freedom, could have occurred due to chance variations in the data (experimental noise). The smaller this probability, the more confidence that we have that the model coefficient is non-zero. For example, probability values of less than 0.05 suggest less than a 5% probability of a chance occurrence due to noise resulted in this regression coefficient. Stated another way, it is the probability of getting an F-value of this size if the term did not have an effect on the response. If this probability is less than a threshold value then the coefficient is considered significant and is retained in the model. Note the large F-value and associated low probability of the *N* term on the normal force response. This is expected since it represents the sensitivity of that particular component and there is a strong correlation between the application of normal force and the associated normal force response. The threshold probability value used for this experiment was 0.05. This means that there must be greater than a 95% probability that a coefficient is non-zero in order to retain it in the model. All coefficients that were below this probability were removed and the reduced model is shown in Table VIII. It is the goal of this phase to minimize the number of coefficients in the model, because each coefficient carries its own uncertainty. Once the model has been determined, then an analysis of the unexplained variance can be performed.

Table VIII. Reduced model of the normal force response.

Source	DF	Mean Square	F-value	Prob > F
Block	3	1.846E+05		
Model	16	5.317E+06	88,666,344	< 0.0001
N	1	8.383E+07	1,397,975,758	< 0.0001
A	1	3.164E-01	5.28	0.0236
P	1	1.937E+03	32,308	< 0.0001
R	1	3.021E+04	503,733	< 0.0001
Y	1	1.013E+01	169	< 0.0001
S	1	1.514E+02	2,525	< 0.0001
N2	1	3.397E-01	5.66	0.0191
R2	1	4.437E-01	7.40	0.0076
Y2	1	2.657E+00	44.30	< 0.0001
S2	1	2.847E-01	4.75	0.0316
NA	1	3.748E+00	62.49	< 0.0001
NP	1	1.719E+00	28.67	< 0.0001
AP	1	1.641E+02	2,736.44	< 0.0001
RY	1	4.367E+01	728.21	< 0.0001
RS	1	3.494E+02	5,826.65	< 0.0001
YS	1	4.763E+00	79.43	< 0.0001
Residual	104	5.997E-02		
Lack of Fit	96	6.038E-02	1.10	0.4894
Pure Error	8	5.502E-02		
Total	123			

An analysis of the unexplained variance was performed in order to partition it into two components, lack-of-fit and pure-error. Typically, in the field of balance calibration, the accuracy is based on the standard deviation of the residual errors obtained by computing the difference between the actual values and the model predicted values. This is normally expressed as two times the standard deviation providing a 95% confidence interval. As previously mentioned, the assumption of statistical independence of the data points from an OFAT calibration is not valid. Even relatively mild correlation can corrupt variance estimates substantially, introducing significant errors into estimates of 95% confidence intervals. More importantly, the total residual error includes two distinct components, lack-of-fit and pure-error. The lack-of-fit relates to the ability of the math model to capture the response of the balance electrical signals as a function of the independent variables. The pure-error is a function of the repeatability of the measurement environment. This includes factors such as the hardware used to set the independent variables, the data acquisition system, the balance instrumentation, the quality of the mechanical interfaces, and the thermal stability of the calibration laboratory. For the purpose of this analysis, the repeatability of the calibration system and the balance will not be separated. We also know that we can not fit the data better than we can repeat the data. MDOE provides a technique to separate the lack-of-fit and pure-error components of the unexplained variance.

The pure-error component is computed from the genuine replicates that are performed throughout the calibration experiment. In the case of the MDOE 124-point design, there were three replicates per block, totaling twelve replicates in four blocks. Subtracting the mean value of each block from these replicates provides eight degrees of freedom (DF) to estimate the pure-error. The sum of the squared (SS) deviations from the means of each block is computed. The mean square error (MSE), variance, can be computed based on Equation 5.

$$MSE_{pure\ error} = \frac{SS_{pure\ error}}{DF_{pure\ error}} \tag{5}$$

Once the pure-error is known, its contribution to the total residual can be determined. This computation involves subtracting the SS of the pure-error from the SS of the total residual as shown in Equation 6.

$$SS_{lack\ of\ fit} = SS_{total\ residual} - SS_{pure\ error} \tag{6}$$

The MSE of all three quantities (total residual, lack-of-fit, and pure-error) can then be computed using the associated degrees of freedom (DF) and the SS according to Equation 5. The ratio of the MSE of the lack-of-fit divided by the MSE of the pure-error forms the F-value as shown below.

$$F_{value} = \frac{MSE_{lack\ of\ fit}}{MSE_{pure\ error}} \tag{7}$$

This F-value is compared against a critical value of the F-distribution that depends on the degrees of freedom for both lack-of-fit and pure-error and the specified significance of the test, 0.01 in our case. This 0.01 significance level means that if our measured F-statistic exceeds the critical F-value, we can reject the null hypothesis with 99% confidence. The null hypothesis in this case is as follows: *Ho:* The variance of the lack-of-fit is not significant relative to the variance of the pure-error. If the F-value is greater than the critical F-value then the null hypothesis can be rejected. In this case, it can be stated that we have 99% confidence that the lack-of-fit is significant. On the other hand, if the F-value is smaller then the critical F-value, then we would not reject the null hypothesis, concluding that we are unable to detect significant lack-of-fit with our required 99% level of confidence. This F-test procedure provides an objective method for determining whether or not the model has significant lack-of-fit. Significant lack-of-fit means that the calibration response function does not adequately represent the data upon which it is based.

A summary of the results of the above analysis is provided in Table IX. This table includes the values of the quantities discussed in the partitioning of the unexplained variance. The columns contain the associated values for each of the response variables.

The data in this table provides insight into the mathematical model and the physical calibration system. The sigma estimates in the table are computed according to Equation 8.

$$\sigma = \sqrt{MSE} \tag{8}$$

The repeatability of the measurement environment can readily be determined from the sigma row of the pure-error section. For example, the standard deviation of the pure-error for the normal force response is 0.23

microvolts per volt. Comparing this to the full-scale response (maximum response) of normal force provides a repeatability of 0.014% of full-scale. This repeatability of 1.4 parts per 10,000 reveals the high-precision of the force balance and calibration system involved in this experiment.

Insight is also provided by the results of the lack-of-fit test. For example, the normal force response does not indicate significant lack-of-fit. This means that adding higher-order coefficients to model the normal force response would not be justified, because the model currently fits within the experimental error of the measurements. Adding higher-order terms would only attempt to fit noise, which is undesirable for a useful model. In classical balance modeling, higher-order models or models with more coefficients have been used to reduce the total residual error without consideration to their justification. It is true that the more terms that are added to the model, the lower the residual error, but what we desire is a model that adequately represents the response of the balance, not the noise. Conversely, the lack-of-fit is considered significant on the rolling moment response, and therefore a higher-order model could be used to provide a better fit.

It is important to realize that the lack-of-fit test is relative to the pure-error. In the limit, as the pure-error goes to zero, the lack-of-fit F-value goes to infinity. The decision to use a higher-order model is also tied to the end use of the force balance. The real question in the application of the force balance to wind tunnel testing is whether the prediction quality (precision requirement) and prediction risk (inference error risk) are sufficient to estimate the aerodynamic performance coefficients on a scaled wind tunnel model.

In order to make this determination, the precision requirement and the inference error risk must be known. Currently at LaRC, a force transducer is considered to be of the highest attainable quality, if the precision is less than 0.10% of full-scale with a 99.15% probability that any one component is within the precision requirement, meaning that there is no less than a 95% probability that all six components are simultaneously within the precision requirement. The balance engineer must make a decision on the need for higher order models based on the aerodynamic load prediction requirements, not just on the pursuit of providing zero error in the balance measurement.

Table IX. Summary of the analyses of unexplained variance.

Quantity	Normal	Axial	Pitch	Roll	Yaw	Side
Maximum Response	1600	1052	1701	1054	875	917
Lack-of-Fit (SS)	5.7965	7.3089	26.1832	30.0050	15.2744	4.0042
Lack-of-Fit (DF)	96	93	100	101	101	99
Lack-of-Fit (MSE)	0.0604	0.0786	0.2618	0.2971	0.1512	0.0404
Lack-of-Fit (sigma)	0.25	0.28	0.51	0.55	0.39	0.20
Pure-Error (SS)	0.4402	0.7222	0.4480	0.0929	0.0859	0.2743
Pure-Error (DF)	8	8	8	8	8	8
Pure-Error (MSE)	0.0550	0.0903	0.0560	0.0116	0.0107	0.0343
Pure-Error (sigma)	0.23	0.30	0.24	0.11	0.10	0.19
Residual (SS)	6.2367	8.0311	26.6312	30.0768	15.3603	4.2785
Residual (DF)	104	101	108	109	109	107
Residual (MSE)	0.0600	0.0795	0.2466	0.2759	0.1409	0.0400
Residual (sigma)	0.24	0.28	0.50	0.53	0.38	0.20
measured F-value	1.10	0.87	4.68	25.57	14.08	1.18
critical F-value (@ α = 0.01)	4.97	4.97	4.96	4.96	4.96	4.96
Significant LoF (@ α = 0.01)?	No	No	No	Yes	Yes	No

units: maximum response is (microvolts/volt), SS is (microvolts/volt)^2, MSE is (microvolts/volt)^2, sigma is (microvolts/volt)

Future Research

Further research in the refinement of the SVS is actively progressing at NASA LaRC. Refinements to the hardware system include a higher load range system and automation. It is believed at this time that four systems will be required to span the range of LaRC balances, from ten pounds to ten thousand pounds of normal force. Automation of the system has been considered in the areas of non-metric positioning, load point positioning, and load application. Maintaining the accuracy in the setting of the independent variables will be paramount in all design decisions with regard to automation. An automated system could reduce calibration to hours, instead of days. It would also provide the ability to perform repeated calibrations and enable the application of statistical process control methods.

MDOE research efforts include higher-order models, and calibration at temperature. MDOE provides a systematic method for research into better mathematical models. A cubic design has been generated and it is planned to be executed on the test balance. It is common practice to include partitioned quadratic coefficients in a balance math model to improve troublesome coefficient fits. These partitioned quadratic coefficients, often referred to as split terms, are more likely higher order terms, such as cubic. A complete cubic model including pure cubic, quadratic times linear, and all possible three-way interactions will provide a more robust means to handle these higher-order terms.

It is generally known that balance calibration models are a function of temperature. At the present time, all LaRC complete balance calibrations are performed at room temperature (70 degrees Fahrenheit). Few balances operate at room temperature in the wind tunnel environment. An abbreviated OFAT sequence of loads is performed at elevated (170 degrees Fahrenheit) or cryogenic (-290 degrees Fahrenheit) temperature depending on the facility in which the balance will be used. The results of the abbreviated temperature loads are difficult to interpret due to the inability to separate the repeatability of the measurement environment from the actual thermal effects. A calibration design that incorporates balance temperature as an independent variable would provide thermally compensated calibration models.

Concluding Remarks

The experimental results from the Single-Vector Balance Calibration System have provided a proof-of-concept. This new integrated hardware system and calibration design optimizes the calibration process. The benefits of this system and the application of MDOE to balance calibration have been presented. The SVS addresses the productivity limitations of current force balance calibration systems while simultaneously improving the data quality. With the application of MDOE, significantly fewer data points were generated, but the quality of the information that was obtained is higher. The experimental results illustrate that a complete six component calibration can be performed on the present SVS in three-days with nearly an order of magnitude reduction in the number of data points. Formal experimental design techniques provide objective methods to review and report the results of the calibration experiment. They also provide new insights into force balance calibration that has not been available with OFAT calibration experiments.

The SVS provides a significant advancement in force balance calibration technology. The insights that it provides will aide in the advancement of other areas of force balance design and production. Ultimately, the benefit to the research community will be an increased accuracy of force and moment measurement during wind tunnel testing.

Acknowledgements

The authors would like to thank Dick DeLoach for his pioneering efforts in the application of MDOE to wind tunnel testing and his willingness to provide expert assistance during this research effort. We would also like to acknowledge the Langley Wind Tunnel Reinvestment Program who provided funding for this effort, and Tom Stokes and Ron Dupont of Modern Machine and Tool Company for their assistance in the design and fabrication of the prototype hardware system.

References

1) Box, G.E.P.; Draper, N.: *Empirical Model-Building and Response Surfaces*. John Wiley & Sons. 1987.

2) Ferris, A. T.: *Strain Gauge Balance Calibration and Data Reduction at NASA Langley Research Center*. Paper DR-2, First International Symposium on Strain Gauge Balances. October, 1996.

3) Levkovitch, Michael: *Automatic Balance Calibrator at Israel Aircraft Industries*. January 1993.

4) Polansky, L.; and Kutney, J. T. Sr.: *A New Working Automatic Calibration Machine For Wind Tunnel Internal Force Balances*. AIAA 93-2467. June 1993.

5) Levkovitch, Michael: *Accuracy Analysis of the IAI Mark III ABCS*. July 1995.

6) Parker, P.A.; and Rhew R.D.: *A Study of Automatic Balance Calibration System Capabilities*. Second International Symposium on Strain Gauge Balances. May 1999.

7) Marshall, R. and Landman, D.: *An Improved Method for Determining Pitch and Roll Angles Using Accelerometers*. AIAA 2000-2384, June 2000.

8) Beer, F.P.; Johnston, E.R. Jr.: *Vector Mechanics for Engineers*. McGraw-Hill. 1962.

9) DeLoach, R.: *Applications of Modern Experiment Design to Wind Tunnel Testing at NASA Langley Research Center*. AIAA 98-0713, 36th Aerospace Sciences Meeting and Exhibit, Reno, NV. Jan. 1998.

10) 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 number NAS1-15648. August 1980.

11) Montgomery, D.C.: *Design and Analysis of Experiments, 4th edition*. John Wiley & Sons. 1997.