The Use of Absolute-Value Terms in Regression Modeling of Multi-Piece Force Balances

Mark E. Kammeyer*
The Boeing Company, St. Louis, MO, 63166

Norbert M. Ulbrich†
Jacobs Technology Inc., Moffett Field, California 94035

Different aspects of the use of absolute-value terms in regression models of the electrical outputs of multi-piece force balance calibration data are discussed. First, characteristics of a variety of regression model term combinations with absolute-value terms are reviewed that are currently used in the aerospace testing community to fit the gage outputs of a balance. Then, a semi-empirical test is presented that quantifies bidirectional characteristics of the balance bridge outputs. Several diagnostic methods are discussed to assess the severity of near-linear dependencies between regressors of models with absolute-value terms. In particular, connections between the linear, absolute-value, quadratic, signed quadratic, and cubic terms are studied in greater detail. Data from an automated calibration of NASA's MK29B force balance are used to illustrate the most important observations and results. Rules of thumb that variance-inflation factors be less than 10 must be relaxed when using absolute-value terms to describe bidirectional balances.

Nomenclature

a	Regression intercept
AF	Axial force
$\mathbf{b}$	Regression coefficient vector
$b1, b2$	Regression coefficients for linear terms
$c1, c2$	Regression coefficients for quadratic terms
$c3, c4, c5, c6$	Regression coefficients for two-way interactions
$d1, d2$	Regression coefficients for cubic terms
F	Load component
m	Number of observations
n	Number of balance components
$N1$	Forward normal force
$N2$	Aft normal force
p	Number of regression terms
r	Correlation coefficient
$rN1$	Forward normal-force bridge output
R	Bridge output
$\mathcal{R}$	Correlation matrix
RM	Roll moment

*Assoc. Technical Fellow, AIAA Associate Fellow
†Aerodynamicist, AIAA Senior Member

S1	Forward side force
S2	Aft side force
$\mathbf{X}$	Regression model matrix
κ	Condition index, $\lambda_{\max}/\lambda$
λ	Eigenvalue
Subscripts	
i, j, k	Term or component number

I. Introduction

The calibration of wind-tunnel force balances presents an advanced application of regression analysis. A force balance has multiple degrees of freedom, typically six bridge outputs responding to six applied forces and moments. Each of the outputs, then, requires a multivariate regression analysis. AIAA Recommended Practice R-091[1] for the calibration and use of wind-tunnel balances notes that the most common regression model for a six-component balance is a 27-term quadratic, consisting of six linear terms, six squared terms, and 15 two-factor interactions. This model has been extended by including cubic terms for the primary component loads. The resulting 33 terms define the basic model for balances with continuous input-output relationships, typically of monolithic or single-piece design:

$$R_i = a_i + \sum_{j=1}^{n} b1_{i,j} F_j + \sum_{j=1}^{n} c1_{i,j} F_j^2 + \sum_{j=1}^{n} \sum_{k=j+1}^{n} c3_{i,j,k} F_j F_k + \sum_{j=1}^{n} d1_{i,j} F_j^3 \qquad (1)$$

In this equation, R_i is the output of bridge i, F is a load component, and $n = 6$ for a six-component balance.

The recommended practice also notes that "it is not uncommon for the load/output relationship of balances, especially those of multi-piece design, to exhibit some dependency on the sign of the strain in the measuring elements."[1] Such a balance is often called *bidirectional*. The situation is illustrated in Figure 1, where a least-squares line has been fit to data with a slope discontinuity at the origin, resulting in a residual pattern of alternating positive/negative signs. The recommended practice addresses the consequences of bidirectional behavior:

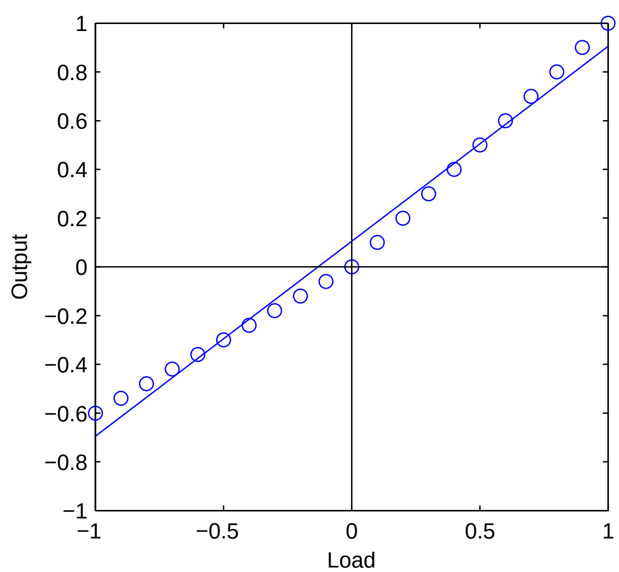

Figure 1. Sign-of-load (bidirectional) sensitivity.

> This asymmetry results in the need to determine and use different calibration coefficients according to the sign of the force or moment acting on the bridge, in order to achieve the best accuracy from the balance. Rather than defining separate calibration coefficients for positive load and negative load and then selecting from these to suit the particular combination of signs in a given instance, this asymmetric load behavior can be modelled [sic] effectively by an extension of the basic math model to include terms combining the component loads with their absolute values. [1, pp. 8–9].

Extending the basic model to include the "sign of load" effect for all coefficients produces the 97-term AIAA recommended model for communication of balance calibration results:

$$
\begin{aligned}
R_i \;=\;& a_i + \sum_{j=1}^{n} b1_{i,j} F_j + \sum_{j=1}^{n} b2_{i,j} |F_j| + \sum_{j=1}^{n} c1_{i,j} F_j^2 + \sum_{j=1}^{n} c2_{i,j} F_j |F_j| \\
&+ \sum_{j=1}^{n} \sum_{k=j+1}^{n} c3_{i,j,k} F_j F_k + \sum_{j=1}^{n} \sum_{k=j+1}^{n} c4_{i,j,k} |F_j F_k| + \sum_{j=1}^{n} \sum_{k=j+1}^{n} c5_{i,j,k} F_j |F_k| + \sum_{j=1}^{n} \sum_{k=j+1}^{n} c6_{i,j,k} |F_j| F_k \\
&+ \sum_{j=1}^{n} d1_{i,j} F_j^3 + \sum_{j=1}^{n} d2_{i,j} |F_j^3| \qquad (2)
\end{aligned}
$$

It is worth emphasizing that Eq. 2 is intended primarily for convenience in communication and computation, and that R-091 does *not* recommend actually evaluating 97 coefficients from any given set of calibration data, instead advising that "it is best to include in the math model only those terms for which there is some underlying physical reasoning" [1, p. 10]. The recommended practice further advises that cubic terms be omitted when including the absolute-value terms since in some respects these function classes serve a similar purpose. But it must be recognized that, even when omitting the cubic terms of the ninth and tenth summations, one is still faced with a multivariate analysis involving 85 regressors.

The properties of the AIAA model with respect to regression analysis were explored by Johnson, Parker, and Landman in Ref. 2. They assessed the conditioning of Eq. 2, in terms of pairwise correlation coefficients and term variance-inflation factors, against a notional calibration load schedule that swept the design space two factors at a time. This load schedule produced variance-inflation factors on the order of 1000 — two orders of magnitude larger than the commonly-suggested threshold value of 10 — indicative of severe collinearity and degradation of coefficient estimates. The source of collinearity was then examined using correlation coefficients, and values in excess of 0.96 were recorded for several pairings of regression terms involving F_j with $F_j|F_j|$, F_j^2 with $|F_j|$, F_j^2 with $|F_j^3|$, and F_j^3 with $F_j|F_j|$. They further showed that the correlation is inherent in the function classes themselves and independent of the load-schedule design.

Johnson et al. suggested four alternatives to reduce the collinearity and improve the model prediction performance, which can be divided into two basic options. The first option is to partition the domain and build separate models based on Eq. 1 without sign-of-load terms. This option, of course, forgoes the benefits of a single model which the absolute-value terms afford. The second option is to omit terms involved in the offending correlations. Addressing this strategy, Montgomery and Peck note "it may not provide a satisfactory solution if the regressors removed have significant explanatory power relative to the response." [3, ch. 15] It is the authors' experience that absolute-value terms do indeed have significant explanatory power for some multi-piece force balances, making their inclusion mandatory. Moreover, cubic terms have not been necessary. This is fortunate, because removing the cubic terms eliminates the most severe collinearities and allows variance-inflation factors for models based on Eq. 2 to be held to the range of 20 to 40. Even though variance-inflation factors of this magnitude exceed commonly recommended thresholds, the resulting regression models have been found to be useful nonetheless.

This paper will illustrate the use of absolute-value terms for balance-calibration regression modeling. First, the general use of absolute-value terms is discussed. Then, a semi-empirical test is presented that quantifies the bidirectional characteristics of balance bridge outputs. Several diagnostic methods are discussed to assess the severity of near-linear dependencies between regressors of models with absolute-value terms, which are then applied to a machine calibration example. Finally, recommendations are given for using absolute-value terms with bidirectional balance calibration data.

II. Regression Model Choices for Bidirectional Balances

The authors observed in the past that many balance users have chosen various subsets of the general model defined in Eq. 2 to perform an analysis of balance calibration data. Table 1 below, for example, lists frequently used term combinations.

Option 1 is recommended for use with single-piece balances as this combination does not use absolute-value terms. Options 2 to 7 are often suggested for use with a multi-piece balance that exhibits bidirectional behavior. They differ in the number of term combinations and types that are used. In addition, all combinations for a multi-piece balance expand the basic term combination that is recommended for a single-piece balance. Option 2, for example, uses only the absolute-value term itself in addition to the basic term com-

Table 1. **Common term group combinations used for the regression analysis of balance calibration data.**

Term group	Option 1	Option 2	Option 3	Option 4	Option 5	Option 6	Option 7
F_j	×	×	×	×	×	×	×
$\|F_j\|$		×	×	×		×	×
F_j^2	×	×		×	×	×	×
$F_j\|F_j\|$			×	×		×	×
F_jF_k	×	×	×	×	×	×	×
$\|F_jF_k\|$						×	×
$F_j\|F_k\|$						×	×
$\|F_j\|F_k$						×	×
F_j^3					×		×
$\|F_j^3\|$							×

bination. Option 3 and option 5 were suggested for use in Ref. 2. Option 3 uses two absolute-value term groups to model the bidirectional behavior. Option 5, on the other hand, uses a third-order term instead of the absolute value terms. Option 4 is another term combination that is frequently applied. Option 6 uses all but third order terms. Finally, Option 7 uses all possible 10 math term groups.

In general, it can be expected that the term combinations listed in Table 1 have both advantages and disadvantages. Therefore, it is necessary to introduce quantitative metrics that help better understand differences between the term combinations. These metrics are discussed in the next section. Afterwards, it will be shown how the metrics may be applied to machine calibration data of NASA's MK29B force balance.

III. Test for Bidirectionality

The role of the absolute-value terms in Eq. 2 is to adjust the base coefficients according to the signs of the component loads (cf. Appendix VI). The first step for the analyst, then, is to determine whether these terms are needed in the calibration model. Unfortunately, Ref. 1 provides only a qualitative description of bidirectional behavior, but no specific guidance on when to include absolute-value terms. Ulbrich[4] developed an empirical criteria to quantify the magnitude of bidirectionality. DeLoach and Ulbrich[5] also explored the use of categorical variables, an idea which was independently discussed in Ref.2. These studies lead to the definition of the test that is currently being used in NASA's *BALFIT* regression analysis tool to assess bidirectional characteristics of balance outputs. The process basically consists of assuming that the absolute-value terms are needed by including them in the model, then assessing whether they are significant in both the statistical and the practical sense.

The *BALFIT* test for bidirectional behavior is based on two metrics that are compared with related thresholds. The test is only valid if a multi-piece force balance is analyzed in force balance format. The first metric is defined as the product of the regression coefficient of the absolute-value term of a primary bridge load (the coefficient of the term $|F_j|$ in Eq. 2) with its design load capacity. The second metric is the p-value of the t-statistic of the coefficient of the term $|F_j|$[a]. The *BALFIT* test simply assumes that a bridge output is bidirectional if first, the absolute value of the first metric exceeds 0.5 percent of the to-capacity-scaled maximum of the difference between gage output and natural zero, and second, if the second metric is less than 0.001. As demonstrated in the appendix, a contribution of 0.5 percent equates to a 1 percent difference in slopes between the F_j half planes.

IV. Regression Diagnostics

The solution of the linear regression model is

$$\mathbf{b} = (\mathbf{X^T X})^{-1}\mathbf{X^T y} \tag{3}$$

[a]A detailed description of this statistical test is given in Ref.6, pp. 84–85.

where $\mathbf{y}$ is an $m \times 1$ vector of responses representing one of the bridge outputs R_i, $\mathbf{X}$ is the $m \times p$ model matrix[b] of the n balance load components expanded into the p regressors (terms) for the m observations (data points) of the calibration data set, $\mathbf{X^T}$ is its transpose, and $\mathbf{b}$ is the $p \times 1$ vector of least-squares estimated parameters. When applied to Eq. 2, $\mathbf{b}$ represents the coefficients $[a|b1 \ldots d2]^T$. Diagnostics are employed to assess the conditioning of the square matrix $(\mathbf{X^T X})$ to be inverted. An *ill-conditioned* case arises when at least one column of X is linearly dependent on the other columns. If the dependence is exact, $(\mathbf{X^T X})$ is singular and there is no unique solution. If instead there is a *near dependency* in the columns of X, a solution exists, but the coefficient estimates $\mathbf{b}$ are unreliable, with high sensitivity to small changes in the data, and possibly magnitudes that are unrealistically large. Two diagnostics will be used to assess the conditioning of regression models for balance calibration using Eq. 2, namely, variance-inflation factors and condition numbers based on eigensystem analysis. The descriptions that follow are based on references 2, 3, 7, 8.

A. Correlation Coefficients

The Pearson correlation coefficient measures the strength of the linear relationship between two vectors. For vectors x and y, the coefficient is defined as [3, ch. 13]

$$r_{xy} = \frac{S_{xy}}{\sqrt{S_{xx} S_{yy}}} \tag{4}$$

where $S_{xy} = \sum xy - (\sum x)(\sum y)/n$, $S_{xx} = \sum x^2 - (\sum x)^2/n$, and $S_{yy} = \sum y^2 - (\sum y)^2/n$. Due to the normalization, values of r_{xy} range from -1 to $+1$. A value of $+1$ indicates perfect positive correlation, whereas a value of -1 indicates perfect negative correlation and values near zero indicate little or no correlation. Noting the similarity to the scalar product, one can interpret a zero coefficient corresponding to orthogonal vectors. Computing correlation coefficients between each column of a model matrix results in the symmetric $p \times p$ correlation matrix $\mathcal{R}$ with diagonal elements equal to one and off-diagonal elements equal to r_{ij}:

$$\mathcal{R}_{ij} = \begin{cases} 1, & i = j \\ r_{ij}, & i \neq j \end{cases} \tag{5}$$

For the purposes of regression diagnostics, the correlation coefficient by itself is of limited use because it only gives indications for pairs of regressors; collinearity involving three or more terms may or may not be readily apparent.

B. Variance-Inflation Factors

The variance-inflation factors (VIF), one per regressor, are computed as the diagonal elements of the inverse of the Pearson correlation matrix:

$$\text{VIF} = \text{diag}(\mathcal{R}^{-1}) \tag{6}$$

The term "variance inflation" arises because the variance of the regression coefficients is proportional to the VIF. The VIFs for regressors involved in a significant correlation will each have correspondingly large values. The presence of two or three large VIFs thus immediately identifies the regressors involved. When four or more VIFs are present, it is not immediately obvious whether one is dealing with a single correlation or multiple collinearities between subsets of the involved regressors. The correlation structure can be revealed by removing one regressor at a time and observing the corresponding changes to the other VIFs. Alternatively, one of the regressors can be regressed against the others; study of the resulting coefficients will often reveal which regressors are participating in the correlation. Better still, an eigensystem analysis will reveal the correlated regressors directly, which is the subject of the next section.

Variance-inflation factors are commonly employed for assessing multicollinearity, perhaps because they are readily produced by a number of regression-analysis software programs. It is often suggested in the literature that VIF values in excess of 10 indicate significant collinearity and a concomitant degradation in the precision of the associated coefficients. It is the authors' experience, however, that VIFs greater than 10 cannot be avoided when using absolute-value terms in combination with linear and quadratic terms in the calibration of multi-piece balances. Regression models may still be useful even when the variance of some of the coefficients is inflated over what it would otherwise be for orthogonal regressors.

[b]Sometimes called the regression design matrix.

C. Eigensystem Analysis

The eigenvalues λ_j of $(\mathbf{X^T X})$ can be used to assess the degree, and the nature, of multicollinearity in the model matrix. If the columns of $\mathbf{X}$ are orthogonal, the eigenvalues will all be equal. For each exact linear dependence in $\mathbf{X}$ there will be one zero eigenvalue. Near-linear dependencies result in small eigenvalues. In practice, the p *condition indices* are defined as

$$\kappa_j = \frac{\lambda_{\max}}{\lambda_j} \tag{7}$$

The maximum condition index is called the *condition number*. The degree of multicollinearity can be judged by the magnitude of the condition indices. If the condition number is less than 100, multicollinearity is not a serious concern. Condition indices between 100 and 1000 indicate moderate to strong collinearity, and values of κ_j greater than 1000 indicate severe collinearity. Furthermore, the eigenvectors corresponding to small eigenvalues (large condition indices) make up the proportions of the parameter variances, which can be used to identify which regressors are involved in the multicollinearity.

An equivalent analysis can be obtained from the singular value decomposition of $\mathbf{X}$, noting that the squares of the singular values of $\mathbf{X}$ are equal to the eigenvalues of $(\mathbf{X^T X})$. The proper threshold values to use with condition indices formed from singular values are therefore the square roots of the values given above for indices based on eigenvalues, i.e. the values 10 and 31 delineate weak, moderate, and severe multicollinearity. Belsley, Kuh, and Welsch[7] advocate that the eigensystem analysis be applied to the original X matrix with the columns scaled to unit length. This allows for one set of threshold values independent of the physical units of the independent variables.

The reader is referred to Refs. 3 and 8 for detailed descriptions of the computation of the condition indices and the scaling and interpretation of the eigenvectors. Draper and Smith's[8] step-by-step implementation using the singular value decomposition was followed for this paper.

D. Regressor Scaling

Many statistics texts employ *centering* of the regressors and dependent variables by subtracting their sample means, and sometimes by normalizing to unit length. One variation is *coding* which maps each independent variable to the range $[-1, 1]$, which may be different from centering if the domain is not symmetric about the origin. One advantage of centering is that it removes the intercept from consideration. It also scales the problem to the domain of the independent variables and normalizes disparate physical units that might span several orders of magnitude. Although it may be argued that numerical accuracy is improved by the normalization, in practice numerical precision has not been an issue. One must be aware when centering is employed because several formulas for analysis of variance tables and confidence intervals are specific to centered data.

It is particularly important that centering or coding of units *not* by employed when using absolute-value terms as the meaning of these functions is dependent on a physical origin. The columns of X are, however, scaled both for computing the Pearson correlation coefficient (which carries over to the computation of the VIFs) and also, as noted above, for the eigensystem analysis. But in these cases the scaling is performed only *after* the independent variables (the load components) are expanded into the columns of the regression model matrix. The conditioning metrics presented in this paper are based on uncentered data. Accordingly, all VIFs correspond to Method 2 of Ref. 9.

E. Coefficient Evaluation

The diagnostics presented thus far assess the conditioning of the candidate regressors with respect to the available data. Once the regressors are chosen and the regression is performed, the resulting model should be assessed by performing the following steps:

1. The residuals should always be plotted. Ideally they should be random with constant variance across the domain of the independent variables and the range of predicted values. Plots can reveal outliers due to data input errors, alignment or other setup problems, or other blunders, and identify points that should be repeated.

2. The coefficients should be checked for statistical significance and unexpected signs. Significance is assessed using the p-value of the t-statistic or the equivalent confidence interval. Removing insignificant terms guards against over-fitting of sparse data sets. A coefficient with an unexpected sign warrants re-examination of the conditioning diagnostics.

V. Example Calibration of a Bidirectional Balance

Data from a 2007 machine calibration of NASA's MK29B balance is used to illustrate different aspects of the application of absolute-value terms in regression models of the bridge outputs of a bidirectional balance. The MK29B is a six-component, multi-piece balance manufactured by Task/Able Corporation. It is a force-type balance with forward and aft normal-force elements *N1* and *N2*, forward and aft side-force elements *S1* and *S2*, a rolling-moment element *RM*, and an axial-force element *AF*. Figure 2 shows the basic layout of the balance. The non-metric end is located on the right-hand side of the figure. Table 2 below summarizes the physical characteristics of the balance.

Figure 2. NASA's 2-inch diameter MK29B six-component multi-piece force balance. (Image courtesy of NASA Ames Research Center, Moffett Field, CA.)

Table 2. Characteristics of the 2-inch MK 29B balance.

Balance type	force
Diameter/length	2.0/11.25 in
Normal/side spacing	9/7 in
Capacities: *N1*, *N2*	2100 lbf
S1, *S2*	700 lbf
RM	3800 in-lbf
AF	350 lbf

The machine calibration of the balance was performed in 2007 in Triumph Aerospace's Automatic Balance Calibration System (ABCS). Figure 3 shows the ABCS with the MK29B balance installed. The ABCS is a non-leveling system that utilizes hydraulic actuators and load cells.[10] The load schedule consists of a number of loading sequences in which a primary load component is cycled over its full positive/negative range, while a secondary load component is held relatively constant. These patterns were repeated for several levels, both positive and negative, of the secondary load. Single-component loads were obtained by holding the secondary load near zero. The schedule totals 1670 load points.

The use of this machine calibration for the current study has three major benefits. First, the number and distribution of load points provides maximum support for all of the regressors in Eq. 2. As shown in Figure 4, the calibration data are well distributed within the four quadrants of all fifteen combined-load plots, which helps fully characterize the cross-product terms in the regression model. Second, the relatively large number of primary load increments in each series maximizes the definition of related regressors such as $|F_j|$ and F_j^2, which in turn allows for an improved assessment of the bidirectional characteristics of the balance bridges. Finally, the MK29B balance exhibits good linearity and small interactions, and residual distributions that are close to normal for all six load components.

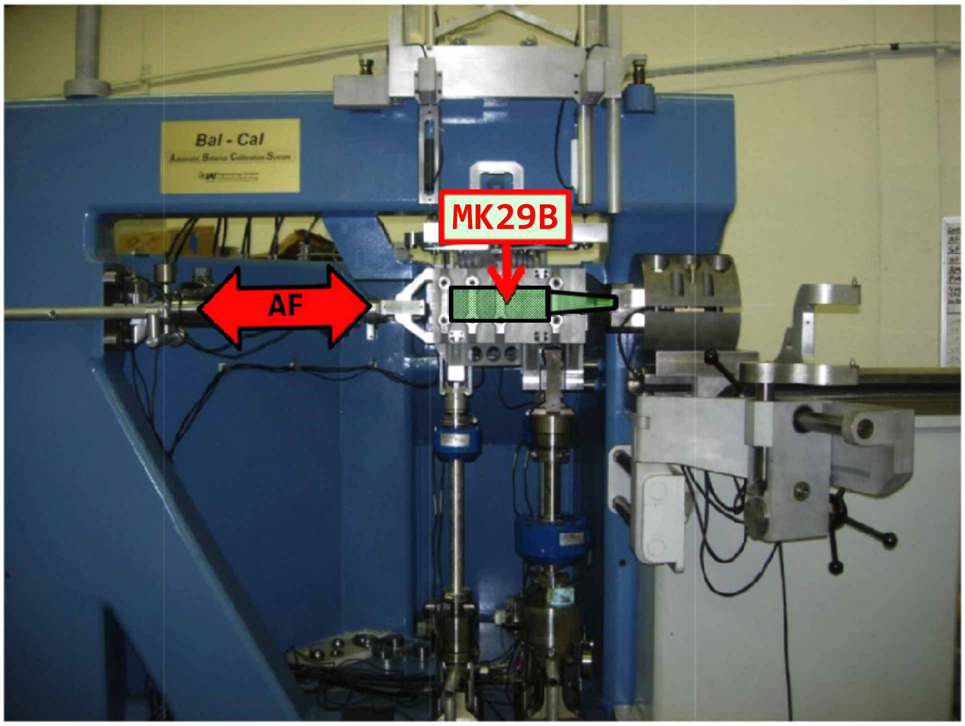

Figure 3. The Automatic Balance Calibration System with MK29B balance installed. (Image courtesy of Triumph Aerospace—Force Measurement Systems, San Diego, CA.)

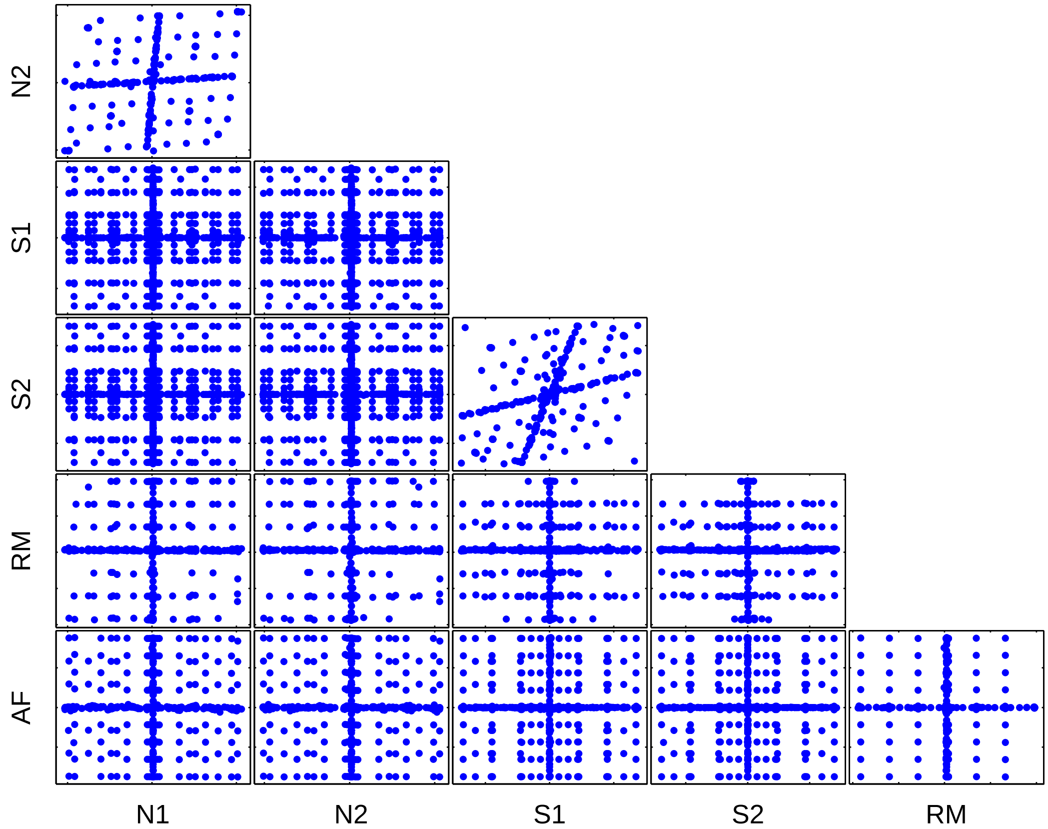

Figure 4. Calibration data set for the MK29B balance.

The analysis of the MK29B calibration data was primarily performed using NASA's *BALFIT* regression analysis software tool that was specifically developed for the processing of wind tunnel balance data.[11] Eigensystem analyses were performed separately with software developed by Boeing. The analysis is done in the design format of the balance, i.e., in force balance format because only in this case can the connection between the bidirectional characteristics of the bridge outputs be directly associated with a single load component.

A. Bidirectionality of the MK29B

First, the bidirectional characteristics of the MK29B balance were investigated using the method described above. Figure 5 shows the results produced by the *BALFIT* software. The output contribution due to each primary absolute-value coefficient is plotted versus the corresponding bridge load. The vertical axis limits are common to all six components to facilitate direct comparison among the components. The plot titles list the coefficient p-values; the coefficients for all six components are less than 0.001 and therefore statistically significant. Each plot also shows the 0.5 percent threshold for that bridge. The outputs of the normal and side force bridges exceed the threshold at approximately 1 and 2.5 percent, respectively. The bidirectional components of the rolling moment and axial force bridges equate to about 0.25 percent each, and are therefore not judged to be significant in a practical sense.

These results are typical for a Task balance, and reflect the physical design. The normal and side force elements are double eccentric columns fastened on either end by screwed connections. The rolling moment and axial elements, on the other hand, attach at the metric and non-metric ends using pins. These results indicate that absolute-value terms are definitely warranted for the calibration model of the MK29B balance.

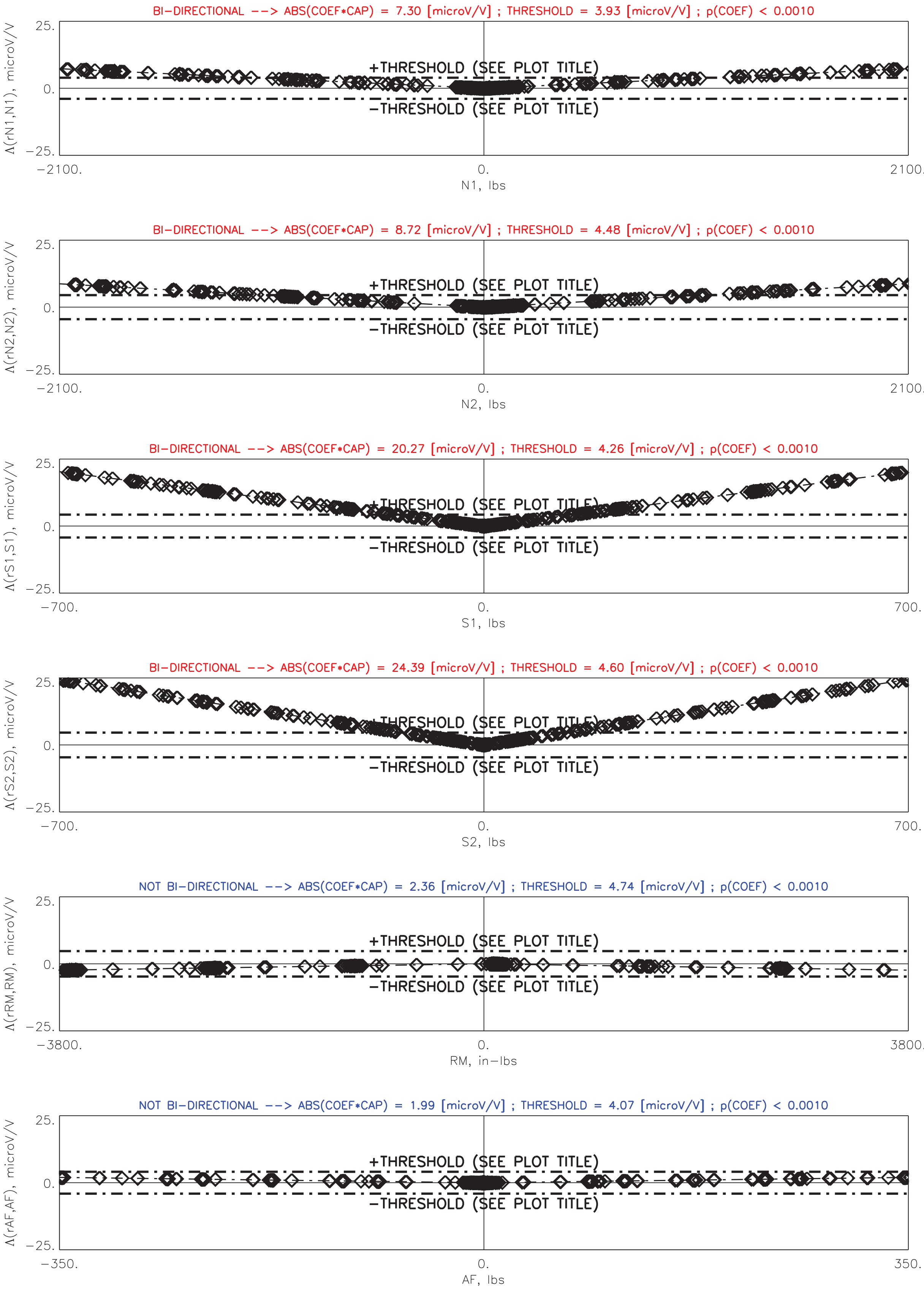

Figure 5. Predicted outputs from the primary absolute-value coefficients, with threshold of 0.5 percent of full scale.

B. Conditioning of the AIAA Model

Different aspects of regression models of multi-piece balances with absolute-value terms were investigated using the MK29B data set. Results will be presented for the forward normal-force bridge output *rN1* as representative of the other components. The first step is to examine the conditioning of the full 96-term model of Eq. 2 against the calibration data set. Figure 6 shows the variance-inflation factors obtained after expanding the applied loads into the 1670×96 model matrix. The maximum VIF is 1180, an alarmingly high

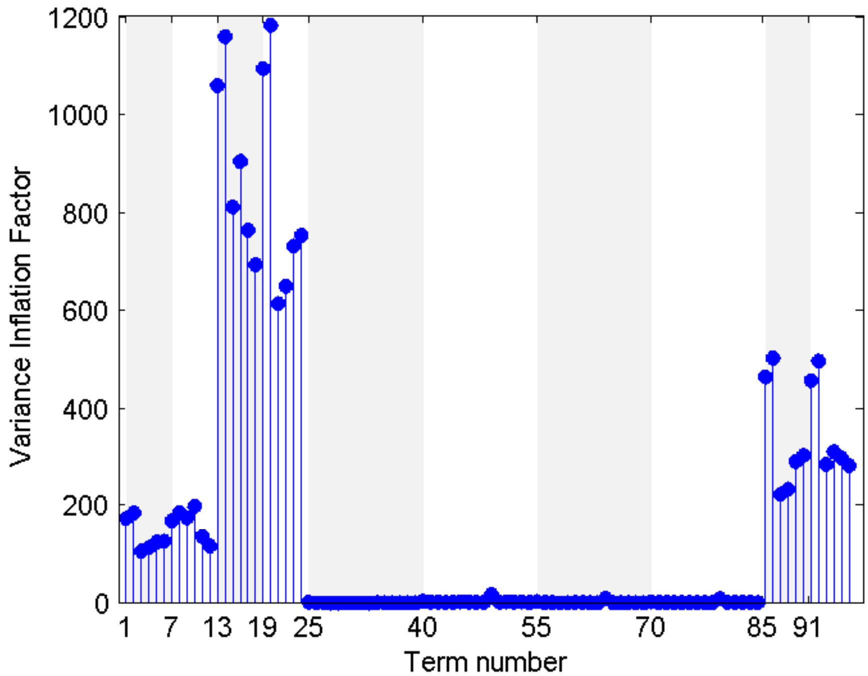

Figure 6. Variance-inflation factors for the 96-term model.

value. The term values labeled on the abscissa correspond to the first terms of each of the ten summation signs in Eq. 2. The VIFs of terms F_j, $|F_j|$, F_j^2, $F_j|F_j|$, F^3, and $|F_j|^3$ all exceed 100. Only the four groups of two-way cross products have VIFs less than 20. The eigensystem analysis confirms the poor conditioning. The condition number is 162, with 12 indices well into the range of severe multicollinearity. The specific collinearities are revealed by examining the variance proportions corresponding to individual condition indices. Figure 7 shows the proportions for the 93rd and 95th condition indices. The 95th condition index $\kappa_{95} = 144.7$ and corresponds to a correlation dominated by $|N1|$, $N1^2$, and $|N1|^3$. Condition index $\kappa_{93} = 131.5$ and corresponds to a correlation dominated by $N1$, $N1 \cdot |N1|$, and $N1^3$.

Figure 8 shows the variance-inflation factors after dropping the cubic terms of the ninth and tenth summations. Removing the cubic terms has a dramatic effect, reducing the maximum VIF from almost 1200 to slightly less than 30. The condition number has also decreased to 29.3, at the upper end of the moderate to strong collinearity range. Clearly, the cubic terms are the primary source of ill-conditioning in Eq. 2. Some moderate collinearity is still indicated by VIFs in excess of 15 for the first 24 terms, which corresponds to F_j, $|F_j|$, F_j^2, and $F_j|F_j|$. The two-factor interactions, terms 25–84, are well conditioned with VIFs mostly less than 5. Based on these results, and assuming suitable balance performance, the 84-term model should be the upper-bound candidate model for bidirectional balances.

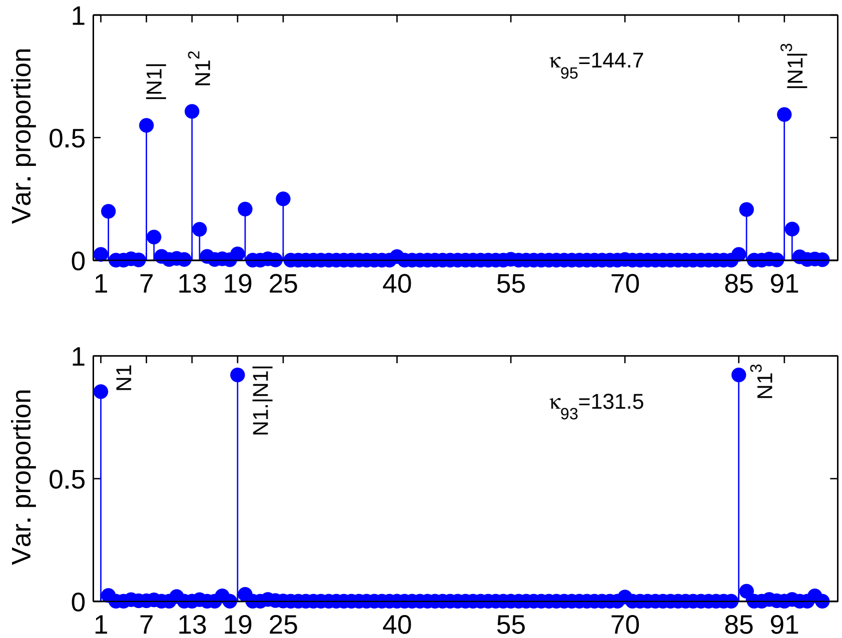

Figure 7. Coefficient variance proportions for the 93rd and 95th condition indices of the 96-term model.

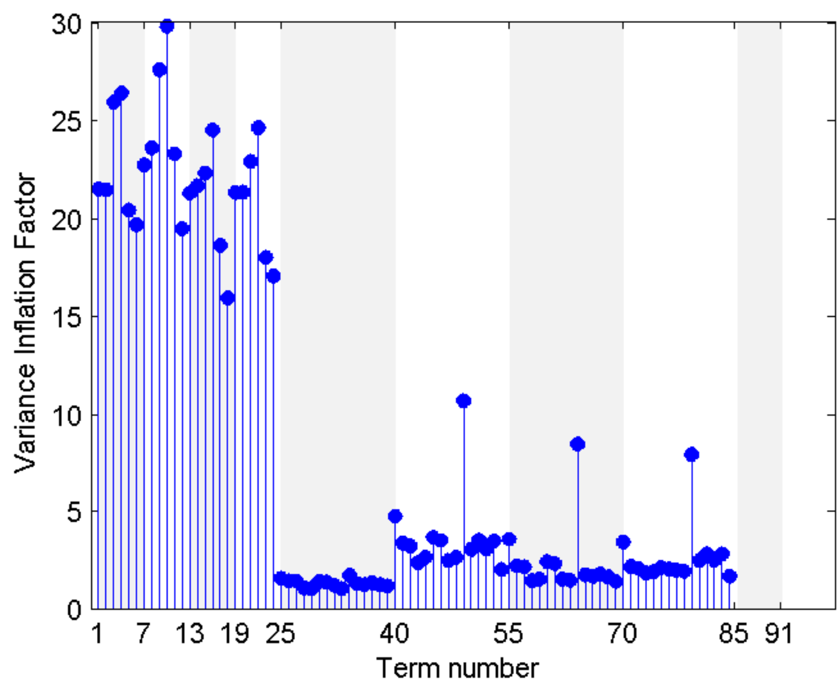

Figure 8. variance-inflation factors for the 84-term model based on term groups 1–8.

C. Comparison of Regression Model Term Choices

Several models with different term choices have been investigated. Table 3 compares the results. The columns are labeled with the basic strategy of each model. The term groups included are noted in the upper half of the table. The lower half summarizes the total number of coefficients, the number of statistically-insignificant coefficients, the conditioning of the regression in terms of VIF and condition number, and finally the standard deviation of the *N1* residuals. The models are sorted roughly in order of increasing collinearity,

Table 3. Summary of regression models of *N1* bridge output.

Term group	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
F_j	×	×	×	×	×	×	×
$\|F_j\|$			×	×	×	×	×
F_j^2	×	×	×	×	×	×	×
$F_j\|F_j\|$				×	×	×	×
F_jF_k	×	×	×	×	×	×	×
$\|F_jF_k\|$				×	×	×	×
$F_j\|F_k\|$				×	×	×	×
$\|F_j\|F_k$				×	×	×	×
F_j^3		×					×
$\|F_j^3\|$							×
Term reduction	none	none	none	BALFIT	backward	none	none
No. coeffs retained	27	33	33	29	44	84	96
No. insignificant	12	15	13	0	2	37	51
Max VIF	1.6	7.2	27	21	23	30	1180
Condition no.	3.9	7.0	23	16	21	29	162
Collinearity	weak	weak	moderate	moderate	moderate	strong	severe
Residual std dev	0.18%	0.17%	0.13%	0.076%	0.068%	0.067%	0.065%

which generally corresponds to decreasing residual standard deviation.

The first two models are based on Eq. 1. The 27-term model was fit using the first, third, and fifth term groups of Eq.2. The maximum variance-inflation factor is less than two, and the condition number is 3.9, indicating that the model is well supported by the data. The residual standard deviation of 0.18 percent is considered moderately successful. Residuals for this model are shown in Figure 9. In this figure, the open circles represent the entire data set of 1670 points. The standard deviation is clearly driven by the concentration of load points near the origin, and significant structure is evident. The 24 points connected by the red line represent a single-component *N1* series during which the other five load components were all less than 10 percent of their respective capacities. This series highlights an alternating pattern of increasing-decreasing-increasing residuals that is not explained by hysteresis. Such a pattern is symptomatic of bidirectional bridge behavior.

The second model adds the six cubic terms of the last summation in Eq. 1. The residuals are shown in Figure 10. Careful comparison with Figure 9 reveals that the addition of the cubic terms has flattened out the residual structure somewhat and reduced the maximum residual by about 0.1 percent. However, significant systematic variation remains, particularly in the *N1* series. Pure cubic terms are clearly unable to describe the bidirectional behavior.

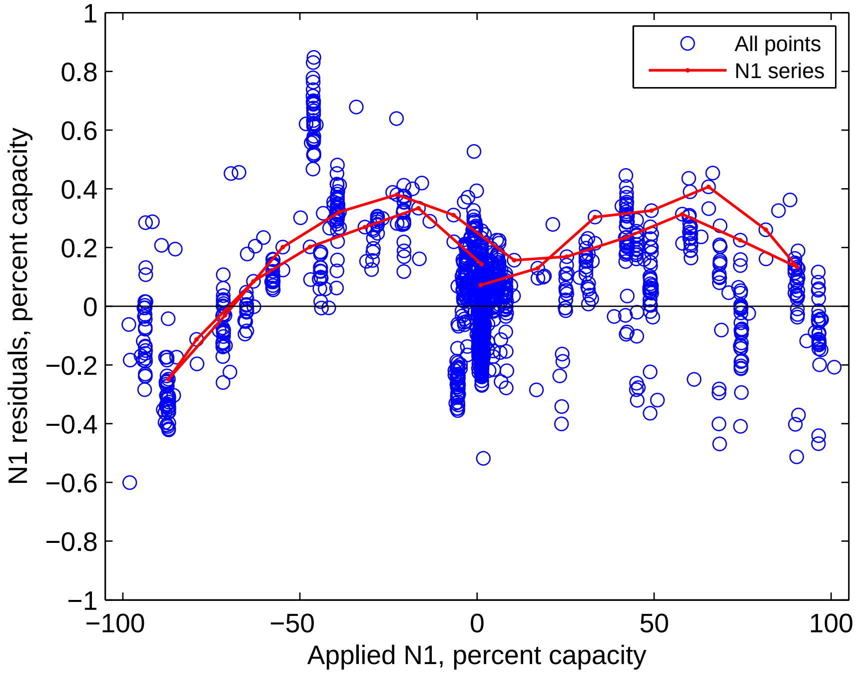

Figure 9. ***N1*** **residuals for the 27-term Model 1.**

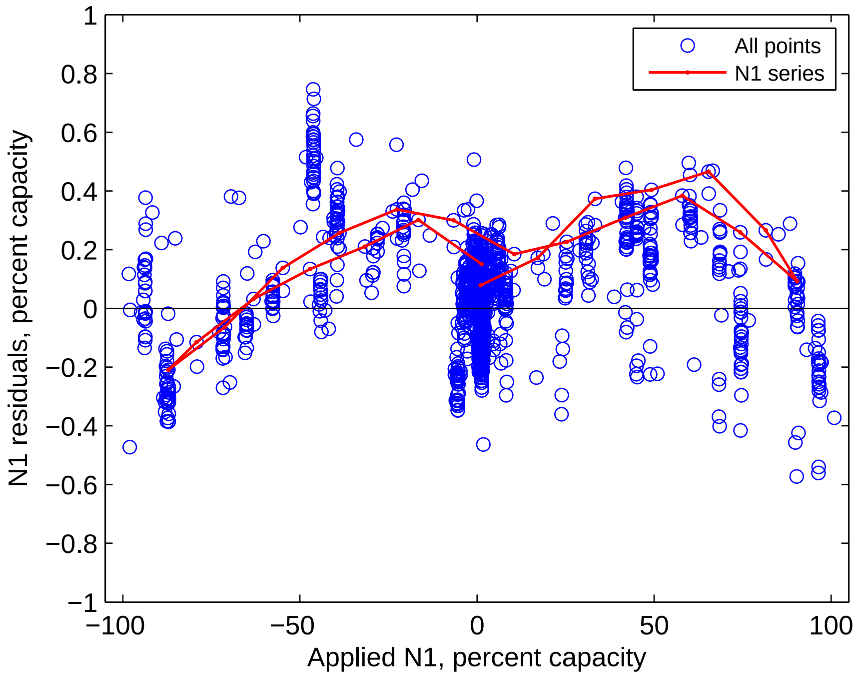

Figure 10. ***N1*** **residuals for the 33-term Model 2.**

The other models in Table 3 all include at least one term group involving absolute values. Model 3 adds $|F_j|$ terms to the first model, and results in the residuals presented in Figure 11. The $|F_j|$ term group has had a significant effect. The residuals for negative applied *N1* have been reduced, and the alternating pattern for the *N1* series has been reduced in amplitude. Still, significant structure remains.

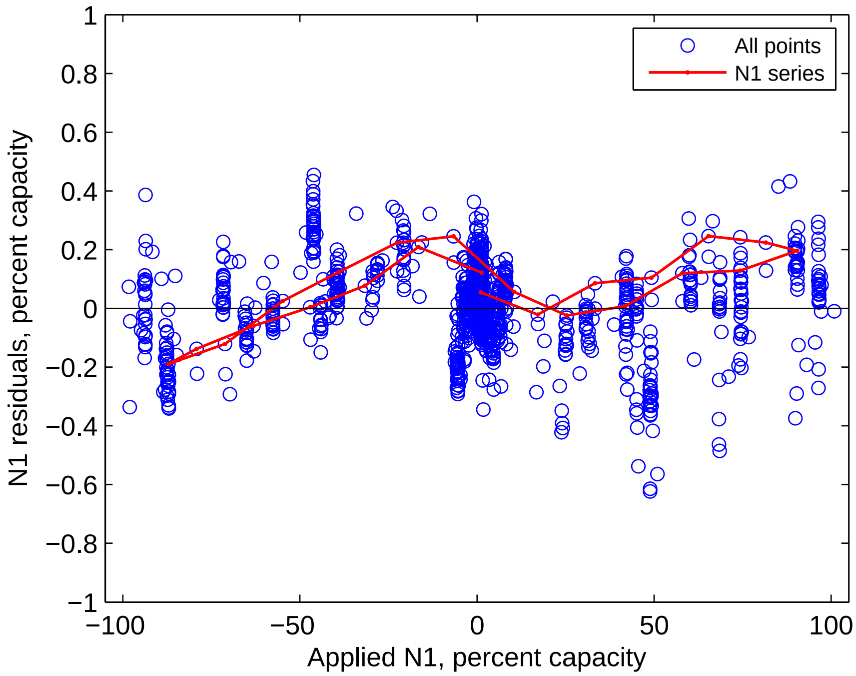

Figure 11. ***N1*** **residuals for the 33-term Model 3.**

The last two models in Table 3 are the full 96-term model and the 84-term baseline model detailed in the previous section. The 96-term model has the lowest residual standard deviation at less than 0.07 percent, however 51 of the 96 coefficients are statistically insignificant, and the model is plagued by severe multicollinearity. The 84-term baseline model, Model 6, shows a marked improvement in conditioning with only a slight increase in the residual standard deviation. Thirty-seven of the 84 terms are statistically insignificant, suggesting that term reduction is appropriate.

Residuals resulting from the 84-term baseline Model 6 are shown in Figure 12. Compared to Figure 11, this model has flattened the residual structure and reduced the maximum residual to about 0.25 percent. The distribution is also consistent with the 0.07 percent standard deviation. The *N1* load series, while biased slightly positive, has a flat, featureless distribution. The improvements in the *N1* load series are attributed to effect of the $N1|N1|$ (0.95 percent contribution). The quadrant-dependent two-way interactions from the sixth, seventh, and eighth summations also contribute to the reductions on the whole.

The fourth and fifth columns of Table 3 summarize two methods of term reduction, Model 4 being the simplified regression model search algorithm built into the *BALFIT* software[12] and Model 5 being a backwards-reduction strategy.[6,8] The backwards-reduction strategy retains 44 of the original 84 terms, two of which are insignificant. The *BALFIT* algorithm reduces the number of terms further to 29, all of which are statistically significant. The conditioning of both term-reduced models has fallen into the moderate range, with VIF and condition-number values both in the low to mid 20s. Even with fewer terms to capture the behavior, the residual standard deviations remain less than 0.1 percent, which is excellent performance. Residuals resulting from the BALFIT optimization and the backwards reduction from 84 terms are shown in Figures 13 and 14. These figures are almost indistinguishable from Figure 12, indicating that the most important terms have indeed been retained.

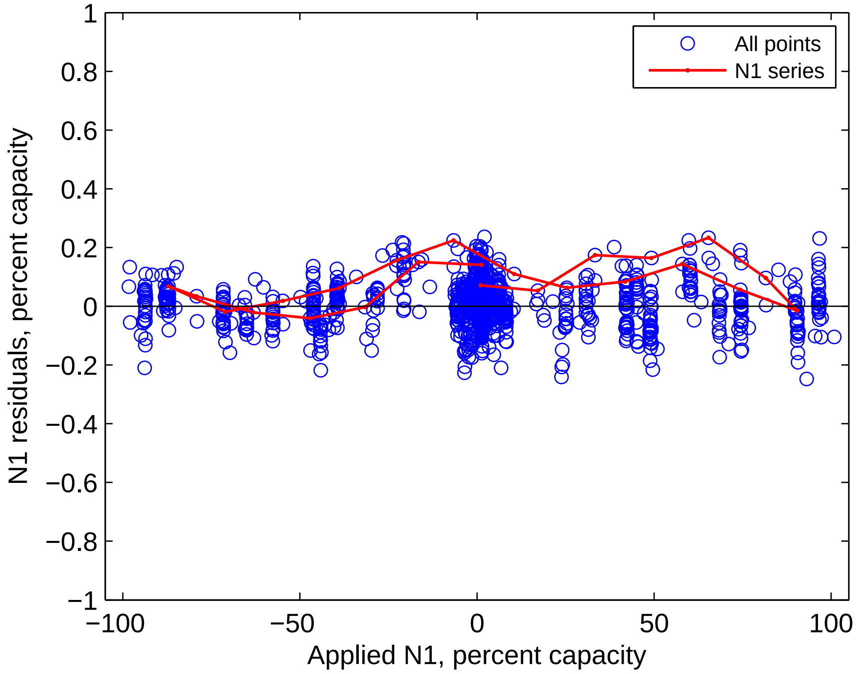

Figure 12. ***N1*** **residuals for the baseline 84-term Model 6.**

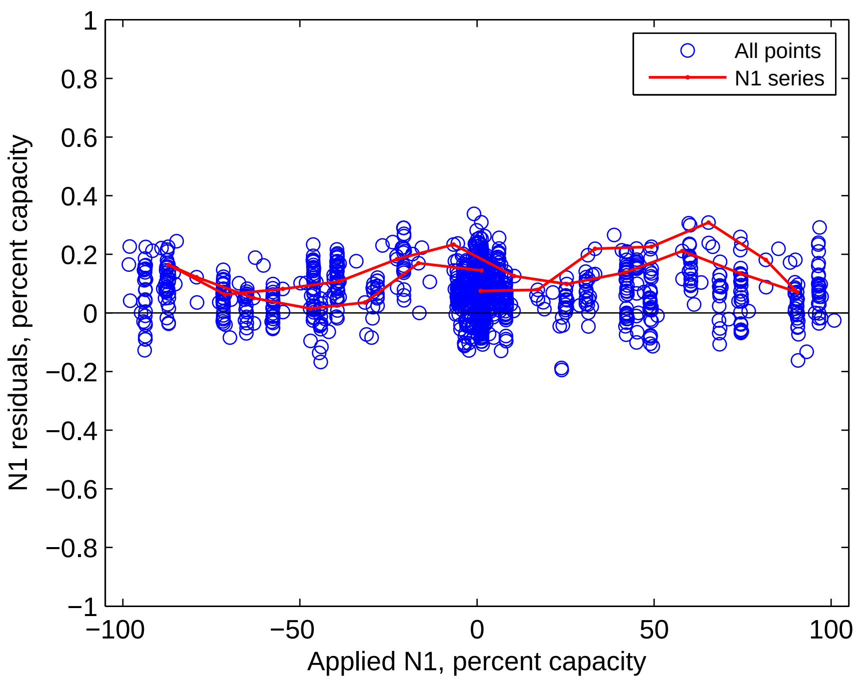

Figure 13. ***N1*** **residuals for the 29-term Model 4 resulting from BALFIT optimization from the 84-term base model.**

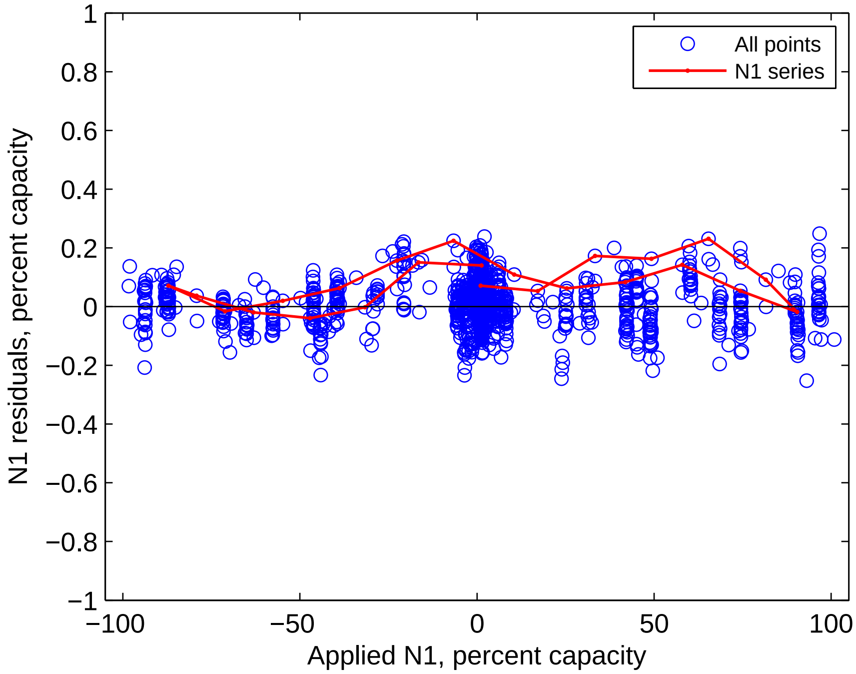

Figure 14. ***N1*** **residuals for the 44-term Model 5 resulting from backwards reduction from the 84-term base model.**

VI. Summary and Recommendations

This paper has shown that ill-conditioning in the AIAA recommended model is primarily due to the cubic terms in combination with the absolute-value terms. Specifically, F_j^3 is inherently correlated with F_j and $F_j \cdot |F_j|$, and $|F_j|^3$ is correlated with $|F_j|$ and F_j^2. These correlations will reveal themselves in the form of variance-inflation factors on the order of 100 or 1000, and condition numbers based on singular values in excess of 31. The proper course of action is to omit one of the offending terms. For bidirectional force balances, the baseline regression model is therefore that of the AIAA model without cubic terms.

Several specific recommendations can be made for the analysis of balance calibration data. For absolute-value terms to be individually meaningful, the component loads must correspond to those components to which the bridge outputs directly respond. In other words, the authors recommend to always analyze multi-piece force balance data in force format.

- Assess the load schedule for support of term groups. Variance-inflation factors can be used for this purpose. Cubic terms should not be included at this stage.

- Assess efficacy of absolute-value terms (bidirectionality) by entertaining $|F_j|$. Retain absolute-value terms if they are statistically significant *and* their contribution to the predicted bridge response is judged to be of practical significance, typically greater than 0.5 percent.

- Compute variance-inflation factors to assess the conditioning of the chosen model. Expect VIFs in the range of 10 to 30 as a natural result of the absolute-value function classes. Values of VIF in excess of 25 or 30 indicate poor conditioning/severe near dependencies and should be addressed by either improving the load schedule with additional load increments and/or combinations, or if necessary, dropping regressors from the analysis. In the latter case, use of the resulting balance matrix should be restricted to avoid extrapolation into regions of the load space not exercised in the calibration. Condition numbers can be computed in place of VIFs, in which case the cutoff value for action is 31 for condition numbers based on singular values (100 if based on eigenvalues).

- Examine residuals of the back-calculated calibration data and, if available, independent confirmation loadings. Bias in the residuals, in the form of linear or higher-order trends, indicates that the chosen model is inadequate. If replicate loadings are available, examine pure-error and lack-of-fit statistics.

- If the lack of fit is significantly high, consider the efficacy of cubic terms, realizing that their inclusion will increase the level of multicollinearity.

Appendix — Discussion of the AIAA Recommended Model

The introduction of absolute-value terms for modeling bidirectional balances can be traced to Galway.[13] The role of the absolute-value terms can be clarified by examining their behavior for the different signs of the load components F_j. When F_j is positive, $|F_j| = F_j$ and $F_j|F_j| = F_j^2$ allowing the first four summation signs in Eq. 2 to be grouped as follows, omitting subscripts and summation signs for clarity:

$$
\begin{aligned}
R &= (b1 + b2)F_j + (c1 + c2)F_j^2 && (8)\\
&= b^+F_j + c^+F_j^2 && (9)
\end{aligned}
$$

Likewise, when F_j is negative, $|F_j| = -F_j$ and $F_j|F_j| = -F_j^2$, which produces

$$
\begin{aligned}
R &= (b1 - b2)F_j + (c1 - c2)F_j^2 && (10)\\
&= b^-F_j + c^-F_j^2 && (11)
\end{aligned}
$$

This shows that the slopes in the positive and negative half planes are equal to

$$
\begin{aligned}
b^+ &= b1 + b2 && (12)\\
b^- &= b1 - b2 && (13)
\end{aligned}
$$

from which it follows that

$$
\begin{aligned}
b1 &= \frac{b^+ + b^-}{2} && (14)\\
b2 &= \frac{b^+ - b^-}{2} && (15)
\end{aligned}
$$

The prime sensitivity *b1* is therefore the average of the slopes, and the primary bidirectional coefficient *b2* is one half of their difference. Likewise the primary quadratic coefficient of F_j^2 is the average of the curvatures in the two half planes, and the bidirectional curvature coefficient of $F_j|F_j|$ is one half of their difference. Similar reasoning can be applied to the cross product terms F_jF_k:

$$
\begin{aligned}
&F_j > 0, F_k > 0; \quad |F_jF_k| = F_j|F_k| = |F_j|F_k = F_jF_k && (16)\\
&F_j < 0, F_k > 0; \quad |F_jF_k| = -F_jF_k,\ F_j|F_k| = F_jF_k,\ |F_j|F_k = -F_jF_k && (17)\\
&F_j < 0, F_k < 0; \quad |F_jF_k| = F_jF_k,\ F_j|F_k| = -F_jF_k,\ |F_j|F_k = -F_jF_k && (18)\\
&F_j > 0, F_k < 0; \quad |F_jF_k| = -F_jF_k,\ F_j|F_k| = -F_jF_k,\ |F_j|F_k = F_jF_k && (19)
\end{aligned}
$$

This allows the fifth through eighth summation terms of Eq. 2 to be collected as follows:

$$
\begin{aligned}
&F_j > 0, F_k > 0; \quad (c3 + c4 + c5 + c6)F_jF_k = c^{++}F_jF_k && (20)\\
&F_j < 0, F_k > 0; \quad (c3 - c4 + c5 - c6)F_jF_k = c^{-+}F_jF_k && (21)\\
&F_j < 0, F_k < 0; \quad (c3 + c4 - c5 - c6)F_jF_k = c^{--}F_jF_k && (22)\\
&F_j > 0, F_k < 0; \quad (c3 - c4 - c5 + c6)F_jF_k = c^{+-}F_jF_k && (23)
\end{aligned}
$$

Term groups five, six, seven, and eight, therefore, are together able to describe the response behavior in all four of the (F_j, F_k) quadrants. By considering specific signs for the component loads F_j and F_k, Eq. 2 can be reduced to Eq. 1.

Finally, Eqs. 20–23 can be solved for the absolute-value coefficients in terms of the quadrant-specific coefficients:

$$c3 = \frac{c^{++} + c^{-+} + c^{--} + c^{+-}}{4} \tag{24}$$

$$c4 = \frac{c^{++} - c^{-+} + c^{--} - c^{+-}}{4} \tag{25}$$

$$c5 = \frac{c^{++} + c^{-+} - c^{--} - c^{+-}}{4} \tag{26}$$

$$c6 = \frac{c^{++} - c^{-+} - c^{--} + c^{+-}}{4} \tag{27}$$

Acknowledgments

The authors wish to thank Mat Rueger and John Fussell of The Boeing Company and Jon Bader of NASA Ames Research Center for their critical review of the final manuscript. The work reported herein was partially supported by NASA Ames Research Center under Jacobs Technology support service contract NNA09DB39C. The calibration data of NASA's MK29B balance was provided courtesy of the Balance Calibration Laboratory at NASA Ames Research Center.

References

[1] AIAA, *Calibration and Use of Internal Strain-Gage Balances with Application to Wind Tunnel Testing*, Recommended Practice R-091-2003, American Institute of Aeronautics and Astronautics, Reston, VA, 2003.

[2] Johnson, T. H., Parker, P. A., and Landman, D., "Calibration Modeling of Nonmonolithic Wind-Tunnel Force Balances," *Journal of Aircraft*, Vol. 47, No. 6, November–December 2010, pp. 1860–1866.

[3] Wadsworth, Jr., H. M., editor, *Handbook of Statistical Methods for Engineers and Scientists*, McGraw-Hill, New York, 1990.

[4] Ulbrich, N., "Detection of Bi-Directionality in Strain-Gage Balance Calibration Data," AIAA Paper 2012-3320, 28th Aerodynamic Measurement Technology and Ground Testing Conference, New Orleans, LA, June 2012.

[5] DeLoach, R. and Ulbrich, N., "A Statistical Theory of Bidirectionality," AIAA Paper 2013-2995, AIAA Ground Testing Conference, San Diego, CA, June 2013.

[6] Montgomery, D. C., Peck, E. A., and Vining, G. G., *Introduction to Linear Regression Analysis*, John Wiley & Sons, New York, 4th ed., 2006.

[7] Belsley, D. A., Kuh, E., and Welsch, R. E., *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*, John Wiley & Sons, New York, 1980.

[8] Draper, N. R. and Smith, H., *Applied Regression Analysis*, John Wiley & Sons, New York, 3rd ed., 1998.

[9] Ulbrich, N., "Regression Model Optimization for the Analysis of Experimental Data," AIAA Paper 2009-1344, 47th AIAA Aerospace Sciences Meeting, Orlando, FL, Jan. 2009.

[10] Booth, D. and King, D., "Automatic Balance Calibration System (ABCS) Upgrades," AIAA Paper 2006-0516, 44th AIAA Aerospace Sciences Meeting, Reno, NV, Jan. 2006.

[11] Software Release Authority, Technology Partnerships Office, NASA Ames Research Center, Moffett Field, CA, *BALFIT — Software Tool for the Regression Analysis of Multivariate Data*, version 3.1.21 (2014/12/02), 2014.

[12] Ulbrich, N., "Analysis of Multivariate Experimental Data Using a Simplified Regression Model Search Algorithm," AIAA Paper 2013-2996, AIAA Ground Testing Conference, San Diego, CA, June 2013.

[13] Galway, R. D., "A Comparison of Methods for Calibration and Use of Multi-Component Strain Gauge Wind Tunnel Balances," Aeronautical Report LR-600, National Research Council, National Aeronautical Establishment, Ottaway, Canada, March 1980.