25 - 28 June 2012, New Orleans, Louisiana

Evaluation of Regression Models of Balance Calibration Data using an Empirical Criterion

(Extended Abstract of Proposed Conference Paper)

N. Ulbrich* and T. Volden**

Jacobs Technology Inc., Moffett Field, California 94035–1000

An empirical criterion is discussed that may be used to assess the significance of a regression model term of fitted strain–gage outputs of a wind tunnel balance. The criterion is based on the percent contribution of a regression model term. It considers a term insignificant if the percent contribution is below the threshold of 0.05 %. The criterion has the advantage that it can easily be computed using the regression coefficients of the strain–gage outputs and the load capacities of the balance. First, a detailed definition of the empirical criterion is provided. Then, the empirical criterion is compared with a more rigorous criterion that is traditionally used in linear regression analysis for the assessment of the statistical significance of a regression model term. Finally, calibration data from a variety of balances is used to illustrate the connection of the empirical criterion to real world data sets. A preliminary review of these results indicates that the percent contribution threshold of 0.05 % may still be too large for the assessment of the significance of terms of some balance calibration data sets. Therefore, it is recommended to apply the rigorous criterion whenever regression model term reduction for the prevention of over–fitting needs to be performed.

Nomenclature

AF	= axial force component of force balance
$b1, b2$	= coefficient of the regression model of a gage output, defined in Ref. [1], Eq. (3.1.3)
$c1, c2, \cdots, c6$	= coefficient of the regression model of a gage output, defined in Ref. [1], Eq. (3.1.3)
$d1, d2$	= coefficient of the regression model of a gage output, defined in Ref. [1], Eq. (3.1.3)
i	= index of a gage output –or– index of a primary gage load
j	= index of a load component
k	= index of a load component
K	= capacity of a primary gage load of a balance
n	= number of load components of a balance
$N1$	= forward normal force component of force balance
$N2$	= aft normal force component of force balance
P	= percent contribution of a coefficient of the regression model of a gage output
$R1, \cdots, R6$	= strain–gage outputs of force balance
RM	= rolling moment component of force balance
$S1$	= forward side force component of force balance
$S2$	= aft side force component of force balance

Summary

Different approaches are used in the wind tunnel testing community to perform a regression analysis of wind tunnel strain–gage balance calibration data. Many analysts prefer the *Iterative Method* as this approach

* Aerodynamicist, Jacobs Technology Inc.
** Computer Engineer, Jacobs Technology Inc.

fits the strain–gage outputs as a function of the calibration loads (see Ref. [1] for a detailed description of the method).

The regression analysis of the strain–gage outputs is a multivariate problem. Figure 1 shows, for example, a typical term selection that an analyst may make for the analysis of a balance calibration data set of a force balance. The term selection takes into account that combined loadings were only applied to the normal and side force components.

Different function classes like linear terms, absolute value terms, square terms, and cross–product terms may be used to assemble a suitable regression model. Ultimately, a set of regression model terms has to be selected such that (i) the behavior of the balance can be characterized correctly, (ii) near–linear dependencies between terms are suppressed, and (iii) the balance calibration data is not overfitted. Therefore, analysts use a host of different empirical and rigorous criteria to select individual terms of the regression model of the gage outputs.

Sometimes, an analyst may use an empirical criterion for the term selection (or reduction) that has originated in the wind tunnel testing community. It is based on the percent contribution of a regression coefficient. This empirical criterion can be summarized as follows:

Empirical Criterion for Regression Model Term Significance

> A term of the regression model of a strain–gage output is considered "significant" if the absolute value of the percent contribution of the term is greater than 0.05%.

The percent contribution is a metric that can easily be computed for a given balance calibration data set using corresponding coefficients of the regression model of the data (see App. 1 for more detail). It is only a function of (i) the load capacities of the balance and (ii) the regression coefficient values of the least squares fit of the balance calibration data. Figure 2 shows, for example, the percent contribution if the terms selected in Fig. 1 are used for the analysis of the calibration data of the Ames MK40 balance.

The origin and justification of the empirical criterion defined above was not obvious to the authors. Therefore, they decided to study the validity of the empirical criterion by using a more rigorous criterion from linear regression analysis that is used in statistics to assess the significance of regression model terms. The application of the rigourous criterion is, of cause, more complex than the application of the empirical criterion. Therefore, only basic ideas behind the rigorous criterion are presented.

In principle, the rigorous criterion assesses the significance of terms of the regression model by looking at the standard error of each regression coefficient. The standard error is an estimate of the standard deviation of the coefficient. It can be thought of as a measure of the precision with which the regression coefficient is measured. A coefficient should be included in the math model if it is large compared to its standard error.

The t–statistic of a coefficient is used to quantitatively compare a regression coefficient with its standard error. It equals the ratio between the coefficient value and its standard error. A coefficient is probably "significant" if its t–statistic is greater than the critical value of a Student's t–distribution (see Ref. [2], p.32). This comparison can also be performed by using the p–value of the coefficient. The p–value of a coefficient is determined from a comparison of the t–statistic with values in a Student's t–distribution. With a p–value of, e.g., 0.0001 (or 0.01 %) one can say with a 99.99 % probability of being correct that the regression coefficient is having some effect. Finally, the rigourous criterion for testing the significance of a regression model term can be summarized as follows:

Rigorous Criterion for Regression Model Term Significance

> A term of the regression model of a strain–gage output is considered "significant" if the p–value of the t–statistic of the regression coefficient is less than 0.0001.

A connection between the rigorous and empirical criterion needs to be established so that the rigorous criterion may be used to evaluate the empirical criterion. A connection can be defined if, for example, the smallest percent contribution is found for a given regression model of strain–gage outputs that still satisfies the rigorous criterion. Consequently, terms below this percent contribution will no longer satisfy the rigorous criterion, i.e., p–value $<$ 0.0001, and, consequently, are considered statistically insignificant.

Balance calibration data sets of different balance types were used to study the connection between the rigorous and empirical criterion. Table 1 below shows some results that were obtained so far.

Table 1: Strain–gage balance data analysis examples.

BALANCE NAME	BALANCE TYPE	NUMBER OF DATA POINTS	REGRESSION MODEL TERM COMBINATION	SMALLEST PERCENT CONTRIBUTION THAT MET RIGOROUS CRITERION: *p–value of t–statistic* < 0.0001
NTF–A	SINGLE–PIECE	410	$F,\ F^2,\ F \cdot G$	0.04 %
MK–40	MULTI–PIECE	164	$F,\ \lvert F \rvert,\ F^2,\ F \cdot G$	0.07 %
MC–60E	HI–CAP	1906	$F,\ \lvert F \rvert,\ F^2,\ F \cdot G$	0.03 %
MC–110	SEMISPAN	1133	$F,\ F^2,\ F \cdot G$	0.02 %
MC–400	SEMISPAN	498	$F,\ F^2,\ F \cdot G$	0.08 %

Overall, it appears that the empirical criterion is close to the arithmetic mean of the smallest percent contributions that still satisfied the rigorous criterion for the investigated data sets. However, it also appears that smaller values of the percent contribution threshold should be chosen for some balance calibration data sets. Therefore, the authors recommend that the rigorous criterion instead of the empirical criterion should be applied whenever regression model term reduction for the prevention of over–fitting of balance calibration data needs to be performed.

A wider variety of balance calibration data sets will be discussed in the final manuscript of the proposed conference paper. In addition, the influence of the regression model term type selection on the percent contribution threshold will be investigated in more detail.

References

[1]AIAA/GTTC Internal Balance Technology Working Group, "Recommended Practice, Calibration and Use of Internal Strain–Gage Balances with Application to Wind Tunnel Testing," AIAA R–091–2003, American Institute of Aeronautics and Astronautics, Reston, Virginia, 2003.

[2]Myers, R. H. and Montgomery, D. C., *Response Surface Methodology: Process and Product Optimization Using Designed Experiments*, 1st ed., John Wiley & Sons, Inc., New York, 1995.

Appendix – Definition of the Percent Contribution

The percent contribution describes the contribution of each term of the regression model of the gage outputs to the total fitted value, expressed as a percentage of the contribution of the principle diagonal term. It is used to assess the degree of linearity of the regression model of the gage outputs.

The percent contribution is defined as the ratio of two numerical values. This ratio is expressed as a percentage. The percent contribution only depends on (i) the known capacities of the load components of the balance and (ii) the regression coefficients that are the result of the regression analysis of the balance calibration data. For convenience, the regression coefficient nomenclature introduced in Ref. [1], Eq. (3.1.3), may be used to illustrate the calculation of the percent contribution. Then, the first numerical value can be expressed as the product of the regression coefficient of the primary linear term "$b1(i,i)$" of the regression model of the gage output with the capacity "$K(i)$" of the related balance load component. We get:

$$Q(i) \quad = \quad b1(i,i) \ \cdot \ K(i) \tag{1}$$

Equation (1) defines the reference value that is used to investigate the linearity of the regression model of the gage output. It is the <u>denominator</u> of the ratio that defines the percent contribution. The <u>numerator</u> of the ratio, on the other hand, is the product of the investigated regression coefficient with the related regressor variable value that is obtained by using the load capacities as applied loads. Now, the percent contribution of the ten math term type groups defined in Ref. [1], Eq. (3.1.3), can be summarized as follows:

$$P(i,b1) \quad = \quad 100\ \% \ \cdot \ [\ b1(i,j) \ \cdot \ K(j)\ ] \ / \ Q(i) \tag{2a}$$

$$P(i,b2) \quad = \quad 100\ \% \ \cdot \ [\ b2(i,j) \ \cdot \ |K(j)|\ ] \ / \ Q(i) \tag{2b}$$

$$P(i,c1) \quad = \quad 100\ \% \ \cdot \ \left[\ c1(i,j) \ \cdot \ K(j)^2\ \right] \ / \ Q(i) \tag{2c}$$

$$P(i,c2) \quad = \quad 100\ \% \ \cdot \ [\ c2(i,j) \ \cdot \ K(j) \ \cdot \ |K(j)|\ ] \ / \ Q(i) \tag{2d}$$

$$P(i,c3) \quad = \quad 100\ \% \ \cdot \ [\ c3(i,j,k) \ \cdot \ K(j) \ \cdot \ K(k)\ ] \ / \ Q(i) \tag{2e}$$

$$P(i,c4) \quad = \quad 100\ \% \ \cdot \ [\ c4(i,j,k) \ \cdot \ |K(j) \ \cdot \ K(k)|\ ] \ / \ Q(i) \tag{2f}$$

$$P(i,c5) \quad = \quad 100\ \% \ \cdot \ [\ c5(i,j,k) \ \cdot \ K(j) \ \cdot \ |K(k)|\ ] \ / \ Q(i) \tag{2g}$$

$$P(i,c6) \quad = \quad 100\ \% \ \cdot \ [\ c6(i,j,k) \ \cdot \ |K(j)| \ \cdot \ K(k)\ ] \ / \ Q(i) \tag{2h}$$

$$P(i,d1) \quad = \quad 100\ \% \ \cdot \ \left[\ d1(i,j) \ \cdot \ K(j)^3\ \right] \ / \ Q(i) \tag{2i}$$

$$P(i,d2) \quad = \quad 100\ \% \ \cdot \ \left[\ d2(i,j) \ \cdot \ |K(j)^3|\ \right] \ / \ Q(i) \tag{2j}$$

where $j = 1, \cdots, n$ and $k = j+1, \cdots, n$. The numerator of the percent contribution, i.e., $Q(i)$, depends on the gage output index i. It changes whenever the index of the gage output changes during the calculation of the percent contribution.

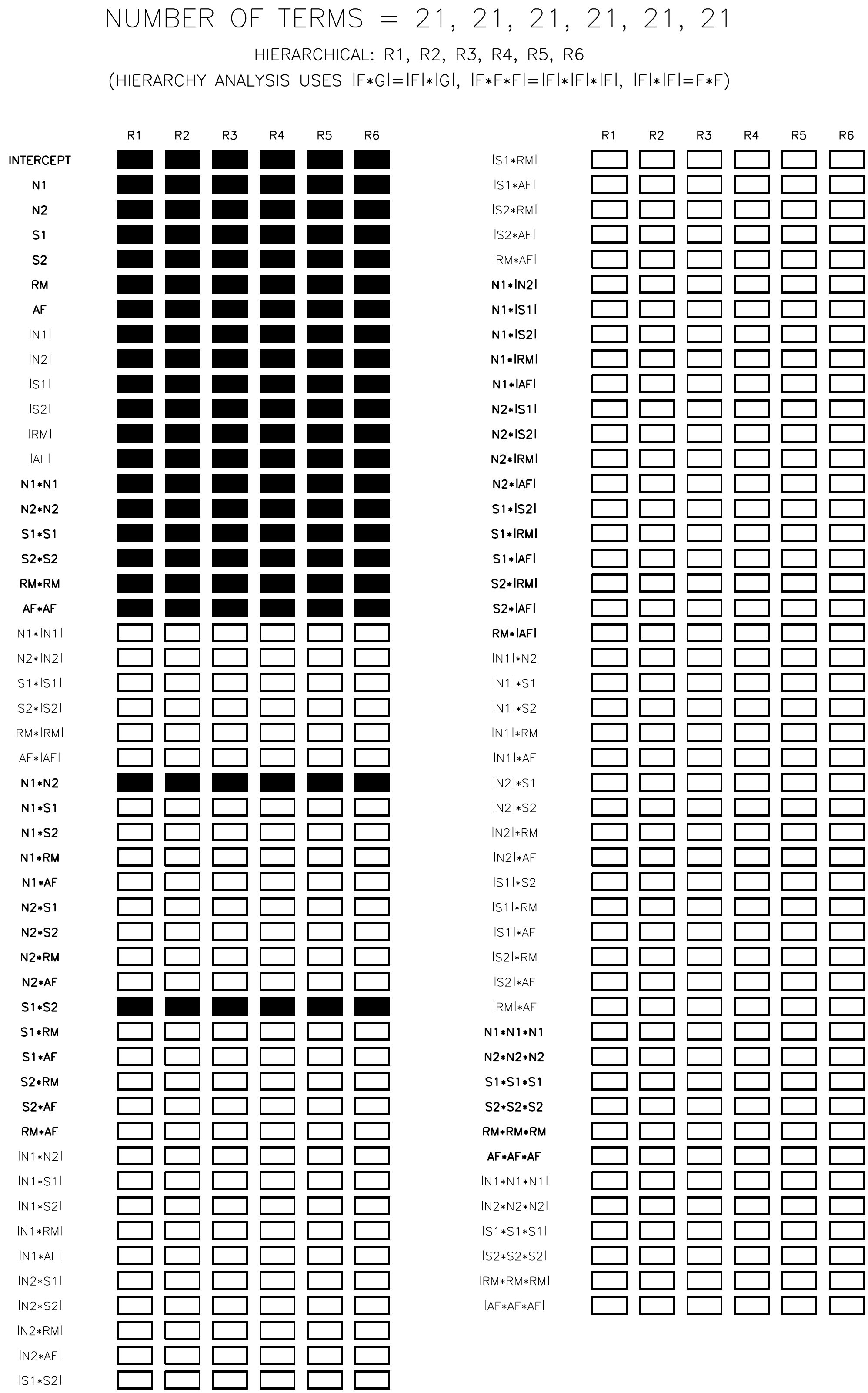


Fig. 1 Regression model term choice for the analysis of the MK40 balance calibration data.

INDEX	TERM	R1	R2	R3	R4	R5	R6
1	INTERCEPT	−1.59 %	−0.15 %	−0.10 %	−4.52 %	−1.50 %	−9.73 %
2	N1	[100.00 %]	−10.95 %	+1.81 %	−0.01 %	−0.03 %	+1.17 %
3	N2	−2.33 %	[100.00 %]	+0.32 %	+0.51 %	−0.10 %	−0.71 %
4	S1	−0.11 %	+0.57 %	[100.00 %]	−4.35 %	+0.09 %	−0.51 %
5	S2	−0.02 %	+0.20 %	−1.42 %	[100.00 %]	−0.07 %	+0.96 %
6	RM	+0.37 %	−0.29 %	+2.48 %	+2.73 %	[100.00 %]	+1.68 %
7	AF	−0.04 %	−0.08 %	+0.03 %	−3.85e−03 %	+0.40 %	[100.00 %]
8	|N1|	+0.71 %	+0.33 %	−0.70 %	−0.42 %	+0.09 %	+0.85 %
9	|N2|	+0.48 %	+0.91 %	−0.15 %	−0.82 %	+0.02 %	−0.17 %
10	|S1|	−0.48 %	−0.23 %	+1.81 %	+0.28 %	−0.06 %	+0.20 %
11	|S2|	−0.16 %	−0.25 %	+0.32 %	+1.99 %	−0.05 %	+0.46 %
12	|RM|	−0.71 %	−0.49 %	+2.30 %	+1.03 %	+0.11 %	−0.15 %
13	|AF|	+0.01 %	+0.02 %	−0.04 %	−0.05 %	−0.02 %	+0.15 %
14	N1*N1	+0.26 %	+0.09 %	+8.78e−05 %	+0.47 %	−0.06 %	−1.01 %
15	N2*N2	+0.10 %	+0.45 %	−0.08 %	+0.78 %	−0.04 %	+1.02 %
16	S1*S1	+0.25 %	+0.13 %	+0.63 %	−0.06 %	+0.05 %	−0.24 %
17	S2*S2	+0.09 %	+0.12 %	+0.01 %	+0.90 %	+0.02 %	−1.03 %
18	RM*RM	+0.29 %	+0.14 %	−1.51 %	−0.58 %	−0.08 %	+1.36 %
19	AF*AF	−8.04e−04 %	−0.01 %	+0.01 %	−0.01 %	+0.01 %	−0.11 %
20	N1*|N1|	0	0	0	0	0	0
21	N2*|N2|	0	0	0	0	0	0
22	S1*|S1|	0	0	0	0	0	0
23	S2*|S2|	0	0	0	0	0	0
24	RM*|RM|	0	0	0	0	0	0
25	AF*|AF|	0	0	0	0	0	0
26	N1*N2	−0.78 %	−1.18 %	+0.95 %	+0.56 %	−0.08 %	+1.31 %
27	N1*S1	0	0	0	0	0	0
28	N1*S2	0	0	0	0	0	0
29	N1*RM	0	0	0	0	0	0
30	N1*AF	0	0	0	0	0	0
31	N2*S1	0	0	0	0	0	0
32	N2*S2	0	0	0	0	0	0
33	N2*RM	0	0	0	0	0	0
34	N2*AF	0	0	0	0	0	0
35	S1*S2	+0.39 %	+0.25 %	−0.16 %	+0.04 %	+0.04 %	+0.29 %
36	S1*RM	0	0	0	0	0	0
37	S1*AF	0	0	0	0	0	0
38	S2*RM	0	0	0	0	0	0
39	S2*AF	0	0	0	0	0	0
40	RM*AF	0	0	0	0	0	0
41	|N1*N2|	0	0	0	0	0	0
42	|N1*S1|	0	0	0	0	0	0
43	|N1*S2|	0	0	0	0	0	0
44	|N1*RM|	0	0	0	0	0	0
45	|N1*AF|	0	0	0	0	0	0
46	|N2*S1|	0	0	0	0	0	0
47	|N2*S2|	0	0	0	0	0	0
48	|N2*RM|	0	0	0	0	0	0
49	|N2*AF|	0	0	0	0	0	0
50	|S1*S2|	0	0	0	0	0	0

BLUE −−−> ABS[PER. CONT.] < 0.05 % (VERY SMALL CONTRIBUTION)   RED −−−> ABS[PER. CONT.] > 25 % (VERY LARGE OFF−DIAGONAL CONTRIBUTION)

[...]. I.E., SQUARE BRACKETS −−−> PERCENT CONTRIBUTION = 100 % (REFERENCE TERM)

Fig. 2 Percent contribution of regression model terms of the strain–gage outputs.

Evaluation of Regression Models of Balance Calibration Data using an Empirical Criterion

N. Ulbrich* and T. Volden**
Jacobs Technology Inc., Moffett Field, California 94035–1000

An empirical criterion for assessing the significance of individual terms of regression models of wind tunnel strain–gage balance outputs is evaluated. The criterion is based on the percent contribution of a regression model term. It considers a term to be significant if its percent contribution exceeds the empirical threshold of 0.05 %. The criterion has the advantage that it can easily be computed using the regression coefficients of the gage outputs and the load capacities of the balance. First, a definition of the empirical criterion is provided. Then, it is compared with an alternate statistical criterion that is widely used in regression analysis. Finally, calibration data sets from a variety of balances are used to illustrate the connection between the empirical and the statistical criterion. A review of these results indicated that the empirical criterion seems to be suitable for a crude assessment of the significance of a regression model term as the boundary between a significant and an insignificant term cannot be defined very well. Therefore, regression model term reduction should only be performed by using the more universally applicable statistical criterion.

Nomenclature

AF	= axial force component of a force balance
$b1, b2$	= coefficients of the regression model of a gage output, defined in Ref. [1], Eq. (3.1.3)
C	= capacity of a primary gage load of a balance
$c1, c2, \cdots, c6$	= coefficients of the regression model of a gage output, defined in Ref. [1], Eq. (3.1.3)
$d1, d2$	= coefficients of the regression model of a gage output, defined in Ref. [1], Eq. (3.1.3)
F, G	= generic strain–gage balance load symbols
i	= index of a gage output –or– index of a primary gage load
j, k	= index of a load component
n	= number of load components of a balance
$N1$	= forward normal force component of a force balance
$N2$	= aft normal force component of a force balance
PC	= percent contribution of a coefficient of the regression model of a gage output
$R1, \cdots, R6$	= strain–gage outputs of a force balance
RM	= rolling moment component of a force balance
$S1$	= forward side force component of a force balance
$S2$	= aft side force component of a force balance
$\alpha, \beta, \epsilon, \lambda$	= regression coefficients of the output of the forward normal force gage
μ, ν, ξ, ρ	= regression coefficients of the output of the axial force gage
Λ	= indicator variable for bi–directionality of a primary gage output

I. Introduction

Different approaches are used in the wind tunnel testing community to perform the regression analysis

* Aerodynamicist, Jacobs Technology Inc.
** Computer Engineer, Jacobs Technology Inc.

of wind tunnel strain–gage balance calibration data. Many analysts prefer the *Iterative Method* as this approach fits strain–gage outputs as a function of calibration loads (see Ref. [1] for a detailed description of the method). The solution of the corresponding global regression analysis problem defines coefficients of a multivariate math model for each gage output as the individual outputs of a typical wind tunnel balance respond to more than one calibration load component.

Function classes like linear terms, absolute value terms, square terms, and cross–product terms may be used to assemble a suitable regression model of the strain–gage outputs. Ultimately, a set of regression model terms has to be selected such that (i) the behavior of the strain–gage outputs of the balance can be characterized correctly, (ii) near–linear dependencies between terms are avoided, and (iii) the balance calibration data is not over–fitted. Analysts use a host of different criteria to both evaluate and select individual terms of the regression model of the gage outputs so that the final regression model meets the chosen quality requirements.

Traditionally, an empirical criterion is applied in the wind tunnel testing community to assess the significance of individual terms of the regression model of the outputs whenever the *Iterative Method* is chosen for the analysis of balance calibration data. This criterion is based on the so–called percent contribution of a regression model term (see the appendix for a definition of the percent contribution). Some analysts prefer to apply a statistical criterion in order to assess the importance of individual terms of a regression model. This alternate criterion is widely used in linear regression analysis.

At this point a question may come up: how does the final term selection obtained from the application of the empirical criterion compare with the term selection that would result from the application of the statistical criterion? The present paper tries to find an answer to this question by applying both term selection criteria to a variety of balance calibration data sets. In addition, the selection of the terms resulting from the application of the two criteria is evaluated by using some knowledge about the physical characteristics of a specific balance design.

In the next section of the paper the empirical and the statistical criterion are defined. Then, data from the calibration of NASA's MK29B balance is used to compare subsets of recommended terms that were obtained after the application of the two criteria with the term selection that results from knowledge about the design of the balance. Finally, results from the analysis of 12 different balances are presented that assess the validity of the percent contribution threshold that the empirical criterion uses.

II. Definition of Empirical and Statistical Criterion

Basic definitions of the empirical and statistical criterion are reviewed in this section so that a meaningful comparison of the results of the application of the criteria can be made. In principle, both criteria may independently be used to eliminate insignificant terms in the regression model of the outputs of a strain–gage balance so that over–fitting of balance calibration data is prevented.

First, the empirical criterion for the assessment of regression model terms is discussed. It is based on the percent contribution of a regression coefficient (see the appendix of this paper for a definition of the percent contribution). The criterion compares the absolute value of the percent contribution of an individual term of a regression model with an empirical threshold in order to decide whether or not the term is significant and should be retained in the model. The empirical criterion can be summarized as follows:

Empirical Criterion for Term Significance

A term of the regression model of a strain–gage output is considered "significant" if the absolute value of the percent contribution of the term is greater than 0.05%.

The use of the percent contribution for the assessment of the significance of regression model terms has an advantage: it is a metric that can easily be computed for a given regression model term. It is only a function of (i) the load capacities of the balance and (ii) the regression coefficient values of the least squares fit of the strain–gage outputs.

The application of the statistical criterion, on the other hand, is more complex than the application of the empirical criterion. The statistical criterion assesses the significance of terms of a regression model by looking at the standard error of each regression coefficient. The standard error is an estimate of the standard deviation of the coefficient. It can be thought of as a measure of the precision with which the regression coefficient is measured. A coefficient should be included in the math model if it is large compared to its standard error.

The t–statistic of a coefficient is used to quantitatively compare a regression coefficient with its standard error. It equals the ratio between the coefficient value and its standard error. A coefficient is probably "significant" if its t–statistic is greater than the critical value of the Student's t–distribution (see Ref. [2], p.32). This comparison can also be performed by using the p–value of the coefficient. The p–value of a coefficient is determined from a comparison of the t–statistic with values in the Student's t–distribution. With a p–value of, e.g., 0.0001 (or 0.01 %) one can say with a 99.99 % probability of being correct that the regression coefficient is having some effect. The statistical criterion for testing the significance of a regression model term can be summarized as follows:

Statistical Criterion for Term Significance

A term of the regression model of a strain–gage output is considered "significant" if the p–value of the t–statistic of its regression coefficient is less than 0.0001, or, if 1.0 minus the p–value of the t–statistic of its coefficient is greater than 0.9999.

The statistical criterion applies the p–value threshold of 0.0001 during the assessment of the significance of a term. This threshold choice is universally accepted and widely used in linear regression analysis (it is, for example, applied in the highly popular regression analysis software package distributed by SAS of Cary, North Carolina). Unfortunately, the authors were unable to find a reference for the justification of the percent contribution threshold of 0.05 % that the empirical criterion uses. Therefore, they decided to study the validity of the empirical criterion's threshold by using results from the application of the statistical criterion as a point of reference.

First, the two criteria's regression model term assessments are compared for a case when information about the design of a strain–gage balance is available. Then, results from the application of the two criteria to a variety of balance calibration data sets are discussed.

III. Balance Calibration Data Example

Machine calibration data of NASA's MK29B balance was selected for the assessment of the two criteria. The MK29B balance is a six–component force balance that was manufactured by the Task Corporation. The calibration data used for the study was obtained in Triumph Aerospace's balance calibration machine. Table 1 below summarizes important characteristics of the balance and calibration data set that was used for the present study.

Table 1: Balance and calibration data set characteristics of the MK29B balance.

BALANCE NAME	2.0 MK29B
BALANCE TYPE	FORCE BALANCE (TASK)
DIAMETER	2.0 [in]
GAGE DISTANCE (NORMAL FORCE GAGES)	7.25 [in]
GAGE DISTANCE (SIDE FORCE GAGES)	6.00 [in]
CALIBRATION DATE	SEPTEMBER 2007
CALIBRATION METHOD	MACHINE CALIBRATION
TOTAL NUMBER OF CALIBRATION POINTS	1751

The calibration data set makes a complete characterization of the physical behavior of the balance possible because it covers the entire use envelope that the balance may experience during a wind tunnel test. It was decided to process the data in its "design" format, i.e., in force balance format. In this case important connections between the primary gage loads and gage outputs become visible. Table 2 below lists load capacities of the balance in force balance format:

Table 2: Capacities of the MK29B balance.

	$N1$, lbs	$N2$, lbs	$S1$, lbs	$S2$, lbs	RM, in–lbs	AF, lbs
CAPACITY	2100	2100	700	700	3800	350

Different math term group combinations may be selected for the development of a regression model of the six strain–gage outputs of the balance. The math term groups themselves are defined in Ref. [1]. One of five group combination choices may be chosen. These five combinations are summarized in Table 3 below.

Table 3: Math Term Group Combination Choices for Initial Data Analysis.

GROUP NUMBER	MATH TERM GROUP COMBINATION
(I)	$F,\ F \cdot G$
(II)	$F,\ F^2,\ F \cdot G$
(III)	$F,\ \lvert F\rvert,\ F^2,\ F \cdot G$
(IV)	$F,\ \lvert F\rvert,\ F^2,\ F \cdot \lvert F\rvert,\ F \cdot G$
(V)	$F,\ \lvert F\rvert,\ F^2,\ F \cdot \lvert F\rvert,\ F \cdot G,\ \ F \cdot \lvert G\rvert,\ \lvert F\rvert \cdot G,\ \lvert F\rvert \cdot \lvert G\rvert$

Some of the gage outputs of the MK29B are bi–directional. Therefore, absolute value terms should be a part of the group combination for an initial analysis of the calibration data. It was decided to use group (V) for the analysis. A numerical technique called Singular Value Decomposition was applied to ensure that the regression model of each gage output would be free of linear dependencies between terms. Then, the initial regression model of each of the six gage outputs of the MK29B balance supports a total of 85 terms. Figure 1a shows the corresponding regression model term selection for the MK29B. It is the authors' experience that only about 1/4 to 1/3 of the terms are needed to correctly describe the physical behavior of the gage outputs of the balance.

The MK29B balance is a Task balance. Therefore, only the normal and side force gage outputs of the balance will have significant levels of bi–directionality (see Ref. [3] for a discussion of bi–directionality characteristics of different balance types). Figure 1b shows the bi–directionality characteristics of the MK29B. It can clearly be seen that the regression models of the normal and side force gage outputs may need absolute value terms as the indicator variable Λ for bi–directionality has the appearance of an absolute value function.

For the present investigations it was decided to focus on the regression models of outputs of the forward normal force gage ($R1$) and the axial force gage ($R6$). First, the application of the empirical and statistical criterion to the outputs of the forward normal force gage is investigated in detail.

Figure 2a shows a comparison of <u>p–value</u> and <u>percent contribution</u> for regression model of the forward normal force gage output. The orange color highlights the values for the primary gage load ($N1$). We see that the p–value is less than the threshold of 0.0001 and the percent contribution is 100.00 %. Therefore, both criteria say that the term is significant. The green color highlights all terms that are significant based on an application of the statistical criterion. A total of 26 of the 85 terms are significant according to the statistical criterion. The empirical criterion is also fulfilled for all 26 terms as the absolute value of the percent contribution of each one of the 26 terms exceeds the threshold of 0.05 %.

It is interesting to study the connection between a set of cross–product terms that are a part of the regression model of the forward normal force gage output. Therefore, the significance of the following four cross–product terms is evaluated: $N1\ N2$, $N1\ \lvert N2\rvert$, $\lvert N1\rvert\ N2$, and $\lvert N1\ N2\rvert$. The yellow color in Fig. 2b highlights values for both the <u>p–value</u> and <u>percent contribution</u> of all terms that are connected with the four cross–product terms (the numerical values reported in Fig. 2a and 2b are identical). For clarity, these values may be listed in table format. First, Table 4a compares the significance of the four linear terms that are indirectly related to the four cross–product terms.

Table 4a: Significance of selected linear terms of regression model of $R1$.

MATH TERM	p–value (p)	$1 - p$	percent contribution
$N1$	$<$ 0.0001	$>$ 0.9999	$>$ 0.05 % (100.00 %)
$N2$	$<$ 0.0001	$>$ 0.9999	$>$ 0.05 % (3.10 %)
$\|N1\|$	$<$ 0.0001	$>$ 0.9999	$>$ 0.05 % (0.98 %)
$\|N2\|$	$<$ 0.0001	$>$ 0.9999	$>$ 0.05 % (0.64 %)

It can be seen that both the statistical criterion (p–value < 0.0001) and the empirical criterion (percent contribution > 0.05 %) are fulfilled for all four terms. They are significant as far as the regression model of the forward normal force gage output is concerned. It remains to inspect the percent contribution and p–value of the cross–product terms themselves. These values are listed in Table 4b.

Table 4b: Significance of selected cross–product terms from regression model of $R1$.

MATH TERM	p–value (p)	$1 - p$	percent contribution
$N1\ N2$	$<$ 0.0001	$>$ 0.9999	$>$ 0.05 % (0.23 %)
$\|N1\ N2\| = \|N1\|\ \|N2\|$	$<$ 0.0001	$>$ 0.9999	$>$ 0.05 % (0.24 %)
$N1\ \|N2\|$	$<$ 0.0001	$>$ 0.9999	$>$ 0.05 % (0.37 %)
$\|N1\|\ N2$	$<$ 0.0001	$>$ 0.9999	$>$ 0.05 % (0.26 %)

Again, as it was the case with the terms listed in Table 4a, it can be seen that both the statistical criterion (p–value < 0.0001) and the empirical criterion (percent contribution > 0.05 %) are fulfilled for all four terms. All four cross–product terms are significant as far as the regression model of the forward normal force gage output is concerned. The reason for the importance of the four cross–product terms can be better understood if we look at their origin. They are the cross–product terms that are obtained after taking the bi–directionality of the gage outputs of the normal force gages into account. Thus, we get:

$$\Big[\ \alpha\ N1\ +\ \beta\ |N1|\ \Big] \cdot \Big[\ \epsilon\ N2\ +\ \lambda\ |N2|\ \Big] \quad = \quad (\alpha\epsilon)\ N1\ N2\ +\ (\alpha\lambda)\ N1\ |N2| \ +\ (\beta\epsilon)\ |N1|\ N2\ +\ (\beta\lambda)\ |N1|\ |N2| \tag{1}$$

The output of the forward normal force gage is simply sensitive to the loads that are applied over both the forward and aft normal force gage. Therefore, all four cross–product terms are needed in the regression model of the forward normal force output.

Situations exist, however, when not all related cross–product terms are needed in the regression model of a gage output. The regression model of the axial gage output may be used to illustrate this statement. Figure 3a shows a comparison of <u>p–value</u> and <u>percent contribution</u> for the regression model of the axial gage output. The orange color highlights the values for the primary gage load (AF). We see that the p–value is less than the threshold of 0.0001 and the percent contribution is 100 %. Therefore, both criteria say that the term is significant. The green color highlights all terms that are significant based on an application of the statistical criterion. A total of 30 of the 85 terms are considered significant based on the statistical criterion. The empirical criterion is also fulfilled for all 30 terms as the absolute value of the percent contribution of each one of the 30 terms exceeds the threshold of 0.05 %.

Now, the significance of the following four cross–product terms is evaluated: $N2\ AF$, $N2\ |AF|$, $|N2|\ AF$, and $|N2|\ |AF|$. The yellow and blue color in Fig. 3b highlights values for both the <u>p–value</u>

and percent contribution of terms that make up the four cross–product terms (the numerical values reported in Fig. 3a and 3b are identical). For clarity, these values may be listed in table format. First, Table 5a compares the significance of the four linear terms that are indirectly related to the four cross–product terms.

Table 5a: Significance of selected linear terms from regression model of $R6$.

MATH TERM	p–value (p)	$1 - p$	percent contribution
$N2$	< 0.0001	> 0.9999	$>$ 0.05 % (5.13 %)
AF	< 0.0001	> 0.9999	$>$ 0.05 % (100.00 %)
$\|N2\|$	< 0.0001	> 0.9999	$>$ 0.05 % (0.77 %)
$\|AF\|$	0.0741	0.9259	$>$ 0.05 % (0.08 %)

The empirical criterion (percent contribution $>$ 0.05 %) is fulfilled for all four terms that are listed above. The result for the statistical criterion, i.e., p–value < 0.0001, is different. Only three of the four terms appear to be significant. The statistical criterion identified the term $|AF|$ as an insignificant term in the regression model of the axial gage output. It remains to inspect the percent contribution and p–value of the cross–product terms themselves before a final evaluation of the discrepancy between the two criteria can be made. These values are listed in Table 5b.

Table 5b: Significance of selected cross–product terms from regression model of $R6$.

MATH TERM	p–value (p)	$1 - p$	percent contribution
$N2\ AF$	< 0.0001	> 0.9999	$>$ 0.05 % (0.30 %)
$\|N2\ AF\| = \|N2\|\ \|AF\|$	0.0074	0.9926	$>$ 0.05 % (0.12 %)
$N2\ \|AF\|$	0.0841	0.9159	$>$ 0.05 % (0.06 %)
$\|N2\|\ AF$	0.0010	0.9990	$>$ 0.05 % (0.13 %)

Again, the empirical criterion (percent contribution $>$ 0.05 %) says that all four cross–product terms are significant. The statistical criterion, on the other hand, is only fulfilled for a single term ($N2\ AF$). It considers the remaining three cross–product terms ($N2\ |AF|$, $|N2|\ AF$, $|N2|\ |AF|$) to be insignificant.

The reason for the discrepancy between the results for the empirical and statistical criterion has to be investigated in more detail. The discrepancy can be better understood if we look at the origin of the four cross–product terms. We get:

$$\Big[\ \mu\ N2\ +\ \nu\ |N2|\ \Big]\ \cdot\ \Big[\ \xi\ AF\ +\ \rho\ |AF|\ \Big]\ =\ (\mu\xi)\ \underbrace{N2\ AF}_{p<0.0001}\ +\ (\mu\rho)\ \underbrace{N2\ |AF|}_{p=0.0841}\ +\ (\nu\xi)\ \underbrace{|N2|\ AF}_{p=0.0010}\ +\ (\nu\rho)\ \underbrace{|N2|\ |AF|}_{p=0.0074} \qquad (2)$$

From Fig. 1b it is known that the axial gage output of Task balance like the MK29B is not bi–directional. Therefore, there is no objective justification for using an absolute value term of the axial force in the regression model. This conclusion is also supported by the corresponding p–value that is listed in Table 5a:

$$p - value\ \ of\ \ |AF|\ =\ 0.0741\ \gg\ 0.0001\ \Longrightarrow\ \rho\ |AF|\ \approx\ 0 \qquad (3)$$

Now, after using the result reported in Eq. (3) in Eq. (2), we get the following result for the cross–product terms:

$$\left[\ \mu\ N2\ +\ \nu\ |N2|\ \right]\ \cdot\ \left[\ \xi\ AF\ +\ 0\ \right]\ \approx\ (\mu\xi)\ \underbrace{N2\ AF}_{p<0.0001}\ +\ (\nu\xi)\ \underbrace{|N2|\ AF}_{p=0.0010} \tag{4}$$

A further investigation of the p–value and the percent contribution of the cross–product term $|N2|\ AF$ was inconclusive. Its p–value is 0.0010 and its percent contribution is 0.13 %. Therefore, the term is a borderline case between being significant and being insignificant as the p–value and the percent contribution are relatively close to their respective thresholds.

The analysis of the cross–product terms of $N2$ and AF of the regression model of the axial gage output also illustrated that the four theoretically possible cross–product terms $N2\ AF$, $N2\ |AF|$, $|N2|\ AF$, and $|N2|\ |AF|$ do not have to be present simultaneously in the regression model of the gage outputs. Some of the four terms may simply be omitted if either the statistical or the empirical criterion indicate they are insignificant. The omission of the two terms $N2\ |AF|$ and $|N2|\ |AF|$ in Eq. (4) is also supported by a known design characteristic of the axial flexure element of the MK29B Task balance. The axial gage output of the MK29B is not bi–directional, i.e., the term $|AF|$ in its regression model is insignificant, because the axial flexure element is joined to the metric and non–metric part of the balance by using tight press pins.

In the next section it is explained how a connection between the threshold of the statistical criterion and the threshold of the empirical criterion can be established. This connection is first illustrated by using the MK29B data set. Then, results of a survey of 12 different balance calibration data sets are discussed. The data sets were processed to evaluate the magnitude of the empirical criterion's threshold.

IV. Survey of Balance Calibration Data

The authors performed a survey of different balance calibration data sets to investigate whether or not the empirical criterion's threshold of 0.05 % is applicable to a wide variety of balance calibration data sets. Therefore, it was decided to define a connection between the threshold of the statistical criterion and the threshold of the empirical criterion so that the statistical criterion may be used to evaluate the empirical criterion.

Different approaches may be used to establish a connection between the statistical and the empirical criterion. One approach, for example, counts the number of regression model terms that simultaneously satisfy the statistical and the empirical criterion for a given percent contribution threshold choice. This process is repeated over the entire range of expected percent contribution thresholds (from $\approx$ 0.001 % to 100.0 %). Then, the resulting term count is plotted versus the corresponding percent contribution threshold choice. The "ideal" percent contribution threshold is the largest threshold that, if chosen, still maximizes the term count.

Results for the regression model of the axial gage output of the MK29B balance may be used to demonstrate the approach that is described in the previous paragraph. The p–values and percent contributions listed in Fig. 3a are used for this study. First, however, in order to highlight important characteristics of the empirical criterion itself, the number of significant terms of the 85 term model of the axial gage output is counted by only applying the empirical criterion over the expected percent contribution threshold range. The intercept term is intentionally not counted during this investigation because its p–value is not computed. Therefore, the maximum of the term count becomes 84. Figure 4a shows the result of these calculation. It can be seen that the term count remains constant as long as the percent contribution threshold is below the minimum that is obtained for the regression model of the axial gage output (the minimum is listed as 0.00406 % in Fig. 3a). Then, the term count monotonically decreases as the threshold increases. It reaches the theoretical minimum of zero when the percent contribution threshold is set to 100 %.

Now, using the approach proposed in the second paragraph of this section, the terms are counted that simultaneously satisfy the statistical and the empirical criterion for a given percent contribution threshold choice. Figure 4b shows corresponding results. It can be seen that the term count first remains constant as the percent contribution threshold increases. At a certain value, i.e., at $\approx$ 0.13 %, the term count suddenly starts to decrease. This threshold is the "ideal" percent contribution threshold that should be used if the empirical criterion is applied to the regression model of the axial gage outputs. It defines the boundary between "over–fitting" and "under–fitting" of the axial gage outputs. A decrease of the percent contribution threshold, for example, would result in the inclusion of insignificant terms in the regression model. These additional terms could cause "over–fitting" of the axial gage outputs. An increase of the percent contribution threshold, on the other hand, would result in the omission of significant terms in the regression model. These missing terms could cause "under–fitting" of the axial gage outputs.

The approach discussed in the previous three paragraphs is complex. An alternate and simpler connection between the statistical and empirical criterion can also be defined if, for example, the smallest percent contribution is found for a given regression model of the strain–gage outputs that still satisfies the statistical criterion. Consequently, terms below this percent contribution will no longer satisfy the statistical criterion, i.e., p–value $<$ 0.0001, and are considered statistically insignificant.

The connection between the two thresholds can be illustrated by using the calibration data set of the MK29B that is discussed in the previous section. Figure 2a shows the percent contributions of the regression model of the forward normal force gage output of the MK29B balance. The smallest absolute value of the percent contribution that satisfied the statistical criterion is the percent contribution of term 29, i.e., 0.09 %. It must be pointed out that the percent contribution of the intercept term must be ignored during the search for the percent contribution minimum as the p–value is not defined in this case. Similarly, the minima of the percent contributions of the regression models of the remaining five primary gage outputs can be determined. All six minima of the 85 term math model, i.e., of math term group combination (V) in Table 3, are listed in Table 6 below.

Table 6: Percent contribution minima of regression models of MK29B calibration data.

	GROUP (TABLE 3)	$R1$	$R2$	$R3$	$R4$	$R5$	$R6$
ABS(MINIMUM)	(V)	**0.09 %**	0.11 %	0.26 %	0.19 %	0.20 %	0.15 %

The minimum of the six minima is 0.09 %. This value, if chosen as the threshold for the empirical criterion, would guarantee that all significant terms of the regression models of the six gage outputs are retained that the statistical criterion would identify for the same data set and math term group combination.

The minimum of the six minima may change if the math term group combination is changed from, say, combination (V) to combination (IV). The MK29B balance is a multi–piece balance that typically needs absolute value terms. Therefore, the minima for group combinations (III), (IV), and (V) were investigated in order to get an idea of the dependency of the percent contribution threshold on the math term group combination. The results of these investigations are shown in the table in Fig. 5a.

The percent contribution threshold of the empirical criterion does not just depend on the math term group combination that is chosen for the regression analysis of the gage outputs. It may also depend on the balance type, whether or not a tare load iteration was performed, and on the calibration process that a balance calibration laboratory uses. Therefore, calibration data sets of a total of 12 different balances were investigated in order to get a more general idea of the percent contribution threshold variation.

Figure 5a shows the percent contribution thresholds that were obtained for multi–piece balance data sets. Figure 5b shows the percent contribution thresholds that were obtained for single–piece and hybrid balance data sets. Figure 5c shows the percent contribution thresholds that were obtained for semispan balance data sets. The results for the recommended math term group combinations of each balance type are highlighted in boldface. The arithmetic mean of all thresholds highlighted in boldface is 0.04 %. This value is very close to the original choice of 0.05 % that was listed in the original definition of the empirical criterion. However, it can also be concluded from the results of the different balance calibration data sets that the percent contribution threshold of 0.05 % is not universally applicable. Higher or lower values of the threshold appear to be appropriate for some of the data sets that were investigated.

V. Summary and Conclusions

An empirical criterion was evaluated that may be used to assess the significance of individual terms of regression models of wind tunnel strain–gage balance outputs. The criterion compares the percent contribution of a term with the threshold of 0.05 % in order to decide whether or not the term is significant.

Calibration data sets from a variety of strain–gage balances were processed to test the criterion. During these tests the term assessment of the empirical criterion was compared with the term assessment of an alternate statistical criterion. Therefore, a connection between the empirical and statistical criterion was first defined by identifying the smallest percent contribution for the regression models of a given balance calibration data set that still satisfied the statistical criterion. Then, the mean value of the percent contribution minima of all tested balance data sets was computed. It was found that this mean value, i.e., 0.04 %, is close to the proposed percent contribution threshold of 0.05 %.

The study of the different balance calibration data sets illustrated that the boundary between a significant and an insignificant term cannot be defined very well for the empirical criterion. Therefore, the empirical criterion seems to be suitable for a crude assessment of the significance of a regression model term. Consequently, regression model term reduction for the prevention of over–fitting of balance calibration data should only be performed by using the more reliable and universally applicable statistical criterion.

Acknowledgements

The authors would like to thank Jon Bader and Robert Gisler of NASA Ames Research Center for their critical and constructive review of the final manuscript. The work reported in this paper was supported by the Wind Tunnel Division at NASA Ames Research Center under contract NNA09DB39C.

References

[1]AIAA/GTTC Internal Balance Technology Working Group, "Recommended Practice, Calibration and Use of Internal Strain–Gage Balances with Application to Wind Tunnel Testing," AIAA R–091–2003, American Institute of Aeronautics and Astronautics, Reston, Virginia, 2003.

[2]Myers, R. H. and Montgomery, D. C., *Response Surface Methodology: Process and Product Optimization Using Designed Experiments*, 1st ed., John Wiley & Sons, Inc., New York, 1995.

[3]Ulbrich, N., "Detection of Bi–Directionality in Strain–Gage Balance Calibration Data," paper presented at the 28th AIAA Aerodynamic Measurement Technology and Ground Testing Conference, June 2012, New Orleans, Louisiana.

Appendix – Definition of the Percent Contribution

The percent contribution describes the contribution of each term of the regression model of the gage outputs to the total fitted value, expressed as a percentage of the contribution of the principle diagonal term. It is used to assess the degree of linearity of the regression model of the gage outputs.

The percent contribution is defined as the ratio of two numerical values. This ratio is expressed as a percentage. The percent contribution only depends on (i) the known capacities of the load components of the balance and (ii) the regression coefficients that are the result of the regression analysis of the balance calibration data. For convenience, the regression coefficient nomenclature introduced in Ref. [1], Eq. (3.1.3), may be used to illustrate the calculation of the percent contribution. Then, the first numerical value can be expressed as the product of the regression coefficient of the primary linear term "$b1(i,i)$" of the regression model of the gage output with the capacity "$C(i)$" of the related balance load component. We get:

$$Q(i) \quad = \quad b1(i,i) \;\cdot\; C(i) \tag{5}$$

Equation (5) defines the reference value that is used to investigate the linearity of the regression model of the gage output. It is the denominator of the ratio that defines the percent contribution. The numerator of the ratio, on the other hand, is the product of the investigated regression coefficient with the related regressor variable value that is obtained by using the load capacities as applied loads. Now, the percent contribution of the ten math term type groups defined in Ref. [1], Eq. (3.1.3), can be summarized as follows:

$$PC(i,b1) \quad = \quad 100\,\% \;\cdot\; [\; b1(i,j) \;\cdot\; C(j) \;] \;/\; Q(i) \tag{6a}$$

$$PC(i,b2) \quad = \quad 100\,\% \;\cdot\; [\; b2(i,j) \;\cdot\; |C(j)| \;] \;/\; Q(i) \tag{6b}$$

$$PC(i,c1) \quad = \quad 100\,\% \;\cdot\; \left[\; c1(i,j) \;\cdot\; C(j)^2 \;\right] \;/\; Q(i) \tag{6c}$$

$$PC(i,c2) \quad = \quad 100\,\% \;\cdot\; [\; c2(i,j) \;\cdot\; C(j) \;\cdot\; |C(j)| \;] \;/\; Q(i) \tag{6d}$$

$$PC(i,c3) \quad = \quad 100\,\% \;\cdot\; [\; c3(i,j,k) \;\cdot\; C(j) \;\cdot\; C(k) \;] \;/\; Q(i) \tag{6e}$$

$$PC(i,c4) \quad = \quad 100\,\% \;\cdot\; [\; c4(i,j,k) \;\cdot\; |C(j) \;\cdot\; C(k)| \;] \;/\; Q(i) \tag{6f}$$

$$PC(i,c5) \quad = \quad 100\,\% \;\cdot\; [\; c5(i,j,k) \;\cdot\; C(j) \;\cdot\; |C(k)| \;] \;/\; Q(i) \tag{6g}$$

$$PC(i,c6) \quad = \quad 100\,\% \;\cdot\; [\; c6(i,j,k) \;\cdot\; |C(j)| \;\cdot\; C(k) \;] \;/\; Q(i) \tag{6h}$$

$$PC(i,d1) \quad = \quad 100\,\% \;\cdot\; \left[\; d1(i,j) \;\cdot\; C(j)^3 \;\right] \;/\; Q(i) \tag{6i}$$

$$PC(i,d2) \quad = \quad 100\,\% \;\cdot\; \left[\; d2(i,j) \;\cdot\; |C(j)^3| \;\right] \;/\; Q(i) \tag{6j}$$

where $j = 1, \cdots, n$ and $k = j+1, \cdots, n$. The numerator of the percent contribution, i.e., $Q(i)$, depends on the gage output index i. It changes whenever the index of the gage output changes during the calculation of the percent contribution.

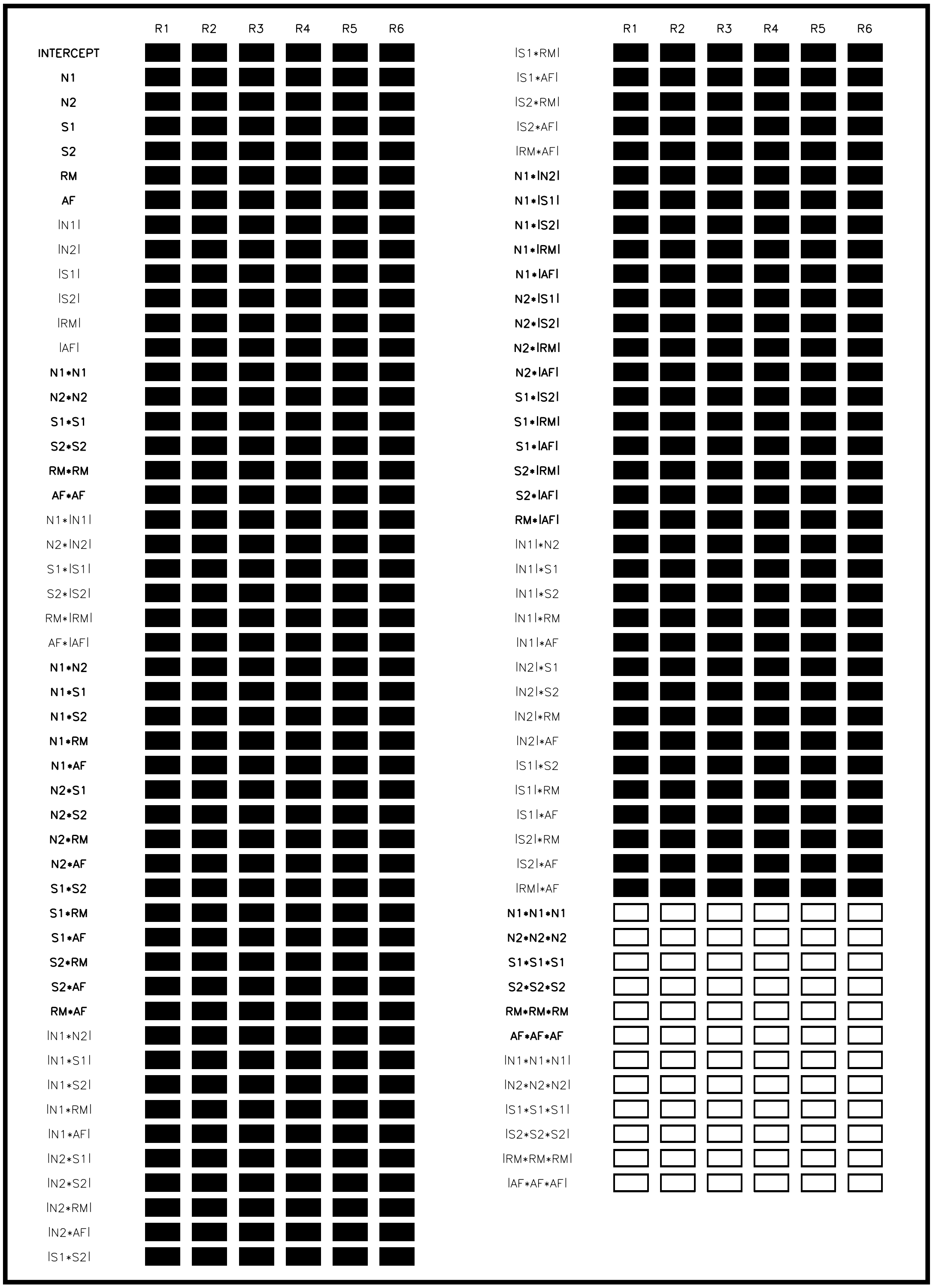
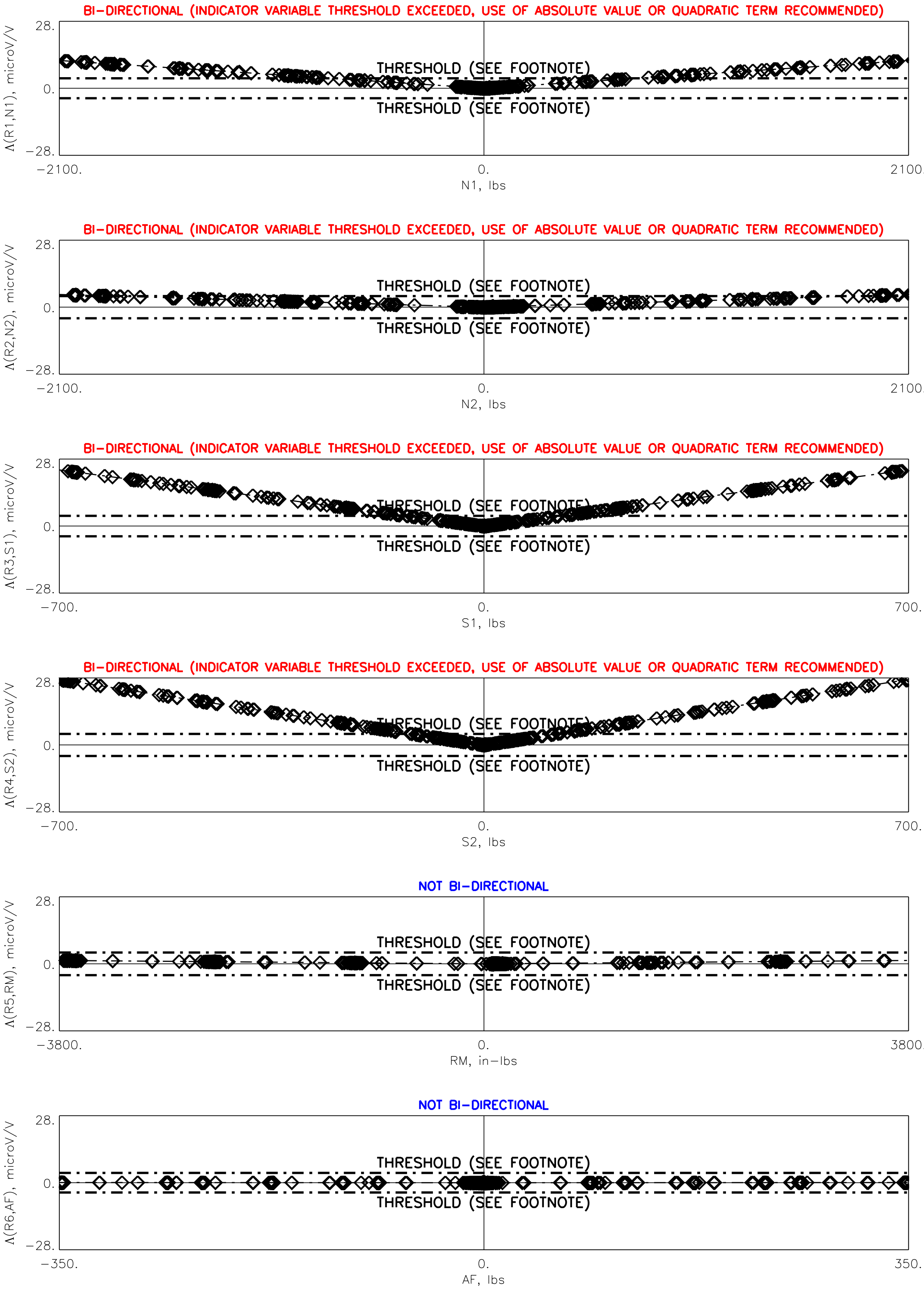


Fig. 1a Initial regression model term choice for the analysis of the MK29B balance calibration data set.



Footnote: *The threshold marked as a dot–dashed line is defined in Ref. [3], Eq. (1b).*

Fig. 1b Bi–directionality characteristics of NASA's MK29B Task balance.

<table>
<tr><th>INDEX</th><th>MATH TERM</th><th>p</th><th>1 - p</th><th>ABS(Percent Cont.)</th></tr>
<tr><th></th><th></th><th>[-]</th><th>[-]</th><th>[%]</th></tr>
<tr><td>1</td><td>INTERCEPT</td><td>-</td><td>-</td><td>0.04</td></tr>
<tr><td>2</td><td>N1</td><td><0.0001</td><td>>0.9999</td><td>[100.00]</td></tr>
<tr><td>3</td><td>N2</td><td><0.0001</td><td>>0.9999</td><td>3.10</td></tr>
<tr><td>4</td><td>S1</td><td><0.0001</td><td>>0.9999</td><td>0.21</td></tr>
<tr><td>5</td><td>S2</td><td>0.0009</td><td>0.9991</td><td>0.07</td></tr>
<tr><td>6</td><td>RM</td><td><0.0001</td><td>>0.9999</td><td>0.52</td></tr>
<tr><td>7</td><td>AF</td><td>0.7814</td><td>0.2186</td><td>0.01</td></tr>
<tr><td>8</td><td>|N1|</td><td><0.0001</td><td>>0.9999</td><td>0.98</td></tr>
<tr><td>9</td><td>|N2|</td><td><0.0001</td><td>>0.9999</td><td>0.64</td></tr>
<tr><td>10</td><td>|S1|</td><td>0.0007</td><td>0.9993</td><td>0.11</td></tr>
<tr><td>11</td><td>|S2|</td><td><0.0001</td><td>>0.9999</td><td>0.12</td></tr>
<tr><td>12</td><td>|RM|</td><td><0.0001</td><td>>0.9999</td><td>0.84</td></tr>
<tr><td>13</td><td>|AF|</td><td>0.9995</td><td>0.0005</td><td><0.01</td></tr>
<tr><td>14</td><td>N1*N1</td><td><0.0001</td><td>>0.9999</td><td>0.14</td></tr>
<tr><td>15</td><td>N2*N2</td><td>0.0008</td><td>0.9992</td><td>0.10</td></tr>
<tr><td>16</td><td>S1*S1</td><td>0.0185</td><td>0.9815</td><td>0.08</td></tr>
<tr><td>17</td><td>S2*S2</td><td>0.1989</td><td>0.8011</td><td>0.04</td></tr>
<tr><td>18</td><td>RM*RM</td><td><0.0001</td><td>>0.9999</td><td>0.27</td></tr>
<tr><td>19</td><td>AF*AF</td><td>0.2502</td><td>0.7498</td><td>0.03</td></tr>
<tr><td>20</td><td>N1*|N1|</td><td><0.0001</td><td>>0.9999</td><td>0.90</td></tr>
<tr><td>21</td><td>N2*|N2|</td><td><0.0001</td><td>>0.9999</td><td>0.80</td></tr>
<tr><td>22</td><td>S1*|S1|</td><td>0.4294</td><td>0.5706</td><td>0.02</td></tr>
<tr><td>23</td><td>S2*|S2|</td><td>0.0791</td><td>0.9209</td><td>0.05</td></tr>
<tr><td>24</td><td>RM*|RM|</td><td>0.0051</td><td>0.9949</td><td>0.07</td></tr>
<tr><td>25</td><td>AF*|AF|</td><td>0.6924</td><td>0.3076</td><td>0.01</td></tr>
<tr><td>26</td><td>N1*N2</td><td><0.0001</td><td>>0.9999</td><td>0.23</td></tr>
<tr><td>27</td><td>N1*S1</td><td><0.0001</td><td>>0.9999</td><td>0.26</td></tr>
<tr><td>28</td><td>N1*S2</td><td>0.8086</td><td>0.1914</td><td><0.01</td></tr>
<tr><td>29</td><td>N1*RM</td><td><0.0001</td><td>>0.9999</td><td>0.09</td></tr>
<tr><td>30</td><td>N1*AF</td><td>0.2444</td><td>0.7556</td><td>0.02</td></tr>
<tr><td>31</td><td>N2*S1</td><td>0.3952</td><td>0.6048</td><td>0.01</td></tr>
<tr><td>32</td><td>N2*S2</td><td>0.9666</td><td>0.0334</td><td><0.01</td></tr>
<tr><td>33</td><td>N2*RM</td><td>0.0742</td><td>0.9258</td><td>0.03</td></tr>
<tr><td>34</td><td>N2*AF</td><td>0.0109</td><td>0.9891</td><td>0.04</td></tr>
<tr><td>35</td><td>S1*S2</td><td>0.0248</td><td>0.9752</td><td>0.03</td></tr>
<tr><td>36</td><td>S1*RM</td><td><0.0001</td><td>>0.9999</td><td>0.88</td></tr>
<tr><td>37</td><td>S1*AF</td><td>0.3179</td><td>0.6821</td><td>0.02</td></tr>
<tr><td>38</td><td>S2*RM</td><td>0.0248</td><td>0.9752</td><td>0.06</td></tr>
<tr><td>39</td><td>S2*AF</td><td>0.9763</td><td>0.0237</td><td><0.01</td></tr>
<tr><td>40</td><td>RM*AF</td><td>0.3229</td><td>0.6771</td><td>0.03</td></tr>
<tr><td>41</td><td>|N1*N2|</td><td><0.0001</td><td>>0.9999</td><td>0.24</td></tr>
<tr><td>42</td><td>|N1*S1|</td><td>0.0396</td><td>0.9604</td><td>0.06</td></tr>
<tr><td>43</td><td>|N1*S2|</td><td>0.0532</td><td>0.9468</td><td>0.05</td></tr>
<tr><td>44</td><td>|N1*RM|</td><td><0.0001</td><td>>0.9999</td><td>0.33</td></tr>
<tr><td>45</td><td>|N1*AF|</td><td><0.0001</td><td>>0.9999</td><td>0.11</td></tr>
<tr><td>46</td><td>|N2*S1|</td><td>0.6743</td><td>0.3257</td><td>0.01</td></tr>
<tr><td>47</td><td>|N2*S2|</td><td>0.1732</td><td>0.8268</td><td>0.04</td></tr>
<tr><td>48</td><td>|N2*RM|</td><td>0.0163</td><td>0.9837</td><td>0.07</td></tr>
<tr><td>49</td><td>|N2*AF|</td><td>0.0322</td><td>0.9678</td><td>0.06</td></tr>
<tr><td>50</td><td>|S1*S2|</td><td>0.0386</td><td>0.9614</td><td>0.07</td></tr>
<tr><td>51</td><td>|S1*RM|</td><td><0.0001</td><td>>0.9999</td><td>0.29</td></tr>
<tr><td>52</td><td>|S1*AF|</td><td>0.0430</td><td>0.9570</td><td>0.06</td></tr>
<tr><td>53</td><td>|S2*RM|</td><td>0.4077</td><td>0.5923</td><td>0.04</td></tr>
<tr><td>54</td><td>|S2*AF|</td><td>0.5938</td><td>0.4062</td><td>0.02</td></tr>
<tr><td>55</td><td>|RM*AF|</td><td>0.0440</td><td>0.9560</td><td>0.08</td></tr>
<tr><td>56</td><td>N1*|N2|</td><td><0.0001</td><td>>0.9999</td><td>0.37</td></tr>
<tr><td>57</td><td>N1*|S1|</td><td>0.1100</td><td>0.8900</td><td>0.03</td></tr>
<tr><td>58</td><td>N1*|S2|</td><td>0.0859</td><td>0.9141</td><td>0.04</td></tr>
<tr><td>59</td><td>N1*|RM|</td><td><0.0001</td><td>>0.9999</td><td>0.40</td></tr>
<tr><td>60</td><td>N1*|AF|</td><td>0.8757</td><td>0.1243</td><td><0.01</td></tr>
<tr><td>61</td><td>N2*|S1|</td><td>0.5701</td><td>0.4299</td><td>0.01</td></tr>
<tr><td>62</td><td>N2*|S2|</td><td>0.8998</td><td>0.1002</td><td><0.01</td></tr>
<tr><td>63</td><td>N2*|RM|</td><td><0.0001</td><td>>0.9999</td><td>0.16</td></tr>
<tr><td>64</td><td>N2*|AF|</td><td>0.9629</td><td>0.0371</td><td><0.01</td></tr>
<tr><td>65</td><td>S1*|S2|</td><td>0.1942</td><td>0.8058</td><td>0.03</td></tr>
<tr><td>66</td><td>S1*|RM|</td><td><0.0001</td><td>>0.9999</td><td>0.21</td></tr>
<tr><td>67</td><td>S1*|AF|</td><td>0.8736</td><td>0.1264</td><td><0.01</td></tr>
<tr><td>68</td><td>S2*|RM|</td><td>0.3284</td><td>0.6716</td><td>0.03</td></tr>
<tr><td>69</td><td>S2*|AF|</td><td>0.3991</td><td>0.6009</td><td>0.02</td></tr>
<tr><td>70</td><td>RM*|AF|</td><td>0.5333</td><td>0.4667</td><td>0.02</td></tr>
<tr><td>71</td><td>|N1|*N2</td><td><0.0001</td><td>>0.9999</td><td>0.26</td></tr>
<tr><td>72</td><td>|N1|*S1</td><td><0.0001</td><td>>0.9999</td><td>0.12</td></tr>
<tr><td>73</td><td>|N1|*S2</td><td>0.0915</td><td>0.9085</td><td>0.03</td></tr>
<tr><td>74</td><td>|N1|*RM</td><td>0.0752</td><td>0.9248</td><td>0.04</td></tr>
<tr><td>75</td><td>|N1|*AF</td><td>0.0016</td><td>0.9984</td><td>0.07</td></tr>
<tr><td>76</td><td>|N2|*S1</td><td>0.4164</td><td>0.5836</td><td>0.02</td></tr>
<tr><td>77</td><td>|N2|*S2</td><td>0.0011</td><td>0.9989</td><td>0.06</td></tr>
<tr><td>78</td><td>|N2|*RM</td><td>0.8703</td><td>0.1297</td><td><0.01</td></tr>
<tr><td>79</td><td>|N2|*AF</td><td>0.1606</td><td>0.8394</td><td>0.03</td></tr>
<tr><td>80</td><td>|S1|*S2</td><td>0.6754</td><td>0.3246</td><td>0.01</td></tr>
<tr><td>81</td><td>|S1|*RM</td><td>0.2967</td><td>0.7033</td><td>0.04</td></tr>
<tr><td>82</td><td>|S1|*AF</td><td>0.0472</td><td>0.9528</td><td>0.05</td></tr>
<tr><td>83</td><td>|S2|*RM</td><td>0.9573</td><td>0.0427</td><td><0.01</td></tr>
<tr><td>84</td><td>|S2|*AF</td><td>0.6288</td><td>0.3712</td><td>0.01</td></tr>
<tr><td>85</td><td>|RM|*AF</td><td>0.0293</td><td>0.9707</td><td>0.08</td></tr>
</table>

Fig. 2a Comparison of *p*–value and percent contribution for regression model of forward normal gage output $R1$.

INDEX	MATH TERM	p	1 - p	ABS(Percent Cont.)
		[-]	[-]	[%]
1	INTERCEPT	-	-	0.04
2	**N1**	**<0.0001**	**>0.9999**	**[100.00]**
3	N2	<0.0001	>0.9999	3.10
4	S1	<0.0001	>0.9999	0.21
5	S2	0.0009	0.9991	0.07
6	RM	<0.0001	>0.9999	0.52
7	AF	0.7814	0.2186	0.01
8	|N1|	<0.0001	>0.9999	0.98
9	|N2|	<0.0001	>0.9999	0.64
10	|S1|	0.0007	0.9993	0.11
11	|S2|	<0.0001	>0.9999	0.12
12	|RM|	<0.0001	>0.9999	0.84
13	|AF|	0.9995	0.0005	<0.01
14	N1*N1	<0.0001	>0.9999	0.14
15	N2*N2	0.0008	0.9992	0.10
16	S1*S1	0.0185	0.9815	0.08
17	S2*S2	0.1989	0.8011	0.04
18	RM*RM	<0.0001	>0.9999	0.27
19	AF*AF	0.2502	0.7498	0.03
20	N1*|N1|	<0.0001	>0.9999	0.90
21	N2*|N2|	<0.0001	>0.9999	0.80
22	S1*|S1|	0.4294	0.5706	0.02
23	S2*|S2|	0.0791	0.9209	0.05
24	RM*|RM|	0.0051	0.9949	0.07
25	AF*|AF|	0.6924	0.3076	0.01
26	N1*N2	<0.0001	>0.9999	0.23
27	N1*S1	<0.0001	>0.9999	0.26
28	N1*S2	0.8086	0.1914	<0.01
29	N1*RM	<0.0001	>0.9999	0.09
30	N1*AF	0.2444	0.7556	0.02
31	N2*S1	0.3952	0.6048	0.01
32	N2*S2	0.9666	0.0334	<0.01
33	N2*RM	0.0742	0.9258	0.03
34	N2*AF	0.0109	0.9891	0.04
35	S1*S2	0.0248	0.9752	0.03
36	S1*RM	<0.0001	>0.9999	0.88
37	S1*AF	0.3179	0.6821	0.02
38	S2*RM	0.0248	0.9752	0.06
39	S2*AF	0.9763	0.0237	<0.01
40	RM*AF	0.3229	0.6771	0.03
41	|N1*N2|	<0.0001	>0.9999	0.24
42	|N1*S1|	0.0396	0.9604	0.06
43	|N1*S2|	0.0532	0.9468	0.05
44	|N1*RM|	<0.0001	>0.9999	0.33
45	|N1*AF|	<0.0001	>0.9999	0.11
46	|N2*S1|	0.6743	0.3257	0.01
47	|N2*S2|	0.1732	0.8268	0.04
48	|N2*RM|	0.0163	0.9837	0.07
49	|N2*AF|	0.0322	0.9678	0.06
50	|S1*S2|	0.0386	0.9614	0.07
51	|S1*RM|	<0.0001	>0.9999	0.29
52	|S1*AF|	0.0430	0.9570	0.06
53	|S2*RM|	0.4077	0.5923	0.04
54	|S2*AF|	0.5938	0.4062	0.02
55	|RM*AF|	0.0440	0.9560	0.08
56	N1*|N2|	<0.0001	>0.9999	0.37
57	N1*|S1|	0.1100	0.8900	0.03
58	N1*|S2|	0.0859	0.9141	0.04
59	N1*|RM|	<0.0001	>0.9999	0.40
60	N1*|AF|	0.8757	0.1243	<0.01
61	N2*|S1|	0.5701	0.4299	0.01
62	N2*|S2|	0.8998	0.1002	<0.01
63	N2*|RM|	<0.0001	>0.9999	0.16
64	N2*|AF|	0.9629	0.0371	<0.01
65	S1*|S2|	0.1942	0.8058	0.03
66	S1*|RM|	<0.0001	>0.9999	0.21
67	S1*|AF|	0.8736	0.1264	<0.01
68	S2*|RM|	0.3284	0.6716	0.03
69	S2*|AF|	0.3991	0.6009	0.02
70	RM*|AF|	0.5333	0.4667	0.02
71	|N1|*N2	<0.0001	>0.9999	0.26
72	|N1|*S1	<0.0001	>0.9999	0.12
73	|N1|*S2	0.0915	0.9085	0.03
74	|N1|*RM	0.0752	0.9248	0.04
75	|N1|*AF	0.0016	0.9984	0.07
76	|N2|*S1	0.4164	0.5836	0.02
77	|N2|*S2	0.0011	0.9989	0.06
78	|N2|*RM	0.8703	0.1297	<0.01
79	|N2|*AF	0.1606	0.8394	0.03
80	|S1|*S2	0.6754	0.3246	0.01
81	|S1|*RM	0.2967	0.7033	0.04
82	|S1|*AF	0.0472	0.9528	0.05
83	|S2|*RM	0.9573	0.0427	<0.01
84	|S2|*AF	0.6288	0.3712	0.01
85	|RM|*AF	0.0293	0.9707	0.08

Fig. 2b Comparison of significance of cross–product terms $N1\ N2$, $|N1\ N2|$, $N1\ |N2|$, and $|N1|\ N2$.

INDEX	MATH TERM	p	1 - p	ABS(Percent Cont.)
		[-]	[-]	[%]
1	INTERCEPT	-	-	0.12
2	N1	<0.0001	>0.9999	4.36
3	N2	<0.0001	>0.9999	5.13
4	S1	0.5555	0.4445	0.02
5	S2	<0.0001	>0.9999	0.28
6	RM	<0.0001	>0.9999	3.14
7	**AF**	**<0.0001**	**>0.9999**	**[100.00]**
8	|N1|	0.4637	0.5363	0.03
9	|N2|	<0.0001	>0.9999	0.77
10	|S1|	0.1105	0.8895	0.08
11	|S2|	<0.0001	>0.9999	0.26
12	|RM|	<0.0001	>0.9999	0.93
13	|AF|	0.0741	0.9259	0.08
14	N1*N1	<0.0001	>0.9999	0.43
15	N2*N2	<0.0001	>0.9999	0.40
16	S1*S1	0.4808	0.5192	0.04
17	S2*S2	0.4806	0.5194	0.04
18	RM*RM	0.0074	0.9926	0.13
19	AF*AF	0.9234	0.0766	4.59E-03
20	N1*|N1|	<0.0001	>0.9999	0.42
21	N2*|N2|	<0.0001	>0.9999	0.99
22	S1*|S1|	0.0089	0.9911	0.12
23	S2*|S2|	0.1153	0.8847	0.07
24	RM*|RM|	0.4797	0.5203	0.03
25	AF*|AF|	0.3186	0.6814	0.05
26	N1*N2	<0.0001	>0.9999	0.52
27	N1*S1	<0.0001	>0.9999	0.27
28	N1*S2	<0.0001	>0.9999	0.15
29	N1*RM	<0.0001	>0.9999	0.95
30	N1*AF	<0.0001	>0.9999	0.16
31	N2*S1	0.5125	0.4875	0.02
32	N2*S2	<0.0001	>0.9999	0.18
33	N2*RM	<0.0001	>0.9999	0.19
34	N2*AF	<0.0001	>0.9999	0.30
35	S1*S2	0.2897	0.7103	0.02
36	S1*RM	<0.0001	>0.9999	0.21
37	S1*AF	0.0146	0.9854	0.07
38	S2*RM	0.0026	0.9974	0.14
39	S2*AF	<0.0001	>0.9999	0.20
40	RM*AF	0.0332	0.9668	0.11
41	|N1*N2|	0.0908	0.9092	0.05
42	|N1*S1|	0.9320	0.0680	**4.06E-03**
43	|N1*S2|	0.1334	0.8666	0.07
44	|N1*RM|	<0.0001	>0.9999	0.62
45	|N1*AF|	0.0485	0.9515	0.09
46	|N2*S1|	0.1293	0.8707	0.07
47	|N2*S2|	0.2923	0.7077	0.05
48	|N2*RM|	0.0905	0.9095	0.08
49	|N2*AF|	0.0074	0.9926	0.12
50	|S1*S2|	<0.0001	>0.9999	0.29
51	|S1*RM|	0.6810	0.3190	0.03
52	|S1*AF|	0.3065	0.6935	0.05
53	|S2*RM|	0.3715	0.6285	0.07
54	|S2*AF|	0.0801	0.9199	0.09
55	|RM*AF|	0.0617	0.9383	0.12
56	N1*|N2|	<0.0001	>0.9999	0.32
57	N1*|S1|	0.0018	0.9982	0.11
58	N1*|S2|	0.1587	0.8413	0.05
59	N1*|RM|	<0.0001	>0.9999	0.83
60	N1*|AF|	0.0237	0.9763	0.08
61	N2*|S1|	0.0588	0.9412	0.07
62	N2*|S2|	0.0033	0.9967	0.10
63	N2*|RM|	0.2958	0.7042	0.04
64	N2*|AF|	0.0841	0.9159	0.06
65	S1*|S2|	0.2751	0.7249	0.04
66	S1*|RM|	<0.0001	>0.9999	0.19
67	S1*|AF|	0.7531	0.2469	0.01
68	S2*|RM|	0.0003	0.9997	0.19
69	S2*|AF|	0.2881	0.7119	0.04
70	RM*|AF|	0.6088	0.3912	0.03
71	|N1|*N2	<0.0001	>0.9999	0.90
72	|N1|*S1	<0.0001	>0.9999	0.16
73	|N1|*S2	0.4139	0.5861	0.03
74	|N1|*RM	0.0061	0.9939	0.11
75	|N1|*AF	<0.0001	>0.9999	0.28
76	|N2|*S1	0.2460	0.7540	0.04
77	|N2|*S2	0.0859	0.9141	0.06
78	|N2|*RM	0.0023	0.9977	0.12
79	|N2|*AF	0.0010	0.9990	0.13
80	|S1|*S2	0.0046	0.9954	0.11
81	|S1|*RM	0.2002	0.7998	0.08
82	|S1|*AF	0.1173	0.8827	0.07
83	|S2|*RM	0.7475	0.2525	0.02
84	|S2|*AF	0.7917	0.2083	0.01
85	|RM|*AF	0.0236	0.9764	0.13

Fig. 3a Comparison of p–value and percent contribution for regression model of axial gage output $R6$.

INDEX	MATH TERM	p	1 - p	ABS(Percent Cont.)
		[-]	[-]	[%]
1	INTERCEPT	-	-	0.12
2	N1	<0.0001	>0.9999	4.36
3	N2	<0.0001	>0.9999	5.13
4	S1	0.5555	0.4445	0.02
5	S2	<0.0001	>0.9999	0.28
6	RM	<0.0001	>0.9999	3.14
7	**AF**	**<0.0001**	**>0.9999**	**[100.00]**
8	|N1|	0.4637	0.5363	0.03
9	|N2|	<0.0001	>0.9999	0.77
10	|S1|	0.1105	0.8895	0.08
11	|S2|	<0.0001	>0.9999	0.26
12	|RM|	<0.0001	>0.9999	0.93
13	|AF|	0.0741	0.9259	0.08
14	N1*N1	<0.0001	>0.9999	0.43
15	N2*N2	<0.0001	>0.9999	0.40
16	S1*S1	0.4808	0.5192	0.04
17	S2*S2	0.4806	0.5194	0.04
18	RM*RM	0.0074	0.9926	0.13
19	AF*AF	0.9234	0.0766	4.59E-03
20	N1*|N1|	<0.0001	>0.9999	0.42
21	N2*|N2|	<0.0001	>0.9999	0.99
22	S1*|S1|	0.0089	0.9911	0.12
23	S2*|S2|	0.1153	0.8847	0.07
24	RM*|RM|	0.4797	0.5203	0.03
25	AF*|AF|	0.3186	0.6814	0.05
26	N1*N2	<0.0001	>0.9999	0.52
27	N1*S1	<0.0001	>0.9999	0.27
28	N1*S2	<0.0001	>0.9999	0.15
29	N1*RM	<0.0001	>0.9999	0.95
30	N1*AF	<0.0001	>0.9999	0.16
31	N2*S1	0.5125	0.4875	0.02
32	N2*S2	<0.0001	>0.9999	0.18
33	N2*RM	<0.0001	>0.9999	0.19
34	N2*AF	<0.0001	>0.9999	0.30
35	S1*S2	0.2897	0.7103	0.02
36	S1*RM	<0.0001	>0.9999	0.21
37	S1*AF	0.0146	0.9854	0.07
38	S2*RM	0.0026	0.9974	0.14
39	S2*AF	<0.0001	>0.9999	0.20
40	RM*AF	0.0332	0.9668	0.11
41	|N1*N2|	0.0908	0.9092	0.05
42	|N1*S1|	0.9320	0.0680	**4.06E-03**
43	|N1*S2|	0.1334	0.8666	0.07
44	|N1*RM|	<0.0001	>0.9999	0.62
45	|N1*AF|	0.0485	0.9515	0.09
46	|N2*S1|	0.1293	0.8707	0.07
47	|N2*S2|	0.2923	0.7077	0.05
48	|N2*RM|	0.0905	0.9095	0.08
49	|N2*AF|	0.0074	0.9926	0.12
50	|S1*S2|	<0.0001	>0.9999	0.29
51	|S1*RM|	0.6810	0.3190	0.03
52	|S1*AF|	0.3065	0.6935	0.05
53	|S2*RM|	0.3715	0.6285	0.07
54	|S2*AF|	0.0801	0.9199	0.09
55	|RM*AF|	0.0617	0.9383	0.12
56	N1*|N2|	<0.0001	>0.9999	0.32
57	N1*|S1|	0.0018	0.9982	0.11
58	N1*|S2|	0.1587	0.8413	0.05
59	N1*|RM|	<0.0001	>0.9999	0.83
60	N1*|AF|	0.0237	0.9763	0.08
61	N2*|S1|	0.0588	0.9412	0.07
62	N2*|S2|	0.0033	0.9967	0.10
63	N2*|RM|	0.2958	0.7042	0.04
64	N2*|AF|	0.0841	0.9159	0.06
65	S1*|S2|	0.2751	0.7249	0.04
66	S1*|RM|	<0.0001	>0.9999	0.19
67	S1*|AF|	0.7531	0.2469	0.01
68	S2*|RM|	0.0003	0.9997	0.19
69	S2*|AF|	0.2881	0.7119	0.04
70	RM*|AF|	0.6088	0.3912	0.03
71	|N1|*N2	<0.0001	>0.9999	0.90
72	|N1|*S1	<0.0001	>0.9999	0.16
73	|N1|*S2	0.4139	0.5861	0.03
74	|N1|*RM	0.0061	0.9939	0.11
75	|N1|*AF	<0.0001	>0.9999	0.28
76	|N2|*S1	0.2460	0.7540	0.04
77	|N2|*S2	0.0859	0.9141	0.06
78	|N2|*RM	0.0023	0.9977	0.12
79	|N2|*AF	0.0010	0.9990	0.13
80	|S1|*S2	0.0046	0.9954	0.11
81	|S1|*RM	0.2002	0.7998	0.08
82	|S1|*AF	0.1173	0.8827	0.07
83	|S2|*RM	0.7475	0.2525	0.02
84	|S2|*AF	0.7917	0.2083	0.01
85	|RM|*AF	0.0236	0.9764	0.13

Fig. 3b Comparison of significance of cross–product terms $N2\ AF$, $|N2\ AF|$, $N2\ |AF|$, and $|N2|\ AF$.

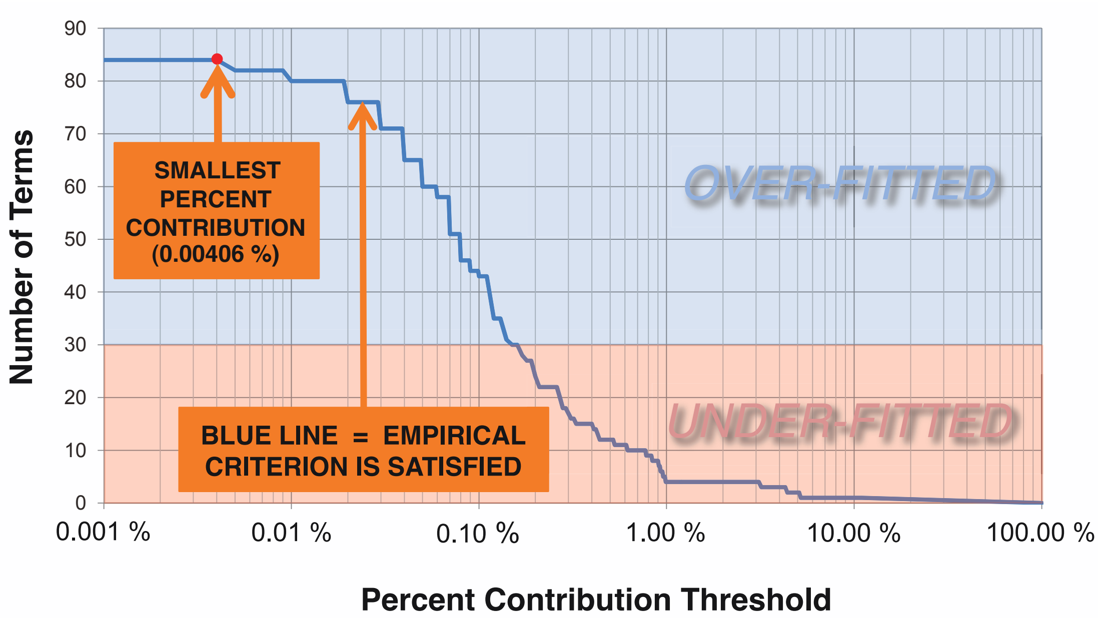


Fig. 4a Number of terms of the math model of the axial gage outputs that satisfy only the empirical criterion.

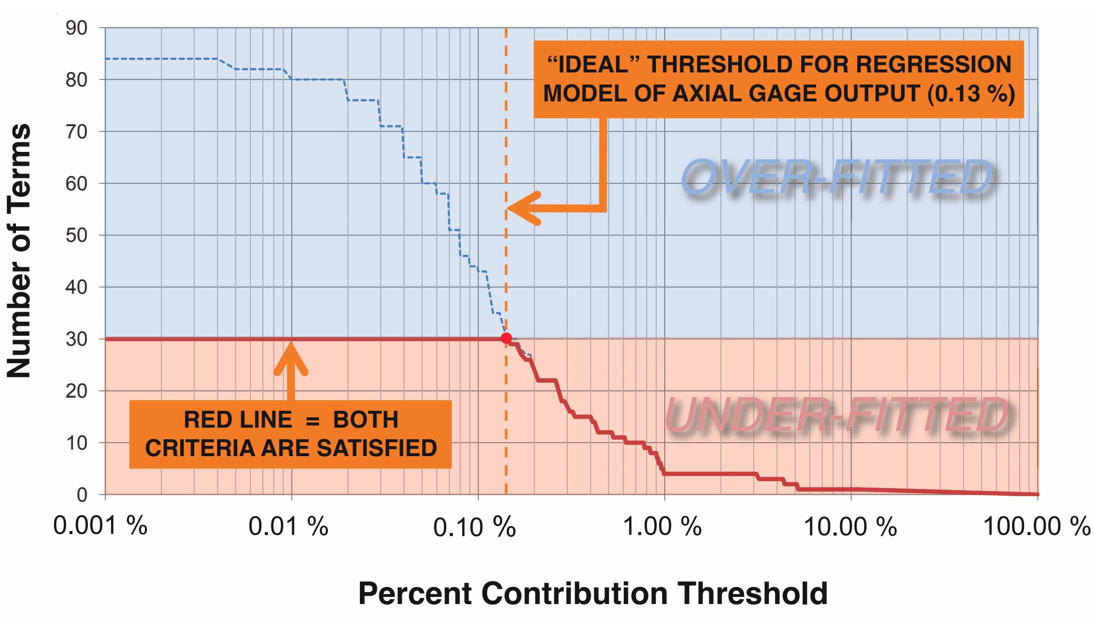


Fig. 4b Number of terms of the math model of the axial gage outputs that satisfy both criteria.

BALANCE NUMBER	BALANCE TYPE	TARE LOAD ITERATION	TESTED MATH TERM GROUP COMBINATION	SMALLEST PERCENT CONTRIBUTION THAT MET STATISTICAL CRITERION: *p–value of t–statistic* $<$ 0.0001
1	MK29B	NO	**(III)**†	**0.08 %**
–	(FORCE)	–	(IV)	0.14 %
–	–	–	(V)	0.09 %
2	FORCE	NO	**(III)**†	**0.06 %**
–	–	–	(IV)	0.10 %
–	–	–	(V)	0.08 %
3	FORCE	NO	**(III)**†	**0.03 %**
–	–	–	(IV)	0.06 %
–	–	–	(V)	0.02 %
4	FORCE	YES	**(III)**†	**0.02 %**
–	–	–	(IV)	0.09 %
–	–	–	(V)	0.09 %
5	FORCE	YES	**(III)**†	**0.07 %**
–	–	–	(IV)	0.12 %
–	–	–	(V)	0.21 %

†Recommended math term group combination, i.e., the smallest group combination that met accuracy expectations (a definition of all group combinations can be found in Table 3).

Fig. 5a Observed percent contribution minima for multi–piece balances.

BALANCE NUMBER	BALANCE TYPE	TARE LOAD ITERATION	TESTED MATH TERM GROUP COMBINATION	SMALLEST PERCENT CONTRIBUTION THAT MET STATISTICAL CRITERION: *p–value of t–statistic* $<$ 0.0001
6	FORCE	NO	**(II)**†	**0.03 %**
–	–	–	(III)	0.03 %
–	–	–	(IV)	0.05 %
–	–	–	(V)	0.04 %
7	DIRECT–READ	YES	**(II)**†	**0.04 %**
–	–	–	(III)	0.07 %
–	–	–	(IV)	0.06 %
–	–	–	(V)	0.06 %
8	DIRECT–READ	NO	**(II)**†	**0.01 %**
–	–	–	(III)	0.02 %
–	–	–	(IV)	0.03 %
–	–	–	(V)	0.08 %
9	MOMENT	YES	**(II)**†	**0.02 %**
–	–	–	(III)	0.02 %
–	–	–	(IV)	0.02 %
–	–	–	(V)	0.02 %

†Recommended math term group combination, i.e., the smallest group combination that met accuracy expectations (a definition of all group combinations can be found in Table 3).

Fig. 5b Observed percent contribution minima for single–piece & hybrid balances.

BALANCE NUMBER	BALANCE TYPE	TARE LOAD ITERATION	TESTED MATH TERM GROUP COMBINATION	SMALLEST PERCENT CONTRIBUTION THAT MET STATISTICAL CRITERION: *p–value of t–statistic* $<$ 0.0001
10	DIRECT–READ	YES	**(I)**†	**0.05 %**
–	–	–	(II)	0.02 %
–	–	–	(III)	0.04 %
11	DIRECT–READ	YES	**(I)**†	**0.03 %**
–	–	–	(II)	0.19 %
–	–	–	(III)	0.09 %
12	DIRECT–READ	YES	**(I)**†	**0.02 %**
–	–	–	(II)	0.04 %
–	–	–	(III)	0.04 %

†Recommended math term group combination, i.e., the smallest group combination that met accuracy expectations (a definition of all group combinations can be found in Table 3).

Fig. 5c Observed percent contribution minima for semispan balances.