

Regression Model Term Selection for the Analysis of Strain–Gage Balance Calibration Data

N. Ulbrich* and T. Volden**

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

The paper discusses the selection of regression model terms for the analysis of wind tunnel strain–gage balance calibration data. Different function class combinations are presented that may be used to analyze calibration data using either a non–iterative or an iterative method. The role of the intercept term in a regression model of calibration data is reviewed. In addition, useful algorithms and metrics originating from linear algebra and statistics are recommended that will help an analyst (i) to identify and avoid both linear and near–linear dependencies between regression model terms and (ii) to make sure that the selected regression model of the calibration data uses only statistically significant terms. Three different tests are suggested that may be used to objectively assess the predictive capability of the final regression model of the calibration data. These tests use both the original data points and regression model independent confirmation points. Finally, data from a simplified manual calibration of the Ames MK40 balance is used to illustrate the application of some of the metrics and tests to a realistic calibration data set.

Nomenclature

A, B	= independent variables
NF	= total normal force acting on a strain–gage balance, [lbs]
$N1$	= normal force component at forward gage, [lbs]
$N2$	= normal force component at aft gage, [lbs]
$R1, R2, \dots, R6$	= electrical outputs of the strain–gages of the balance, [microV/V]
ΔNF	= normal force residual, i.e., the difference between the measured and fitted value

I. Introduction

New regression model term selection criteria have been proposed for the analysis of wind tunnel strain–gage balance calibration data since the publication of the AIAA Recommended Practice entitled “Calibration and Use of Internal Strain–Gage Balances” in 2003 (see Ref. [1], pp. 7–10, for a description of the previously applied term selection criteria). A numerical technique called Singular Value Decomposition (SVD), for example, was introduced in balance calibration data analysis in 2005 that helps screen a regression model of calibration data for unwanted linear dependencies (for more detail see Ref. [2]). In addition, more advanced metrics have been suggested for the evaluation of a regression model of balance calibration data. These metrics have been used in statistics or response surface methodology since the 1970s. They test the statistical significance of individual terms of a regression model or screen a regression model for near–linear dependencies (see Ref. [3], p. 84, pp. 323–341 for a description of the new criteria).

The new term selection criteria were successfully combined with traditional strain–gage balance calibration data analysis methods in a regression model optimization process that was developed to analyze calibration data faster and more efficiently at NASA Ames’ Balance Calibration Laboratory (see Refs. [4], [5], [6], [7], and [8]). The new term selection criteria can, of course, also be applied if an analyst has to choose terms of a regression model by inspection. Therefore, the most important steps of the regression model term selection process are discussed in the present paper in order to introduce a wider group of users within the

* Aerodynamicist; Jacobs Technology Inc.

** Computer Engineer; Jacobs Technology Inc.

wind tunnel testing community to the new set of metrics.

II. Math Term Selection Process

A. General Remarks

Different steps can be identified that help an analyst to build a regression model of balance calibration data by inspection. Figure 1 summarizes the most important steps that will be discussed in the paper. Initially, the analyst has to decide if the calibration data is to be analyzed using a *non-iterative* or an *iterative* method. This decision determines if, for example, the balance loads or the gage outputs are the independent variables of the global regression analysis problem. Then, a suitable combination of function classes (math term groups) has to be chosen that will best represent the calibration data. The analyst must also decide whether to include an intercept (constant) term in the regression model. Afterwards, using the selected function class combination as a starting point, SVD should be applied to the vector space defined by the regressors in order to remove linearly dependent regressors that, if included, would lead to a singular solution of the global regression analysis problem. In the next step, the regression model should be screened for near-linear dependencies between terms using the variance inflation factor. In addition, the p -value of the t -statistic of the terms should be compared with a suitable threshold so that statistically insignificant terms can be removed from the regression model. The analyst should also determine if subject matter knowledge about (i) the balance or (ii) the calibration data supports the use of a “hierarchical” regression model. In that case, missing lower order terms should be added to make the regression model “hierarchical.” Finally, three different tests should be applied to the regression model at the end of the term selection process. These tests compare (i) the standard deviation of the residuals of the dependent variables, (ii) the standard deviation of the PRESS residuals of the dependent variables (see Ref. [3], pp. 125–126 for a definition of PRESS residuals), and (iii) the standard deviation of the confirmation point residuals of the dependent variables with accuracy expectations for the wind tunnel test. In the next sections of the paper the different steps are discussed in more detail.

B. Regression Analysis Approach

During a calibration experiment balance loads are “applied” and gage outputs are “measured.” Therefore, an analyst would naturally pick the loads to be the *independent* variables and the gage outputs to be the *dependent* variables. The resulting regression model would predict the gage outputs as a function of the loads. Unfortunately, the reverse operation is needed during a wind tunnel test: loads need to be predicted as a function of gage outputs. Therefore, the wind tunnel testing community has historically developed two types of data analysis methods in order to make it possible to predict loads as a function of gage outputs. They are called *non-iterative* and *iterative* methods.

The analyst needs to decide which one of the two methods is to be used for the regression analysis of the balance calibration data. This choice determines if the gage outputs or the balance loads are the *independent* or the *dependent* variables of the global regression analysis problem. Table 1 below compares the two different analysis options with other variable naming conventions that are used in mathematics, science, engineering, and response surface methodology.

Table 1: Definition of Independent and Dependent Variables.

MATHEMATICS, SCIENCE, & ENGINEERING	RESPONSE SURFACE METHODOLOGY	<u>NON-ITERATIVE</u> ANALYSIS OF BALANCE CALIBRATION	<u>ITERATIVE</u> ANALYSIS OF BALANCE CALIBRATION
INDEPENDENT VARIABLE	FACTOR	GAGE OUTPUT	BALANCE LOAD
DEPENDENT VARIABLE	RESPONSE	BALANCE LOAD	GAGE OUTPUT

A *non-iterative* method is a simpler method. This one-step approach directly fits balance loads as a function of gage outputs. It ignores the fact that the balance loads are the *independent* variables of the calibration experiment and simply exchanges *independent* with *dependent* variables.

An *iterative* method, on the other hand, is a two-step approach. First, the measured gage outputs are fitted as a function of the applied calibration loads. Then, the resulting regression coefficients are transformed to so-called data reduction matrix coefficients. They are used in an iteration scheme so that loads can be

predicted from gage outputs during a wind tunnel test (see, e.g., Ref. [1] or Ref. [9] for a description of different *iterative* methods).

C. Math Term Group or Function Class Combination Selection

Different combinations of function classes are used in the wind tunnel testing community in order to model the physical behavior of a strain-gage balance in the least squares sense. The final function class combination choice depends primarily on the design of the balance and the expected characteristics of the calibration data. It is assumed, as an example, that an *iterative* method is chosen for the analysis of some balance calibration data. Then, the six function class combination choices shown in Fig. 2 are available. Combination I is commonly used if the calibrated balance is a 6-component single-piece balance. Combinations II to VI are often used if the balance is a multi-piece design (e.g., Task balance). In principle, the same function class combinations may be applied if a *non-iterative* method is applied to the calibration data.

Now a question comes up. Should an intercept term (constant) be included in the regression model of the data? In general, the authors recommend to include an intercept in the regression model as the presence of the intercept will ensure that unexpected shifts in the calibration data are not erroneously charged to coefficients that scale with the independent variables. Often a tare load iteration scheme is used during the application of an *iterative* method in order to better assess the influence of the weight of the balance shell, calibration body, and calibration hardware on the gage outputs. This tare load iteration may be performed using the raw gage outputs in combination with the natural zeros of the balance as a gage output datum. In that case, the intercept must be included in the regression model of the gage outputs as the intercept will be the least squares approximation of the natural zeros.

D. Linear Dependency Test (Singular Value Decomposition)

Regressors have to be computed as a function of the independent variables after suitable function classes were selected for the analysis. However, some of regressors may not be supported by the calibration data. They need to be identified and removed from the model so that a singular solution of the global regression analysis problem is avoided.

In general, it is recommended to apply a numerical process called Singular Value Decomposition (SVD) to the combination of regressors of the regression model in order to identify and eliminate linear dependencies between the terms. The application of SVD ultimately helps define the largest regression model for a given calibration data set and function class combination that will lead to a unique solution of the regression analysis problem. References [2] and [7] provide more information about SVD, its implementation, and its benefits.

E. Near-Linear Dependency Test (Variance Inflation Factors)

At the end of 2006 *DeLoach* pointed out to *Ulbrich* that “near-linear” dependencies (also called “collinearity” or “multicollinearity”) between terms of a regression model have to be avoided at all cost as they could diminish the predictive capabilities of a regression model (see Ref. [10] for more detail). In addition, *Ulbrich* observed a direct connection between the divergence of the iteration equation that an *iterative* method uses for the regression analysis of balance calibration data and the presence of massive “near-linear” dependencies in a regression model (see Ref. [5], p. 4). Therefore, it is recommended to identify and remove terms in a regression model that cause “near-linear” dependencies.

Different techniques are given in the literature that help to diagnose and avoid “near-linear” dependencies in a regression model (see Ref. [3], pp. 334–340). The authors suggest to use the variance inflation factor (VIF) for this purpose (see Ref. [4] for a detailed explanation of steps that are needed to compute the VIFs of a regression model). The VIF is computed for each individual term of a regression model. “Near-linear” dependencies in a math model are negligible if the largest VIF of all math model terms of a tested regression model is smaller than the threshold of 5 or 10. A decrease of the VIF threshold from, e.g., 10 to 5, tightens the constraint that is applied. This tightening results in a smaller and more robust regression model of the data.

F. Test of Statistical Significance of Terms (p -Value of t -Statistic)

This test allows the analyst to identify and remove insignificant terms of the regression model that may lead to overfitting of the calibration data. The test of the statistical significance of individual coefficients of

a regression model looks at the standard error of each regression coefficient of a math model. The standard error is an estimate of the standard deviation of the coefficient. It is 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.

Traditionally, the t -statistic of a coefficient is used in order to quantitatively compare a regression coefficient with its standard error. The t -statistic equals the ratio between the coefficient value and its standard error (for more detail see, e.g., Ref. [3], p. 84). A coefficient is probably “significant” if its t -statistic is greater than the critical value of a Student’s t -distribution. This comparison can also be performed 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., 0.01 (or 1 %) one can say with a 99 % probability of being correct that the regression coefficient is having some effect. The threshold for the p -value may range from a conservative value of 0.0001 to a liberal value of 0.01. A decrease of the p -value threshold, e.g., from 0.001 to 0.0001, tightens the term rejection criterion. Therefore, the final regression model will have fewer terms and is expected to have smaller extrapolation errors if it is applied close to but outside of the boundaries of the fitted data set.

G. Hierarchy Rule

Some analysts prefer to use a “hierarchical” math model, i.e., a regression model that does not have any missing lower order terms, for the analysis of balance calibration data. These “hierarchical” math models have the ability to correctly model experimental data whenever the independent variables are known to have a constant shift (see also the discussion of math model “hierarchy” in Ref. [4], pp. 6–8). Often, however, it is unclear if a given experimental data set needs to be modeled using a “hierarchical” math model. An analytic first principles description of the experimental data may not be known that could be used as the basis for the regression model term selection. Then, it is left to the analyst to decide if (i) the math model should be made “hierarchical” to prevent a suspected problem with a “non-hierarchical” model or if (ii) the analysis is performed using a “non-hierarchical” model (see, e.g., Refs. [4] and [11] for a discussion of different aspects of “hierarchical” regression models).

H. Predictive Capability Tests

The predictive capability of the final regression model still needs to be tested after the completion of the term selection process. These tests make sure that the balance calibration data and its regression model meet expected accuracy requirements. The predictive capability may be tested using two types of points: data points that are used to compute the regression model coefficients and confirmation points that are independent of the regression model (see Ref. [8] for a discussion of the two types of points).

The data points and the confirmation points may be used to compute three types of residuals of the dependent variables. The first residual type is defined as the difference between the measured and fitted values at the data points. This residual is widely used in the wind tunnel testing community for the assessment of the predictive capability of a regression model. The second residual type is the so-called PRESS residual of the data points. PRESS residuals are very useful as they are recommended in the literature for the comparison of the predictive capability of different regression models (see, e.g., the discussion of PRESS residuals in Ref. [3], pp. 125–126, pp. 141–142). The third residual type is defined by the regression model independent confirmation points. These residuals are computed as the difference between the corresponding measured and fitted value at a confirmation point.

The three residual types may be used to compute three test metrics: (i) the standard deviation of the residuals, (ii) the standard deviation of the PRESS residuals, and (iii) the standard deviation of the confirmation point residuals. These three standard deviations can easily be compared with the accuracy requirements of the balance load predictions during a wind tunnel test in order to decide if the calibration data and the regression model meet expectations. The authors believe that the standard deviation of the confirmation point residuals is the most important test metric as the regression model’s performance is explicitly tested at points that were not used to compute the regression model coefficients.

Sometimes, it may happen that certain accuracy requirements of the balance load predictions cannot be met. This observation could have several explanations: (i) a suboptimal function class combination was selected for the regression analysis of the balance calibration data that could not model important characteristics of the calibration data; (ii) a suboptimal load schedule was used during the calibration of the balance that omitted important loads or load combinations that are needed for a proper characterization of

the physical behavior of the balance; (iii) large load and/or gage output measurement errors occurred that were not detected during the calibration of the balance; (iv) balance design characteristics (like, e.g., bolted joints) cause hysteresis effects that cannot easily be modeled using a global regression analysis approach. A detailed examination of the regression model, the load schedule design, measurement error sources, and the balance design may be needed in order to improve the analysis results and/or the calibration data so that the desired accuracy of the load predictions can be met.

III. Application of Term Selection Process to a Data Example

Data from a preliminary and simplified manual calibration of the Ames MK40 Task balance is used in this section in order to illustrate several steps of the regression model term selection process. The MK40 balance is a force type balance that has six gage outputs. For the discussion of the example the balance loads were transformed from the original 5-force-1-moment format to a 3-force-3-moment format. The calibration data set consisted of a total of 164 points that were taken in 16 load series. Tare load corrections were applied to the calibration data prior to the term selection and the regression analysis so that a tare load iteration could be avoided. In addition, gage outputs were given as differences relative to the natural zeros of the balance.

For simplicity, a *non-iterative* method was selected for the regression analysis of the balance calibration data. It was also decided to only perform a regression analysis of the normal force component (NF). The six gage outputs were considered to be the independent variables of the regression analysis problem. The normal force component was considered to be the dependent variable. Figure 3a shows the six gage outputs that were used for the regression analysis. Figure 3b shows the corresponding tare corrected normal force that was applied at each data point.

In the next step, a suitable function class combination had to be selected for the analysis. Combination II in Fig. 2 was chosen for this purpose. Then, SVD was applied to the regressors that are defined by the function class selection in order to remove unsupported terms from the regression model. Figure 4a shows the upper bound of the supported regression models that was obtained after the application of SVD. The model consists of a total of 23 terms. Figure 4b shows the final regression model that was obtained after the application of the proposed regression model term selection process. This model has only a total of 7 terms and does not use any square or combined terms.

It is interesting to compare some of the statistical metrics of the two regression models that are depicted in Figs. 4a and 4b. First, the 23 term regression model given in Fig. 4a is investigated. Figure 5a shows the analysis of variance results, the p -value of the t -statistic, and the VIFs of the regression model terms. Figure 5b shows the confidence intervals of the regression model coefficient values. Most of the confidence intervals of the coefficients experience a sign change (blue color) which indicates the presence of many insignificant terms in the regression model.

We notice immediately in Fig. 5a that the combined terms $R1 \cdot R5$ and $R2 \cdot R5$ appear to be the cause for massive near-linear dependencies in the regression model as their VIFs are significantly greater than the recommended threshold of 10. This observation can be explained if we plot the corresponding gage output combinations side-by-side. Figure 5c shows the result of this comparison. The regressor values of the two combined terms are essentially identical. Therefore, they should not be used simultaneously in a regression model.

After looking at the t -statistic of the coefficients in Fig. 5a we also see that the terms $R1$ and $R2$ appear to dominate the regression model of the normal force ($R1$ has a t -statistic of 1016 and $R2$ has a t -statistic of 1335; the magnitude of the t -statistic of all other terms is significantly smaller). This observation can be explained by the fact that the balance is a Task balance. Therefore, by design, the gage output $R1$ is proportional to the $N1$ load component at the forward gage and the gage output $R2$ is proportional to the $N2$ load component at the aft gage. Consequently, as the normal force NF is the sum of the two normal force components $N1$ and $N2$, the behavior of the normal force NF may be described by using a linear combination of the two gage outputs $R1$ and $R2$. This hypothesis can easily be tested if, for example, the sum of the two gage outputs $R1$ and $R2$ is determined. Figure 5d compares this sum with the tare corrected normal force NF . As expected, the sum of the gage outputs is almost proportional to the total normal force which explains the dominating influence of $R1$ and $R2$ in the regression model of the normal force NF . The actual coefficient values for $R1$ and $R2$ are listed in Fig. 5a. They are given as 3.2319 and 2.7362 which

confirms that NF is primarily proportional to the sum of $R1$ and $R2$.

Figure 6a shows the analysis of variance results for the 7 term regression model that is depicted in Fig. 4b. All statistical metrics fall into recommended ranges. In addition, the confidence intervals of the regression coefficients listed in Fig. 6b are very small and the interval boundaries do not show any unwanted sign change. The “F-Value of Regression” of the 7 term model is also listed on the top of Fig. 6a. It is significantly larger than the corresponding value of the 23 term model that is listed in Fig. 5a (393303 versus 206052). All these observations indicate that the 7 term model is the better of the two regression models.

No confirmation points were available that could have been used for an independent test of the predictive capability of the two regression models. However, residuals and PRESS residuals of the two regression models can be computed and compared. Figures 7a and 7b show the residuals and the PRESS residuals for the 23 term model. Figures 8a and 8b show the corresponding results for the 7 term model. The residuals of the 23 term model are only slightly smaller than the residuals of the 7 term model even though the 23 term model uses more than 3 times as many terms for the fit. This observation and the comparison of the statistical metrics discussed earlier (variance inflation factors, F-Value of Regression, confidence interval boundaries) indicates that the 23 term model is probably overfitting the data and should not be used as a least squares representation of the calibration data.

IV. Summary and Conclusions

Different steps were reviewed in the paper that may be used for the selection of terms of a regression model of strain-gage balance calibration data. These steps will help develop a regression model of the calibration data that (i) meets strict statistical quality requirements, (ii) satisfies expected load prediction accuracy requirements, and (iii) prevents overfitting. Data from a simplified manual calibration of the Ames MK40 balance was used to illustrate the application of some of the metrics and tests that are discussed in the paper.

Commercially available software tools exist that support the recommended regression model term selection process (e.g., *SAS/STAT* distributed by the *SAS Institute Inc.*, or, *Design Expert*, distributed by *Stat-Ease Inc.*). These software tools require a certain degree of understanding and training in order to be applied with confidence during the regression analysis of strain-gage balance calibration data. Ultimately, the application of more advanced regression analysis software tools and the use of more advanced metrics during the analysis of strain-gage balance calibration data is an advantage. It allows an analyst to gain a better understanding of the quality of the calibration load schedule design and the experimental data. In addition, it leads to a more realistic assessment of the predictive capability of the final regression model as the use of more advanced statistical metrics helps to suppress the overfitting of balance calibration data.

V. Acknowledgements

The work reported in this paper was partially supported by NASA’s Aeronautic Test Program and the Wind Tunnel Division at Ames Research Center under contract NNA09DB39C.

VI. References

- ¹AIAA/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, sponsored and published by the American Institute of Aeronautics and Astronautics, Reston, Virginia, 2003.
- ²Ulbrich, N. and Volden, T., “Strain-Gage Balance Calibration Analysis Using Automatically Selected Math Models,” AIAA 2005-4084, paper presented at the 41st AIAA/ASME/SAE/ASEE Joint Propulsion Conference and Exhibit, Tucson, Arizona, July 2005.
- ³Montgomery, D. C., Peck, E. A., and Vining, G. G., *Introduction to Linear Regression Analysis*, 4th ed., John Wiley & Sons, Inc., New York, 2006, p. 84, pp. 125-126, pp. 141-142, pp. 323-341.
- ⁴Ulbrich, N., “Regression Model Optimization for the Analysis of Experimental Data,” AIAA 2009-1344, paper presented at the 47th AIAA Aerospace Sciences Meeting and Exhibit, Orlando, Florida, January 2009.
- ⁵Ulbrich, N. and Volden, T., “Regression Analysis of Experimental Data Using an Improved Math Model Search Algorithm,” AIAA 2008-0833, paper presented at the 46th AIAA Aerospace Sciences Meeting

and Exhibit, Reno, Nevada, January 2008.

⁶Ulbrich, N. and Volden, T., “Development of a User Interface for a Regression Analysis Software Tool,” AIAA 2010–0932, paper presented at the 48th AIAA Aerospace Sciences Meeting and Exhibit, Orlando, Florida, January 2010.

⁷Ulbrich, N., and Bader, J., “Analysis of Sting Balance Calibration Data Using Optimized Regression Models,” AIAA 2009–5372, paper presented at the 45th AIAA/ASME/SAE/ASEE Joint Propulsion Conference and Exhibit, Denver, Colorado, August 2009, pp. 3–4.

⁸Ulbrich, N., “Optimization of Regression Models of Experimental Data using Confirmation Points,” AIAA 2010–0930, paper presented at the 48th AIAA Aerospace Sciences Meeting and Exhibit, Orlando, Florida, January 2010.

⁹Ulbrich, N. and Volden, T., “Application of a New Calibration Analysis Process to the MK–III–C Balance,” AIAA 2006–0517, paper presented at the 44th AIAA Aerospace Sciences Meeting and Exhibit, Reno, Nevada, January 2010, pp. 9–12.

¹⁰DeLoach, R. and Ulbrich, N., “A Comparison of Two Balance Calibration Model Building Methods,” AIAA 2007–0147, paper presented at the 45th AIAA Aerospace Sciences Meeting and Exhibit, Reno, Nevada, January 2007.

¹¹DeLoach, R., “The Role of Hierarchy in Response Surface Modeling of Wind Tunnel Data,” AIAA 2010–0931, paper presented at the 48th AIAA Aerospace Sciences Meeting and Exhibit, Orlando, Florida, January 2010.

STEP 1 - REGRESSION ANALYSIS APPROACH

METHOD 1 (NON-ITERATIVE): ---> INDEP. VAR. = GAGE OUTPUTS ---> DEP. VAR. = LOADS

METHOD 2 (ITERATIVE): ---> INDEP. VAR. = LOADS ---> DEP. VAR. = GAGE OUTPUTS

STEP 2 - MATH TERM GROUP (FUNCTION CLASS) COMBINATION

FUNCTION CLASS OPTIONS ---> INTERCEPT, LINEAR, ABS, SQUARE, CUBIC, COMBINED

STEP 3 - LINEAR DEPENDENCY TEST (PROBLEM OF REGRESSION MODEL UNIQUENESS)

SINGULAR VALUE DECOMPOSITION (SVD) -----> IDENTIFIES LINEARLY DEPENDENCIES BETWEEN TERMS

STEP 4 - NEAR-LINEAR DEPENDENCY TEST (PROBLEM OF REGRESSION MODEL UNIQUENESS)

VARIANCE INFLATION FACTOR (VIF) -----> IDENTIFIES NEAR-LINEAR DEPENDENCIES BETWEEN TERMS

VARIANCE INFLATION FACTOR THRESHOLDS ---> 5 & 10

STEP 6 - STATISTICAL SIGNIFICANCE OF INDIVIDUAL TERMS

P-VALUE OF T-STATISTIC OF TERM -----> IDENTIFIES & REMOVES INSIGNIFICANT TERMS

P-VALUE OF T-STATISTIC THRESHOLDS -----> 0.0001 & 0.001

STEP 7 - APPLICATION OF THE HIERARCHY RULE (OPTIONAL)

OPTION 1 ---> NOT APPLIED, I.E., NO KNOWLEDGE EXISTS THAT SUGGESTS HIERARCHICAL MODEL

OPTION 2 ---> APPLIED DURING THE TERM SELECTION PROCESS

OPTION 3 ---> APPLIED AFTER THE TERM SELECTION PROCESS

STEP 8 - DIRECT TEST OF THE PREDICTIVE CAPABILITY OF THE REGRESSION MODEL

TEST 1: STANDARD DEVIATION OF RESIDUALS OF THE DEPENDENT VARIABLES

TEST 2: STANDARD DEVIATION OF PRESS RESIDUALS OF THE DEPENDENT VARIABLES

TEST 3: STANDARD DEVIATION OF CONFIRMATION POINT RESIDUALS OF THE DEPENDENT VARIABLES

Fig. 1 Regression model term selection process steps for the analysis of balance calibration data.

FUNCTION CLASS COMBINATION \Rightarrow	I	II	III	IV	V	VI
INTERCEPT [†]	(\times)	(\times)	(\times)	(\times)	(\times)	(\times)
A^{\ddagger}	\times	\times	\times	\times	\times	\times
$ A $		\times		\times	\times	\times
$A \cdot A$	\times	\times	\times	\times	\times	\times
$A \cdot A $				\times	\times	\times
$A \cdot B$	\times	\times	\times	\times	\times	\times
$ A \cdot B $					\times	\times
$A \cdot B $					\times	\times
$ A \cdot B$					\times	\times
$A \cdot A \cdot A$			\times			\times
$ A \cdot A \cdot A $						\times

[†] See detailed discussion of the intercept in the text; [‡] Symbols A and B represent the independent variables.

Fig. 2 Function class combination choices for the analysis of balance calibration data.

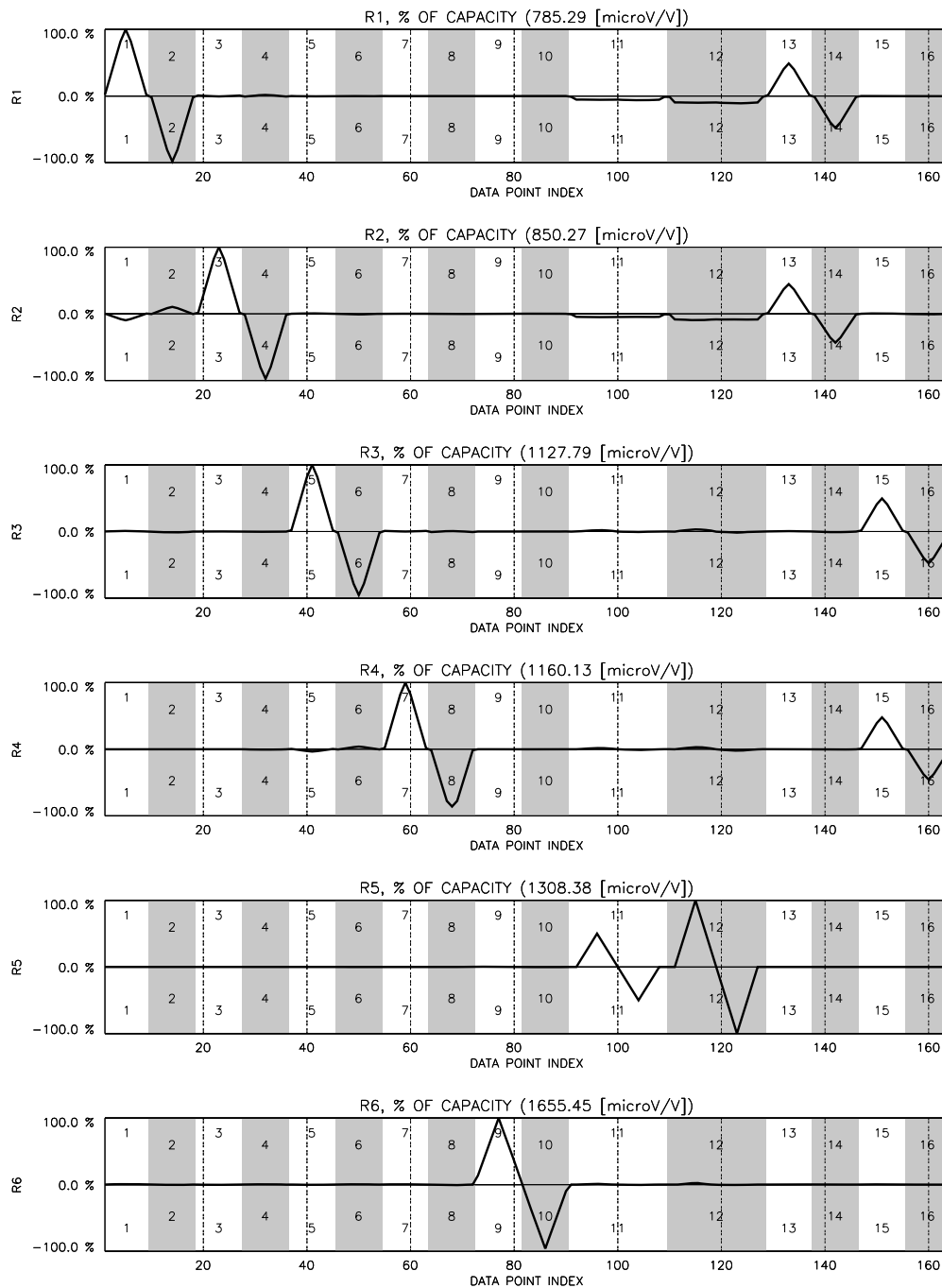


Fig. 3a Independent Variables: Gage outputs of calibration data of the Ames MK40 balance.

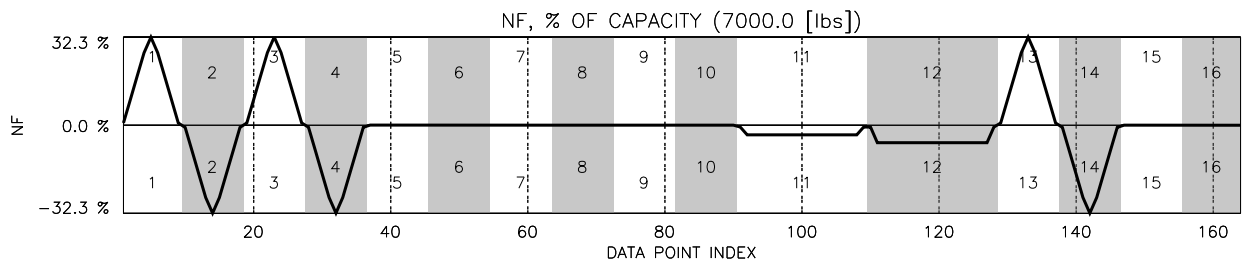


Fig. 3b Dependent Variable: Normal force of calibration data of the Ames MK40 balance.

INDEX	COEFFICIENT	NF
1	INTERCEPT	
2	R1	
3	R2	
4	R3	
5	R4	
6	R5	
7	R6	
8	IR1I	
9	IR2I	
10	IR3I	
11	IR4I	
12	IR5I	
13	IR6I	
14	R1*R1	
15	R2*R2	
16	R3*R3	
17	R4*R4	
18	R5*R5	
19	R6*R6	
26	R1*R2	
27	R1*R3	
28	R1*R4	
29	R1*R5	
30	R1*R6	
31	R2*R3	
32	R2*R4	
33	R2*R5	
34	R2*R6	
35	R3*R4	
36	R3*R5	
37	R3*R6	
38	R4*R5	
39	R4*R6	
40	R5*R6	

LINEAR (rows 2-7)

ABS (rows 8-13)

SQUARED (rows 14-19)

COMBINED (rows 26-40)

selected (points to R1*R2, R1*R5, R2*R5)

massive near-linear dependency (points to R1*R2, R1*R5, R2*R5)

Fig. 4a Upper bound of regression models after the application of Singular Value Decomposition.

INDEX	COEFFICIENT	NF
1	INTERCEPT	
2	R1	
3	R2	
4	R3	
5	R4	
6	R5	
7	R6	
8	IR1I	
9	IR2I	
10	IR3I	
11	IR4I	
12	IR5I	
13	IR6I	
14	R1*R1	
15	R2*R2	
16	R3*R3	
17	R4*R4	
18	R5*R5	
19	R6*R6	
26	R1*R2	
27	R1*R3	
28	R1*R4	
29	R1*R5	
30	R1*R6	
31	R2*R3	
32	R2*R4	
33	R2*R5	
34	R2*R6	
35	R3*R4	
36	R3*R5	
37	R3*R6	
38	R4*R5	
39	R4*R6	
40	R5*R6	

LINEAR (rows 2-7)

ABS (rows 8-13)

SQUARED (rows 14-19)

COMBINED (rows 26-40)

eliminated (points to R1*R2, R1*R3, R1*R4, R1*R5, R1*R6)

Fig. 4b Final regression model after the application of the regression model term selection process.

SOURCE OF VARIATION	SUM OF SQUARES	PRESS STATISTIC	DEGREES OF FREEDOM	MEAN SQUARE	F-VALUE OF REGRESSION	P-VALUE OF REGRESSION
REGRESSION	9.9684e+07	—	22	4.5311e+06	206051.5514	< 0.0001
RESIDUAL	3100.6091	5501.4784	141	21.9901	—	—
TOTAL	9.9687e+07	—	163	—	—	—
R-SQUARE		ADJ. R-SQUARE		PRESS R-SQUARE		
0.999969		0.999964		0.999945		

test of significance

MATH TERM CONTRIBUTES SIGNIFICANTLY IF THE T-STATISTIC OF ITS COEFFICIENT IS GREATER THAN 1.977
(SIGNIFICANCE = 2.50 %, DEGREES OF FREEDOM OF RESIDUAL = 141)

DIAGNOSTIC DESCRIBING NEAR-LINEAR DEPENDENCIES (MULTICOLLINEARITY) BETWEEN MATH MODEL TERMS
*** WARNING *** MAXIMUM OF VARIANCE INFLATION FACTOR (PRIMARY VALUE) = 98110.45 > UPPER LIMIT OF 10.0 *** WARNING ***

PHYSICAL VARIABLES & UNITS – REGRESSION COEFFICIENT ESTIMATES AND STATISTICAL METRICS (NF)
REGRESSION MODEL HIERARCHY CHARACTERISTICS = HIERARCHICAL

MATH TERM INDEX	MATH TERM NAME	COEFFICIENT VALUE	STANDARD ERROR	T-STATISTIC OF COEFFICIENT	P-VALUE OF COEFFICIENT	VIF (PRIMARY)	VIF (ALTERNATE)
1	INTERCEPT	-0.2057	+0.8599	-0.2392	—	—	—
2	R1	+3.2319	+0.0032	+1015.6838	< 0.0001	+1.0847	+2.2119
3	R2	+2.7362	+0.0020	+1334.9270	< 0.0001	+2.9167	+1.0346
4	R3	-0.0135	+0.0016	-8.7128	< 0.0001	+1.5685	+1.0436
5	R4	-0.0058	+0.0016	-3.7187	+0.0003	+1.5718	+1.0783
6	R5	-0.1253	+0.0163	-7.7056	< 0.0001	+38.2112	[+141.6188]
7	R6	+0.0035	+0.0012	+3.0021	+0.0032	+1.0251	+1.0122
8	IR11	-0.0000	+0.0000	-2.6935	+0.0079	+18.2658	[+18.2658]
9	IR21	-0.0000	+0.0000	-4.3902	< 0.0001	+18.4802	[+18.4802]
10	IR31	+0.0219	+0.0071	+3.0645	+0.0026	+18.7159	[+18.7159]
11	IR41	+0.0114	+0.0073	+1.5678	+0.1192	+19.4980	[+19.4980]
12	IR51	+0.0624	+0.0060	+10.3272	< 0.0001	+17.0164	[+17.0164]
13	IR61	-0.0000	+0.0000	-2.6311	+0.9083	+16.0782	[+16.0782]
14	R1•R1	-2.6253e-05	+7.8003e-06	-3.3771e-03	+0.1171	+18.8054	[+18.8384]
15	R2•R2	+4.0459e-05	+3.5756e-06	+1.1293e-02	+0.0095	+21.4352	[+21.4879]
16	R3•R3	-9.9441e-06	+7.6786e-06	-1.2890e-03	+0.2016	+16.5171	[+16.5844]
17	R4•R4	-3.8003e-06	+2.8700e-05	-0.4949	+0.6214	+16.4972	[+16.6415]
18	R5•R5	+0.0002	+5.0402e-08	+6.6537	< 0.0001	+380.2354	[+380.2363]
19	R6•R6	-5.0402e-08	+3.5756e-06	-0.0141	+0.9888	+14.8402	[+14.9170]
26	R1•R2	-1.2377e-05	+2.3169e-05	-0.5342	+0.5940	+2.5092	+2.5015
29	R1•R5	-0.0369	+0.0052	-7.1068	< 0.0001	+96785.5116	[+80462.3788]
33	R2•R5	+0.0373	+0.0053	+7.0560	< 0.0001	+98110.4470	[+74681.1173]
35	R3•R4	-1.3300e-05	+1.2202e-05	-1.0900	+0.2776	+2.9114	+2.4924

most important terms

massive near-linear dependency

Fig. 5a Analysis of variance results for upper bound of regression models (23 terms).

PHYSICAL VARIABLES & UNITS – REGRESSION COEFFICIENT ESTIMATES AND CONFIDENCE INTERVALS (NF)
REGRESSION MODEL HIERARCHY CHARACTERISTICS = HIERARCHICAL

MATH TERM INDEX	MATH TERM NAME	COEFFICIENT VALUE	95 % CI LOWER	95 % CI UPPER	99 % CI LOWER	99 % CI UPPER	VIF (PRIMARY)
1	INTERCEPT	-2.0571e-01	-1.9056e+00	+1.4942e+00	-2.4510e+00	+2.0395e+00	—
2	R1	+3.2319e+00	+3.2256e+00	+3.2382e+00	+3.2236e+00	+3.2402e+00	+1.0847
3	R2	+2.7362e+00	+2.7321e+00	+2.7403e+00	+2.7308e+00	+2.7416e+00	+2.9167
4	R3	-1.3546e-02	-1.6620e-02	-1.0473e-02	-1.7606e-02	-9.4866e-03	+1.5685
5	R4	-5.8488e-03	-8.9582e-03	-2.7395e-03	-9.9557e-03	-1.7420e-03	+1.5718
6	R5	-1.2530e-01	-1.5744e-01	-9.3150e-02	-1.6775e-01	-8.2838e-02	+38.2112
7	R6	+3.5028e-03	+1.1961e-03	+5.8094e-03	+4.5611e-04	+6.5494e-03	+1.0251
8	IR11	-2.7359e-02	-4.7439e-02	-7.2787e-03	-5.3881e-02	-8.3678e-04	+18.2658
9	IR21	-4.2487e-02	-6.1619e-02	-2.3355e-02	-6.7757e-02	-1.7217e-02	+18.4802
10	IR31	+2.1855e-02	+7.7561e-03	+3.5953e-02	+3.2331e-03	+4.0476e-02	+18.7159
11	IR41	+1.1368e-02	-2.9661e-03	+2.5702e-02	-7.5646e-03	+3.0301e-02	+19.4980
12	IR51	+6.2386e-02	+5.0444e-02	+7.4329e-02	+4.6612e-02	+7.8160e-02	+17.0164
13	IR61	-5.6280e-04	-1.0209e-02	+9.0832e-03	-1.3303e-02	+1.2178e-02	+16.0782
14	R1•R1	-2.6253e-05	-5.9171e-05	+6.6654e-06	-6.9732e-05	+1.7226e-05	+18.8054
15	R2•R2	+4.0459e-05	+1.0059e-05	+7.0858e-05	+3.0696e-07	+8.0611e-05	+21.4352
16	R3•R3	-9.9441e-06	-2.5266e-05	+5.3779e-06	-3.0182e-05	+1.0293e-05	+16.5171
17	R4•R4	-3.8003e-06	-1.8980e-05	+1.1380e-05	-2.3850e-05	+1.6250e-05	+16.4972
18	R5•R5	+1.9096e-04	+1.3422e-04	+2.4770e-04	+1.1602e-04	+2.6590e-04	+380.2354
19	R6•R6	-5.0402e-08	-7.1190e-06	+7.0182e-06	-9.3867e-06	+9.2859e-06	+14.8402
26	R1•R2	-1.2377e-05	-5.8180e-05	+3.3427e-05	-7.2875e-05	+4.8121e-05	+2.5092
29	R1•R5	-3.6922e-02	-4.7193e-02	-2.6652e-02	-5.0488e-02	-2.3357e-02	+96785.5116
33	R2•R5	+3.7338e-02	+2.6876e-02	+4.7799e-02	+2.3520e-02	+5.1155e-02	+98110.4470
35	R3•R4	-1.3300e-05	-3.7423e-05	+1.0823e-05	-4.5162e-05	+1.8562e-05	+2.9114

Fig. 5b Confidence intervals for terms of upper bound of regression models (23 terms, blue indicates sign change).

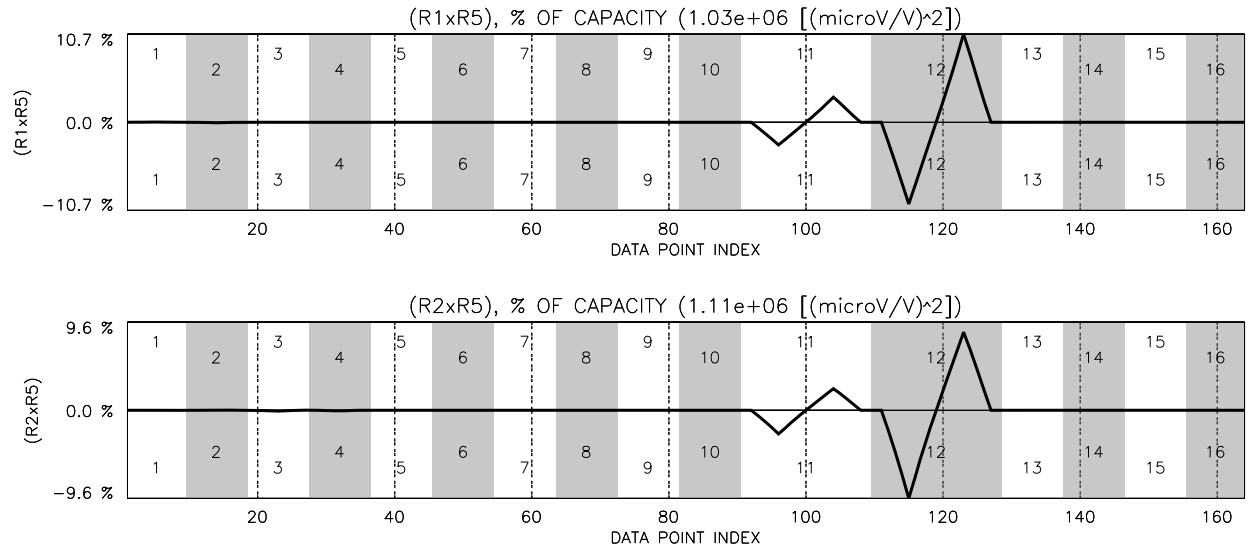


Fig. 5c Comparison of two products of the independent variables: $R1 \cdot R5$ with $R2 \cdot R5$.

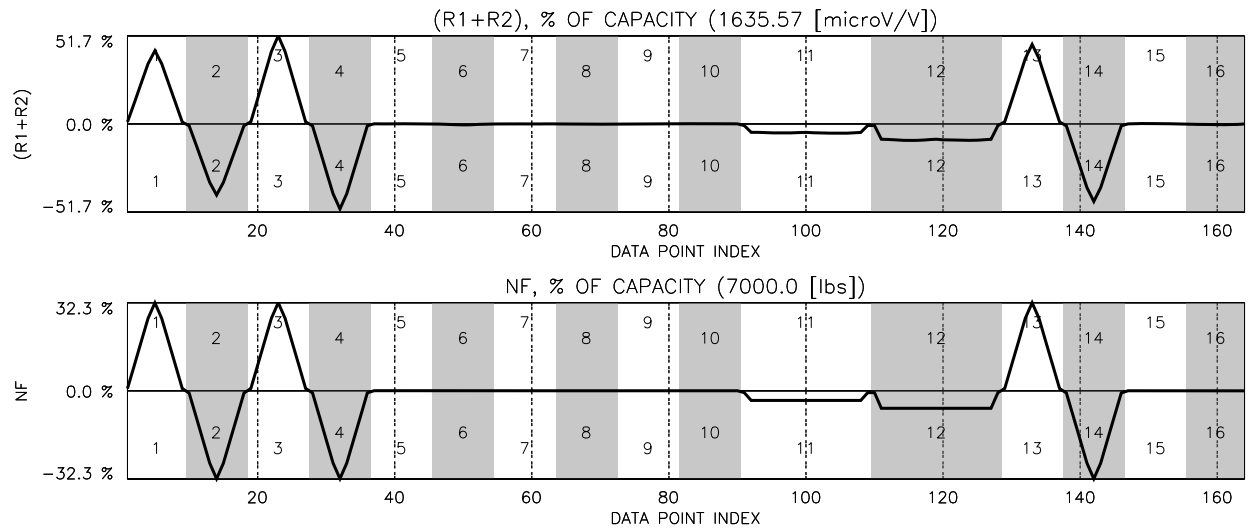


Fig. 5d Comparison of the sum of the independent variables $R1$ and $R2$ with the dependent variable NF .

SOURCE OF VARIATION	SUM OF SQUARES	PRESS STATISTIC	DEGREES OF FREEDOM	MEAN SQUARE	F-VALUE OF REGRESSION	P-VALUE OF REGRESSION
REGRESSION	9.9682e+07	—	7	1.4240e+07	393303.4283	< 0.0001
RESIDUAL	5684.4585	6868.4601	157	36.2067	—	—
TOTAL	9.9687e+07	—	163	—	—	—
R-SQUARE		ADJ. R-SQUARE		PRESS R-SQUARE		
0.999943		0.999941		0.999931		

test of significance

MATH TERM CONTRIBUTES SIGNIFICANTLY IF THE T-STATISTIC OF ITS COEFFICIENT IS GREATER THAN 1.975
(SIGNIFICANCE = 2.50 %, DEGREES OF FREEDOM OF RESIDUAL = 157)

DIAGNOSTIC DESCRIBING NEAR-LINEAR DEPENDENCIES (MULTICOLLINEARITY) BETWEEN MATH MODEL TERMS
MAXIMUM OF VARIANCE INFLATION FACTOR (PRIMARY VALUE) = 1.09 < UPPER LIMIT OF 10.0

PHYSICAL VARIABLES & UNITS – REGRESSION COEFFICIENT ESTIMATES AND STATISTICAL METRICS (NF)
REGRESSION MODEL HIERARCHY CHARACTERISTICS = HIERARCHICAL

MATH TERM INDEX	MATH TERM NAME	COEFFICIENT VALUE	STANDARD ERROR	T-STATISTIC OF COEFFICIENT	P-VALUE OF COEFFICIENT	VIF (PRIMARY)	VIF (ALTERNATE)
2	R1	+3.2151	+0.0028	+1156.5240	< 0.0001	+1.0261	+1.0261
3	R2	+2.7364	+0.0026	+1046.5711	< 0.0001	+1.0231	+1.0231
4	R3	-0.0150	+0.0020	-7.6726	< 0.0001	+1.0024	+1.0024
8	IR1I	-0.0321	+0.0029	-11.1648	< 0.0001	+1.0416	+1.0416
9	IR2I	-0.0453	+0.0027	-16.7269	< 0.0001	+1.0481	+1.0481
10	IR3I	+0.0154	+0.0020	+7.8510	< 0.0001	+1.0887	+1.0887
12	IR5I	+0.0258	+0.0018	+14.2922	< 0.0001	+1.0486	+1.0486

Fig. 6a Analysis of variance results for final regression model (7 terms).

PHYSICAL VARIABLES & UNITS – REGRESSION COEFFICIENT ESTIMATES AND CONFIDENCE INTERVALS (NF)							
REGRESSION MODEL HIERARCHY CHARACTERISTICS = HIERARCHICAL							
MATH TERM INDEX	MATH TERM NAME	COEFFICIENT VALUE	95 % CI LOWER	95 % CI UPPER	99 % CI LOWER	99 % CI UPPER	VIF (PRIMARY)
2	R1	+3.2151e+00	+3.2096e+00	+3.2206e+00	+3.2078e+00	+3.2223e+00	+1.0261
3	R2	+2.7364e+00	+2.7313e+00	+2.7416e+00	+2.7296e+00	+2.7432e+00	+1.0231
4	R3	-1.5001e-02	-1.8862e-02	-1.1139e-02	-2.0099e-02	-9.9028e-03	+1.0024
8	IR1I	-3.2117e-02	-3.7799e-02	-2.6436e-02	-3.9618e-02	-2.4617e-02	+1.0416
9	IR2I	-4.5256e-02	-5.0600e-02	-3.9912e-02	-5.2311e-02	-3.8202e-02	+1.0481
10	IR3I	+1.5356e-02	+1.1492e-02	+1.9219e-02	+1.0256e-02	+2.0456e-02	+1.0887
12	IR5I	+2.5848e-02	+2.2276e-02	+2.9420e-02	+2.1132e-02	+3.0564e-02	+1.0486

Fig. 6b Confidence intervals for terms of final regression model (7 terms).

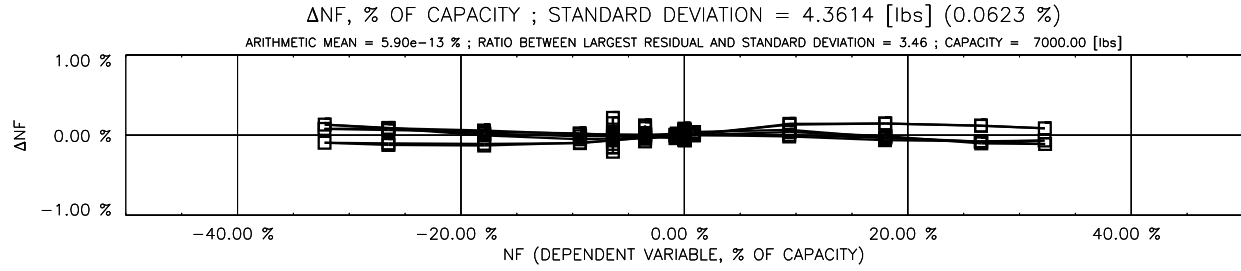


Fig. 7a Normal force residuals for upper bound of regression models (23 terms).

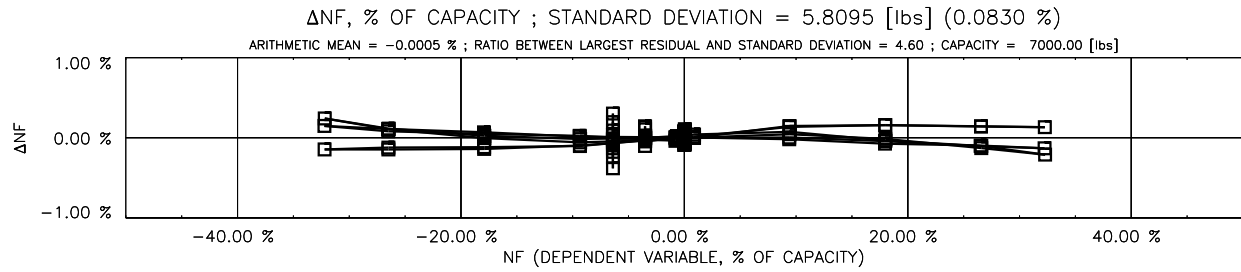


Fig. 7b PRESS residuals of the normal force for upper bound of regression models (23 terms).

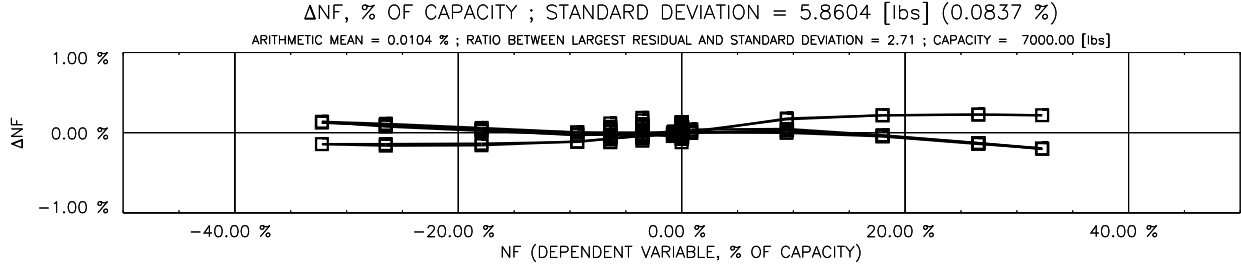


Fig. 8a Normal force residuals for final regression model (7 terms).

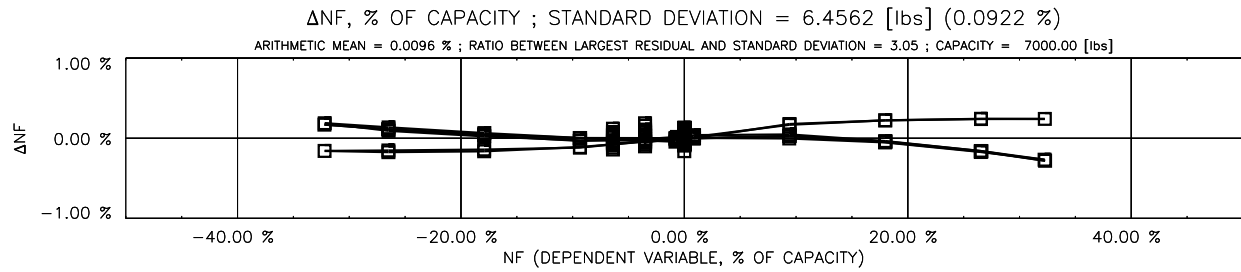


Fig. 8b PRESS residuals of the normal force for final regression model (7 terms).