Analysis of Multivariate Experimental Data Using A Simplified Regression Model Search Algorithm

N. Ulbrich[*]

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

A new regression model search algorithm was developed that may be applied to both general multivariate experimental data sets and wind tunnel strain–gage balance calibration data. The algorithm is a simplified version of a more complex algorithm that was originally developed for the NASA Ames Balance Calibration Laboratory. The new algorithm performs regression model term reduction to prevent overfitting of data. It has the advantage that it needs only about one tenth of the original algorithm's CPU time for the completion of a regression model search. In addition, extensive testing showed that the prediction accuracy of math models obtained from the simplified algorithm is similar to the prediction accuracy of math models obtained from the original algorithm. The simplified algorithm, however, cannot guarantee that search constraints related to a set of statistical quality requirements are always satisfied in the optimized regression model. Therefore, the simplified algorithm is not intended to replace the original algorithm. Instead, it may be used to generate an alternate optimized regression model of experimental data whenever the application of the original search algorithm fails or requires too much CPU time. Data from a machine calibration of NASA's MK40 force balance is used to illustrate the application of the new search algorithm.

Nomenclature

AF	=	axial force
$N1$	=	normal force component at the forward normal force gage of a strain–gage balance
$N2$	=	normal force component at the aft normal force gage of a strain–gage balance
$R1$	=	electrical output of the forward normal force gage
$R2$	=	electrical output of the aft normal force gage
$R3$	=	electrical output of the forward side force gage
$R4$	=	electrical output of the aft side force gage
$R5$	=	electrical output of the rolling moment gage
$R6$	=	electrical output of the axial force gage
RM	=	rolling moment
$S1$	=	side force component at the forward side force gage of a strain–gage balance
$S2$	=	side force component at the aft side force gage of a strain–gage balance

I. Introduction

During the past few years a regression model search algorithm was developed for the NASA Ames Balance Calibration Laboratory. The search algorithm performs regression model term reduction in order to identify regression models of multivariate experimental data sets that meet strict statistical quality requirements and prevent overfitting of the dependent variable (see Refs. [1] to [5] for a description of different aspects of the original search algorithm). The original search algorithm was successfully implemented in NASA's BALFIT software package (see Ref. [6]). Currently, the regression model search algorithm is applied on a regular basis to wind tunnel strain–gage balance calibration data sets at both NASA Ames Research

---

[*] Aerodynamicist, Jacobs Technology Inc.

American Institute of Aeronautics and Astronautics

Center and NASA Langley Research Center.

Experience showed over the years that the original regression model search algorithm is not always able to successfully complete a search. A failure of a regression model search occurs whenever the algorithm detects a discontinuity in its search metric, i.e., the standard deviation of the PRESS residuals of the data. Several causes for the failure of a regression model search were identified. They were related (i) to search restrictions imposed by the algorithm itself or (ii) to shortcomings and imperfections in experimental data sets. Consequently, the search restrictions of the original algorithm were relaxed by giving an analyst a greater choice of (i) math term group combinations, (ii) search constraint options, and (iii) search directions. In addition, improvements of BALFIT's analysis report format were made so that shortcomings in experimental data sets can more easily be spotted.

In 2010, after analyzing problems associated with a sparse wind tunnel balance calibration data set, it was concluded that the application of the search algorithm is also sometimes unsuccessful because tare load corrections have to be applied to some balance data sets <u>before</u> the regression model search can be started (see Ref. [7] for a detailed description of the tare load iteration process). Unfortunately, a sufficiently accurate tare load correction can often only be obtained if a good estimate of the final optimized regression model of the balance calibration data is already known ("chicken–and–egg" problem). Therefore, the author attempted to simplify several elements of the original regression model search algorithm such that the CPU time needed for the search would be significantly reduced while still enforcing the three search constraints of the original search algorithm (i.e., variance inflation factor constraint, p–value constraint, hierarchy rule constraint). These efforts lead to the development of the simplified regression model search algorithm that is discussed in the present paper.

Important elements of the simplified regression model search algorithm are briefly summarized in the next section. Then, data from the machine calibration of a wind tunnel strain–gage balance is used as an example in order (i) to determine the CPU time that the simplified algorithm needs for the completion of a typical regression model search and (ii) to assess the predictive capability of its optimized regression models.

II. Simplified Search Algorithm

The simplified regression model search algorithm was derived from a more complex algorithm that was developed for the NASA Ames Balance Calibration Laboratory (see Refs. [1] to [4] for details about the original algorithm). The development of the simplified search algorithm became necessary because some wind tunnel strain–gage balance calibration data sets require tare corrections before the original more complex search algorithm can be applied. Accurate tare corrections, however, can often only be obtained if a good initial "guess" of the final optimized regression model is already available before the beginning of the search. This "guess" has to be determined within a fraction of the CPU time that the original algorithm needs for the completion of a regression model search. The simplified search algorithm was designed to fulfill all these requirements.

The simplified search algorithm has matured since it was first developed in 2011. It is now included in NASA's BALFIT software package as a new math model type choice that an analyst can select for analysis (i.e., the math model type choice *Suggested Math Model*). The simplified search algorithm uses many of the statistical quality metrics and constraints that the original search algorithm applies (see Ref. [5] for a general discussion of different metrics that may be used to evaluate a regression model of experimental data). However, the quality metrics and constraints are applied differently in order to reduce the total number of numerical operations that have to be performed during a typical regression model search. These differences are explained in more detail below. In addition, the simplified search algorithm omits the search metric minimization that the original algorithm performs (see Ref. [3] for details about the search metric minimization used by the original algorithm).

The flowchart in Fig. 1 summarizes basic elements of the simplified search algorithm. The search starts by first selecting a combination of math term groups, i.e., function classes, that help define an upper limit for the regression model search. The analyst should select this math term group combination by using some subject matter knowledge about the given experimental data set. The chosen combination, e.g., may consist of linear terms, cross–product terms, square terms, and absolute value terms. Math term groups suggested for the analysis of wind tunnel strain–gage balance calibration data are listed in Ref. [7].

Not every possible term of the chosen math term groups may be supported by the given experimental data set. Therefore, the upper bound of the math model, i.e., the so–called *Permitted Math Model*, needs to be determined so that only non–singular solutions of the least squares problem can exist when different regression models are tested during the search. This requirement is enforced by using a numerical technique called Singular Value Decomposition (SVD). At this point, after completion of the SVD process, the *Permitted Math Model* is known. It defines the upper bound of the search space that the simplified regression model search algorithm needs for the identification of the optimized math model.

In the next step, the variance inflation factor constraint has to be applied in order to prevent near–linear dependencies in the optimized regression model. The variance inflation factor constraint requires that the largest variance inflation factor (VIF) of a tested math model is less than a threshold. The literature recommends a conservative threshold of 10 for the near–linear dependency test (see, e.g., Ref. [8], p. 111; some analysts believe that a threshold of up to 50 is still acceptable). Then, the tested math model will not have unwanted near–linear dependencies. The simplified algorithm enforces the VIF constraint iteratively. At first, the VIFs of all terms of the *Permitted Math Model* are computed (Step 1). Then, the term of the maximum VIF is identified and removed if its VIF exceeds the chosen threshold (Step 2). Now, the VIFs of all terms of the updated math model are computed (Step 3). Finally, Step 2 and Step 3 are repeated until the largest VIF of the remaining math model is below the VIF threshold (Step 4). The remaining math model is a first estimate of the optimized regression model (Step 5).

If needed, the first estimate of the optimized regression model, is made "hierarchical" by adding missing lower order terms. Enforcing the "hierarchy" rule in the first estimate guarantees that the math model can represent possible "hidden" constant shifts in the independent variables correctly.

Finally, the p–values of the t–statistic of all regression coefficients of the "hierarchical" first estimate of the optimized regression model are computed. The calculation of the p–values of the coefficients helps identify and remove statistically "insignificant" terms. A term is removed from the first estimate of the optimized regression model if its p–value exceeds the literature recommended threshold of 0.001 (threshold is taken from Ref. [8], p. 85; some analysts prefer the more conservative threshold of 0.0001). The p–values are processed by examining higher order terms before lower order terms. In addition, a term is only removed if the remaining math model remains "hierarchical." The resulting regression model is the final optimized regression model that the simplified regression model search algorithm recommends for the regression analysis of the given data.

Extensive studies of the performance of the simplified regression model search algorithm have shown that the CPU time needed for a typical search drops by one order of magnitude when compared with the CPU time that the original algorithm requires. This advantage of the simplified algorithm becomes important whenever large data sets need to be processed. The significant reduction in CPU time became possible because the simplified algorithm performs an iterative reduction of math terms by exclusively using the VIFs. Afterwards, the removal of "insignificant" terms is performed by analyzing a single math model, i.e., the "hierarchical" first estimate of the optimized math model (see also Fig. 1). The original regression model search algorithm, on the other hand, minimizes a search metric and enforces constraints at the same time. Therefore, it has to compute the search metric, the VIFs, and the p–values of the t–statistic of the regression coefficients for each regression model that is tested during a search. This alternate approach requires more numerical operations, i.e., CPU time, because the calculation of the p–values is always preceeded by a least squares fit of the entire data set.

The next section of the paper discusses the regression analysis of a wind tunnel strain–gage balance calibration data example in order to illustrate different aspects of the application of the simplified search algorithm.

III. Discussion of Example

Data from the calibration of NASA's MK40 wind tunnel strain–gage balance was chosen to illustrate the application of the simplified regression model search algorithm. The MK40 balance was manufactured by the *Task Corporation* (see Ref. [7] for a discussion of strain–gage balances). It is a six–component force balance that measures five forces ($N1$, $N2$, $S1$, $S2$, AF) and one moment (RM). The balance has a diameter of 2.5 inches and a length of 17.31 inches. Table 1 shows the load capacity of each load component.

Table 1: Load capacities of the NASA's 2.5 inch MK40 balance.

	$N1$, *lbs*	$N2$, *lbs*	$S1$, *lbs*	$S2$, *lbs*	RM, *in–lbs*	AF, *lbs*
CAPACITY	3500	3500	2500	2500	8000	400

The MK40 balance was calibrated in 2008 in Triumph Aerospace's *Automatic Balance Calibration System* (ABCS). Triumph's ABCS is a balance calibration machine that can apply complex combined loadings to a balance. These combined loadings are needed for an appropriate characterization of interactions between the six balance gages. The final balance calibration data set consisted of the applied loads and the measured electrical outputs of the strain–gages of the balance. The supplied calibration data was already corrected for the tare loads caused by the metric part of the balance and other calibration fixtures. Therefore, the tare load iteration process was omitted during the analysis of the data. Table 2 below summarizes important characteristics of the balance and its calibration data set.

Table 2: Balance and calibration data set characteristics of the MK40 balance.

BALANCE TYPE	FORCE BALANCE
DIAMETER	2.5 [in]
CALIBRATION DATE	JULY 2008
CALIBRATION METHOD	MACHINE CALIBRATION
TOTAL NUMBER OF CALIBRATION POINTS	1863
LOAD FORMAT	TARE CORRECTED LOADS
GAGE OUTPUT FORMAT	GAGE OUTPUT DIFFERENCES

It was decided to analyze the balance calibration data by applying the *Iterative Method*. This approach fits the strain–gage outputs as a function of the applied balance loads. Afterwards, an iteration scheme is applied so that loads can be predicted from the measured gage outputs during a wind tunnel test. For simplicity, only the fit of the gage outputs as a function of the applied balance loads is discussed in the paper. Detailed information about the load iteration scheme may be found in Ref. [7].

First, in order to start the search for the optimized regression models of the six gage outputs of the balance, it was necessary to define an upper bound, i.e., the *Permitted Math Model*. The balance is a multi–piece balance. Therefore, as the data is to be processed using the *Iterative Method*, it was decided to include absolute value terms in the math term group combination for the regression models of the six gage outputs of the balance. The following five groups were chosen for the regression models of the gage outputs:

Group 1 $\equiv$ F **Group 2** $\equiv$ $|F|$ **Group 3** $\equiv$ F^2 **Group 4** $\equiv$ $F \cdot |F|$ **Group 5** $\equiv$ $F \cdot G$

The chosen math term groups are a subset of groups that are recommended in Ref. [7] for the analysis of strain–gage balance calibration data. The symbols F and G represent the load components of the balance (i.e., $N1$, $N2$, $S1$, $\cdots$). The balance calibration data set was obtained in a calibration machine that supports all load combinations that can be constructed from the five chosen math term groups. Therefore, after including the intercept in the regression models of the gage outputs, the initial upper bound for the search has a total of 40 terms for each gage output. The corresponding regression models of the gage outputs $R1$ to $R6$ are shown in Fig. 2. The individual terms of the six regression models are constructed from the given balance loads.

Specific knowledge about characteristics of a *Task* balance may be used to further reduce the number of regression model terms in the initial upper bound. In a recent paper it was shown that not all gages of a *Task* balance have the bi–directional characteristic that justifies the selection of absolute value terms in the regression model of calibration data of a *Task* balance (see Ref. [9]). Figure 3a shows, for example, the bi–directionality characteristic of the MK40 balance by using the indicator variable that is defined in Ref. [9].

It can clearly be seen that the rolling moment gage output $R5$ and the axial force gage output $R6$ are <u>not</u> bi–directional when plotted versus the corresponding primary gage load (see the two plots near the bottom of Fig. 3a). The plots of the indicator variable for the outputs $R5$ and $R6$ simply do not look like an absolute value function as the indicator variable at capacity stays below an empirical threshold. Consequently, there is no justification for the use of the terms $|AF|$ and $|RM|$ in the regression models of the gage outputs of the MK40 balance. A revised 36–term upper bound of the regression models is obtained after removing corresponding terms from the initial upper bound (see Fig. 3b).

An initial analysis of the variance inflation factors of the 36–term upper bound for the regression model of the forward normal force gage output $R1$ is shown in Fig. 3c. A moderate near–linear dependency between the absolute value terms of the normal/side force components and corresponding quadratic terms appears to exist as terms of those two function classes have variance inflation factors that are between the literature recommended threshold of 10 and an intermediate threshold of 50 (see two red boxes in Fig. 3c). The connection between the absolute value and quadratic terms comes from the fact that both the absolute value function and a parabola are even functions. The variance inflation factor analysis of the MK40 calibration data indicates that either the absolute value terms or the quadratic terms of the normal/side force components may be omitted in the regression models. A *Task* balance is, by design, a very linear device. Therefore, it was decided to omit the quadratic instead of the absolute value terms of the normal and side force components. The corresponding second revision of the upper bound for the regression model search is shown in Fig. 3d.

Now, the original and the simplified search algorithms can be applied to the MK40 calibration data set by using the reduced math model shown in Fig. 3d as the upper bound for the regression models of the gage outputs. Figure 4a shows the optimized models of the six gage outputs that were obtained after applying the original search algorithm to the calibration data. It took approximately 28 minutes of CPU time to complete the search. Figure 4b shows the calibration load residuals of the six load components. The loads residuals, i.e., the difference between applied and calculated loads, were obtained after applying the load iteration scheme that can be constructed from the fitted coefficients of regression models shown in Fig. 4a. (see again Ref. [7] for a description of the load iteration process).

Figure 5a shows the optimized regression models of the gage outputs that were obtained after applying the simplified search algorithm to the calibration data. In this case, it took approximately 3 minutes of CPU time to complete the search. Figure 5b shows the calibration load residuals of the six load components. Again, the load residuals were obtained after applying the load iteration scheme that can be constructed from the coefficients of the regression models shown in Fig. 5a.

At this point, it is interesting to compare the standard deviation of the load residuals for the five different regression model sets that were discussed above. This metric may be used for a preliminary assessment of the predictive capability of the five regression models whenever regression model independent confirmation points are not available for analysis. Table 3 below lists the corresponding standard deviations as a percentage of the load capacity for the six load components.

Table 3: Standard deviation of the load residuals in percent of the load capacity.

MATH MODEL	$N1$	$N2$	$S1$	$S2$	RM	AF
Fig. 2	0.1008 %	0.0967 %	0.2115 %	0.1762 %	0.1780 %	0.1354 %
Fig. 3b	0.1101 %	0.1032 %	0.2402 %	0.1822 %	0.2011 %	0.1709 %
Fig. 3d	0.1207 %	0.1063 %	0.2456 %	0.1931 %	0.2016 %	0.1760 %
Fig. 4a	0.1216 %	0.1065 %	0.2473 %	0.1986 %	0.2055 %	0.1808 %
Fig. 5a	0.1229 %	0.1075 %	0.2481 %	0.1942 %	0.2065 %	0.1782 %

Overall, standard deviations of the load residuals compare very well for the five regression model sets. The comparison also illustrates that differences between the predictive capability of the regression models of the original search algorithm (Fig. 4a) and the regression models of the simplified search algorithm (Fig. 5a) are very small if the fitted data itself is chosen to assess their predictive capability. The observed differences are less than 0.005 % of the load capacity.

IV. Summary and Conclusions

A new simplified regression model search algorithm was described. The new algorithm was derived from a more complex algorithm that was originally developed for the NASA Ames Balance Calibration Laboratory. The new simplified search algorithm tries to rapidly identify a regression model for a given experimental data set that meets strict statistical quality requirements and simultaneously prevents overfitting of the data. The simplified search algorithm has the advantage that it requires only about one tenth of the original algorithm's CPU time for the completion of a regression model search. However, the simplified search algorithm cannot guarantee that the final optimized regression model will always meet the chosen statistical quality requirements. Therefore, the simplified search algorithm is intended to be used for the generation of an alternate optimized regression model whenever the application of the original search algorithm fails or requires too much CPU time. The simplified algorithm is also good for an initial assessment of an experimental data set as it allows an analyst to catch potential problems sooner. In addition, the simplified algorithm may be used to get better initial estimates of tare corrected loads of wind tunnel strain–gage balance calibration data. These improved estimates are needed as input whenever the original search algorithm is applied to strain–gage balance calibration data that has not yet been corrected for the weight of the metric part of the balance, the calibration body, and other calibration hardware components.

Machine calibration data of NASA's MK40 six–component force balance was used to illustrate the application of the simplified search algorithm to a strain–gage balance calibration data set. First, it was explained how a suitable upper bound for the regression model search can be defined for the given data set. New information about the bi–directionality characteristics of the individual gages of a *Task* balance was used for this purpose. Therefore, all absolute value terms of the rolling moment and axial force and the quadratic terms of the normal and side force components were omitted in the regression models of the gage outputs. Then, both the original and the simplified search algorithm were applied to the calibration data set using the upper bound to limit the regression model search. This analysis demonstrated that prediction accuracy estimates of the optimized regression models show excellent agreement whenever results for the optimized models of the original algorithm were compared with corresponding results for the optimized models of the simplified algorithm. However, the simplified algorithm has the key advantage that it can generate an optimized math model much faster than the original algorithm. It required only about one tenth of the original algorithm's CPU time for the completion of the regression model search.

Acknowledgements

References

[1]Ulbrich, 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.

[2]Ulbrich, 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 2006.

[3]Ulbrich, 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.

[4]Ulbrich, 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.

[5]Ulbrich, N. and Volden, T., "Regression Model Term Selection for the Analysis of Strain–Gage Balance Calibration Data," AIAA 2010–4545, paper presented at the 27th AIAA Aerodynamic Measurement Technology and Ground Testing Conference, Chicago, Illinois, June/July 2010.

[6]Ulbrich, 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.

[7]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.

[8]Montgomery, D. C., Peck, E. A., and Vining, G. G., *Introduction to Linear Regression Analysis*, 4th ed., John Wiley & Sons, Inc., New York, 2006, p. 111.

[9]Ulbrich, N., "Detection of Bi–Directionality in Strain–Gage Balance Calibration Data," AIAA 2012–3320, paper presented at the 28th AIAA Aerodynamic Measurement Technology, Ground Testing, and Flight Testing Conference, New Orleans, Louisiana, June 2012.

American Institute of Aeronautics and Astronautics

EXPERIMENTAL DATA SET

I. SUBJECT MATTER KNOWLEDGE OF ANALYST

... SELECT A MATH TERM GROUP COMBINATION (FUNCTION CLASS SET)
THAT DEFINES THE UPPER LIMIT OF THE PERMITTED MATH MODEL

II. REMOVAL OF LINEAR DEPENDENCIES BETWEEN MATH TERMS

... APPLY SINGULAR VALUE DECOMPOSITION (SVD)

– USE SVD TO DETERMINE THE PERMITTED MATH MODEL, I.E., THE LARGEST MATH
MODEL OF THE GIVEN DATA THAT LEADS TO A NON–SINGULAR LEAST SQUARES FIT

– PERMITTED MATH MODEL = UPPER BOUND OF THE REGRESSION MODEL SEARCH

III. REMOVAL OF NEAR–LINEAR DEPENDENCIES BETWEEN MATH TERMS

... APPLY VARIANCE INFLATION FACTOR (VIF) CONSTRAINT

STEP 1:   COMPUTE VIFs OF THE PERMITTED MATH MODEL (UPPER BOUND OF SEARCH)
STEP 2:   REMOVE MATH TERM OF THE LARGEST VIF IF ITS VIF EXCEEDS THRESHOLD
STEP 3:   COMPUTE VIFs OF THE UPDATED MATH MODEL AFTER THE TERM IS REMOVED
STEP 4:   REPEAT STEPS 2 & 3 UNTIL THE LARGEST VIF IS BELOW THE THRESHOLD
STEP 5: USE REMAINING MATH MODEL AS FIRST ESTIMATE OF THE FINAL MATH MODEL

IV. APPLICATION OF THE "HIERARCHY" RULE

... ADD MISSING LOWER ORDER TERMS TO FIRST ESTIMATE

V. REMOVAL OF STATISTICALLY "INSIGNIFICANT" MATH TERMS

... APPLY P–VALUE CONSTRAINT

STEP 1: DETERMINE P–VALUES OF T–STATISTIC OF ALL REGRESSION
COEFFICIENTS OF THE "HIERARCHICAL" FIRST ESTIMATE

STEP 2: REMOVE TERMS FROM FIRST ESTIMATE THAT EXCEED P–VALUE THRESHOLD
(REMOVAL IS PERFORMED FROM HIGHER ORDER TO LOWER ORDER TERMS;
A TERM IS ONLY REMOVED IF MATH MODEL REMAINS "HIERARCHICAL")

OPTIMIZED MATH MODEL
(ALSO CALLED SUGGESTED MATH MODEL)

Fig. 1 Basic elements of the simplified regression model search algorithm.

American Institute of Aeronautics and Astronautics

NUMBER OF TERMS = 40, 40, 40, 40, 40, 40

HIERARCHICAL: R1, R2, R3, R4, R5, R6

(HIERARCHY ANALYSIS USES |F*G|=|F|*|G|, |F*F*F|=|F|*|F|*|F|, |F|*|F|=F*F)

	R1	R2	R3	R4	R5	R6
INTERCEPT	■	■	■	■	■	■
N1	■	■	■	■	■	■
N2	■	■	■	■	■	■
S1	■	■	■	■	■	■
S2	■	■	■	■	■	■
RM	■	■	■	■	■	■
AF	■	■	■	■	■	■
|N1|	■	■	■	■	■	■
|N2|	■	■	■	■	■	■
|S1|	■	■	■	■	■	■
|S2|	■	■	■	■	■	■
|RM|	■	■	■	■	■	■
|AF|	■	■	■	■	■	■
N1*N1	■	■	■	■	■	■
N2*N2	■	■	■	■	■	■
S1*S1	■	■	■	■	■	■
S2*S2	■	■	■	■	■	■
RM*RM	■	■	■	■	■	■
AF*AF	■	■	■	■	■	■
N1*|N1|	■	■	■	■	■	■
N2*|N2|	■	■	■	■	■	■
S1*|S1|	■	■	■	■	■	■
S2*|S2|	■	■	■	■	■	■
RM*|RM|	■	■	■	■	■	■
AF*|AF|	■	■	■	■	■	■
N1*N2	■	■	■	■	■	■
N1*S1	■	■	■	■	■	■
N1*S2	■	■	■	■	■	■
N1*RM	■	■	■	■	■	■
N1*AF	■	■	■	■	■	■
N2*S1	■	■	■	■	■	■
N2*S2	■	■	■	■	■	■
N2*RM	■	■	■	■	■	■
N2*AF	■	■	■	■	■	■
S1*S2	■	■	■	■	■	■
S1*RM	■	■	■	■	■	■
S1*AF	■	■	■	■	■	■
S2*RM	■	■	■	■	■	■
S2*AF	■	■	■	■	■	■
RM*AF	■	■	■	■	■	■
|N1*N2|	□	□	□	□	□	□
|N1*S1|	□	□	□	□	□	□
|N1*S2|	□	□	□	□	□	□
|N1*RM|	□	□	□	□	□	□
|N1*AF|	□	□	□	□	□	□
|N2*S1|	□	□	□	□	□	□
|N2*S2|	□	□	□	□	□	□
|N2*RM|	□	□	□	□	□	□
|N2*AF|	□	□	□	□	□	□
|S1*S2|	□	□	□	□	□	□

	R1	R2	R3	R4	R5	R6
|S1*RM|	□	□	□	□	□	□
|S1*AF|	□	□	□	□	□	□
|S2*RM|	□	□	□	□	□	□
|S2*AF|	□	□	□	□	□	□
|RM*AF|	□	□	□	□	□	□
N1*|N2|	□	□	□	□	□	□
N1*|S1|	□	□	□	□	□	□
N1*|S2|	□	□	□	□	□	□
N1*|RM|	□	□	□	□	□	□
N1*|AF|	□	□	□	□	□	□
N2*|S1|	□	□	□	□	□	□
N2*|S2|	□	□	□	□	□	□
N2*|RM|	□	□	□	□	□	□
N2*|AF|	□	□	□	□	□	□
S1*|S2|	□	□	□	□	□	□
S1*|RM|	□	□	□	□	□	□
S1*|AF|	□	□	□	□	□	□
S2*|RM|	□	□	□	□	□	□
S2*|AF|	□	□	□	□	□	□
RM*|AF|	□	□	□	□	□	□
|N1|*N2	□	□	□	□	□	□
|N1|*S1	□	□	□	□	□	□
|N1|*S2	□	□	□	□	□	□
|N1|*RM	□	□	□	□	□	□
|N1|*AF	□	□	□	□	□	□
|N2|*S1	□	□	□	□	□	□
|N2|*S2	□	□	□	□	□	□
|N2|*RM	□	□	□	□	□	□
|N2|*AF	□	□	□	□	□	□
|S1|*S2	□	□	□	□	□	□
|S1|*RM	□	□	□	□	□	□
|S1|*AF	□	□	□	□	□	□
|S2|*RM	□	□	□	□	□	□
|S2|*AF	□	□	□	□	□	□
|RM|*AF	□	□	□	□	□	□
N1*N1*N1	□	□	□	□	□	□
N2*N2*N2	□	□	□	□	□	□
S1*S1*S1	□	□	□	□	□	□
S2*S2*S2	□	□	□	□	□	□
RM*RM*RM	□	□	□	□	□	□
AF*AF*AF	□	□	□	□	□	□
|N1*N1*N1|	□	□	□	□	□	□
|N2*N2*N2|	□	□	□	□	□	□
|S1*S1*S1|	□	□	□	□	□	□
|S2*S2*S2|	□	□	□	□	□	□
|RM*RM*RM|	□	□	□	□	□	□
|AF*AF*AF|	□	□	□	□	□	□

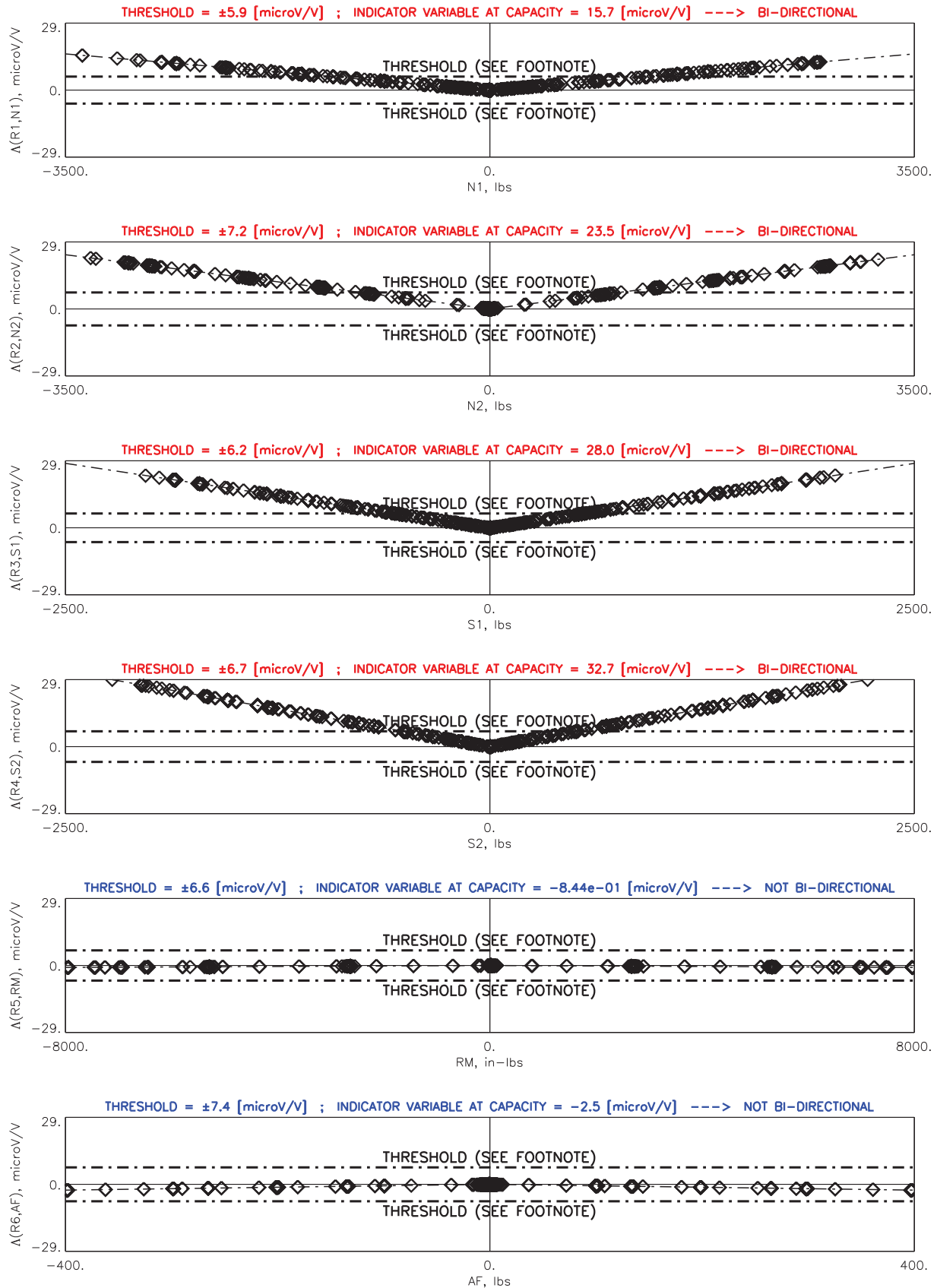**Fig. 2** Initial version of the permitted math model of the MK40 calibration data set.

Fig. 3a Indicator variable for bi–directionality of the MK40 balance.

American Institute of Aeronautics and Astronautics

NUMBER OF TERMS = 36, 36, 36, 36, 36, 36

HIERARCHICAL: R1, R2, R3, R4, R5, R6

(HIERARCHY ANALYSIS USES |F*G|=|F|*|G|, |F*F*F|=|F|*|F|*|F|, |F|*|F|=F*F)

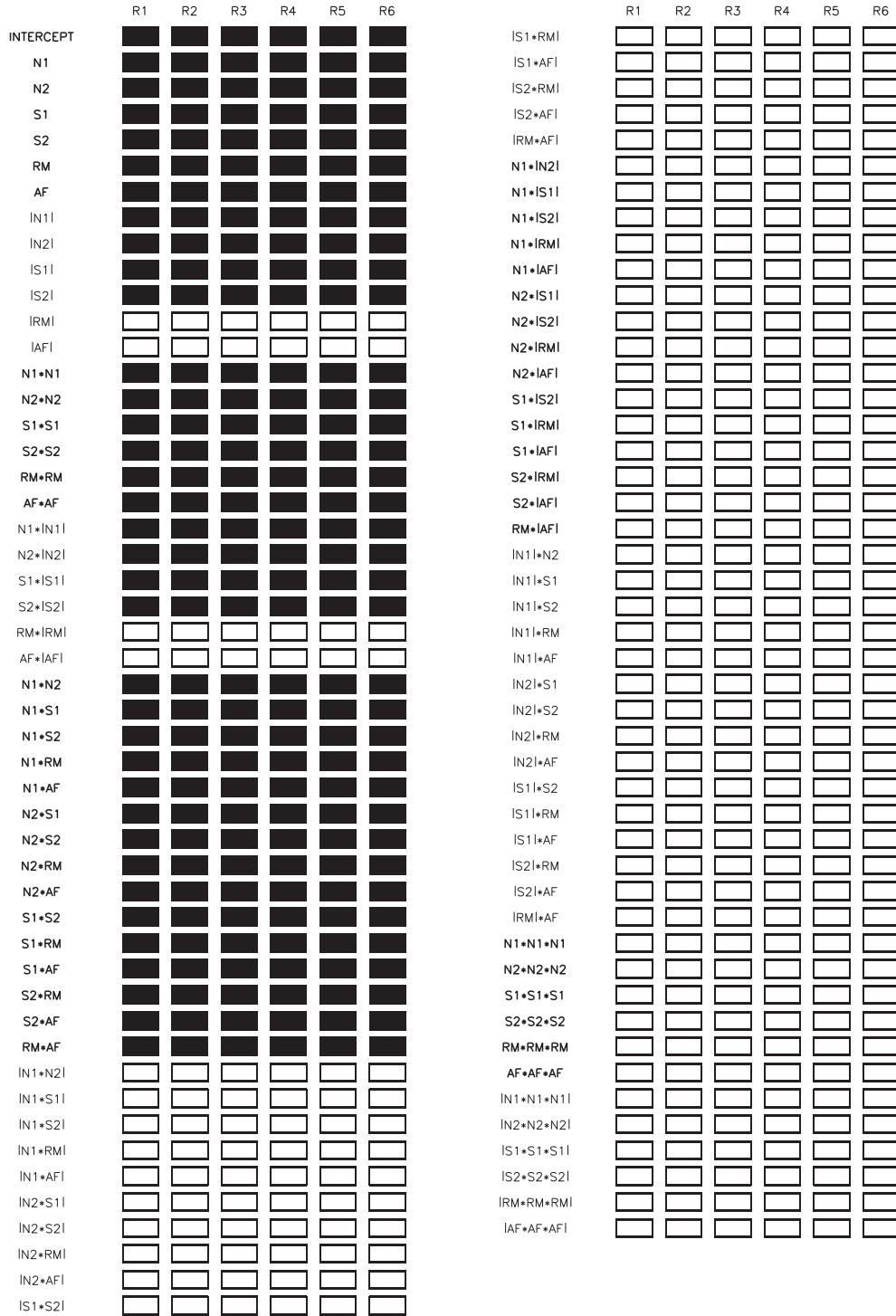| Term	R1	R2	R3	R4	R5	R6
INTERCEPT	■	■	■	■	■	■
N1	■	■	■	■	■	■
N2	■	■	■	■	■	■
S1	■	■	■	■	■	■
S2	■	■	■	■	■	■
RM	■	■	■	■	■	■
AF	■	■	■	■	■	■
|N1|	■	■	■	■	■	■
|N2|	■	■	■	■	■	■
|S1|	■	■	■	■	■	■
|S2|	■	■	■	■	■	■
|RM|	□	□	□	□	□	□
|AF|	□	□	□	□	□	□
N1*N1	■	■	■	■	■	■
N2*N2	■	■	■	■	■	■
S1*S1	■	■	■	■	■	■
S2*S2	■	■	■	■	■	■
RM*RM	■	■	■	■	■	■
AF*AF	■	■	■	■	■	■
N1*|N1|	■	■	■	■	■	■
N2*|N2|	■	■	■	■	■	■
S1*|S1|	■	■	■	■	■	■
S2*|S2|	■	■	■	■	■	■
RM*|RM|	□	□	□	□	□	□
AF*|AF|	□	□	□	□	□	□
N1*N2	■	■	■	■	■	■
N1*S1	■	■	■	■	■	■
N1*S2	■	■	■	■	■	■
N1*RM	■	■	■	■	■	■
N1*AF	■	■	■	■	■	■
N2*S1	■	■	■	■	■	■
N2*S2	■	■	■	■	■	■
N2*RM	■	■	■	■	■	■
N2*AF	■	■	■	■	■	■
S1*S2	■	■	■	■	■	■
S1*RM	■	■	■	■	■	■
S1*AF	■	■	■	■	■	■
S2*RM	■	■	■	■	■	■
S2*AF	■	■	■	■	■	■
RM*AF	■	■	■	■	■	■
|N1*N2|	□	□	□	□	□	□
|N1*S1|	□	□	□	□	□	□
|N1*S2|	□	□	□	□	□	□
|N1*RM|	□	□	□	□	□	□
|N1*AF|	□	□	□	□	□	□
|N2*S1|	□	□	□	□	□	□
|N2*S2|	□	□	□	□	□	□
|N2*RM|	□	□	□	□	□	□
|N2*AF|	□	□	□	□	□	□
|S1*S2|	□	□	□	□	□	□
|S1*RM|	□	□	□	□	□	□
|S1*AF|	□	□	□	□	□	□
|S2*RM|	□	□	□	□	□	□
|S2*AF|	□	□	□	□	□	□
|RM*AF|	□	□	□	□	□	□
N1*|N2|	□	□	□	□	□	□
N1*|S1|	□	□	□	□	□	□
N1*|S2|	□	□	□	□	□	□
N1*|RM|	□	□	□	□	□	□
N1*|AF|	□	□	□	□	□	□
N2*|S1|	□	□	□	□	□	□
N2*|S2|	□	□	□	□	□	□
N2*|RM|	□	□	□	□	□	□
N2*|AF|	□	□	□	□	□	□
S1*|S2|	□	□	□	□	□	□
S1*|RM|	□	□	□	□	□	□
S1*|AF|	□	□	□	□	□	□
S2*|RM|	□	□	□	□	□	□
S2*|AF|	□	□	□	□	□	□
RM*|AF|	□	□	□	□	□	□
|N1|*N2	□	□	□	□	□	□
|N1|*S1	□	□	□	□	□	□
|N1|*S2	□	□	□	□	□	□
|N1|*RM	□	□	□	□	□	□
|N1|*AF	□	□	□	□	□	□
|N2|*S1	□	□	□	□	□	□
|N2|*S2	□	□	□	□	□	□
|N2|*RM	□	□	□	□	□	□
|N2|*AF	□	□	□	□	□	□
|S1|*S2	□	□	□	□	□	□
|S1|*RM	□	□	□	□	□	□
|S1|*AF	□	□	□	□	□	□
|S2|*RM	□	□	□	□	□	□
|S2|*AF	□	□	□	□	□	□
|RM|*AF	□	□	□	□	□	□
N1*N1*N1	□	□	□	□	□	□
N2*N2*N2	□	□	□	□	□	□
S1*S1*S1	□	□	□	□	□	□
S2*S2*S2	□	□	□	□	□	□
RM*RM*RM	□	□	□	□	□	□
AF*AF*AF	□	□	□	□	□	□
|N1*N1*N1|	□	□	□	□	□	□
|N2*N2*N2|	□	□	□	□	□	□
|S1*S1*S1|	□	□	□	□	□	□
|S2*S2*S2|	□	□	□	□	□	□
|RM*RM*RM|	□	□	□	□	□	□
|AF*AF*AF|	□	□	□	□	□	□

Fig. 3b First revision of the permitted math model after removal of terms related to $|AF|$ and $|RM|$.

PHYSICAL VARIABLES & UNITS — REGRESSION COEFFICIENT ESTIMATES AND STATISTICAL METRICS (R1)

REGRESSION MODEL HIERARCHY CHARACTERISTICS = HIERARCHICAL

TERM INDEX	TERM NAME	COEFFICIENT VALUE	STANDARD ERROR	T−STATISTIC OF COEFFICIENT	P−VALUE OF COEFFICIENT	VIF (PRIMARY)	VIF (ALTERNATE)
1	INTERCEPT	−1.3915	+0.0724	−19.2079	−	−	−
2	N1	+0.3435	+0.0001	+3218.6915	< 0.0001	+3.1535	[+11.1063]
3	N2	−0.0081	+0.0001	−73.9053	< 0.0001	+2.7146	[+17.9385]
4	S1	+0.0002	+0.0001	+1.5301	+0.1262	+2.6630	[+12.7765]
5	S2	+0.0002	+0.0001	+1.6099	+0.1076	+3.1455	[+16.5158]
6	RM	+0.0005	+1.2244e−05	+40.4820	< 0.0001	+1.2300	+1.0116
7	AF	+0.0009	+0.0003	+3.5283	+0.0004	+1.0939	+1.0052
8	|N1|	−0.0004	+0.0002	−2.2613	+0.0239	+16.6365	[+16.1960]
9	|N2|	+0.0021	+0.0001	+15.7623	< 0.0001	+18.6616	[+18.6444]
10	|S1|	−0.0016	+0.0003	−6.0404	< 0.0001	+24.2112	[+24.2108]
11	|S2|	−0.0003	+0.0002	−1.4333	+0.1519	+24.0623	[+24.0623]
14	N1*N1	+1.2580e−06	+7.0086e−08	+17.9486	< 0.0001	+17.4855	[+16.6215]
15	N2*N2	−1.3589e−08	+5.0617e−08	−0.2685	+0.7884	+17.4322	[+17.4844]
16	S1*S1	−9.7985e−08	+1.3672e−07	−0.7167	+0.4737	+18.8256	[+18.8276]
17	S2*S2	+5.1810e−08	+1.1008e−07	+0.4707	+0.6379	+17.8914	[+17.8914]
18	RM*RM	−1.3117e−07	+2.2172e−09	−59.1601	< 0.0001	+1.1603	+1.1606
19	AF*AF	+2.6983e−06	+8.1051e−07	+3.3291	+0.0009	+1.0432	+1.0432
20	N1*|N1|	+1.2456e−06	+5.1676e−08	+24.1036	< 0.0001	+1.3995	[+11.4671]
21	N2*|N2|	+1.6883e−07	+4.5120e−08	+3.7419	+0.0002	+2.6652	[+18.2346]
22	S1*|S1|	−3.3343e−07	+9.3268e−08	−3.5750	+0.0004	+1.9403	[+11.9933]
23	S2*|S2|	−2.4281e−07	+8.6584e−08	−2.8043	+0.0051	+2.3194	[+15.1907]
26	N1*N2	−1.4454e−07	+2.2080e−08	−6.5461	< 0.0001	+1.5788	+1.5701
27	N1*S1	−1.1806e−06	+6.1998e−08	−19.0419	< 0.0001	+1.7335	+1.5130
28	N1*S2	+2.2981e−08	+5.5697e−08	+0.4126	+0.6799	+1.7532	+1.5240
29	N1*RM	−3.0356e−07	+1.9196e−08	−15.8139	< 0.0001	+1.2554	+1.0358
30	N1*AF	−1.6866e−07	+2.8800e−07	−0.5856	+0.5582	+1.1131	+1.0233
31	N2*S1	+1.1083e−07	+4.9463e−08	+2.2408	+0.0252	+1.5110	+1.5093
32	N2*S2	−1.0591e−07	+4.4563e−08	−2.3766	+0.0176	+1.5228	+1.5204
33	N2*RM	+1.3387e−08	+1.4845e−08	+0.9018	+0.3673	+1.0344	+1.0323
34	N2*AF	+9.7052e−08	+2.4748e−07	+0.3922	+0.6950	+1.0199	+1.0227
35	S1*S2	+3.2285e−08	+5.0522e−08	+0.6390	+0.5229	+1.8467	+1.8467
36	S1*RM	−8.8355e−07	+2.6275e−08	−33.6266	< 0.0001	+1.4044	+1.4048
37	S1*AF	−3.8047e−07	+4.1734e−07	−0.9117	+0.3621	+1.3182	+1.3183
38	S2*RM	−8.6731e−08	+2.5668e−08	−3.3790	+0.0007	+1.4104	+1.4108
39	S2*AF	+4.6004e−09	+3.8899e−07	+0.0118	+0.9906	+1.3182	+1.3183
40	RM*AF	−3.4890e−08	+1.2391e−07	−0.2816	+0.7783	+1.0000	+1.0001

Fig. 3c Variance inflation factors of the math model of gage output $R1$ for the revised permitted math model.

American Institute of Aeronautics and Astronautics

NUMBER OF TERMS = 32, 32, 32, 32, 32, 32

HIERARCHICAL: R1, R2, R3, R4, R5, R6

(HIERARCHY ANALYSIS USES |F*G|=|F|*|G|, |F*F*F|=|F|*|F|*|F|, |F|*|F|=F*F)

	R1	R2	R3	R4	R5	R6			R1	R2	R3	R4	R5	R6
INTERCEPT	■	■	■	■	■	■		|S1*RM|	□	□	□	□	□	□
N1	■	■	■	■	■	■		|S1*AF|	□	□	□	□	□	□
N2	■	■	■	■	■	■		|S2*RM|	□	□	□	□	□	□
S1	■	■	■	■	■	■		|S2*AF|	□	□	□	□	□	□
S2	■	■	■	■	■	■		|RM*AF|	□	□	□	□	□	□
RM	■	■	■	■	■	■		N1*|N2|	□	□	□	□	□	□
AF	■	■	■	■	■	■		N1*|S1|	□	□	□	□	□	□
|N1|	■	■	■	■	■	■		N1*|S2|	□	□	□	□	□	□
|N2|	■	■	■	■	■	■		N1*|RM|	□	□	□	□	□	□
|S1|	■	■	■	■	■	■		N1*|AF|	□	□	□	□	□	□
|S2|	■	■	■	■	■	■		N2*|S1|	□	□	□	□	□	□
|RM|	□	□	□	□	□	□		N2*|S2|	□	□	□	□	□	□
|AF|	□	□	□	□	□	□		N2*|RM|	□	□	□	□	□	□
N1*N1	□	□	□	□	□	□		N2*|AF|	□	□	□	□	□	□
N2*N2	□	□	□	□	□	□		S1*|S2|	□	□	□	□	□	□
S1*S1	□	□	□	□	□	□		S1*|RM|	□	□	□	□	□	□
S2*S2	□	□	□	□	□	□		S1*|AF|	□	□	□	□	□	□
RM*RM	■	■	■	■	■	■		S2*|RM|	□	□	□	□	□	□
AF*AF	■	■	■	■	■	■		S2*|AF|	□	□	□	□	□	□
N1*|N1|	■	■	■	■	■	■		RM*|AF|	□	□	□	□	□	□
N2*|N2|	■	■	■	■	■	■		|N1|*N2	□	□	□	□	□	□
S1*|S1|	■	■	■	■	■	■		|N1|*S1	□	□	□	□	□	□
S2*|S2|	■	■	■	■	■	■		|N1|*S2	□	□	□	□	□	□
RM*|RM|	□	□	□	□	□	□		|N1|*RM	□	□	□	□	□	□
AF*|AF|	□	□	□	□	□	□		|N1|*AF	□	□	□	□	□	□
N1*N2	■	■	■	■	■	■		|N2|*S1	□	□	□	□	□	□
N1*S1	■	■	■	■	■	■		|N2|*S2	□	□	□	□	□	□
N1*S2	■	■	■	■	■	■		|N2|*RM	□	□	□	□	□	□
N1*RM	■	■	■	■	■	■		|N2|*AF	□	□	□	□	□	□
N1*AF	■	■	■	■	■	■		|S1|*S2	□	□	□	□	□	□
N2*S1	■	■	■	■	■	■		|S1|*RM	□	□	□	□	□	□
N2*S2	■	■	■	■	■	■		|S1|*AF	□	□	□	□	□	□
N2*RM	■	■	■	■	■	■		|S2|*RM	□	□	□	□	□	□
N2*AF	■	■	■	■	■	■		|S2|*AF	□	□	□	□	□	□
S1*S2	■	■	■	■	■	■		|RM|*AF	□	□	□	□	□	□
S1*RM	■	■	■	■	■	■		N1*N1*N1	□	□	□	□	□	□
S1*AF	■	■	■	■	■	■		N2*N2*N2	□	□	□	□	□	□
S2*RM	■	■	■	■	■	■		S1*S1*S1	□	□	□	□	□	□
S2*AF	■	■	■	■	■	■		S2*S2*S2	□	□	□	□	□	□
RM*AF	■	■	■	■	■	■		RM*RM*RM	□	□	□	□	□	□
|N1*N2|	□	□	□	□	□	□		AF*AF*AF	□	□	□	□	□	□
|N1*S1|	□	□	□	□	□	□		|N1*N1*N1|	□	□	□	□	□	□
|N1*S2|	□	□	□	□	□	□		|N2*N2*N2|	□	□	□	□	□	□
|N1*RM|	□	□	□	□	□	□		|S1*S1*S1|	□	□	□	□	□	□
|N1*AF|	□	□	□	□	□	□		|S2*S2*S2|	□	□	□	□	□	□
|N2*S1|	□	□	□	□	□	□		|RM*RM*RM|	□	□	□	□	□	□
|N2*S2|	□	□	□	□	□	□		|AF*AF*AF|	□	□	□	□	□	□
|N2*RM|	□	□	□	□	□	□								
|N2*AF|	□	□	□	□	□	□								
|S1*S2|	□	□	□	□	□	□								

Fig. 3d Second revision of the permitted math model after removal of square terms of $N1$, $N2$, $S1$, and $S2$. (Second revision is used as the upper bound for both the <u>original</u> and <u>simplified</u> search algorithm.)

HIERARCHICAL: R1, R2, R3, R4, R5, R6
(HIERARCHY ANALYSIS USES |F*G|=|F|*|G|, |F*F*F|=|F|*|F|*|F|, |F|*|F|=F*F)

Left column labels (rows), with columns R1, R2, R3, R4, R5, R6:

INTERCEPT, N1, N2, S1, S2, RM, AF, |N1|, |N2|, |S1|, |S2|, |RM|, |AF|, N1*N1, N2*N2, S1*S1, S2*S2, RM*RM, AF*AF, N1*|N1|, N2*|N2|, S1*|S1|, S2*|S2|, RM*|RM|, AF*|AF|, N1*N2, N1*S1, N1*S2, N1*RM, N1*AF, N2*S1, N2*S2, N2*RM, N2*AF, S1*S2, S1*RM, S1*AF, S2*RM, S2*AF, RM*AF, |N1*N2|, |N1*S1|, |N1*S2|, |N1*RM|, |N1*AF|, |N2*S1|, |N2*S2|, |N2*RM|, |N2*AF|, |S1*S2|

Right column labels (rows), with columns R1, R2, R3, R4, R5, R6:

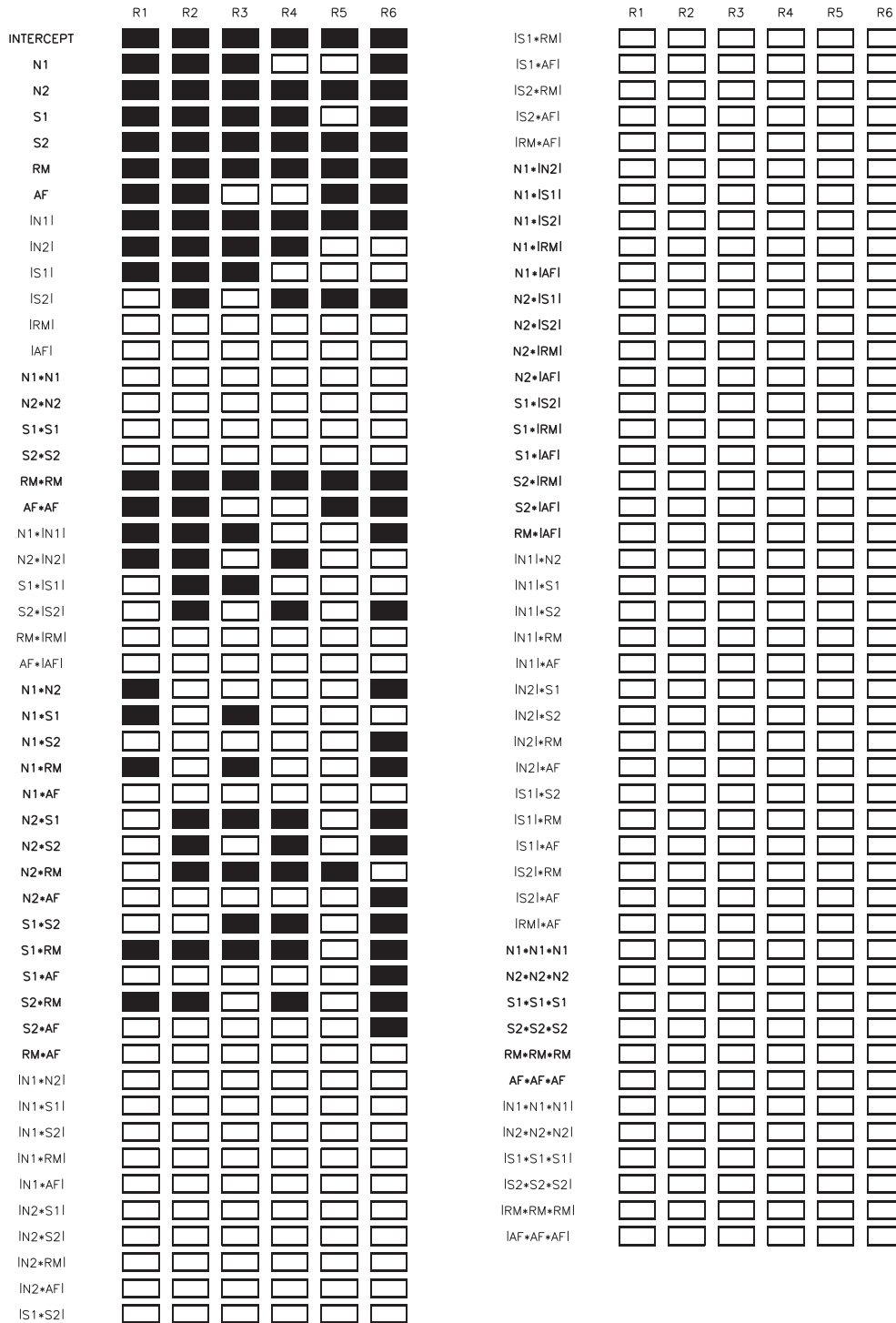|S1*RM|, |S1*AF|, |S2*RM|, |S2*AF|, |RM*AF|, N1*|N2|, N1*|S1|, N1*|S2|, N1*|RM|, N1*|AF|, N2*|S1|, N2*|S2|, N2*|RM|, N2*|AF|, S1*|S2|, S1*|RM|, S1*|AF|, S2*|RM|, S2*|AF|, RM*|AF|, |N1|*N2, |N1|*S1, |N1|*S2, |N1|*RM, |N1|*AF, |N2|*S1, |N2|*S2, |N2|*RM, |N2|*AF, |S1|*S2, |S1|*RM, |S1|*AF, |S2|*RM, |S2|*AF, |RM|*AF, N1*N1*N1, N2*N2*N2, S1*S1*S1, S2*S2*S2, RM*RM*RM, AF*AF*AF, |N1*N1*N1|, |N2*N2*N2|, |S1*S1*S1|, |S2*S2*S2|, |RM*RM*RM|, |AF*AF*AF|

Fig. 4a Optimized math model obtained after application of the <u>original</u> search algorithm. (Regression model search was completed after ≈ 28 minutes of CPU time.)

ΔN1, % OF CAPACITY ; STANDARD DEVIATION = 4.2569 [lbs] OR 0.1216 %
ARITHMETIC MEAN = −0.0002 % ; RATIO BETWEEN LARGEST RESIDUAL AND STANDARD DEVIATION = 4.30 ; CAPACITY = 3500.00 [lbs]

ΔN2, % OF CAPACITY ; STANDARD DEVIATION = 3.7267 [lbs] OR 0.1065 %
ARITHMETIC MEAN = −0.0001 % ; RATIO BETWEEN LARGEST RESIDUAL AND STANDARD DEVIATION = 4.23 ; CAPACITY = 3500.00 [lbs]

ΔS1, % OF CAPACITY ; STANDARD DEVIATION = 6.1836 [lbs] OR 0.2473 %
ARITHMETIC MEAN = 0.0013 % ; RATIO BETWEEN LARGEST RESIDUAL AND STANDARD DEVIATION = 4.63 ; CAPACITY = 2500.00 [lbs]

ΔS2, % OF CAPACITY ; STANDARD DEVIATION = 4.9641 [lbs] OR 0.1986 %
ARITHMETIC MEAN = 0.0007 % ; RATIO BETWEEN LARGEST RESIDUAL AND STANDARD DEVIATION = 4.76 ; CAPACITY = 2500.00 [lbs]

ΔRM, % OF CAPACITY ; STANDARD DEVIATION = 16.4424 [in−lbs] OR 0.2055 %
ARITHMETIC MEAN = −2.56e−05 % ; RATIO BETWEEN LARGEST RESIDUAL AND STANDARD DEVIATION = 7.73 ; CAPACITY = 8000.00 [in−lbs]

ΔAF, % OF CAPACITY ; STANDARD DEVIATION = 0.7233 [lbs] OR 0.1808 %
ARITHMETIC MEAN = −9.60e−05 % ; RATIO BETWEEN LARGEST RESIDUAL AND STANDARD DEVIATION = 6.14 ; CAPACITY = 400.00 [lbs]

Fig. 4b Load residuals for optimized math model of the <u>original</u> search algorithm.

HIERARCHICAL: R1, R2, R3, R4, R5, R6
(HIERARCHY ANALYSIS USES |F*G|=|F|*|G|, |F*F*F|=|F|*|F|*|F|, |F|*|F|=F*F)

Term	R1	R2	R3	R4	R5	R6
INTERCEPT	■	■	■	■	■	■
N1	■	■	■	■	■	■
N2	■	■	■	■	■	■
S1	■	■	■	■	□	■
S2	□	■	■	■	■	■
RM	■	■	■	■	■	■
AF	■	■	□	□	■	■
|N1|	■	■	■	■	□	■
|N2|	■	■	■	■	□	■
|S1|	■	■	■	□	□	□
|S2|	□	■	□	□	■	■
|RM|	□	□	□	□	□	□
|AF|	□	□	□	□	□	□
N1*N1	□	□	□	□	□	□
N2*N2	□	□	□	□	□	□
S1*S1	□	□	□	□	□	□
S2*S2	□	□	□	□	□	□
RM*RM	■	■	■	■	□	■
AF*AF	■	□	□	□	□	□
N1*|N1|	■	■	■	□	□	■
N2*|N2|	□	■	□	■	□	■
S1*|S1|	□	□	■	□	□	□
S2*|S2|	□	■	□	■	□	■
RM*|RM|	□	□	□	□	□	□
AF*|AF|	□	□	□	□	□	□
N1*N2	□	□	□	□	□	■
N1*S1	■	□	■	□	□	□
N1*S2	□	□	□	□	□	■
N1*RM	■	□	■	■	■	■
N1*AF	□	□	□	□	□	□
N2*S1	□	□	□	■	□	□
N2*S2	□	■	□	■	□	□
N2*RM	□	■	■	■	■	□
N2*AF	□	□	□	□	□	■
S1*S2	□	□	■	■	□	■
S1*RM	■	□	■	□	□	■
S1*AF	□	□	□	□	□	■
S2*RM	□	■	□	■	□	□
S2*AF	□	□	□	□	□	■
RM*AF	□	□	□	□	□	□
|N1*N2|	□	□	□	□	□	□
|N1*S1|	□	□	□	□	□	□
|N1*S2|	□	□	□	□	□	□
|N1*RM|	□	□	□	□	□	□
|N1*AF|	□	□	□	□	□	□
|N2*S1|	□	□	□	□	□	□
|N2*S2|	□	□	□	□	□	□
|N2*RM|	□	□	□	□	□	□
|N2*AF|	□	□	□	□	□	□
|S1*S2|	□	□	□	□	□	□

Term	R1	R2	R3	R4	R5	R6
|S1*RM|	□	□	□	□	□	□
|S1*AF|	□	□	□	□	□	□
|S2*RM|	□	□	□	□	□	□
|S2*AF|	□	□	□	□	□	□
|RM*AF|	□	□	□	□	□	□
N1*|N2|	□	□	□	□	□	□
N1*|S1|	□	□	□	□	□	□
N1*|S2|	□	□	□	□	□	□
N1*|RM|	□	□	□	□	□	□
N1*|AF|	□	□	□	□	□	□
N2*|S1|	□	□	□	□	□	□
N2*|S2|	□	□	□	□	□	□
N2*|RM|	□	□	□	□	□	□
N2*|AF|	□	□	□	□	□	□
S1*|S2|	□	□	□	□	□	□
S1*|RM|	□	□	□	□	□	□
S1*|AF|	□	□	□	□	□	□
S2*|RM|	□	□	□	□	□	□
S2*|AF|	□	□	□	□	□	□
RM*|AF|	□	□	□	□	□	□
|N1|*N2	□	□	□	□	□	□
|N1|*S1	□	□	□	□	□	□
|N1|*S2	□	□	□	□	□	□
|N1|*RM	□	□	□	□	□	□
|N1|*AF	□	□	□	□	□	□
|N2|*S1	□	□	□	□	□	□
|N2|*S2	□	□	□	□	□	□
|N2|*RM	□	□	□	□	□	□
|N2|*AF	□	□	□	□	□	□
|S1|*S2	□	□	□	□	□	□
|S1|*RM	□	□	□	□	□	□
|S1|*AF	□	□	□	□	□	□
|S2|*RM	□	□	□	□	□	□
|S2|*AF	□	□	□	□	□	□
|RM|*AF	□	□	□	□	□	□
N1*N1*N1	□	□	□	□	□	□
N2*N2*N2	□	□	□	□	□	□
S1*S1*S1	□	□	□	□	□	□
S2*S2*S2	□	□	□	□	□	□
RM*RM*RM	□	□	□	□	□	□
AF*AF*AF	□	□	□	□	□	□
|N1*N1*N1|	□	□	□	□	□	□
|N2*N2*N2|	□	□	□	□	□	□
|S1*S1*S1|	□	□	□	□	□	□
|S2*S2*S2|	□	□	□	□	□	□
|RM*RM*RM|	□	□	□	□	□	□
|AF*AF*AF|	□	□	□	□	□	□

Fig. 5a Optimized math model obtained after application of the <u>simplified</u> search algorithm. (Regression model search was completed after ≈ 3 minutes of CPU time.)
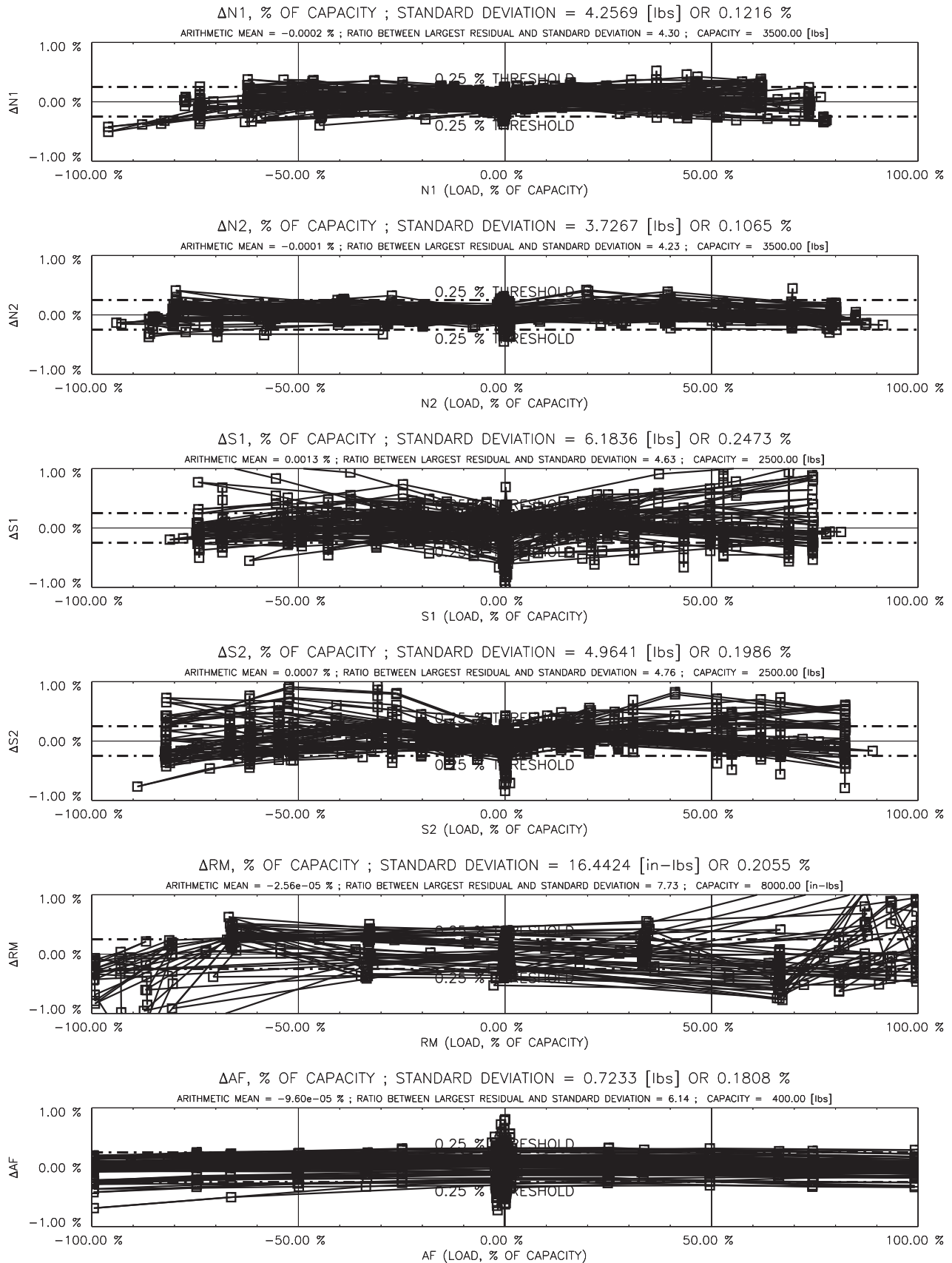
ΔN1, % OF CAPACITY ; STANDARD DEVIATION = 4.3004 [lbs] OR 0.1229 %
ARITHMETIC MEAN = −0.0002 % ; RATIO BETWEEN LARGEST RESIDUAL AND STANDARD DEVIATION = 4.32 ; CAPACITY = 3500.00 [lbs]

ΔN2, % OF CAPACITY ; STANDARD DEVIATION = 3.7632 [lbs] OR 0.1075 %
ARITHMETIC MEAN = −0.0002 % ; RATIO BETWEEN LARGEST RESIDUAL AND STANDARD DEVIATION = 4.21 ; CAPACITY = 3500.00 [lbs]

ΔS1, % OF CAPACITY ; STANDARD DEVIATION = 6.2024 [lbs] OR 0.2481 %
ARITHMETIC MEAN = 0.0013 % ; RATIO BETWEEN LARGEST RESIDUAL AND STANDARD DEVIATION = 4.62 ; CAPACITY = 2500.00 [lbs]

ΔS2, % OF CAPACITY ; STANDARD DEVIATION = 4.8544 [lbs] OR 0.1942 %
ARITHMETIC MEAN = 0.0007 % ; RATIO BETWEEN LARGEST RESIDUAL AND STANDARD DEVIATION = 4.88 ; CAPACITY = 2500.00 [lbs]

ΔRM, % OF CAPACITY ; STANDARD DEVIATION = 16.5232 [in−lbs] OR 0.2065 %
ARITHMETIC MEAN = 8.09e−07 % ; RATIO BETWEEN LARGEST RESIDUAL AND STANDARD DEVIATION = 8.13 ; CAPACITY = 8000.00 [in−lbs]

ΔAF, % OF CAPACITY ; STANDARD DEVIATION = 0.7128 [lbs] OR 0.1782 %
ARITHMETIC MEAN = −9.09e−05 % ; RATIO BETWEEN LARGEST RESIDUAL AND STANDARD DEVIATION = 6.29 ; CAPACITY = 400.00 [lbs]
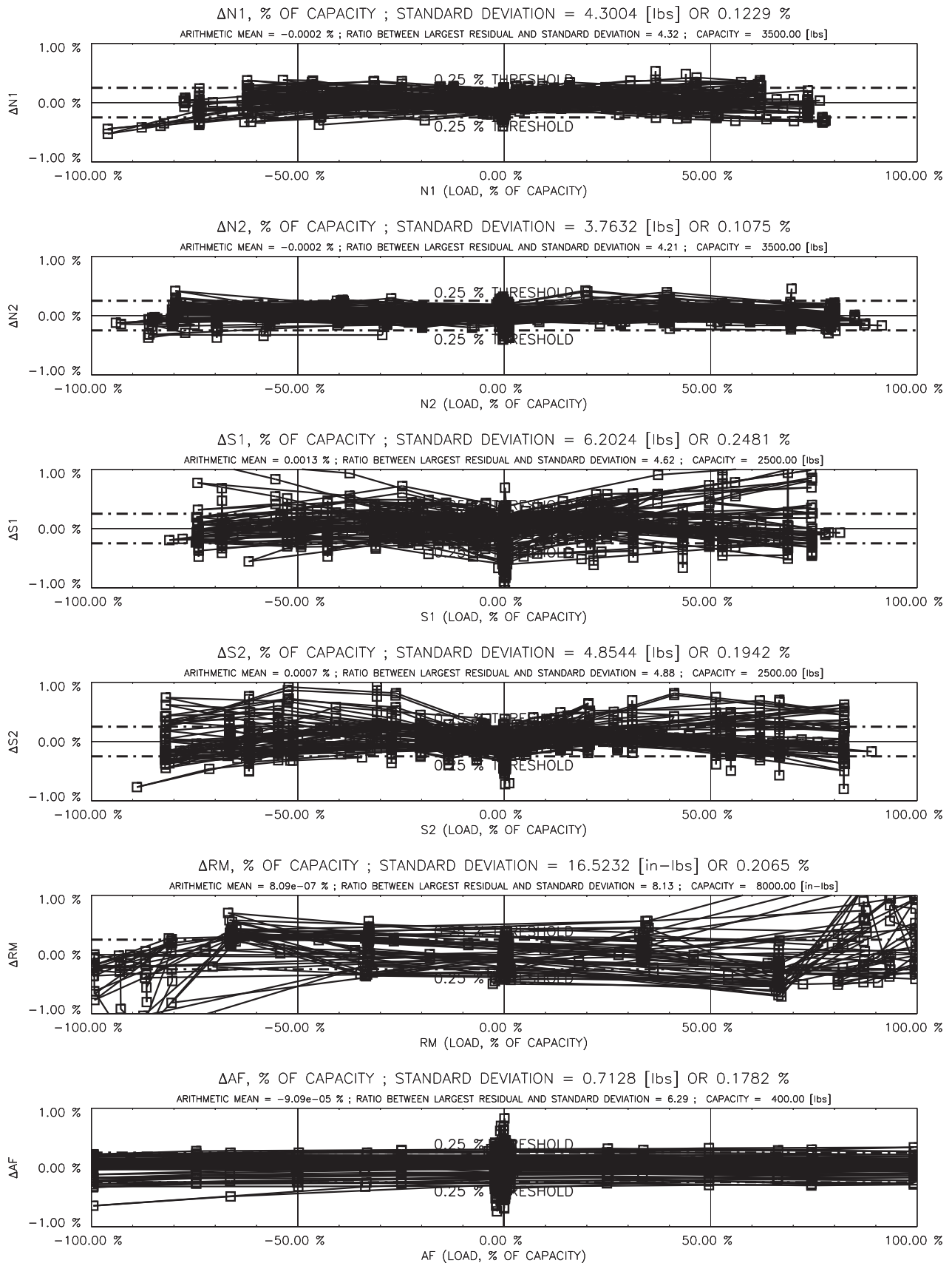
Fig. 5b Load residuals for optimized math model of the <u>simplified</u> search algorithm.

American Institute of Aeronautics and Astronautics