

# Wind Tunnel Force Balance Calibration Study – Interim Results

Ray D. Rhew<sup>1</sup>

*NASA Langley Research Center, Hampton, Virginia, 23681, USA*

Wind tunnel force balance calibration is preformed utilizing a variety of different methods and does not have a direct traceable standard such as standards used for most calibration practices (weights, and voltmeters). These different calibration methods and practices include, but are not limited to, the loading schedule, the load application hardware, manual and automatic systems, re-leveling and non-re-leveling. A study of the balance calibration techniques used by NASA was undertaken to develop metrics for reviewing and comparing results using sample calibrations. The study also includes balances of different designs, single and multi-piece. The calibration systems include, the manual, and the automatic that are provided by NASA and its vendors. The results to date will be presented along with the techniques for comparing the results. In addition, future planned calibrations and investigations based on the results will be provided.

## Nomenclature

<i>N</i>	=	Normal Force component
<i>A</i>	=	Axial Force component
<i>P</i>	=	Pitch Moment component
<i>R</i>	=	Roll Moment component
<i>Y</i>	=	Yaw Moment component
<i>S</i>	=	Side Force component
<i>N*A</i>	=	Normal times Axial model coefficient (similar form for other 2-factor terms)
<i>N^2</i>	=	Normal force squared model coefficient (similar form for other squared terms)

## I. Introduction

Wind tunnel force balances utilized for testing in NASA programs include single and multi-piece designs and are calibrated using several different systems and methodologies. The calibration systems are operated by both NASA and balance vendors. Since a “*standard*” calibration system, such as those maintained by the National Institute for Standards and Technology (NIST) for many measurement or instrumentation systems, does not exist, organizations have developed their own. This study was undertaken to assess the systems in their current operating configurations to answer the fundamental question, “Which systems are adequate to provide calibrations for NASA programs?”. The term adequate does not refer to which systems are better but which ones are acceptable based on the requirements of the wind tunnel users or programs.

Two metrics to consider in evaluating the systems against the fundamental question are:

- the back computed residuals (standard deviation)
- the resultant mathematical models

While the standard deviation of the residuals is a typical metric used in the evaluation of balance calibrations, the author will place more emphasis on the mathematical models in this paper. The reason for the model emphasis is that the user is provided a balance to use for a test along with the math models. If the models are statistically equivalent, the results of applying the model to the wind tunnel data will be essentially the same or within the uncertainty of the calibration.

---

<sup>1</sup> Branch Head, Aeronautics Systems Engineering Branch, MS #238, first author.

## II. Approach

The overall goal of the study is to evaluate the calibration systems for each of the balance types used in NASA programs. The calibration systems include multiple manual and multiple automatic systems. The data presented will not identify the systems until the study is complete. The balances include multi-piece (or task type), single-piece, and hi-capacity designs. Another important element of the study is the load schedule. Since the ability of the calibration data to compute or generate a math model is directly related to the loading schedule, this is an integral part of the evaluation. Ideally the same load schedule would be executed for each balance on each calibration system. However, the calibration systems each have their own unique constraints making this somewhat impractical. Yet, attention needs to be directed to this important piece of the evaluation. Unfortunately for the results presented here, the calibration provider's load schedules were used and limits the evaluation due to the unequal math models that are capable of being computed. This approach does provide examples of the current models the NASA programs are using and the potential differences in models that result. Table 1 shows the matrix of systems and balances to be evaluated. The balances discussed in this report are a single-piece design and a multi-piece Task balance. These balances are of different load capacities and sizes and were calibrated on two and three systems respectively.

Calibration Systems	Balance Design Types		
	Single-Piece Design	Multi-Piece (Task)	Triumph Hi-Cap
NASA Ames ABCM			
Triumph Manual			
NASA Langley Single Vector System (SVS)			
NASA Langley Manual			
Triumph ABCS			

**Table 1. Calibration Study Systems and Balances**

The calibration data for each system and balance, in this interim report, were analyzed using the following approach since the load schedules were not equal.

- A full second order model was computed for each measurement response
- No model reduction was performed based on statistically insignificant terms
- The only terms removed were those with very large (>10) variance inflation factors (VIF, see reference 1 for a discussion on the VIF)
- The analyses were performed using the software package Design Expert
- Only the forward regression models were computed and compared (the models were not inverted as is typically done in production)

References 2 and 3 discuss methods for optimizing the math models for balance calibrations using statistical tools. These techniques will be used in the future analyses to explore the impacts on the results.

### III. Interim Results

Table 2 displays the standard deviation of the back-computed residuals for a single-piece balance across two systems.

	Back Computed Residuals (standard deviation)	
	System 1	System 2
	(% Full-Scale)	(% Full-Scale)
Normal	0.04%	0.03%
Axial	0.14%	0.09%
Pitch	0.06%	0.05%
Roll	0.16%	0.08%
Yaw	0.12%	0.07%
Side	0.05%	0.06%

**Table 2. Single-Piece Balance Back Computed Residuals**

The back computed residuals are within 0.1% full-scale or less when compared across the two systems for the single-piece balance. Based on the residual results, the calibrations appear to be adequate based on historical practices. Data were also acquired for a Task balance. The differences were quite large, >0.5%. Since the Task balance used here typically has larger uncertainties than the single-piece, this observation is expected. However, the author was not expecting the Task balance differences to be quite this large indicating some issue may need to be investigated into the calibration hardware systems and the balance itself. The remaining analyses discussion will focus on the single-piece balance.

Table 3 lists the balance sensitivities (linear model coefficient for that component) and the percent full-scale differences from an average across the systems. This average is used for comparison purposes since, again, the correct answer is unknown. The large differences in the sensitivities, larger than the differences in the back computed residuals, illustrates the first area that needs to be investigated and the importance of reviewing the math models as part of this study. The sensitivities differences are fundamental to the calibration math model and if these do not agree, investigating other differences prior to resolving this one seems impractical. Items such as local gravity, hardware quality assurance measurements and data system settings are some of the fundamental areas that could contribute to these differences. Additional discussion on next steps will be provided in the summary and next steps section.

	Sensitivities (difference)	
	System 1	System 2
	(% Full-Scale)	(% Full-Scale)
Normal	0.09%	-0.09%
Axial	0.04%	-0.04%
Pitch	0.05%	-0.05%
Roll	-0.10%	0.10%
Yaw	-0.03%	0.03%
Side	0.04%	-0.04%

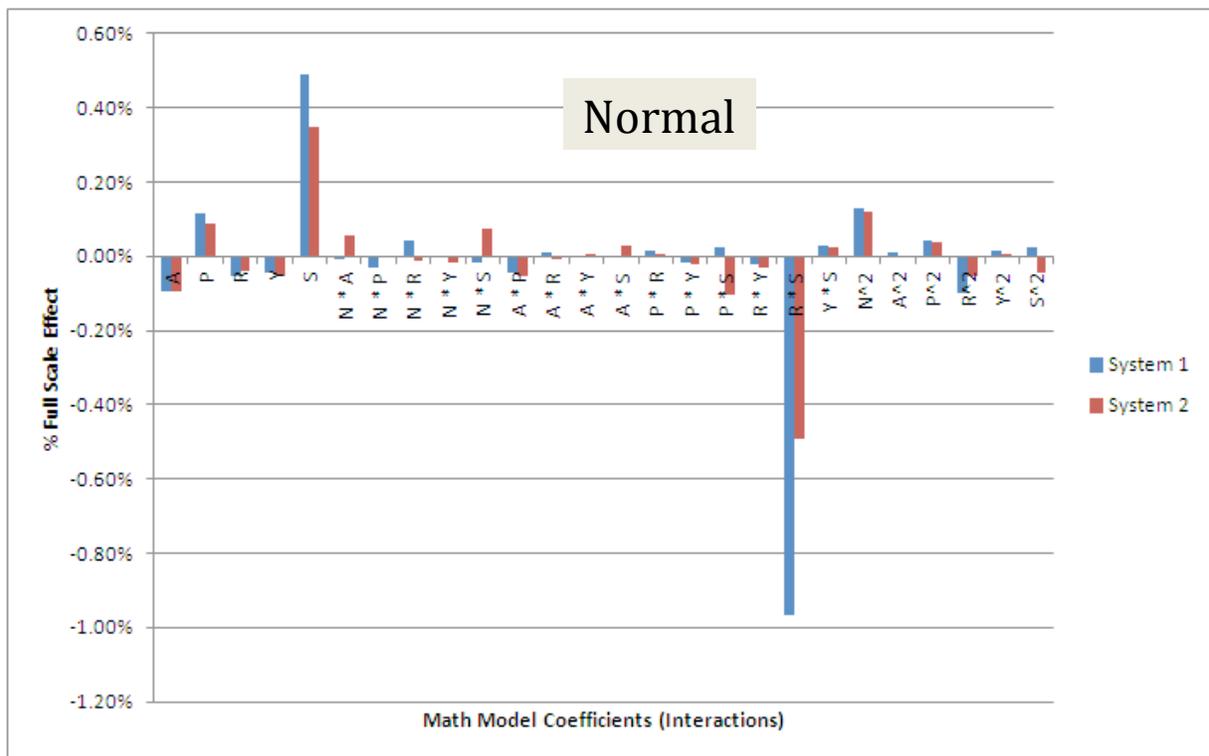
**Table 3. Single-Piece Balance Sensitivity Differences from the Average**

Figures 1a-1d shows a comparison of the math model terms based on the percent full-scale effect. These graphs are to provide insight into the type of balance used for this study and to begin to illustrate where the terms are different for the two system calibrations. The y-axis scale is the percent full-scale effect while the x-axis lists each model term in the full second-order model as listed in Table 4. Excluded from the plot are the component sensitivities.

Designation	Term
N	Normal
A	Axial
P	Pitch
R	Roll
Y	Yaw
S	Side

**Table 4. Math Model Coefficient Designations**

The balance used for this part of the study has significantly large interactions or math model terms, some are on the order of 20% full-scale. Therefore, the balance is very sensitive to off axis loads and a good candidate for exercising the calibration systems.



**Figure 1a. Single-Piece Balance % Full-Scale Math Model Coefficients – Normal Force (Note All y-axis scales are different)**

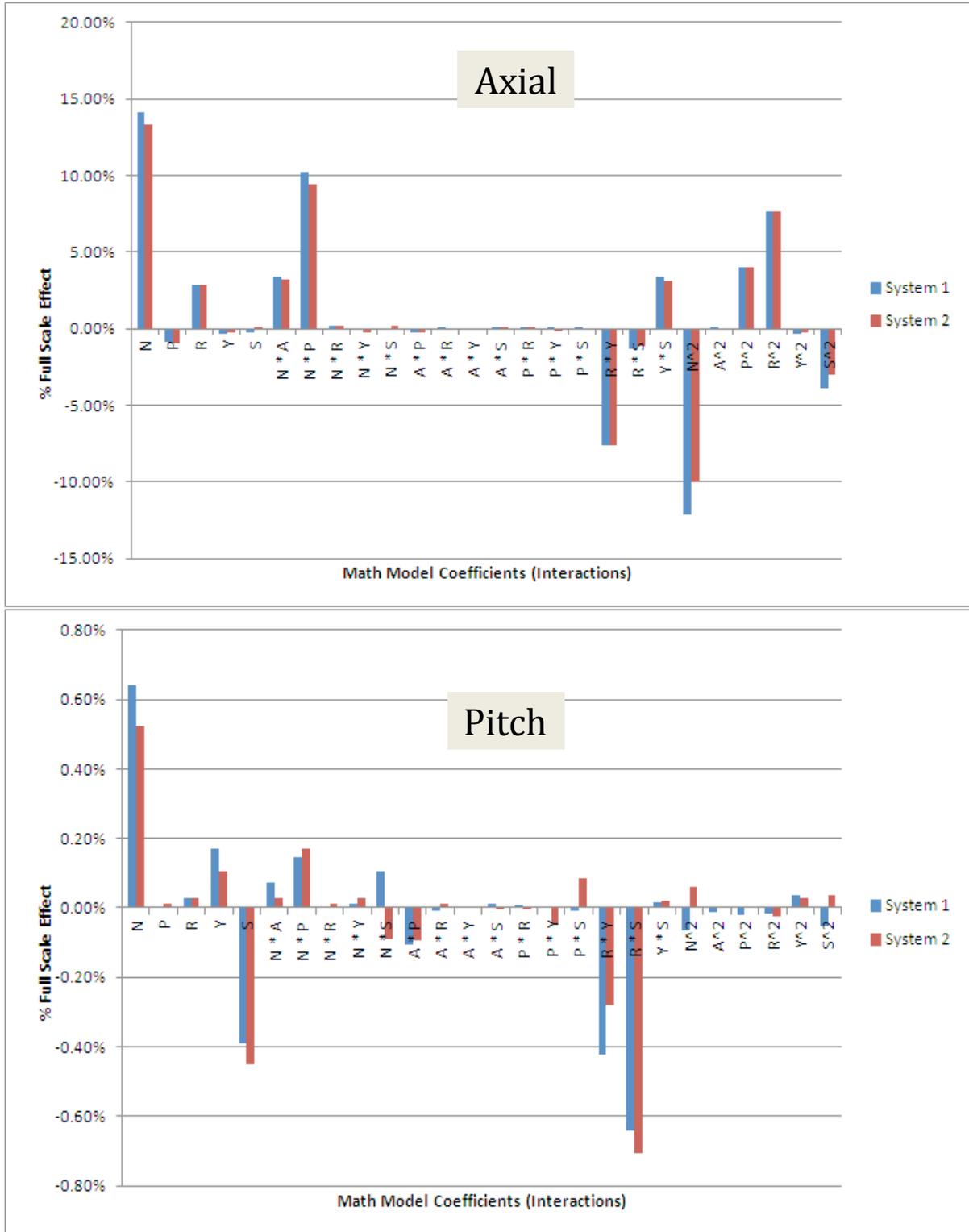


Figure 1b. Single-Piece Balance % Full-Scale Math Model Coefficients – Axial Force and Pitch Moment (Note All y-axis scales are different)

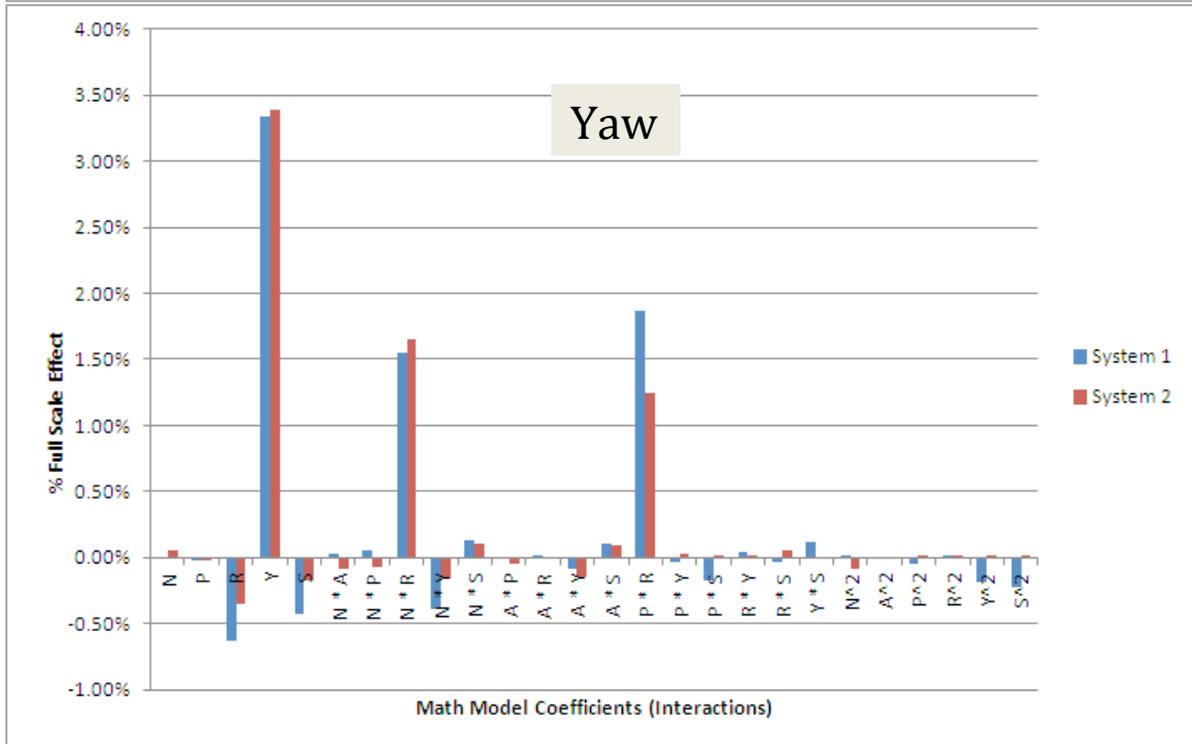
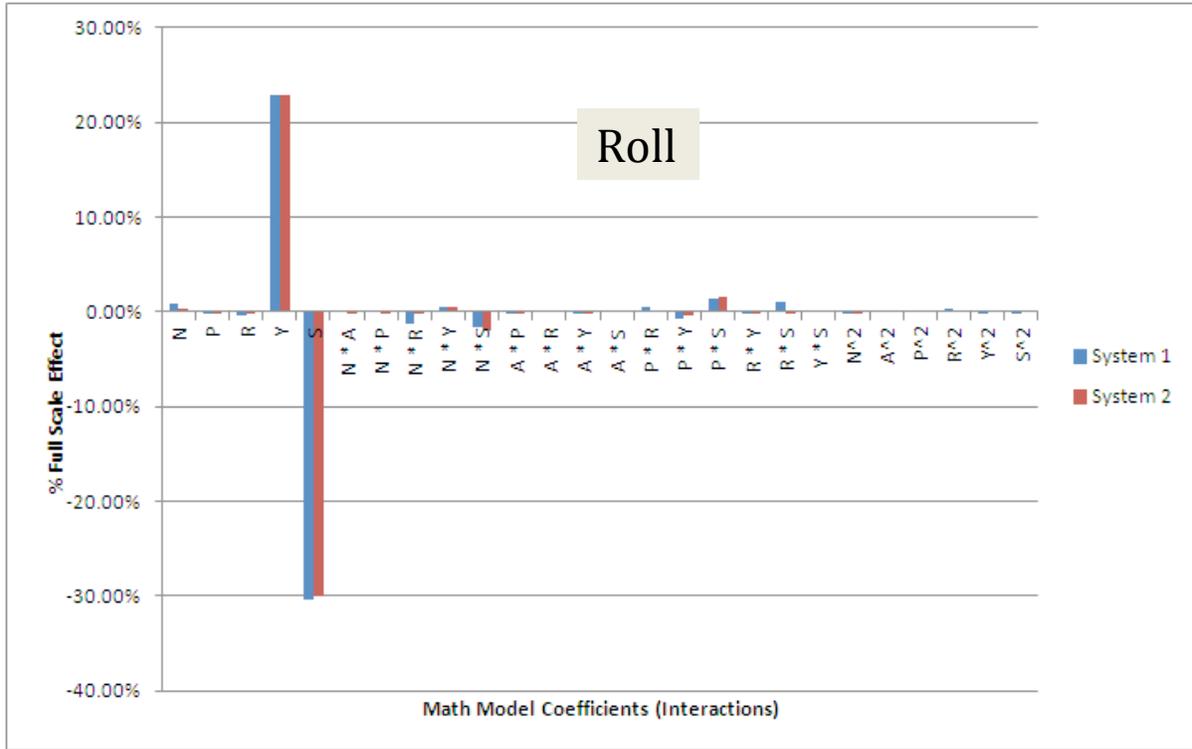
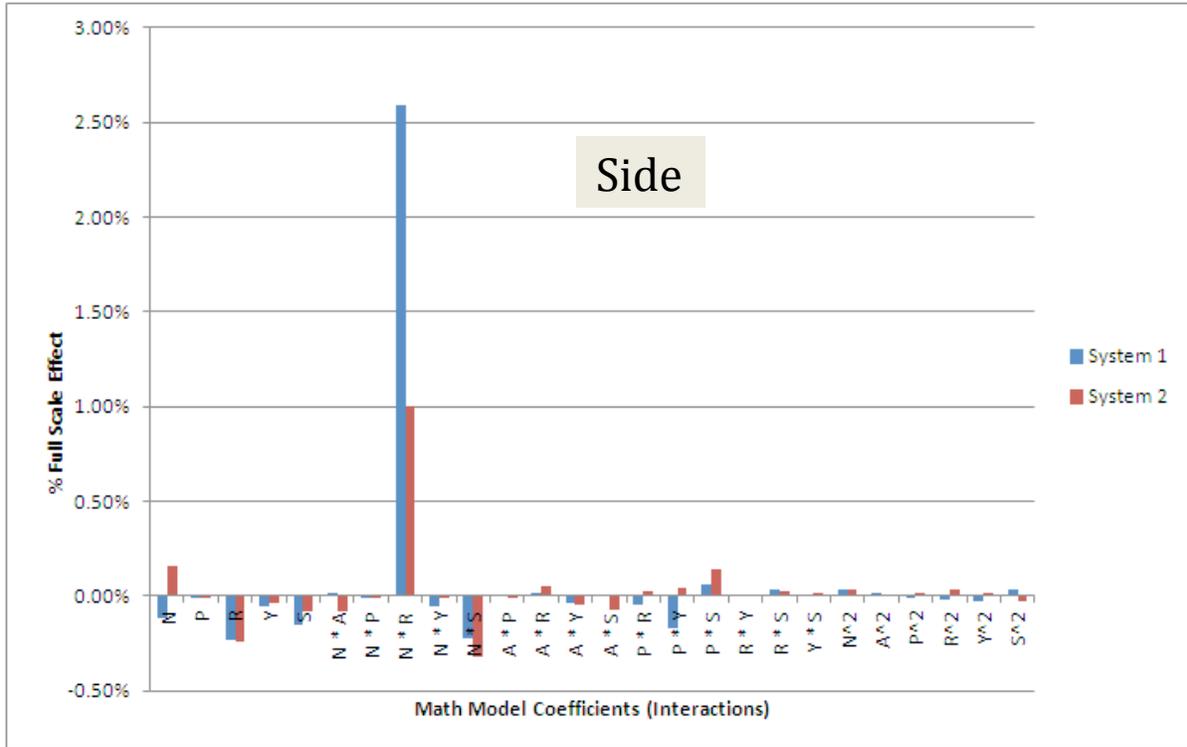


Figure 1c. Single-Piece Balance % Full-Scale Math Model Coefficients – Roll Moment and Yaw Moment (Note All y-axis scales are different)



**Figure 1d. Single-Piece Balance % Full-Scale Math Model Coefficients - Side Force (Note All y-axis scales are different)**

Figures 2a-2c plot the differences or deltas for each component's math model coefficients providing better resolution. The scales are all the same to enable the largest differences to be noted easily. Some differences are on the order of 2.5%. This resolution can be used to start investigating the differences in the calibration systems. Figures 2a-2c also begins to illustrate the differences in loads the two models would compute if applied to the same set of user or wind tunnel data. One typical comparison approach is to cross-apply the math models from one calibration to the calibration data from another. While this approach provides insights into the math model differences, it also tends to favor the model that was computed from the data. A different approach is discussed in the next paragraph.

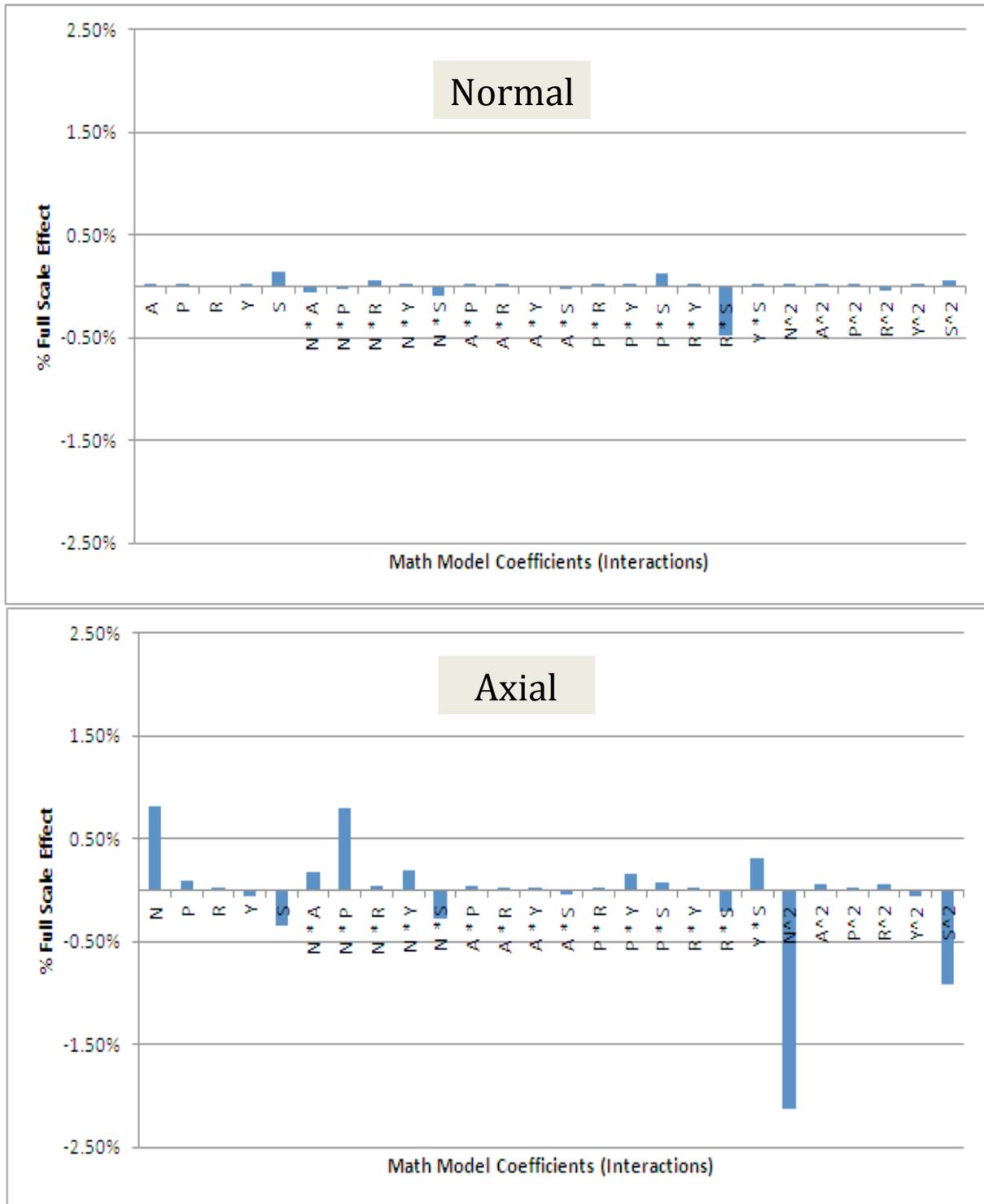


Figure 2a. Single-Piece Balance % Full-Scale Math Model Coefficient Differences – Normal Force and Axial Force (Note All y-axis scales are the same)

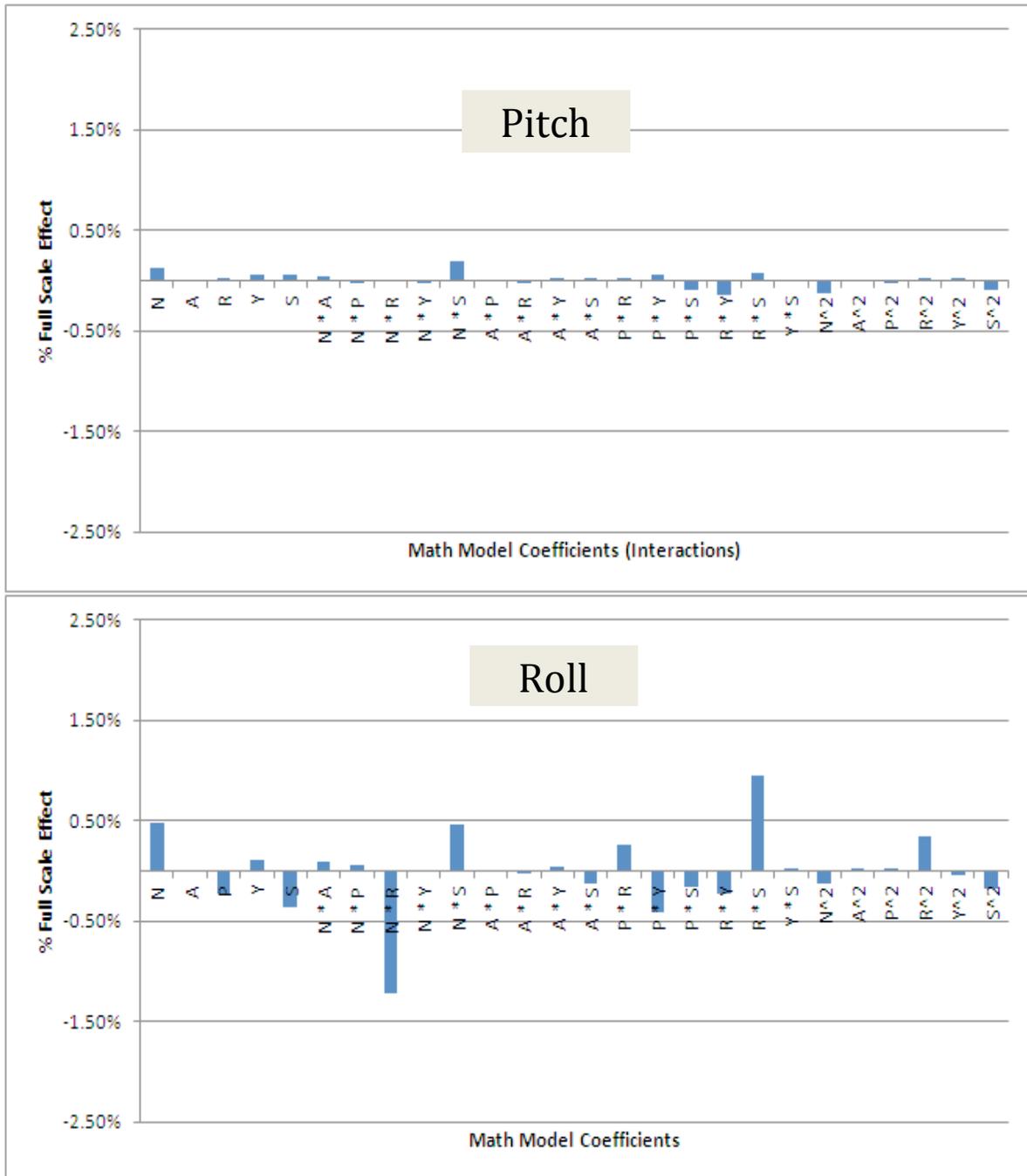
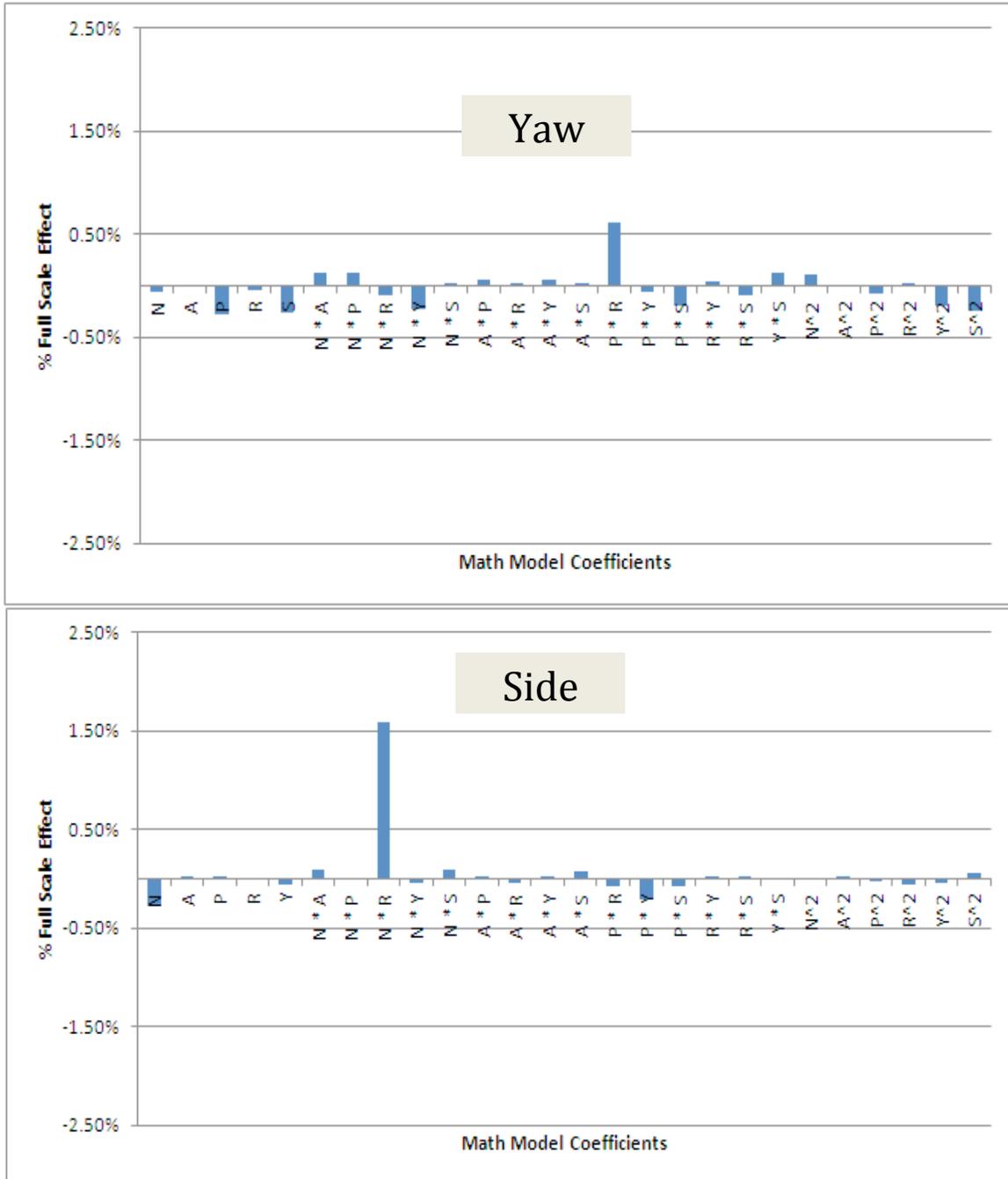


Figure 2b. Single-Piece Balance % Full-Scale Math Model Coefficient Differences – Pitch Moment and Roll Moment (Note All y-axis scales are the same)



**Figure 2c. Single-Piece Balance % Full-Scale Math Model Coefficient Differences – Yaw Moment and Side Force (Note All y-axis scales are the same)**

Figures 3a-3c show 3-D surface plots to illustrate the differences in the two models computed from the two calibration systems. The lower axes are from +1, representing positive full-scale of that component, to -1, or negative full-scale. The vertical axis is the component under review and is the % full-scale difference. The legend displays the load levels of the other components for each case. All six of the components can be varied to study the impact on the differences. Additionally, depending on the location of the users data within the load space, the % difference can vary dramatically. While neither is known to be better than the other, this approach is illustrating what the user would experience if they analyzed the same set of wind tunnel data with the two different math models for the same balance. In some cases the % full-scale differences are as large as 3% that is an order of magnitude different than any of the quoted back computed residuals.

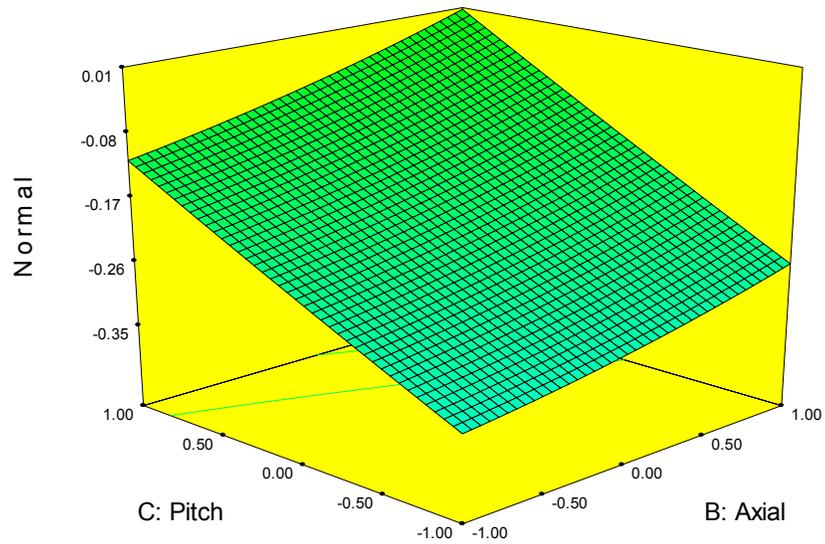
Design-Expert® Software

Normal



X1 = B: Axial  
X2 = C: Pitch

Actual Factors  
A: Normal = -1.00  
D: Roll = 0.50  
E: Yaw = 0.50  
F: Side = 0.50



**Figure 3a. Normal Force % Full-Scale Difference Surface Plot**

In figure 3a with Normal on the vertical axis, the surface is a plane indicating that the differences for this part of the load space are dominated by linear effects. While the plots for Axial and Pitch (figures 3b and 3c) are dominated by non-linear terms evident by the shape of the surface. The most dramatic percent difference is for axial and is expected since the model term differences are largest for this component. However, the plot begins to show how large these can be depending on where the users data is acquired.

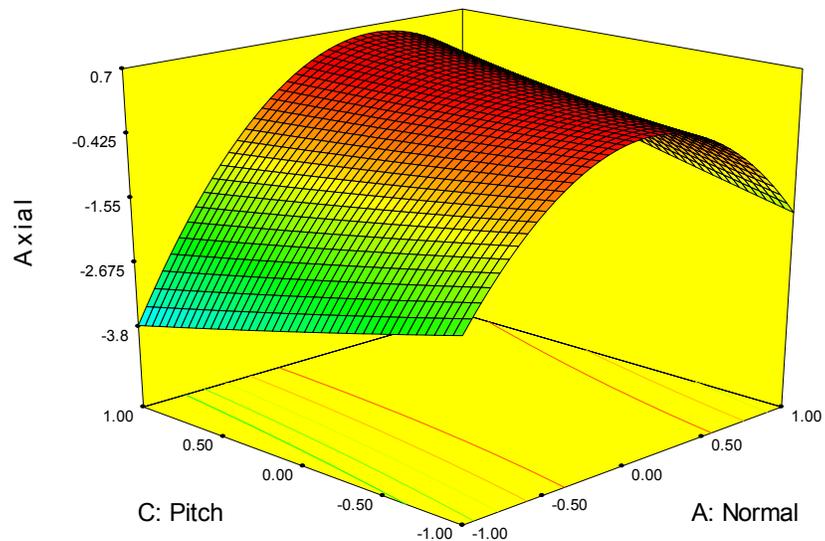
Design-Expert® Software

Axial



X1 = A: Normal  
X2 = C: Pitch

Actual Factors  
B: Axial = 1.00  
D: Roll = -0.50  
E: Yaw = -0.50  
F: Side = -0.50



**Figure 3b. Axial Force % Full Scale Difference Surface Plot**

Design-Expert® Software

Pitch  
0.402802  
-0.839307

X1 = A: Normal  
X2 = B: Axial

Actual Factors  
C: Pitch = 1.00  
D: Roll = 0.00  
E: Yaw = 0.00  
F: Side = 0.00

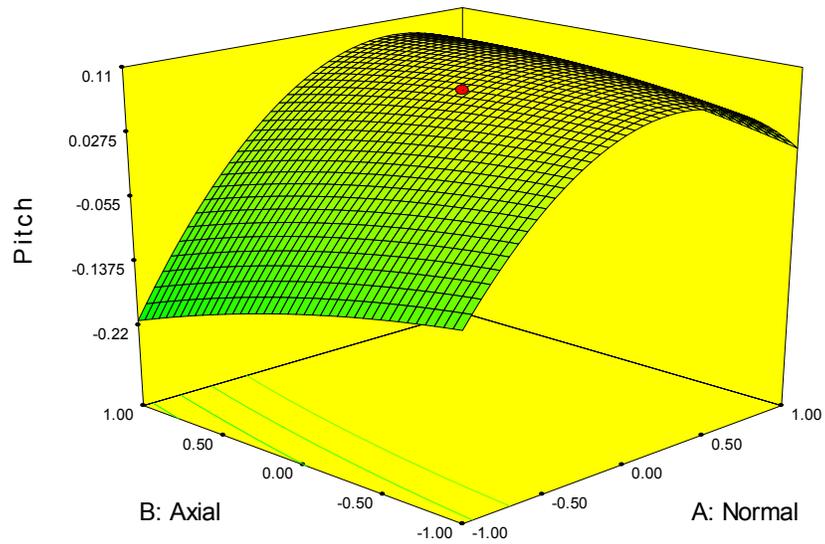


Figure 3c. Pitch Moment % Full-Scale Difference Surface Plot

#### IV. Summary and Next Steps

In summary, this paper presents the status of the calibration system study undertaken by NASA. To date, calibrations have been performed on two balance types and across three calibration systems. However, the data presented here is for the single-piece balance only due to the concerns on the Task balance data acquired. The analyses focuses on the math models generated from each calibration since these are the deliverables to the user. Reviewing the model coefficient differences in percent full-scale begins to provide some knowledge on the size and possible impact. Further, the surface plots are a new way to view the differences across the load space and with respect to what a user would experience if they analyzed their wind tunnel data with both math models for the same balance and test. The surface plots also help to illustrate the dominant differences within each component and areas to investigate why the systems provide different results.

The next steps for this study are outlined below and attempt to provide more structure to this effort towards answering the fundamental question posed at the beginning, “Which systems are adequate to provide calibrations for NASA programs?”.

- Design an experimental load schedule to assess the results consistently (same load schedule for part of the calibration and add points unique to a system where appropriate)
- Develop fundamental experiments for the weights, data systems, dimensional inspections
- Add statistical tools to the analyses plan to provide additional insight such as confidence and prediction intervals to assess if, statistically, model coefficients are different
- Provide a comprehensive report on the results

#### Acknowledgments

The Aeronautics Test Program (ATP) through the National Force Measurement Technology Capability (NFMTC) project funded this study and is greatly appreciated by the author. Triumph Force Measurement Systems, Modern Machine and Tool Co., NASA Ames and members of the Langley NFMTC team all provided calibration data and input to the study and will be part of the next steps.

## References

<sup>1</sup>Myers, Raymond H., and Montgomery, Douglas C., *Response Surface Methodology*, 2<sup>nd</sup> ed., John Wiley and Sons, New York, 2002, Appendix 2.

<sup>2</sup>Ulbrich, N., and Volden, T., "Regression Model Term Selection for the Analysis of Strain-Gage Balance Calibration Data," AIAA-2010-4545, 2010.

<sup>3</sup>Ulbrich, N., and Volden, T., "Regression Analysis of Experimental Data Using an Improved Math Model Search Algorithm," AIAA-2008-0833, 2008.