ESTIMATION OF SUPERSONIC STAGE SEPARATION AERODYNAMICS OF WINGED-BODY LAUNCH VEHICLES USING RESPONSE SURFACE METHODS

Gary E. Erickson and Richard DeLoach
NASA Langley Research Center

Keywords: *response surface methods, stage separation aerodynamics, multi-stage launch vehicles, supersonic speeds, shock waves, wind tunnel testing*

Abstract

A collection of statistical and mathematical techniques referred to as response surface methodology was used to estimate the longitudinal stage separation aerodynamic characteristics of a generic, bimese, winged multi-stage launch vehicle configuration using data obtained on small-scale models at supersonic speeds in the NASA Langley Research Center Unitary Plan Wind Tunnel. The simulated Mach 3 staging was dominated by multiple shock wave interactions between the orbiter and booster vehicles throughout the relative spatial locations of interest. This motivated a partitioning of the overall inference space into several contiguous regions within which the separation aerodynamics were presumed to be well-behaved and estimable using cuboidal and spherical central composite designs capable of fitting full second-order response functions. The primary goal was to approximate the underlying overall aerodynamic response surfaces of the booster vehicle in belly-to-belly proximity to the orbiter vehicle using relatively simple, lower-order polynomial functions that were piecewise-continuous across the full independent variable ranges of interest. The quality of fit and prediction capabilities of the empirical models were assessed in detail, and the issue of subspace boundary discontinuities was addressed. The potential benefits of augmenting the central composite designs to full third-order using computer-generated D-optimality criteria were also evaluated. The usefulness of central composite designs, the subspace sizing, and the practicality of fitting lower-order response functions over a partitioned inference space dominated by highly nonlinear and possibly discontinuous shock-induced aerodynamics are discussed.

1 General Introduction

A general discussion of response surface methodology (RSM) and the motivation for its use in aerospace research are presented in this section. The specific application of RSM to wind tunnel testing and the experimental design modifications that were adopted in the current investigation are also discussed.

1.1 What is RSM?

RSM is a family of statistical and mathematical techniques by which response variables can be represented as empirical functions of the independent variables that influence them [1]. The term RSM is attributed to the graphical perspective of the problem environment, in which a response is often plotted as a surface constructed over a plane defined by selected independent variables oriented in a Cartesian coordinate system. An example of a response surface obtained in a low-speed wind tunnel test of a military aircraft configuration is shown in figure 1, which depicts the aerodynamic drag coefficient response surface plotted as a function of the horizontal tail incidence angle and the vehicle angle of attack. In most scientific applications, the underlying relationship between a system response and

selected independent variables is unknown. Consequently, experimental data are required to fit an appropriate mathematical model that estimates the true but unknown response function. Most RSM applications involve multiple independent variables and, as a result, there is a close connection between RSM and regression analysis, which is a branch of statistical model building. The true underlying relationship is often suitably approximated for each response variable of interest with a low-order Taylor series, or polynomial, model applied to a relatively small region of the independent variable space. A first-order polynomial function in three independent variables (or factors in RSM terminology) can be expressed as:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{23} x_2 x_3 + \varepsilon \quad (1)$$

The response variable, y, is expressed as a function of the three main factor effects (x_1, x_2, and x_3), three two-factor interactions (x_1x_2, x_1x_3, and x_2x_3), and a random error component (ε). The acquisition of experimental data and the application of multiple linear regression techniques [2] featuring the method of least squares yield unbiased estimates of the unknown parameters (β's) and the random error in equation (1). The interaction terms represent first-order curvature in the system response, and their inclusion in the model indicates that the effect of one factor on the response depends on the level to which the other factor is set. In this example, the independent variable space, typically referred to in RSM as the inference space or design space, is defined as a volume (or a cube when the variables are coded using a linear transformation in the experimental design). The response surface resides "above" the inference space in a fourth dimension, which is not easy to visualize graphically. If the curvature in the true response surface is of a higher order, the polynomial function can be augmented to include second-order curvature as depicted in the full quadratic model:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{23} x_2 x_3 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \beta_{33} x_3^2 + \varepsilon \quad (2)$$

The second-order model is very flexible and is an effective approximation to the true response surface in many applications. The model in equation (2) is a linear function of the unknown β's, hence, it is still a linear regression model regardless of the shape of the response surface that it generates. The quality of the regression model is critically dependent on the manner in which the data are collected, and RSM employs many quality assurance tactics and statistical experimental designs to produce high-quality empirical functions. A family of experimental designs referred to as Central Composite Designs (CCD) [1] is used extensively in building second-order response surface models. The first- and second-order models are often very effective in providing relatively simple and interpretable empirical response functions in many industrial and aerospace applications.

1.2 Why Use RSM?

RSM is an alternative to the data-intensive one-factor-at-a-time (OFAT) experimental method [3] in which all factors are typically held constant in an experiment except for one in an attempt to identify the relationship between the response and independent variables. The OFAT approach is vulnerable to the effects of systematic variation that occur in experimental investigations, and it does not allow a reliable estimation of interactions that are frequently a key to understanding the system response. RSM is based on well-established statistical and mathematical techniques that are essential to the design and execution of the experiment and the analysis of the results. The pseudo three-dimensional display of response surfaces relative to the inference space (see figure 1, for example) provides a more global perspective of the aerodynamic responses. Estimates of the mean response can be obtained for any combination of the independent variables, including those combinations that were not physically set during the experiment. Each point in the experimental data set used to fit the regression model contributes in some degree to the estimates of the response and the uncertainty at that point. Furthermore, every data point beyond some minimum threshold contributes to

the uncertainty assessment at every location in the inference space. Certain quality assurance tactics such as replication, randomization, and blocking [3] can be invoked to defend against systematic experimental variation, thereby increasing the precision and reproducibility of the experiment. Testing of all combinations of the independent variables allows the estimation of interactions, which are often the key to interpreting the responses obtained in a complex aerodynamic environment. Furthermore, it is not necessary to know the mathematical form of the response surface a priori in order to fit it with experimental data, since a Taylor series model provides an approximation of the unknown relationship via regression analysis. The successful application of RSM requires careful consideration of many issues. The knowledge that is expected to be gained from an experiment is clearly defined during pretest planning. Obviously, the aerodynamic responses, independent variables, and variable ranges of interest must be identified. The desired experimental precision and resolution should be specified, which determines the minimum volume of data that must be acquired. The wind tunnel facility, test article, and instrumentation capabilities, the testing environment, and the available resources significantly influence the design and execution of an experiment. Furthermore, the expected aerodynamic response behavior within the inference space of interest will affect the experimental design, execution, model building, and analysis, particularly if the aerodynamic flow field exhibits different physical mechanisms throughout the inference space (e.g., linear versus nonlinear, symmetric versus asymmetric, shock-free versus shock-dominated, etc.).

1.3 Specific Application of RSM

The statistically-designed experimental approach of RSM was used in an exploratory wind tunnel test in the NASA Langley Research Center (NASA LaRC) Unitary Plan Wind Tunnel (UPWT) of a generic bimese wing-body launch vehicle configuration to develop empirical models of the stage separation aerodynamic forces and moments at supersonic speeds. This work was performed in support of on-going NASA stage separation and ascent aerodynamics and aerothermodynamics research programs focusing on future two-stage-to-orbit (TSTO), fully reusable launch vehicle (RLV) systems [4]. A generic TSTO configuration developed by NASA LaRC and referred to as the Langley Glideback Booster (LGBB) was selected as the test article. The LGBB is a wing-body configuration, and it was tested in a bimese, belly-to-belly arrangement consisting of an orbiter and a booster as illustrated in figure 2. A bimese TSTO vehicle is one in which both the booster and the orbiter have the same outer-mold-lines, that is, identical external geometries. The proximity aerodynamics of multi-stage launch vehicle systems are highly configuration-dependent [4]. The LGBB configuration was considered a worst-case in wing-to-wing proximity effects, known to provide significant aerodynamic interference, which was useful in assessing the stage separation experimental techniques and computational analysis tools that were applied in these in-house programs.

The current application assumes a Mach 3 staging, where the orbiter is at full thrust at separation, the booster is at no thrust, and the booster glides back to the launch site. However, the wind tunnel testing was performed with power-off. No sideslip or asymmetric separation conditions were tested, so the longitudinal aerodynamic forces and moments were the focus of this investigation. The region within which supersonic flow was expected to persist during separation is relatively large, and the aerodynamic interference between the orbiter and booster vehicles in this region is dominated by multiple shock wave development and interactions [4]. These phenomena can introduce highly nonlinear and, possibly, discontinuous, changes in the stage separation aerodynamics, particularly for the booster vehicle, which remains in the influence of the orbiter throughout the region of interest [4]. For example, the pressure field about the booster vehicle changes significantly as a result of the impingement of the orbiter bow shock. The corresponding large changes in the booster lift,

drag, and pitching moment characteristics are dependent on the relative spatial location and attitude of the two vehicles. Given these conditions, a practical application of response surface methods in a wind tunnel experiment at Mach 3 was formulated. The centerpiece of this RSM application was a central composite design capable of fitting a full second-order model and certain mixed cubic terms to each of six response variables (lift, drag, and pitching moment coefficients for the orbiter and booster vehicles) for ranges of three independent variables (relative axial and vertical positions and relative angle of attack). The sequential nature of RSM easily permitted design augmentation to full third-order by acquiring additional data, as necessary. However, the expectation of highly nonlinear aerodynamics over a broad inference space led to modifications of the experimental design as outlined in the next section.

1.4 Truncated Inference Subspaces

A disadvantage of applying a single CCD to nonlinear responses over a broad inference space is that important features of the aerodynamic responses may not be effectively resolved. There is an even greater concern if different flow phenomena are present depending on the location within the inference space. If the underlying response functions are not adequately estimated with a second-order model, the CCD is capable of adding a limited number of higher-order terms. Design augmentation will allow fitting even higher-order models to the experimental data. The danger of this approach, however, is that the higher-order response functions can behave unrealistically between points and immediately outside of the inference space, and they have higher prediction variance, since the average prediction variance within the inference space is proportional to the number of model coefficients [5]. In many cases, the response surfaces are forced to fit the experimental data points exactly and are, therefore, poor predictors of the mean responses at any other combinations of the independent variables for which data were not acquired. An alternative to adding additional higher-order terms when the response is a complex function of the independent variables is to trim the inference space [6]. Specifically, the overall inference space can be partitioned into multiple smaller contiguous regions within which even low-order response functions may be adequate, *if the ranges of the independent variables are sufficiently limited*. As a result, piecewise continuous models of the overall responses can be developed over the full ranges of the independent variables. In principle, this method permits relatively simple and well-behaved functions to be fitted locally and the resolution of different physical phenomena that may prevail throughout the inference space. A key to the successful application of this approach is to define the most expansive subspace ranges possible within which relatively low-order polynomial models will adequately estimate the true underlying responses. The use of truncated inference subspaces was the final element in the current investigation to estimate the supersonic stage separation aerodynamics of a bimese, like-scale, winged vehicle configuration using statistically-designed experimental methods.

2 Previous Work

2.1 RSM Applications to Wind Tunnel Testing

RSM techniques have not been widely exploited in experimental aeronautics, although its use has increased in recent years as researchers have been more exposed to the discipline of experimental design through formal education and practical applications. RSM is a key component of a unified design, execution, and analysis process of scientific research under development at NASA LaRC known as the Modern Design of Experiments (MDOE) [6]. An exhaustive review of MDOE applications in wind tunnel experiments is not within the scope of this paper. However, the following references are examples of the versatility of statistically-designed experiments in a diverse range of aerospace investigations: aeroelastic deformation on slender wings [7]; control

surface effects on commercial and military aircraft [8]; assessment of aerodynamic intrusiveness of optical measurement techniques [9]; wind tunnel calibrations [10]; temperature measurements in supersonic combusters [11]; aerodynamics of micro air vehicles with flexible wings [12]; optimization of vortex flow control devices for tail buffet mitigation [13]; noise reduction techniques for high bypass ratio turbo-fan engines [14]; proximity aerodynamics of an airborne-launched hypersonic research vehicle [15]; and laboratory calibrations of wind tunnel model instrumentation [16]. Of particular relevance to the current RMS application to stage separation aerodynamics are the following two investigations. A hybrid experimental design was developed in reference [17] to estimate the low-speed aerodynamic characteristics of a high-performance military aircraft over a large inference space, which was partitioned into a low angle-of-attack, linear portion of the design space and a high angle-of-attack, nonlinear region characterized by stall/post-stall aerodynamics. A nested, 5-level face-centered design was applied in each subspace and was developed specifically for this test to estimate expected higher-order control surface effects. The method of truncated inference subspaces was also applied to estimate the supersonic aerodynamic characteristics of a slender-wing fighter configuration [6]. Similar to the experiment in reference [17], the goal was to permit separate empirical models to be developed for regions of the inference space in which different physical phenomena were expected. The two-dimensional inference space in this experiment encompassed large ranges of the angles of attack and sideslip. Linear aerodynamics were expected at low angles of attack and sideslip, whereas nonlinear, symmetric and asymmetric, vortex-dominated aerodynamics were anticipated at other combinations of the independent variables. The overall inference space was partitioned into a total of six subregions. The sideslip axes were both divided into negative and positive regions, and the positive region for angle of attack was further divided into low and high segments. A spherical central composite design [1] was applied in all subspaces, which provided piecewise-continuous response models for the six-component aerodynamic forces and moments over the full ranges of the independent variables. This investigation included a detailed assessment of methods to assess the adequacy of the empirical models, including predictions at the subspace boundaries.

2.2 Stage Separation Wind Tunnel Testing

Studies of the aerodynamics of separating bodies have been performed for decades [4]. In the 1980s, post-Shuttle development, NASA continued to examine earth-to-orbit space transportation concepts to cover a wide range of mission requirements, which included two-stage, fully reusable systems. An experimental program was initiated by NASA in the early 2000s to develop experimental tools and testing methodologies to apply to supersonic stage separation problems for future multi-stage launch vehicle systems [4]. This program featured supersonic proximity testing in the NASA LaRC UPWT of the generic, bimese TSTO configuration previously shown in figure 2. Specially designed stage separation support hardware and specifically-scaled models were fabricated to accomplish this testing. Extensive modifications to the UPWT automated model control system algorithms and new support system calibration procedures were implemented to enable a new two-model proximity testing capability in this facility. Proximity data were acquired within a two-dimensional grid of spatial locations representing relative axial and vertical positions of the orbiter and booster vehicles during the supersonic portion of the staging event. Proximity data were obtained at a total of 328 grid points, which are illustrated in figure 3. The origin of the proximity matrix corresponds to the location of the booster model moment reference center (MRC) when the booster and orbiter model are aligned in a near-docking position. The axial and vertical displacements of the booster model expressed in inches, Δx and Δz, respectively, are defined relative to this origin, while the orbiter model remained in a fixed position. The quantities X/L and Z/L are also shown in figure 3, which correspond to the

relative axial and vertical positions nondimensionalized by the model reference length, *L*. An OFAT testing technique was used in reference [4], whereby data were acquired in a "sweep" of axial positions of the booster vehicle while all other independent variables of interest were held constant. Figure 4 shows the variation of the booster model normal force and pitching moment coefficients plotted against the axial displacement, Δx, at fixed relative vertical position (Δz = 1.313 inches) and relative angle of attack ($\Delta\alpha$ = 0 degrees) at Mach = 3.0. Interference increments are seen to be highly nonlinear, changing signs several times as the booster moves aft of the orbiter (increasing values of Δx). It was speculated during this experiment that the flow field region was dominated by multiple shock reflections and flow separation, which contributed to the nonlinear aerodynamics. This was confirmed in the schlieren flow visualization images shown in figure 5 obtained in a complementary test program in the 14-Inch Tri-Sonic Tunnel at the NASA Marshall Space Flight Center (NASA MSFC) [18]. In the mated conditions, mutual interference is characterized by a channel-like flow between the bodies, and the bow shock waves of each body impinge on the other resulting in multiple reflections. As the two bodies move a short distance apart in the axial and vertical directions, the channel-like flow is not observed. Instead, mutual interference is mainly determined by bow shock impingements and the reflections. The results from the UPWT investigation in reference [4] showed that the booster remained in the influence of the orbiter throughout the entire proximity matrix at Mach = 3.0. The ability to leverage off of this work was the primary driver in the application of formal experimental design to estimate the stage separation aerodynamics of like-scale bodies in a belly-to-belly arrangement. The Mach 3 staging conditions, the OFAT proximity matrix and measured interference increments, and the schlieren flow visualization results provided guidance in the definition of the formal experimental design and the partitioning of the overall inference space.

3 Test Information

This section will describe the wind tunnel facility and test conditions, the models and instrumentation, and the test techniques relevant to the current investigation.

3.1 Wind Tunnel Facility and Test Conditions

The NASA LaRC UPWT is a closed-circuit, continuous-flow, pressure tunnel with two test sections that are nominally 4-feet by 4-feet in cross section and 7 feet long. The nozzle throat-to-test section area ratio is varied by a lower asymmetric sliding nozzle block that provides continuous variation of the Mach number. The Mach number range is nominally 1.5 to 2.86 in Test Section 1 and 2.3 to 4.63 in Test Section 2. Reynolds numbers from 1.0 to 5.0 million per foot can be routinely run with a capability to attain 6.0 million per foot on a transient basis. The basic model support mechanism is a horizontal wall-mounted strut that is capable of forward and aft travel of over 3 feet in the streamwise direction. A main sting support attached to the strut can traverse laterally +/-20 inches and can provide yaw capability of +/-12 degrees. Forward of the main sting support is the angle-of-attack mechanism that provides pitch motion from -15 degrees to +30 degrees. A roll mechanism can be installed ahead of the pitch mechanism to provide continuous roll motion over a nearly 360-degree range.

Proximity testing of the LGBB configuration was accomplished by installing the booster model onto the main sting support using a standard straight sting and the orbiter model onto a fixed horizontal blade strut mounted to the test section sidewall. Figure 6 shows photographs of the LGBB configuration installed in the wind tunnel test section. The traverse and rotational capabilities of the main support system allowed positioning of the booster model in the test section to within +/-0.050 inches of the relative axial and vertical distance setpoints and to within +/-0.1 degrees of the relative angle of attack setpoints. The custom horizontal blade strut was mounted to the sidewall because of constraints in attaching

support hardware to the ceiling or floor of the test section. The standard sidewall with multiple optical-quality windows was replaced with a solid sidewall in order to install the blade strut. This precluded the use of the facility schlieren flow visualization system during the stage separation testing. The fixed horizontal blade strut positioned the orbiter model centerline 18 inches from the test section sidewall at pitch and yaw angles of 0 degrees and a roll angle of -90 degrees (wings vertical). A slot was machined into the lee side of the orbiter in place of the vertical tail to accommodate the blade strut entry into the model. The booster model was rolled +90 degrees to a wings-vertical orientation to simulate a belly-to-belly arrangement of this bimese TSTO configuration.

The experiment was conducted in Test Section 2 at a free-stream Mach number of 3.0, a Reynolds number per foot of 2.0 million, and a stagnation temperature of 125 degrees Fahrenheit. The three independent variables were the relative axial distance, Δx, relative vertical distance, Δz, and relative angle of attack $\Delta\alpha$ of the orbiter and booster models. Changes in the independent variables were accomplished exclusively through the translation and rotation of the booster model relative to the fixed orbiter model. A near-docking position of the orbiter and booster models was previously shown in the OFAT proximity grid point matrix in figure 3, which corresponded to $\Delta x = \Delta z = 0$ inches. At this position, the models were axially aligned, and the minimum vertical spacing between the bellies of each model was 0.41 inches. This spacing was established from static and wind-on verification testing as a reasonable buffer to account for small-amplitude model dynamics and booster model angles of attack up to 2.5 degrees during the OFAT testing. A more conservative spacing of approximately 1.066 inches was established for the current response surface investigation corresponding to $\Delta z = 0.656$ inches, since the maximum booster model angle of attack was increased to 5 degrees. The relative axial distance, Δx, was varied from 0 inches to 28 inches, which corresponded to a range of approximately 2.13 model body lengths. The relative vertical distance, Δz, was varied from 0.656 inches to 13.13 inches, which represented a range of approximately 0.95 body lengths. The range of $\Delta\alpha$ was 0 degrees to 5 degrees. The positioning of the booster model in the test section was corrected for deflections due to aerodynamic loads. Flow angularity corrections were not applied to the data, since estimates of the local flow angle throughout the test section were not obtained in this experiment. A brief discussion of possible flow angle gradients in UPWT Test Section 2 and the impact on stage separation aerodynamic measurements is provided in reference [4].

3.2 Models and Instrumentation

Two 0.0175-scale stainless steel force models of the LGBB configuration were fabricated for the stage separation testing in the NASA LaRC UPWT. The reference dimensions used to calculate the aerodynamic coefficient data are included in the sketch of the LGBB in figure 2. The MRC of each model was located at 68% of the distance along the reference body length measured from the nose. The models installed in UPWT Test Section 2 were previously shown in figure 6. The external geometries of the booster and orbiter models were identical with the exception of the vertical tail arrangement. The booster model accommodated a single centerline vertical tail, whereas the orbiter model accepted the blade strut in place of the vertical tail.

The six-component aerodynamic forces and moments were measured simultaneously on both models using internally-mounted strain-gage balances. Both balances were calibrated using a single-vector calibration system [16] that applied MDOE to the calibration and data analysis processes. Model base/chamber pressures were acquired, but no corrections based on these measurements were applied to the balance forces and moments.

Grit particles were applied to both models to encourage boundary layer transition. A 0.10-inch wide ring consisting of #40 grit was applied 1.2 inches aft of the nose. Strips of 0.10-inch width #40 grit were also applied 0.28

inches aft of the wing leading edges. A similar grit strip was applied at the same distance from the leading edge of the booster model vertical tail.

3.3 Test Techniques

Modifications were made to the facility calibration procedures and the automated model control system algorithms to ensure accurate positioning of the booster model relative to the orbiter model. In addition, the control system software was modified to allow input and automation of the random setpoint order of the independent variables that were required in each subspace of the experimental design. These modifications were validated in static and wind-on testing of two inexpensive aluminum orbiter and booster models installed on non-instrumented balances ("dummy" balances) built specifically for this purpose. All setpoint matrices were run to ensure satisfactory positioning of the models. In addition, these check-out runs confirmed the absence of undesired model dynamics at all locations within the proximity matrix. The booster model was automatically moved to a "home" position at the conclusion of each run that corresponded to the positive extreme in the relative axial location (Δx = 28 inches) and the minimum values of the relative vertical location (Δz = 0.656 inches) and the relative angle of attack ($\Delta\alpha$ = 0 degrees). If the random run order resulted in consecutive replicate points, the booster model was again moved to the home position between points. The gains in the automated control system were tailored to allow the booster to establish a desired setpoint without overshooting. Data were acquired in a "setpoint-pause" mode, with sufficient time allowed for flow stabilization once the setpoint was established (approximately 2 seconds) and two seconds for data acquisition (30 samples per second). The automated data acquisition system would acquire data only when the setpoints and tunnel flow conditions were simultaneously within prescribed tolerances. The experimental design concepts of replication, randomization, and blocking [1] were key components in the current stage separation experiment. More details on these data quality assurance tactics are provided in the next section on experimental designs.

This investigation was considered exploratory in nature, consequently, only proximity data were acquired. That is, data were not acquired on the isolated booster or orbiter models, which would have allowed the estimation of aerodynamic interference increments. Isolated and proximity data were acquired on the same LGGB configuration at Mach = 3.0 in prior OFAT testing [4]. Blade strut interference effects were also not accounted for. Inferences regarding strut interference effects based on comparisons of the OFAT data to computational fluid dynamics predictions are briefly discussed in reference [4].

4 Experimental Design

4.1 Inference Subspace Definition

The inference subspace boundaries were established with the goal of compartmentalizing the overall flow field into smaller regions within which the aerodynamic responses were presumably well-behaved and estimable using relatively low-order polynomial functions. Preliminary rapid screening tests can effectively identify the regions in the inference space where the linear and nonlinear aerodynamics prevail. Prior experience or data from similar tests and engineering judgment are alternative means of estimating the "topography" of the overall inference space to characterize regions rich in structure or broad and featureless [6]. The prior OFAT testing [4] of the LGBB configuration indicated that the interference effects of the orbiter model on the booster model were highly nonlinear throughout the entire inference space at Mach = 3.0. These results suggested that a more uniform distribution of subspaces and design points were appropriate in an attempt to capture the nonlinear flow-field effects. It was decided to align the subspaces primarily along existing OFAT proximity matrix grid points (figure 3) and to define a total of 11 contiguous, or

adjoining, regions. Figure 7 is a schematic representation of the inference space with subspace boundaries labeled. The subspace boundaries are listed in Table I.

4.2 Precision and Resolution Requirements

A high-precision response surface predicts response values that have a high probability of being within a small interval of the true response. Given the magnitudes of the interference effects observed in the prior OFAT testing [4], precision intervals of +/-0.01 for lift coefficient, +/-0.001 for drag coefficient, and +/-0.001 for pitching moment coefficient were specified for the current exploratory investigation. The upper level of acceptable inference error probability was specified as 0.05, corresponding to 95% confidence. The two elements of the precision requirement were combined into specifications of a 95% confidence interval of +/-0.01 for lift coefficient and 95% confidence intervals of +/-0.001 for drag coefficient and pitching moment coefficient. Consequently, a sufficient volume of data was required to ensure that the response functions that were fit to these data points would predict lift coefficients that had at least a 95% probability of being within +/-0.01 of the true lift coefficient and drag coefficients and pitching moment coefficients that had at least a 95% probability of being within +/-0.001 of the true drag coefficient and pitching moment coefficient.

Invoking the Central Limit Theorem [1], it was assumed that the probability distribution of RSM response predictions for any specified combination of the independent variables was Gaussian, where the distribution was assumed to be based on models developed from numerous *n*-point data sets drawn from the same population (*n* is the sample size). Assuming the distribution was centered on the true response value, the current test required that the difference between the predicted and true responses be no greater than the specified precision limits with at least 95% probability. The precision limits, γ, can be expressed as the product of the standard deviation of this distribution and a constant which, for an "infinite" number of prior measurements, is 1.96, if 95% confidence is specified. The RSM prediction variance depends on the location in the inference space as well as the variance in the data [1], with the exact distribution of prediction variances directly related to the distribution of data points throughout the inference space. However, a single formula (3) represents the average variance for a polynomial regression across all measured points in the inference space [5], independent of the order of the polynomial or the number of independent variables, where *p* is the number of parameters in the model, *n* is the number of data points used to fit the model, and $\hat{\sigma}$ is the estimated standard deviation:

$$Var\left(\hat{y}\right) = \frac{p\hat{\sigma}^2}{n} \tag{3}$$

The square root of this quantity is the standard deviation of the distribution of response predictions, and the precision limit can then be expressed as

$$\gamma = t_{\alpha_s}\hat{\sigma}\sqrt{\frac{p}{n}} \tag{4}$$

Solving for *n* in equation (4) yields

$$n = p\left(\frac{t_{\alpha_s}\hat{\sigma}}{\gamma}\right)^2 \tag{5}$$

Here, the confidence interval half width, γ, is expressed as the number of standard deviations of the distribution of response surface predictions. It depends on the acceptable probability of error, α_s , and the amount of data used to estimate σ. Tables of values for the Student's *t* critical value, t_{α_s} , are found in most statistical texts and handbooks, and it has the assumed value of 1.96 as indicated earlier. Based on the prior OFAT testing, the estimated σ values for lift, drag, and pitching moment coefficients (*CL1*, *CD1*, and *CM1*) are .005, .0007, and .0007, respectively. A full second-order response model in three independent variables requires the estimation of 10 model parameters ($p = 10$). Given the precision limits of γ = 0.01, 0.001, and 0.001 for the lift, drag, and pitching moment coefficients, *n* is computed from equation (5) as follows:

Lift coefficient:

$$n = 10\left(\frac{1.960(0.01)}{0.005}\right)^2 = 9.6 \approx 10 \text{ points}$$

Drag coefficient:

$$n = 10\left(\frac{1.960(0.001)}{0.0007}\right)^2 = 18.8 \approx 19 \text{ points}$$

Pitching moment coefficient:

$$n = 10\left(\frac{1.960(0.001)}{0.0007}\right)^2 = 18.8 \approx 19 \text{ points}$$

These calculations reveal that if the standard deviation of each measurement of lift, drag, and pitching moment coefficient is no greater than 0.005, 0.0007, and 0.0007, respectively, a full second-order model in three independent variables must be fitted to *at least* 10, 19, and 19 points, respectively in order for the average 95% confidence interval half-width associated with predictions at those points to be no greater than 0.01, 0.001, and 0.001, respectively. The aerodynamic responses were measured simultaneously by the wind tunnel data acquisition system. Specifying the largest volume of data required for any one response variable, in this case 19 points for the drag coefficient and pitching moment coefficient, ensured that the minimum precision requirements were achieved for those variables, and it provided additional margin for the lift coefficient response. A total of 20 points were actually specified in each subspace, for reasons discussed in the next section

A high resolution response surface is one for which there is a high probability that a small indicated difference in response is real. The probability of correctly resolving the difference in two response levels (typically referred to as the power) depends on the size of that difference. This would be a consideration in the experimental design if, for example, the aerodynamic interference increments (proximity data minus the isolated data) were the response variables of interest. Figure 8 shows a power curve for lift coefficient corresponding to a full second-order response model in three independent variables developed by fitting 20 data points. The assumed measurement standard deviation is 0.005 and the level of significance is 0.05. In this case, the response surface is expected to resolve changes in the lift coefficient of 0.0094 with 95% confidence (the precision requirement for lift coefficient is 0.01). In contrast, figure 8 shows the power to resolve differences in the drag coefficient (or pitching moment coefficient) of 0.001 (corresponding to the precision requirement) is approximately 0.80 with 95% confidence. The power to detect a difference of 0.001 at 95% confidence would be 0.95 if the number of data points was increased to 29.

4.3 Selection of Points

The prediction variance distribution over the inference space depends on how the data points are distributed [1]. There are a number of established experimental designs that distribute the data points to achieve good prediction variance properties. A known point distribution called a cuboidal or face-centered design (FCD) [1] was selected as the primary design for this investigation and was applied in each of the 11 subspaces previously shown in figure 7. In the case of three independent variables, the FCD specifies 15 unique sites in a three-dimensional (3-D) inference space distributed in such a way that reveals a certain symmetry when the independent variables are coded by a linear transformation that maps them into an interval from -1 to +1. Details of the 15 sites in a 3-D FCD are shown in figure 9. Note that the design is partitioned into two cubes, which corresponds to the blocking scheme used in this experiment. One cube represents the factorial block containing 8 points at the cube corners or vertices (all combinations of the independent variables) and 1 point at the center of the cube. The factorial points allow estimation of main effects and first-order interaction terms as in equation (1). The axial block contains the 6 axial points at the face centers and 1 point at the center of the cube. The axial points at the face centers allow estimation of second-order or quadratic terms. Two additional center points were included in each block such that the total number of points in the FCD in each subspace was 20. The three points at the center of the

design in each block represent replicates, which were used to estimate the pure error in the experiment, to perform lack of fit tests [1], and to contribute to the estimation of quadratic effects. The location of the replicates at the center of the design also produced reasonable stability of the scaled prediction variance [1]. Blocking is a quality assurance tactic that defends against unexplained variance in the data [6]. It entails clustering the data into groups or blocks. All points in one block are executed before any of the points in another block, and the blocks are executed in random order. Blocking was applied throughout the day in the current experiment, and the block boundaries were aligned with natural candidates for systematic variation, such as shift and operator changes and end-of-daily operations. Setting the independent variable levels in a randomized order within blocks defended against the adverse impact of correlated errors and effectively decoupled systematic errors from the true factor effects [6].

Orthogonal blocking for a cuboidal design requires disproportionate block sizes [1]. If the blocking is not orthogonal, then presence of significant block effects can affect the numerical values of the coefficients in the response model. Consider the following equation (6) for orthogonal blocking, which expresses the distance of the axial points from the center of a CCD to the number of factorial points, number of independent variables, and number of center points in the factorial and axial blocks:

$$\alpha_a = \sqrt{\frac{F\left(2k + a_o\right)}{2\left(F + F_o\right)}} \qquad (6)$$

In the cuboidal design used in this experiment, the number of factorial points, F, is eight; the number of independent variables, k, is three; and the number of center points is three in both the factorial and axial blocks (F_o and a_o, respectively). This yields an axial distance, α_a, of 1.809 for orthogonal blocking. However, the axial distance in a cuboidal design is α_a = 1.0. One way of achieving orthogonal blocking in this case is to assign twenty four center points to the factorial block and two center points to the axial block. This was not considered an acceptable option in this experiment. A more detailed description of blocking, including orthogonal blocking, is available in the literature [1].

Of the 20 degrees of freedom in each subspace, ten were allocated to a full second-order model in three independent variables, leaving ten residual degrees of freedom to assess the quality of the fit. This provided a 50-50 split between the resources allocated to estimating the response models and resources allocated to assessing them. Since there were 15 unique sites in each subspace and ten parameters to fit for a full second-order model, five of the ten residual degrees of freedom were available to assess the lack of fit. Up to five additional terms could be added to the response model to improve fit if needed, without having to specify any additional data. However, there would be no available degrees of freedom to assess lack of fit. The five remaining residual degrees of freedom were available to assess unexplained variance in the data. However, one of these was allocated to assessing and defending against block effects. The four remaining variance degrees of freedom per subspace were available to assess pure error, or random variations in the data. Table II summarizes the degrees of freedom budget in the full second-order model applicable to the current experiment.

The cuboidal design was the design of choice compared to the spherical CCD illustrated in figure 9 because of physical constraints imposed on the spacing of the orbiter and booster models. It was desired to obtain a sufficient array of measurements at unique sites in the inference space where the orbiter and booster models were in close proximity, within the safety constraints imposed by the facility. As previously stated, the minimum vertical spacing for this test corresponded to Δz = 0.656 inches, which established the lower range of this independent variable. Consequently, the lower plane of the first row of subspaces was aligned with Δz = 0.656 inches, which yielded a desirable density of observations at the minimum allowable vertical separation distance. The FCD provided additional measurements at this minimum separation, since the axial points

in the cuboidal design were at the centers of the cube faces. In contrast, a spherical CCD would require a larger displacement of the factorial cubes from the minimum separation location in order to incorporate the axial points that were outside of the cube faces in this design.

The FCD could also support response surface models with a limited number of higher-order terms (e.g., mixed cubic). A decision was made during the testing to augment the FCD in subspaces 1 through 5 (see figure 7) to accommodate full third-order models. A computer-generated D-optimal design [1] was defined by selecting design points from a candidate list that minimized the variance of the regression coefficients. Fourteen unique design points were specified at the cube edges (edge centers, edge thirds) and at locations interior to the cube along with 3 center points in each subspace, and these 17 additional data points were run in random order in a separate block for each of the selected subspaces. This yielded a total of 37 data points in each of the 5 subspaces for which the D-optimal design was created. Available resources also allowed the inclusion of spherical CCDs in four subspaces with trimmed ranges of the spatial factors. A sketch of the subspace layout within which a spherical CCD was applied is shown in figure 10. The subspaces are denoted 1′, 2′, 4′, and 5′ to distinguish them from the different subspaces to which the FCD was applied. For clarity, only the axial points for 1′ and 5′ are shown. These designs also supported full second-order models and a limited number of mixed cubic terms.

The FCD, D-optimal, and spherical CCD effectively yielded a total of six basic empirical models for comparison: (1) up to full second-order using the FCD, (2) full second-order and additional mixed cubic and quartic terms using the FCD, (3) full third-order using D-optimality, (4) full third-order using D-optimality and additional quartic and mixed quartic terms, (5) full second-order using the spherical CCD, and (6) full second-order and additional mixed cubic and quartic terms using the spherical CCD. Each of these cases yielded a different distribution of the prediction variance throughout the inference space. Figure 11 illustrates the distributions of the unit standard error of prediction [1] for all six cases. The full second-order FCD exhibits a desirable broad, flat region of nearly constant, relatively low error in the center of the design with the larger errors pushed toward the boundaries of the design. A similar description applies to the full second-order spherical CCD, which is a rotatable design since the error depends only on distance from the center, not the direction. The full third-order D-optimal design exhibits less desirable prediction variance properties with a more irregular and undulating error distribution. For all the three previous designs, adding higher-order terms degrades the error distribution, particularly for the D-optimal design. Figure 11 shows that the addition of model terms of higher order increases the levels and local gradients of the prediction variance distributions. Furthermore, the error distributions are less predictable.

Two confirmation points were also obtained in each of the 11 subspaces with the FCD, and 4 confirmation points were acquired in each of the 4 subspaces with the spherical CCD. These points were not used to generate the response surface functions but, instead, were held in reserve to test the empirical models. The models were used to predict the responses at these points, and it was a requirement that a sufficient number of confirmation points be adequately predicted for the models to be considered a proper representation of the system response.

The data point total for the current experiment is now summarized. The FCD in three variables featuring 22 points (including confirmation points) was specified in all 11 inference subspaces for a total of 242 points. There were 103 unique sites in the overall inference space, 11 of which were replicated 6 times. Augmenting the FCD to a full third-order D-optimal design required 17 additional points in 5 subspaces totaling 85 points. There were 70 additional unique sites, and 5 sites were replicated 3 times. The spherical CCD featured 24 points (including confirmation points) in each of 4 subspaces totaling 96 points. There were 48 unique sites, and 4 sites were replicated

3 times. A grand total of 375 points was budgeted for the entire experiment.

5 RSM Analysis

5.1 Model Building

Multiple linear regression methods [1], [2] were used to develop the empirical response functions relating the aerodynamic coefficients to the independent variables in all subspaces. A total of 12 responses were measured (6 aerodynamic forces and moments for the booster model and orbiter model) in 11 subspaces for a total of 132 response functions. The FCD supported the inclusion of up to 13 model parameters. Assuming at least one degree of freedom was required to assess the quality of fit, there were up to 12 terms that could either be included in the model or not (two possible states), or 2^{12}-1 = 4095 candidate models in each case, or a total of 540,540 candidate models to consider, in principle. The task was reduced by half, since only the three longitudinal aerodynamic coefficients for each model were of interest. Furthermore, the focus of this exploratory investigation was the booster model lift, drag, and pitching moment coefficients, so "only" 135,135 candidate models remained. Fortunately, a number of model-building strategies exist to render this task more feasible, and the present analysis used a backward elimination method as a guide to selection of final models. A detailed description of this strategy is provided in references [1] and [2]. All analyses were performed using commercially-available experimental design and statistical analysis software [19]. It is noted that backward elimination does not guarantee a "best model." Any model building strategy is affected by the presence of correlated independent variables, or multicollinearity, where the contribution that a given model term makes to the explained variance depends on the presence of another variable in the model. An advantage of the second-order FCD and spherical CCD is that this linear dependence is minimized, so the use of a model-building strategy such as backward elimination is more likely to yield an adequate model. The initial approach that was taken with data acquired using the FCD was to specify a full second-order model and to apply backward elimination to identify a model for more detailed consideration. A formal analysis of variance (ANOVA) [1] partitioned the total variance in the experimental data into a component that could be explained by the model and a residual or unexplained component not explained by the model. The ANOVA output included an *F*-test [1] to determine if the regression model was significant at an assumed significance level (α_s-level) of 0.05. Individual *F*-tests were also performed on all regression coefficients. A lack of fit test provided a metric to assess how well the model fit the data. Acquiring data at more unique combinations of the independent variables than there were coefficients in the model provided an estimate of the unexplained variation due to lack of fit, and the estimate of pure error from the center point replicates provided a basis with which to assess its significance. Excessive lack of fit suggested that a different model may be necessary. It is noted, however, the pure error component in most wind tunnel tests is typically extremely small. Since this variance component comprises the denominator in the lack-of-fit *F*-test, even quite reasonable lack of fit variance components are declared statistically significant. Consequently, the regression models will often fail the lack of fit test not because this variance component is large in a meaningful absolute sense but because there is so little pure error. The R^2_{adj} and R^2_{pred} statistics [1] provided measures of the explained variance relative to the total variance and the amount of variation in new data explained by the model, respectively. A guideline [19] was that these two values should be within about 0.20 of each other. It is possible to obtain very high values of R^2_{adj} by overfitting the model and correspondingly small values of R^2_{pred} since the model is a poor predictor of the response at other locations within the inference space. The main problem with using the *R*-squared statistics for wind tunnel response surface model evaluation is the

precision of the measurements is generally so good that these values are often very close to 1. As a result, they are not very clear discriminators. The analysis also included an Adequate Precision value [19], which was basically a signal-to-noise ratio or another measure of regression model discrimination.

Important information that was provided in the post-ANOVA output was contained in the variation inflation factors (VIF) [1] for the individual terms. VIF is a measure of how much the variance of the model is inflated by the lack of orthogonality in the design. A value of 1 indicates the factor is orthogonal to all other factors in the model. Values greater than 10 indicate the factors are too correlated with each other, that is, multicollinearity is sufficiently strong to adversely affect the estimates of the regression coefficients and the ability to discern which factors truly affect the aerodynamic responses.

5.2 Model Validation

5.2.1 Residuals Analysis

Absent any other information external to the experiment, all the information that is available about the quality of the regression model fit is contained in the residuals. The residuals are the difference between the observed response and the predicted response for the same combination of the independent variables. Underlying assumptions in the regression model include normally and independently distributed errors having a mean of zero and a constant variance. The residuals for a well-formed model should be due only to random error in the data, not a systematic departure of the model from the data. A comprehensive residuals analysis was performed for all promising models from the backward elimination method. This process was very rapid and efficient using the diagnostic plots and influence plots described in reference [19]. A detailed discussion of residuals analysis is provided in the literature (see references [1] and [2], for example). The diagnostic plots included a normal probability plot of the residuals, residuals versus predicted response values, residuals versus experimental run order, predicted versus actual response values, Box-Cox plots [1], and residuals versus factor level. The influence plots included externally studentized residuals, leverage, DFFITS and DFBETAS [19], and Cook's D [1]. Upon completion of the residuals analysis, 3-D response surface plots and 2-D contour plots were created to visualize the estimated aerodynamic responses throughout each inference subspace.

A manual review was performed of all recommended models resulting from backward elimination. Selected model terms were added and/or deleted, and the analysis of the updated regression model was repeated. Engineering judgments were made regarding the usefulness of certain model terms. For example, a model term could be declared statistically significant based on the 5% level of significance assumed in these analyses, but it may not have been judged practically significant. The analysis was initially driven by the concept of model parsimony [1], that is, to develop easily-interpretable response functions with the minimum number of terms that were good predictors of the aerodynamic responses. Model hierarchy [1] was also maintained throughout the analyses. The hierarchy principle indicates that if a model contains significant higher-order terms, it should also contain all of the lower-order terms that comprise it, even if the *F*-tests on these terms are insignificant. In the end, the models were largely the product of the model-building algorithm and user input.

5.2.2 Subspace Boundary Discontinuities

The partitioning of the overall inference space resulted in empirical response surfaces that did not meet seamlessly at the subspace boundaries. Discontinuities in the levels of the estimated response and the slopes at the subspace boundaries were observed. This occurred because the responses at the boundaries were estimated from two different regression models (built from different data sets) applied to the same combinations of the independent variables. If the regression models have adequately captured the underlying aerodynamic response behavior in the adjoining subspaces, then the boundary discontinuities

should simply reflect model prediction uncertainty, which exists throughout the inference space and not just at inference subspace boundaries. The boundary phenomenon is simply the RSM equivalent of replicating data. For example, two different response values would be obtained on the boundary at the same combination of the independent variables if two direct measurements were acquired instead of predictions from two different response surface functions. This is due to random error in any data set, which is in fact the same reason that the response models do not perfectly agree. In assessing subspace boundary discontinuities, the question was whether the variance in the predictions was consistent with experimental uncertainty. The claim is not that the response surfaces themselves represent the true response of the system, but that the true response lies with some probability within locally parallel surfaces representing the limits of some prescribed precision interval. Those surfaces can have discontinuities at the boundaries. The only claim is that the true surface is continuous at the boundary and lies everywhere between the precision intervals, even if those intervals do not meet smoothly at the boundary.

The differences in the response predictions for the lift, drag, and pitching moment coefficients on the subspace boundaries obtained from the regression functions on either side of the boundaries were compared to precision requirements. The precision requirements for each response variable presented in an earlier section describe the acceptable uncertainties for individual response estimates. To compute acceptable corresponding uncertainties for differential response estimates, these values were simply multiplied by the square root of two, since under commonly occurring conditions the variance of the difference between two random variables is the sum of their variances [2]. Given the layout of the 11 inference subspaces in this experiment, there were a total of 14 common boundaries. There were a total of 5 common design points (4 factorial points and 1 axial point) at each boundary. Consequently, 70 comparisons were made for each of the three longitudinal aerodynamic coefficient responses, resulting in 210 total comparisons. Fewer comparisons were made in the analyses of the FCD design with added higher-order terms, the D-optimal design, and the spherical CCD since fewer subspaces were investigated. The total comparisons in these cases are summarized in Table III.

5.2.3 Confirmation Points

Another metric in assessing the adequacy of the empirical response surface models was to compare the predicted responses to the observed responses at different combinations of the independent variables than those for which data were acquired to build the models. Confirmation points were acquired at randomly-selected locations within the inference subspaces at the same time as the design points. These points were interleaved with the design points in each block and were therefore acquired in random order. The confirmation points were not used to generate the models but held in reserve until promising regression functions had been identified. Twenty-two confirmation points, 2 points per subspace, were obtained during the acquisition of data required by the FCD. Ten of these confirmation points were also used for comparisons to the predictions obtained in five selected subspaces using the FCD with added higher-order terms and the D-optimal design. A total of 16 confirmation points, 4 points per subspace, were obtained along with the design points for the spherical CCD. Table IV summarizes the confirmation points for all designs. A successful confirmation point was obtained if the measured response fell within a prediction interval centered about the predicted response, with the interval width corresponding to a 95% probability. An objective measure of how many successful confirmations were required to consider the empirical functions as reasonable predictors of the aerodynamic responses was proposed in reference [6]. In this approach, the model confirmation was considered a Bernoulli process in which there were a prescribed number of trials with exactly two outcomes, success and failure, and a clearly defined criterion for success. A successful trial was

defined as having occurred in the confirmation process when a confirmation point was situated within the 95% prediction interval of the model. A failure occurred when the confirmation point fell outside this interval. The true probability of success on any one trial was assumed to be 95% corresponding to the size of the prediction interval. The cumulative binomial probability distribution was used to determine a "critical binomial number," call it *m*, for an *n*-trial process for which the success rate was 95% for each trial. Specifically, there was at least a 95% probability of *m* or more successes in *n* trials if the probability of success on each trial was 95%. For the 22-trial process in the current experiment, the critical binomial number was 19. Consequently, 19 or more confirmation points falling within the 95% prediction intervals was the success criterion.

6 Discussion of Results

A significant amount of information was acquired in the stage separation experiment, which included 12 aerodynamic force and moment responses (6 responses per model) as functions of 3 independent variables, 3 experimental designs, and up to 11 inference subspaces. The focus of the current investigation was narrowed to three aerodynamic responses for the booster model corresponding to the lift, drag, and pitching moment coefficients within the 3-D inference space defined by the independent variables, Δx, Δz, and $\Delta\alpha$. The results presented in this section are confined to the lift coefficient response, since the discussion is also applicable to the drag coefficient and pitching moment coefficient responses.

6.1 FCD

Initial full second-order models for the booster model lift coefficient response, *CL1*, were specified in all 11 subspaces, and backward elimination was then used to build reduced models. Additional terms were often deleted in a manual editing process to include only the "heaviest hitters" in the model. For example, a model recommended from backward elimination may have included a quadratic term in Δz with a "*p*-value" (probability value) [1] much less than the cutoff level of significance of 0.05 (that is, Δz^2 highly significant). However, its contribution to the explained variance may have been two or three orders of magnitude less than the main effect of $\Delta\alpha$. In the interest of model parsimony, the quadratic term was removed from the model, as long as the ANOVA output and model diagnostics were satisfactory. The terms that were included in the final models for *CL1* are listed in Table V. There were at most 5 model terms (not including the intercept), which left 9 degrees of freedom for lack of fit. The main effect of $\Delta\alpha$ was common to all models, and this term accounted for a significant portion of the explained variation in the data. This was an expected effect, since the lift coefficient is known to be a strong function of the angle of attack [20]. Several of the empirical models included the main effects of Δx, Δz, the first-order interaction of Δx and Δz, and a quadratic term in Δz. Three models from subspaces where the booster model was more distant from the orbiter model included only the main effect of $\Delta\alpha$. Lack of fit was the only *F*-test in the ANOVA output that was a concern, since it was statistically significant in all cases. This was a common result, even with the higher-order models that are discussed in later sections. The model diagnostics did not reveal any significant issues other than a limited number of points being flagged with high residuals or as potentially influential observations.

Figure 12 presents a composite plot of the lift coefficient (*CL1*) response surfaces in subspaces 1-11. In each subspace, the response surface is situated above an inference plane defined by the independent variables Δx and Δz. The third independent variable, $\Delta\alpha$, is fixed at its mid-level of 2.5 degrees. The 11 response surfaces are arranged approximately in their relative locations within the overall design space to provide a global qualitative view of the estimated lift coefficient response. The response surfaces are plotted on a common scale for all subspaces. The numerical values on the coordinate axes are not clearly visible.

However, the common plot scale provides a useful means of qualitatively assessing significant changes in the empirical response surfaces across the inference space and to compare local changes in the response levels and slopes at the subspace boundaries. Although not presented in this paper, the lift coefficient generally exhibited a linear dependence on $\Delta\alpha$ for any combination of Δx and Δz. The dependence of *CL1* on Δx and Δz at a given $\Delta\alpha$ was less interpretable, which can be inferred from figure 12.

Several of the response surfaces in subspaces closer to the fixed orbiter model position exhibit a quadratic effect in Δz at a given level of Δx. In addition, a twisting of the response surfaces indicates a first-order interaction involving Δx and Δz, that is, the effect of one factor, say Δz, on the lift coefficient response depends on the level at which the other factor, Δx, is set. The effect of Δx on *CL1* is typically linear. In the three subspaces where contours of constant response are not shown in the inference plane defined by Δx and Δz, the predicted lift coefficient is a function of $\Delta\alpha$ only.

The linear dependence of *CL1* on Δx conflicts with the nonlinear trends that were previously illustrated in figure 4 corresponding to the OFAT testing [4] of the LGBB. In addition, the discontinuities in the response levels and slopes at the subspace boundaries are problematic, particularly where quadratic curvature is opposite in sign or where a strong quadratic trend is predicted from the response function on one side of the boundary and a linear trend from the response function on the other side. Comparisons of the predictions from response surface models in adjoining subspaces at all 70 common boundary points are summarized in Table III. A total of 61 comparisons for *CL1* were within the precision limits, 57 for drag coefficient, *CD1*, and 39 for pitching moment coefficient, *CM1*. The booster model pitching moment coefficient was expected to be particularly sensitive to the location and strength of impinging shock waves from the orbiter model [4]. If the comparisons of the predicted responses at the subspace boundaries are regarded as a Bernoulli process as outlined in the previous section, then the critical binomial number is 63 for this 70-trial process. That is, 63 or more comparisons falling within the precision intervals is the success criterion. This criterion was not met for any of the aerodynamic response comparisons. The simplified models from the FCD design are judged inadequate, since the variance in the predictions was not considered within the experimental uncertainty.

6.2 FCD with Higher-Order Terms

The regression model building process was repeated for subspaces 1-5 to determine if the addition of higher-order terms that were supported by the model would improve the fits. The FCD supported up to 14 model terms (exclusive of the intercept) that were unaliased, that is, not confounded with any other terms in the model [1]. The initial model that was specified prior to the application of backward elimination was a full second-order model plus mixed cubic and a single mixed quartic term. The issue that arose in these cases was that backward elimination frequently identified all fourteen model terms as statistically significant, and the full model was retained. The pure error estimate from the centerpoint replicates was the sole source of information with which to estimate the experimental error, and this term was so small that all statistical tests of significance for the model terms yielded very low *p*-values (much less than the cutoff value of 0.05). With no lack of fit degrees of freedom, there was no means of assessing the quality of the fit. In essence, the model was being forced through the mean of the responses at every site in the inference space for which data was acquired. Consequently, all models that were selected as promising candidates from backward elimination were manually edited, and certain higher-order terms were deleted from the model on the basis of their relative contribution to the explained variance. This approach provided at least 2 degrees of freedom for lack of fit and a more meaningful analysis of the residuals. All models except for the pitching moment coefficient response model in subspace 2 failed

the lack of fit test. At this point, it was concluded that the lack of fit test was too misleading, since the extremely small estimates of pure error consistently rendered any acceptably small lack of fit component as highly significant. Table V shows the models that were developed for the lift coefficient response in subspaces 1-5. The models are significantly more complex than their counterparts from the previous section, and they include main effects, two- and three-factor interactions, quadratic effects in Δx and Δz, mixed cubic terms, and a single mixed quartic term. The higher-order effects of Δx and Δz seem more consistent with the trends that were exhibited in the OFAT data set in reference [4]. The more nonlinear lift coefficient response is also apparent in figure 13, which shows a composite plot of the response surfaces estimated from the higher-order models in the 5 adjoining subspaces in the bottom half of the figure. For reference, the lower-order response surfaces from the previous section are repeated in the upper half of the figure. All plots feature a common scale in figure 13. Side-by-side comparisons of the response surfaces estimated from the two different sets of empirical functions are shown in figure 14, where the plot scales are adjusted from subspace-to-subspace but are the same within a given subspace. The second-order curvature along the Δx and Δz axes is apparent, and the first- and higher-order interaction terms result in complex twisting of the surfaces. Visual inspection of the response surfaces at the subspace boundaries also suggest a better match in the local response levels and slopes.

The post-ANOVA output indicated that the variance inflation factor for the mixed quartic term was greater than the recommended cutoff value of 10 [19]. The presence of multicollinearity, or lack of orthogonality in the design, inflated the variance of the corresponding model regression coefficients. The most significant outcome from the residuals analysis was that several points exhibited very high leverage or leverage values of 1, and other influence diagnostics showed many points well outside of the recommended limits. A leverage value of 1 indicated that the predicted response was forced to exactly match the observed data at that point. It is concluded, then, that too many higher-order terms were included in the models, since the models were essentially being fit to noise. The results in Tables III and IV illustrate this point. A comparison of the predicted responses at the subspace boundaries using the regression models from adjoining subspaces yields 100% agreement. In contrast, most of the observed confirmation points fail to fall within the 95% prediction intervals centered about the predicted responses.

6.3 FCD Augmented to D-Optimal

The prior results motivated an augmentation of the FCD to a full third-order D-optimal design by obtaining additional data points at 14 unique sites in separate randomized blocks in subspaces 1-5. This model supported terms beyond full third-order including selected mixed quartic and pure quartic terms. The general form of the final models is summarized in Table V. All models were hierarchical and included terms up through pure quartic. The inclusion of pure cubic, mixed cubic, pure quartic, and mixed quartic terms involving Δx and Δz is not unreasonable given the highly nonlinear behavior that was observed or inferred in the OFAT testing [4]. Interpretation of many of the higher-order mixed terms is that the cubic effects of Δx and Δz depend on the level of the other spatial factor or the relative angle of attack. The same approach to model building that was taken in the previous section was applied to the current data set. Specifically, all available unaliased terms were included in the model as a starting point for backward regression. If all available terms were retained in the model after backward regression, engineering judgment was used to delete at least 2 terms to provide degrees of freedom for lack of fit and to better assess the residuals. Figure 15 presents side-by-side comparisons of the response surfaces estimated from the simplified FCD models and from the D-optimal models at each of the 5 selected subspaces. Similar to figure 14, the plots scales are different from subspace-to-subspace but are the same within a given subspace to allow a direct qualitative comparison of the two designs. The

increased complexity of the D-optimal response surfaces in all subspaces is obvious. The highly-warped response surfaces are less interpretable, which may be indicative of significant spatial variability of the multiple shock wave impingements on the booster model.

The response surface models developed in subspaces 1, 2, and 4 were used to predict piecewise continuous distributions of *CL1* versus Δx at fixed values of Δz = 1.313 inches and $\Delta\alpha$ = 0 degrees as shown in figure 16. This distribution is directly comparable to the OFAT distribution that was previously shown in figure 4. It is noted that the response surface distribution corresponds to combinations of the independent variables for which data were not acquired. The response surface estimates appear to capture the general character of the lift coefficient distribution that was observed in the OFAT testing, although quantitative agreement is lacking.

The post-ANOVA results indicated that all models included several terms ranging from main effects to mixed quartic terms with VIF values indicative of extreme multicollinearity (VIF ranging from 50 to greater than 30,000!). The residuals analysis revealed many points with high leverage values, several with leverage of 1, and extremely large DFFITS, DFBETAS, and Cook's D. These influence diagnostics indicated that the response surfaces were being forced to fit the data points, and that the models would be poor predictors of the responses at other points in the inference space. This assessment is confirmed in Tables III and IV, which show that the predicted values at the subspace boundaries using the regression models from adjoining regions are universally within the 95% precision intervals, but the models are poor predictors when compared to the observed confirmation points.

6.4 Spherical CCD

A spherical CCD was applied to the four subspaces previously shown in figure 10 to determine any potential advantages to an alternate, rotatable CCD with orthogonal blocking applied to inference subspaces that were slightly trimmed compared to the FCD. The subspaces were labeled $1', 2', 4',$ and $5'$ to distinguish them from the original FCD subspaces (some overlap exists between the subspaces). The minimum Δz for the factorial blocks was increased to 1.00 inch from 0.656 inches to accommodate the axial points that extended outside the inference subspace cubes. The spherical design provided five levels of the independent variables, so it was more flexible, had better standard error of design characteristics, and was nearly orthogonal. In addition, orthogonal blocking was supported. The trimming of the subregions was an initial attempt to better resolve the shock-rich inference space navigated by the booster model. The spherical design supported a full second-order model and certain mixed cubic and mixed quartic terms. The application of backward elimination typically returned all unaliased terms that were originally specified in the model. Manual editing of the models led to the deletion of one or two model terms that were statistically significant but represented the smallest contributions to the explained variance in the data set. Consequently, there were at most 2 degrees of freedom for lack of fit. The deletion of any higher-order terms had very little effect on the regression coefficients because of the near-orthogonality of the spherical CCD. This was in contrast to the severe multicollinearity encountered in the FCD with higher-order terms and the D-optimal design. Table V shows the terms that were included in the final regression models for *CL1* in the four subspaces. Model terms common to the response surface functions in subspaces $1', 2', 4',$ and $5'$ included all main effects, all two-factor interactions, and quadratic, mixed cubic, and mixed quartic terms in Δx and Δz. Depending on the subspace, the quadratic effect of $\Delta\alpha$ and certain higher-order interaction terms involving $\Delta\alpha$ were included in the models. Figure 17 shows a composite layout of the lift coefficient response surfaces, where the independent variable $\Delta\alpha$ is set to its mid-level of 2.5 degrees. A common plot scale was used in all subspaces. The second-order curvature in the spatial variables is apparent in all response

surfaces and, qualitatively, the subspace boundary discontinuities appear reasonable. The predictions of the lift coefficient at the subspace boundaries using the regression functions in adjoining subspaces are all within the specified precision intervals with 95% confidence (see Table III). Unfortunately, the models are poor predictors of the response at other locations within the inference subspaces. As shown in Table IV, very few of the observed confirmation points fell within the 95% prediction intervals. The ANOVA output indicated significant lack of fit for most models, the only exceptions being *CM1* in subspaces 1′ and 5′ and *CD1* in subspace 2′. The post-ANOVA output did not reveal any issues with multicollinearity. The residuals analysis showed a large percentage of the design points with very high leverage, or leverage values of 1 in addition to large values of the externally studentized residuals, DFFITS, DFBETAS, and Cook's D. Consequently, the response surfaces were once again being forced to fit the data, and the functions behaved unrealistically at other combinations of the independent variables.

Figure 18 compares the lift coefficient response surfaces at $\Delta\alpha$ = 2.5 degrees for all of the designs and response models in subspace 1, where the subspaces corresponding to the simplified FCD, FCD with higher-order terms, and D-optimal have been trimmed to match the spherical CCD inference space 1′. Significant differences in the response surface behavior are apparent in the D-optimal and spherical designs compared to the FCD design models, particularly near the boundary defined by the Δz = 1.00-inch plane. These differences are also apparent in the 2-D plots of *CL1* versus Δx at Δz = 1.00 inch and $\Delta\alpha$ = 2.5 degrees in figure 19. Better agreement in the response levels and slopes is obtained inside the inference space, as shown in figure 20 corresponding to Δz = 2.5 inches and $\Delta\alpha$ = 2.5 degrees.

7 Conclusions

Response surface methodology was used in the design, execution, and analysis of a wind tunnel investigation to estimate the supersonic stage separation aerodynamics of a generic, bimese two-stage-to-orbit vehicle in a belly-to-belly arrangement. The current paper focused on the development of empirical response surface models to estimate the longitudinal aerodynamics of a winged booster vehicle as it navigated through a three-dimensional inference space dominated by shock waves originating from a like-scale orbiter vehicle. The inference space was defined by two spatial variables, the relative axial and vertical locations of the booster and orbiter, and one rotational variable, the relative angle of attack. Experimental data to support the response surface modeling was obtained in the NASA Langley Unitary Plan Wind Tunnel at Mach = 3.0 using 0.0175-scale models of the booster and orbiter vehicles. The overall inference space was partitioned into several contiguous subspaces within which data were acquired using formal design of experiments. The experimental design principles of randomization, replication, and blocking were key elements in this investigation. The partitioning was guided by prior data using a one-factor-at-a-time testing technique, which showed highly nonlinear booster vehicle aerodynamics that persisted throughout the inference space because of orbiter vehicle shock wave impingements. The data obtained in the multiple subspaces were used in an attempt to build piecewise-continuous empirical response functions relating the aerodynamic coefficients to the independent variables over the entire inference space. Face-centered central composite, D-optimal, and spherical central composite designs were executed that supported models ranging from full second-order plus mixed cubic and mixed quartic terms to full third-order plus pure quartic and mixed quartic terms. A combination of backward elimination and engineering judgment was used to build the multiple linear regression models, and established statistical techniques were applied to judge the adequacy of all models. The model evaluation was augmented by objective comparisons of the predicted responses at the subspace boundaries using empirical response surface models in adjoining regions. In addition, response surface model predictions were compared to confirmation data

points, which were measurements specifically acquired for model confirmation.

Reduced second-order response surface models estimated from data obtained using the three-level, face-centered, or cuboidal, design were judged inadequate because of systematic discontinuities in the response levels and slopes at the subspace boundaries. The dependence of the aerodynamic responses on the relative angle of attack was adequately modeled, but the dependence on the spatial variables appeared to be of a higher-order than assumed in these relatively simple models. The addition of higher-order mixed cubic and mixed quartic terms to these models mitigated the subspace boundary issue. However, new concerns were introduced related to multicollinearity in the independent variables and high-influence points that produced empirical models that were poor predictors of the response at other locations within the inference space. Augmentation of the face-centered designs in selected subspaces to full third order using D-optimality criteria was accomplished by acquiring additional data. More complex response functions were developed that included up to pure quartic terms in the spatial variables. These models were also poor predictors of the aerodynamic responses within the inference space because of severe multicollinearity and multiple high-influence data points.

A 5-level spherical central composite design was applied to selected subspaces with trimmed ranges of the independent variables to better capture the complex shock-induced aerodynamic responses. The spherical design produced response models with terms up to mixed quartic. This near-orthogonal design avoided the issue of multicollinearity, but multiple high-leverage points resulted in empirical functions that failed objective model confirmation tests.

The results obtained in this investigation suggest the shock-rich environment traversed by the booster model in the simulated Mach 3 staging is more complex than originally assumed. Shock waves represent an aerodynamic discontinuity across which the local pressure and temperature abruptly change. The presence of multiple shock waves, shock wave interactions, and shock impingement induces nonlinear and possibly discontinuous aerodynamic flow-field effects on the booster. If the booster model aerodynamics are nonlinear and discontinuous, then the experimental design must have sufficient spatial resolution to effectively capture these phenomena. The partitioning of the overall inference space in the current experiment into 11 contiguous regions was insufficient to resolve these features. Adding higher-order terms to regression models built from data sets acquired over too-expansive ranges of the independent variables resulted in empirical functions that were fit to noise and were, therefore, poor predictors of the aerodynamic responses throughout the inference subspace. In addition, the distribution of points in central composite designs may not be appropriate to capture the very complex responses that are often encountered in aerospace testing. Significant improvements to the response surface modeling in aerodynamic flow fields similar to the current stage separation scenario may be possible by higher-level partitioning of the overall inference space and the application of full 4^{th}-order D-optimal designs within the trimmed subspaces.

8 References

[1] Myers R H and Montgomery D C. *Response Surface Methodology: Process and Product Optimization Using Designed Experiments*. 2nd Edition. John Wiley & Sons, Inc., New York, 2002.

[2] Neter J, Kutner M H, Nachtsheim C J and Wasserman, W. *Applied Linear Statistical Models.* 4th edition. McGraw-Hill, Inc. 1996.

[3] Deloach R. Improved Quality in Aerospace Testing Through the Modern Design of Experiments (Invited). AIAA-2000-0825. 38th AIAA Aerospace Sciences Meeting & Exhibit, Reno, Nevada, January 2000.

[4] Murphy K J, Erickson G E and Goodliff S L. Experimental Stage Separation Tool Development in NASA Langley's Unitary Plan Wind Tunnel. AIAA-2004-4727. 22nd AIAA Applied Aerodynamics Conference & Exhibit, Providence, Rhode Island, August 2004.

[5] Box G E P and Draper N R. *Empirical Model Building and Response Surfaces.* John Wiley & Sons, Inc., New York, 1987.

[6] DeLoach R and Erickson G E. Low-Order Response Surface Modeling of Wind Tunnel Data Over

Truncated Inference Subspaces. AIAA-2003-0456. 41st Aerospace Sciences Meeting & Exhibit, Reno, Nevada, January 2003.

[7] DeLoach R. Applications of Modern Experiment Design to Wind Tunnel Testing at NASA Langley Research Center. AIAA-1998-0713. 36th AIAA Aerospace Sciences Meeting & Exhibit, Reno, Nevada, January 1998.

[8] Landman D, Simpson J, Vicroy D, and Parker P. Efficient Methods for Complex Aircraft Configuration Aerodynamic Characterization using Response Surface Methodologies. AIAA-2006-0922, 44th AIAA Aerospace Sciences Meeting & Exhibit, Reno, Nevada, January 2006.

[9] Amer T R, Liu T, Oglesby D M. Characterization of Pressure Sensitive Paint Intrusiveness Effects on Aerodynamic Data. AIAA-2001-0556, 39th AIAA Aerospace Sciences Meeting & Exhibit, Reno, Nevada, 2001.

[10] Rhode M N and DeLoach R. Hypersonic Wind Tunnel Calibration Using the Modern Design of Experiments. AIAA-2005-4274. 41st AIAA/ASME/SAE/ASEE Joint Propulsion Conference & Exhibit, Tucson, Arizona, 2005.

[11] Danehy P M, DeLoach R, and Cutler A D. Application of Modern Design of Experiments to CARS Thermography in a Supersonic Combustor. AIAA-2002-2914, 22nd AIAA Aerodynamic Measurement Technology and Ground Testing Conference, St. Louis, Missouri, June 2002.

[12] Albertani R, DeLoach R, Stanford B, Hubner J P, and Ifju P. Wind Tunnel Data Base Development and Nonlinear Modeling Applied to Powered Micro Air Vehicles with Flexibly Wing. AIAA-2006-6640, AIAA Atmospheric Flight Mechanics Conference and Exhibit, Keystone, Colorado, August 2006.

[13] Luner J J and Healey M D. Modern Design of Experiments Techniques to Optimize a Leading Edge Extension. AIAA-2003-0655, 41st Aerospace Sciences Meeting & Exhibit, Reno, Nevada, January 2003.

[14] Henderson B, Norum T and Bridges J. An MDOE Assessment of Nozzle Vanes for High Bypass Ratio Jet Noise Reduction. AIAA-2006-2543, 12th AIAA/CEAS Aeroacoustics Conference (27th AIAA Aeroacoustics Conference), Cambridge, Massachusetts, May 2006.

[15] Davis M C, Sim A G, Rhode M N and Johnson K D. Wind-Tunnel Results of the B-52B with the X-43A Stack. Journal of Spacecraft and Rockets, Vol. 44, No. 4, 2007.

[16] Parker P A, Morton M, Draper N and Line W. A Single-Vector Force Calibration Method Featuring the Modern Design of Experiments. AIAA-2001-0170, 39th Aerospace Sciences Meeting and Exhibit, Reno, Nevada, January 2001.

[17] Landman D, Simpson J, Mariani R, Ortiz F and Britcher C. Hybrid Design for Aircraft Wind-Tunnel Testing Using Response Surface Methodologies. Journal of Aircraft, Vol. 44, No. 4, pp. 1214-1221, 2007.

[18] Bordelon W J, Frost A L and Reed D K. Stage Separation Wind Tunnel Tests of a Generic Two-Stage-to-Orbit Launch Vehicle. AIAA-2003-4227, 21st Applied Aerodynamics Conference, Orlando, Florida, 2003.

[19] Staff of Stat-Ease, Inc. *Handbook for Experimenters, Version 07.3*. Stat-Ease, Inc., Minneapolis, Minnesota, 2007.

[20] Küchemann D. *The Aerodynamic Design of Aircraft*. 1st Edition. Pergamon Press Inc., New York, 1978.

Copyright Statement

The authors confirm that they, and/or their company or institution, hold copyright on all of the original material included in their paper. They also confirm they have obtained permission, from the copyright holder of any third party material included in their paper, to publish it as part of their paper. The authors grant full permission for the publication and distribution of their paper as part of the ICAS 2008 proceedings or as individual off-prints from the proceedings.

Subspace	*Face-Centered Design (FCD)*					
	Δx		*Δz*		*Δα*	
	Low	*High*	*Low*	*High*	*Low*	*High*
1	0	5.25	0.656	4.594	0	5
2	5.25	10.50	0.656	4.594	0	5
3	5.25	10.50	4.594	9.188	0	5
4	10.50	16.00	0.656	4.594	0	5
5	10.50	16.00	4.594	9.188	0	5
6	16.00	22.00	0.656	4.594	0	5
7	16.00	22.00	4.594	9.188	0	5
8	16.00	22.00	9.188	13.13	0	5
9	22.00	28.00	0.656	4.594	0	5
10	22.00	28.00	4.594	9.188	0	5
11	22.00	28.00	9.188	13.13	0	5
Subspace	*Spherical CCD*					
	Δx		*Δz*		*Δα*	
	Low	*High*	*Low*	*High*	*Low*	*High*
1′	1.165	4.835	1.0	4.0	1	4
2′	4.835	9.165	1.0	4.0	1	4
4′	9.165	14.835	1.0	4.0	1	4
5′	9.165	14.835	4.0	7.0	1	4

Table I. Inference subspace boundaries.

Model Development	**10**
Grand Mean	*1*
Regressors	*9*
Model Assessment	**10**
Lack of Fit	*5*
Data Quality	*5*
Block Effects	*1*
Pure Error	*4*
Total	**20**

Table II. Degrees of freedom budget in a full second-order model.

Design	Aerodynamic Coefficient	Comparisons within Precision Limits	Total Comparisons	Critical Binomial Number
FCD	CL1	61	70	63
FCD	CD1	57	70	63
FCD	CM1	39	70	63
FCD+added terms	CL1	25	25	22
FCD+added terms	CD1	25	25	22
FCD+added terms	CM1	25	25	22
FCD augmented to D-optimal	CL1	25	25	22
FCD augmented to D-optimal	CD1	25	25	22
FCD augmented to D-optimal	CM1	25	25	22
Spherical CCD	CL1	12	12	10
Spherical CCD	CD1	12	12	10
Spherical CCD	CM1	12	12	10

Table III. Subspace discontinuity comparisons for all designs.

Design	Aerodynamic Coefficient	Confirmation Points within 95% Prediction Interval	Total Confirmation Points	Critical Binomial Number
FCD	CL1	21	22	19
FCD	CD1	20	22	19
FCD	CM1	18	22	19
FCD+added terms	CL1	3	10	8
FCD+added terms	CD1	4	10	8
FCD+added terms	CM1	1	10	8
FCD augmented to D-optimal	CL1	3	10	8
FCD augmented to D-optimal	CD1	1	10	8
FCD augmented to D-optimal	CM1	3	10	8
Spherical CCD	CL1	2	16	14
Spherical CCD	CD1	2	16	14
Spherical CCD	CM1	0	16	14

Table IV. Confirmation point comparisons for all designs.

Response: CL1 (Face-Centered Design)

	A	B	C	AB	AC	BC	A^2	B^2	C^2
SS	Δx	Δz	α	$\Delta x\Delta z$	$\Delta x\alpha$	$\Delta z\alpha$	Δx^2	Δz^2	α^2
1	X	X	X	X				X	
2	X	X	X					X	
3		X	X					X	
4	X	X	X	X				X	
5	X	X	X	X					
6			X						
7	X	X	X	X				X	
8	X	X	X	X					
9	X	X	X	X				X	
10			X						
11			X						

Response: CL1 (Face-Centered Design + Additional Higher-Order Terms)

	A	B	C	AB	AC	BC	A^2	B^2	C^2	ABC	A^2B	A^2C	AB^2	A^2B^2
SS	Δx	Δz	α	$\Delta x\Delta z$	$\Delta x\alpha$	$\Delta z\alpha$	Δx^2	Δz^2	α^2	$\Delta x\Delta z\alpha$	$\Delta x^2\Delta z$	$\Delta x^2\alpha$	$\Delta x\Delta z^2$	$\Delta x^2\Delta z^2$
1	X	X	X	X	X	X	X	X			X	X	X	X
2	X	X	X	X	X	X	X	X		X	X		X	X
3	X	X	X	X			X	X			X		X	X
4	X	X	X	X	X		X	X			X		X	X
5	X	X	X	X	X	X	X	X		X	X		X	X

Response: CL1 (Face-Centered Design Augmented to D-Optimal)

	A	B	C	AB	AC	BC	A^2	B^2	C^2	ABC	A^2B	A^2C
SS	Δx	Δz	α	$\Delta x\Delta z$	$\Delta x\alpha$	$\Delta z\alpha$	Δx^2	Δz^2	α^2	$\Delta x\Delta z\alpha$	$\Delta x^2\Delta z$	$\Delta x^2\alpha$
1	X	X	X	X	X	X	X	X	X	X	X	
2	X	X	X	X	X	X	X	X	X	X	X	X
3	X	X	X	X		X	X	X			X	
4	X	X	X	X	X	X	X	X	X		X	
5	X	X	X	X	X	X	X	X	X	X	X	X

	AB^2	B^2C	A^3	B^3	C^3	A^2B^2	A^3B	A^3C	AB^3	B^3C	A^4	B^4
SS	$\Delta x\Delta z^2$	$\Delta z^2\alpha$	Δx^3	Δz^3	α^3	$\Delta x\Delta z^2$	$\Delta x^3\Delta z$	$\Delta x^3\alpha$	$\Delta x\Delta z^3$	$\Delta z^3\alpha$	Δx^4	Δz^4
1	X	X	X	X		X	X		X			
2	X	X	X	X		X			X	X	X	
3	X		X	X		X	X					
4	X	X		X	X	X				X		
5	X	X	X	X	X	X	X	X		X		X

Response: CL1 (Spherical CCD)

	A	B	C	AB	AC	BC	A^2	B^2	C^2	ABC	A^2B	A^2C	AB^2	A^2B^2
SS	Δx	Δz	α	$\Delta x\Delta z$	$\Delta x\alpha$	$\Delta z\alpha$	Δx^2	Δz^2	α^2	$\Delta x\Delta z\alpha$	$\Delta x^2\Delta z$	$\Delta x^2\alpha$	$\Delta x\Delta z^2$	$\Delta x^2\Delta z^2$
1′	X	X	X	X	X	X	X	X			X	X	X	X
2′	X	X	X	X	X	X	X	X		X	X		X	X
4′	X	X	X	X	X	X	X	X	X	X	X		X	X
5′	X	X	X	X	X	X	X	X	X	X	X		X	X

Table V. Final regression model terms.

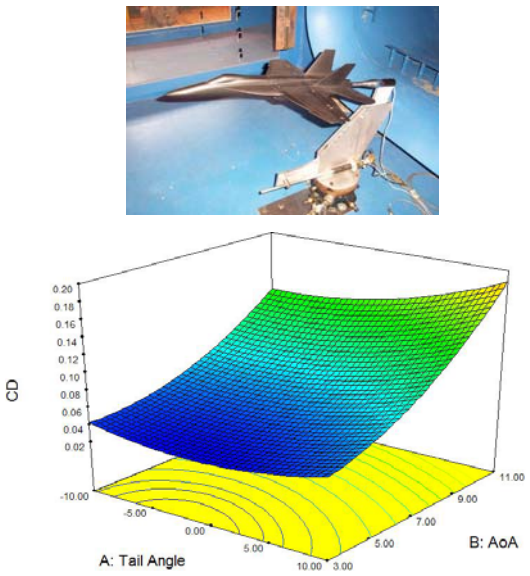


Fig. 1. Drag coefficient response surface.

Reference Dimensions:
Area (Sref): 19.18 in.^{2}
Chord (Lref): 13.13 in.
Span (Bref): 6.46 in.
Moment Reference Center (MRC)
8.93 in. from nose (0.68Lref)

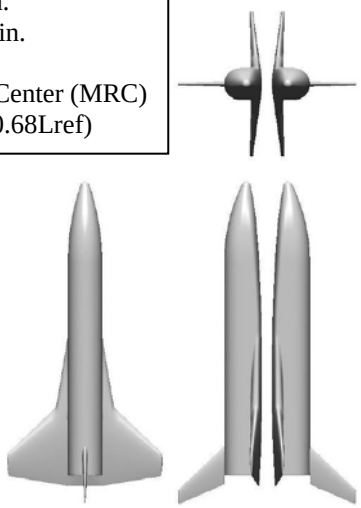


Fig. 2. LGBB bimese wing-body configuration in belly-to-belly arrangement.

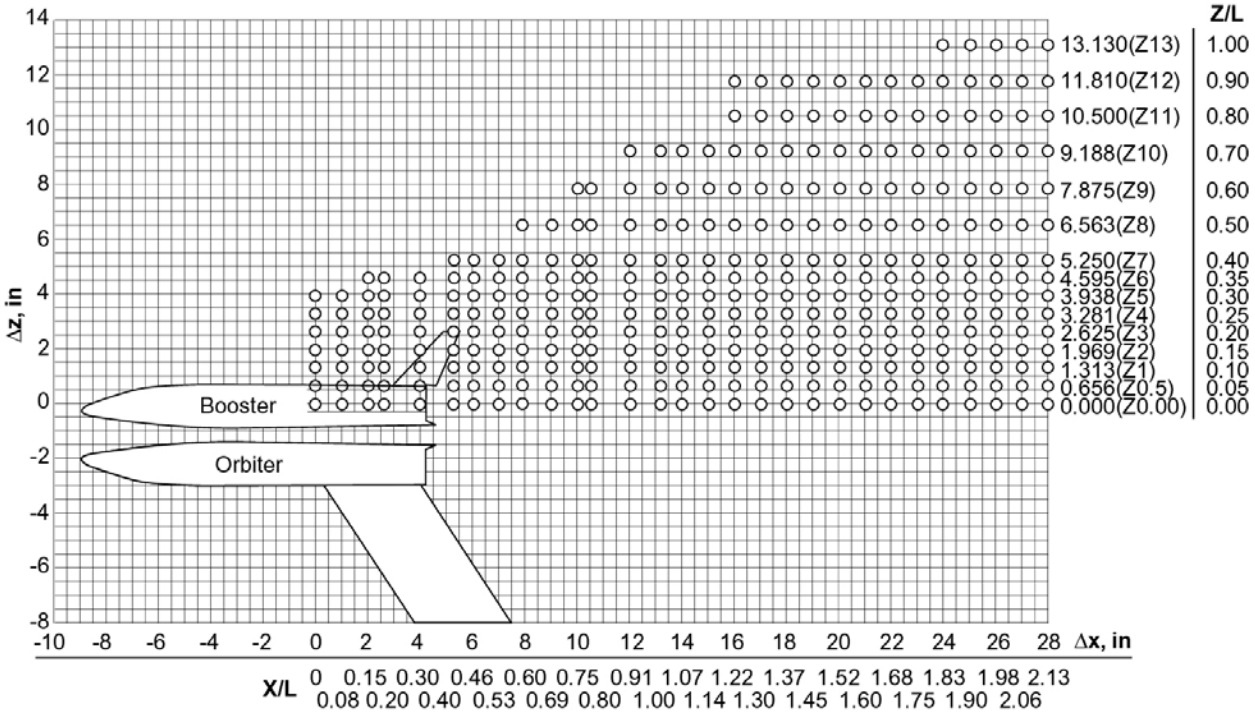


Fig. 3. Proximity matrix for OFAT testing of the NASA Langley LGBB bimese wing-body configuration in belly-to-belly arrangement.

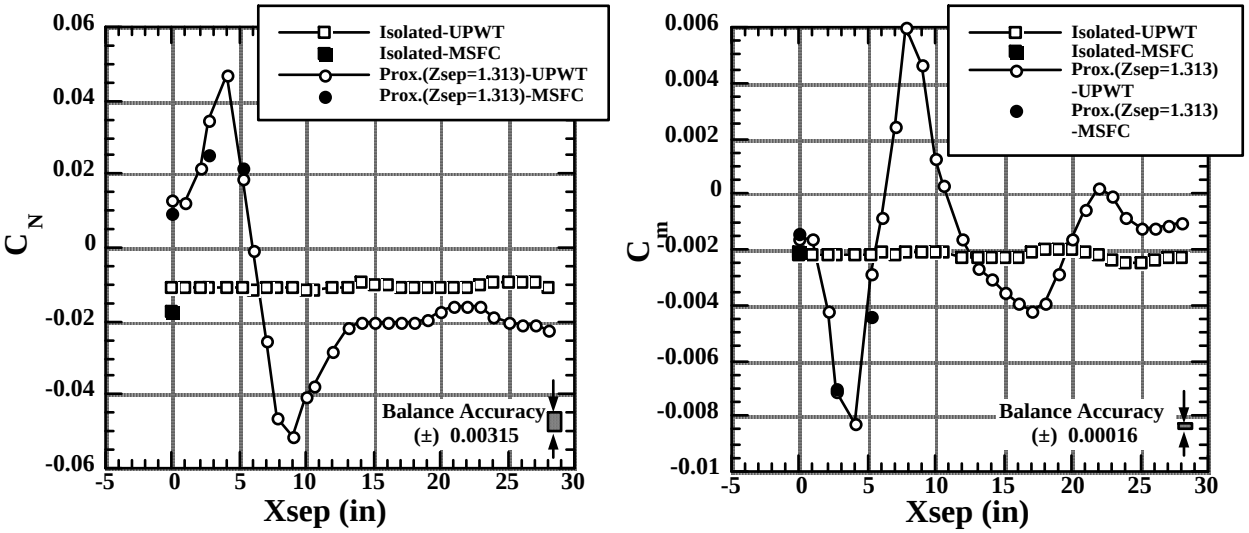


Fig. 4. Booster model normal force and pitching moment coefficients for isolated and proximity data [4] for Zsep=4.594 in. at *M*=3.

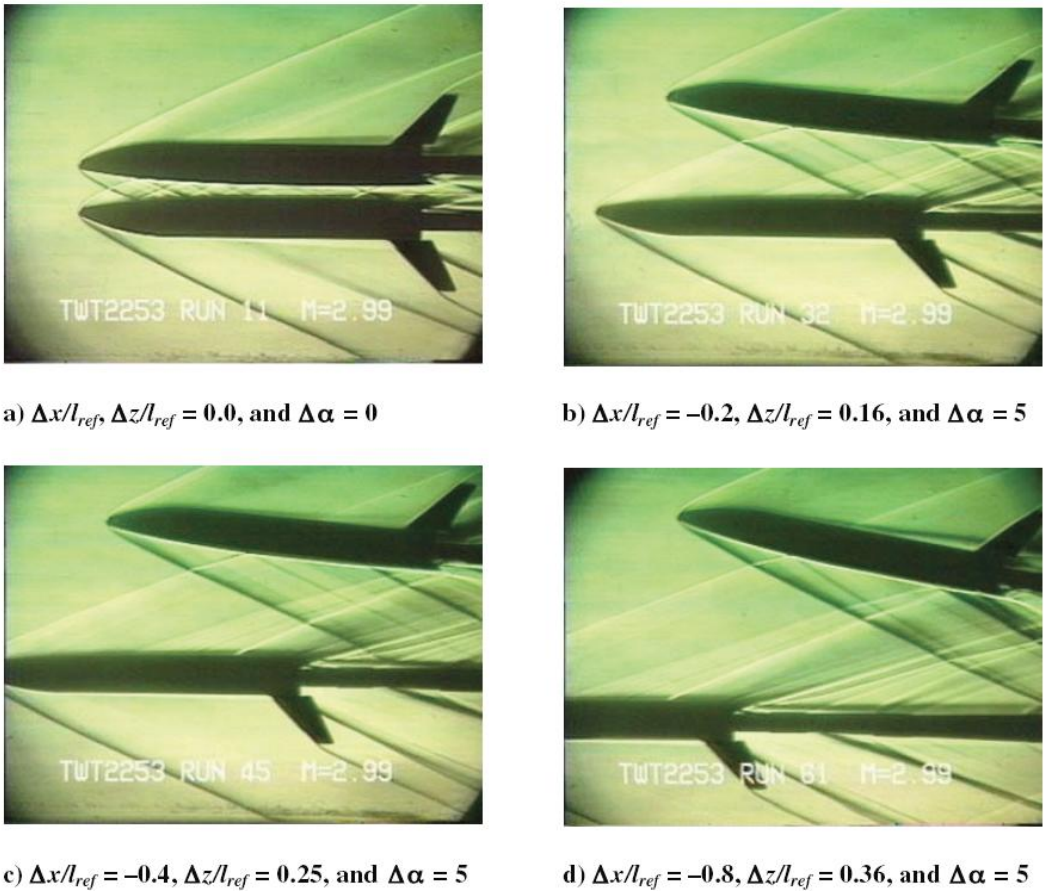


Fig. 5. Schlieren flow visualization [18] of the LBGG configuration in the NASA MSFC 14-Inch Tri-Sonic Tunnel at *M*=3.

Fig. 6. Photographs of the LBGG configuration installed in Test Section 2 of the NASA LaRC UPWT.

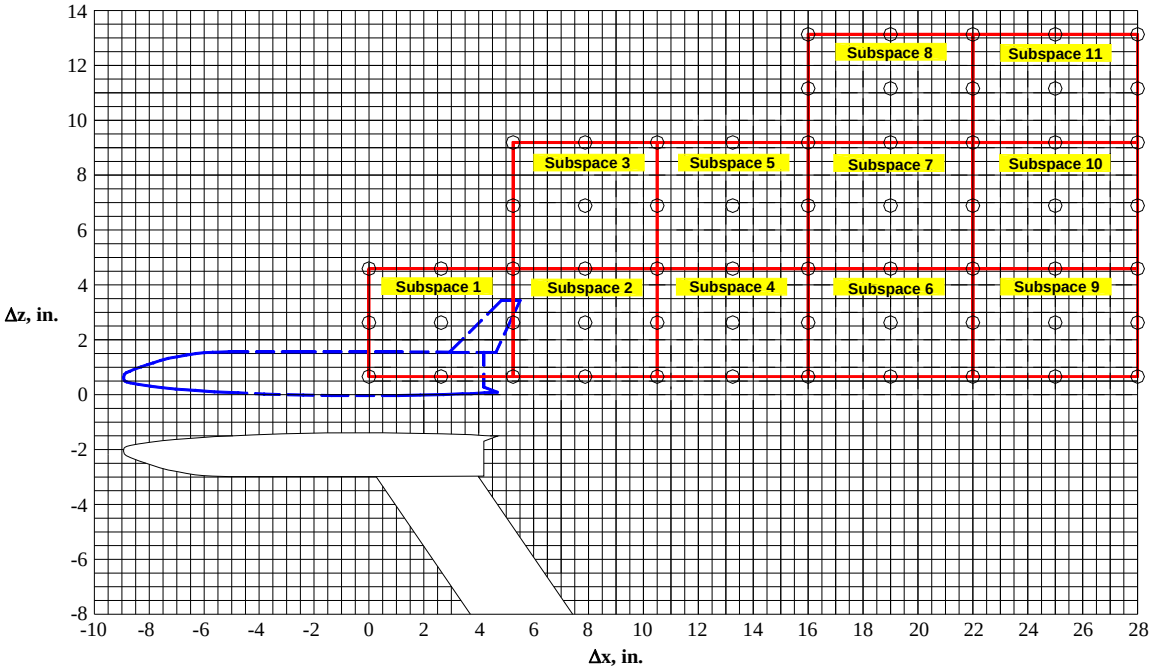


Fig. 7. Partitioning of the overall inference space for the LBGG stage separation experiment.

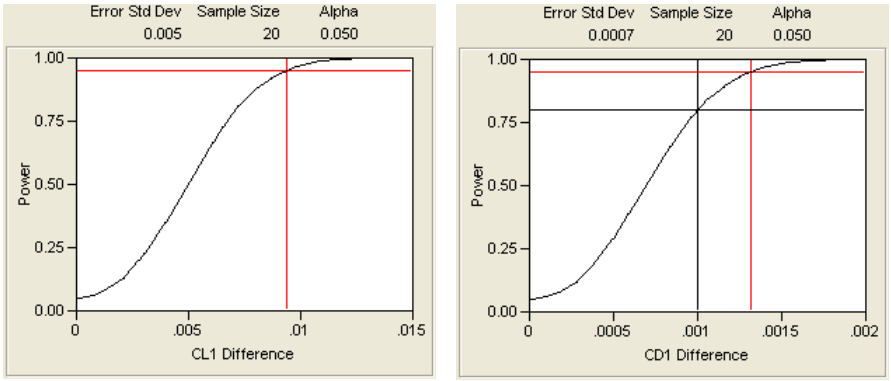
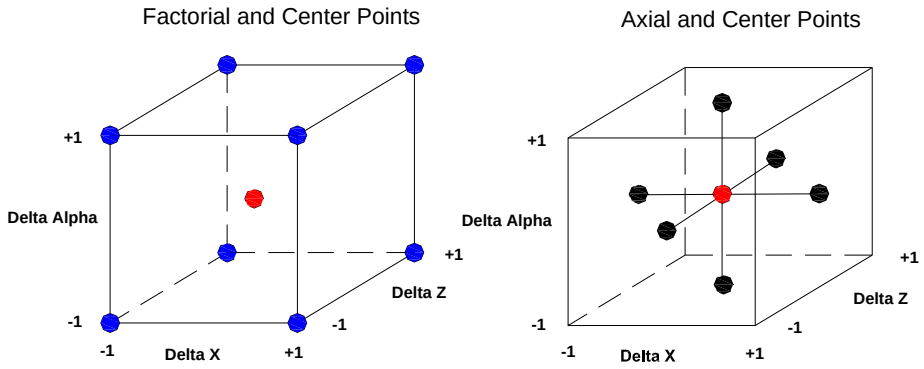
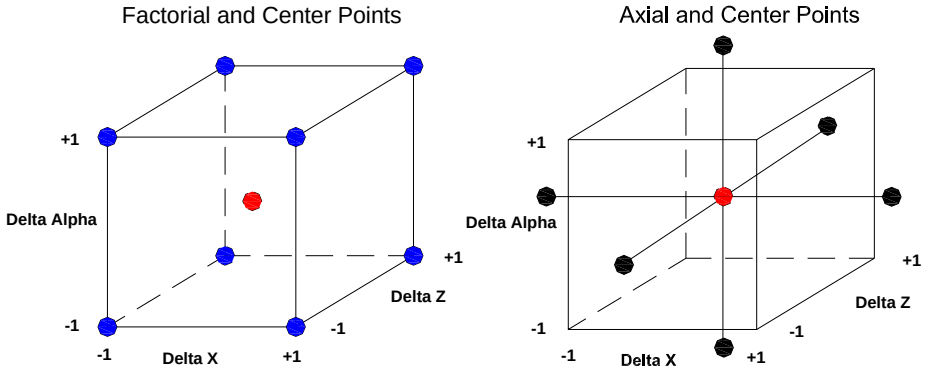


Fig. 8. Power curves for the lift and drag coefficients.



(a) face-centered design (cuboidal CCD)



(b) spherical CCD

Fig. 9. Distribution of points in the experimental designs.

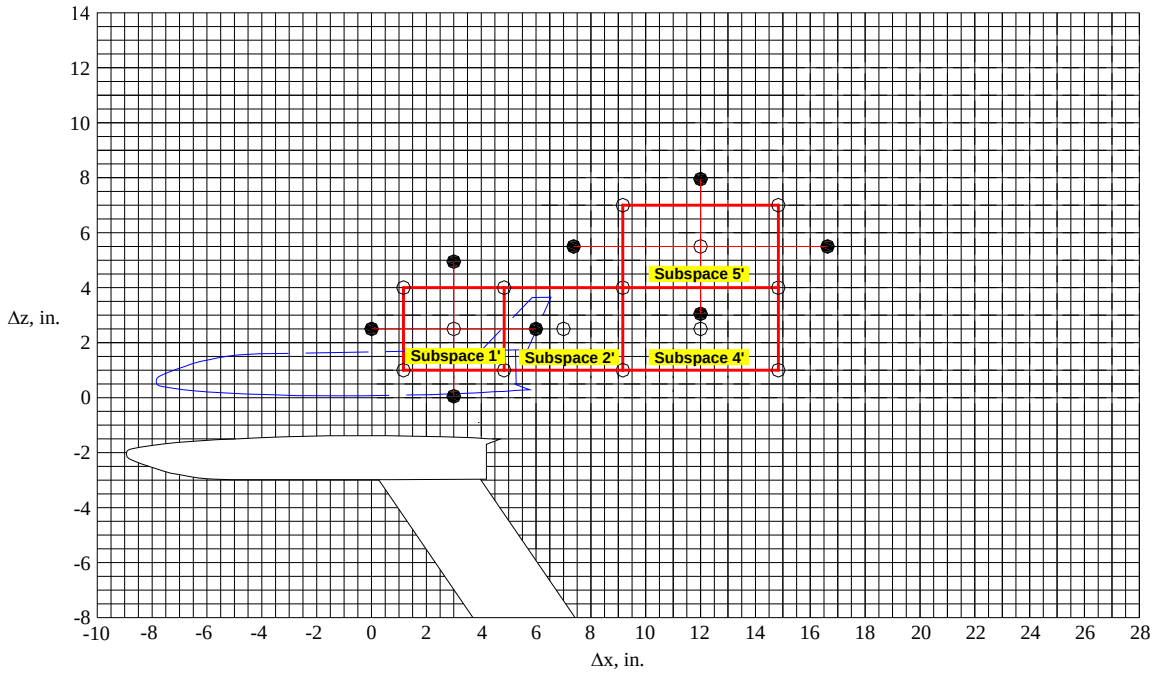


Fig. 10. Four subspaces with spherical CCD.

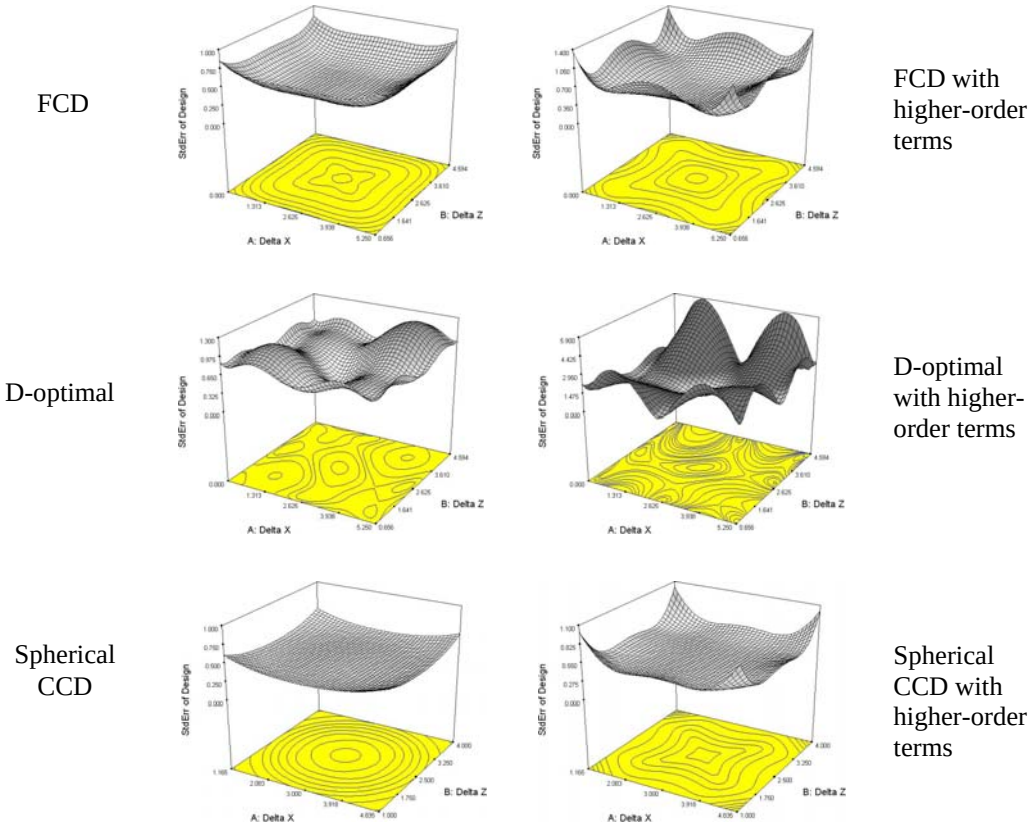


Fig. 11. Unit standard error of design for FCD, D-optimal, and spherical CCD.

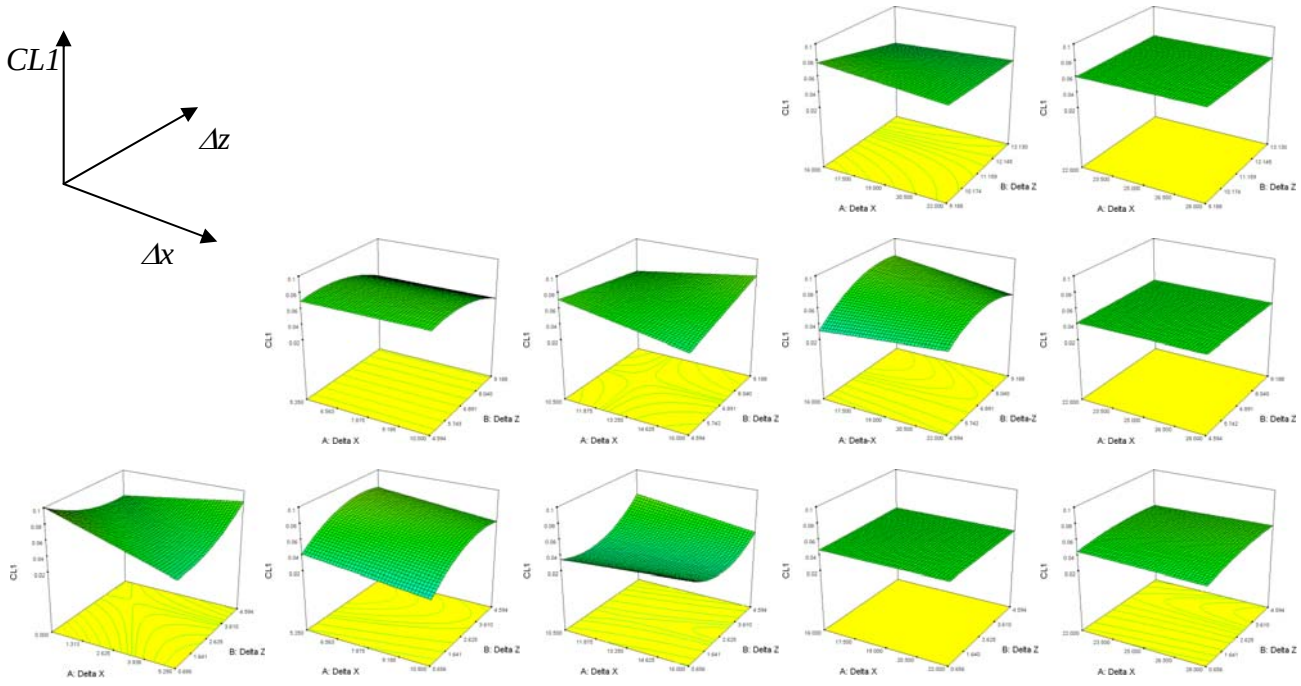
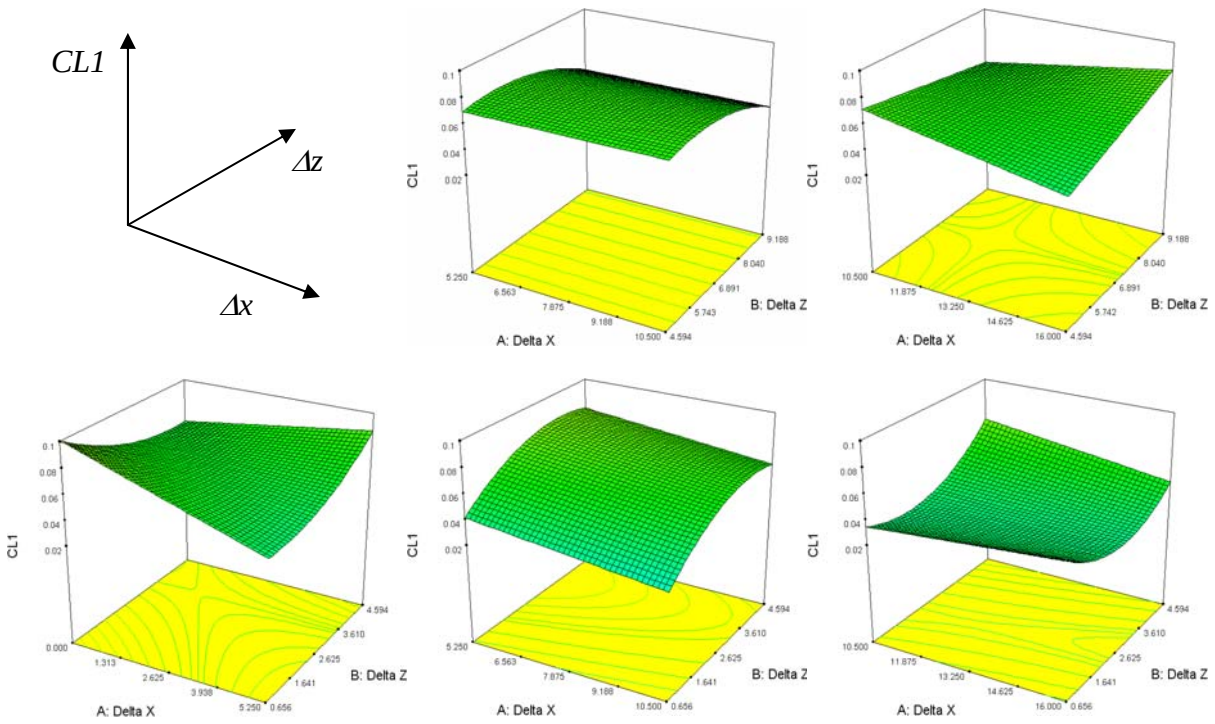
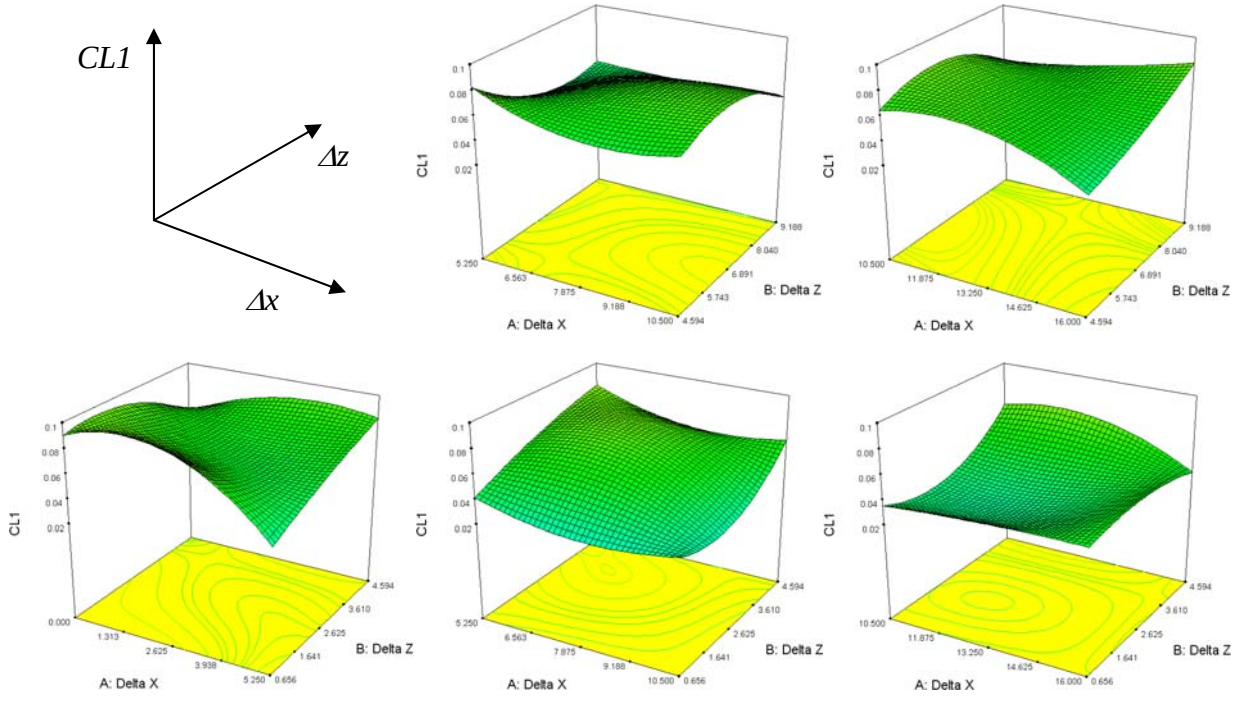


Fig. 12. Composite plot of booster model lift coefficient response surfaces in subspaces 1-11; $\Delta\alpha$ = 2.5 degrees.



(a) simplified FCD



(b) FCD with additional higher-order terms

Fig. 13. Comparison of lift coefficient response surfaces in subspaces 1-5 from simplified FCD and FCD with additional higher-order terms; $\Delta\alpha$ = 2.5 degrees.

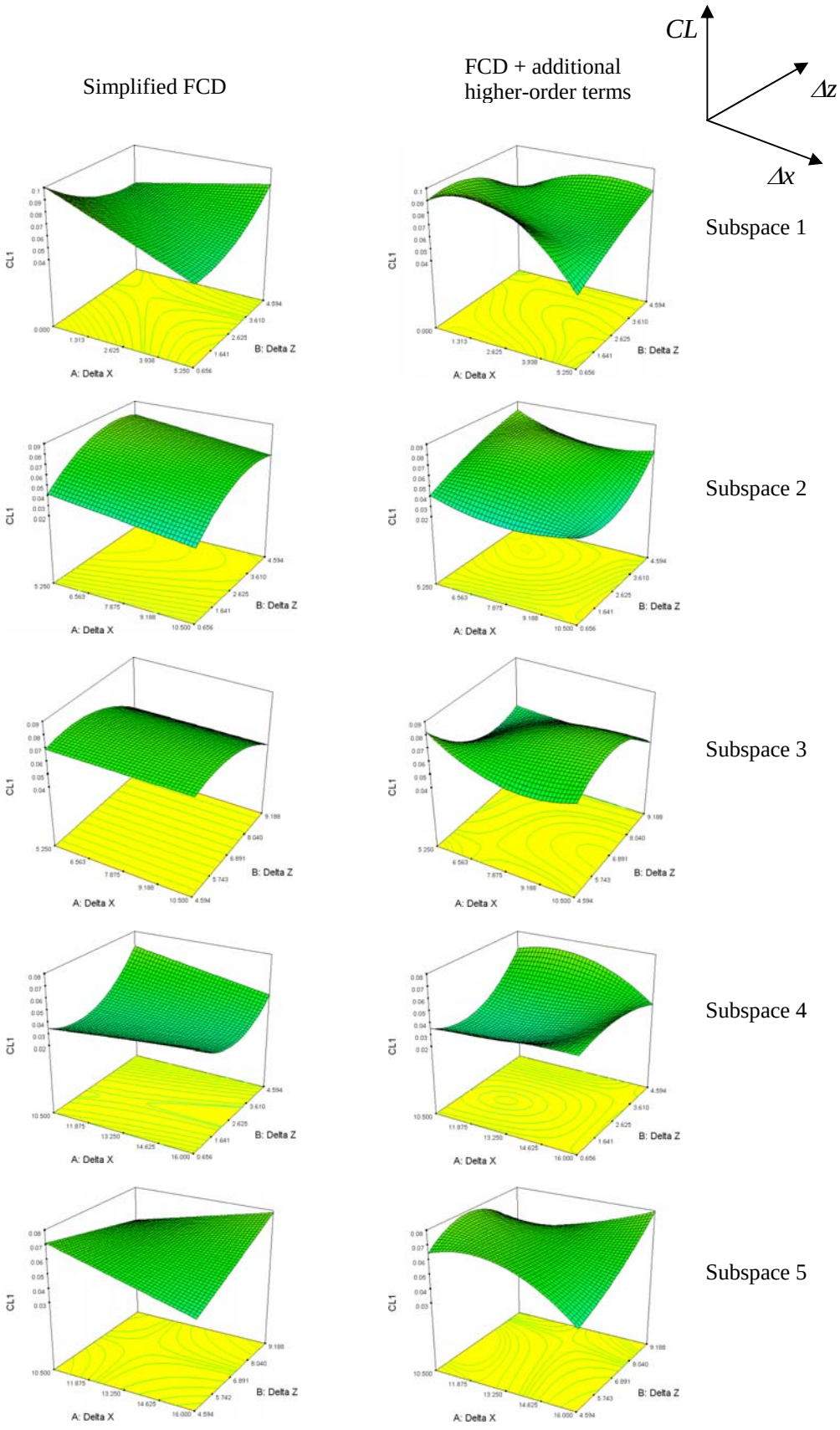


Fig. 14. Comparison of lift coefficient response surfaces with simplified FCD and FCD with additional higher order terms; $\Delta\alpha$ = 2.5 degrees.

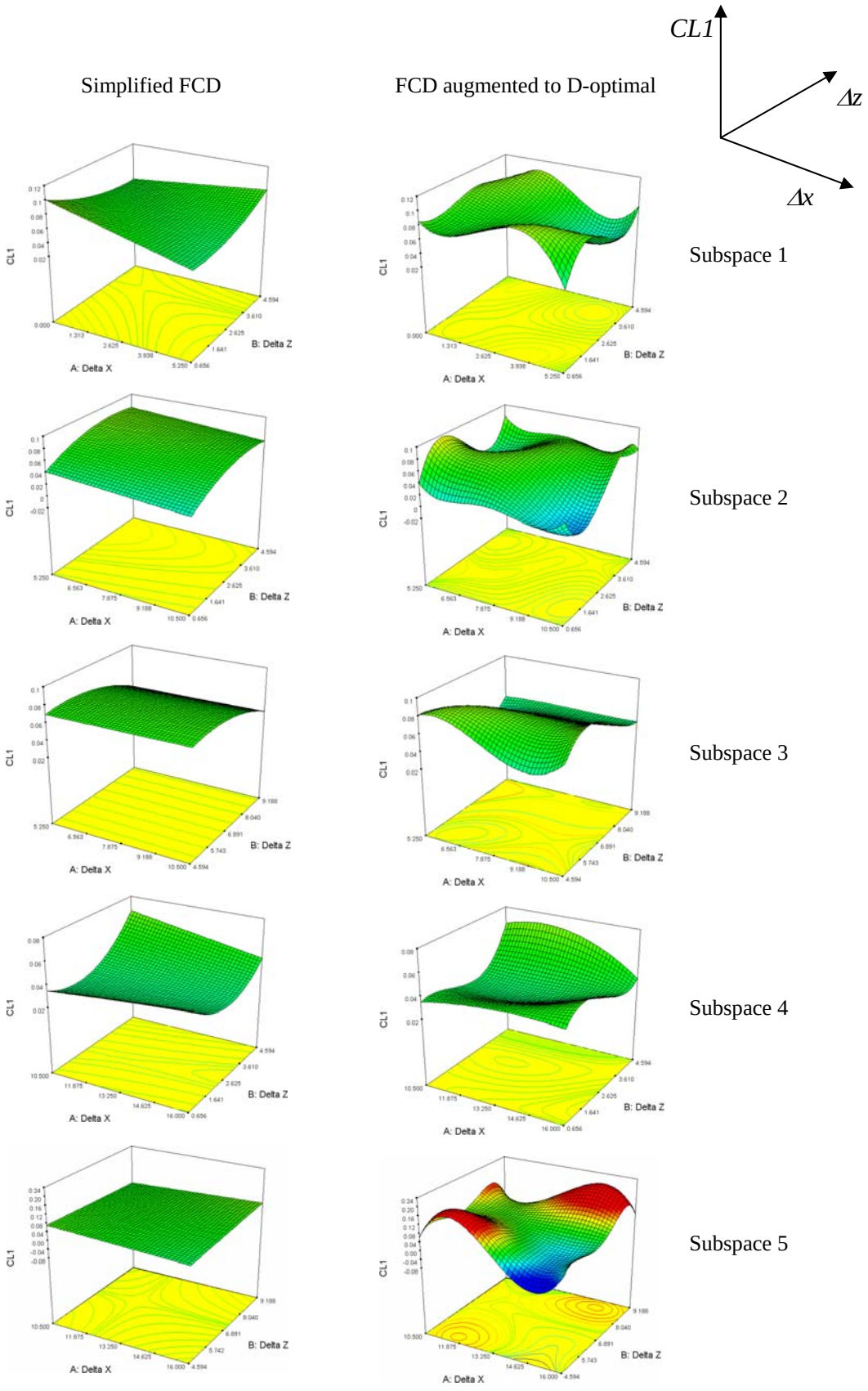


Fig. 15. Comparison of lift coefficient response surfaces with simplified FCD and FCD augmented to D-optimal; $\Delta\alpha$ = 2.5 degrees.

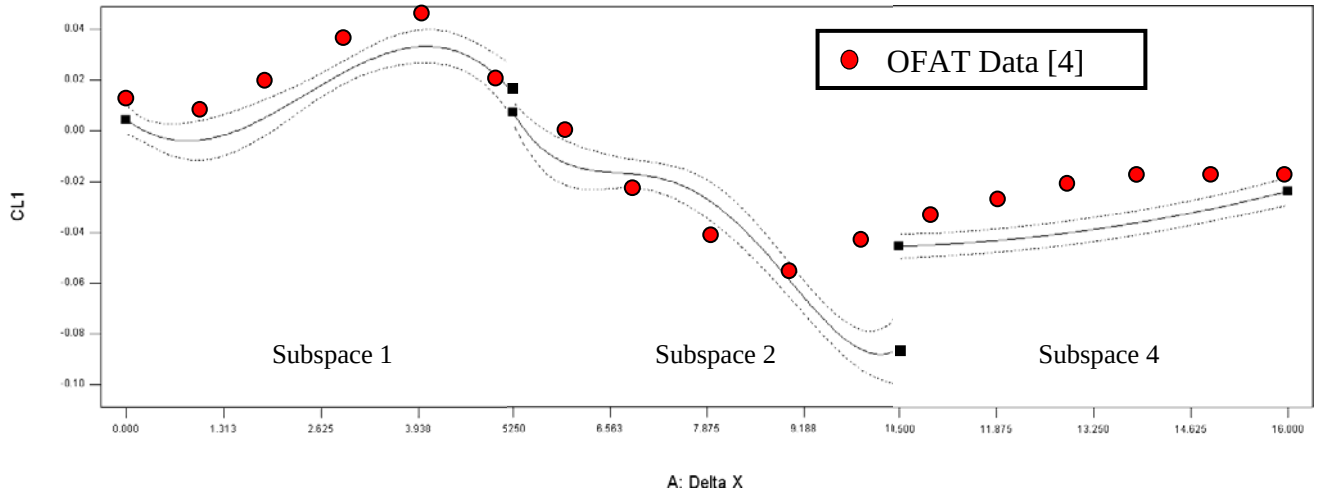


Fig. 16. Comparison of predicted lift coefficient from D-optimal design and observed data from OFAT test [4]; subspaces 1, 2, and 4; Δz = 1.313 inches, $\Delta\alpha$ = 0 degrees.

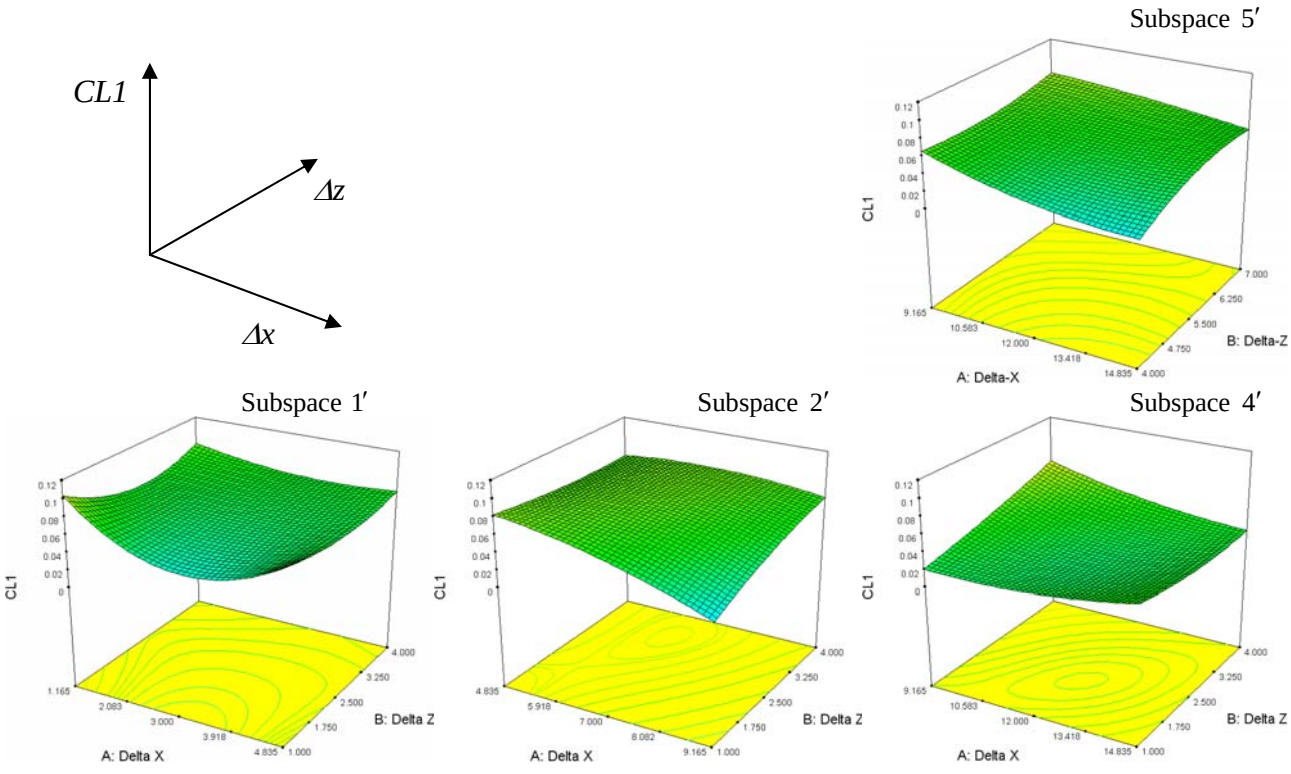


Fig. 17. Composite plot of booster model lift coefficient response surfaces in subspaces 1′, 2′, 4′, and 5′ ; spherical CCD; $\Delta\alpha$ = 2.5 degrees.

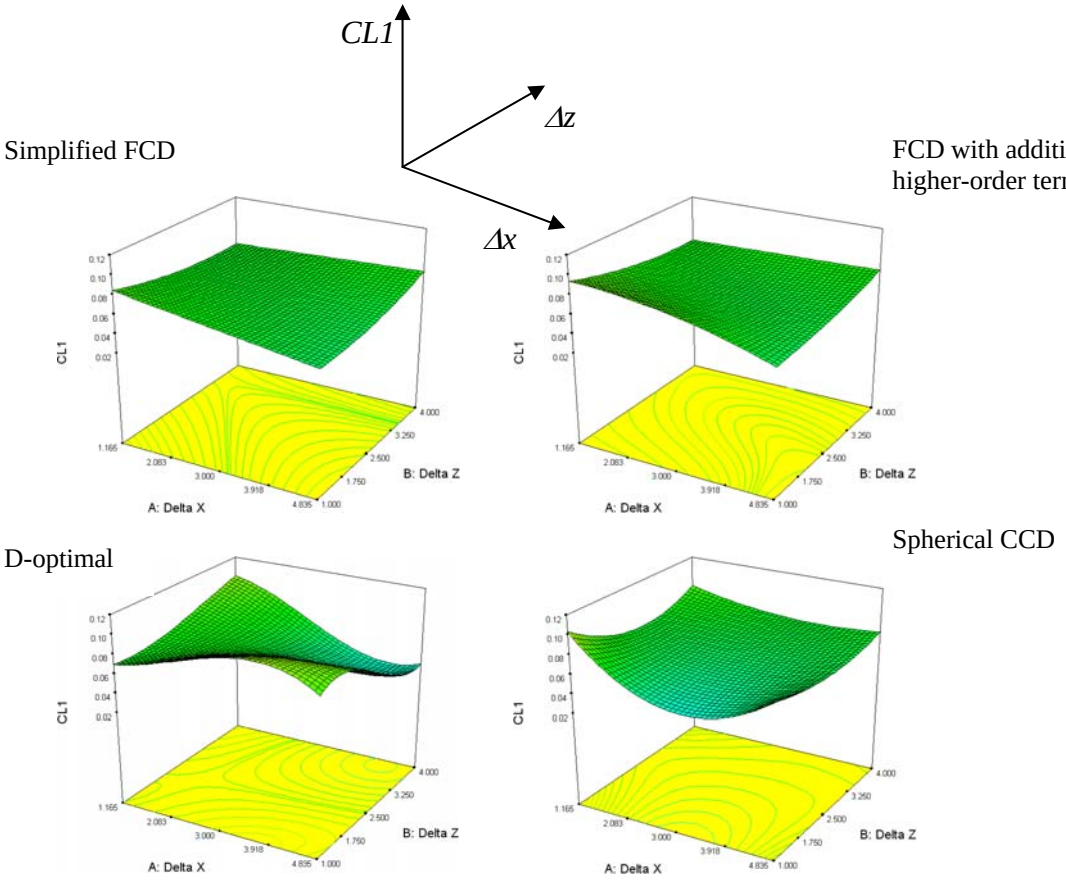


Fig. 18. Comparison of lift coefficient response surfaces for all designs and empirical models in subspace 1; inference space trimmed to match the spherical CCD; $\Delta\alpha$ = 2.5 degrees.

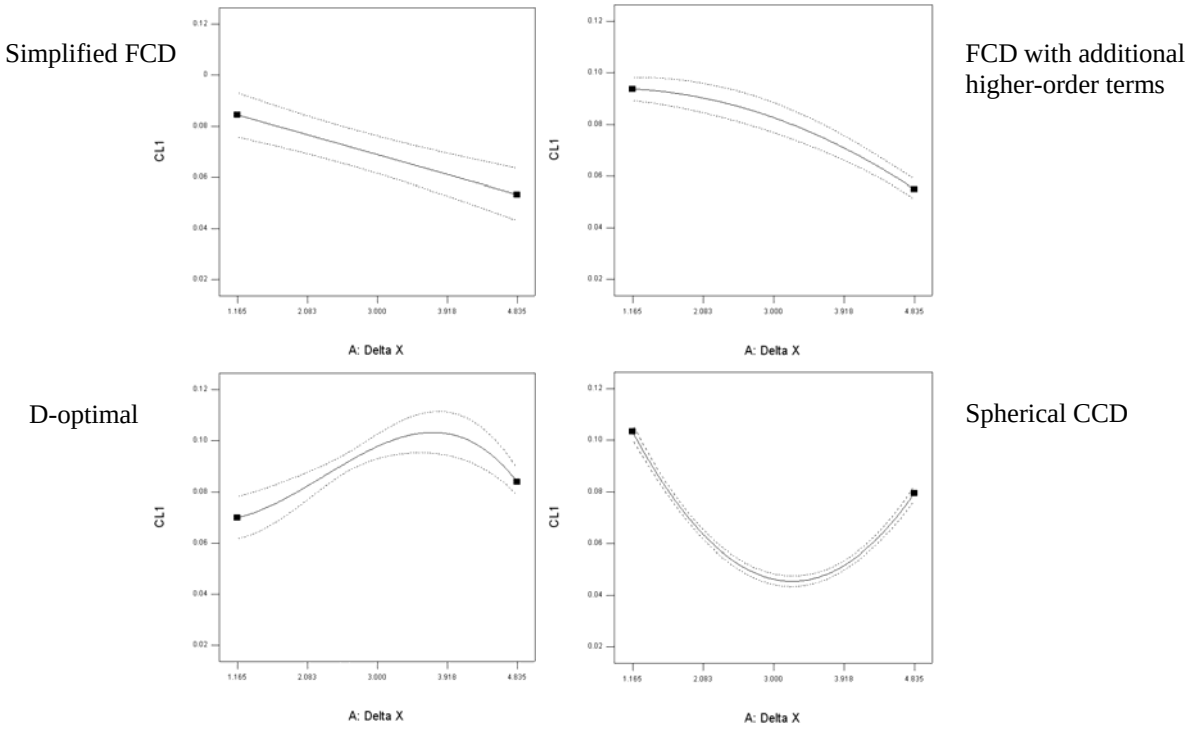


Fig. 19. Comparison of predicted lift coefficient versus Δx for different experimental designs; Δz = 1.0 inches, $\Delta\alpha$ = 2.5 degrees.

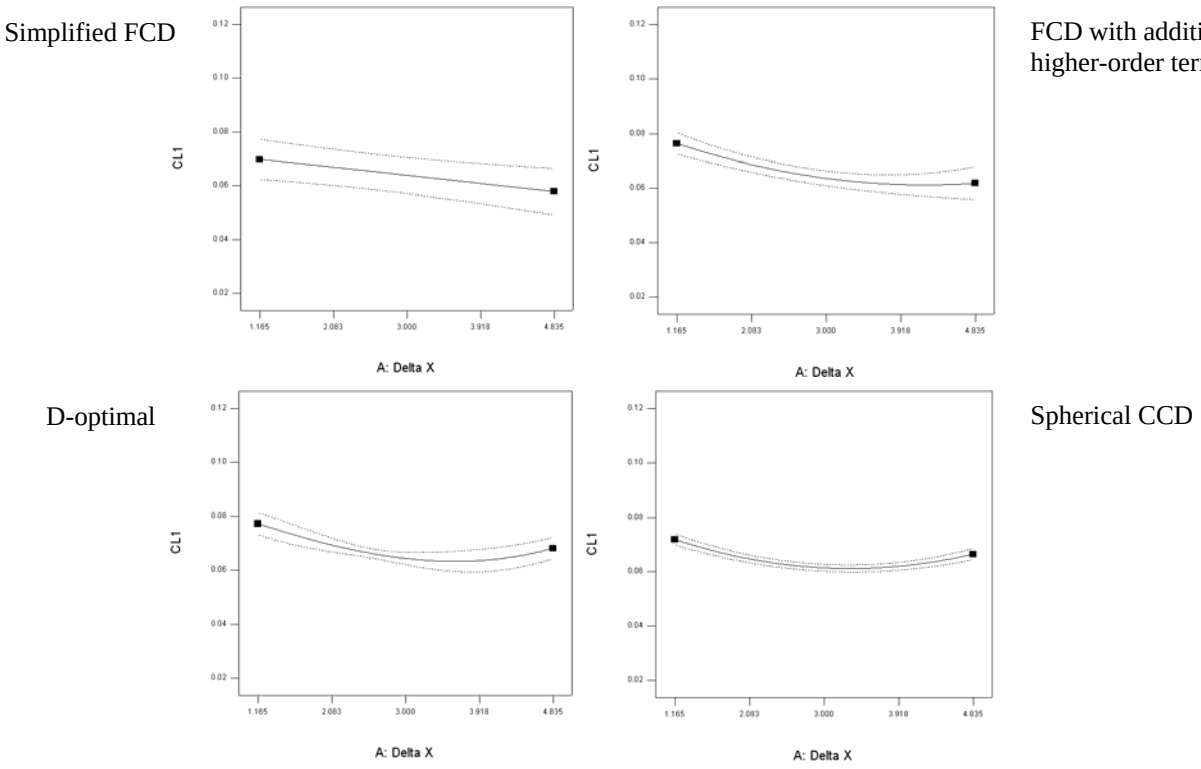


Fig. 20. Comparison of predicted lift coefficient versus Δx for different experimental designs; Δz = 2.5 inches, $\Delta\alpha$ = 2.5 degrees.