Innovative Fusion of Experiment and Analysis for Missile Design and Flight Simulation

Patrick H. Reisenthel, John F. Love, Daniel J. Lesieutre, and Marnix F. E. Dillenius
Nielsen Engineering & Research, Inc.,
605 Clyde Avenue, Suite 200
Mountain View, CA 94043-2241, USA

{phr,jlove,lesieutre,mfed}@nearinc.com

ABSTRACT

The integration of experimental data and computational databases is key to supporting decisions during the development of missile systems. An innovative technique is demonstrated to increase the accuracy of databases used for comprehensive flight simulations of missiles. This technique uses multidimensional response surface technology to mutually enhance heterogeneous data sets. An important application of this technology is the use of sparse data points from limited wind tunnel tests to correct/calibrate computational databases used in flight simulations.

1.0 LIST OF SYMBOLS AND ABBREVIATIONS

b	RBF scale parameter
c_k	basis function coefficients, elements of $[C]$
$[C]$	solution vector ($[c_1, c_2, \ldots, c_p]^T$)
$[\hat{C}]$	regression estimate for $[C]$
C	load coefficient (generic)
$C_{l,n}$	nose rolling moment coefficient
CFD	computational fluid dynamics
cov	covariance
$[e]$	regression error
$E[.]$	expected value
f	shape function
F	global interpolant (output of RBF network)
$\hat{F}$	regression estimate for F
$[I]$	identity matrix
M	Mach number
N	dimension (number of independent variables)
NEAR	Nielsen Engineering & Research

p	number of support vectors
$\lfloor Q_{1,2} \rfloor$	orthogonal matrices
RBF	radial basis function
RMS	root mean square
RS	response surface
SVD	singular value decomposition
SVM	support vector machine
var	variance
x_i	i-th independent variable (i-th coordinate of X vector)
X	independent variables vector
Y	dependent variable
α_c	total incidence angle
δ_{ij}	Kroenecker delta tensor
δY_k	seed uncertainty of dependent variable at point k
ε	auxiliary variable
φ	roll angle
φ_k	k-th radial basis function
χ_k	radial basis function center for φ_k
σ^2	regression error variance
$[\Sigma]$	diagonal matrix of singular values
$[\Sigma^+]$	truncated inverse of $[\Sigma]$

2.0 MOTIVATION

It is well-known that pointwise, or even line plot, comparisons between various data sets can be deceiving. This is particularly true in regions where large gradients exist. The goal of the present paper is to present a data processing technique that helps the engineer develop a global understanding of the data, specifically limited experimental test data, with the aid of physics-based computational information. A data fusion technique is used to produce a response surface acting as a global interpolant of the data, both computational and experimental. The advantage of this technique, as opposed to conventional interpolation or data fitting techniques, is that the interpolation of the experimental data can be regarded as essentially computational (model) based. In other words, the physics of numerical simulations can be used to interpolate (and, possibly, extrapolate) the experimental data where the sampling is sparse or even absent, while still respecting the integrity of the experimental data. Vice versa, the resulting metamodel representation can also be regarded as a calibration of the computational model based on experimental data.

Global metamodels and response surface technology have been used in a variety of fields, including structural reliability, instrument calibration, and aerodynamic and trajectory optimization, to name a few [1-10]. These models are a critical part of surrogate-based analysis and optimization [11,12]. A lesser known application, however, is a rational process for fusing data from disparate sources. Designers are frequently confronted with the problem of effectively integrating data from multiple sources (theoretical, numerical, experimental) [13], while appropriately weighing uncertainty, past experience, and prior knowledge.

A systematic framework for aiding the designer and analyst in achieving this variable fidelity, multisource, multidimensional integration has been developed. This framework uses robust multidimensional data generalization techniques which have their roots in machine learning methods such as neural networks [14,15], support vector machines [16], and other kernel methods [17]. The particular approach used in this paper is based on self-training radial basis function networks which form the basis of the NEAR-RS (response surface) technology.

NEAR-RS is a software system consisting of two modules: a metamodel (response surface) identification module, and a metamodel evaluation/interrogation module. A graphical user interface included in this second module serves as a multidimensional viewer facilitating the visualization of trends in high-dimensional data [18]. A key aspect of the technology is the ability to estimate further sampling needs and model quality, based on automatic uncertainty estimation. The application discussed in this paper illustrates the data adaptivity and data fusion capabilities of the method by considering the problem of assimilating missile data from a wind tunnel test into a comprehensive aerodynamic database for guidance and control.

3.0 TECHNICAL BACKGROUND

Response surface methods can be used to perform data fusion operations in order to enhance the usefulness of limited experimental data. The problem is akin to interpolating and extrapolating the data outside of the range where these data were collected, a task which, without any regularizing assumptions, constitutes a fundamentally ill-posed problem [13]. Regularizing assumptions can come in various forms: physics based models, mathematical equations (such as splines), implicit smoothness assumptions, or other empiricisms. The method used here employs a particular form of regularization, in which a hypersurface going through the experimental data is "supported" by additional computational constraints. We present, first, the basic theory behind the response surface identification and its uncertainty, and, second, examine how it can be applied to the problem of data fusion.

3.1 Theory

The task of formulating a response surface in *N*-dimensional space amounts to identifying a smooth mapping $F : \mathrm{R}^N \rightarrow \mathrm{R}$ on the basis of *p* available data points. If this response surface acts as an interpolant, then the function *F* must satisfy the constraints

$$F(\mathbf{X}_i) = Y_i, \quad i = 1, \ldots, p \quad \textbf{(1)}$$

where each X_i represents a vector of independent variables (for example, spatial coordinates, flow conditions, and/or configuration parameters), and each Y_i is a dependent variable (for example, pressure). In the case where *F* represents, instead, a fit to the data, then the response surface is required to minimize the distance $\| F(X_i) - Y_i \|$, typically in the least squares sense.

This goal can be achieved by a number of different means, for example Kriging [19,20], which is used in the popular DACE stochastic process model [21], multivariate adaptive splines [22], and Support Vector Machine (SVM) [16] algorithms. The goal of this paper is not to compare these methods to each other, but to illustrate how this class of methods can be used to achieve the goal of fusing experimental and computational data. NEAR's approach uses a radial basis function (RBF) network to represent the function *F*. In this approach, *F* is expanded into basis functions φ_k which are radially symmetric about their control point, χ_k. By analogy with SVMs, we will refer to $(\chi_k ; F(\chi_k))$ as the *support vectors* for the response surface. Thus,

$$F(X) = \sum_k c_k \varphi_k(X), \quad \varphi_k(X) = f\left(\| X - \chi_k \| ; b \right) \qquad \textbf{(2)}$$

where f is a scalar shape function, *b* is an adjustable [23] scale or stiffness parameter, and $\| . \|$ designates the Euclidean norm. The c_k are the basis function coefficients. They are parameters to be identified. Note that, if the basis functions (i.e., their shape, centers, and number) are known, then the determination of the nonlinear response surface boils down to an identification problem which is linear-in-the-parameters. In other words, the coefficients c_k are the solution of a least-squares linear problem

$$[A][C] = [Y] \qquad \textbf{(3)}$$

where each row of Eq. (3) is an instantiation of the constraints expressed in Eq. (1). Radial basis function models, such as Eqs. (1) and (2), can be viewed [24] as a three-layer feedforward neural network with linear output mapping. This is shown schematically in **Figure 1**, where the number of nodes in the hidden layer is equal to the number of basis functions, the inputs x_i are the coordinates of X **,** and the weights of the output layer are the coefficients c_k . Also, the shape function (f) is the activation function of the hidden layer nodes, which can take a number of forms, for example, Gaussian, thin plate spline, multiquadric, or reciprocal multiquadric [25,26].

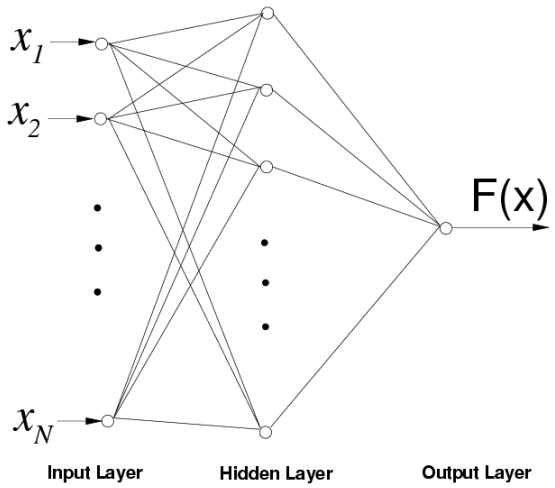

Figure 1: Radial Basis Function Network.

There are, *a priori*, a number of different ways of selecting the control points χ_k . One possible approach is the use of sequential approximation and optimization methods. This can be quite expensive, and a more efficient approach to control point selection consists of using a fixed subset of the existing training data. Algorithms such as generalized cross-validation (GCV) [27] can be used for this purpose, resulting in parsimonious networks with good generalization properties. While the use of a small number of regressors φ_k is indeed desirable from the point of view of model robustness, we will confine the present discussion to simple networks where the training data are assumed deterministic and sparse. Thus, in this particular implementation, the number of basis functions is equal to the number of training data points, and the centers (control points ξ_k) of the RBFs coincide with the data points. As a result of this simplification, there is no need for stepwise regression algorithms: the structure of the equivalent neural network (Figure 1) is automatically determined by the data.

Multidimensional response surface identification in NEAR-RS is a three-step process: (1) preconditioning/classification, (2) formation of the [A] matrix in Eq. (3), and (3) solution method for [C]. The preconditioning step uses a classification algorithm to associate data points which have the same (or substantially similar) values of the independent variables. This step is a mechanism for mitigating problems associated with overfitting dense point clusters. At present, this operation is performed on the basis of a user-defined tolerance in the independent variables. If no tolerance is prescribed, then strict equality is required in order for two points to be associated. The purpose of preconditioning is to improve the solution characteristics by improving the condition number of the [A] matrix. The formation of the [A] matrix is relatively straightforward: it involves the calculation of distances between all training data points. If the response surface uncertainty is desired, then an additional step (weighted least squares) is used. This situation is described below. Finally, the solution method uses robust pseudoinversion technology, the purpose of which is to take care of pathological situations, such as the handling of inconsistent data. Such data can occur, for example, as a result of improper or incomplete parameterization, such as repeatability tests or the existence of data from various sources (different codes, different fidelity level, algorithms, etc.) at the same or substantially similar condition. These data "inconsistencies" amount to an ill-posed problem in terms of interpolation, a difficulty which is circumvented using regularization techniques.

One important addition to these ideas is the concept of response surface uncertainty. In the following, it is shown that, due to linearity-in-the-parameters, it is possible to make use of well-established statistical results to propagate the uncertainty of the support data onto an uncertainty of the response surface itself.

Consider the original equation $[A][C]=[Y]$ as a regressor model for the data. Rewrite Eq. (3) as

$$[A][C] \;=\; [Y]+[e] \qquad \textbf{(4)}$$

where $[e]$ is the modeling error. Let $[A]=[Q_1][\Sigma][Q_2^T]$ designate the singular value decomposition of the matrix [A]. The pseudoinverse solution is then given by [28]

$$[\hat{C}] \;=\; ([Q_2][\Sigma^+][Q_1^T])[Y]$$

It can then be shown [29], under certain simplifying assumptions, that the covariance matrix of the solution vector is

$$cov[C] \;\equiv\; E\left[(\hat{C}-C)(\hat{C}-C)^T\right] \;=\; \sigma^2\, E[Q_2\Sigma^+\Sigma Q_2^T] \qquad \textbf{(5)}$$

where $E[.]$ designates the expected value, and σ^2 is the variance, presumed uniform, of $[e]$. Thus, Eq. **(5)** propagates the uncertainty in $[Y]$ onto the solution vector [C]. Alternatively, the matrix $[Q_2\Sigma^+\Sigma Q_2^T]$ can be interpreted as a sensitivity matrix which redistributes the measurement noise onto the solution vector components. This uncertainty, in turn, ties into the uncertainty on the response surface itself. Let $F(X)=\sum_k c_k\varphi_k(X)$ where the φ_k are the members of a radial basis function set. It then follows that

$$E\left[(\hat{F}-F)^2\right] \;=\; \sum_i\sum_j (cov[C])_{ij}\;\varphi_i(X)\,\varphi_j(X) \qquad \textbf{(6)}$$

Equation **(6)** represents the variance $var(F)$ of the response surface. Assume a Gaussian probability distribution. The resulting uncertainty on $F(X)$, defined as $\Delta F = \pm 3\sqrt{var(F)}$ ("three-sigma" uncertainty),

is shown in the hypothetical example of **Figure 2**. The dependent variable uncertainty of the training data is indicated in the form of vertical error bars. The resulting uncertainty on the response surface is indicated in the form of upper and lower bounds, $F \pm \Delta F$, using thin dashed lines.

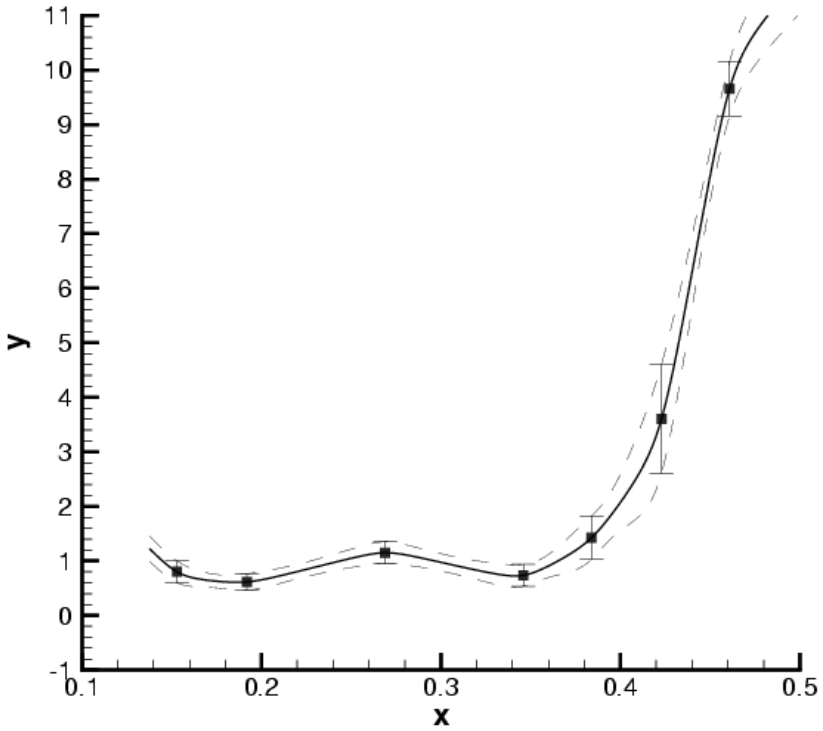

Figure 2: Response Surface (solid line) Plus/Minus Uncertainty (dashed lines).

Note that, in order to propagate uncertainty according to the method described above, the variable $[e]$ must be a stochastic variable such that

$$cov[e] = \sigma^2[I] \quad \textbf{(7)}$$

where $[I]$ is the identity matrix. In other words, $[e]$ must have zero cross-correlation, and must be of uniform variance σ^2 across all of its components. While the zero cross-correlation assumption is typically not justified if there is a deterministic bias between the regressor model and the data, it is still possible in practice to use the above equations to propagate uncertainty, by assuming that $[e]$ represents a vector of random measurement errors $[\delta Y_1, \delta Y_2, \ldots, \delta Y_p]^T$. When these seed uncertainty levels differ from support vector to support vector, such as in the example of Figure 2, then both left- and right-hand sides of Eq. **(3)** are multiplied by a weighting matrix $[W]$, where $[W]$ is defined in tensor notation as $W_{ij} = \delta_{ij} / \sqrt{var(Y_i)}$. This simple algorithm ensures that the variance of the transformed variables is uniform. At present, the uncertainties are assumed to be uncorrelated between data points. If this were not the case, then more sophisticated techniques, such as Markov estimators and/or instrumental variables [29] could be used.

3.2 Application to Data Fusion

While there are many definitions of data fusion, consider the general notion (Li et al. [30]) of data fusion defined as "the combination of a group of inputs with the objective of producing a single output of greater quality and reliability." In the present paper, we assume the existence of two data streams, one computational, the other experimental. Instead of using these data streams to validate each other directly, the point of view adopted here is to recognize and accept that there will always be differences between them, due to experimental limitations, as well as approximations in the physical models used. Fusion of these multiple data streams is then used to enhance data understanding. We will confine the analysis to the case of experimental and computational data. Specifically, we focus on the situation where the computational data are reasonably affordable to obtain, in contrast to the experimental data, which will be assumed to result from expensive wind-tunnel tests at a limited number of configurations and flow conditions. Thus, not only are the experimental and computational data not sampled at the same conditions, but the typical situation is one where the experimental data are sparse, with respect to the computational data.

The basic idea behind the use of response surface technology for the fusion of experimental and computational data is to take advantage of the radial symmetry of the basis functions to construct a metamodel that incorporates all the data. This can be done by adding one auxiliary variable $\varepsilon \equiv x_{N+1}$ to the multidimensional design space $(x_1, x_2, \ldots, x_N)$. This extra variable is binary in nature, and is used to tag whether the data are computational ($\varepsilon = 0$) or experimental ($\varepsilon = 1$). A single global response surface is then calculated in N+1 dimensions. By querying the response surface projected along $\varepsilon = 1$ one obtains a model representation which respects the integrity of the experimental data, while simultaneously "inheriting" the essential features of the computational model. To understand how the method works, consider the sketch shown in **Figure 3**.

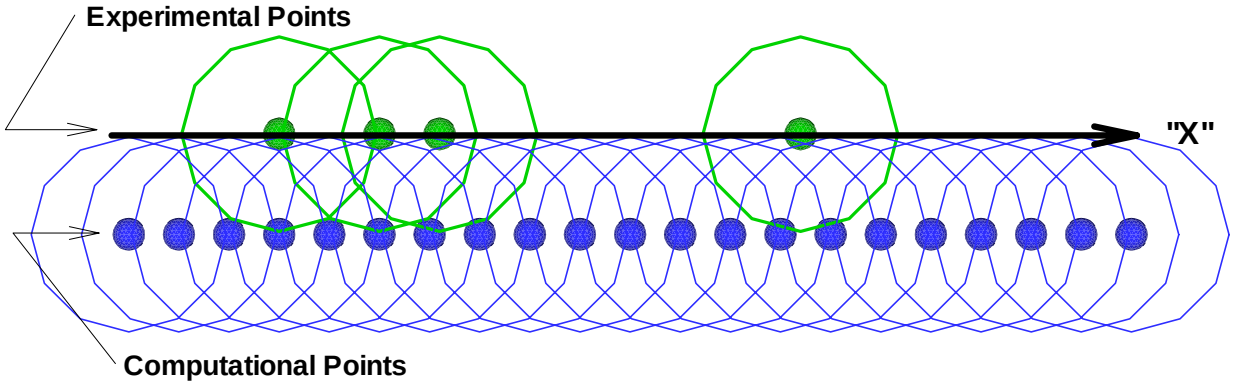

Figure 3: Schematic Illustrating the Layout of Computational and Experimental Support Vectors.

The schematic lays out the position of the experimental and computational support vectors relative to each other. The horizontal coordinate "X" symbolizes the independent variables $(x_1, x_2, \ldots, x_N)$. The vertical coordinate represents the auxiliary variable ε. The circles around each point symbolize a "region of influence" or spatial correlation associated with each radial basis function. The radius of these circles is related to the scale parameter b in Eq. (2). Thus, wherever the data sampling is high, the interpolant will be mostly influenced by the basis functions whose centers are in the immediate vicinity. On the other hand, when the experimental data points are widely separated relative to the width of the basis functions, the interpolation will be affected primarily by the computational points. The evaluation of the response surface in the experimental plane has the

effect of interpolating the experimental data in a way that is rooted not in mathematics or simple-minded smoothness assumptions, but, rather, in whatever physics are included in the computational model.

4.0 RESULTS

Two missile aerodynamics applications of the method are presented. The first is one-dimensional, with the benefit of abundant experimental data, thus allowing the illustration of the method using different data samplings. The second application is multidimensional, and concerns the enhancement of a MISL3 aerodynamic database using limited experimental data.

4.1 A One-Dimensional Example

For purposes of illustration, we now consider the experimental data from a series of wind tunnel tests carried out by Shorts Missile Systems Ltd. in the 1990s (Ref. [31]) for a free rolling missile body with a decoupled canard-controlled nose section (see **Figure 4**).

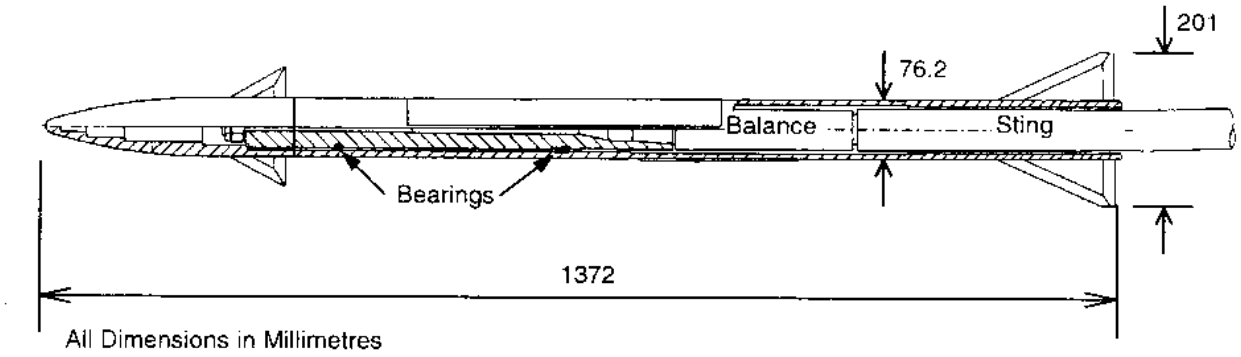

Figure 4: Details of Wind Tunnel Model (reproduced from Ref. [31], with permission).

Figure 5 shows the static rolling moment $C_{l,n}$ measured on the nose, as a function of the roll angle φ at a freestream Mach number $M = 3.5$, total incidence angle $\alpha_c = 8\,deg$, and canard fins canted at 8 and 12 degrees (leading edge up) for the port and starboard fins, respectively. The data shown in Figure 5 were taken at static conditions in order to compare them to CFD predictions.

This configuration happens to present an interesting case where, at certain roll angles, the upper canard experiences partial shielding from the windward flow, due to the expansion over the nose, an effect which was correctly predicted by the CFD calculations of Ref. [31]. In this paper, we repeat these CFD calculations with a finer roll angle increment, in order to better capture the nonlinearities in the rolling moment variation. For data fusion comparison purposes, a lower-fidelity method which does not incorporate all of the proper physics is also used. The results of both methods are described next.

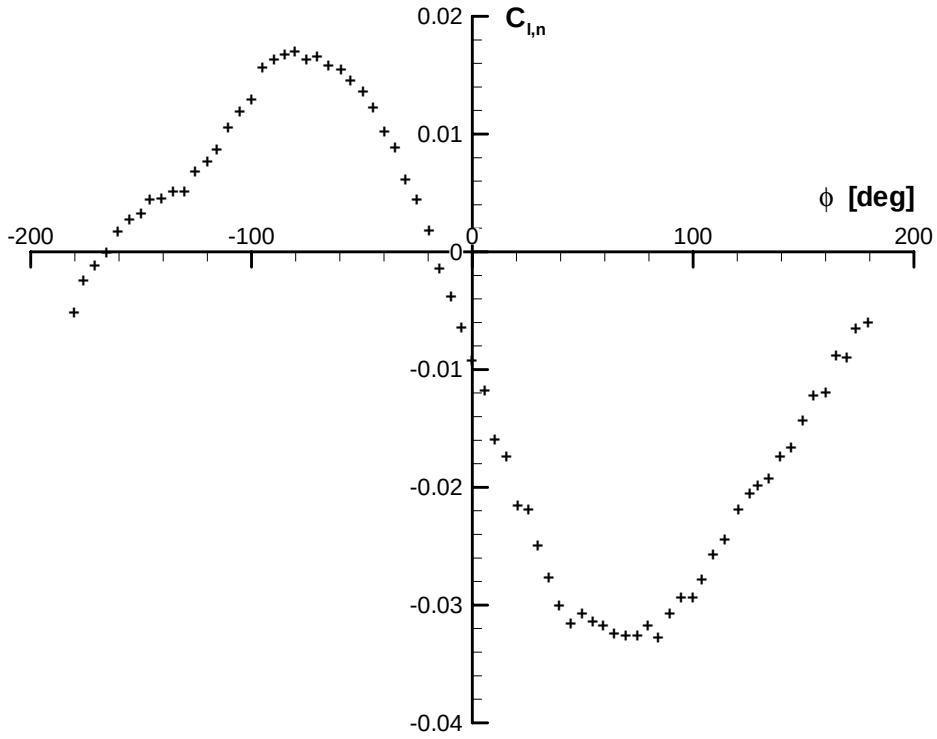

Figure 5: Variation of Nose Rolling Moment as a Function of Roll Angle, $M = 3.5$, $\alpha_c = 8\ deg$ (digitized from Ref. [31], Fig. 3, with permission).

We begin with the use of the low-fidelity computational data, not with the intent of showing what happens when one attempts to fuse experimental data with computations of inadequate fidelity, but simply as a more visually interesting case, and as a teaching tool for how the data fusion method works. **Figure 6** illustrates the results of the data fusion process when augmenting the computation with sparsely sampled subsets of the experimental data. Each plot in the figure portrays three entities: (1) the inputs, both computational (dashed line) and experimental (red dots), (2) the single fusion model output (solid line), and (3) the complete set of wind-tunnel measurements ("+" symbols). It must be stressed that only the input data are used in the data fusion calculation. The verification data are presented for comparison purposes only, i.e., as a reference against which to judge the quality of the prediction. The inadequacy of the computational model taken by itself is evident. The fusion of a single experimental data point ($\varphi = 0\ deg$) with the computational data stream is shown in the upper left graph of Figure 6. As expected, the fusion produces a slight shift in the prediction in order to accommodate the experimental support vector. Let us assume that a second experimental data point is acquired at $\varphi \approx -135\ deg$ (upper right graph). The fused prediction at negative roll angles is now tilted upward. The prediction maintains the overall character of the computation, but it has "learned" from the significant correction/improvement at $\varphi \approx -135\ deg$. Suppose a third experimental data point is added (lower left graph, $\varphi \approx -120\ deg$). The prediction locally adapts to reflect the new information, eliminating much of the undershoot in the $-135\ deg \le \varphi \le -100\ deg$ roll angle range. In the lower right graph of Figure 6, the addition of a fourth experimental data point at $\varphi \approx 80\ deg$ results in an upward tilt of the prediction to, once again, accommodate the new information.

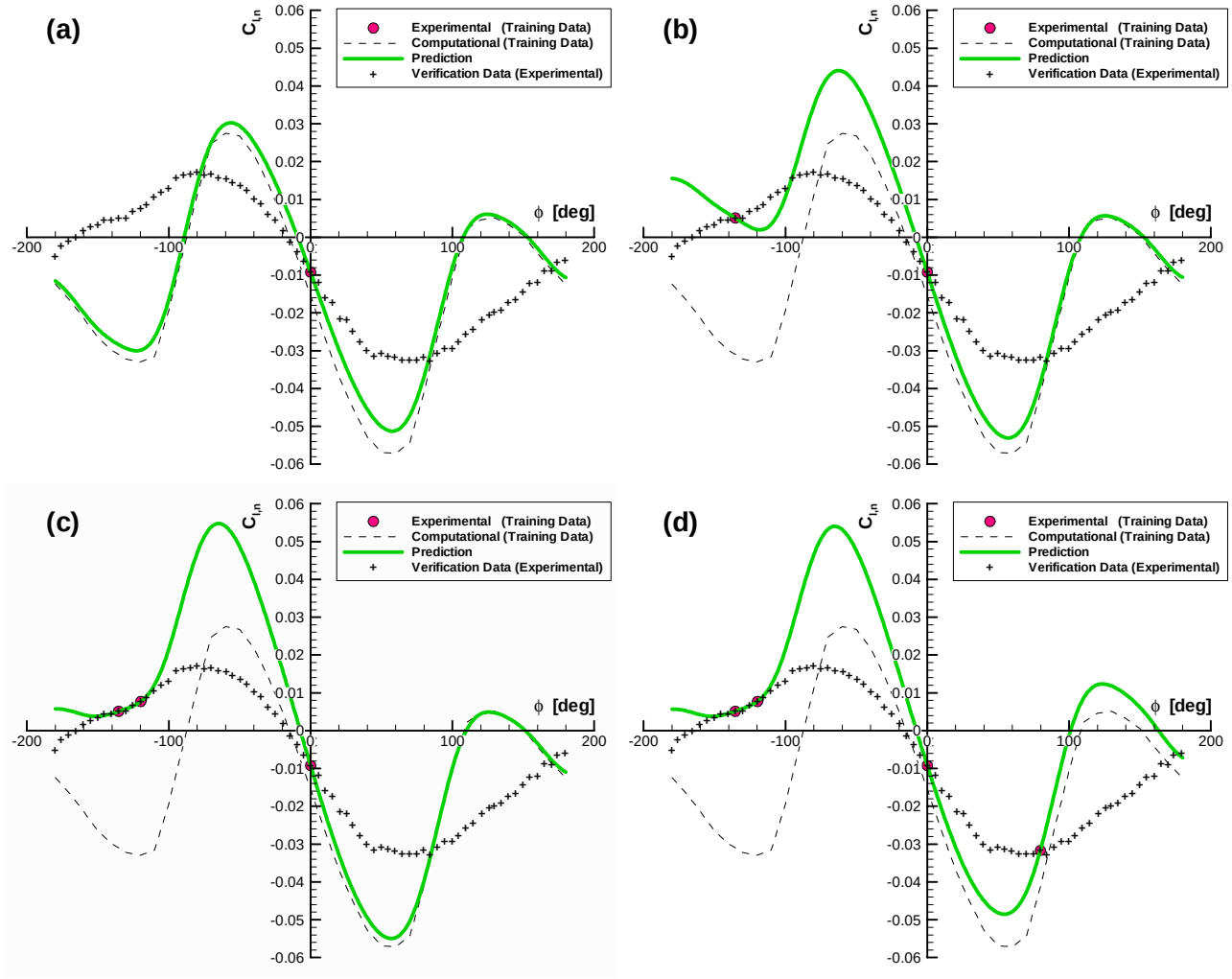

Figure 6: Data Fusion Predictions Using One (upper left), Two (upper right), Three (lower left), and Four (lower right) Experimental Data Points.

As anticipated from the geometric interpretation of Figure 3, the response surface respects and adjusts to the experimental support vectors, while maintaining the overall character of the computational tool, whether in interpolation or extrapolation mode. It is worth noting, however, that given a sufficient number of experimental data constraints, the influence of the computational data stream will eventually become insignificant. This has been demonstrated on this particular example: by using experimental data every 20 degrees (not shown), the RMS difference between the fusion result and the full data set was reduced from 1.7×10^{-2} (with four points) to 6.6×10^{-4} (using 19 points). Clearly, it is always possible to overcome the limitations of a poorly chosen computational model, given enough experimental data. The main interest, however, concerns the case where the experimental data are sparse, because this will inevitably be the case when the number of independent variables is large. What happens when one uses a computational model that contains the appropriate physics is shown next.

Instead of the lower fidelity computational model used in Figure 6, the data fusion experiments depicted above can be repeated using a computational methodology of the appropriate level (in this instance, an Euler CFD code, NEARZEUS, Ref. [32]). **Figure 7** depicts the data fusion prediction using the CFD data stream augmented with only two experimental points, $\varphi = 0\ deg$ and $\varphi = \pm 180\ deg$.

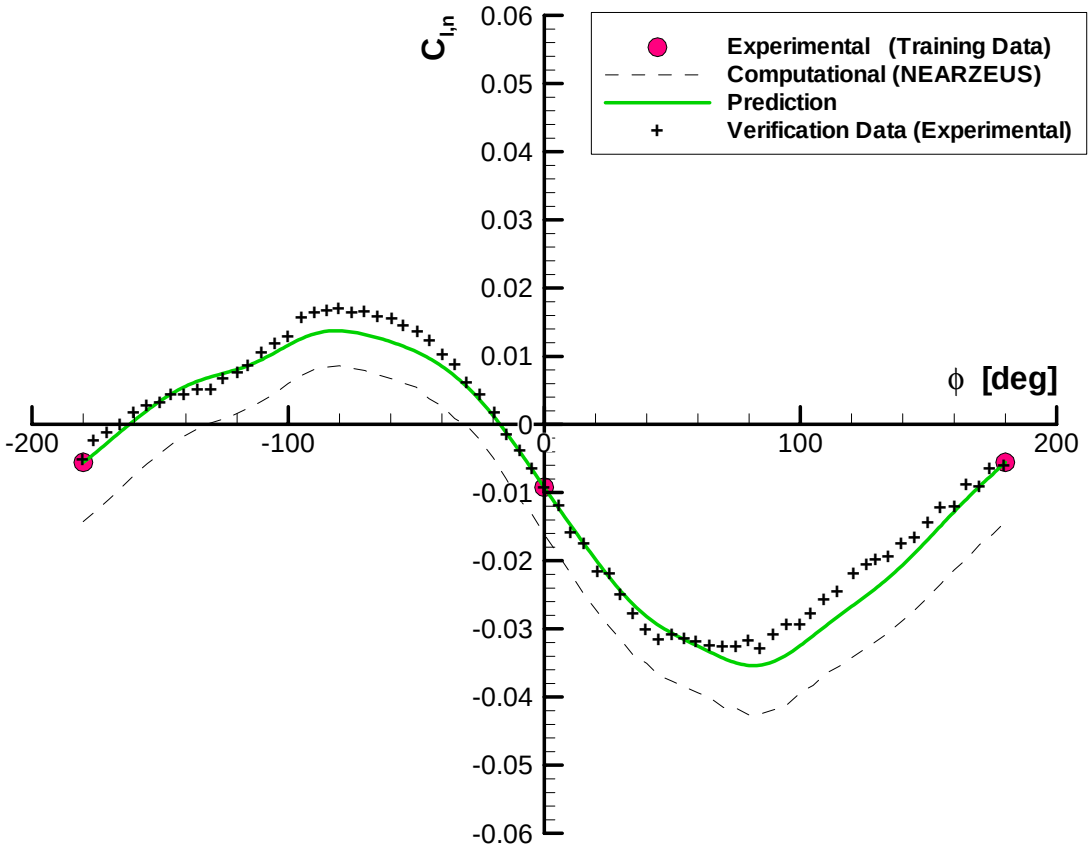

Figure 7: Data Fusion Prediction Based on Euler CFD Prediction and Two Experimental Data Points.

The comparison between the data fusion prediction and the complete experimental data set indicates a high degree of correlation. In particular, details of the aerodynamic nonlinearities predicted by NEARZEUS are visible in the response surface. With sparse experimental data, the response surface model has been tailored to *learn* from the computational data stream, while simultaneously adjusting to accommodate the experimental observations. Note that this particular implementation assumes the experimental data to be correct, which is the rationale for evaluating the global response surface in the experimental "plane" ($\varepsilon = 1$, see Section 3.2). Figure 7 corresponds to the nominal prediction F in Eq. (2). It is worth mentioning that the variance on the prediction (not shown) is also automatically computed by NEAR-RS, a topic that will be addressed in a separate paper in the future.

As a final note of caution, the importance of performing the data fusion operations with the correct analysis is stressed in **Figure 8**, which compares the data fusion predictions obtained by using the same two experimental data constraints as in Figure 7, namely $\varphi = 0\ deg$ and $\varphi = \pm 180\ deg$, but with different computational analyses.

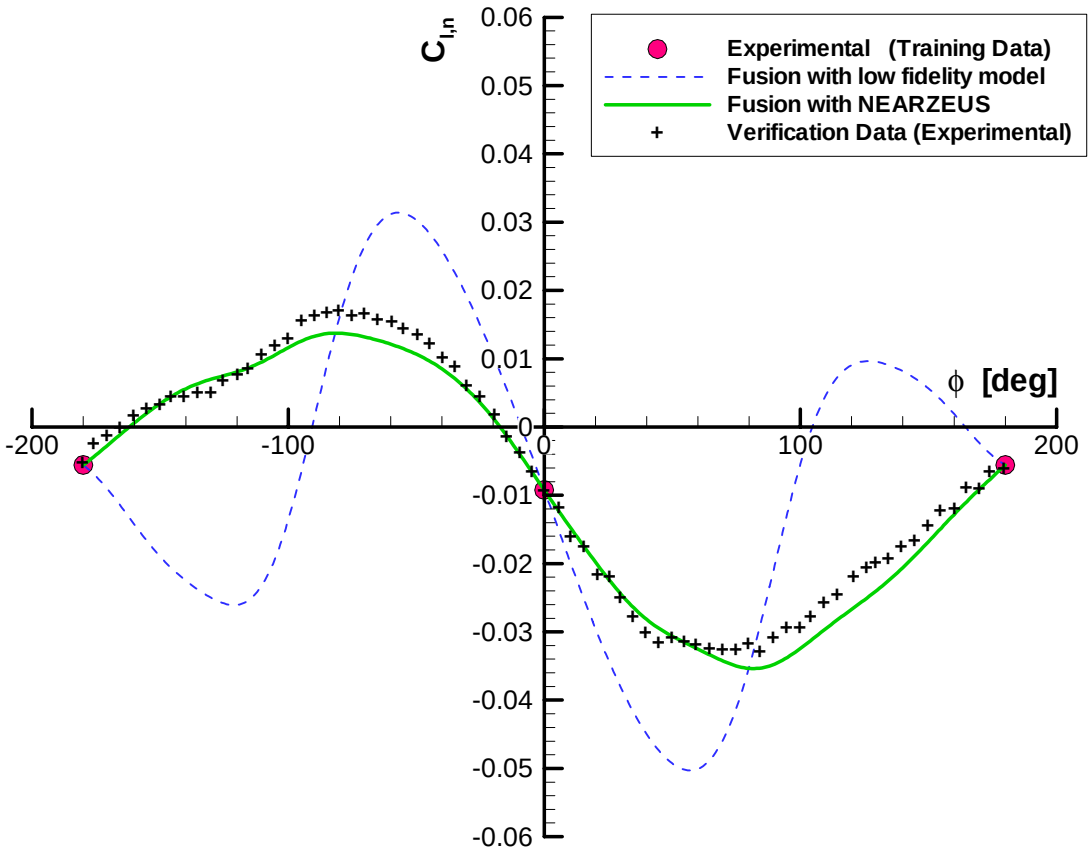

Figure 8: Comparison of Data Fusion Predictions Based on Two Computational Data Sets.

The results shown in Figure 8 emphasize the importance of using computational models which incorporate the correct physics. This is especially true when performing sparse data interpolation, since the reliance on the computation becomes greater. While the example of taking only two experimental data points may appear extreme and somewhat academic, it is in reality highly relevant to the case of multidimensional data. When data are characterized by a large number of independent variables, finite resources (time and budget) impose limitations on the number of conditions that can be acquired. Modern design-of-experiment techniques can be used to maximize the amount of information that can be harvested from a given number of tests. But when the number of dimensions is large, "filling-in" the space in all variables remains a physical impossibility. Out of necessity, the data sampling will be sparse in at least *some* directions or regions of the parameter space. Having illustrated the basic characteristics of the present data fusion method in one dimension, we now turn to another missile aerodynamics application, this time involving three independent variables and very limited quantities of experimental data.

4.2 Correction of MISL3 Database Using Experimental Data

The goal of this program was to assimilate limited wind-tunnel data, with the goal of increasing the accuracy of comprehensive flight simulations of a missile. The data shown here correspond to a generic body-tail configuration (not shown). This application merges two data sets: an experimental (wind tunnel) data set, and a computational data set. These data are used as the support vectors of a global response surface. The "computational" support vectors are supplied by the MISL3 code [33]. This MISL3 database consists of

forces and moments predictions for a wide range of angles of attack, roll angles, and Mach numbers in the subsonic, transonic, and supersonic range. The experimental support vectors were supplied by a wind tunnel test for a much smaller range of conditions consisting of three Mach numbers, four roll angles, and a subset of the angle-of-attack range.

To produce an "error database," to be used as an experimental correction to the MISL3 prediction, the difference between a fit to the MISL3 database and a fit to the experimental data was calculated. For each force or moment coefficient (generically denoted C) the fit was obtained by constructing a single low-order analytic (smoothly varying) four-dimensional response surface $RS_C(\alpha_c, \varphi, M, \varepsilon)$ based on both MISL3 and experimental training data sets. As described earlier, this is done by introducing the auxiliary variable ε as a fourth independent variable. This additional variable is used to separate the support vectors as distinct projections, or "planes," of the parameter space, as illustrated in **Figure 9**.

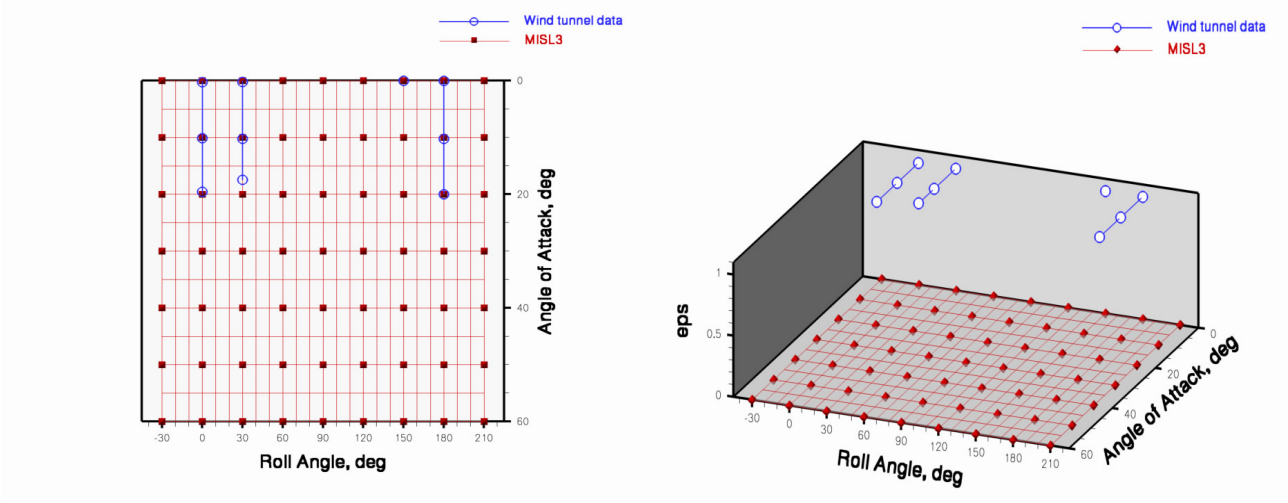

Figure 9: Schematic Illustrating Dimensionality Augmentation Prior to Data Fusion.

For ease of representation, the Mach number direction is omitted from Figure 9. The symbols indicate the locations in parameter space where wind tunnel and MISL3 data are available at a Mach number common to both data sets. Note that, for many of the Mach numbers, data are available in one of the two planes only. In addition, the angle of attack range of the experimental data was further limited at some Mach numbers. Therefore, the data do not lie on a regular matrix, a situation common to most high-dimensional data sets. Even in the rare cases (mainly low-dimensional) where a regular matrix of test points can be afforded, one is frequently confronted with the necessity of dealing with exceptions, i.e., "holes" in the data, or, as in the present case, unanticipated limitations in the range of some variables. These conditions are precisely what makes conventional structured data interpolators fail, yet are eminently suitable for constructive approximation via radial basis functions.

From a user perspective, this process is automatic and does not require the specification of any equations. Only support vectors from experimental and computational sources of data are needed. In this particular case, the computational source of data is the MISL3 database. **Figures 10** and **11** depict, respectively, the rolling moment and side force coefficient predictions. In order to compare in the same graph the data fusion predictions at $\varepsilon = 1$, the MISL3 database, and the experimental data, the Mach number of Figures 10 and 11 corresponds to a case where data common to both the experiment and MISL3 were available. Recall that the surface produced is not the result of two-dimensional interpolation, but a two-dimensional projection of a

four-dimensional response surface. In terms of data interpolation/extrapolation, Figures 10 and 11 make it clear that the shape of the prediction surface with respect to angle of attack and roll angle is "inherited" primarily from the MISL3 database, which was the desired intent.

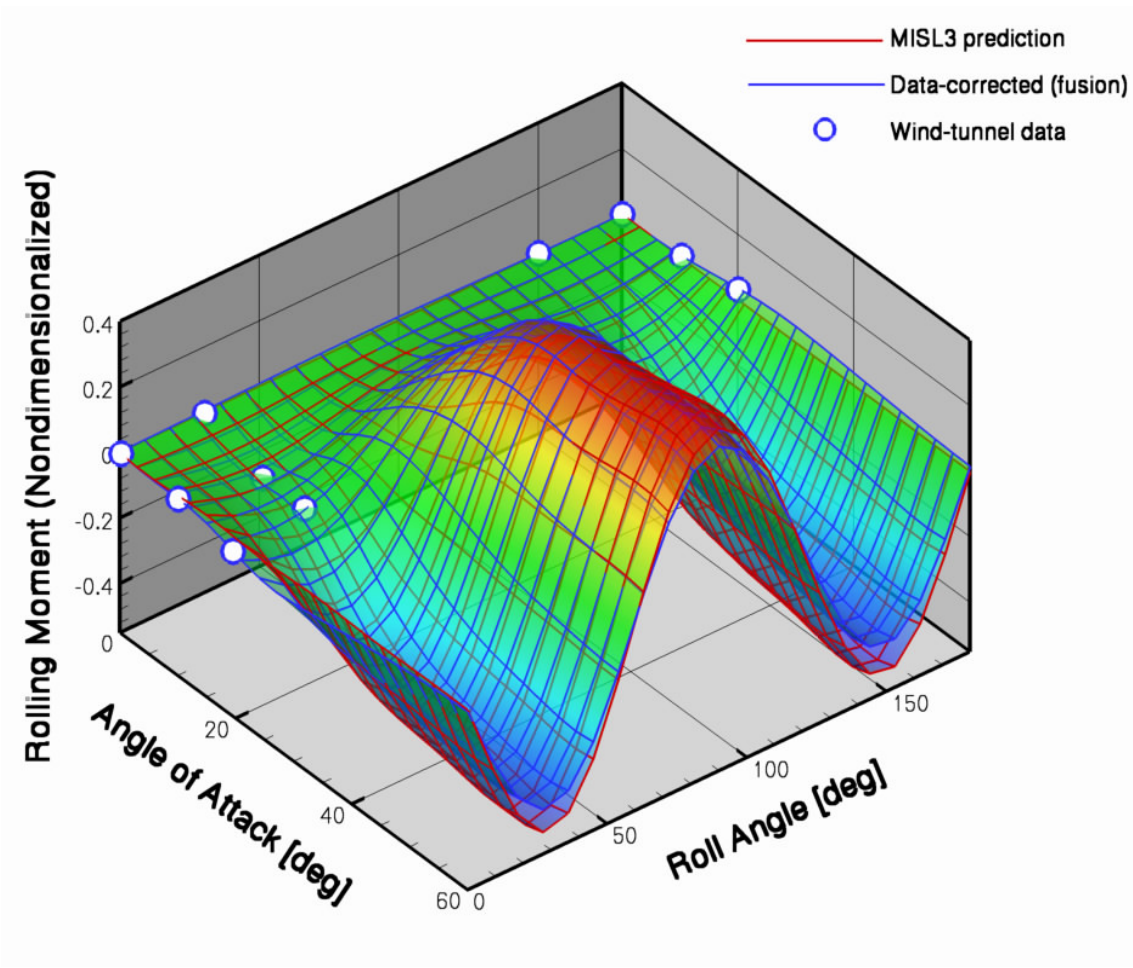

Figure 10: MISL3 and Experimental Rolling Moment Predictions.

Note, in addition, that this process can be used to create an "error database," which is of interest to assess the effects of aerodynamic errors in flight simulations. By arbitrarily taking the MISL3 prediction as the reference base, the error database, defined as $\delta C(\alpha_c,\varphi,M) = RS_C(\alpha_c,\varphi,M,\varepsilon=1) - RS_C(\alpha_c,\varphi,M,\varepsilon=0)$, can be used to "correct" the MISL3 database ($C^{corrected} = C^{MISL3} + \delta C$) so as to take into account the experimental measurements. With the exception of minor differences pertaining to the sampling of the MISL3 database for selecting the support vectors, $C^{corrected}$ is equivalent to $RS_C(\alpha_c,\varphi,M,\varepsilon=1)$. In other words, the corrected database is the result of the data fusion response surface, evaluated in the experimental plane. Similarly to the example shown earlier (Section 4.1), the data fusion prediction can be interpreted as a calibration of the MISL3 output, based on experimental data. Conversely, this method is a way of constructing smart interpolation and extrapolation schemes in cases where only limited quantities of experimental data are available.

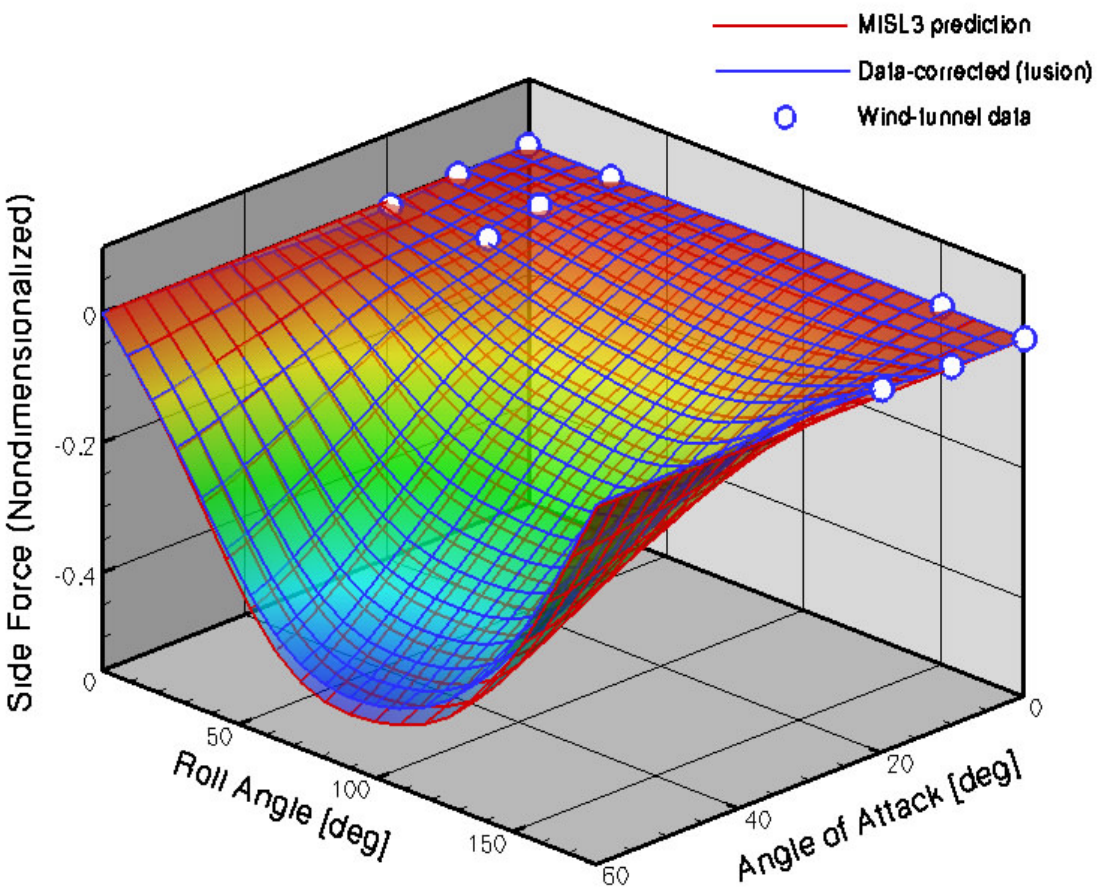

Figure 11: MISL3 and Experimental Side Force (rolled coordinate system) Predictions.

NEAR has also used this method to improve pressure drag predictions for a different application, based on the fusion of high-resolution CFD calculations with limited pressure tap measurements in a wind-tunnel test. The technique outlined above is quite general and can be used, either to "fill-in" where limited quantities of experimental data are available, or to "fine tune" the results of computational analyses using the limited data available.

5.0 CONCLUSION

An innovative method for fusing experimental and computational data was presented. We have shown how, using this method, limited wind tunnel data for a missile can be used to increase the accuracy of databases in comprehensive flight simulation programs. This method allows data structure flexibility and the use of heterogeneous data sets, provides a fully analytic, mathematical description that can be easily manipulated and shared between applications, and provides a rational basis for propagating uncertainty estimates.

6.0 ACKNOWLEDGEMENTS

The authors gratefully acknowledge the Office of Naval Research and Program Officer Brian Almquist, Code 321OE, for its financial support, Thales Air Defense Ltd. for the use of their data, and Stanley Perkins of NEAR for providing the NEARZEUS calculations.

7.0 REFERENCES

[1] Burman, J., Papila, N., Shyy, W., and Gebart, B. R., "Assessment of Response Surface-Based Optimization Techniques for Unsteady Flow Around Bluff Bodies," AIAA 2002-5596.

[2] DeLoach, R. and Erickson, G. E., "Low-Order Response Surface Modeling of Wind Tunnel Data Over Truncated Inference Subspaces," AIAA 2003-0456.

[3] Forrester, A. I. J., Bressloff, N. W., and Keane, A. J., "Response Surface Model Evolution," AIAA 2003-4089.

[4] Hirokawa, N., Fujita, K., and Iwase, T., "Voronoi Diagram Based Blending of Quadratic Response Surfaces for Cumulative Global Optimization," AIAA 2002-5460.

[5] Knill, D. L., Giunta, A., Baker, C. A., Grossman, B., Mason, W. H., Haftka, R. T., and Watson, L. T., "Response Surface Models Combining Linear and Euler Aerodynamics for Supersonic Transport Design," *Journal of Aircraft*, Vol. 36, No. 1, Jan.-Feb. 1999, pp. 75-86.

[6] Krishnamurphy, T., "Response Surface Approximation With Augmented and Compactly Supported Radial Basis Functions," AIAA 2003-1748.

[7] Papila, N., Shyy, W., Griffin, L., and Dorney, D. J., "Shape Optimization of Supersonic Turbines Using Response Surface and Neural Network Methods," AIAA 2001-1065.

[8] Walker, E. L., "Statistical Calibration and Validation of a Homogeneous Ventilated Wall-Interference Correction Method for the National Transonic Facility," NASA TP-2005-213947.

[9] Rais-Rohani, M. and Singh, M. N., "Efficient Response Surface Approach for Reliability Estimation of Composite Structures," AIAA 2002-5604.

[10] Vittal, S. and Hajela, P., "Confidence Intervals for Reliability Estimated Using Response Surface Methods," AIAA 2002-5475.

[11] Jones, D. R., Schonlau, M., and Welch, W. J., "Efficient Global Optimization of Expensive Black-Box Functions," *Journal of Global Optimization*, Vol. 13, 1998, pp. 455-492.

[12] Queipo, N. V., Haftka, R. T., Shyy, W., Goel, T., Vaidyanathan, R., and Tucker, P. K., "Surrogate-Based Analysis and Optimization," *Progress in Aerospace* Sciences, Vol. 41, 2005, pp. 1-28.

[13] Zeldin, B. A. and Meade, A. J., "Integrating Experimental Data and Mathematical Models in Simulation of Physical Systems," *AIAA Journal*, Vol. 35, No. 11, 1998, pp. 1787-1790.

[14] Poggio, T. and Girosi, F., "Network for Approximation and Learning," *Proc. IEEE*, Vol. 78, No. 9, 1990, pp. 1481-1497.

[15] Raeth, P. G., Gustafson, S. C., Little, G. R., and Puterbaugh, T. S., "Stretch and Hammer Neural Networks For N-Dimensional Data Generalization," Air Force Wright Laboratory Report WL-TR-91-1146, Jan. 1992.

[16] Vapnik, V. N., *The Nature of Statistical Learning Theory*. Springer, 1995.

[17] Gunn, S. R., "Structural Modelling with Sparse Kernels," Proc. 13th IFAC Symposium on System Identification, Rotterdam, Netherlands, 2003.

[18] Rodman, L. C., Reisenthel, P. H., and Childs, R. E., "An Automated Documentation and Reporting System For CFD," AIAA 2002-0986.

[19] Matheron, G., "Principles of Geostatistics," *Economic Geology*, Vol. 58, 1963, pp. 1246-1266.

[20] Cressie, N., "The Origins of Kriging," *Mathematical Geology*, Vol. 22, No. 3, 1990, pp. 239-252.

[21] Sacks, J., Welch, W. J., Mitchell, T. J., and Wynn, H. P., « Design and Analysis of Computer Experiments, » *Statistical Science*, Vol. 4, 1989, pp. 409-435.

[22] Friedman, J. H., "Multivariate Adaptive Regression Splines," *The Annals of Statistics*, Vol. 19, No. 1, 1991, pp. 1-67.

[23] Orr, M., "Optimising the Widths of Radial Basis Functions," *Proceedings of the Fifth Brazilian Symposium on Neural Networks*, Belo Horizonte, Brazil, 1998.

[24] Pottmann, M. and Seborg, D. E., "Identification of Non-Linear Processes Using Reciprocal Multiquadric Functions," *Journal of Process and Control*, Vol. 2, No. 4, 1992, pp. 189-203.

[25] Hardy, R. L., "Multiquadric Equations of Topography and Other Irregular Surfaces.," *J. Geophys. Res.*, Vol. 76, No. 8, Mar. 1971, pp. 1905-1915.

[26] Franke, R., "Scattered Data Interpolation: Test of Some Methods," *Math. of Comp.*, Vol. 38, No. 157, Jan. 1982, pp. 181-200.

[27] Orr, M. J. L., "Regularisation in the Selection of Radial Basis Function Centres," *Neural Computation*, Vol. 7, No. 3, 1995, pp. 606-623.

[28] Strang, G., *Linear Algebra and Its Applications*, Academic Press, New York, 1980, p. 142.

[29] Norton, J. P., *An Introduction to Identification*, Academic Press, New York, 1986, pp. 87-119.

[30] Li, H., Manjunath, B. S., and Mitra, S. K., "Multisensor Image Fusion Using the Wavelet Transform," *Computer Vision, Graphics, and Image Processing: Graphical Models and Image Processing*, Vol. 57, 1993, pp. 235-245, 1993.

[31] McIlwain, S. T., Mallon, P. C. G., Fleming, R. J., McConnell, G., "An Experimental and Numerical Study of a Supersonic Spinning Missile," *Proceedings of the RTO AVT Symposium on Missile Aerodynamics*, Sorrento, Italy, May 11-14, 1998, RTO MP-5, p. 35-1.

[32] Perkins, S. C., Jr., Wardlaw, A. B., Jr., Priolo, F., and Baltakis, F., "NEARZEUS User Manual, Vol. I, Operational Instructions; Vol. II, Sample Cases; Vol. III, Boundary Layer Code, ZEUSBL," NEAR TR-459, Nielsen Engineering & Research, Oct. 1997.

[33] Lesieutre, D. J., Love, J. F., and Dillenius, M. F. E., "MISL3 Aerodynamic Analysis For Finned Vehicles With Axisymmetric Bodies," NEAR TR 561, Nielsen Engineering & Research, Mar. 2004.