Multi-fidelity Aerodynamic Dataset Generation of a Fighter Aircraft

Huseyin Burak KURT[1] and Murat MILLIDERE[2]
Turkish Aerospace, Ankara, Turkey
Middle East Technical University, Ankara, Turkey

Fazıl Selcuk GOMEC [3]
Turkish Aerospace, Ankara, Turkey

Omur UGUR [4]
Middle East Technical University, Ankara, Turkey

High-fidelity aerodynamic dataset generation is one of the most significant components of the aircraft flight simulation, while it is a time consuming and costly process. Data fusion techniques suggest that, instead of using high fidelity data for entire aerodynamic dataset, a incorporating combination of high-fidelity and low-fidelity data is a more cost-effective one. The objective of data fusion is to obtain high-fidelity dataset accuracy by combining less amount of high-fidelity dataset and more amount of low-fidelity dataset. In this paper, two different data fusion approaches, namely modified Variable-Complexity Modelling and co-Kriging, are applied to F-16 fighter aircraft. Wind tunnel test data is utilized as the high-fidelity dataset while data obtained by a semi-empirical approach (Digital Datcom) is used as the low-fidelity dataset.

Nomenclature

h_f	=	High-Fidelity
l_f	=	Low-Fidelity
C_L	=	Lift Force Coefficient
C_D	=	Drag Force Coefficient
C_m	=	Pitching Moment Coefficient
α	=	Angle of Attack
β	=	Angle of Sideslip
δ_h	=	Horizontal Tail Deflection
δ_{LEF}	=	Leading Edge Flap Deflection
GA	=	Genetic Algorithm
X	=	Set of Sample Data
Y	=	Set of Observed Response Data
n	=	Number of samples in the Dataset
k	=	Dimension Number of One Sample Data

[1] Flight Dynamics, Simulation & Control Engineer.
[2] Ph.D. Candidate, Department of Engineering Sciences.
[3] Aeroacoustic Chief Engineer.
[4] Prof. Dr., Institute of Applied Mathematics.

μ = Mean Value of Dataset
σ = Variance of Dataset
cor = Correlation of Dataset
Ψ = Correlation Matrix of nxn Dataset
MLE = Maximum Likelihood Estimation
CFD = Computational Fluid Dynamics
VCM = Variable Complexity Modelling
RMSE = Root Mean Square Error

I. Introduction

Aircraft flight simulation is a key component of modern aircraft design, certification process and aircraft pilots training. How much the aircraft in flight simulation reveals the actual aircraft flight characteristic depends on various factors, one of which is the aerodynamic dataset accuracy. The aerodynamic model is the core of a flight simulator. An error in the aerodynamic model can lead to a simulation which might fail in the qualification process [1]. Therefore, flight simulator needs the high-fidelity datasets to provide the real aircraft behavior to the pilots. There are several sources to obtain the aerodynamic dataset. These are semi-empirical datasheet methods, linear flow solvers, nonlinear flow solvers, small scale wind tunnel tests, full scale wind tunnel tests and flight tests in the increasing order of fidelity. As fidelity increases, processing time and cost increase.

Data fusion is one of the functional approaches to obtain aerodynamic dataset. Since using only high-fidelity wind tunnel test or high-fidelity computational analysis for the aerodynamic dataset generation is time consuming and expensive. The main motivation of data fusion is to achieve high-fidelity data in a less time-consuming fashion. In this study we implemented two different data fusion techniques being "Variable-Complexity Modelling" (VCM) [2].and co-Kriging method. In the modified VCM approach, there is an increment function which computes that how much value should be added to low-fidelity data, in order to obtain high-fidelity data when desired. In this approach, Kriging method is used for interpolation. Although Kriging is a global interpolation technique, it does not assume a global form for the function. Kriging try to interpolate function which goes exactly given data points. Kriging model could interpolate merged dataset around high-fidelity data; hence, Kriging such an interpolation property is best suited for data fusion. In addition, introducing new data points to Kriging will increase the accuracy of predicted model [2].

The co-Kriging approach developed for multi-fidelity analysis is another data fusion algorithm. In order to use low-fidelity data, some form of correction process is formulated [3] [4]. This correction process models the differences between high-fidelity and low-fidelity data. co-Kriging is a form of Kriging which correlates the multiple-fidelity dataset [3] [4].

In this paper, the wind tunnel test data is considered as high-fidelity dataset and the data from semi-empirical approach (Datcom) is considered as low-fidelity dataset. Aerodynamic dataset is generated using the two data fusion approaches described above. First modified VCM approach is applied to high-fidelity and low-fidelity datasets, then the co-Kriging approach is also applied to both datasets.

The rest of the paper is organized as follows: Section II presents data evaluation. Section III and IV discuss the modified variable complexity modelling approach and co-Kriging approach respectively. The results of both approaches are presented and compared in Section V and, conclusions are given in Section VI. Finally, the appendix shows how to obtain the partitioned inverse of matrix.

II. Data Evaluation

For aerodynamic database generation, various computational and experimental methods can be utilized with respect to the cost and accuracy requirements of the design phase. A summary of these methods with sample tools and classifications are provided in [5] [6]. As expected, the higher the cost, the more accurate the aerodynamic database. The resource and time cost of each method are essentially tens to hundreds times higher than those of the less accurate method [7], as seen in Figure 1.

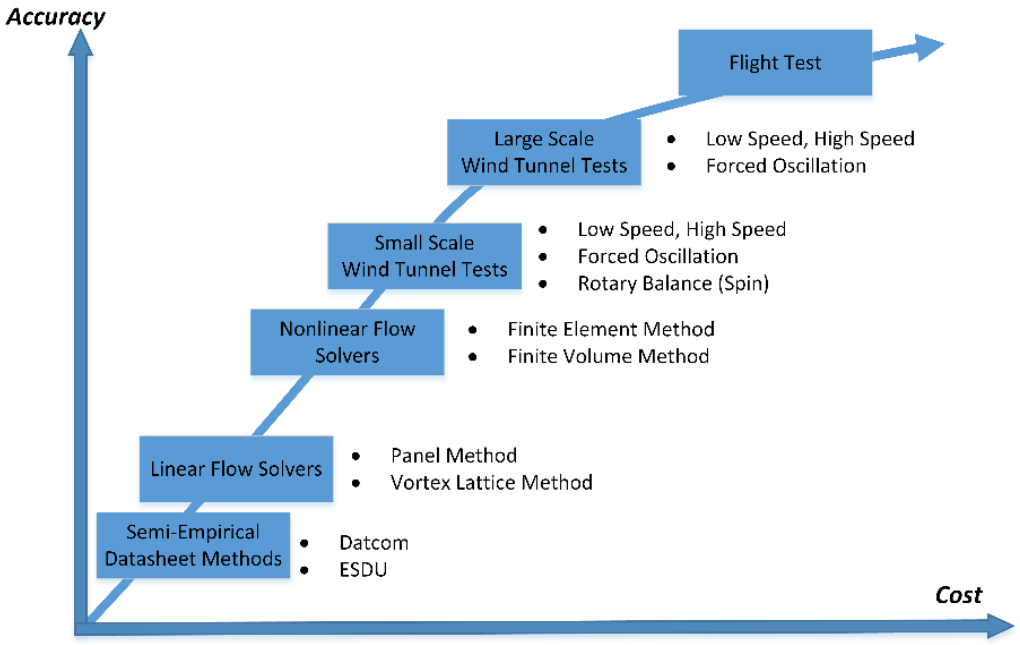

Figure 1. Aerodynamic database generation methods with examples

Semi-empirical methods are based on the experimental results that are generalized and formulized with these methods. Datcom and ESDU are typical examples of this approach. The computational effort is significantly small as compared to other approaches. In fact, a full flight dataset can be obtained in a couple of minutes. They do not need any computational grid generation.

Linear flow solvers are based on potential flow theory. Panel and vortex lattice methods (VLM) are typical examples of this approach. They may have some corrections for the compressibility, boundary layer, dynamic motions or nonlinear flow characteristics. They require computational grids, however, the size of grids are much smaller than nonlinear flow solvers. VSaero, Panair, Tranair and AVL are typical examples of Panel and vortex lattice methods. The linear flow regimes can be computed more accurate than semi empirical methods while the accuracy significantly decreases as the flow non-linearity increases [5]. Semi-empirical methods and linear flow solvers are generally used in conceptual design phase.

Nonlinear flow solvers attempt to solve all conservation equations including continuity, momentum and energy equations. They can cover all turbulence models, compressibility and boundary layer related problems. Finite volume and finite element methods are considered in the nonlinear flow solver methods. Ansys-Fluent, Star-CCM+, OpenFoam, CFL3D and SU2 are widely known nonlinear flow solvers. Such methods have been studied accurately within at low angle of attack sweeps (linear region) in subsonic-supersonic regions but the results in transonic region including shock-induced separations and/or at high angle of attack sweeps (nonlinear region) are different from the actual aerodynamic data due to highly unsteady flow characteristic. It is difficult to analyze dynamic effects including forced oscillations and spin conditions using such methods. They require high computational power than linear flow solvers due to more effort in pre-processing and computations [5].

As for wind tunnel tests, based on the principle of scale similarity, aerodynamic database are generated through experiments. A wind tunnel test model should be designed and manufactured in order to conduct these experiments which result in a much more costly process when compared with the computational approaches. Nevertheless, highly unsteady flight conditions in nonlinear flow solvers may only be solved with wind tunnel tests. In small scale wind tunnels, some scaling issues such as Reynolds number, aeroelastic characteristics, and so on exist so that full scale wind tunnel is preferred [8]. Supersonic aerodynamics, forced oscillation and spin characteristics can be obtained in small scale wind tunnels while take off-landing characteristics are accurately observed in large scale wind tunnels. Forced oscillation and high-speed tests can also be conducted in large scale wind tunnels. Surely, there exists some

deviation from actual aerodynamic data because of Reynolds correction, the experimental system errors, the influence of wind tunnel airflow quality, the interference of tunnel walls/supports and so on [7].

To eliminate the errors in wind tunnel tests, flight tests are the final alternatives. Aerodynamic database can be generated from a flight test data using a comprehensive system identification methods. Aerodynamic data should be validated with flight tests for further aircraft development.

In this study, aerodynamic database of F-16 aircraft is constructed by Digital Datcom as a function of the angle of attack and the horizontal tail deflection. Moreover, wind tunnel tests are adapted to improve the database. The former represents the low-fidelity dataset while the latter does high-fidelity dataset. The wind tunnel results are obtained from the open source F-16 aircraft data [9]. Digital Datcom includes empirical methods for aerodynamic data predictions. It enables fast and reasonably accurate computations in conceptual design phases.

The main geometry of F16 is obtained from OpenVSP-Hangar as shown in Figure 2 [10]. The fuselage geometry is prepared from this geometry while the wing, horizontal tail and vertical tail geometries are modified with respect to a supersonic wind tunnel test campaign as in [11]. The model is in 1/15 scale.

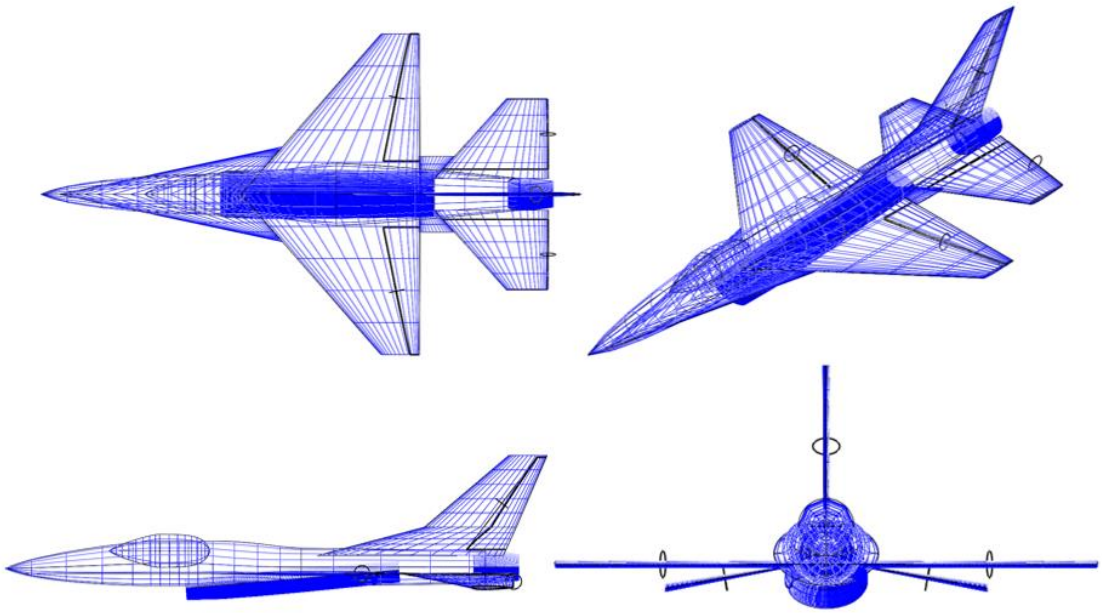

Figure 2. Top, isometric, bottom and left-side view of F16-OpenVSP geometry

The aerodynamic database is created for clean aircraft and horizontal tail deflections. The angle of attack ranges between 0° and 30° while the Mach number is set to 0.2 at sea level flight condition. For the non-dimensional coefficients, the mean aerodynamic chord is selected as 3.45 m while the span is set to 9.15 m. Additionally, the wing area is set to 27.86 m^2.

In Figures 3 - 8, lift, drag and pitch moments coefficients of F-16 aircraft geometry are compared. The coefficients are obtained from Datcom and wind tunnel test reference, which is an open source data of F-16 aircraft. Since the main idea is to improve the data quality of a low-fidelity model, these figures are helpful to observe differences between low- and high-fidelity models.

Particularly, in Figures 3 - 5, angle of attack sweeps of drag, lift and pitch moment coefficients are provided for different horizontal tail deflections. As seen from these figures, lift and drag coefficients of Datcom are much more reliable than the pitch moment coefficient. Since the section-wise pressure distribution data are not computed in Datcom, pitch moment distribution on the wing and horizontal tails may deviate from the experimental data. Additionally, the horizontal tail of the experimental data is smaller than the Datcom model since the TP1538 data is

based upon the YF-16 geometry. Nevertheless, Datcom model can be improved with respect to the reference wind tunnel data.

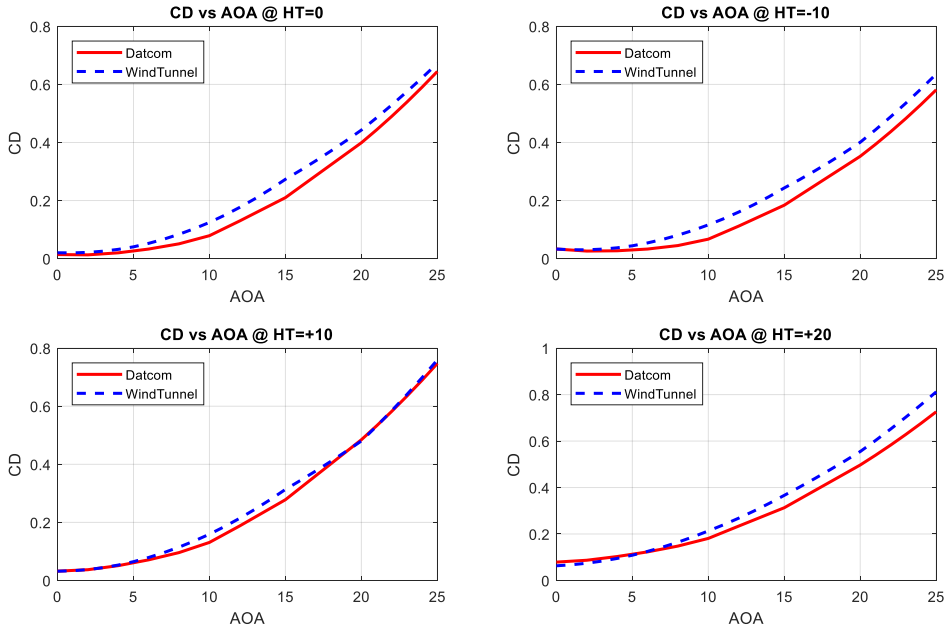

Figure 3. Comparison of drag coefficient of TP1538 and Datcom for various horizontal tail deflections

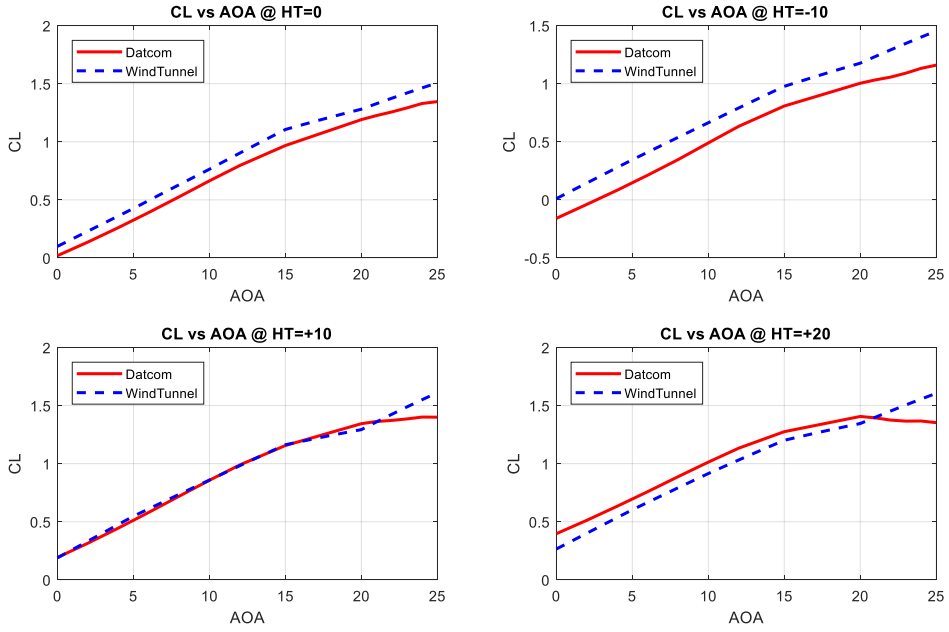

Figure 4. Comparison of lift coefficient of TP1538 and Datcom for various horizontal tail deflections

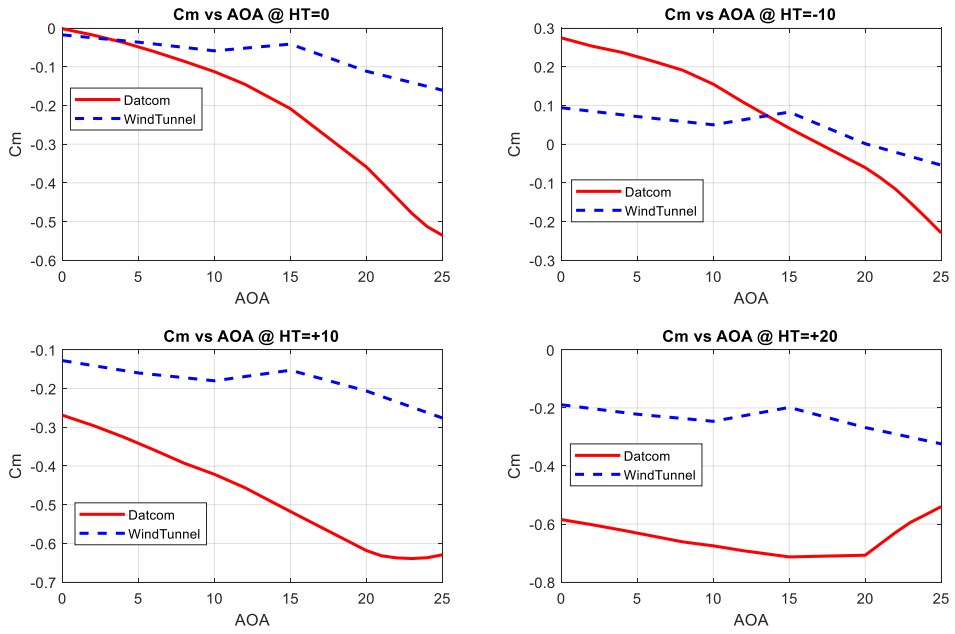

Figure 5. Comparison of pitch moment (at 25% of MAC) coefficient of TP1538 and Datcom for various horizontal tail deflections

On the other hand, Figures 6- 8, horizontal tail deflection sweeps of drag, lift and pitch moment coefficients are provided for different angle of attack conditions. The deflection effects in drag coefficient are closer at 0° of angle of attack while the remaining results are all deviate from the wind tunnel data. Nevertheless, the deviations are observed as shifts in drag coefficient and as rotations in pitch moment and lift force coefficients.

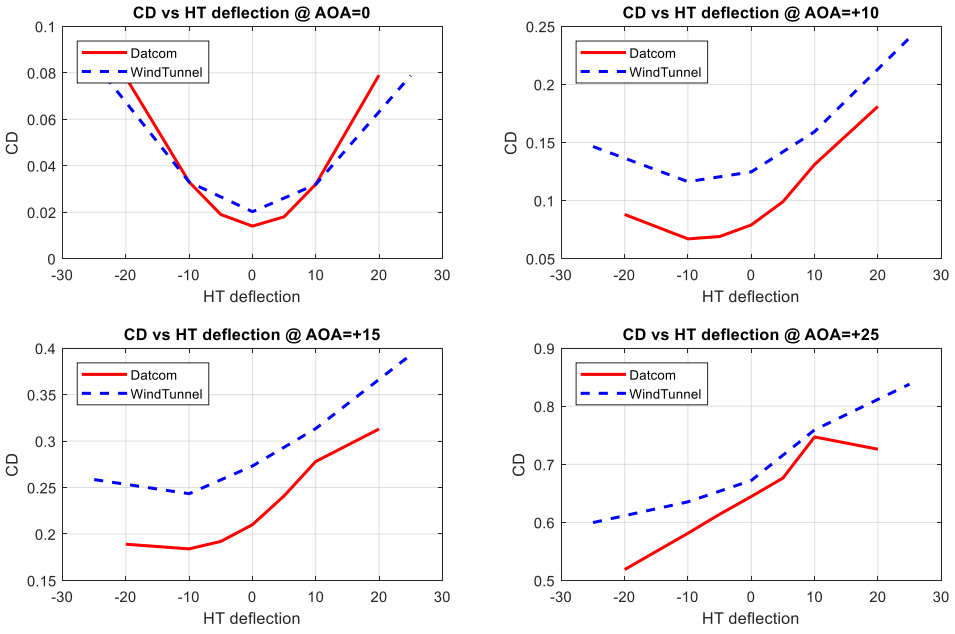

Figure 6. Comparison of drag coefficient of TP1538 and Datcom for various horizontal tail deflections

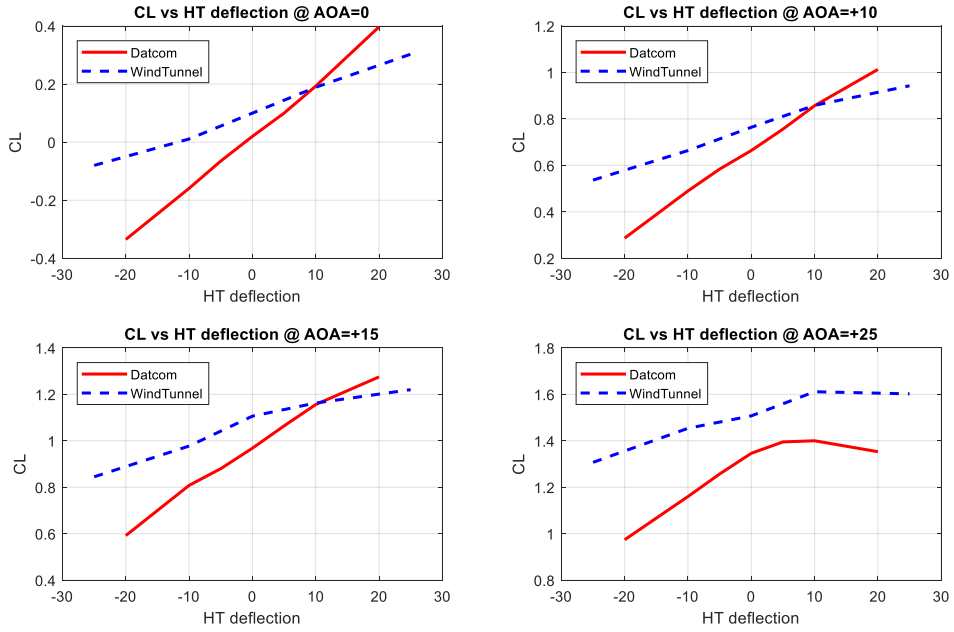

Figure 7. Comparison of lift coefficient of TP1538 and Datcom for various horizontal tail deflections

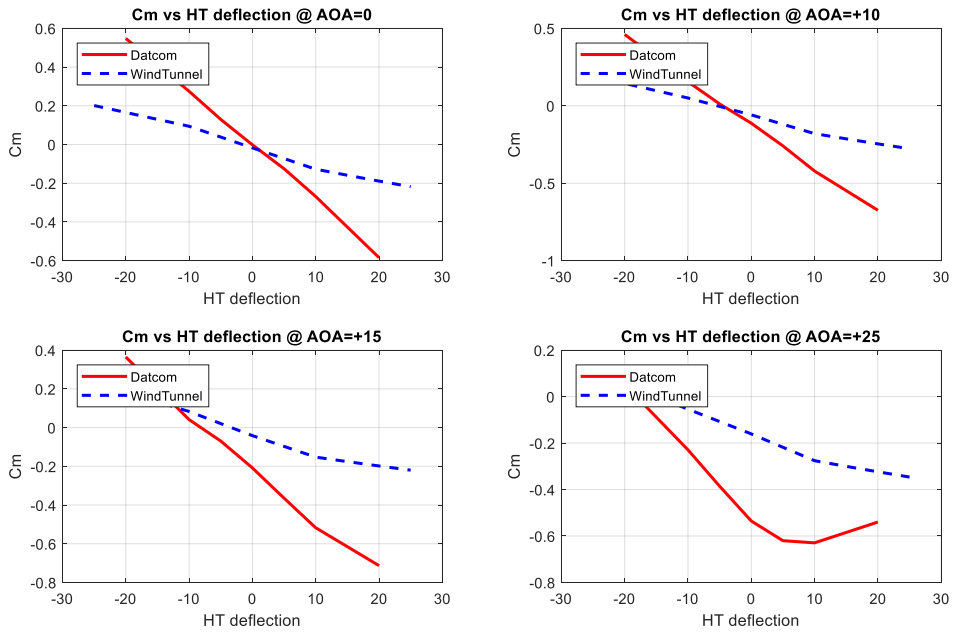

Figure 8. Comparison of pitch moment (at 25% of MAC) coefficient of TP1538 and Datcom for various angle of attack values

III. Modified Variable Complexity Modelling Approach

A. Building Kriging Model

In this section, Kriging interpolation is constructed in order to use the modified Variable-Complexity Modelling (VCM) model and the formulation follows [4].

Let *X* be the set of sample data and *y* be the observed responses defined as

$$X = \left\{X^{(1)}, X^{(2)}, X^{(3)}, \dots, X^{(n)}\right\} \tag{1}$$

$$y = \left\{y^{(1)}, y^{(2)}, y^{(3)}, \dots, y^{(n)}\right\} \tag{2}$$

where *n* is the number of sample data.

Observed responses are thought as if they are from a stochastic process. So, observed responses are denoted by set of random vectors as follows:

$$Y = \begin{pmatrix} Y(X^{(1)}) \\ \vdots \\ Y(X^{(n)}) \end{pmatrix} \tag{3}$$

These random variables are correlated with each other using basis function expression:

$$cor\left[Y\left(X^{(i)}\right), Y\left(X^{(l)}\right)\right] = \exp\left(-\sum_{j=1}^{k} \theta_j \left|X_j^{(i)} - X_j^{(l)}\right|^{p_j}\right) \tag{4}$$

Basis function structure mentioned in Eq. (4), it is special form of Gaussian radial basis function used in Kriging models. Where Gaussian radial basis function has $1/\sigma^2$, Kriging basis function has a vector $\theta = \{\theta_1, \theta_2, \theta_3, \dots, \theta_k\}^T$ which allows the width of each basis function to vary from variable to variable. Also, Gaussian basis function p value is fixed at 2. However, Kriging basis function p values also vary $p \in [1,2]$. Kriging exponent can be written as vector $p = \{p_1, p_2, p_3, \dots, p_k\}^T$, where k is dimension number of each sample in dataset. In other words, k expresses the number of dependencies of each aerodynamic coefficient. In this study, Kriging basis function p values are taken as 2 for each dimension in order to reduce the computational complexity. Here, i and l are the numbers of samples in the dataset, and such that $1 \leq i, l \leq n$.

Having defined the Kriging basis function, the n x n correlation matrix of all the observed data becomes

$$\Psi = \begin{bmatrix} cor\left[Y\left(X^{(1)}\right), Y\left(X^{(1)}\right)\right] & cor\left[Y\left(X^{(1)}\right), Y\left(X^{(2)}\right)\right] & \cdots & cor\left[Y\left(X^{(1)}\right), Y\left(X^{(n)}\right)\right] \\ cor\left[Y\left(X^{(2)}\right), Y\left(X^{(l)}\right)\right] & cor\left[Y\left(X^{(2)}\right), Y\left(X^{(2)}\right)\right] & \cdots & cor\left[Y\left(X^{(2)}\right), Y\left(X^{(n)}\right)\right] \\ \vdots & \vdots & \ddots & \vdots \\ cor\left[Y\left(X^{(n)}\right), Y\left(X^{(l)}\right)\right] & cor\left[Y\left(X^{(n)}\right), Y\left(X^{(2)}\right)\right] & \cdots & cor\left[Y\left(X^{(n)}\right), Y\left(X^{(n)}\right)\right] \end{bmatrix} \tag{5}$$

To represent the correlation matrix of the set of random variables (Y). Each sample point is correlated with each other in the given dataset with this correlation matrix Ψ. Each sample is correlated with other sample point using absolute distance between them using $\left|X_j^{(i)} - X_j^{(l)}\right|$ formulation and parameters θ_j and p_j. Again, in this formulation i and l express the number of samples in dataset and j represents each dimension in one sample point. Note that, we have $1 \leq i, l \leq n$, $1 \leq j \leq k$ and n is the number of samples and k is the number of dimensions.

In Kriging basis function, θ and p values are to be estimated. Thus, the genetic algorithm (GA) optimization method is used, in which, θ and p are the design parameters. Basically, GA minimizes a cost function using these design parameters in order to find the optimum values for the design parameters. For this purpose, maximum likelihood formulation is chosen for cost function.

In maximum likelihood estimation (MLE), it is assumed that the likelihood function $L\left(Y^{(1)}, Y^{(2)}, \dots, Y^{(n)} \mid \mu, \sigma\right)$, where μ and σ are distribution parameters, is of the form

$$L\left(Y^{(1)}, Y^{(2)}, \dots, Y^{(n)} \mid \mu, \sigma\right) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left[-\frac{\sum(Y^i - \mu)^2}{2\sigma^2}\right] \tag{6}$$

and that it can be expressed in terms of the sample data, y,

$$L = \frac{1}{(2\pi\sigma^2)^{n/2}|\Psi|^{1/2}} \exp\left[-\frac{(y - 1\mu)^T \Psi^{-1}(y - 1\mu)}{2\sigma^2}\right] \tag{7}$$

where 1 is the n x 1 vector of ones. We also assume that each observation is independent and identical distributed.

The mathematical problem will be simpler, if we use the log-likelihood function by invoking the natural logarithm to obtain

$$\ln(L) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma^2) - \frac{1}{2}\ln(\Psi) - \frac{(y - 1\mu)^T \Psi^{-1}(y - 1\mu)}{2\sigma^2} \tag{8}$$

Thence, in order to find maxima of Eq. (8), by taking derivatives of Eq. (8) with respect to μ and σ^2 and setting the to zero: maximum likelihood estimates are then obtained for μ and σ^2.

$$\hat{\mu} = \frac{1^T \Psi^{-1} y}{1^T \Psi^{-1} 1} \tag{9}$$

$$\hat{\sigma}^2 = \frac{(y - 1\mu)^T \Psi^{-1}(y - 1\mu)}{n} \tag{10}$$

When Eq. (9) and Eq. (10) are substituted back into Eq. (8) and constant terms are removed, we obtain the log-likelihood function as follows:

$$\ln(L) \approx -\frac{n}{2}\ln(\sigma^2) - \frac{1}{2}\ln|\Psi| \tag{11}$$

Therefore, we use the log-likelihood function as the cost function as the objective in GA. As seen in Eq. (11), there is the correlation matrix which depends on θ and p, parameters. In order to reduce the complexity and calculation time, in this paper, besides taking p as 2, θ is constrained within the upper and lower limit values 10^2 and 10^{-3}, respectively.

Implementing the genetic algorithm in order to find the optimal values for θ, the correlation matrix Ψ is therefore obtained optimally. In next stage, Kriging prediction will be mentioned.

B. Kriging Prediction

In this section, Kriging correlation matrix is used to make prediction for new values based on the observed data. The derivation of the Kriging predictor follows the idea given in [12].

As the correlation parameters are found by GA to maximize log-likelihood function of observed data y, the new prediction $\hat{y}$ at X should be consistent with the observed data and therefore with the correlation. So, prediction should maximize the likelihood of the sample data and prediction itself by using the correlation parameters which are found in the previous section.

In order to achieve this, augmented observed data and augmented correlation matrix are defined by

$$\boldsymbol{\psi} = \begin{pmatrix} cor\left[Y\left(X^{(1)}\right), Y(X)\right] \\ cor\left[Y\left(X^{(2)}\right), Y(X)\right] \\ \vdots \\ cor\left[Y\left(X^{(n)}\right), Y(X)\right] \end{pmatrix} = \begin{pmatrix} \psi^{(1)} \\ \psi^{(2)} \\ \vdots \\ \psi^{(n)} \end{pmatrix} \tag{12}$$

and

$$\widetilde{\Psi} = \begin{pmatrix} \Psi & \boldsymbol{\psi} \\ \boldsymbol{\psi}^T & 1 \end{pmatrix} \tag{13}$$

respectively. Augmented observed data $\tilde{y} = \{y^T, \hat{y}\}$ and the vector of correlation are defined using the observed and the new prediction. Hence, the log-likelihood of augmented data becomes

$$\ln(L) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\hat{\sigma}^2) - \frac{1}{2}\ln\left|\widetilde{\Psi}\right| - \frac{(\tilde{y} - 1\hat{\mu})^T \widetilde{\Psi}^{-1}(\tilde{y} - 1\hat{\mu})}{2\hat{\sigma}^2} \tag{14}$$

Here, only the last term of log-likelihood expression depends on the new prediction $\hat{y}$, so we can consider just last term in the log-likelihood maximization process. Substituting the augmented variable, $\tilde{y}$ and $\widetilde{\Psi}$ yields

$$\ln(L) \approx \frac{-\begin{pmatrix} y - 1\hat{\mu} \\ \hat{y} - \hat{\mu} \end{pmatrix}^T \begin{pmatrix} \Psi & \psi \\ \psi^T & 1 \end{pmatrix}^{-1} \begin{pmatrix} y - 1\hat{\mu} \\ \hat{y} - \hat{\mu} \end{pmatrix}}{2\hat{\sigma}^2} \tag{15}$$

In order to maximize this log-likelihood expression, firstly, augmented correlation matrix should be inversed using partitioned inverse method [13], described also in Appendix-A:

$$\widetilde{\Psi}^{-1} = \begin{pmatrix} \Psi^{-1} + \Psi^{-1}\boldsymbol{\psi}(1 - \boldsymbol{\psi}^T\Psi^{-1}\boldsymbol{\psi})^{-1}\boldsymbol{\psi}^T\Psi^{-1} & -\Psi^{-1}\boldsymbol{\psi}(1 - \boldsymbol{\psi}^T\Psi^{-1}\boldsymbol{\psi})^{-1} \\ -(1 - \boldsymbol{\psi}^T\Psi^{-1}\boldsymbol{\psi})^{-1}\boldsymbol{\psi}^T\Psi^{-1} & (1 - \boldsymbol{\psi}^T\Psi^{-1}\boldsymbol{\psi})^{-1} \end{pmatrix} \tag{16}$$

which, then, simplifies Eq. (15) to

$$\ln(L) \approx \left(\frac{-1}{2\hat{\sigma}^2(1 - \boldsymbol{\psi}^T\Psi^{-1}\boldsymbol{\psi})}\right)(\hat{y} - \hat{\mu})^2 + \left(\frac{\boldsymbol{\psi}^T\Psi^{-1}(y - 1\hat{\mu})}{\hat{\sigma}^2(1 - \boldsymbol{\psi}^T\Psi^{-1}\boldsymbol{\psi})}\right)(\hat{y} - \hat{\mu}) \tag{17}$$

Hence, the maximum of log-likelihood estimate $\hat{y}$ can be found by differentiating and setting to zero:

$$\left(\frac{-1}{\hat{\sigma}^2(1 - \boldsymbol{\psi}^T\Psi^{-1}\boldsymbol{\psi})}\right)(\hat{y} - \hat{\mu}) + \left(\frac{\boldsymbol{\psi}^T\Psi^{-1}(y - 1\hat{\mu})}{\hat{\sigma}^2(1 - \boldsymbol{\psi}^T\Psi^{-1}\boldsymbol{\psi})}\right) = 0 \tag{18}$$

which results in

$$\hat{y} = \hat{\mu} + \boldsymbol{\psi}^T\Psi^{-1}(y - 1\hat{\mu}) \tag{19}$$

C. Modified Variable Complexity Model Implementation

In this section, Modified Complexity Model implementation is constructed using Kriging interpolation of the previous sections.

In modified VCM approach, data fusion function is approximated using Eq. (20):

$$f(x) \approx f_{lf}(x) - \beta(x) \tag{20}$$

where f_{lf} is low-fidelity Kriging interpolation model which created using low-fidelity dataset, β is the increment function which describes the difference between the low-fidelity Kriging interpolation model prediction and high-fidelity data points for each point in high fidelity dataset; it is defined by

$$\beta(x_o) = f_{hf}(x_o) - f_{lf}(x_o) \tag{21}$$

The modified VCM algorithm (taken from [2]):

1. Low-fidelity Kriging interpolation model is build using low-fidelity dataset.
2. Predict low-fidelity values for each data point in high-fidelity dataset. This step is necessary, because given dataset probably is not taken in same aerodynamic condition.
3. At each high-fidelity data point, the increments or differences are calculated between low-fidelity data which calculated in step-2 and high-fidelity data.
4. Increment function in Eq. (21) is calculated by interpolating the set of increments or differences data which are found in step-3 using Kriging interpolation approach.
5. Data fusion function in Eq. (20) is simple summation of low-fidelity Kriging and increment Kriging model for desired data point.

IV. Co-Kriging Approach

Using combination of large amount of low-fidelity data and small amount of high-fidelity data increases the accuracy of high-fidelity dataset function. In order to use low-fidelity data with high-fidelity data, some correction process should be defined and applied to low-fidelity data. This correction models the difference between low-fidelity and high-fidelity data. Actually, the modified VCM is a kind of correction processes. The modified VCM models the relation between low and high-fidelity data. Before, derivation of complex correction process, the correction process is simplified. If sample location of low-fidelity data, and high-fidelity data are same $(X_{hf} = X_{lf})$. The correction process takes the simple form:

$$y_{hf} = Z_\rho y_{lf} + Z_d \tag{22}$$

where, high-fidelity values y_{hf} at X_{hf} and low-fidelity values y_{lf} at X_{lf}. As seen in Eq. (22), basically, low-fidelity y_{lf} values multiply with scaling factor Z_ρ and adding the shifting factor to obtain high-fidelity y_{lf} values. If the correction process is simple like in Eq. (22), it can be said that using large amount of low-fidelity data with correction could make more accurate prediction than less amount of high-fidelity data [3]. However, correction process is not always simple. Generally, sample point locations of high and low-fidelity data are not the same. Therefore, we need a more powerful method in order to calculate correction between low- and high-fidelity data. Therefore in the next section, a more powerful, and more complex co-Kriging method, is introduced and formulated following the work in [3].

A. Building Co-Kriging Model

Co-Kriging is modified form of Kriging method. Additional to Kriging method, co-Kriging also correlates multiple sets of data. High-fidelity data with values y_{hf} at X_{hf} and low-fidelity data with values y_{lf} at X_{lf}. These data are concatenated as follows:

$$X = \begin{pmatrix} X_{lf} \\ X_{hf} \end{pmatrix} = \begin{pmatrix} {X_{lf}}^{(1)} \\ \vdots \\ {X_{lf}}^{(n_{lf})} \\ {X_{hf}}^{(1)} \\ \vdots \\ {X_{hf}}^{(n_{hf})} \end{pmatrix} \tag{23}$$

As with Kriging derivation, sample points X are considered as stochastic processes. Therefore, observed response are presented by a set of random vectors

$$Y = \begin{pmatrix} Y_{lf}(X_{lf}) \\ Y_{hf}(X_{hf}) \end{pmatrix} = \begin{pmatrix} Y_{lf}({X_{lf}}^{(1)}) \\ \vdots \\ Y_{lf}({X_{lf}}^{(n_{lf})}) \\ Y_{hf}({X_{hf}}^{(1)}) \\ \vdots \\ Y_{hf}({X_{hf}}^{(n_{hf})}) \end{pmatrix} \tag{24}$$

In co-Kriging derivation, we use auto-regressive model from [14]. Briefly, if the high-fidelity value y_{hf} is known at $x^{(i)}$, there is no contribution of low-fidelity value y_{lf} at point $x^{(i)}$. That means, we assume that the high-fidelity data is true and low-fidelity data contains all of the error terms.

Using auto-regressive model, high fidelity values can be approximated by multiplying low-fidelity values with constant scaling factor ρ plus a Gaussian process Z_d:

$$Z_{hf}(X) = \rho Z_{lf}(X) + Z_d(X) \tag{25}$$

where, $Z_{hf}(.)$, $Z_{lf}(.)$ and $Z_d(.)$ are Gaussian processes. They represent the local features of high- and low-fidelity data and the correlation between them is given in Eq. (4). Covariance matrix is defined as $cov[Y(X), Y(X)] = \sigma^2 \Psi(X, X)$, where Ψ is correlation matrix defined in Eq. (5). For co-Kriging, covariance matrix is constructed below in Eq. (26), Eq. (27) and Eq. (28). We note here that there are multiple covariance matrix, because of multiple dataset.

$$\begin{aligned} cov\left[Y_{lf}(X_{lf}), Y_{lf}(X_{lf})\right] = \quad & cov\left[Z_{lf}(X_{lf}), Z_{lf}(X_{lf})\right] \\ & {\sigma_{lf}}^2 \Psi_{lf}(X_{lf}, X_{lf}) \end{aligned} \tag{26}$$

$$\begin{aligned} cov\left[Y_{hf}(X_{hf}), Y_{lf}(X_{lf})\right] = \quad & cov\left[\rho Z_{lf}(X_{hf}) + Z_d(X_{hf}), Z_{lf}(X_{lf})\right] \\ & \rho {\sigma_{lf}}^2 \Psi_{lf}(X_{hf}, X_{lf}) \end{aligned} \tag{27}$$

$$\begin{aligned} cov\left[Y_{hf}(X_{hf}), Y_{hf}(X_{hf})\right] = \quad & cov\left[\rho Z_{lf}(X_{hf}) + Z_d(X_{hf}), \rho Z_{lf}(X_{hf}) + Z_d(X_{hf})\right] \\ & \rho^2 cov\left[Z_{lf}(X_{hf}), Z_{lf}(X_{hf})\right] + cov\left[Z_d(X_{hf}), Z_d(X_{hf})\right] \\ & \rho^2 {\sigma_{lf}}^2 \Psi_{lf}(X_{hf}, X_{hf}) + {\sigma_d}^2 \Psi_d(X_{hf}, X_{hf}) \end{aligned} \tag{28}$$

Complete covariance matrix is therefore

$$C = \begin{bmatrix} cov\left[Y_{lf}(X_{lf}), Y_{lf}(X_{lf})\right] & cov\left[Y_{lf}(X_{lf}), Y_{hf}(X_{hf})\right] \\ cov\left[Y_{hf}(X_{hf}), Y_{lf}(X_{lf})\right] & cov\left[Y_{hf}(X_{hf}), Y_{hf}(X_{hf})\right] \end{bmatrix} \tag{29.1}$$

or equivalently,

$$C = \begin{bmatrix} {\sigma_{lf}}^2 \Psi_{lf}(X_{lf}, X_{lf}) & \rho {\sigma_{lf}}^2 \Psi_{lf}(X_{lf}, X_{hf}) \\ \rho {\sigma_{lf}}^2 \Psi_{lf}(X_{hf}, X_{lf}) & \rho^2 {\sigma_{lf}}^2 \Psi_{lf}(X_{hf}, X_{hf}) + {\sigma_d}^2 \Psi_d(X_{hf}, X_{hf}) \end{bmatrix} \tag{29.2}$$

Correlations are the same as in Eq. (4) and there are two correlations, ψ_{lf} and ψ_d, as seen in Eq. (29.2). Therefore, we have more parameters to estimate than there are in Kriging method. To be explicit, these parameters are $\theta_{lf}, \theta_d, p_{lf}, p_d$ and scaling factor ρ.

Low-fidelity and high-fidelity data are independent of each other. Therefore, optimal values for $\theta_{lf}, p_{lf}, \mu_{lf}$ and ${\sigma_{lf}}^2$ can be found by using the maximum likelihood estimation formulation of low-fidelity data. Taking the natural logarithm, we obtain, without considering the constant terms,

$$\ln(L_{lf}) = -\frac{n_{lf}}{2} \ln({\sigma_{lf}}^2) - \frac{1}{2} \ln\left|\Psi_{lf}(X_{lf}, X_{lf})\right| - \frac{(y_{lf} - 1\mu_{lf})^T \Psi_{lf}(X_{lf}, X_{lf})^{-1} (y_{lf} - 1\mu_{lf})}{2{\sigma_{lf}}^2} \tag{30}$$

By setting the derivatives of Eq. (30) with respect to μ_{lf} and ${\sigma_{lf}}^2$ to zero and solving, Eq. (31) and Eq. (32) are obtained:

$$\hat{\mu}_{lf} = \frac{1^T \Psi_{lf}(X_{lf}, X_{lf})^{-1} y_{lf}}{1^T \Psi_{lf}(X_{lf}, X_{lf})^{-1} 1} \tag{31}$$

$${\hat{\sigma}_{lf}}^2 = \frac{(y_{lf} - 1\hat{\mu}_{lf})^T \Psi_{lf}(X_{lf}, X_{lf})^{-1} (y_{lf} - 1\hat{\mu}_{lf})}{n_{lf}} \tag{32}$$

Substituting Eq. (31) and Eq. (32) into Eq. (30) yields the concentrated log-likelihood:

$$\ln(L_{lf}) = -\frac{n_{lf}}{2} \ln({\hat{\sigma}_{lf}}^2) - \frac{1}{2} \ln|\Psi_{lf}(X_{lf}, X_{lf})| \tag{33}$$

As seen that there are two unknown variables in Eq. (33). These unknown variables are $\hat{\theta}_{lf}$ and $\hat{p}_{lf}$ (if not being set to 2) inside correlation matrix Ψ_{lf}. In order to find this unknown variable, one of global optimization method, such as the genetic algorithm may again be used.

Having found out the estimations $\hat{\theta}_{lf}$ and $\hat{p}_{lf}$, in order to find the rest of the unknown parameters, θ_d, p_d, μ_d and ${\sigma_d}^2$ and ρ, we first define

$$d = y_{hf} - \rho y_{lf}(X_{hf}) \tag{34}$$

where, $y_{lf}(X_{hf})$ are values of y_{lf} at the locations of high-fidelity data samples X_{hf}. We consider just high-fidelity data sample with respect to auto-regressive model, even if there are low-fidelity samples in the same locations. If there are no low-fidelity data samples at same location with high-fidelity samples We should estimate y_{lf} at location of high-fidelity data samples X_{hf}. In order to estimate ρ, Kriging prediction can be applied using the already determined hyper-parameters $\hat{\theta}_{lf}$ and $\hat{p}_{lf}$. After estimating ρ for each high-fidelity data, d can be calculated for each of these high-fidelity data points.

Note that the log-likelihood of the high-fidelity data can be written as

$$\ln(L_d) = -\frac{n_{hf}}{2} \ln({\sigma_d}^2) - \frac{1}{2} \ln|\Psi_d(X_{hf}, X_{hf})| - \frac{(d - 1\mu_d)^T \Psi_d(X_{hf}, X_{hf})^{-1} (d - 1\mu_d)}{2{\sigma_d}^2} \tag{35}$$

and, by setting the derivatives of Eq. (35) with respect to μ_d and ${\sigma_d}^2$ to zero, we obtain

$$\hat{\mu}_d = \frac{1^T \Psi_d(X_{hf}, X_{hf})^{-1} d}{1^T \Psi_d(X_{hf}, X_{hf})^{-1} 1} \tag{36}$$

$${\hat{\sigma}_d}^2 = \frac{(d - 1\hat{\mu}_d)^T \Psi_d(X_{hf}, X_{hf})^{-1} (d - 1\hat{\mu}_d)}{n_{hf}} \tag{37}$$

Then, substituting Eq. (36) and Eq. (37) into Eq. (35) yields the concentrated log-likelihood

$$\ln(L_d) = -\frac{n_{hf}}{2} \ln({\hat{\sigma}_d}^2) - \frac{1}{2} \ln|\Psi_d(X_{hf}, X_{hf})| \tag{38}$$

As seen that there are two unknown variables, again, in Eq. (38). These are $\hat{\theta}_d$ and $\hat{p}_d$ (if not being set to 2) inside correlation matrix Ψ_d and $\hat{p}$. Hence, once again, we use GA with log-likelihood cost function to estimate the optimal parameter values.

Next, we use this co-Kriging methodology in predictions.

B. Co-Kriging Prediction

After estimation of all hyper-parameters, the co-Kriging prediction step is derived as in [3]. Like Kriging prediction, co-Kriging prediction process follows the same methodology. The main idea is that prediction of a new high-fidelity sample point should be correlative with observed data and the parameters, which are found by using maximum likelihood estimation. Like Kriging, we augment the objective data with a predicted value to obtain the maximum likelihood estimation of the new high-fidelity point $\hat{y}_{hf}(x)$. To do so, we maximize the likelihood of the augmented dataset by varying prediction while keeping the hyper-parameters fixed.

Augmented dataset is defined as $\tilde{X} = \{X_{lf}^T X_{hf}^T x\}^T$ and $\tilde{y} = \{Y_{lf}^T Y_{hf}^T \hat{y}_{hf}(x)\}^T$. Augmented covariance matrix is given by

$$\tilde{C} = \begin{bmatrix} C & c \\ c^T & \rho^2 {\hat{\sigma}_{lf}}^2 + {\hat{\sigma}_d}^2 \end{bmatrix} \tag{39}$$

Furthermore, when we define the column vector of the covariance of X and x by

$$c = \begin{bmatrix} \rho {\sigma_{lf}}^2 \psi_{lf}(X_{lf}, x) \\ (\rho^2 {\hat{\sigma}_{lf}}^2 + {\hat{\sigma}_d}^2)\psi_d(X_{hf}, x) \end{bmatrix} \tag{40}$$

the augmented covariance matrix turns to be

$$\tilde{C} = \begin{bmatrix} {\sigma_{lf}}^2 \Psi_{lf}(X_{lf}, X_{lf}) & \rho {\sigma_{lf}}^2 \Psi_{lf}(X_{lf}, X_{hf}) & \rho {\sigma_{lf}}^2 \psi_{lf}(X_{lf}, x) \\ \rho {\sigma_{lf}}^2 \Psi_{lf}(X_{hf}, X_{lf}) & \rho^2 {\sigma_{lf}}^2 \Psi_{lf}(X_{hf}, X_{hf}) + {\sigma_d}^2 \Psi_d(X_{hf}, X_{hf}) & (\rho^2 {\hat{\sigma}_{lf}}^2 + {\hat{\sigma}_d}^2)\psi_d(X_{hf}, x) \\ \rho {\sigma_{lf}}^2 \psi_{lf}(X_{lf}, x) & (\rho^2 {\hat{\sigma}_{lf}}^2 + {\hat{\sigma}_d}^2)\psi_d(X_{hf}, x) & \rho^2 {\hat{\sigma}_{lf}}^2 + {\hat{\sigma}_d}^2 \end{bmatrix} \tag{41}$$

As shown in Eq. (30) and Eq. (35), ,it is seen that only the last term of the log-likelihood contains the sample data $\tilde{y}$, and hence to find the maximum likelihood estimations $\hat{y}_{hf}(x)$ we need to maximize last term of log-likelihood equation given by,

$$\ln(L) = -\frac{1}{2}(\tilde{y} - 1\mu)^T \tilde{C}^{-1}(\tilde{y} - 1\mu) \tag{42}$$

which may further be written in detail as follows:

$$\ln(L) \approx -\frac{1}{2}\begin{pmatrix} y - 1\hat{\mu} \\ \hat{y}_{hf}(x) - \hat{\mu} \end{pmatrix}^T \begin{pmatrix} C & c \\ c^T & \rho^2 {\hat{\sigma}_{lf}}^2 + {\hat{\sigma}_d}^2 \end{pmatrix}^{-1} \begin{pmatrix} y - 1\hat{\mu} \\ \hat{y}_{hf}(x) - \hat{\mu} \end{pmatrix} \tag{43}$$

Inverse of augmented covariance matrix can also be found using partitioned inverse formula described in appendix. Thus, substituting the inverse

$$\tilde{C}^{-1} = \begin{bmatrix} C^{-1} + C^{-1}c(\rho^2 {\hat{\sigma}_{lf}}^2 + {\hat{\sigma}_d}^2 - c^T C^{-1} c)^{-1} c^T C^{-1} & -C^{-1}c(\rho^2 {\hat{\sigma}_{lf}}^2 + {\hat{\sigma}_d}^2 - c^T C^{-1} c)^{-1} \\ -(\rho^2 {\hat{\sigma}_{lf}}^2 + {\hat{\sigma}_d}^2 - c^T C^{-1} c)^{-1} c^T C^{-1} & (\rho^2 {\hat{\sigma}_{lf}}^2 + {\hat{\sigma}_d}^2 - c^T C^{-1} c)^{-1} \end{bmatrix} \tag{44}$$

into Eq. (43) and ignoring the constant terms, we simplify the objective to

$$\ln(L) \approx \left(\frac{-1}{2(\rho^2 {\hat{\sigma}_{lf}}^2 + {\hat{\sigma}_d}^2 - c^T C^{-1} c)}\right)\left(\hat{y}_{hf}(x) - \hat{\mu}\right)^2 + \left(\frac{c^T C^{-1}(y - 1\hat{\mu})}{\rho^2 {\hat{\sigma}_{lf}}^2 + {\hat{\sigma}_d}^2 - c^T C^{-1} c}\right)\left(\hat{y}_{hf}(x) - \hat{\mu}\right) \tag{45}$$

to be maximized. Again, by taking the derivative with respect to $\hat{y}_{hf}(x)$ and setting the result to zero, we obtain

$$\left(\frac{-1}{\rho^2 {\hat{\sigma}_{lf}}^2 + {\hat{\sigma}_d}^2 - c^T C^{-1} c}\right)\left(\hat{y}_{hf}(x) - \hat{\mu}\right) + \left(\frac{c^T C^{-1}(y - 1\hat{\mu})}{\rho^2 {\hat{\sigma}_{lf}}^2 + {\hat{\sigma}_d}^2 - c^T C^{-1} c}\right) = 0 \tag{46}$$

so that the optimal solution

$$\hat{y}_{hf}(x) = \hat{\mu} + c^T C^{-1}(y - 1\hat{\mu}) \tag{47}$$

$$\hat{\mu} = \frac{1^T C^{-1} y}{1^T C 1}, y = (y_{lf}, y_{hf}) \tag{47}$$

is achieved.

V. Comparison of Kriging and Co-Kriging Method Results

In this section, the results obtained from the modified VCM and co-Kriging methods are compared. Results are plotted in form C_L, C_D, C_m vs α (angle of attack) for various horizontal tail deflections at Mach 0.2. In order to measure the success of the predictions, root mean square error (RMSE) of predicted and observed data values is implemented. Therefore, we need a test dataset to use RMSE criteria: high-fidelity data are used for this purpose, since we accept high-fidelity data as correct and low-fidelity data contain all possible error terms. Firstly, high-fidelity dataset is separated as train and test dataset. Then, train dataset is used for building co-Kriging model and Modified VCM model. In order to calculate RMSE for each approach, high-fidelity test dataset is used. RMSE formulation is therefore given by

$$RMSE = \sqrt{\frac{(y_{pred} - y_{obs})^2}{number\ of\ data}} \tag{48}$$

To lower the cost of data generation, the total number of test cases to construct high fidelity dataset should be decreased. Design of Experiment methods were developed to maximize the amount of information from a limited number of samples. The another issue is to select the number of ratio between high fidelity and low fidelity datasets. Total number of high and low fidelity dataset is in order 341 and 774. The RMSE value for ratio of high fidelity and low fidelity datasets from 10% to 90% is observed for each approach. 60% of ratio yields the most reasonable results as shown in Figure 9.

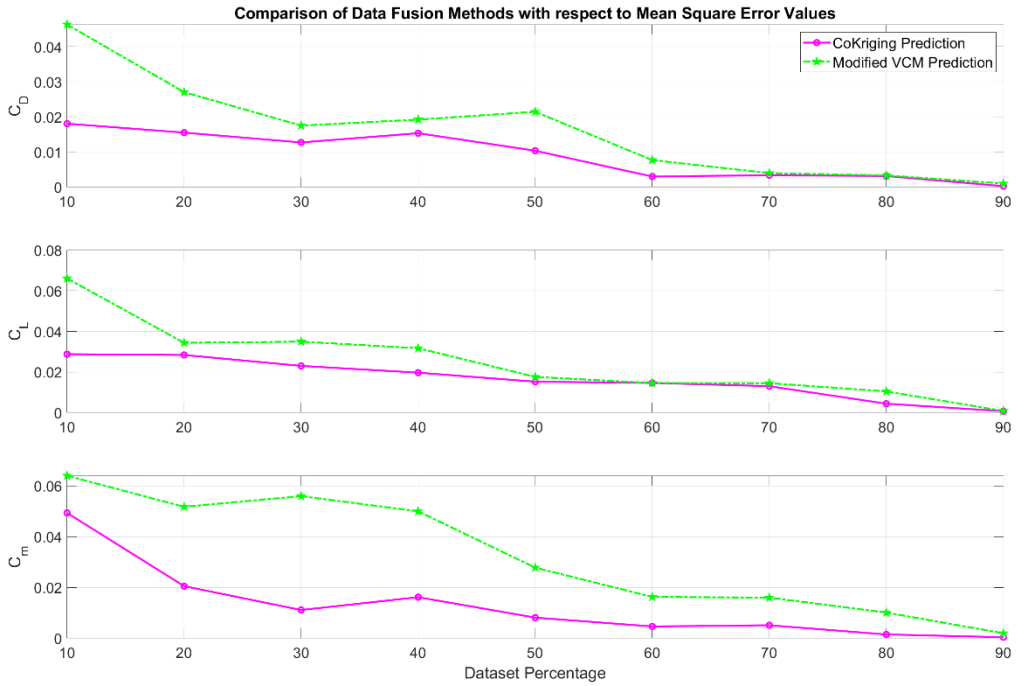

Figure 9. Comparison of data fusion methods with respect to mean square error values.

For the rest of study, we continue with 60% of high fidelity dataset with respect to low fidelity dataset.

Below, C_L, C_D, C_m vs α plots of various horizontal tail deflection conditions at Mach 0.2 are shown for the two data fusion methods; their performances are compared in terms of RMSE and illustrated in figures.

C_L vs α plot for Mach 0.2 and $-20, -10, 0, 10, 20$ horizontal tail deflections are given from Figure 10 to Figure 14,
C_L vs α plot for Mach 0.2 and $-20, 0, 20$ horizontal tail deflections are given in same plot in Figure 15,
C_D vs α plot for Mach 0.2 and $-20, -10, 0, 10, 20$ horizontal tail deflections are given from Figure 16 to Figure 19,
C_D vs α plot for Mach 0.2 and $-20, 0, 20$ horizontal tail deflections are given in same plot in Figure 20,
C_m vs α plot for Mach 0.2 and $-20, -10, 0, 10, 20$ horizontal tail deflections are given from Figure 21 to Figure 25,
C_m vs α plot for Mach 0.2 and $-20, 0, 20$ horizontal tail deflections are given in same plot in Figure 26.

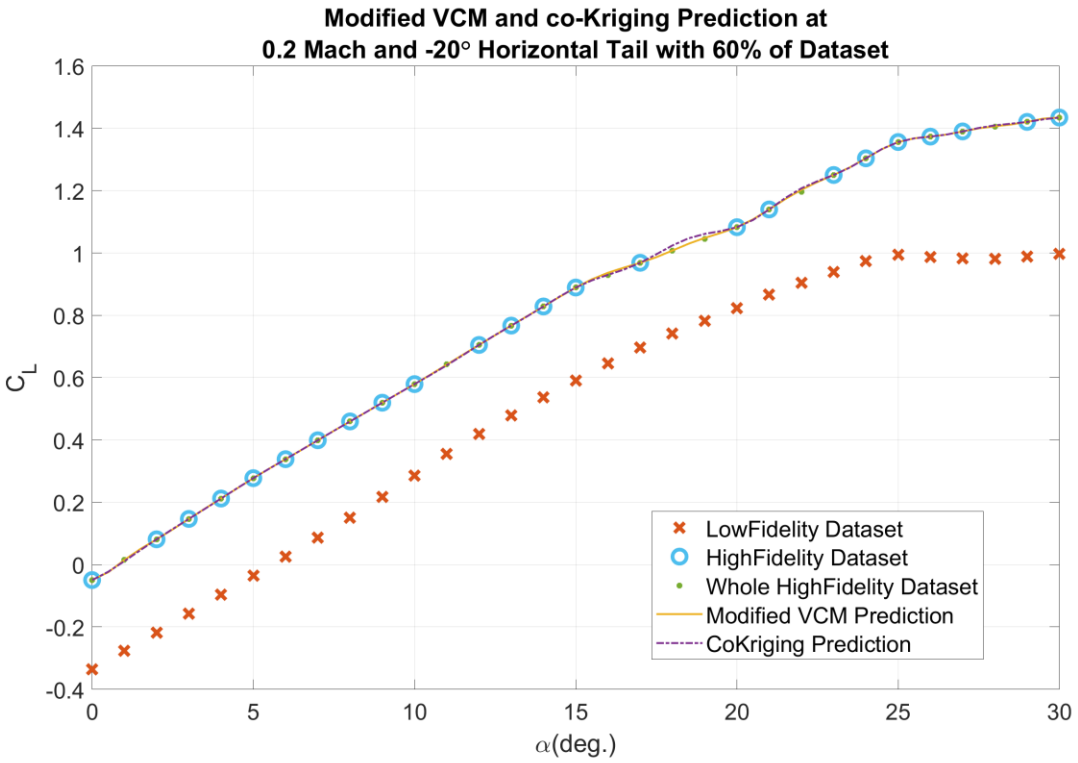

Figure 10. Comparison of Modified VCM and co-Kriging prediction for C_L vs α.

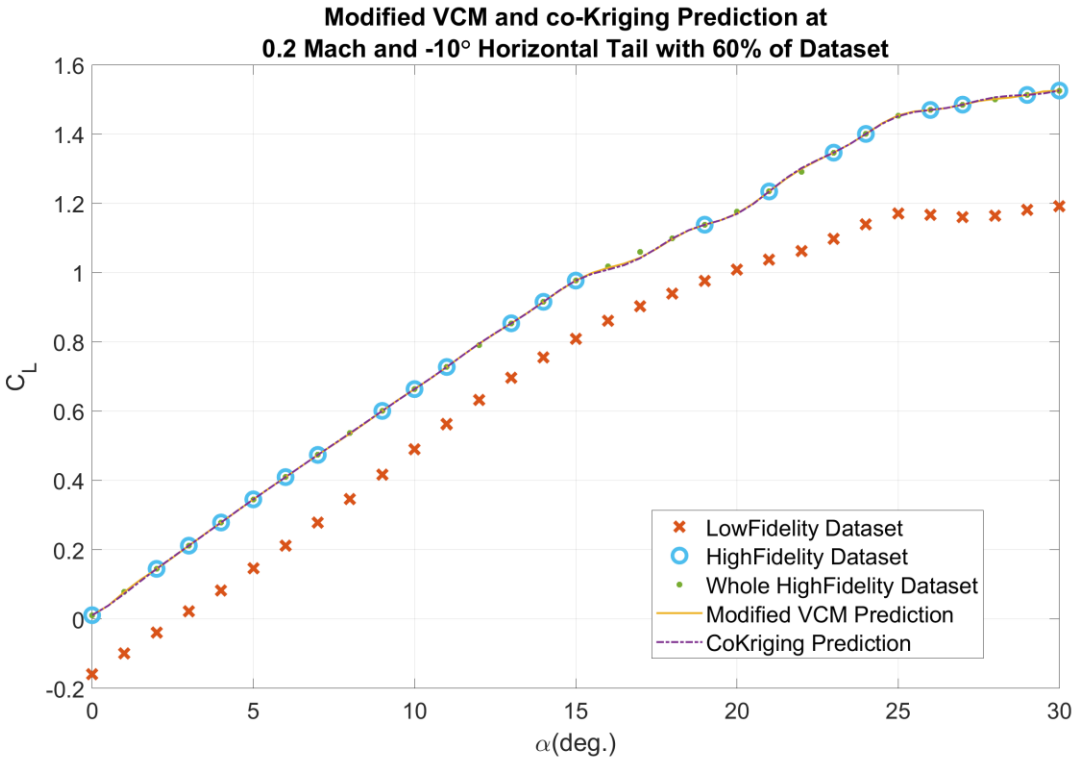

Figure 11. Comparison of Modified VCM and co-Kriging prediction for C_L vs α.

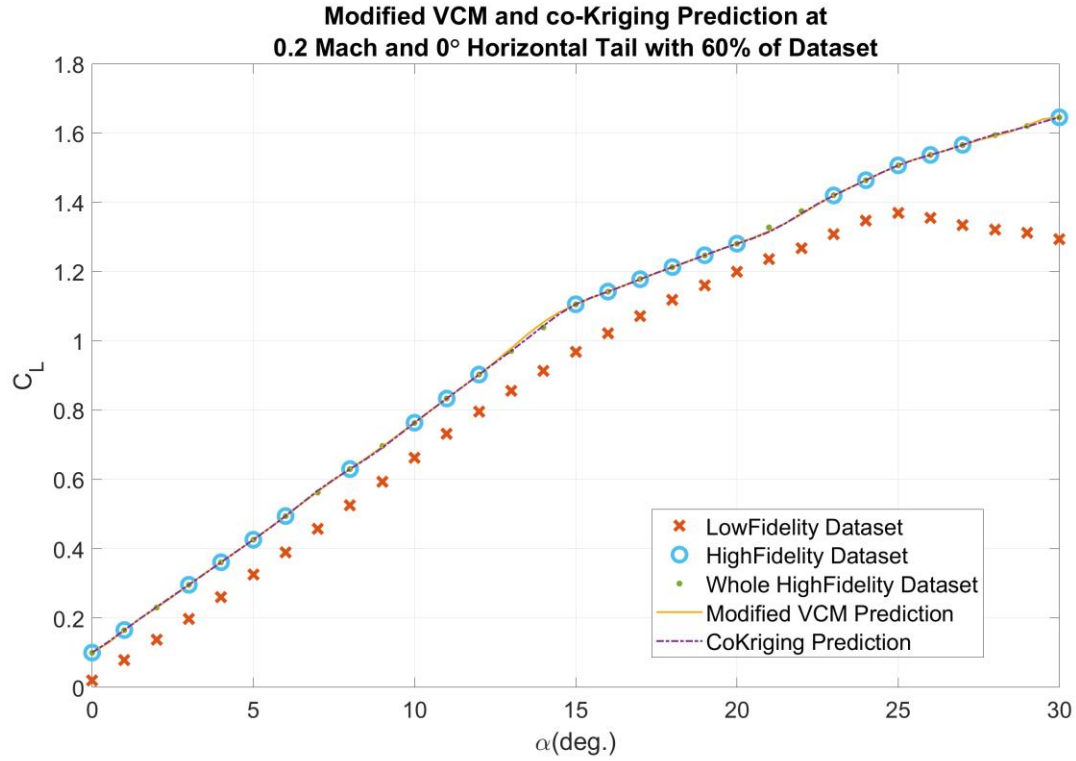

Figure 12. Comparison of Modified VCM and co-Kriging prediction for C_L vs α.

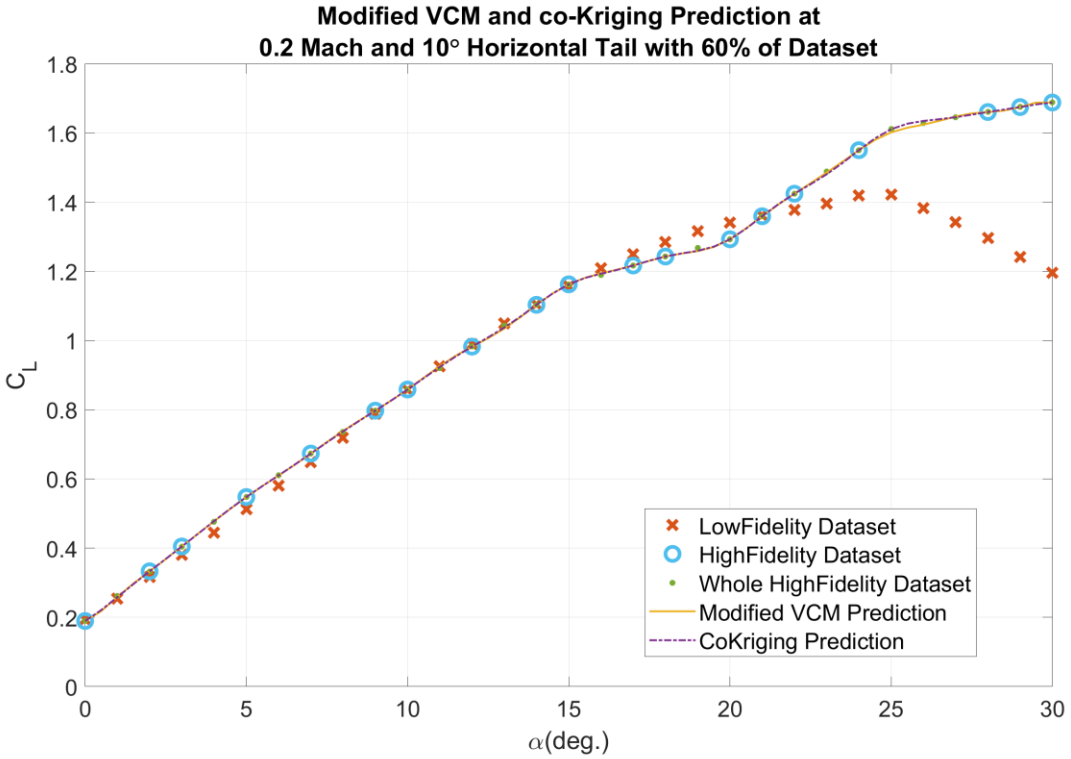

Figure 13. Comparison of Modified VCM and co-Kriging prediction for C_L vs α.

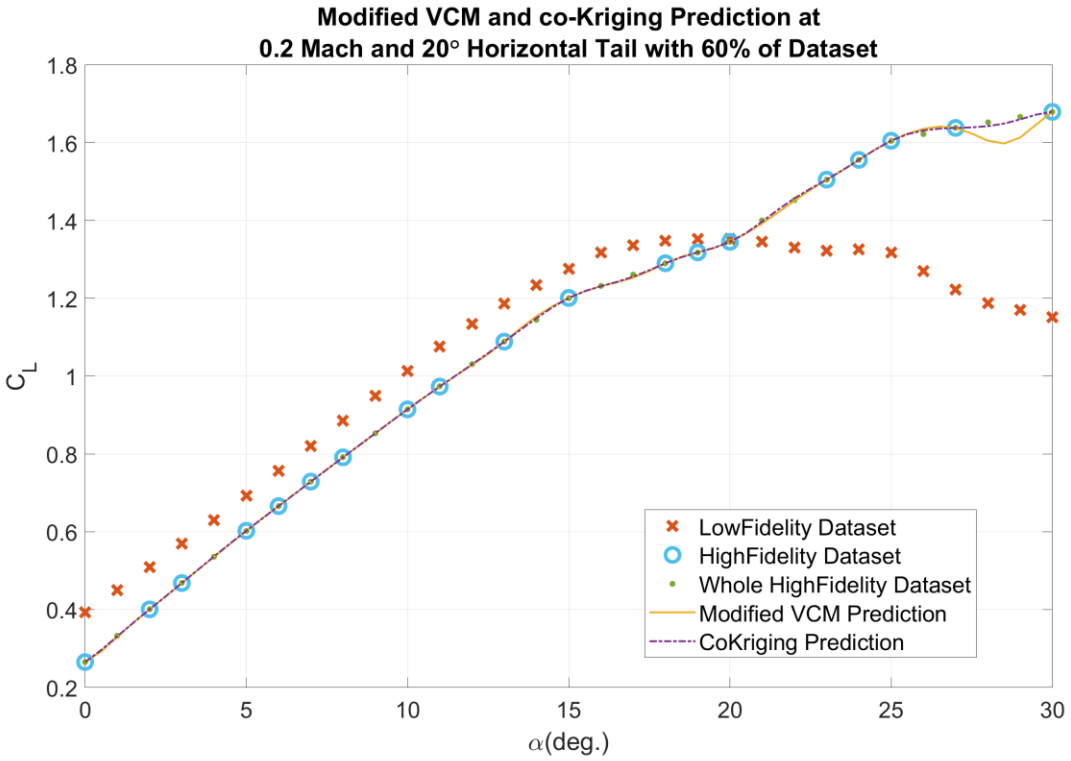

Figure 14. Comparison of Modified VCM and co-Kriging prediction for C_L vs α.

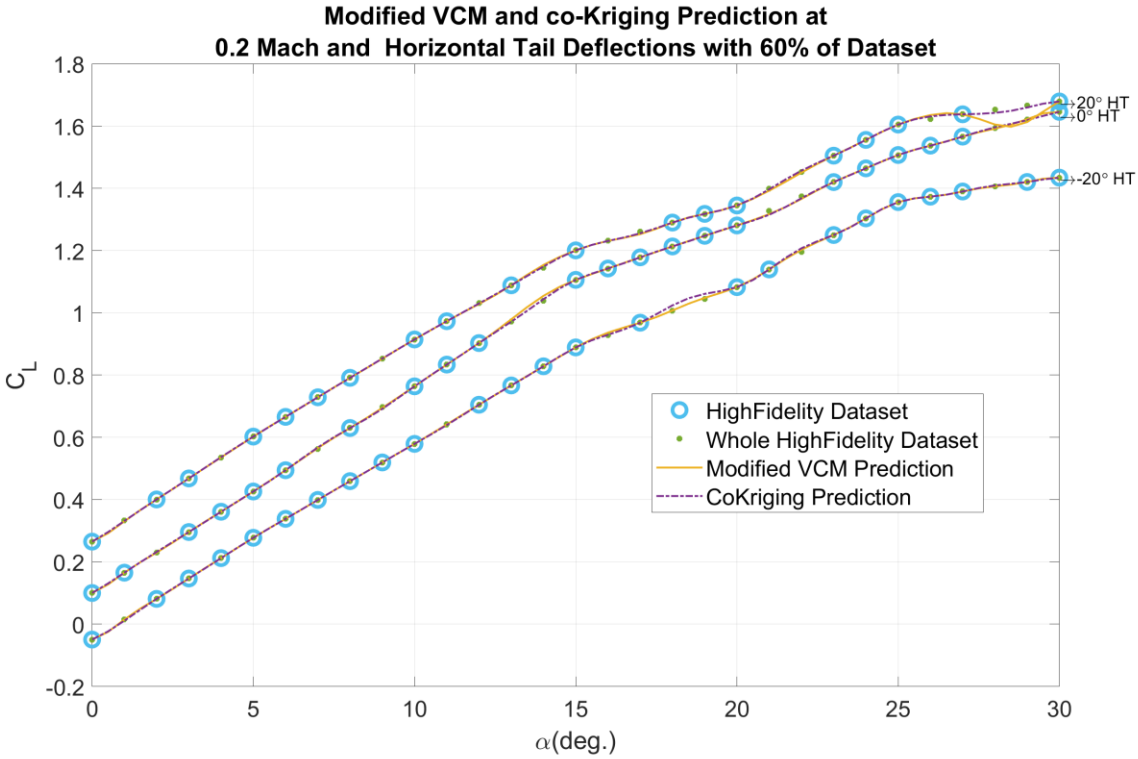

Figure 15. Comparison of Modified VCM and co-Kriging prediction for C_L vs α.

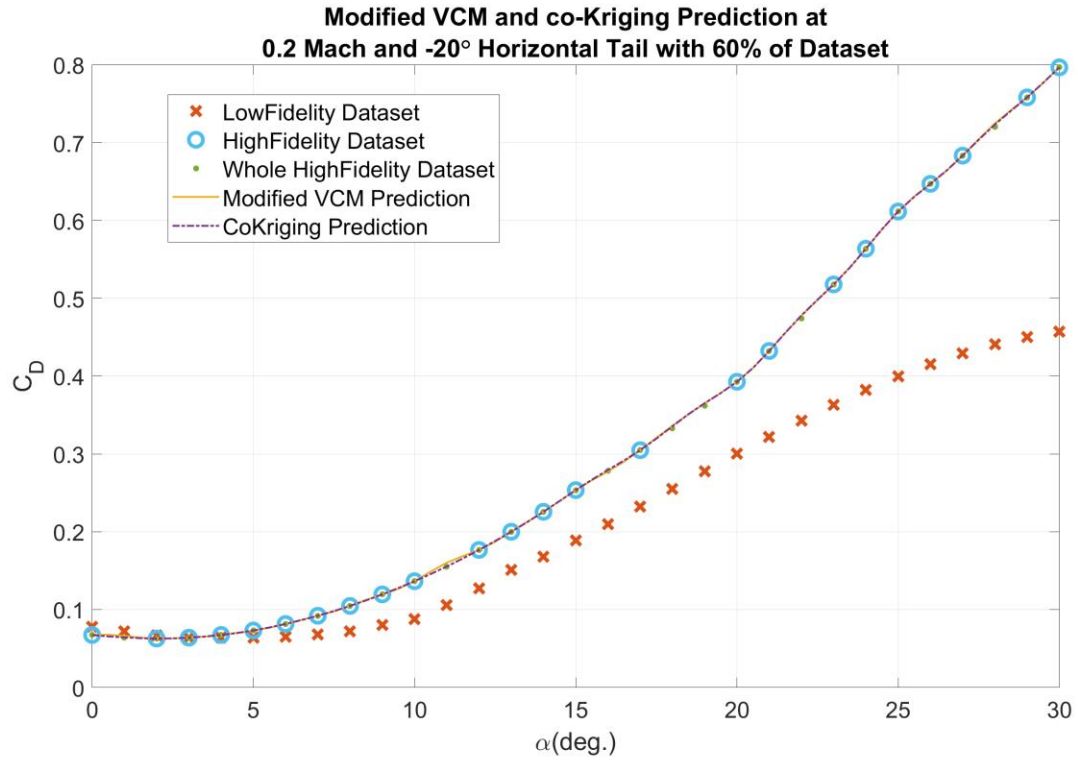

Figure 16. Comparison of Modified VCM and co-Kriging prediction for C_D vs α.

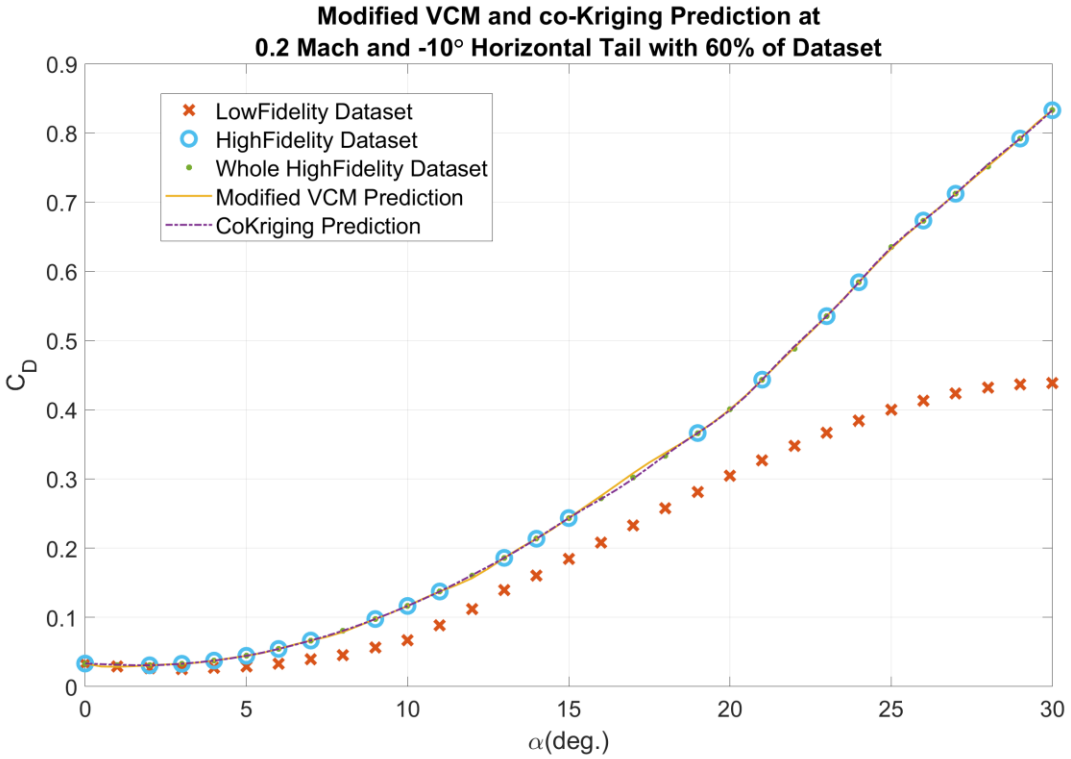

Figure 17. Comparison of Modified VCM and co-Kriging prediction for C_D vs α.

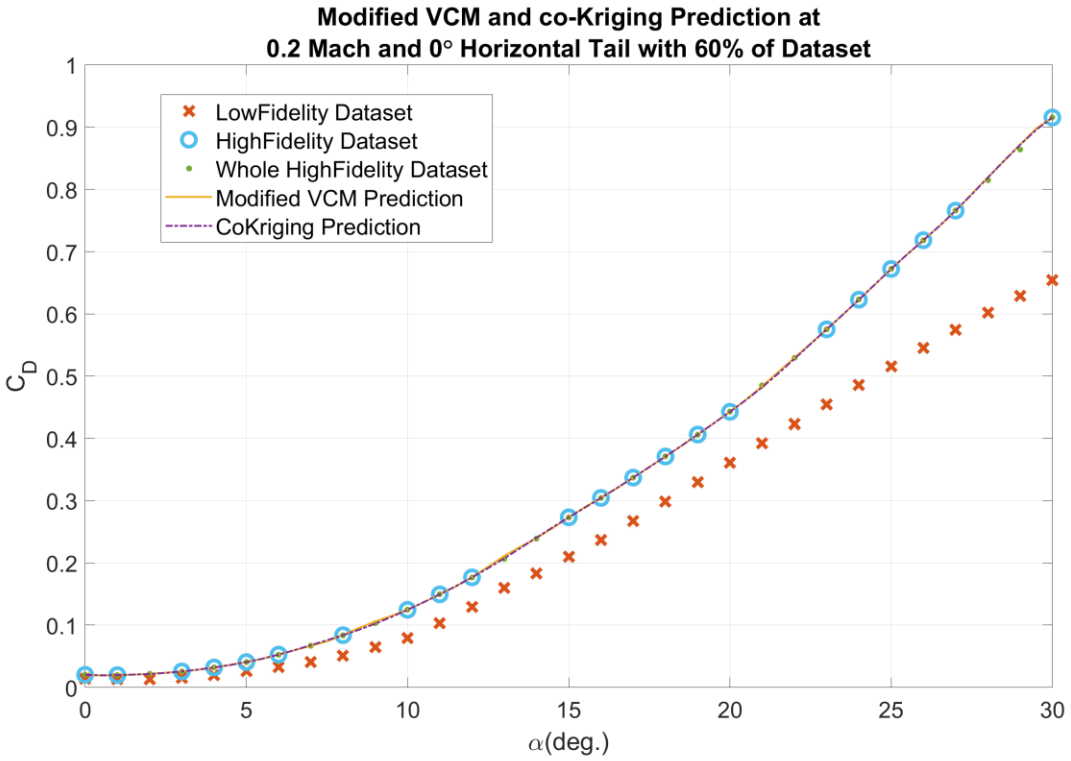

Figure 17. Comparison of Modified VCM and co-Kriging prediction for C_D vs α.

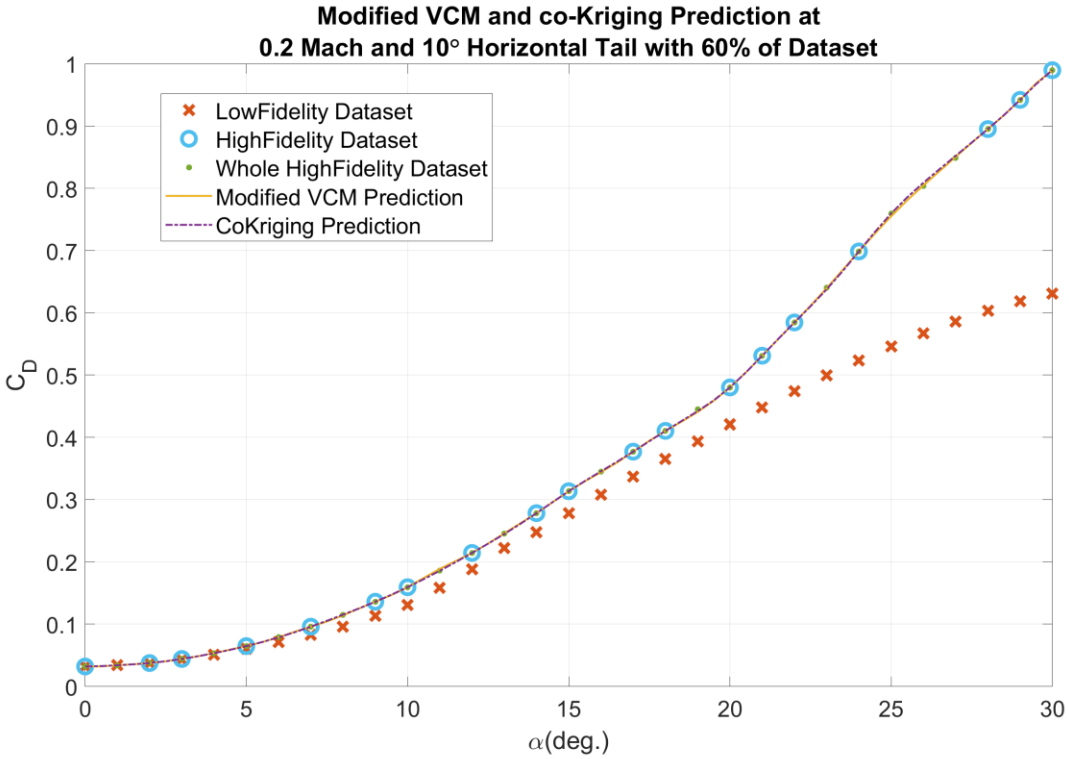

Figure 18. Comparison of Modified VCM and co-Kriging prediction for C_D vs α.

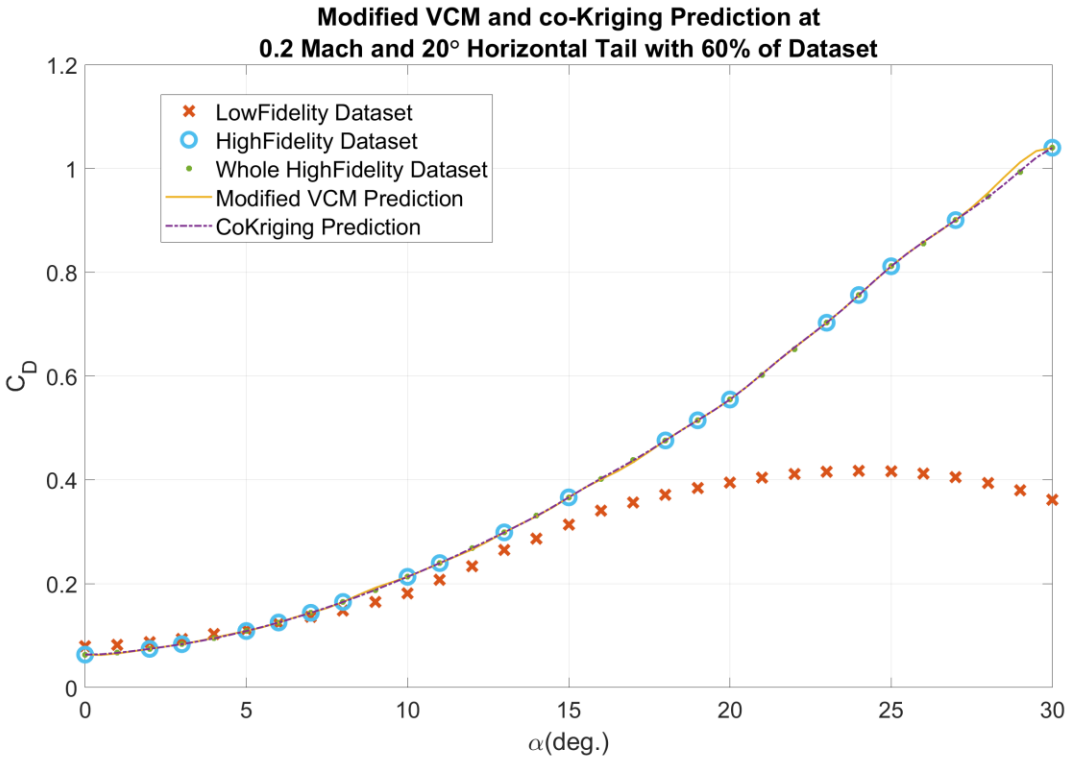

Figure 19. Comparison of Modified VCM and co-Kriging prediction for C_D vs α.

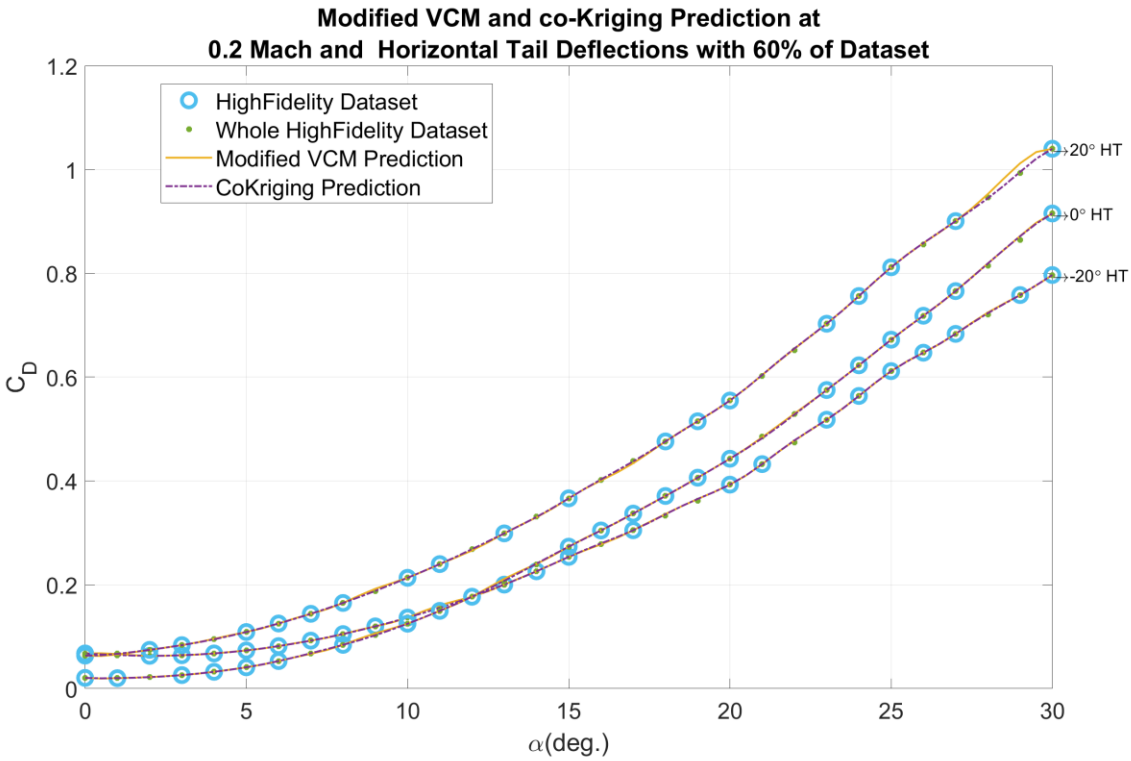

Figure 20. Comparison of Modified VCM and co-Kriging prediction for C_D vs α.

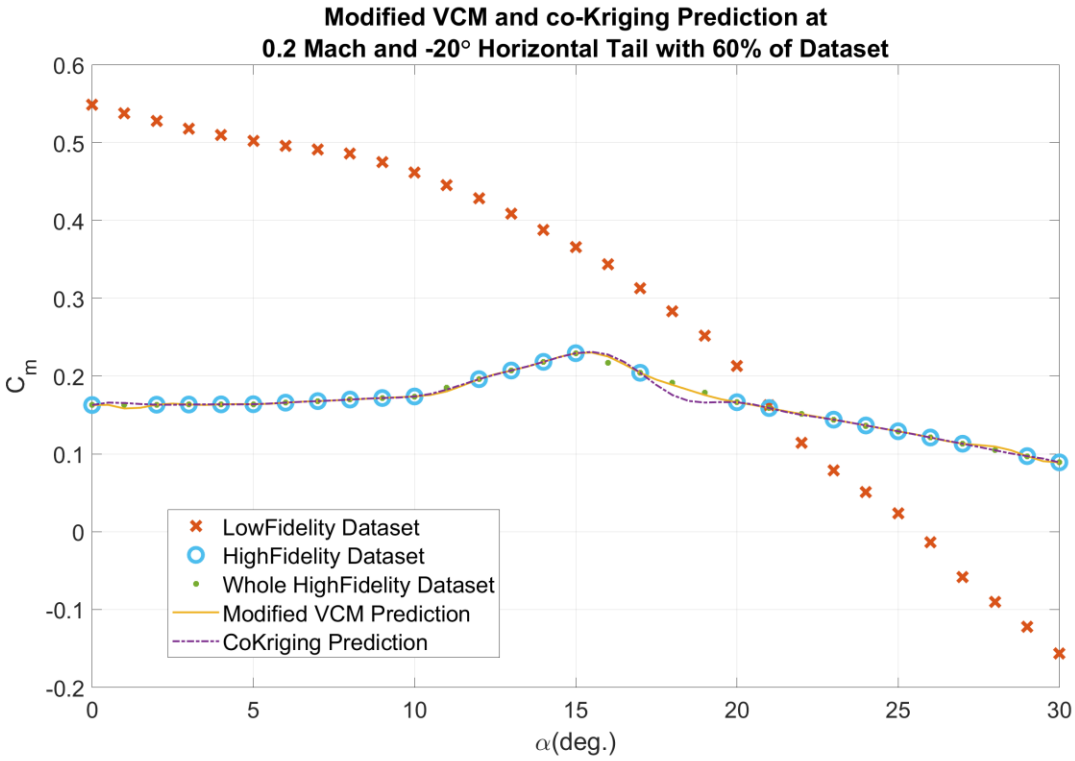

Figure 21. Comparison of Modified VCM and co-Kriging prediction for C_m vs α.

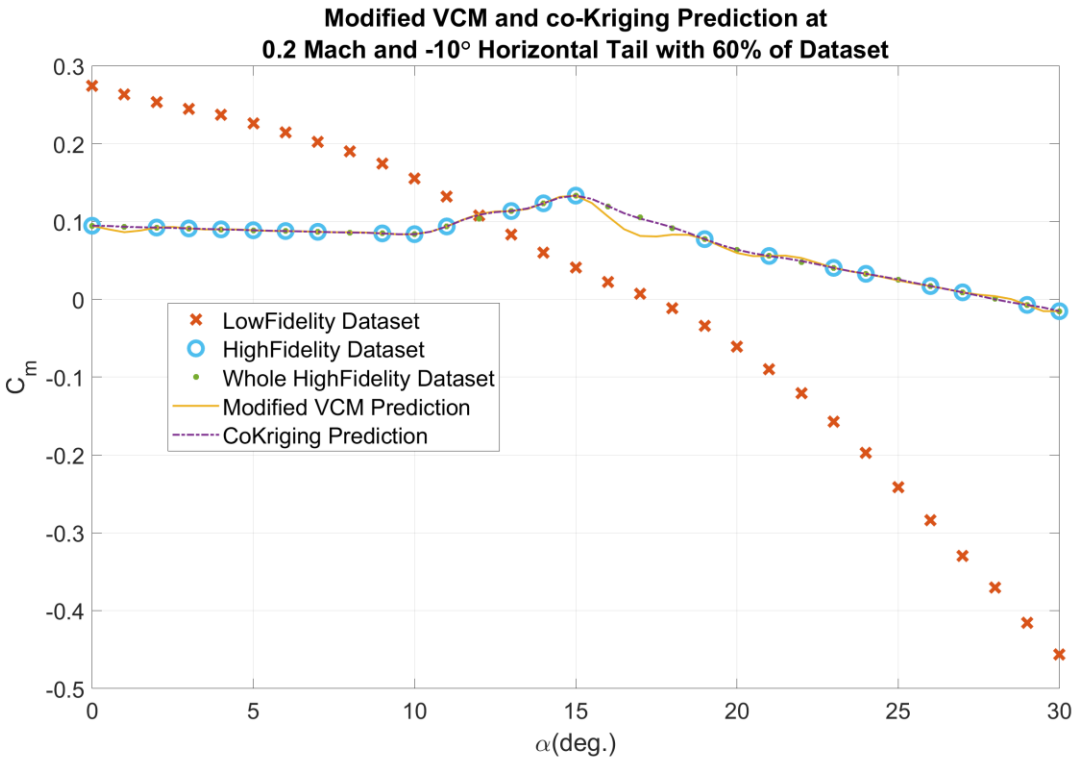

Figure 22. Comparison of Modified VCM and co-Kriging prediction for C_m vs α.

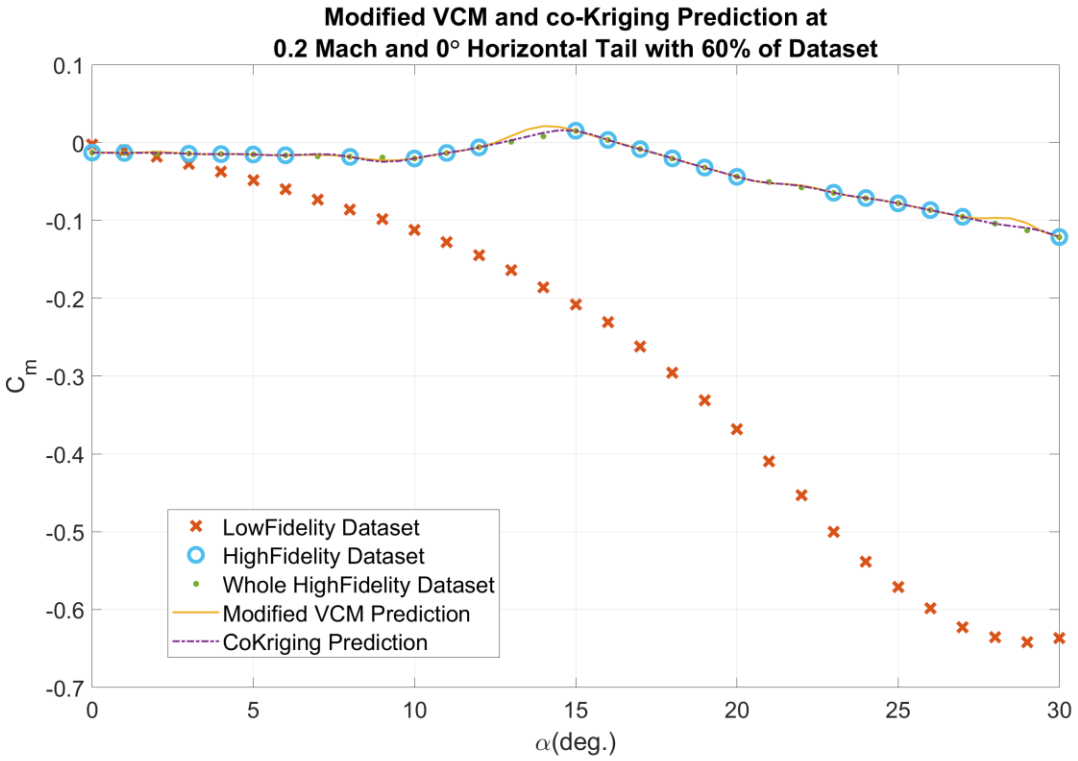

Figure 23. Comparison of Modified VCM and co-Kriging prediction for C_m vs α.

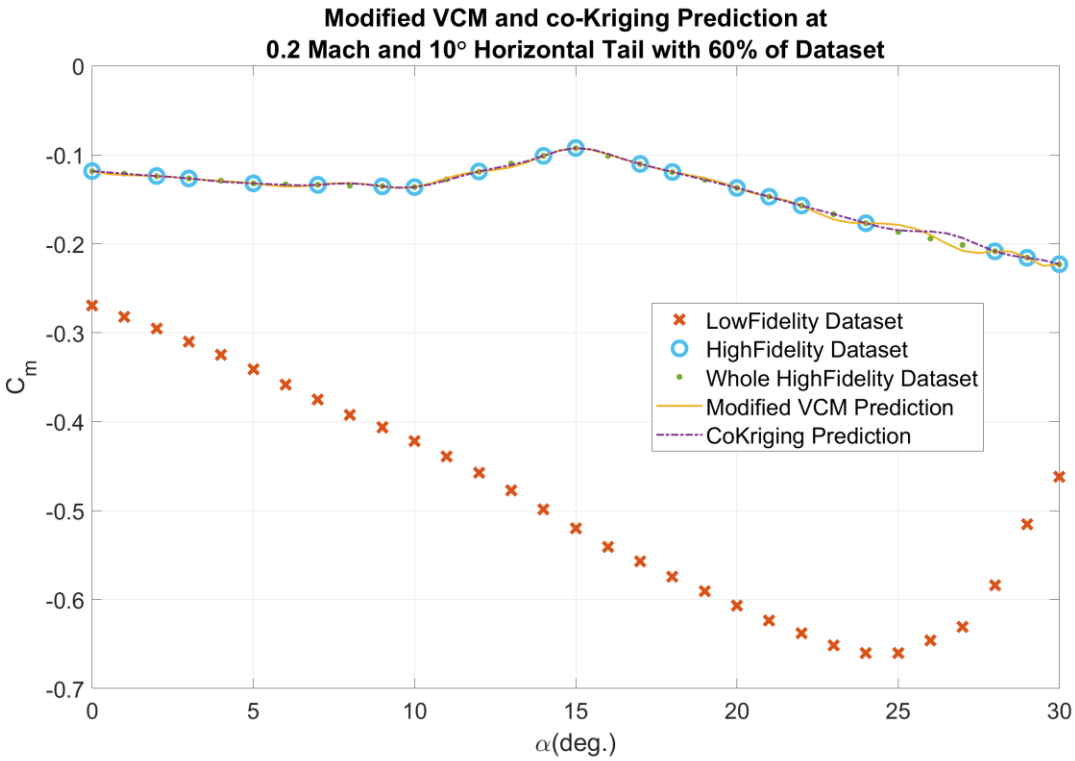

Figure 24. Comparison of Modified VCM and co-Kriging prediction for C_m vs α.

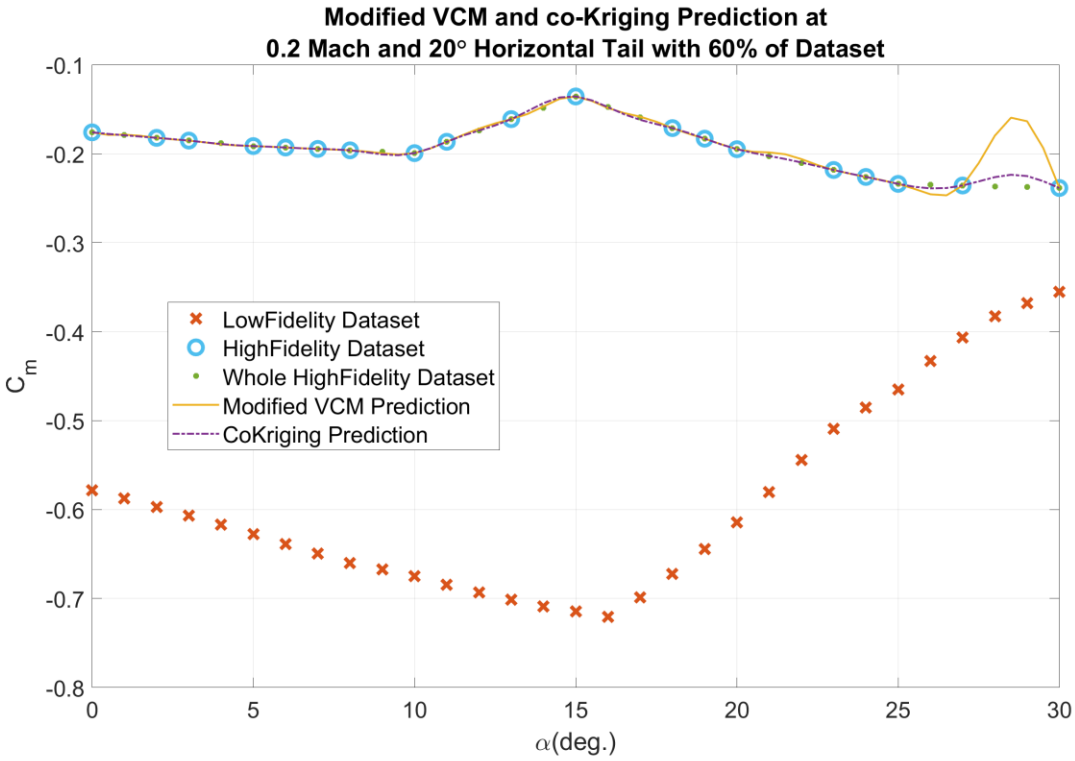

Figure 25. Comparison of Modified VCM and co-Kriging prediction for C_m vs α.

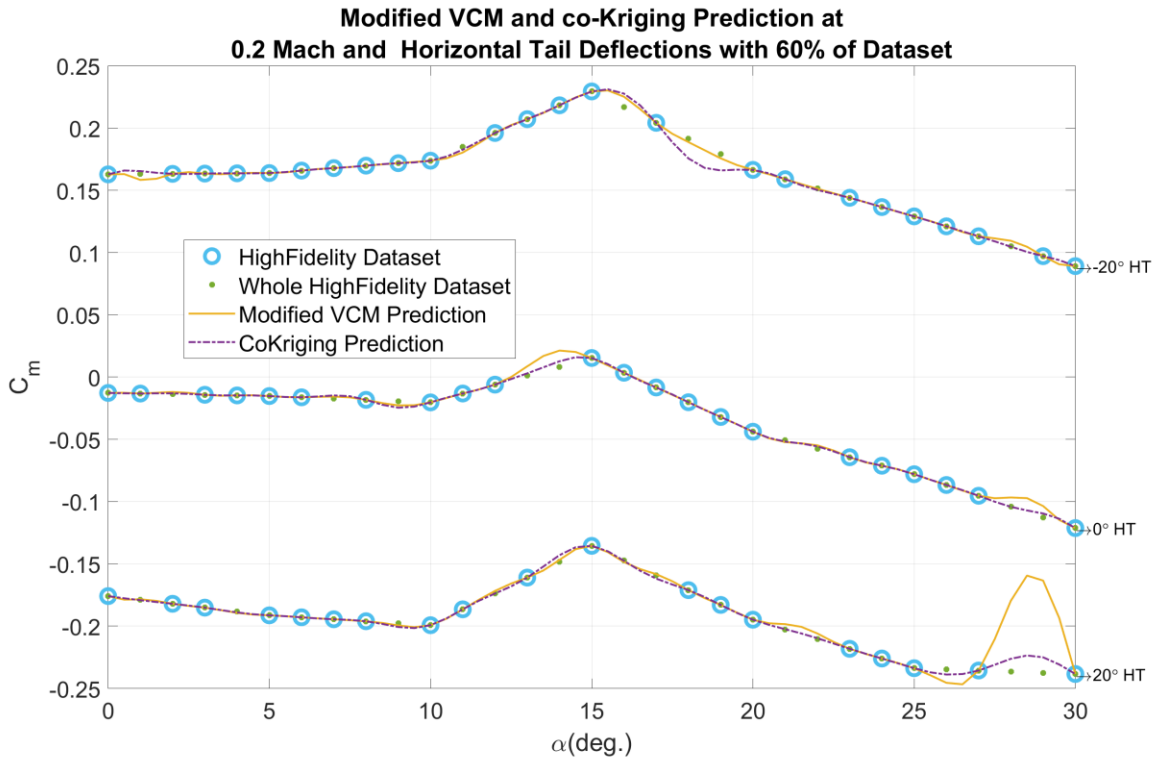

Figure 26. Comparison of Modified VCM and co-Kriging prediction for C_m vs α.

In Figures 10-26, comparison of the modified VCM and co-Kriging approach predictions are given. As seen from these figures, co-Kriging method provides more accurate than the modified VCM. To be specific, the RMSEs of both approaches for high fidelity data points are plotted at Figure 9. We should remark again that 60% of high fidelity gives reasonable results for both methods as seen in Figure 9.

In the modified VCM approach, the low-fidelity data provides the overall shape of the function whereas the high-fidelity data adjusts the data [15]. On the other hand, co-Kriging approach tries to correlate both set of data. The modified VCM approach uses Kriging model for prediction. Therefore, this approach passes exactly from high-fidelity data. However, if there are no high-fidelity data, the modified VCM shows an oscillatory trend. Conversely, co-Kriging approach uses correlation of high and low dataset. Thus, this approach becomes more robust and does not show an oscillatory behavior.

VI. Conclusion and Future Work

A simple F16 fighter aircraft geometry is obtained from the publicly open resources and its geometric details are corrected. Datcom input file is created and subsonic aerodynamic database of F16 is generated for the baseline and deflected horizontal tail configurations. This database is assigned as low-fidelity dataset. Subsonic wind tunnel data of F16 is obtained as the high-fidelity data. The differences in low and high-fidelity datasets are clarified. Similar trends are observed for lift and drag coefficients although there exist reasonably variable shifts between these datasets. However, the trends in pitch moment are not compatible with lift and drag coefficients in similarity of low- and high-fidelity data.

Co-kriging and the modified VCM approaches are used to derive high-fidelity datasets using low-fidelity datasets. The results obtained for the lift, drag and pitch moment coefficient are compared for various horizontal tail deflections. co-Kriging approach provides more accurate results than the modified VCM due to the lower rate of oscillations in the results. That is, the modified VCM tends to follow low fidelity data behavior, if there are no data samples around.

In fact, the RMSE values verify such a conclusion as well. Overall RMSE value of the modified VCM is approximately twice of co-Kriging approach, when number of high fidelity data is small as seen in Figure 9.

Additionally, both approaches are prone to diverge after the lower and upper bounds of fit. In other words, these approaches could not be used effective for extrapolations purposes. Therefore, for design purpose, high fidelity dataset should include upper and lower bounds.

Acknowledgements

The authors would like to thank Turkish Aerospace for providing technical data used in development of this study.

VII.References

[1] D. Allerton, Principles of Flight Simulation, West Sussex, UK: A John Wiley and Sons Inc., 2009.

[2] C. Tang, K. Gee and S. Lawrence, "Generation of Aerodynamic Data Using a Design of Experiment and Data Fusion Approach," in *AIAA Aerospace Science Meeting and Exhibit*, Reno, NV, 2005.

[3] A. Forrester, A. Sobester and A. Keane, "Multi-fidelity optimization via surrogate modelling," *Proceedings of the Royal Society: A Mathematical, Physical and Engineering Sciences*, vol. 463, no. 2088, pp. 3251-3269, 2007.

[4] A. Forrester, A. Sobester and A. Keane, Engineering Design via Surrogate Modelling: A Practical Guide, A John Wiley and Sons Inc, 2008.

[5] B. Peerlings, "A review of aerodynamic flow models, solution methods and solvers and their applicability to aircraft conceptual design," Delft University of Technology, Delft, 2018.

[6] M. Tyan, M. Kim, V. Pham, C. Choi, L. Nguyen and J. W. Lee, "Development of Advanced Aerodynamic Data Fusion Techniques for Flight Simulation Database Construction," in *AIAA Aviation Forum*, Atlanta, GA, 2018.

[7] Q. Liu, P. Zhang, W. Xiao and J. D. Diao, "Correction Methods of Aerodynamic Force and Moment Coefficients Based on the Identification Data," in *Chinese Control Conference*, Guangzhou, China, 2019.

[8] M. R. Malik and D. M. Bushnell, "Role of Computational Fluid Dynamics and Wind Tunnels in Aeronautics R&D," Nasa Technical Paper 217062, Hampton,VA, 2012.

[9] L. Nguyen, M. Ogburn, W. P. Gilbert, K. S. Kibler, P. W. Brown and P. L. Deal, "Simulator Study of Stall/PostStall Characteristics of a Fighter Airplane with Relaxed Longitudinal Static Stability," NASA Technical Paper 1538, Hampton, VA, 1979.

[10] William2002730. [Online]. Available: hangar.openvsp.org/vspfiles/343. [Accessed 05 06 2020].

[11] M. Fox and D. Forrest, "Supersonic Aerodynamic Characteristics of an Advanced F-16 Derivative Aircraft Configuration," NASA Technical Paper 3355, 1993.

[12] D. R. Jones, "A Taxonomy of Global Optimization Methods Based on Response Surfaces," *Journal of Global Optimization*, vol. 21, no. 4, pp. 345-383, 2001.

[13] H. Theil, Principles of Econometrics, A John Wiley and Sons, Inc., 1971.

[14] M. C. Kennedy and A. O'Hagan, "Predicting the Output from a Complex Computer Code When Fast Approximations Are Available," *Biometrika Trust*, vol. 87, no. 1, pp. 1-13, 2000.

[15] J. Bowles, W. Henline, L. Hyunh, C. Davies, C. Roberts and L. Yang, "Development of an Aerothermodynamic Environments Database for the Integrated Design of the X-33 Prototype Flight Test Vehicle," in *AIAA Aerospace Sciences Meeting and Exhibit*, Reno,NV, 1998.

Appendix

A. Partitioned Inverse of Matrix

For a nonsingular $n \times n$ matrix, A there is a unique $n \times n$ inverse matrix A^{-1}, which satisfies $AA^{-1} = A^{-1}A = I$ (I is an identity matrix). Therefore, given the nonsingular matrix

$$A = \begin{pmatrix} P_1 & R_1 \\ {R_1}^T & Q_1 \end{pmatrix} \tag{A1}$$

where P_1 and Q_1 are nonsingular (block) submatrices, we wish to solve

$$A^{-1}A = I = \begin{pmatrix} P_2 & R_2 \\ {R_2}^T & Q_2 \end{pmatrix}\begin{pmatrix} P_1 & R_1 \\ {R_1}^T & Q_1 \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} \tag{A1.1}$$

This splits into the four submatrix equations:

$$P_2P_1 + R_2{R_1}^T = I \tag{A1.2}$$

$$P_2R_1 + R_2Q_1 = 0 \tag{A1.3}$$

$${R_2}^TP_1 + Q_2{R_1}^T = 0 \tag{A1.4}$$

$${R_2}^TR_1 + Q_2Q_1 = I \tag{A1.5}$$

Hence, from Eq. (A1.4), we obtain ${R_2}^T = -Q_2{R_1}^T{P_1}^{-1}$, which is then substituted into Eq. (A1.5) to give $Q_2\left(Q_1 - {R_1}^T{P_1}^{-1}R_1\right) = I$, so

$$Q_2 = \left(Q_1 - {R_1}^T{P_1}^{-1}R_1\right)^{-1} \tag{A1.6}$$

and, therefore,

$${R_2}^T = -\left(Q_1 - {R_1}^T{P_1}^{-1}R_1\right)^{-1}{R_1}^T{P_1}^{-1} \tag{A1.7}$$

$$R_2 = -{P_1}^{-1}R_1\left(Q_1 - {R_1}^T{P_1}^{-1}R_1\right)^{-1} \tag{A1.8}$$

By substituting Eq. (A1.8) into Eq. (A1.2), clearly one gets $P_2P_1 - {P_1}^{-1}R_1\left(Q_1 - {R_1}^T{P_1}^{-1}R_1\right)^{-1}{R_1}^T = I$ and hence,

$$P_2 = {P_1}^{-1} + {P_1}^{-1}R_1\left(Q_1 - {R_1}^T{P_1}^{-1}R_1\right)^{-1}{R_1}^T{P_1}^{-1} \tag{A1.9}$$

Finally, by using Eq. (A1.6), (A1.7), (A1.8), and (A1.9) all in one to form A^{-1} we get

$$A^{-1} = \begin{pmatrix} {P_1}^{-1} + {P_1}^{-1}R_1\left(Q_1 - {R_1}^T{P_1}^{-1}R_1\right)^{-1}{R_1}^T{P_1}^{-1} & -{P_1}^{-1}R_1\left(Q_1 - {R_1}^T{P_1}^{-1}R_1\right)^{-1} \\ -\left(Q_1 - {R_1}^T{P_1}^{-1}R_1\right)^{-1}{R_1}^T{P_1}^{-1} & \left(Q_1 - {R_1}^T{P_1}^{-1}R_1\right)^{-1} \end{pmatrix} \tag{A1.10}$$