Aerodynamic model identification of the flying V from wind tunnel data

Ruiz-García, Alberto; Vos, Roelof; de Visser, Coen C.

DOI
10.2514/6.2020-2739

Publication date
2020

Document Version
Final published version

Published in
AIAA AVIATION 2020 FORUM

AIAA AVIATION Forum
June 15-19, 2020, VIRTUAL EVENT
AIAA AVIATION 2020 FORUM

10.2514/6.2020-2739

Aerodynamic Model Identification of the Flying V from Wind Tunnel Data

Alberto Ruiz-García*, Roelof Vos†, Coen C. de Visser‡
Delft University of Technology, Delft, 2629 HS, The Netherlands

The aerodynamic model identification of a novel aircraft configuration known as the "Flying V" is presented. A global longitudinal aerodynamic model is estimated using static wind tunnel data from a 4.6% sub-scale model. The aerodynamic model structure, unknown *a priori*, is determined from the data using a modified stepwise regression technique. Orthogonal polynomial models using Multivariate Orthogonal Functions and non-orthogonal spline models in the angle-of-attack dimension are defined for the estimation of the measured aerodynamic coefficients. The estimated models are validated against a partition of the data not used for the estimation, which shows that an adequate model fit and good prediction capabilities are attained. Spline models achieve better results in terms of fitting and show a better matching between the estimation and validation data. All estimated models are considered adequate, with a maximum relative Root Mean Square error below 8% for the polynomial models and below 3% for the spline models.

Nomenclature

Latin Symbols

b	Wing span of the half model, [m]
$\bar{c}$	Reference chord, [m]
C_X, C_Y, C_Z	Force coefficients in body axes, [N]
C_l, C_m, C_n	Moment coefficients in body axes, [N·m]
F	Force [N]
G	Orthogonalization matrix
J	OLS cost function
L, M, N	Roll, pitch and yaw moment, [N·m]
M	Moment, [N·m]
n	Number of regressors in model
N	Number of samples
$\mathbf{p}_j$	Original regressors, $j = 1, 2, ..., n$
q	Dynamic pressure, [kg·m/s^2]
S	Wing area, [m^2]
$\mathbf{v}$	Residual vector
V	Wind speed, [m/s]
$\hat{V}$	Dimensionless wind speed, [-]
$\mathbf{x}$	State vector
X	Regression matrix
X, Y, Z	Aerodynamic forces in body axes, [N]
$\mathbf{y}$	Output vector
$\mathbf{z}$	Measurement vector

Greek Symbols

α	Angle of attack, [deg, rad]
δ	Control surface deflection [deg]
$\boldsymbol{\theta}$	Parameter vector
$\boldsymbol{\xi}_j$	Orthogonal regressors, $j = 1, 2, ..., n$

Subscripts

0	original
$Xb,\ Yb,\ Zb$	Balance reference frame
$i,\ j$	integers

Acronyms

BLUE	Best Linear Unbiased Estimator
CFD	Computational Fluid Dynamics
MAC	Mean Aerodynamic Chord
OJF	Open Jet Facility
OLS	Ordinary Least Squares
PSE	Predicted Square Error
RMS	Root Mean Square Error
SVD	Singular Value Decomposition
VIF	Variance Inflation Factors

*Researcher, Faculty of Aerospace Engineering
†Assistant Professor, Faculty of Aerospace Engineering, Associate Fellow AIAA
‡Assistant Professor, Faculty of Aerospace Engineering, AIAA Member

I. Introduction

In pursuit of a strong reduction of the global-warming impact of large, long-range transport aircraft, a new aircraft configuration was proposed in 2015 [1]. This configuration, dubbed "Flying V", is essentially a highly-swept flying-wing aircraft with the passenger cabin inside the wing. Initial studies carried out at Airbus showed how this conceptual design could theoretically outperform the Airbus A350, improving by 10% its aerodynamic efficiency and having a slightly lower operating empty weight. This conceptual design was improved by Faggiano et al. [2], who optimized the planform geometry and aerodynamic surfaces increasing the aerodynamic efficiency of the aircraft to 23. Moreover, the wing parameterization utilized an oval fuselage cross section, a structurally efficient pressure vessel devised by Vos et al. [3]. Rubio Pascual and Vos showed that the installation of over-the-wing turbofan engines is very sensitive to location and could cause a penalty in aerodynamic efficiency of 10% [4]. An impression of the Flying V is shown in Fig. 1. Based on the optimized aerodynamic shape of Ref. [2], Palermo and Vos performed wind tunnel tests on a 4.6% scaled wind tunnel modelt at a Reynolds number of one million. They demonstrated an untrimmed maximum lift coefficient of 1.02, an aerodynamic center excursion over angle of attack as much as $55\%\bar{c}$ due to vortex formation, and a trimmed maximum lift coefficients are 0.68 [5].

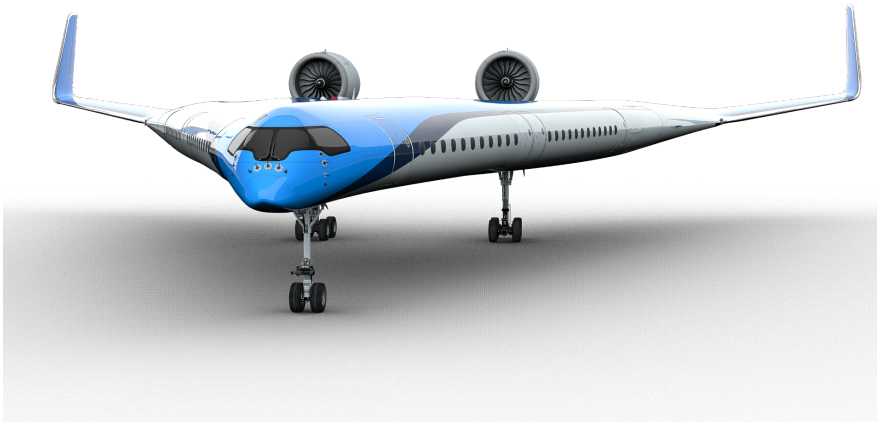

Fig. 1 Artist impression of the Flying V, a long-haul passenger aircraft, seating more than 300 passengers in a standard two-class configuration.

While Palermo and Vos studied the effect of control power and the mutual interaction of the trailing-edge control surfaces, an aerodynamic model was not derived from the data. An aerodynamic model allows to predict the response of the aircraft at conditions that might not have been tested. At the same time it provides a better understanding of the aerodynamic behavior without detailed knowledge of the underlying physical effects such as flow separation or vortex formation. Such a model is needed if one wants to predict the performance and handling qualities of the aircraft. It also opens different lines of work for control system design, such as the design of an Automatic Flight Control System or control allocation studies.

To derive the aerodynamic model of an airplane, different theoretical approaches are possible such as panel methods, strip theory, high-fidelity CFD, or a variety of methods developed from the experience acquired throughout the years (e.g. U.S. Air Force DATCOM) [6]. Nevertheless, even with the increased computational power available nowadays, wind tunnel tests and flight tests are still needed to assess the nonlinear aerodynamics that happen, for example, at high angle-of-attack.The main objective of the present work is therefore *to derive a global longitudinal aerodynamic model of a sub-scale Flying V aircraft from wind-tunnel data.*

To derive the aerodynamic model, a System Identification approach is proposed. The field of System Identification deals with the creation of a mathematical model from data gathered during experiments in order to describe a system based on the observed input and output behavior. To determine the aerodynamic model of the Flying-V subscale test article, two main problems are addressed in the present work: the determination of an adequate model structure, and the estimation of its parameters. Because of the novelty of this aircraft configuration, no existing models are available in the open literature. Therefore, the model

structure is determined from experimental data obtained during a dedicated wind tunnel test campaign.

To determine an adequate model structure, stepwise regression stands out as a good candidate to tackle this problem, as it is a well known statistical technique which has been successfully used in the field of aerodynamic model identification to derive an adequate model structure [7–10]. Several statistical metrics can be used to help with the model selection process. In some references where Multivariate Orthogonal Functions are used the model selection process is automatic and based on a single metric, the Predicted Square Error (PSE), and regressors are added to the model in order of greatest to smallest reduction of the cost function [11–13]. On the other hand, usual stepwise regression is based on correlations and on the partial and total F-statistic[14] for the addition of regressors and final model selection. More complex estimation methods such as B-splines[15, 16], Multivariate Simplex Splines[17], or Neural Networks[18] are not considered for this early phase of the study. The main reasons are the limited data available and the uncertainty if such powerful methods are indeed required.

In the present work orthogonal polynomial models using Multivariate Orthogonal Functions [11] and non-orthogonal spline models are postulated. To compare their prediction accuracy, independent model identifications are performed. It should be noted that only static wind tunnel data is used for the derivation of the aerodynamic model. Therefore, a dynamic model could not be estimated. Nevertheless, once dynamic data from a scaled flight test article is available, the presented technique can also be used to estimate a complete dynamic model of the Flying V.

The paper is structured as follows. A description of the wind tunnel model and the conducted experiments is given in Section II. The mathematical models for the polynomial and spline estimation are presented in Section III. The methodology for the model structure determination and the metrics used for the model selection are discussed in Section IV. The estimated polynomial and spline models are presented in Section V and mutually compared according to their statistical qualities. Finally, some concluding remarks are given in Section VI.

II. Experiment Description

The data for the present work come from a series of wind tunnel tests performed in January 2019. A 4.6%-scaled half model of the Flying V was built to conduct these tests. It was designed using the so-called Froude scaling law in order to duplicate the inertial and gravitational effects in a later flight test and hence to be able to predict the rigid-body modes for the full-scale aircraft [19]. The scaling factor of 4.6% was the maximum possible given the current drone regulations in the Netherlands. It corresponds with a model weight of 25 kg. Engine and winglet were left out of the study and were not implemented in the model. However, the methodology proposed in this work does not depend on the aircraft configuration, and can be used again once these components are added to estimate their contributions.

The wind tunnel used for the tests was the Open Jet Facility (OJF) from the Faculty of Aerospace Engineering of Delft University of Technology. Its cross section is a 285×285 cm octagon, and the wind speed ranges from 3 to around 30 m/s. The velocity deviations in the vertical plane measured two meters from the outlet are smaller than 0.5%, and the longitudinal turbulence intensity level is smaller than 0.24% [20]. A CAD model of the setup and the basic dimensions of the wind tunnel model are shown in Fig. 2. The different components of the setup can be distinguished in the figure. The model is elevated from the bottom wall of the wind tunnel in order to avoid interactions with the wind tunnel-wall boundary layer. A reflection plane was placed in order to aerodynamically isolate the wing from the connection to the balance. The connection between wing and balance was shielded from the wind by means of an aerodynamic cover, which connected to both the splitting plane and the reflection plane and prevented the air from influencing the balance measurements.

The test-article is a thin-walled fiber-glass wing, with three control surfaces. The exact dimensions and positions of this model are reported in a previous study [5]. The reference area of the model is $S = 0.9345$ m^2, has a semi span of $b/2 = 1.495$ m, and a reference chord of $\bar{c} = 0.729$ m. The wing consists of two tapered wing trunks with a leading-edge kink positioned at 63% of the semi-span. Inboard of the kink, the leading-edge sweep angle measures 64.4 degrees, while outboard of the kink, the leading-edge sweep angle is 37.8 degrees. The aspect ratio of the wing is 4.78 with a taper ratio of 0.131. The outer wing is twisted linearly, with a washout angle of 4.3 degrees at the wing tip.

The aerodynamic forces and moments were measured using a six-axis balance. The wind speed was

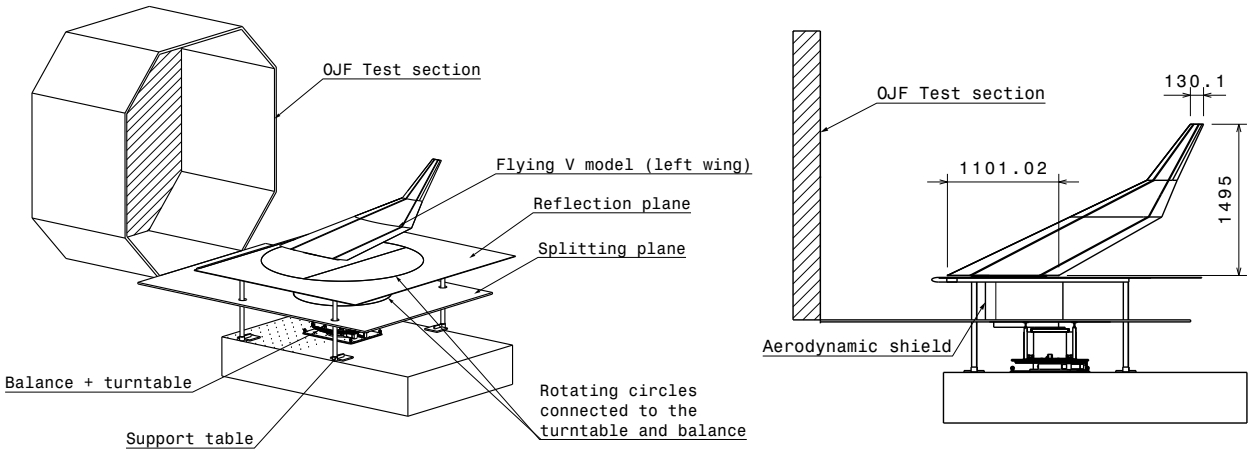

Fig. 2 Drawing of the experimental setup in TU Delft Open Jet Facility (OJF). Isometric view (left) and side view with basic dimensions (right). All dimensions in millimeters.

automatically controlled by the wind tunnel through the fan rotational speed and measured at the beginning of the test section by means of a pitot tube. The control surfaces deflections were measured using calibrated potentiometers connected to the actuators' shafts. The angle of attack was set by a turntable connected to the wing. As no other angle-of-attack measurements were available, the commanded position was assumed to equal the actual angle-of-attack of the wing. A summary of the instrumentation used for the test and its basic characteristics are shown in Table 1, where the values for the standard deviation of the measurements come from the calibration of the balance and the potentiometers. No calibration data is available for the wind speed measurements or the angle of attack from the turntable.

Table 1 Available instrumentation for the wind tunnel tests

Measurement	Sample rate	Std. deviation	Remarks
Force (balance X-axis) F_{Xb} [N]	2 kHz	0.02	Average over 10 s.
Force (balance Y-axis) F_{Yb} [N]	2 kHz	0.05	Average over 10 s.
Force (balance Z-axis) F_{Zb} [N]	2 kHz	0.05	Average over 10 s.
Moment (balance X-axis) M_{Xb} [N·m]	2 kHz	0.01	Average over 10 s.
Moment (balance Y-axis) M_{Yb} [N·m]	2 kHz	0.01	Average over 10 s.
Moment (balance Z-axis) M_{Zb} [N·m]	2 kHz	0.07	Average over 10 s.
Wind speed V [m/s]	2 kHz	-	Average over 10 s.
Angle of attack α [deg]	-	-	Set by the turntable
Control surface deflections δ_i [deg]	100 Hz	0.5	Average over 10 s.

III. Parameter Estimation

As discussed in the introduction, the aerodynamic model identification of the Flying V presents two main challenges: a *model structure determination*, in which the dependencies of each aerodynamic coefficient with the independent variables must be found, and a *parameter estimation*, where the values of the set of parameters that best describe the system of interest are estimated. The latter is treated in this section, since the model structure determination makes use of the parameter estimation to find a suitable model structure.

In the wind tunnel experiment, both the dependent and independent variables can be measured directly. Therefore, the estimation can be done using an equation-error method. It is assumed that the measurements of the dependent variables (the aerodynamic forces and moments) are corrupted with noise, and that the

independent variables (airspeed, angle of attack, and control surface deflections) are known without error. Each observation can be described by the measurement equation:

$$\mathbf{z}(i) = \mathbf{y}(i) + \mathbf{v}(i), \qquad i = 1, 2, ..., N \tag{1}$$

where $\mathbf{z}$ is the measurement vector, $\mathbf{y}$ is the true value of the dependent variable, and $\mathbf{v}$ is a vector of measurement errors. These three vectors have dimensions of $[N \times 1]$, where N is the number of samples. The dependent variables are the dimensionless aerodynamic coefficients of forces and moments in body axes. The forces and moments are assumed to depend on a state vector which includes the independent variables (also called regressors) that are varied during the tests.

$$\begin{aligned} C_X(\mathbf{x}) &= \frac{X(\mathbf{x})}{qS}, \quad C_l(\mathbf{x}) = \frac{L(\mathbf{x})}{qSb} \\ C_Y(\mathbf{x}) &= \frac{Y(\mathbf{x})}{qS}, \quad C_m(\mathbf{x}) = \frac{M(\mathbf{x})}{qS\bar{c}} \\ C_Z(\mathbf{x}) &= \frac{Z(\mathbf{x})}{qS}, \quad C_n(\mathbf{x}) = \frac{N(\mathbf{x})}{qSb} \end{aligned} \tag{2}$$

The state vector is given by: $\mathbf{x} = [\alpha, \hat{V}, \delta_1, \delta_2, \delta_3]$, where $\hat{V}$ is the wind speed normalized with a reference speed in order to avoid large differences in the magnitudes of the regressors.

A. Model Postulation

An estimator is a function that finds the values of the parameter vector $\hat{\boldsymbol{\theta}}$ that best describes the observed model behavior, that is, the fitting of the model output to the measured data when the same inputs are applied. The estimator to construct is limited by the available information about the parameters and noise statistics. Since the probability density of the parameters and the noise is unknown, the model is formulated according to the least squares assumptions: $\boldsymbol{\theta}$ is considered to be a vector of unknown constant parameters, and $\mathbf{v}$ is considered to be a random vector of measurement noise. With these assumptions, the least squares regression finds the best estimate in a single shot for a given model structure. It is assumed that the dependent variable, that is, each aerodynamic coefficient, can be written as:

$$\mathbf{y} = \theta_0 + \sum_{j=1}^{n} \theta_j \mathbf{p}_j, \tag{3}$$

or:

$$\mathbf{y} = X \cdot \boldsymbol{\theta}; \qquad X = [1, \mathbf{p}_1, \mathbf{p}_2, ..., \mathbf{p}_n], \tag{4}$$

where $\mathbf{p}_j$ are functions of the independent variables, X is the regression matrix, and $\boldsymbol{\theta}$ is the parameter vector to be estimated. Each aerodynamic coefficient C_a is modeled as:

$$C_a = C_{a0} + \sum_{n=1}^{n_\alpha} C_{\alpha^n} \cdot \alpha^n + \sum_{n=1}^{n_{\delta_1}} C_{\delta_1^n} \cdot \delta_1^n + \sum_{n=1}^{n_{\delta_2}} C_{\delta_2^n} \cdot \delta_2^n + \sum_{n=1}^{n_{\delta_3}} C_{\delta_3^n} \cdot \delta_3^n + \sum_{n=1}^{n_{\hat{V}}} C_{\hat{V}^n} \cdot \hat{V}^n + [\text{coupling terms}] \tag{5}$$

An underlying assumption with this approach is that the aerodynamic forces and moments can be approximated arbitrarily close by a multivariate Taylor series in the independent variables. Because of this, they become inaccurate far from zero, especially for high angles-of-attack as it is confirmed experimentally in Sec. V).

Because of the known shortcomings of polynomial models, spline models are developed such that larger regions of the flight envelope can be covered with a single global model. The simple "+ basis function" is used for this spline implementation because of its simplicity and its reported adequacy [7, 8, 15]. Only spline terms in the angle of attack are included. Each aerodynamic coefficient $C_a = C_X, C_Y, ...$ is modeled

as follows:

$$
\begin{aligned}
C_a &= C_{a0} + C_{a\alpha}(\alpha) + C_{a\delta_1}(\alpha,\delta_1) + C_{a\delta_2}(\alpha,\delta_2) + C_{a\delta_3}(\alpha,\delta_3) + C_{a\hat{V}}(\alpha,\hat{V}) + [\text{coupling terms}];\\
C_{a\alpha}(\alpha) &= \sum_{n=0}^{n_\alpha} C_{\alpha^n}\cdot\alpha^n + \sum_{l=1}^{k} C_{\alpha_l^d}(\alpha-\alpha_l)_+^d;\\
C_{a\delta_i}(\alpha,\delta_i) &= \sum_{n=0}^{n_{\delta_i}} C_{\delta_i^n}\cdot\delta_i^n + \sum_{l=1}^{k} C_{\delta_i\alpha_l}\cdot\delta_i\cdot(\alpha-\alpha_l)_+^0, \quad i=1,2,3;\\
C_{a\hat{V}}(\alpha,\hat{V}) &= \sum_{n=0}^{n_{\hat{V}}} C_{\hat{V}^n}\cdot\hat{V}^n + \sum_{l=1}^{k} C_{\hat{V}\alpha_l}\cdot\hat{V}\cdot(\alpha-\alpha_l)_+^0.
\end{aligned}
\tag{6}
$$

where $n_\bullet$ is the maximum order considered for each of the regressors, k is the number of knots, and $(\alpha-\alpha_l)_+^d$ is defined as:

$$
(\alpha-\alpha_l)_+^d = \begin{cases} 0 & \alpha < \alpha_l \\ (\alpha-\alpha_l)^d & \alpha > \alpha_l \end{cases}
\tag{7}
$$

A knot of order d has C^{d-1} continuity; that is, the model has $d-1$ continuous derivatives at that knot location. The knots are placed every three degrees in the angle-of-attack range to model local nonlinearities. To implement the spline model, each knot is modeled as a different regressor, and, as such, has its own associated coefficient $\hat{\theta}_j$.

Coupling terms between the angle-of-attack and the other independent variables are included in the spline terms $C_{a\delta_i}$ and $C_{a\hat{V}}$. However, because of the high collinearity among regressors with knots in close locations, only a few of these spline couplings are included. Usual polynomial coupling terms such as $(\alpha^d\cdot\delta_i^e)$ are also allowed in the model in order to improve the fitting without such large penalties in the parameter variances.

B. Parameter Estimation Method

To find the parameter vector estimate $\hat{\boldsymbol{\theta}}$, Ordinary Least Squares (OLS) is used. It is the simplest form of least-squares regression, and has been successfully applied in the past for aerodynamic model identification [11, 21, 22]. To derive the OLS estimator, it is assumed that 1) the independent variables are not contaminated with noise, 2) the dependent variable is contaminated with uniformly distributed noise (white noise), and 3) the residuals are uncorrelated with the independent variables. If these assumptions hold, the Gauss-Markov Theorem states that the OLS estimator is the BLUE [23] (Best Linear Unbiased Estimator, where "best" implies minimum variance) and it can also be shown that the estimates are consistent and efficient[24].

The errors (residuals) are assumed to be zero-mean and to have a constant covariance:

$$
\mathrm{E}[\mathbf{v}] = 0, \quad \mathrm{E}[\mathbf{v}\mathbf{v}^T] = \sigma^2 I
\tag{8}
$$

The well-known cost function for OLS and the parameter vector estimate that minimizes it are:

$$
J = \frac{1}{2}[\mathbf{z}-X\boldsymbol{\theta}]^T[\mathbf{z}-X\boldsymbol{\theta}] = \frac{1}{2}[\mathbf{v}^T\mathbf{v}], \qquad \hat{\boldsymbol{\theta}} = (X^TX)^{-1}X^T\mathbf{z}
\tag{9}
$$

In order to derive confidence bounds for the parameters and the model output the noise is assumed to be given by a normal distribution, an assumption which is checked afterwards. The measurement error variance σ^2 is unknown and needs to be estimated from the measured data. An unbiased but model-dependent estimate for σ^2 can be found based on the residuals:

$$
\hat{\sigma}^2 = \frac{\sum_{i=1}^{N}[z(i)-\hat{\mathbf{y}}(i)]^2}{N-n},
\tag{10}
$$

which only gives adequate results if the model structure is adequate [14].

IV. Model Structure Determination

Up to this point, it has been assumed that the model structure, that is, the functions $\mathbf{p}_j$ that define the regression matrix, is known. However, they also need to be determined since the result of the estimation differs greatly from one model structure to another. An adequate model structure is such that 1) sufficiently fits the data, 2) allows for a successful parameter estimation, and 3) has good prediction capabilities[22]. The number of parameters in the model should be kept as small as possible since large numbers of regressors lead to overparameterization issues. The residual between the model and the measurements always decreases with the addition of regressors, so the model fit to the data always improves. However, noise and outliers eventually enter the model, which degrades its prediction capabilities. This section presents the method to determine the model structure.

A. Modified Stepwise Regression

The core of the model structure determination routine is the so-called Modified Stepwise Regression as it is presented by Klein et al. [25]. Usually, stepwise regression relies only on statistical metrics to select which regressors are to enter or leave the model. However, experience shows that this leads to a large number of terms in the model, which increases overfitting and worsens the prediction capabilities. Therefore, a user-guided selection is done, and priority is given to regressors that carry some useful meaning (e.g. $C_{m\alpha}$) and to regressors with low degrees (that is, α^2 before α^7). The model selection process starts by including the bias in the model, and then the following steps are repeated at every iteration:

1) *Forward selection:* correlations of the regressors with the dependent variable and their partial F-statistic are calculated to select a significant regressor to enter the model.
2) *Backward elimination:* significance of all regressors is evaluated by means of a newly calculated F-statistic after the addition of the last regressor. If some regressor's significance is below a certain threshold, it is removed from the model to keep it as simple as possible.
3) *Definition of new dependent variable:* a new dependent variable is calculated subtracting the current model output to focus on the characteristics that have not been modeled yet.
4) *Model analysis:* at the end of each iteration, a model analysis is performed in order to evaluate each model's performance and select the best one. This analysis includes the calculation of covariances, correlations, collinearity among regressors, and residual statistics.

The stepwise regression routine stops when no more regressors are available or when the remaining regressors are not significant enough to be included. In the specific case of the spline models, since each knot location and degree is modeled as an independent regressor, this method finds the most significant knot locations and the most significant degrees. These are the only ones present in the final models.

B. Regressors Orthogonalization

Since only five independent variables can be used to construct the regressors (wind speed, angle of attack, and the deflections of the three control surfaces), the pool of regressors needs to be defined with combinations of these regressors up to different orders, such as $\alpha^p, V^p, \alpha^p \cdot \delta_i^q, \ldots$. With such a model, collinearity problems arise during the model structure determination routine, leading to the ill-conditioning of $(X^T X)^{-1}$, which worsens the accuracy of the estimation. To solve these collinearity problems, an orthogonal pool of regressors is created using the method presented by Morelli and DeLoach, which is based on a Gram-Schmidt orthogonalization procedure and allows for an easy de-orthogonalization [26]. Starting from the bias term ($\mathbf{p}_0 = \boldsymbol{\xi}_0 = \mathbf{1}$), all regressors are successively orthogonalized with respect to the previous ones. The j^{th} orthogonal function is given by:

$$\boldsymbol{\xi}_j = \mathbf{p}_j - \sum_{k=0}^{j-1} \gamma_{kj} \boldsymbol{\xi}_k; \quad \gamma_{kj} = \frac{\boldsymbol{\xi}_k^T \mathbf{p}_j}{\boldsymbol{\xi}_k^T \boldsymbol{\xi}_k}, \qquad j = 1, 2, ..., n \tag{11}$$

By rearranging the indices γ_{kj} in a matrix G, one can express the orthogonal regressors (now the columns of the regression matrix X, in order to keep the same notation) in terms of the original regressors (columns of X_0) as:

$$X = X_0 \cdot G^{-1}; \qquad X = [\boldsymbol{\xi}_0, \boldsymbol{\xi}_1, ..., \boldsymbol{\xi}_n], \quad X_0 = [\mathbf{p}_0, \mathbf{p}_1, ..., \mathbf{p}_n] \tag{12}$$

C. Metrics and Variables for Model Selection

The decision for the model selection is based on several statistical metrics instead of relying on a single one. The same methodology for model selection is applied for both the orthogonal polynomial models (Eq.5) and the non-orthogonal spline models (Eq. 6).The paragraphs below discuss the various statistical metrics that are used in the present work for model selection.

The overall fit to the data is assessed with the relative Root Mean Square error ($\mathrm{RMS_{rel}}$), normalized with the difference between the maximum and minimum values of the variable being estimated:

$$\mathrm{RMS_{rel}}(\mathbf{v}) = \frac{\sqrt{\frac{1}{N}\sum_{i=1}^{N}\mathbf{v}(i)^2}}{|\max(\mathbf{z}) - \min(\mathbf{z})|}\,. \tag{13}$$

A relative RMS value of 1 indicates that the error is of the order of magnitude of the range of the dependent variable. Morelli suggests in Ref. [11] to use the Predicted Square Error as an useful metric for model selection, based on the work of Barron [27]. It measures the model fit through the Mean Square Error (the square of the RMS), and includes an overfit penalty which increases linearly with the number of regressors in the model n:

$$\mathrm{PSE} = \underbrace{\frac{1}{N}\sum_{i=1}^{N}\mathbf{v}(i)^2}_{\text{Mean Square Error}} + \underbrace{\hat{\sigma}^2_{\max}\frac{n}{N}}_{\text{Overfit penalty}} \tag{14}$$

The value of $\hat{\sigma}_{max}$ needs to be estimated from the data. Barron proposes in the same reference an estimate which is independent of the model structure:

$$\hat{\sigma}^2_{\max} = \frac{1}{N}\sum_{i=1}^{N}(\mathbf{z}(i) - \bar{z})^2 \tag{15}$$

The well-known coefficient of determination R^2 is also used, which measures the proportion of the variation explained by the terms other than the bias [28] :

$$R^2 = 1 - \frac{\sum_{i=1}^{N}\mathbf{v}(i)^2}{\sum_{i=1}^{N}(\mathbf{z}(i) - \bar{z})^2} \tag{16}$$

Finally, the overall statistical significance of the regression is measured with the F-statistic:

$$F = \frac{N-n}{n-1}\cdot\frac{R^2}{1-R^2} \tag{17}$$

High values of both the F-statistic and the coefficient of determination are indicators of good model quality.

The parameter variances and covariances are also used to decide on the model selection. They can be regarded as a measure of the sensitivity of the estimates to the noise, so they should be kept as small as possible. Finally, a residual analysis for each model is performed to guarantee that the assumptions made for the OLS estimation are consistent.

V. Results

In the results of this study are presented. First, the dataset is described in subsection A. Subsequently, the polynomial model and spline model are presented in subsections B and C, respectively. In subsection D, the performance of these two models is compared.

A. Data Coverage

The designed test matrix is constrained by the maximum speed achievable by the wind tunnel ($\sim$ 30 m/s) and the balance load limits stemming from the large size of the model. A fine coverage in the $\alpha - \delta_i$ planes is achieved, and a coarser one for the airspeed, which is limited by the available test time and the wind tunnel operation. In a large number of runs the test conditions are randomized in order to minimize the effect of systematic bias errors[29].

The data is split into an estimation and a validation dataset. The validation dataset is defined using roughly half of the runs from the wind tunnel test, and the model from the estimation dataset is tested against it to minimize overparameterization issues. The region of validity of the model is defined by the convex hull of the estimation dataset, which is the smallest convex set that contains the data. Outside this region no measurements are taken, so one should not extrapolate outside it since new, unmodeled effects might occur. Figure 3 shows two cuts of the estimation convex hull, along with the points from both datasets. It can be seen that all validation data points are inside the estimation convex hull so the estimated models are valid inside the validation dataset as well.

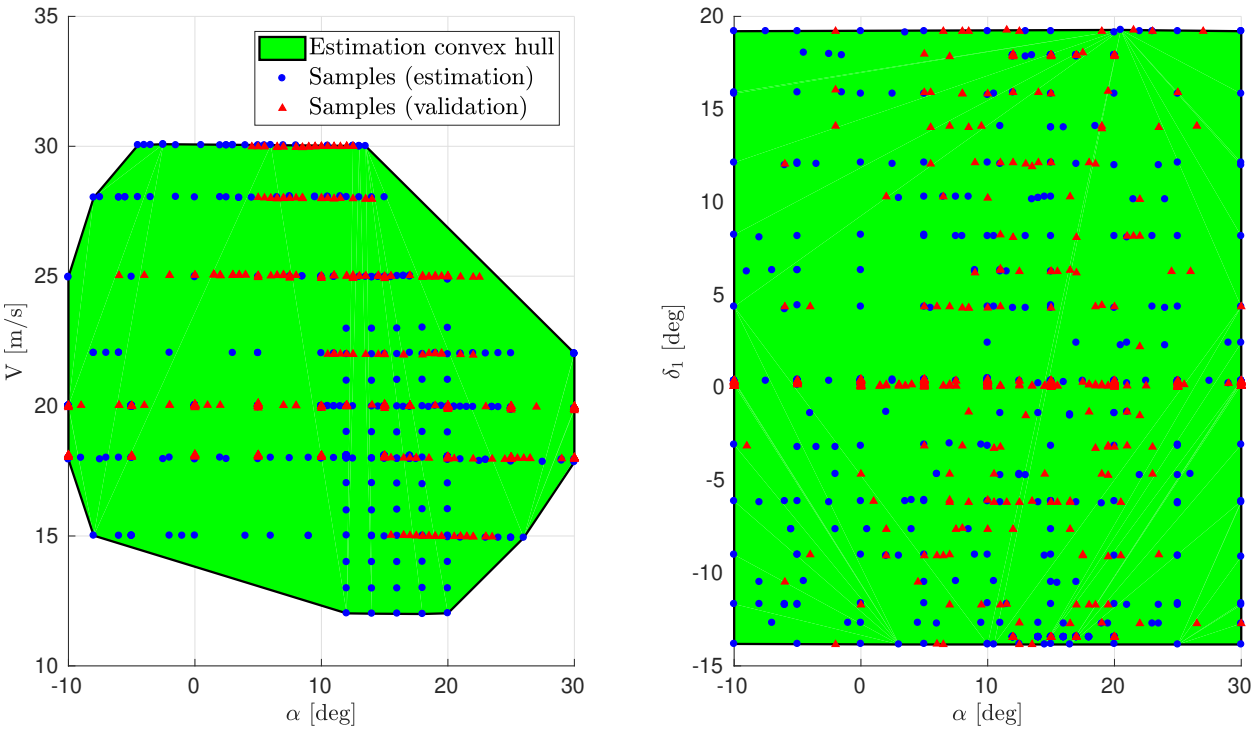

Fig. 3 Convex hull of the estimation dataset ($\alpha - V$ and $\alpha - \delta_1$ cuts)

B. Polynomial Model Estimation

Different pools of regressors were used for the different coefficients, and the order of the regressors for the orthogonalization is carefully chosen for each of them, by applying a previous stepwise regression to the data without orthogonalizing the regressors. The most correlated regressors found with the stepwise regression routine are good candidates to be the first ones to be orthogonalized to keep the model as compact as possible. The maximum degree of the regressors is driven by the value of the parameter variances, which increases with higher order regressors.

Once the model structure has been determined and fixed for each coefficient, the de-orthogonalization is done in order to recover the dependencies in terms of the original independent variables. The final aerodynamic model structure for the longitudinal coefficients after de-orthogonalizating is shown below, and the values of the parameters including their confidence bounds (confidence level $\alpha = 0.05 \rightarrow 95\%$) are presented in Tables 2, 3, and 4.

$$C_X = C_{X0} + C_{X\alpha}\cdot\alpha + C_{X\alpha^2}\cdot\alpha^2 + C_{X\alpha^3}\cdot\alpha^3 + C_{X\alpha^4}\cdot\alpha^4 + C_{X\delta_1}\cdot\delta_1 + \\ + C_{X\delta_2}\cdot\delta_2 + C_{X\delta_3}\cdot\delta_3 + C_{X\delta_1^2}\cdot\delta_1^2 + C_{X\delta_2^2}\cdot\delta_2^2 + C_{X\hat{V}}\cdot\hat{V} + C_{X\hat{V}^2}\cdot\hat{V}^2 \tag{18}$$

$$C_Z = C_{Z0} + C_{Z\alpha}\cdot\alpha + C_{Z\alpha^2}\cdot\alpha^2 + C_{Z\alpha^3}\cdot\alpha^3 + C_{Z\delta_1}\cdot\delta_1 + C_{Z\delta_2}\cdot\delta_2 + \\ + C_{Z\delta_3}\cdot\delta_3 + C_{Z\hat{V}}\cdot\hat{V} \tag{19}$$

$$C_m = C_{m0} + C_{m\alpha}\cdot\alpha + C_{m\alpha^2}\cdot\alpha^2 + C_{m\alpha^3}\cdot\alpha^3 + C_{m\alpha^4}\cdot\alpha^4 + C_{m\delta_1}\cdot\delta_1 + \\ + C_{m\delta_2}\cdot\delta_2 + C_{m\delta_3}\cdot\delta_3 + C_{m\delta_1^2}\cdot\delta_1^2 + C_{m\delta_2^2}\cdot\delta_2^2 + C_{m\delta_1\delta_2}\cdot(\delta_1\cdot\delta_2) + \\ + C_{m\alpha\delta_1^2}\cdot(\alpha\cdot\delta_1^2) + C_{m\alpha\delta_2^2}\cdot(\alpha\cdot\delta_2^2) + C_{m\alpha^2\delta_1}\cdot(\alpha^2\cdot\delta_1) + \\ + C_{m\alpha^2\cdot\delta_2}\cdot(\alpha^2\cdot\delta_2) + C_{m\hat{V}}\cdot\hat{V} + C_{m\hat{V}\delta_1}\cdot(\hat{V}\cdot\delta_1) \tag{20}$$

Table 2 C_X polynomial model coefficients

Parameter	Value ± confidence bounds
C_{X0}	$-2.029\cdot10^{-2} \pm (8.556\cdot10^{-4})$
$C_{X\alpha}$	$4.288\cdot10^{-2} \pm (2.662\cdot10^{-3})$
$C_{X\alpha^2}$	$9.792\cdot10^{-1} \pm (2.227\cdot10^{-3})$
$C_{X\alpha^3}$	$-1.617\cdot10^{-2} \pm (7.731\cdot10^{-2})$
$C_{X\alpha^4}$	$-1.519 \pm (1.639\cdot10^{-1})$
$C_{X\delta_1}$	$1.060\cdot10^{-2} \pm (6.473\cdot10^{-4})$
$C_{X\delta_2}$	$7.162\cdot10^{-3} \pm (1.193\cdot10^{-3})$
$C_{X\delta_3}$	$6.893\cdot10^{-4} \pm (1.294\cdot10^{-4})$
$C_{X\delta_1^2}$	$-3.928\cdot10^{-2} \pm (2.945\cdot10^{-3})$
$C_{X\delta_2^2}$	$-2.047\cdot10^{-2} \pm (4.027\cdot10^{-3})$
$C_{X\hat{V}}$	$8.727\cdot10^{-3} \pm (3.708\cdot10^{-4})$
$C_{X\hat{V}^2}$	$-3.075\cdot10^{-3} \pm (3.318\cdot10^{-4})$

Table 3 C_Z polynomial model coefficients

Parameter	Value ± confidence bounds
C_{Z0}	$-7.583\cdot10^{-2} \pm (5.310\cdot10^{-3})$
$C_{Z\alpha}$	$-1.943 \pm (1.564\cdot10^{-2})$
$C_{Z\alpha^2}$	$-1.026 \pm (1.288\cdot10^{-1})$
$C_{Z\alpha^3}$	$1.866 \pm (2.343\cdot10^{-1})$
$C_{Z\delta_1}$	$-1.353\cdot10^{-1} \pm (5.239\cdot10^{-3})$
$C_{Z\delta_2}$	$-8.462\cdot10^{-2} \pm (6.915\cdot10^{-3})$
$C_{Z\delta_3}$	$-5.012\cdot10^{-3} \pm (7.192\cdot10^{-4})$
$C_{Z\hat{V}}$	$3.747\cdot10^{-2} \pm (7.564\cdot10^{-3})$

Table 4 C_m polynomial model coefficients

Parameter	Value ± confidence bounds
C_{m0}	$3.361 \cdot 10^{-2} \pm (2.788 \cdot 10^{-3})$
$C_{m\alpha}$	$-1.069 \cdot 10^{-1} \pm (5.938 \cdot 10^{-3})$
$C_{m\alpha^2}$	$-2.883 \cdot 10^{-1} \pm (6.506 \cdot 10^{-3})$
$C_{m\alpha^3}$	$-8.492 \cdot 10^{-1} \pm (1.839 \cdot 10^{-1})$
$C_{m\alpha^4}$	$4.088 \quad \pm (3.731 \cdot 10^{-1})$
$C_{m\delta_1}$	$-1.482 \cdot 10^{-1} \pm (2.139 \cdot 10^{-3})$
$C_{m\delta_2}$	$-1.092 \cdot 10^{-1} \pm (1.433 \cdot 10^{-3})$
$C_{m\delta_3}$	$-6.377 \cdot 10^{-2} \pm (3.673 \cdot 10^{-3})$
$C_{m\delta_1^2}$	$1.039 \cdot 10^{-1} \pm (1.608 \cdot 10^{-2})$
$C_{m\delta_2^2}$	$3.400 \cdot 10^{-2} \pm (1.127 \cdot 10^{-2})$
$C_{m\delta_1\delta_2}$	$6.120 \cdot 10^{-2} \pm (1.248 \cdot 10^{-2})$
$C_{m\alpha\delta_1^2}$	$-5.459 \cdot 10^{-1} \pm (5.529 \cdot 10^{-2})$
$C_{m\alpha\delta_2^2}$	$-2.125 \cdot 10^{-1} \pm (4.211 \cdot 10^{-2})$
$C_{m\alpha^2\delta_1}$	$1.729 \cdot 10^{-1} \pm (4.450 \cdot 10^{-2})$
$C_{m\alpha^2\delta_2}$	$1.775 \cdot 10^{-1} \pm (3.443 \cdot 10^{-2})$
$C_{m\hat{V}}$	$8.071 \cdot 10^{-3} \pm (3.664 \cdot 10^{-3})$
$C_{m\hat{V}\delta_1}$	$2.203 \cdot 10^{-3} \pm (4.492 \cdot 10^{-4})$

C. Spline Model Estimation

Splines are used to cope with the inability of polynomials to model the complete measured range of angles of attack. Once the spline estimation is completed, Multivariate Orthogonal Functions are introduced using the already estimated model structure to sort the regressors orthogonalization in order of significance. Although the orthogonal spline models exhibit smaller variances and no collinearity, the non-orthogonal models show a closer similarity between the estimation and validation dataset (fit, whiteness, and normality) keeping the variances sufficiently small, so they are accepted as better models and are the only ones shown here.

Since non-orthogonal regressors are used for the spline estimation, the collinearity among them needs to be addressed to validate the estimation. The equations for the longitudinal coefficients are presented below, and the value of the coefficients in Tables 5, 6, and 7.

$$\begin{aligned} C_X = {} & C_{X\alpha} \cdot \alpha + C_{X\alpha^2} \cdot \alpha^2 + C_{X\delta_1} \cdot \delta_1 + C_{X\delta_2} \cdot \delta_2 + C_{X\delta_3} \cdot \delta_3 + \\ & + C_{X\hat{V}} \cdot \hat{V} + C_{X\hat{V}^2} \cdot \hat{V}^2 + C_{X\alpha^2_{20}} \cdot (\alpha - \alpha_{20^\circ})^2_+ \end{aligned} \tag{21}$$

$$\begin{aligned} C_Z = {} & C_{Z0} + C_{Z\alpha} \cdot \alpha + C_{Z\delta_1} \cdot \delta_1 + C_{Z\delta_2} \cdot \delta_2 + C_{Z\delta_3} \cdot \delta_3 + C_{Z\hat{V}} \cdot \hat{V} + \\ & + C_{Z\alpha^1_{11}} \cdot (\alpha - \alpha_{11^\circ})_+ + C_{Z\alpha^2_{17}} \cdot (\alpha - \alpha_{17^\circ})^2_+ \end{aligned} \tag{22}$$

Table 5 C_X spline model coefficients

Parameter	Value ± confidence bounds
$C_{X\alpha}$	$4.590 \cdot 10^{-2} \pm (3.49 \cdot 10^{-3})$
$C_{X\alpha^2}$	$7.484 \cdot 10^{-1} \pm (1.29 \cdot 10^{-2})$
$C_{X\delta_1}$	$8.103 \cdot 10^{-3} \pm (1.86 \cdot 10^{-3})$
$C_{X\delta_2}$	$6.727 \cdot 10^{-3} \pm (1.50 \cdot 10^{-3})$
$C_{X\delta_3}$	$7.246 \cdot 10^{-3} \pm (1.69 \cdot 10^{-3})$
$C_{X\hat{V}}$	$2.276 \cdot 10^{-2} \pm (2.03 \cdot 10^{-3})$
$C_{X\hat{V}^2}$	$1.011 \cdot 10^{-2} \pm (1.66 \cdot 10^{-3})$
$C_{X\alpha_{20}^2}$	$1.921 \pm (7.88 \cdot 10^{-2})$

Table 6 C_Z spline model coefficients

Parameter	Value ± confidence bounds
C_{Z0}	$8.554 \cdot 10^{-2} \pm (5.53 \cdot 10^{-3})$
$C_{Z\alpha}$	$1.869 \pm (8.62 \cdot 10^{-3})$
$C_{Z\delta_1}$	$1.284 \cdot 10^{-1} \pm (5.31 \cdot 10^{-3})$
$C_{Z\delta_2}$	$7.395 \cdot 10^{-2} \pm (3.70 \cdot 10^{-3})$
$C_{Z\delta_3}$	$4.303 \cdot 10^{-2} \pm (4.77 \cdot 10^{-3})$
$C_{Z\hat{V}}$	$3.126 \cdot 10^{-2} \pm (5.14 \cdot 10^{-3})$
$C_{Z\alpha_{11}}$	$5.559 \cdot 10^{-1} \pm (2.44 \cdot 10^{-2})$
$C_{Z\alpha_{17}^2}$	$3.176 \pm (1.22 \cdot 10^{-1})$

$$
\begin{aligned}
C_m = {} & C_{m0} + C_{m\alpha} \cdot \alpha + C_{m\delta_1} \cdot \delta_1 + C_{m\delta_2} \cdot \delta_2 + C_{m\delta_3} \cdot \delta_3 + C_{m\hat{V}} \cdot \hat{V} + \\
& + C_{m\delta_1\delta_2} \cdot \delta_1\delta_2 + C_{m\delta_2\delta_3} \cdot \delta_2 \cdot \delta_3 + C_{m\delta_2^2} \cdot \delta_2^2 + C_{m\alpha\delta_2^2} \cdot \alpha \cdot \delta_2^2 + \\
& + C_{m\alpha_{20}^2} \cdot (\alpha - \alpha_{20^\circ})_+^2 + C_{m\alpha_{26}^2} \cdot (\alpha - \alpha_{26^\circ})_+^2 + \\
& + C_{m\delta_1\alpha_2} \cdot \delta_1 \cdot (\alpha - \alpha_{2^\circ})_+^0 + C_{m\delta_1\alpha_{23}} \cdot \delta_1 \cdot (\alpha - \alpha_{23^\circ})_+^0 + \\
& + C_{m\delta_2\alpha_{-7}} \cdot \delta_2 \cdot (\alpha - \alpha_{-7^\circ})_+^0 + C_{m\delta_2\alpha_{14}} \cdot \delta_2 \cdot (\alpha - \alpha_{14^\circ})_+^0 + \\
& + C_{m\delta_3\alpha_8} \cdot \delta_3 \cdot (\alpha - \alpha_{8^\circ})_+^0
\end{aligned}
\tag{23}
$$

D. Model Comparison and Discussion

The characteristics of the estimated models are presented in Table 8, which includes the errors from the estimation and validation data sets as well as the all the other metrics used for the selection. All models are considered adequate in terms of model fit. Their high values of the F-statistic and the coefficient of determination are indicators of a good model. The maximum absolute relative residuals for the models are presented in Table 9, where it can be seen that the pitching moment coefficient has relatively large values for the maximum residuals (especially in the validation dataset), which could indicate some model deficiencies or the presence of outliers in the data due to measurement errors. The spline models show a better fitting except for the forward force coefficient C_X. It is also clear that the values of the studied metrics for the estimation and validation datasets match better for the spline models, which suggests a degree of overfitting present in the polynomial ones.

Because of the large range of angles-of-attack that the models aim to approximate ($\alpha \in [-10, 30]^\circ$),

Table 7 C_m spline model coefficients

Parameter	Value ± confidence bounds
C_{m0}	$2.440 \cdot 10^{-2} \pm (3.02 \cdot 10^{-3})$
$C_{m\alpha}$	$1.461 \cdot 10^{-1} \pm (3.41 \cdot 10^{-3})$
$C_{m\delta_1}$	$9.921 \cdot 10^{-2} \pm (7.30 \cdot 10^{-3})$
$C_{m\delta_2}$	$8.065 \cdot 10^{-2} \pm (9.34 \cdot 10^{-3})$
$C_{m\delta_3}$	$7.077 \cdot 10^{-2} \pm (4.85 \cdot 10^{-3})$
$C_{m\hat{V}}$	$1.383 \cdot 10^{-2} \pm (2.81 \cdot 10^{-3})$
$C_{m\delta_2^2}$	$4.729 \cdot 10^{-2} \pm (1.05 \cdot 10^{-2})$
$C_{m\alpha\delta_2^2}$	$2.243 \cdot 10^{-1} \pm (3.75 \cdot 10^{-2})$
$C_{m\delta_1\delta_2}$	$4.165 \cdot 10^{-2} \pm (1.11 \cdot 10^{-2})$
$C_{m\delta_2\delta_3}$	$3.312 \cdot 10^{-2} \pm (1.08 \cdot 10^{-2})$
$C_{m\alpha_{20}^2}$	$5.545 \pm (2.50 \cdot 10^{-1})$
$C_{m\alpha_{26}^2}$	$8.353 \pm (1.48 \cdot 10^{0})$
$C_{m\delta_1\alpha_2}$	$4.691 \cdot 10^{-2} \pm (7.99 \cdot 10^{-3})$
$C_{m\delta_1\alpha_{23}}$	$5.040 \cdot 10^{-2} \pm (9.99 \cdot 10^{-3})$
$C_{m\delta_2\alpha_{-7}}$	$3.080 \cdot 10^{-2} \pm (9.87 \cdot 10^{-3})$
$C_{m\delta_2\alpha_{-7}}$	$3.383 \cdot 10^{-2} \pm (3.87 \cdot 10^{-3})$
$C_{m\delta_3\alpha_8}$	$1.476 \cdot 10^{-2} \pm (5.73 \cdot 10^{-3})$

Table 8 Model characteristics comparison

	RMS_{rel} (est.)	RMS_{rel} (val.)	F-statistic	R^2	PSE	$\hat{\sigma}^2$
C_X polynomial	1.42 %	2.42 %	$2.452 \cdot 10^4$	99.74%	$5.422 \cdot 10^{-5}$	$7.193 \cdot 10^{-6}$
C_X spline	1.99%	1.97%	$1.602 \cdot 10^4$	99.49%	$5.164 \cdot 10^{-5}$	$1.411 \cdot 10^{-5}$
C_Z polynomial	1.42 %	2.38 %	$3.870 \cdot 10^4$	99.70%	$1.881 \cdot 10^{-3}$	$4.236 \cdot 10^{-4}$
C_Z spline	0.75%	0.96%	$1.249 \cdot 10^5$	99.92%	$1.9 \cdot 10^{-3}$	$1.165 \cdot 10^{-4}$
C_m polynomial	2.54 %	7.24 %	$2.516 \cdot 10^3$	98.00%	$9.416 \cdot 10^{-5}$	$4.732 \cdot 10^{-5}$
C_m spline	2.13 %	2.75 %	$2.894 \cdot 10^3$	99.27%	$9.084 \cdot 10^{-5}$	$3.362 \cdot 10^{-5}$

Table 9 Maximum absolute relative residuals

	$\max(\lvert\mathbf{v}_{\text{rel}}\rvert)$ (est.)	$\max(\lvert\mathbf{v}_{\text{rel}}\rvert)$ (val.)
C_X polynomial	4.96%	6.52%
C_X spline	7.11%	6.57%
C_Z polynomial	3.99%	6.20 %
C_Z spline	2.35%	2.92%
C_m polynomial	9.51%	17.60%
C_m spline	9.32%	10.62%

polynomial models fail to explain some of the nonlinearities present in the data. A solution could be to split the data using the value of the angle of attack, which would lead to a piecewise continuous model.

Figure 4 shows two slices from the polynomial and spline model of the pitching moment coefficient, along with the calculated prediction confidence bounds. Normality of the residuals was assumed to calculate these, and normality tests were passed with positive results. The normality assumption is also reinforced since, as can be seen in the figure, most data points fall within the calculated confidence bounds.

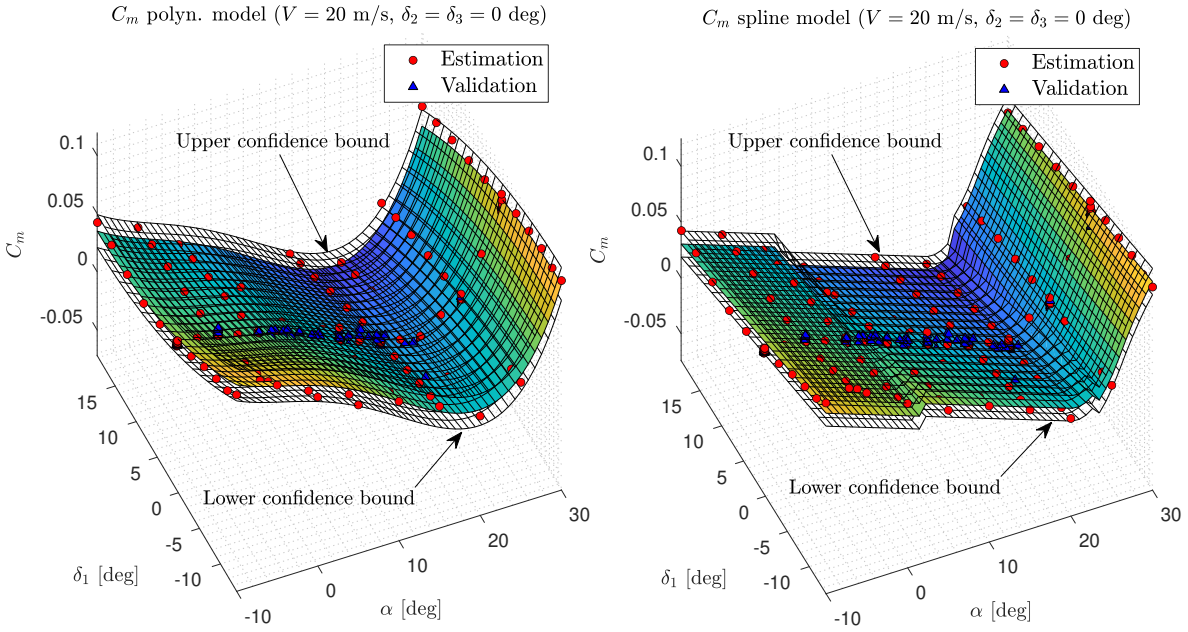

Fig. 4 $\alpha - \delta_1$ slice of the polynomial and spline model for the pitching moment coefficient (colored surface) and prediction confidence bounds (transparent surfaces)

It also is interesting to compare the "local model quality" of the spline models with respect to the polynomial models. That is, how the model quality changes for the different values of the independent variables. In this case, the angle-of-attack is the variable of choice since it is the most significant one. To do so, the validation residuals are split according to their value of the angle-of-attack in segments of 5 degrees, and the relative RMS is calculated for each of these segments. The results are shown in Fig. 5, where it can be seen that the overall local quality of spline models is better, and has a smaller spatial variability than in the polynomial models. It can also be observed that the assumption that the residuals are uncorrelated with the independent variables (angle-of-attack in this case) holds much better for the spline models than for the polynomial ones.

From the fitting point of view, it is clear that spline models are superior. The spline models used in this work are, however, susceptible to collinearity issues. Therefore, it is worth to check a limited set of indicators to ensure that collinearity remain within acceptable bounds. Commonly used indicators are the Variance Inflation Factors (VIF), which measure the collinearity of the j^{th} regressor with respect to the rest of the regressors in the model, and how large the parameter confidence intervals become with respect to those from an orthogonal framework [14] . The condition indices of the Singular Value Decomposition (SVD) of the regression matrix are also used, which measure the ratio between the maximum singular value and the remaining ones. A large value indicates that the associated eigenvalue is very small compared to the largest one, or in other words, that two regressors are almost linearly dependent. The collinearity indicators from the spline model for the pitching moment coefficient are shown in Fig. 6 as an example, since it is the coefficient with the highest levels of collinearity. It can be seen how some of the parameters exhibit reasonably large VIFs, which show that there is room for improvement by using an orthogonal framework. On the other hand, all SVD condition indices are smaller than the limit proposed by Klein for moderate collinearity [21].

Even though all collinearity indicators remain within acceptable bounds *a priori*, a sensitivity analysis is

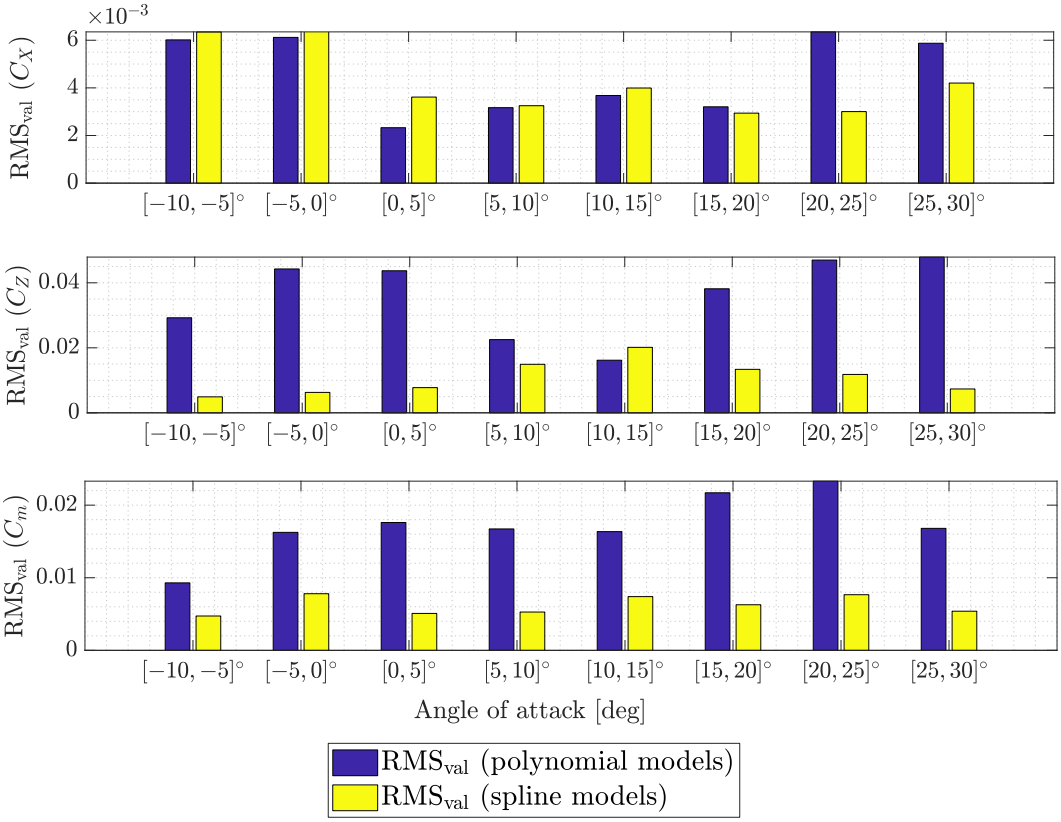

Fig. 5 RMS of five-degree segments in the angle-of-attack range to assess local spatial quality

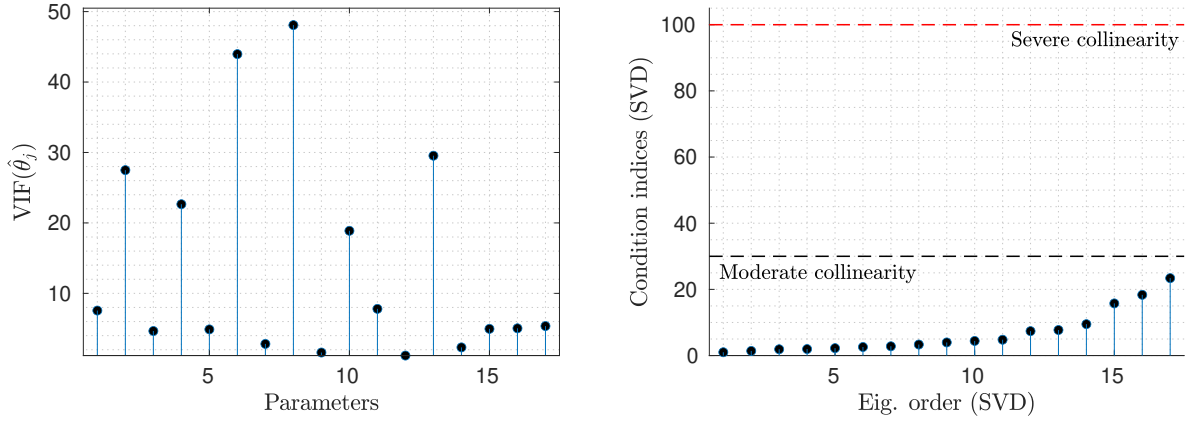

Fig. 6 Variance Inflation Factors and SVD condition indices of the spline model for C_m

conducted on the parameters to see how their change would affect the accuracy of the estimation. To do so, they are perturbed with a value equal to their estimated confidence bounds and the difference in the value of the RMS is recorded. The results are presented in Fig. 7, along with the values for the polynomial model as a comparison with an orthogonal framework.

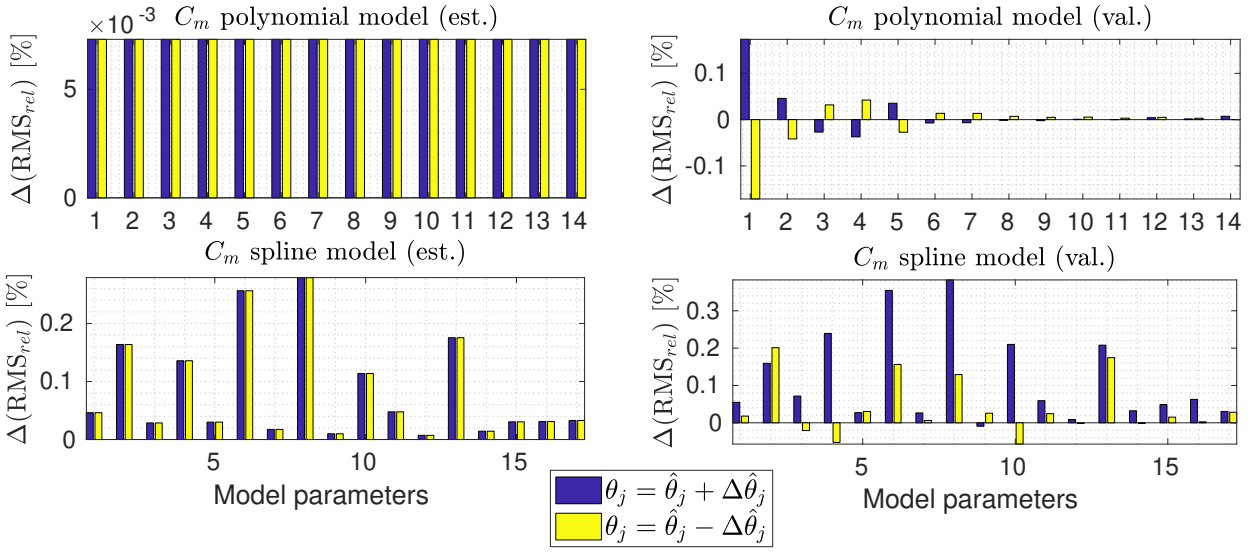

Fig. 7 Parameters sensitivity analysis results for the polynomial and spline model for C_m

The results, as one may expect, are better for the polynomial model, with a maximum increase of $\sim 0.15\%$ of the RMS on the validation dataset, whereas for the spline model the maximum increase is more than double this value. On the other hand, the increase in RMS in the estimation and validation datasets behaves more similarly for the spline model, which is, of course, a desirable outcome, which hints to overfitting of the polynomial models. All in all, the increase of the error from these perturbations is fairly small, and the effect of the present collinearity in the spline models is considered relatively harmless.

VI. Conclusions

A first global aerodynamic model for the Flying V has been estimated from longitudinal, static wind tunnel data from a 4.6% half model. Orthogonal polynomial models and non-orthogonal spline models in the angle-of-attack dimension were used for the model estimation, achieving compact models with tight confidence bounds. They were validated against a partition of the data showing good prediction capabilities and reasonably random, uncorrelated residuals. The polynomial models attained good fitting, with a relative RMS for the validation dataset being around 2.5% for C_X and C_Z, and about 7% for C_m. Spline models outperformed the polynomial models, with a relative RMS below 2% for C_X, 1% for C_Z, and 3% for C_m, which demonstrates their higher approximation capability. The resulting models are analytical functions in the state and control variables, which can be used to determine the aerodynamic forces and moments of the subscale model at any given flight condition inside the models' region of validity.

The polynomial models show some model deficiencies due to the large range of the measured angles of attack. Data partitioning is recommended to keep adequate polynomial models, which would lead to a piecewise continuous aerodynamic model. On the other hand, the proposed spline implementation is rather limited since it only considers splines in the angle-of-attack dimension and no continuity constraints are enforced at the knots. Increasingly complex spline models, such as B-splines or multivariate simplex splines, seem a natural next step to improve the estimation. Since no previous validated data from the Flying V is available, it is uncertain whether the dataset is free of systematic errors and errors in the independent variables, which would introduce a bias in the estimates. Sub-scale flight tests, or further investigations in the wind tunnel are required in order to gain more confidence in the measured data.

Acknowledgments

The authors would like to thank Malcom Brown, Marco Palermo, and Rob Viet for the shared effort in the design, building, and testing of the wind tunnel model and the experimental setup. The authors would also like to acknowledge the help and guidance of the lab technicians Fred Bosch and Victor Horbowiec for helping with the production of the wind tunnel model.

References

[1] Benad, J., "The Flying V - A new aircraft configuration for commercial passenger transport," *Deutscher Luft-und Raumfahrtkongress*, 2015.

[2] Faggiano, F., Vos, R., Baan, M., and van Dijk, R., "Aerodynamic Design of a Flying V Aircraft," *17th AIAA Aviation Technology, Integration and Operations Conference*, American Institute of Aeronautics and Astronautics, 2017.

[3] Vos, R., Geuskens, F. J. J. M. M., and Hoogreef, M. F. M., "A New Structural Design Concept for Blended Wing Body Cabins," *53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference*, American Institute of Aeronautics and Astronautics, 2012.

[4] Rubio Pascual, B., and Vos, R., "The Effect of Engine Location on the Aerodynamic Efficiency of a Flying-V Aircraft," *AIAA Scitech 2020 Forum*, Orlando, FL, 2020. https://doi.org/10.2514/6.2020-1954.

[5] Palermo, M., and Vos, R., "Experimental Aerodynamic Analysis of a 4.6%-Scale Flying-V Subsonic Transport," *AIAA Scitech 2020 Forum*, Orlando, FL, 2020. https://doi.org/10.2514/6.2020-2228.

[6] Klein, V., and Morelli, E. A., *Aircraft System Identification: Theory and Practice*, American Institute of Aeronautics and Astronautics, 2006.

[7] Klein, V., and Batterson, J. G., "Aerodynamic parameters estimated from flight and wind tunnel data," *Journal of Aircraft*, Vol. 23, No. 4, 1986, pp. 306–312. https://doi.org/10.2514/3.45304.

[8] Klein, V.; Batterson, J. G., "Determination of Airplane Model Structure From Flight Data Using Splines and Stepwise Regression," NASA Technical Paper 2126, National Aeronautics and Space Administration, March 1983. https://doi.org/10.1007/978-90-481-9482-7.

[9] Klein, V., Batterson, J., and Smith, P., "On the Determination of Airplane Model Structure From Flight Data," *IFAC Identification and System Parameter Estimation*, Vol. 15, 1982, pp. 1163–1168. https://doi.org/10.1016/s1474-6670(17)63154-9.

[10] Batterson, J., G., and Klein, V., "Partitioning of flight data for aerodynamic modeling of aircraft at high angles of attack," *Journal of Aircraft*, Vol. 26, No. 4, 1989, pp. 334–339. https://doi.org/10.2514/3.45765.

[11] Morelli, E. A., "Global Nonlinear Aerodynamic Model Using Multivariate Orthogonal Functions," *Journal of Aircraft*, Vol. 32, No. 2, 1995, pp. 270–277.

[12] Morelli, E. A., "Global Nonlinear Parametric Modeling with Application To F-16 Aerodynamics," *Proceedings of the 1998 American Control Conference*, Vol. 2, 1998, pp. 997–1001.

[13] Grauer, J. A., and Morelli, E. A., "Generic Global Aerodynamic Model for Aircraft," *Journal of Aircraft*, Vol. 52, No. 1, 2015, pp. 13–20. https://doi.org/10.2514/1.C032888.

[14] Montgomery, D. C., Peck, E. A., and Vining, G. G., *Introduction to Linear Regression Analysis*, 5th ed., Wiley Series in Probability and Statistics, Wiley, 2012.

[15] Smith, P. L., "Curve Fitting and Modeling with Splines using Statistical Variable Selection Techniques," NASA Contractor Report 166034, National Aeronautics and Space Administration, 1982.

[16] Bruce, P., and Kellet, M., "Modelling and identification of non-linear aerodynamic functions using B-splines," *Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering*, Vol. 214, 2000, pp. 27–40.

[17] C. C. de Visser, "Global Nonlinear Model Identification with Multivariate Splines - Application to Aerodynamic Model Identification of the Cessna Citation II," Ph.D. thesis, Delft University of Technology, 2011.

[18] Jategaonkar, R. V., "Flight Vehicle System Identification: A Time Domain Methodology," *Progress in Aeronaunatics and Astronautics*, Vol. 216, 2006.

[19] Wolowicz, C. H., Bowman-Jr, J. S., and Gilbert, W. P., "Similitude requirements and scaling relationships as applied to model testing," NASA Technical Paper 1435, National Aeronautics and Space Administration, 1979.

[20] Yu, W., Hong, V. W., Simao Ferreira, C. J., and van Kuik, G. A. M., "Experimental analysis on the dynamic wake of an actuator disc undergoing transient loads," *Thrust and Drag: Its Prediction and Verification*, Vol. 58, No. 10, 2017. https://doi.org/https://doi.org/10.1007/s00348-017-2432-9.

[21] Klein, V., "Estimation of Aircraft Aerodynamic Parameters From Flight Data," *Progress in Aerospace Sciences*, Vol. 26, No. 1, 1989, pp. 1–77.

[22] Murphy, P. C., "A Methodology for Airplane Parameter Estimation and Confidence Interval Determination in Nonlinear Estimation Problems," NASA Reference Publication 1153, National Aeronautics and Space Administration, apr 1986.

[23] Mithra, S. K., and Rao, C. R., "Conditions for Optimality and Validity of Simple Least Squares Theory," *The Annals of Mathematical Statistics*, Vol. 40, No. 5, 1969, pp. 1617–1624.

[24] Hsia, T. C., *System Identification*, Lexington Books, 1977.

[25] Klein, V., Batterson, J. G., and Murphy, P. C., "Airplane Model Structure Determination from Flight Data," *Journal of Aircraft*, Vol. 20, No. 5, 1983, pp. 469–474.

[26] Morelli, E. A., and Deloach, R., "Wind Tunnel Database Development using Modern Experiment Design and Multivariate Orthogonal Functions," *41st AIAA Aerospace Science Meeting and Exhibit*, American Institute of Aeronautics and Astronautics, Reno, 2003.

[27] Barron, A., "Predicted squared error: a criterion for automatic model selection," *Self-Organizing Methods in Modeling*, Marcel Dekker, New York, 1984.

[28] Klein, V., Batterson, J. G., and Murphy, P. C., "Determination of Airplane Model Structure From Flight Data by Using Stepwise Regression," NASA Technical Paper 1916, National Aeronautics and Space Administration, 1981.

[29] Montgomery, D. C., *Design and Analysis of Experiments*, 7$^{\text{th}}$ ed., John Wiley & Sons, 2009.