



Delft University of Technology

Stall Model Identification of a Cessna Citation II from Flight Test Data Using Orthogonal Model Structure Selection

van Ingen, J.B.; de Visser, C.C.; Pool, D.M.

DOI

[10.2514/6.2021-1725](https://doi.org/10.2514/6.2021-1725)

Publication date

2021

Document Version

Final published version

Published in

AIAA Scitech 2021 Forum

Citation (APA)

van Ingen, J. B., de Visser, C. C., & Pool, D. M. (2021). Stall Model Identification of a Cessna Citation II from Flight Test Data Using Orthogonal Model Structure Selection. In *AIAA Scitech 2021 Forum: 11–15 & 19–21 January 2021, Virtual Event Article AIAA 2021-1725* American Institute of Aeronautics and Astronautics Inc. (AIAA). <https://doi.org/10.2514/6.2021-1725>

Important note

To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.



Stall Model Identification of a Cessna Citation II from Flight Test Data Using Orthogonal Model Structure Selection

Joost B. van Ingen*, Coen C. de Visser†, and Daan M. Pool‡
Delft University of Technology, Delft, The Netherlands

Since 2019, a key element of simulator-based training of airline pilots is stall training. A major and still largely open research question is which level of model fidelity is required for effective training. As part of an effort to answer this question, a model of the quasi-steady stall dynamics of a Cessna Citation II aircraft is identified from flight test data that was specifically collected for this research at an altitude of 5,500 m. To ensure any reductions in elevator and aileron effectiveness during stall were also explicitly measured, the test pilots used additional quasi-random flight test inputs. The considered stall model structure is based on Kirchhoff's theory of flow separation. During identification, the nonlinear and linear parameters of the model are estimated in separate, recursively executed, steps. This separation enables the application of a semi-objective model structure selection method using multivariate orthogonal functions for the aerodynamic coefficients included in the model. This approach shows that stall-related effects should be included in the model equations for lift, drag, and pitch moment. Overall, it is found that the model parameters were consistently estimated from the flight test data and that the model accurately describes the aircraft's stall dynamics in the considered flight condition. The developed methodology is concluded to be well-suited for the direct identification of stall models from flight test data.

Nomenclature

<i>Roman</i>			
A	Matrix of regression variables	U	Theil statistic
A_*	Specific force in $*$ -direction [m/s^2]	V_{TAS}	True airspeed [m/s]
$\mathbf{a}$	Regression variable vector	x	Regression variables
a_1	X-parameter for shape [-]	x_E, y_E, z_E	Cartesian position in F_E [m]
b	Aircraft span [m]	$\dot{x}_E, \dot{y}_E, \dot{z}_E$	Velocity components in F_E [m/s]
C_*	Force/moment coefficient [-]	X	Flow separation point variable [-]
$\bar{c}$	Average chord length [m]	$\mathbf{y}$	Measurement vector
I_*	Angular moment of inertia around $*$ -axis [kgm^2]	$\hat{\mathbf{y}}$	Model output vector
J	Cost function value	<i>Greek</i>	
M	Mach number [-]	α	Angle of attack [rad]
m	Mass [kg]	α_*	X-parameter for scheduling [rad]
N	Amount of time samples	β	Angle of sideslip [rad]
n	Amount of terms in a model structure	$\gamma_{k,j}$	Gram-Schmidt scaling parameter
p	Roll rate [rad/s]	ϵ	Vector of remnant
q	Pitch rate [rad/s]	δ_a	Aileron deflection [rad]
R^2	Coefficient of determination	δ_e	Elevator deflection [rad]
r	Yaw rate [rad/s]	δ_r	Rudder deflection [rad]
S	Wing surface area [m^2]	θ	Pitch angle [rad] or parameter vector
$\mathbf{p}$	Orthogonalized regression variable vector	$\hat{\theta}$	Optimal estimate of parameter vector

*MSc student, Control & Simulation Section, Faculty of Aerospace Engineering, P.O. Box 5058; joostingen@gmail.com.

†Assistant Professor, Control & Simulation Section, Faculty of Aerospace Engineering, P.O. Box 5058, 2600GB Delft, The Netherlands; c.c.devissier@tudelft.nl. Member AIAA.

‡Assistant Professor, Control & Simulation Section, Faculty of Aerospace Engineering, P.O. Box 5058, 2600GB Delft, The Netherlands; d.m.pool@tudelft.nl. Senior Member AIAA.

ρ	Correlation	<i>Subscripts</i>	
σ	Standard deviation	D	Drag
τ_1	X-parameter for lag [s]	L	Lift
τ_2	X-parameter for hysteresis [s]	l	Roll moment
ϕ	Roll angle [rad] <i>or</i> orthogonal parameter vector	m	Pitch moment
$\hat{\phi}$	Optimal estimate of orthogonal parameter vector	n	Yaw moment
		T	Thrust force
ψ	Heading angle [rad]	Y	Lateral force

I. Introduction

Loss of control in-flight currently is the largest category of fatal civil and general aviation accidents [1], and aerodynamic stall is an important contributor [2]. Better training of pilots in recognizing, preventing, and recovering of stalls is seen as an important step for increasing safety [3–5]. From March 2019 onwards, such training will become mandatory for all airline pilots [6]. Due to cost and safety considerations, this stall training will take place in simulators. As a result, there is an increased demand for stall models of civil aviation aircraft.

Currently, aerodynamic stall models are difficult to obtain. Stalls are characterized as highly nonlinear, unsteady, configuration-dependent, and fundamentally unpredictable [7]. Current stall modeling methods generally fall in either of two categories. The first category is based on modeling methods that were originally developed for fighter jets. Starting in the 1980s, these were also applied to civil aviation aircraft. NASA [8, 9] and the EU [10, 11] both have had dedicated research programs based around this “fighter” method. This method relies on extensive wind tunnel testing, is complex and expensive to implement, but can be used for modeling even extreme maneuvers such as spins. The second category is based on Kirchhoff’s theory of flow separation [12]. It involves modeling of an internal flow separation point variable X , which in turn affects lift and possibly other forces or moments. The potential range of application of “Kirchhoff’s method” is smaller – spins cannot be modeled – but it is simpler to implement. Research has shown that the flow separation point parameters can be identified from flight test data. [13–15] Both these approaches have resulted in stall models of varying fidelity. Recently, novel modeling methods have appeared, such as those based on CFD [16] or representative methods based on aircraft geometry and/or configuration. [17, 18] Regardless the used modeling method, a major unanswered question is what level of fidelity is actually required for effective pilot stall training. While pilots are found to be able to notice even small changes in key stall model parameters in some experiments [19], recent pilot-in-the-loop evaluations of stall models of varying fidelity levels found no differences in subjective ratings from a group of experienced pilots [20].

This paper is a preliminary result of an effort at the Faculty of Aerospace Engineering of TU Delft, where the available research simulator (SIMONA Research Simulator) and laboratory aircraft (PH-LAB, a Cessna Citation II) are used to develop improved stall modeling methodologies. Previous research [21] has resulted in a partial stall model based on Kirchhoff’s method, which includes the longitudinal dynamics only. Because flight data of only quasi-static stall maneuvers were available, it was shown to be difficult to identify dynamic stall effects. Moreover, a thorough substantiation of which terms to include in the model remained to be done, specifically considering control surface degradation, which was not yet included in the model.

Therefore, the goal of this paper is to retrieve an extended Citation II stall model. For this, dedicated flight tests were performed to gather data that contains sufficient dynamic excitation as well as control surface perturbations during stall maneuvers. An air data boom is installed that measures side slip angle, enabling the identification of lateral model terms. Furthermore, aerodynamic model identification is enhanced by using an objective model structure selection based on multivariate orthogonal functions modeling [22]. Finally, the nonlinear parameter estimation problem is approached efficiently, by optimizing the flow separation point parameters separately from the rest of the model.

The structure of this paper is as follows. Section II describes the flight test vehicle and the maneuvers that were flown to gather the data. Section III presents the stall modeling methodology. The parameter estimation and model structure selection techniques that were used are discussed separately. Results are presented in Section IV. Section V follows with a discussion of the work, and Section VI presents the conclusions.

II. Flight Test Data

A. Flight Test Vehicle

Flight tests were performed in a Cessna Citation II (callsign PH-LAB). This aircraft is jointly owned by the Faculty of Aerospace Engineering of Delft University of Technology, and the Netherlands Aerospace Center (NLR). The aircraft is equipped with an advanced flight test instrumentation system (FTIS), which connects and logs data from installed sensors. Schematic views of the aircraft can be found in Fig. 1. Tables 1 and 2 describe general mass and geometric properties, and list the flight test equipment relevant to this research. Critical to this research is the air data boom, which is mounted on the nose of the aircraft, and is illustrated on the left-hand side of Fig. 1. The boom enables accurate measurements of the angle of attack and angle of sideslip [23].

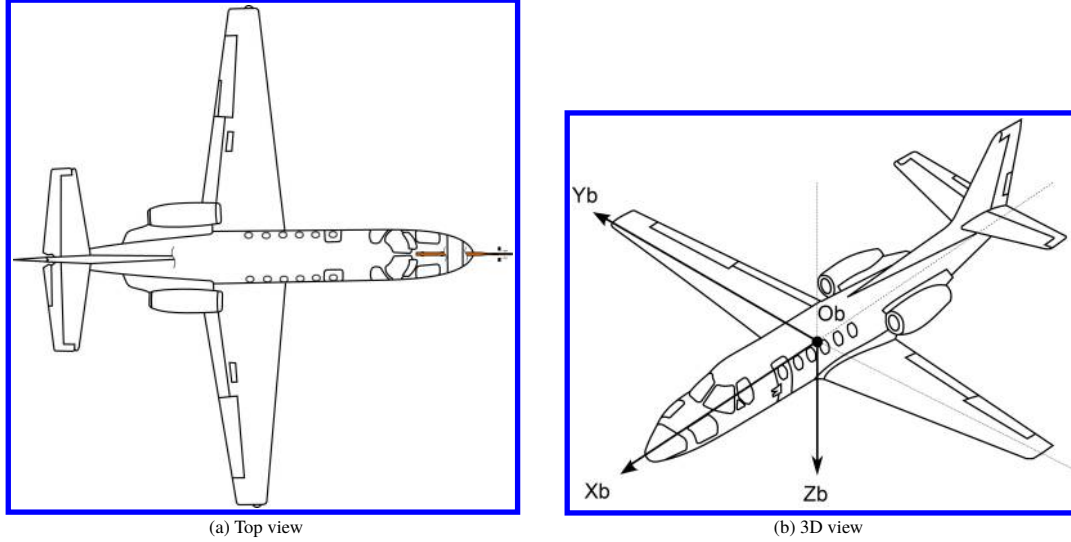


Fig. 1 Schematic views of the PH-LAB, including the body-fixed reference frame definition. (a) also shows the installed air data boom.

Table 1 PH-LAB dimensions and mass properties.

Dimensions	
b	15.9 m
$\bar{c}$	2.09 m
S	30.0 m ²
Mass and inertia	
m	4,157 kg
I_{xx}	12,392 kgm ²
I_{yy}	31,501 kgm ²
I_{zz}	41,908 kgm ²
I_{xz}	2,252.2 kgm ²

Table 2 Flight test equipment installed on the PH-LAB Cessna Citation II aircraft, including the measured variables relevant to this research.

Name	Explanation	Measures	Variables	Units
GPS	Global positioning system	Position in F_E	x_E, y_E, z_E	m
		Velocity in F_E	$\dot{x}_E, \dot{y}_E, \dot{z}_E$	m/s
DADC	Digital airdata computer	Total airspeed	V_{TAS}	m/s
AHRS	Attitude & heading reference system	Aircraft attitude	ϕ, θ, ψ	rad
		Body rotation rates	p, q, r	rad/s
		Body specific forces	A_x, A_y, A_z	m/s ²
Synchro	Angle measurements	Contr. surf. defl.	$\delta_a, \delta_e, \delta_r$	rad
Boom	Air data boom	Air incidence ang.	α, β	rad

B. Flight Test Maneuvers

Two dedicated flights were conducted in December 2018 in order to gather the stall data required for this research. Special maneuvers were flown for the specific goal of stall model identification. In total, 34 stall maneuvers were recorded, all in clean configuration (i.e., no flaps, and gear retracted). Each stall maneuver was trimmed to include stall entry, stall itself, as well as recovery phases.

Fig. 2 visualizes several key characteristics of the data sets. Fig. 2(a) shows that the majority of data sets were gathered around 5,500 m flight altitude. As a result, this is the only condition for which the model output could be validated. Fig. 2(b) shows that angles of attack up to 22 degrees were achieved, where the Citation would stall around 12 degrees. Fig. 3 shows the gathered flight data in a V - n diagram.

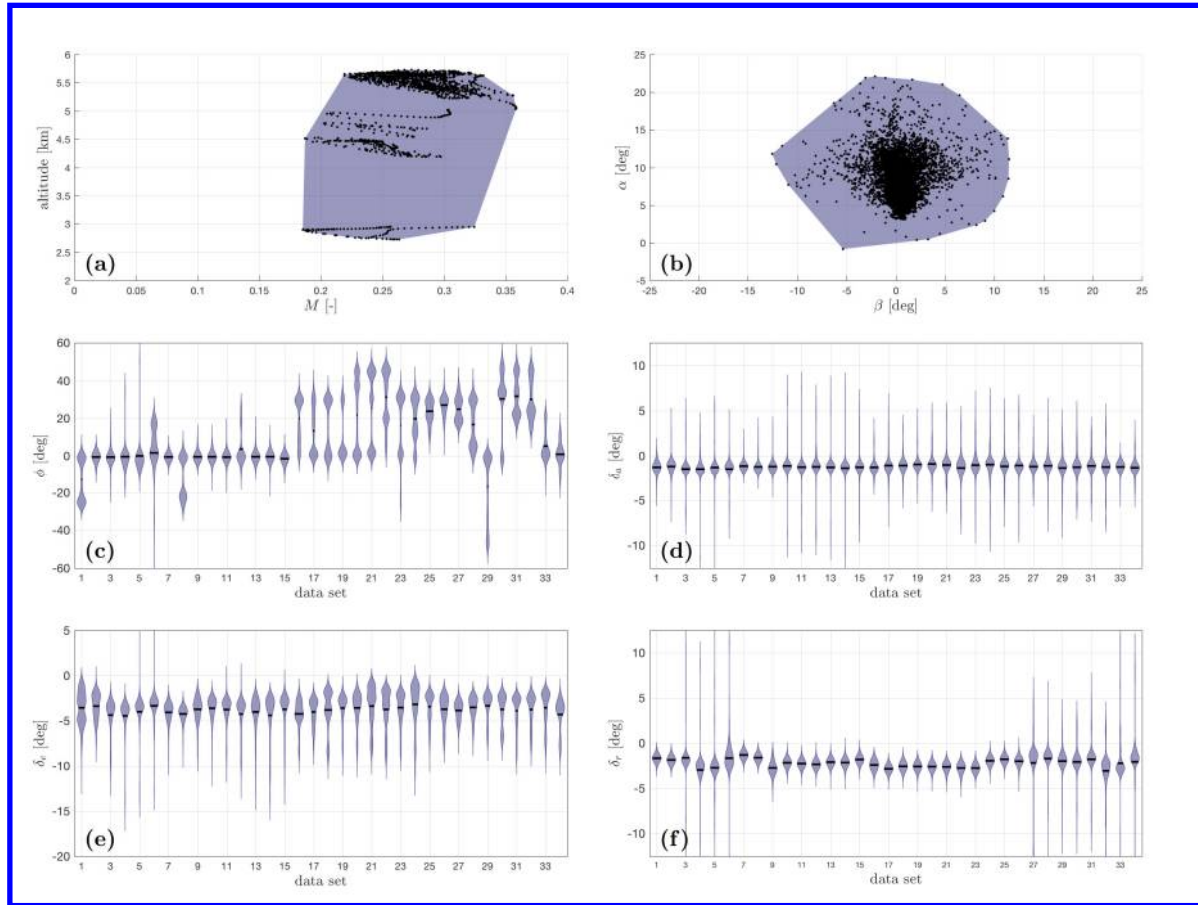


Fig. 2 Visualization of several properties of the gathered flight test data. Subfigures (a) and (b) show the Mach-altitude and α - β envelope, respectively. Plots (c)-(f) are violin plots, which show approximations of the distribution of the bank angle and control inputs, where black marks show means. The difference between longitudinal and lateral maneuvers can clearly be seen, as well as the fact that the rudder was used very little in maneuvers 7-26.

Two types of quasi-steady stall maneuvers were flown: **wings-level symmetric** and **accelerated** stalls. This distinction was made in an attempt to distinguish $\dot{\alpha}$ -related effects from q -related ones. The difference between these maneuver types can be seen in ϕ in Fig. 2(c). The wings-level symmetric stalls were entered by achieving trimmed, level flight, and then reducing airspeed by 1 kts/s by closing the throttle. During the stall, effort was put into keeping the wings-level condition with aileron inputs. The accelerated stall entry was similar, but started from a coordinated turn of either about 30 or 45 degrees bank (approximately 1.1 or 1.3 g). In the stall, the pilots aimed to keep the same bank angle and turn rate. Maneuvers 1-15, 33, and 34 as shown in Fig. 2 are wings-level, the rest are accelerated stalls. All accelerated stalls were right-hand turns, except maneuver 29, which was a left-hand turn.

A piloting technique was developed in collaboration with the test pilots, based on one described by Morelli [22]. The main principle is that pilot inputs are composed of two components: one part to keep the aircraft at or close to the desired flight condition, and a second component with (semi-)random disturbances to excite the aircraft. Figures 2(d), (e), and (f) visualize the use of control surfaces during each maneuver. In Fig. 2(f) it can be seen that large rudder inputs were only applied during maneuvers 3-6 and 27-34. This is because rudder inputs were only applied as “disturbance”. The response of the aircraft to rudder excitations during stall was unknown, which is why rudder deflections were carefully increased during the experiments. The end result is that not all data sets contain sufficient rudder excitations for identifying a yaw moment model. The aileron and elevator were used during all maneuvers.

For the conditions that were tested, the aircraft’s stall behavior was quite benign and showed a strong self-recovering tendency. Active pilot elevator inputs were required to direct the Citation II back into a stalled condition after the aircraft would recover itself. As a result, some of the stall maneuvers are actually a sequence of short stalls. Examples of this can be seen later on in this paper, for instance in Fig. 26.

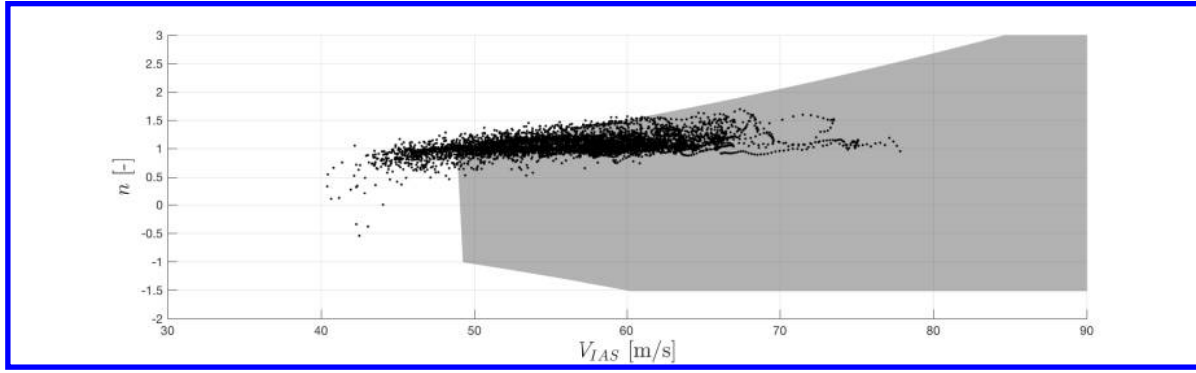


Fig. 3 Decimated view of all the time samples in the flight data sets on the maneuvering load factor limit diagram of the Cessna Citation 550 II, defined for a take-off weight of 11,500 lbs. All airspeeds were converted to indicated airspeeds. The grey shaded area marks the flight envelope where normal flight maneuvers are possible.

C. Data Pre-Processing

Before the model identification, two data pre-processing steps were taken. First, all recorded signals were filtered using a zero-phase low-pass filter using Matlab's `filtfilt` function. A Butterworth filter of order 4 was used. There were two main disturbance sources: vibrations due to the stall buffet, especially for measurements obtained with the air data boom, and amplified noise in signals that are obtained through numerical differentiation (e.g., $\dot{p}$, or $\dot{\alpha}$). The buffet vibrations are vital to a realistic stall model implementation in a simulator, but these are modeled by a separate buffet model [21], and are not in the scope of this research. The cut-off frequencies for the filtered signals are listed in Table 3. An example of filtering can be seen in Figures 4 and 5, where part of a time history and a power spectral density estimate are shown both before and after low-pass filtering.

After applying the low-pass filter, the second pre-processing step was to apply an optimal state reconstruction using an Unscented Kalman Filter (UKF). This filter was developed and shown to be effective in previous research [21] and was adjusted to use the α and β signals from the air data boom.

Table 3 Low-pass filter cut-off frequencies.

Signal(s)	f_c [Hz]
α, β	4.0
$\dot{\alpha}, \dot{\beta}$	4.0
$\delta_a, \delta_e, \delta_r$	4.0
A_x, A_y, A_z	1.5
p, q, r	1.5
$\dot{p}, \dot{q}, \dot{r}$	4.0

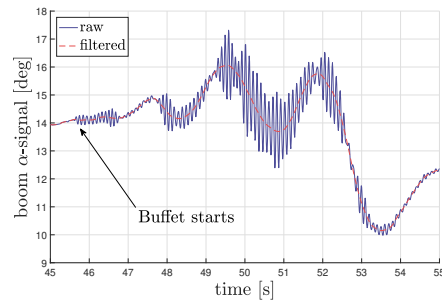


Fig. 4 Detail of boom α -signal time history showing oscillatory vibrations due to buffet.

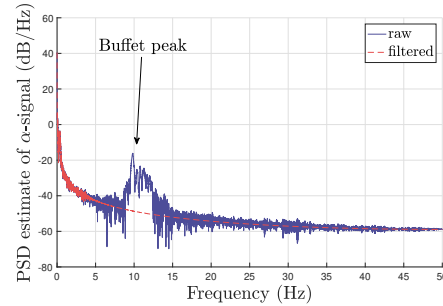


Fig. 5 Power spectral density estimate of α -signal before and after low-pass filtering.

III. Methodology

The stall model identified in this paper consists of six separate models for the force (C_L , C_D , C_Y) and moment (C_l , C_m , C_n) coefficients. The methodology that was followed for the identification of these models is explained in this section. First, the general modeling principle is explained, which is based on Kirchoff's theory of flow separation. After that, a general overview of the system identification process is given. Finally, the methods used for model structure selection and parameter estimation, which form the core of the novel approach in this research, are presented separately.

A. Kirchoff's Theory of Flow Separation

In the early 1990s, a relatively simple model structure was published that is able to explain the nonlinear, dynamic, and time-dependent phenomena caused by aerodynamic stall [12]. Previous research has shown the validity of this

modeling approach for several different aircraft [14, 15]. This model is based on the assumption that the effect of stall on airfoil lift can be described by Kirchoff's theory of flow separation, which states that the relation between airfoil lift and the flow separation point X , which represents the fraction of the chord length affected by flow separation, can be modeled by:

$$C_L = C_{L\alpha} \left(\frac{1 + \sqrt{X}}{2} \right)^2 \alpha. \quad (1)$$

The variable X represents the distance along the wing chord where the flow separates, and ranges from 1 (flow is fully attached) to 0 (flow is fully separated). It was shown that the dynamics of X can be adequately modeled by a first-order ordinary differential equation (ODE):

$$\tau_1 \frac{dX}{dt} + X = \frac{1}{2} (1 - \tanh [a_1 (\alpha - \tau_2 \dot{\alpha} - \alpha^*)]) . \quad (2)$$

Only four parameters are needed for this ODE, and each represents a physical effect of the stall. τ_1 characterizes the lag in the dynamics of X , which is caused by inertia in the flow; it takes time before the air flow has reacted to sudden changes in α . τ_2 encodes hysteresis effects on X . The static mapping between X and α is parametrized by a_1 , which is a shape parameter that determines the ‘‘abruptness’’ of the stall, and by α^* , which defines the angle of attack at which the stall occurs. These four parameters will from here on be referred to as the X -parameters. Visualizations of the effect of varying these parameters on C_L and X can be found in previous research [19, 21].

Because of Eq. (2), solving for the model parameters is a nonlinear optimization problem, which makes it sensitive to initial conditions, and computationally more demanding to solve. Another challenge is that a direct measurement of X is not available. The solution is to estimate the parameters via another measurement that is available, and which is influenced by X . The most obvious candidate for this is C_L , which is directly affected by X as described in Eq. (1). A special system identification methodology, detailed in the next section, was used to deal with these challenges.

B. System Identification Methodology

A key idea in this research is to split the parameter estimation problem into two parts. First, the X -parameters are estimated using nonlinear methods. After that, these model parameters are assumed fixed and thus X can be calculated as a potential regressor for all data sets. This is followed by aerodynamic model structure selection, and finally the estimation of the remaining parameters using conventional linear methods. These three steps are visualized in Fig. 6.

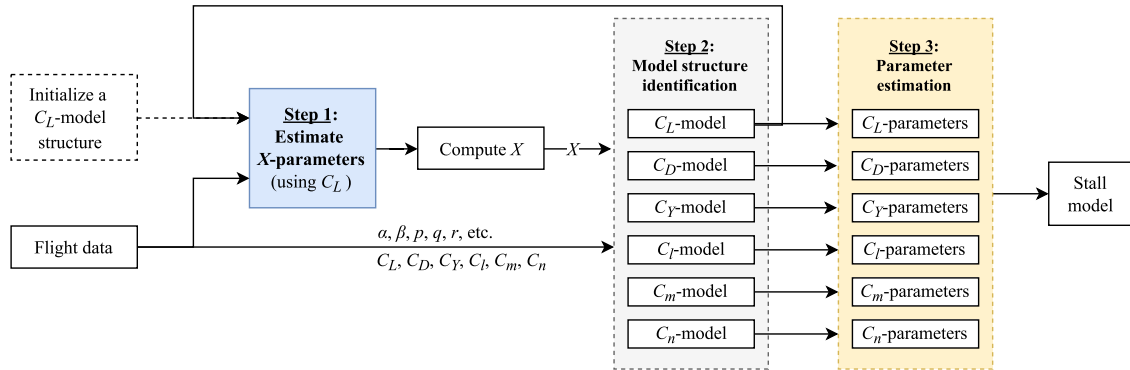


Fig. 6 Block diagram of the steps and flow in the system identification approach.

The most important advantage of this approach is that X can simply be regarded as a regressor during model structure selection. This enables quick iterations between selecting a model structure, estimating its parameters, and evaluating the result of any changes. Due to the iterative and interactive nature of a system identification task, this is seen as a highly desirable advantage. Moreover, if this distinction was not made, the semi-objective model structure selection method would not have been possible.

The X -parameters are identified using the measured C_L , as this relation is best documented in literature [12–15]. An initial assumption on the model structure of C_L needs to be made. The implicit assumption is made that the dependence

of the X -parameters on the model structure is mild. When, during the later model structure selection step, a different model for C_L is found, the X -parameters will be re-estimated, and this process is repeated until it has converged. The next sections describe the steps in Fig. 6 in detail.

C. Parameter Estimation

As shown in Fig. 6, the estimation of our model's parameters is split into two parts. The linear parameter estimation in step 3 is done using the familiar ordinary least-squares methods, whilst the nonlinear estimation of step 1 is done using a gradient-based solver available in Matlab.

1. Linear Parameter Estimation

Once the model structures of the force and moment coefficient equations are set, the parameter estimation problem is straightforward. For this research, it was chosen to use the model structures in polynomial form, i.e.,

$$\hat{\mathbf{y}} = \theta_1 \mathbf{a}_1 + \theta_2 \mathbf{a}_2 + \dots + \theta_n \mathbf{a}_n, \quad (3)$$

where $\hat{\mathbf{y}}$ is the model output vector, $\mathbf{a}_i$ are the regression variable vectors, and θ_i are parameters. Note that $\mathbf{a}_i$ is allowed to contain non-linear polynomial terms, such as α^2 . Eq. (3) can be written in matrix notation as:

$$\hat{\mathbf{y}} = A\theta, \quad \text{where } A = [\mathbf{a}_1 \ \mathbf{a}_2 \ \dots \ \mathbf{a}_n], \quad \text{and } \theta = [\theta_1 \ \theta_2 \ \dots \ \theta_n]^\top. \quad (4)$$

To minimize the error between measurement and model output $\epsilon = \mathbf{y} - \hat{\mathbf{y}}$, ordinary least squares (OLS) is used. This leads to the well-known closed-form solution for the parameter estimate $\hat{\theta}$:

$$\hat{\theta} = (A^\top A)^{-1} A^\top \mathbf{y}. \quad (5)$$

This procedure is the same for all aerodynamic model equations.

2. Nonlinear X -Parameter Estimation

A second parameter estimation method was used for estimating the X -parameters. An optimization problem was defined with the goal of minimizing the mean squared error between the measured lift coefficient C_L , and the model output $\hat{C}_L(\theta, x)$. Both are defined as $N \times 1$ vectors, and x indicates the regressors obtained from the flight test data.

$$\hat{\theta} = \arg \min_{\theta} J(\theta, x) \quad \text{with} \quad J(\theta, x) = \frac{1}{N} \left(C_L - \hat{C}_L(\theta, x) \right)^\top \left(C_L - \hat{C}_L(\theta, x) \right). \quad (6)$$

This method requires a choice for the model structure of $\hat{C}_L$. *Only during Step 1* of the system identification procedure, see Fig. 6, the following model was used:

$$\hat{C}_L(\theta, x) = C_{L_0} + C_{L_\alpha} \left(\frac{1 + \sqrt{X}}{2} \right)^2 \alpha + C_{L_q} \frac{q\bar{c}}{V}, \quad (7)$$

with X as in Eq. (2). Upper and lower bounds were set on all parameters as constraints. The initial conditions were randomly and uniformly sampled from the parameter space constrained by these boundaries. The active set algorithm, as available from Matlab's `fmincon` function was used to solve for the X -parameters. Multiple runs were done on each data set. Since the cost function surface tended to be flat near the solution, the best few percent of runs were averaged to obtain a point estimate for each data set. This final step reduced the variance of the results. To deal with the numerical difficulties of the parameter sensitivity of the solution of an ODE [24], at each iteration, the gradient of the cost function with respect to the parameters was explicitly computed using the method described in Appendix A.

For the initial step, which used the model structure as shown in Eq. (7), the upper and lower bounds are presented in Appendix B. For the final optimization results, the boundaries are presented in Table 4.

Regarding the choice for the initial model structure: the only hard requirement is that some dependency on X is included, otherwise estimating the X -parameters is impossible. The Kirchoff term found in literature was selected, as its effectiveness has been shown before. The q -term was added because it is commonly found in models for lift. On subsequent iterations, the model structure that followed from the model structure selection step was used instead, so the initial model structure of $\hat{C}_L$ was not seen as critical.

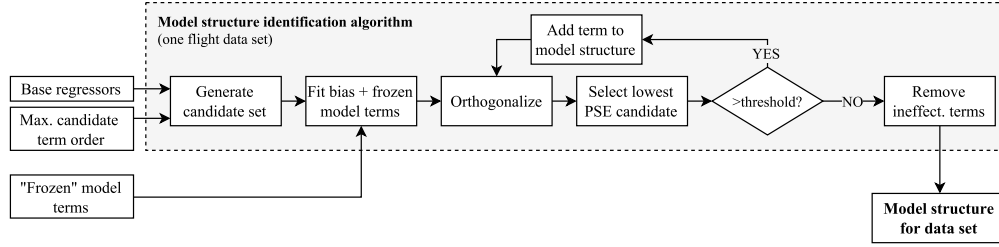


Fig. 7 Block diagram visualization of the multivariate orthogonal function modeling algorithm used for model structure selection.
During step 2 in Figure 6, this algorithm is run for each model (C_L , C_D , ...) for each training data set.

D. Model Structure Selection Method

Before estimating the model parameters, a model structure needs to be selected. In other words, this means identifying which regressors are required for capturing the dynamics in the measured data. Note that it is during this step in the system identification process that choices are made *if* e.g., changes in control surface effectiveness or dynamic effects need to be included, and if so, *how* this is done.

A good model structure must be effective at explaining the patterns in the data, and be parsimonious. The model structure has a large influence on the trade-off between bias (model cannot represent data) and variance (model does not generalize well to previously unseen data) [25]. In general, evaluating the usefulness of separate model terms is hard, since these influence each other i.e., adding an extra term to some model structure is very likely to change the estimated parameter values of the other terms. Further complexities are caused by the fact that the flight data will inevitably contain errors, have interdependencies/correlations, and is limited in quantity.

An multivariate orthogonal function modeling algorithm was developed as a tool to deal with these challenges. Its block diagram is presented in Fig. 7, and will be explained in the next sections. Sections III.D to III.D.3 will explain the way it is used and how its results were interpreted.

1. Multivariate Orthogonal Function Modeling Algorithm

An algorithm for creating and selecting model terms that lead to a good model was developed based on the multivariate orthogonal function (MOF) modeling method described by Morelli [22]. To avoid correlations of model terms, all candidate model terms are orthogonalized. The most important consequence of this is that it becomes possible to compute the effect of adding terms independently of each other. This makes it an objective and (semi-)automated procedure for model structure selection.

The algorithm's basic principle is that it iteratively "builds up" model structures from scratch, using a pool of orthogonalized candidate regressor terms. This iterates until a cost function is minimized, which penalizes both errors between model output and the dependent variable, as well as model complexity. In the following, orthogonalized regression variables will be denoted by $\mathbf{p}_i$ to distinguish them from the regular regression variables $\mathbf{a}_i$. The algorithm starts by generating a set of candidate model terms, based on a set of base regressors and a maximum term order. This will be discussed in the next section. After that, the first step is to include a bias term:

$$\mathbf{a}_1 = \mathbf{p}_1 = \mathbf{1} . \quad (8)$$

Then at each iteration, a Gram-Schmidt orthogonalization procedure is used for making all remaining candidate model terms orthogonal to the terms that are already in the selection.

$$\mathbf{p}_j = \mathbf{a}_j - \sum_{k=1}^{j-1} \gamma_{k,j} \mathbf{p}_k , \quad j = 2, 3, \dots, n , \quad \text{where:} \quad \gamma_{k,j} = \frac{\mathbf{p}_k^T \mathbf{a}_j}{\mathbf{p}_k^T \mathbf{p}_k} . \quad (9)$$

Note that the orthogonal regressors are linear combinations of the original regressors. Once orthogonalized, all candidates are evaluated based on the effect that they would have on the predicted square error (PSE). This cost metric consists of two parts: a fit error (equal to the mean squared error), and a penalty term for complexity:

$$\text{PSE} = \frac{(\mathbf{y} - \hat{\mathbf{y}})^T (\mathbf{y} - \hat{\mathbf{y}})}{N} + \sigma_y^2 \frac{n}{N} . \quad (10)$$

In Eq. (10) N is the number of data points, n is the number of terms currently used in the model, and σ_y^2 is the variance of the modeled signal $\mathbf{y}$ (e.g., C_m , C_Y , etc.), which is used as a scaling term. Because of the orthogonality, the change in

PSE which would result from adding candidate j can be computed independently for each candidate as:

$$\Delta \text{PSE}_j = -\frac{(\mathbf{p}_j^\top \mathbf{y})^2}{\mathbf{p}_j^\top \mathbf{p}_j} + \sigma_y^2 \frac{1}{N}. \quad (11)$$

The candidate leading to the greatest reduction in cost is selected and added to the model. After that, the procedure is repeated: all candidates are again made orthogonal to the selected terms, the best one is selected, etc. This process is repeated until $\Delta \text{PSE}_j \geq 0$ for all j , which means that the added penalty of increasing the model complexity is no longer offset by a decrease in mean square error.

At this point, a matrix $P = [\mathbf{p}_1 \ \mathbf{p}_2 \ \dots \ \mathbf{p}_n]$ has been constructed, which can be used to find the vector of maximum-likelihood parameters $\hat{\phi}$ connected to the orthogonalized regressors. For this, the OLS procedure that was explained can be used in exactly the same way. The result $\hat{\mathbf{y}} = P\hat{\phi}$ can then be transformed back to the “normal” non-orthogonalized regressors and associated parameters $\hat{\mathbf{y}} = A\hat{\theta}$. When the $\gamma_{j,k}$ are collected in a matrix:

$$\Gamma = \begin{bmatrix} 1 & \gamma_{1,2} & \gamma_{1,3} & \dots & \gamma_{1,n} \\ 0 & 1 & \gamma_{2,3} & \dots & \gamma_{2,n} \\ 0 & 0 & 1 & \dots & \gamma_{3,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}, \quad (12)$$

then $A = P\Gamma$, and thus the optimal parameter vector $\hat{\phi}$ can easily be transformed back to the original regressor definition by applying the following substitution:

$$\hat{\mathbf{y}} = P\hat{\phi} = PI\hat{\phi} = P(\Gamma\Gamma^{-1})\hat{\phi} = A\Gamma^{-1}\hat{\phi}, \quad \text{hence: } \hat{\theta} = \Gamma^{-1}\hat{\phi}. \quad (13)$$

Once A and $\hat{\theta}$ are obtained, the algorithm checks all the terms that have been selected for their contribution to the final model output. It does so by computing the root mean square $\text{RMS} = \frac{1}{N}\sqrt{\hat{\mathbf{y}}^\top \hat{\mathbf{y}}}$ of the model output $\hat{\mathbf{y}}$, and comparing that with the RMS of the model output that would occur if that term would be simply ignored, without making any other changes to the parameters or regressors. If the change in RMS is lower than a threshold, for which here 0.5% was used, the term would be regarded as ineffective and removed.

The output of the MOF modeling algorithm of Fig. 7 is a selection of regressor terms A for the specific flight data set that was used as input. As a final step, the results of all data sets need to be generalized. This is done based on a simple count of how many times each model term appears in the output for each flight data set. If a term is selected in 50% of data sets or more, it generally is concluded to be useful, and included in the general model structure. This threshold was used since it was found to work well in the identification of the regular flight envelope model of the Citation II [26]. However, the results were always interpreted alongside other analyses such as the effect on MSE, consistency in the estimate of the associated parameter, or correlations between estimated parameters.

2. Candidate Model Terms

The described algorithm iteratively searches within a set of candidate terms for the best contribution to a model. Any decision that influences the terms that are included in this set will have a strong influence on the final results. If the pool does not contain any useful terms, the quality of the model will never be good. Or, if many similar terms are included, it will be impossible to choose between them since the results will be inconsistent across data sets. This section explains how the candidate model terms are generated, and explains several design choices that were made.

The most straightforward approach for creating candidates is to simply use the common regressors such as α , δ_e , or p . In this research, the thrust force coefficient C_T and Mach number M were also included. Appendix C explains why including C_T is important in the current application. This approach is extended by also including X and several mathematical transformations of X as regressors. These will play a crucial role in including the effects of aerodynamic stall on the aircraft dynamics. Again, note that this is only possible because the parameter estimation problem is split into two separate parts, which enables the assumption that the X -parameters can be considered constants during model structure selection. The following terms were used as “base regressors”:

$$\mathbf{1}, \alpha, \dot{\alpha}, \beta, \dot{\beta}, p, q, r, \delta_a, \delta_e, \delta_r, C_T, M, \\ X, (1 - X), \left(\frac{1+\sqrt{X}}{2}\right)^2, \max(0.5, X).$$

These are named “base regressors” since they form the basis for the automatic generation of any potential higher-order candidate regressors. A routine was implemented that automatically generates all unique product-wise combinations of the base regressors, up to and including a user-defined maximum order n .

Four base regressors that contain X were included and it is important to consider why these were selected. In this, it is relevant to re-state that for attached flow $X = 1$, and for fully separated flow $X = 0$. The regressor X itself can be used for explaining an effect that reduces or disappears during stall. The second regressor, $(1 - X)$, does the opposite, it only takes effect during flow separation. The third term is part of the term appearing in Kirchoff’s classical theory of flow separation, which would be created by taking the product of this term with α . Finally, the term $\max(0.5, X)$ was used to take into account some effects that change during stall, but do not completely disappear for fully separated flow.

Obviously, many more potential base regressors may be selected. However, it was found that it is important to avoid similar candidate regressors, i.e., those that have strong correlations with each other. If two regressors are correlated, they explain the same patterns in the data. The result of this is that it is not clear which regressor to select, which makes the results ambiguous. For example, consider a feature that is present in 90% of the data. Then, say there are three regressors that are equally suitable for explaining this feature. Each will then, on average, be selected on 30% of the data sets, which is below the 50% threshold. When viewing these results, the user should recognize that these terms are equivalent.

3. Practical Use of Algorithm

Following the approach outlined in Section III.D.2, it was found that in practice setting the maximum model order higher than $n = 2$ caused correlated terms being included in the candidate pool, resulting in less useful and consistent model outcomes. Next to that, the amount of terms increases very quickly for increasing model order, which complicates the interpretation of the results by a user. As a result, the algorithm was used in quite a specific way, which will be explained in the next section.

Ideally, the model structure selection algorithm converges to the same model structure for all data sets. It was found that this was not the case. When setting the maximum order of candidate terms higher than two, models with many higher-order model terms were the result. This conflicted with the goal of selecting a parsimonious model structure.

The suspected reason for why this occurs is that the flight data contains nonlinear effects that cannot easily be modeled, as well as significant random disturbances. While simpler regressor terms might be able to explain general trends just fine, it is likely that there exists a higher-order term that does the job better on that specific data set alone. The disturbances differ for each data set. Hence if the described algorithm is given the choice, it will select different higher-order terms for each data set. This makes it hard to identify model terms that generalize well across data sets.

The solution that was implemented was to change the way in which the algorithm was used. This is visualized in Fig. 8. The algorithm is run for both $n = 1$ and $n = 2$ separately. The results are then interpreted manually to make conclusions about the useful model terms. In most cases, useful new model terms are only found for $n = 1$, but there are two exceptions (C_L and C_m) that will be discussed in the results. The usefulness of terms is judged by the effect on the total model fit quality (MSE and R^2) on the validation data sets.

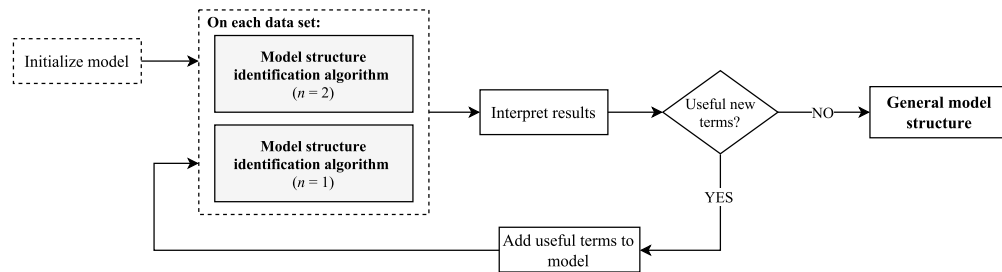


Fig. 8 Flow chart of the practical way that the model structure selection algorithm was used. Usefulness of model terms is judged by the effect on MSE and R^2 on the validation data sets.

Any useful terms that are found are “frozen” into the model structure such that they will always be included during future iterations. Iterations are run until no more new good model terms are found. In all cases, only 1 or 2 iterations were needed. The resulting model structure is passed on to the parameter estimation step of the system identification process.

IV. Results

To identify the aerodynamic model through the proposed methodology, 34 flight data sets were available. It was chosen to split this data into 27 training and 7 validation sets (roughly a 80-20 split). Sets 5, 7, 12, 16, 22, 23, and 25 were used for validation. The split was made randomly, but it was ensured that both the training and validation data sets contained wings-level as well as accelerated stalls. First, the final model structure that was obtained from the model structure selection method described in previous section will be presented. Next, the results of the nonlinear X -parameter optimization are presented. Third, the parameter estimates for the aerodynamic model equations are shown. The final subsection gives the results of model validation, which showcases the current quality of the stall model.

A. Stall Model Structure

The final stall model structure, which resulted from multiple iterations of the process described in Fig. 6 is given in Eq. (14). The model for each of the aerodynamic model equations will be discussed separately.

$$\begin{aligned}
 \hat{C}_L &= C_{L_0} + C_{L_\alpha} \left(\frac{1+\sqrt{X}}{2} \right)^2 \alpha + C_{L_{\alpha^2}} (\alpha - 6^\circ)_+^2 \\
 \hat{C}_D &= C_{D_0} + C_{D_\alpha} \alpha + C_{D_{\delta_e}} \delta_e + C_{D_X} (1 - X) + C_{D_{C_T}} C_T \\
 \hat{C}_Y &= C_{Y_0} + C_{Y_\beta} \beta + C_{Y_p} \frac{pb}{2V} + C_{Y_r} \frac{rb}{2V} + C_{Y_{\delta_a}} \delta_a \\
 \hat{C}_l &= C_{l_0} + C_{l_\beta} \beta + C_{l_p} \frac{pb}{2V} + C_{l_r} \frac{rb}{2V} + C_{l_{\delta_a}} \delta_a \\
 \hat{C}_m &= C_{m_0} + C_{m_\alpha} \alpha + C_{m_{X\delta_e}} \max(\frac{1}{2}, X) \delta_e + C_{m_{C_T}} C_T \\
 \hat{C}_n &= C_{n_0} + C_{n_\beta} \beta + C_{n_r} \frac{rb}{2V} + C_{n_{\delta_r}} \delta_r
 \end{aligned} \tag{14}$$

1. C_L -model

The obtained C_L -model contains the classical Kirchoff term, and a spline term using the angle of attack squared, $(\alpha - 6^\circ)_+^2$, which will be discussed later. The model structure selection algorithm was run as explained in Fig. 8. The results of the first iteration are shown in Fig. 9(a). The grey bar for the bias term indicates that it was frozen into the model; it was not left up to the algorithm itself to select it. Three terms were found to be useful (i.e., occurring in more than 50% of data sets): α , $((1 + \sqrt{X})/2)^2$, and the combination of the two, the Kirchoff term: $((1 + \sqrt{X})/2)^2 \alpha$. It was decided to include the latter, since it resulted in the largest improvement in model validation fit quality.

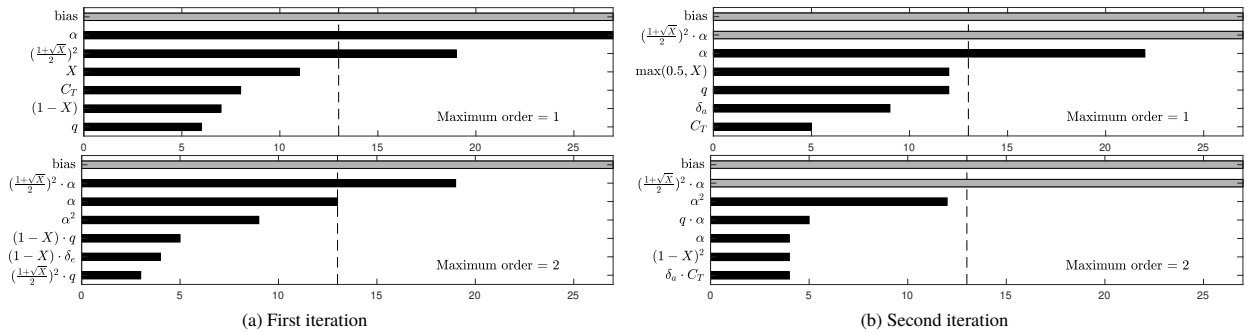


Fig. 9 Results of model structure selection algorithm for C_L displaying the count of how many times a model structure is selected. Dashed bar marks 50% of training data sets. Grey bars indicate that a term was forced to the model structure manually.

Next, the algorithm was re-run with the Kirchoff term frozen in the model structure. The results of this second iteration are in Fig. 9(b), and they indicate that another dependency on the angle of attack is a potentially good term: either as α or α^2 . This was surprising, as initially it was expected that all dependency on the angle of attack would be modeled by the Kirchoff term. However, after experimenting with several variations of α -related terms, the term $(\alpha - 6^\circ)_+^2$ was added to the model. This notation indicates a univariate spline in α with zero-order continuity:

$$(\alpha - 6^\circ)_+^2 = \begin{cases} (\alpha - 6^\circ)^2 & \text{when } \alpha \geq 6^\circ \\ 0 & \text{when } \alpha < 6^\circ \end{cases} \quad (15)$$

In understanding why this extra term is useful to have in addition to the Kirchhoff term, consider what happens when the aircraft enters a quasi-steady stall. Kirchhoff's theory of flow separation predicts the lift due to α will decrease by a factor of 0.25 when X goes to zero. However, this reduction is too strong for our data, as the measured reduction in lift is less. Discrepancies such as these make sense: Kirchhoff's theory was derived for airfoils, whereas the flight data is of a full aircraft in three-dimensional flow. The term $(\alpha - 6^\circ)_+^2$ contributes a positive effect to the aircraft lift at high angles of attack, even when X goes to zero. It thus mitigates part of the lift reduction due to stall. The beneficial effect on the C_L -model is visualized in Fig. 10, where during the deepest part of the stall ($t = 40$ -50 s), the spline corrects part of the lift reduction modeled by the Kirchhoff term.

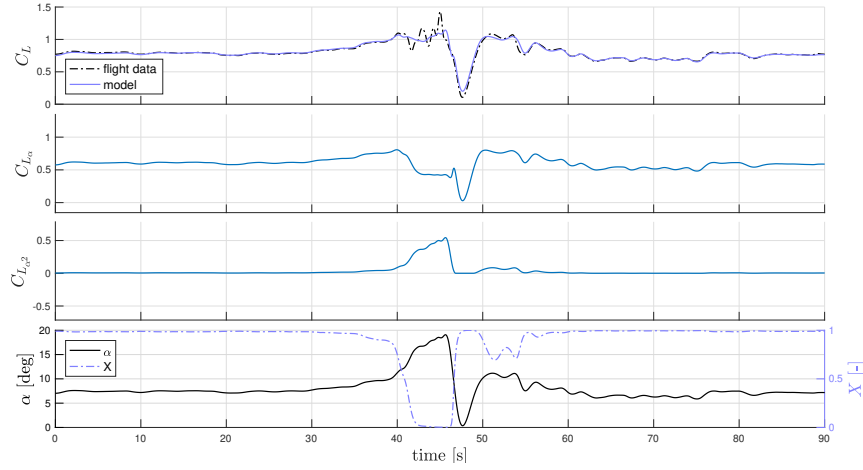


Fig. 10 Contributions of individual C_L -model terms for data set 5 (validation).

Although the difference in model validation fit was modest (difference is $<1\%$ on MSE), it was decided to use the spline term instead of simply α^2 . There were two reasons for this. First, the spline term's effects are limited to higher angles of attack ($\alpha > 6^\circ$) which reduces the correlation with the Kirchhoff term. This led to lower variance in the parameter estimates. Secondly, if α^2 was used, this term would contribute positive lift for negative angle of attack – which can occur during a stall recovery maneuver – which is clearly unphysical. The spline term prevents this. The threshold of 6° resulted in the best fit, and was found using trial and error.

During a third iteration, no new model terms were found. The way the model structure selection algorithm was used for C_L is a good illustration of both its strengths and limitations. It provided no clear-cut answer, but its results were insightful to guide further analysis.

2. C_D -model

For the drag coefficient model structure selection process, Fig. 11(a) shows the results of the first iteration. Many contributing model terms were identified, including C_T . The appearance of the thrust coefficient term has to do with suspected errors in the engine model. This is discussed in Appendix C in detail. The useful second-order terms in the first iteration all contained C_T . No clear underlying physical phenomenon was found for this. Also, further analysis showed that these terms generalized badly across data sets, so they were not selected.

During the first iteration, multiple X -related terms were found. It was decided that including $(1 - X)$ was most desirable, for two reasons. First, the difference in model quality between the X -terms was very small. This is interesting to note, and is indicative of the difficulties in selecting the best model terms. Secondly, the effect of $(1 - X)$ is constrained to just the stalled flight conditions, which means that the effect on the other model terms (especially the bias term) is limited as much as possible. To conclude, after the first iteration C_T , δ_e , α , and $(1 - X)$ were selected as regressors for C_D .

Fig. 11(b) shows the results of the second iteration. The part of C_D not modeled by the selected terms does not correlate clearly to any remaining candidate. This can be interpreted either as that the current model structure captures

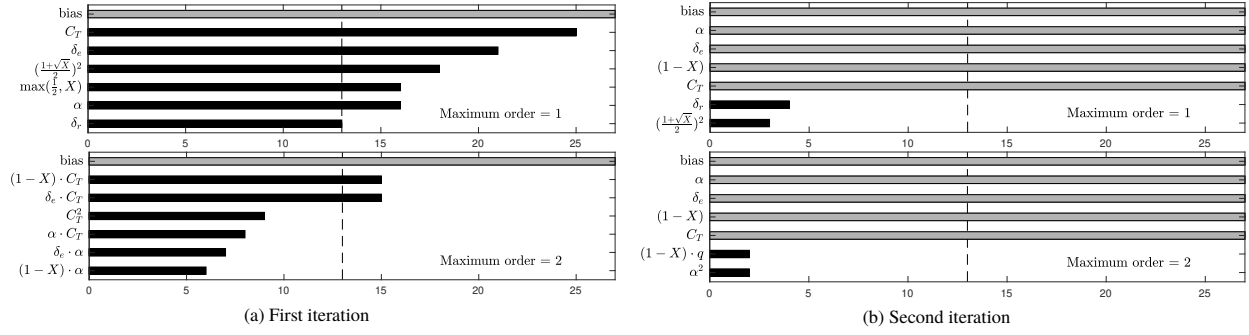


Fig. 11 Results of model structure selection algorithm for C_D displaying the count of how many times a model structure is selected. Dashed bar marks 50% of training data sets. Grey bars indicate that a term was forced to the model structure manually.

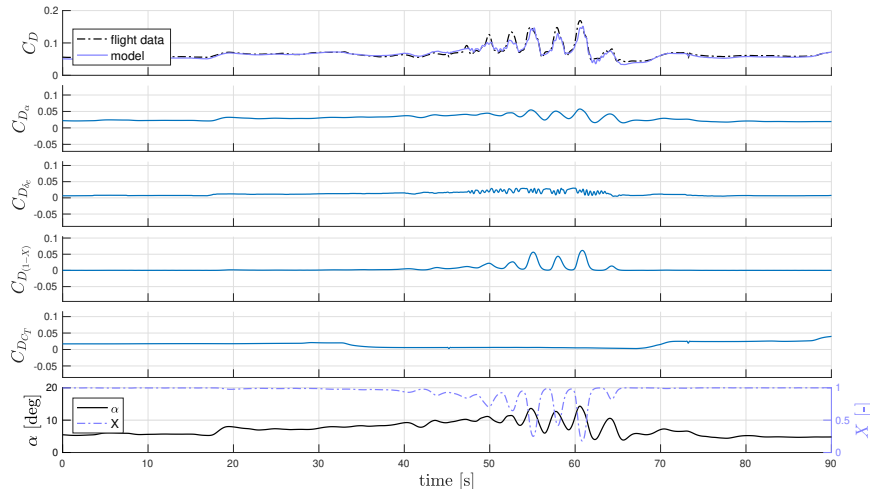


Fig. 12 Contributions of individual C_D -model terms for data set 23 (validation).

all relevant dynamics, or as that the candidate set is missing suitable candidates. Either way, no new terms were added to the model.

Fig. 12 shows the contributions of the individual terms in the model for C_D . It is seen that the largest influences are produced by the angle of attack and by flow separation point X . Also the effect of including C_T is clear, none of the other terms can explain the sudden rise in drag that occurs after the stalled flight. Finally, the elevator term is used to explain high-frequency fluctuations in drag, which are also seen in the measured C_D data.

3. C_Y -model

The obtained stall model structure for the lateral force contains no stall-related terms. The data does not suggest that special additions to the conventional aerodynamic modeling terms are necessary. Furthermore, only first-order terms are included in the model, as no support for more complex terms is found in the results of the model structure selection algorithm. β , p , r , and δ_a were included after the first model iteration. In the second iteration δ_r was also identified as a potentially good contribution, but further analysis showed a small negative effect on the validation fit quality. Plots of individual model contributions are presented in Appendix D.

4. C_l -model

Fig. 14 shows the results of the model structure selection algorithm for the roll moment coefficient. As for the lateral force, only first-order terms were found to be useful. Interesting to see is that both β and r are found to be good options, of which only the latter is selected since it led to a (slightly) lower MSE. On the second iteration, the angle of attack and

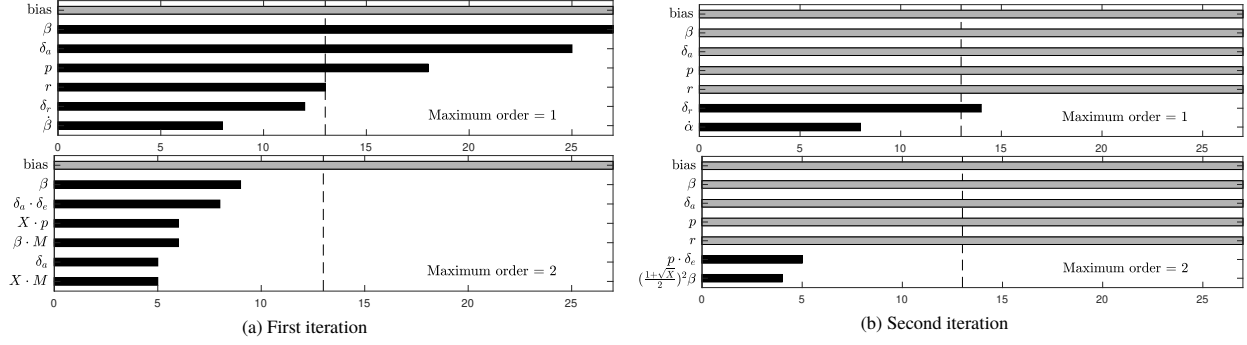


Fig. 13 Results of model structure selection algorithm for C_Y displaying the count of how many times a model structure is selected. Dashed bar marks 50% of training data sets. Grey bars indicate that a term was forced to the model structure manually.

elevator also received good scores. Since there is little physical justification for including these terms in the roll moment model, and because they did not increase the fit of the model on validation data, they were not added.

No stall-related terms were selected; the data does not suggest that being in a stall results in changes to the roll dynamics. This only partially agrees with experience from the flight tests. On the one hand, it was commented by the test pilots that roll control was not noticeably affected during stall. This suggests that changes to aileron effectiveness are not needed in the model. On the other hand, during some maneuvers the stall resulted in significant roll-off motions, and these are currently not included in the model. Plots of individual model contributions are presented in Appendix D.

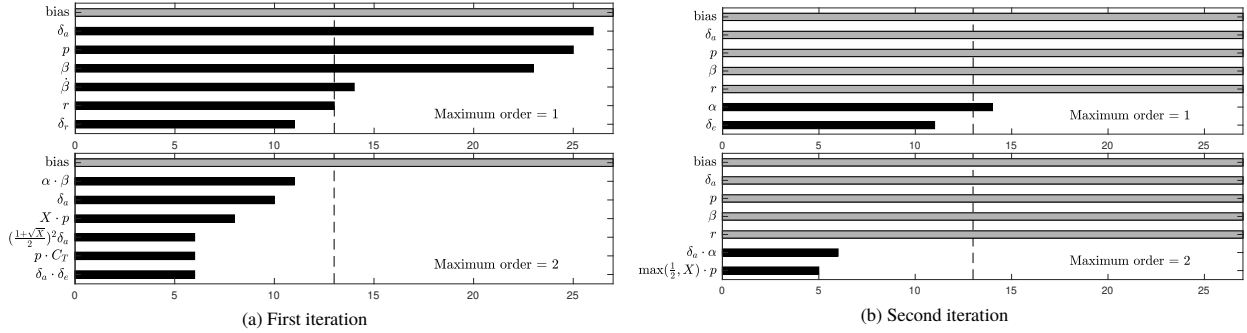


Fig. 14 Results of model structure selection algorithm for C_l displaying the count of how many times a model structure is selected. Dashed bar marks 50% of training data sets. Grey bars indicate that a term was forced to the model structure manually.

5. C_m -model

The obtained model for the pitch moment includes the flow separation point variable in the form $\max(\frac{1}{2}, X) \cdot \delta_e$. From Fig. 16(a), it can be seen that this second-order term is selected in the majority of data sets, indicating a strong match with flight data. This model structure reduces the elevator effectiveness by up to half its original value during stall. Physically, this makes sense. In the stall, the horizontal tail can be in the wake of the main wing, reducing its effectiveness. This also agrees with comments from the test pilots, who stated that in stall, pitch controllability was reduced but still clearly present.

Next to the modified control term, terms depending on α and C_T are selected. Due to the vertical offset of the engines, the thrust setting influences the pitch moment. The contributions of the individual model terms are visualized in Fig. 15. During the stall onset (40-50 s), a (small) negative bias occurs. This is seen in more data sets. One possible reason for this is that the modeled change in control effectiveness is inaccurate during stall onset, which is possible since $\max(\frac{1}{2}, X)$ is a crude mapping. Another possible reason is that an important effect is not included in the model. However, no solutions to this issue were found.

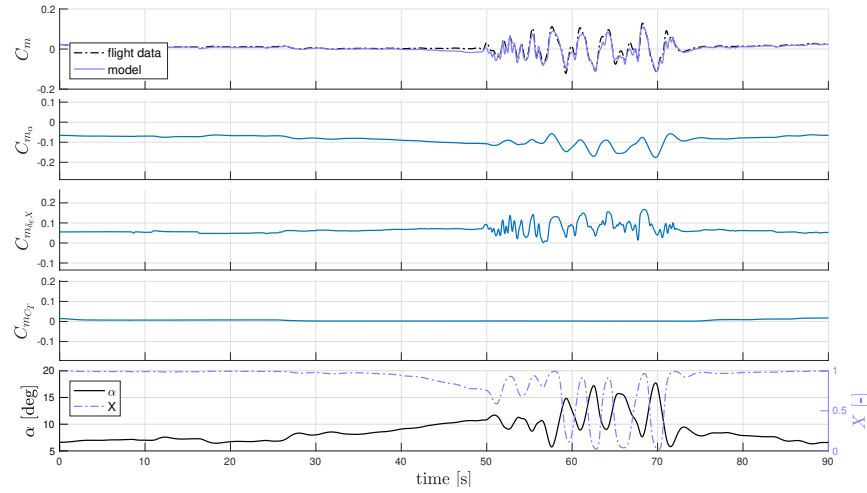


Fig. 15 Contributions of individual C_m -model terms for data set 12 (validation).

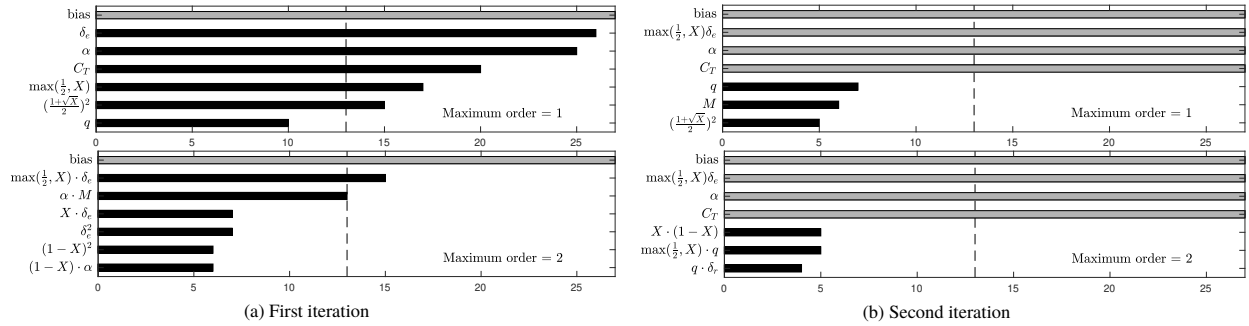


Fig. 16 Results of model structure selection algorithm for C_m displaying the count of how many times a model structure is selected. Dashed bar marks 50% of training data sets. Grey bars indicate that a term was forced to the model structure manually.

6. C_n -model

For training the yaw moment model, only 11 out of the 27 training sets could be used. As shown in Fig. 2, elevator and ailerons were used in all stall maneuvers to keep the aircraft at desired pitch and bank angles, but rudder inputs were only applied as deliberate disturbance inputs, and only for some of the data sets. It was found that without the rudder disturbances, C_n identification results were not reliable.

Using only the 11 data sets that were suitable, the same model structure selection procedure was applied. After the first iteration, β and δ_r were added, as is shown in Fig. 17(a). The second iteration of the algorithm did not select any new model terms. Based on engineering insight, the effect of adding r as a term was also tested. Although it was not picked by the model structure selection algorithm, it did lead to a small improvement in model quality. Plots of individual model contributions are presented in Appendix D.

B. X -parameter Estimates

As shown in Fig. 7, the estimation of X -parameters was performed iteratively. After the C_L model structure had converged to the form in Eq. (14), one final optimization of X was performed. The results of this optimization are presented in this section. Table 4 lists the numerical results and the outcomes of several statistical tests. Fig. 19 is a matrix plot of the parameter estimates, which visualizes the distributions and correlations of the estimates. In each figure, the diagonal plots are histograms that visualize an approximation of the parameter distributions. The off-diagonals are scatter plots of each of the parameter combinations. If correlations are present, the data points will form a clear diagonal. The red star marks the median value, and the values of the correlations are shown (ρ).

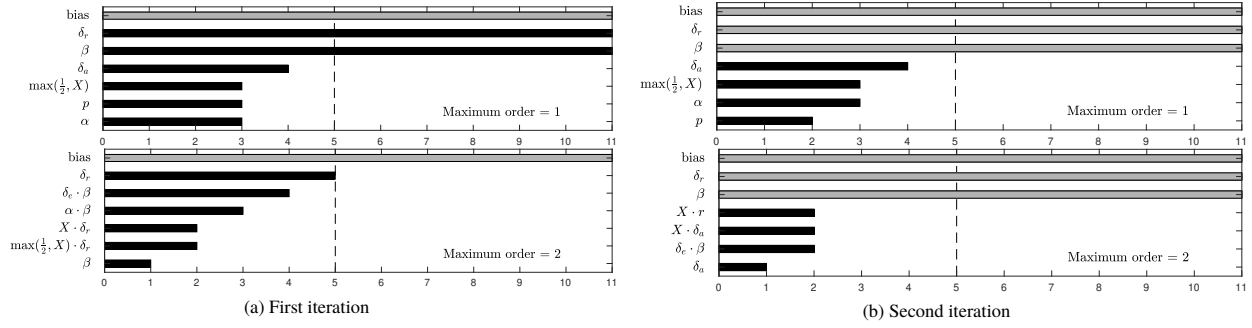


Fig. 17 Results of model structure selection algorithm for C_n displaying the count of how many times a model structure is selected. Dashed bar marks 50% of training data sets. Grey bars indicate that a term was forced to the model structure manually.

Table 4 Results of estimating the X -parameters and several statistical tests. For the KS-test a significance of $\alpha = 0.1$ was used. $h = N$ means that the distribution can be assumed to be approximately normal, $h = \times$ when that this is not possible. The t -test and signed rank test were run with $\alpha = 0.01$ and a Bonferroni correction (here $n = 4$). For these, $h = *$ if the null hypothesis is rejected, and the parameter value is expected to differ from zero.

Parameter		Results				KS-test		t -test		Signed rank	
Name	Unit	$\hat{\theta}$	θ_{lb}	θ_{ub}	$\sigma(\hat{\theta})$	p	h	p	h	p	h
τ_1	[s]	0.2547	0.001	0.80	0.1565	0.147	N	0.000	*	0.000	*
τ_2	[s]	0.0176	0.000	0.50	0.0819	0.020	$\times$	0.002	*	0.000	*
a_1	[-]	27.6711	15.000	40.00	6.7177	0.781	N	0.000	*	0.000	*
α^*	[rad]	0.2084	0.100	0.35	0.0202	0.594	N	0.000	*	0.000	*

For each of the 27 training data sets, the optimization was run 300 times from random initial conditions. These initial conditions were uniformly sampled from the parameter space defined by the upper and lower parameter boundaries (θ_{ub} and θ_{lb}), see Table 4. For each data set, the results of the 300 runs were ordered by the MSE of the model output. It was found that the cost function surface was relatively flat near the optimal solution. Therefore, to reduce the variance of the parameter estimates, it was decided to take the average of all runs that reached an MSE within 2% of the best run as the point estimate for that data set. The medians across all data sets were taken to obtain the final parameter estimates $\hat{\theta}$.

Three statistical tests were performed. The Kolmogorov-Smirnov (KS) test was used to determine whether it is reasonable to assume that the distribution of the estimates is approximately normal. Next, both parametric (t -test) and non-parametric (Wilcoxon's signed-rank) one-sample tests were done to check with what degree of certainty it can be stated that the *real* parameter value is different from zero, given the data available. For the one-sample tests, a Bonferroni correction was applied on their significance level to determine the result h .

The outcome of the statistical tests shows that only the distribution of τ_2 cannot be considered normal, which agrees with observing the histograms in Fig. 19. The skewed distribution is likely to be caused by the lower bound of the parameter: the parameters are clustered around $\tau_2 = 0$. It was decided to keep the lower bound in place because negative values of τ_1 and τ_2 are unphysical. Furthermore, all four X -parameters were found to be significantly different from zero. This was found surprising, considering the low value of τ_2 . Also this result is likely to have been influenced by the lower parameter bound on τ_2 . If negative values of τ_2 had been allowed, its distribution is expected to be spread around the zero value, which would have reduced the significance of the results of the one-sample tests.

In general, the static parameters (a_1 and α^*) were easier to estimate than the dynamic parameters (τ_1 and τ_2). Two main reasons were found for this. First, the cost function surface tended to become irregular for large values of τ_1 and τ_2 , as is shown in Fig. 20. Since a gradient-based optimization algorithm was used, this reduced its effectiveness. This problem was mostly solved by setting θ_{ub} small enough for τ_1 and τ_2 . A second reason is that the sensitivity of the model output to changes in a_1 or α^* is greater than the sensitivity to changes in τ_1 or τ_2 , as is visualized in Fig. 18. Whereas the static parameters influence the output during almost all times of the maneuver, the dynamic parameters only influence C_L during the stall itself.

Finally, the estimated X -parameters were compared to those found in literature. Table 5 contains the results from several other studies. The dynamic parameters were made dimensionless for better comparison. Compared to the other aircraft types, a_1 and α^* are similar. However, τ_1 is about a factor 2 lower, and τ_2 is about a factor 5 lower compared to the values found for other aircraft. These differences in sensitivity between τ_1 and τ_2 are a direct result of the flight test data used for model identification. For the (benign) quasi-steady stalls performed in the flight tests, the hysteresis

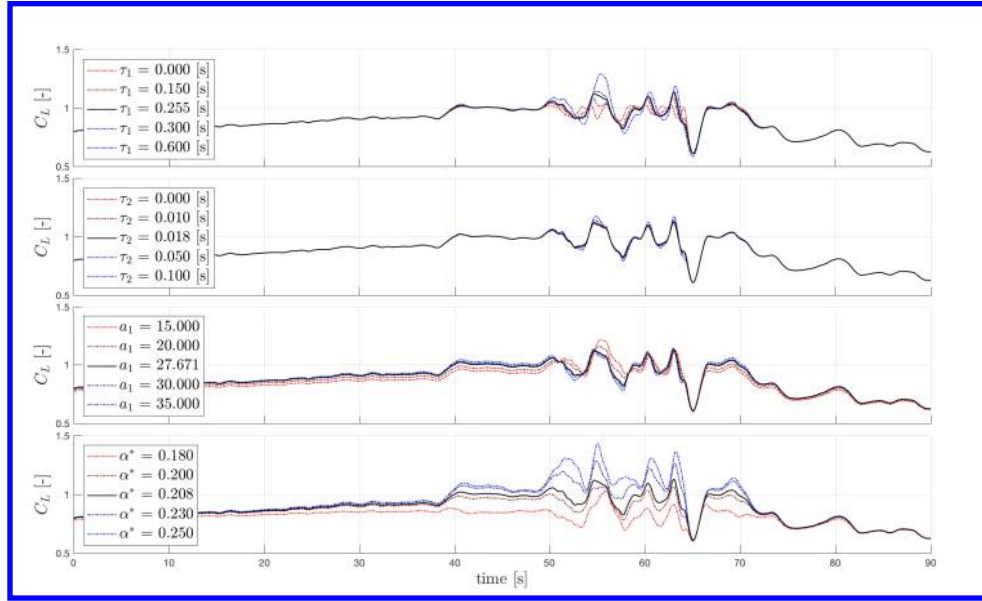


Fig. 18 Visualization of the sensitivity of the model output to the X -parameters. Whereas the static parameters influence almost the entire time-history, the dynamic parameters only affect the part where the aircraft actually stalls. Data set 3 (training) is shown.

Table 5 Comparison of estimated X -parameter values to literature, where for the Cessna Citation II: $\bar{c} = 2.06$ m, and $V_{\text{stall}} \approx 75$ m/s (approximate average from flight test data).

Parameter	Unit	Citation II	VFW-614[13]	C-160 [13]	AT-26[27]	Citation II[21]
$\tau_1 \frac{V}{\bar{c}}$	[-]	9.27	15.6	14.5	-	22.44
$\tau_2 \frac{V}{\bar{c}}$	[-]	0.64	4.45	3.46	-	13.18
a_1	[-]	27.67	15.00	25.70	25.00	25.87
α^*	[rad]	0.21	0.34	0.36	0.25	0.25

modeled by τ_{u_2} was hardly excited, while this is a key influential parameter for modelling deep stalls. The final column contains the results of previous research on the same aircraft [21]. This previous research used only data on quasi-static stall maneuvers; data with very little dynamic excitation. Furthermore, a very different approach to estimating τ_1 and τ_2 was used, namely through the stall buffet vibrations. Since in the current research a *direct* estimation of these parameters is used, instead of via the buffet, it is concluded that the newly obtained X -parameter values are more accurate. The static parameters are equivalent between the two Citation II stall models.

C. Aerodynamic Model Parameter Estimates

The results of the parameter estimation for the aerodynamic model equations can be found in Table 6. This table shows the estimated aerodynamic coefficient values, the standard deviations of the estimates over the data sets, and the outcomes of the same statistical tests as for the X -parameters. Matrix plots of the parameter estimates are shown in Fig. 21.

All parameters in Table 6 can be assumed to come from a normal distribution, except for $C_{m_{\delta_{eX}}}$. This agrees with the histograms for this parameter, found in Fig. 21(e). Furthermore, the one-sample statistical tests indicate that all model parameters are statistically different from zero, except for C_{D_0} . Because this is a bias parameter, it was still kept in the model. The fact that the drag bias is estimated to be small is surprising. It is hypothesized that this can be explained by the inclusion of $C_{D_{C_T}}$ into the model. Because the thrust is constant for the most part of the maneuvers, this term already accounts for a large part of the bias effect. This is problematic for the current model for C_D , but is expected to be solved once an updated engine model will become available.

In the lift, drag, and moment coefficient models, the bias and α parameters show clear correlations. Especially for the drag coefficient model, this can be regarded as a problem affecting the accuracy of the estimate (see Fig. 21(b), $\rho = -0.95$). These correlations are caused by the fact that the data segments used for model identification (intentionally) included considerable stretches of normal (unstalled) flight prior to and following the stall. During these segmented the angle of attack is approximately constant, i.e., in most cases for the largest part of the identification maneuvers. Hence,

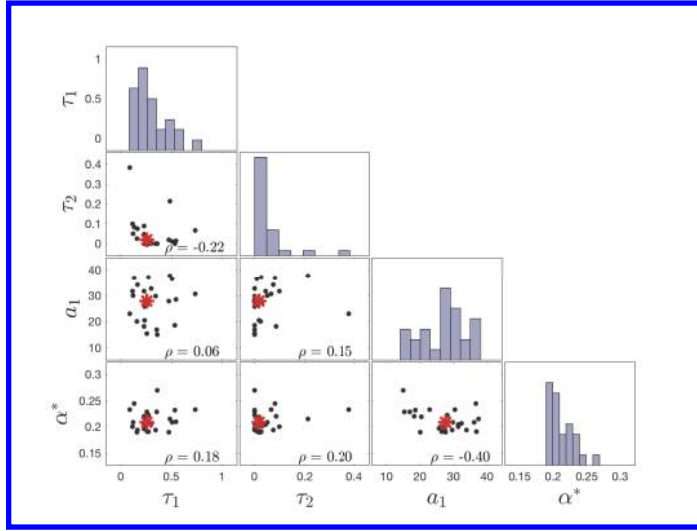


Fig. 19 Matrix plot of the estimated X-parameters

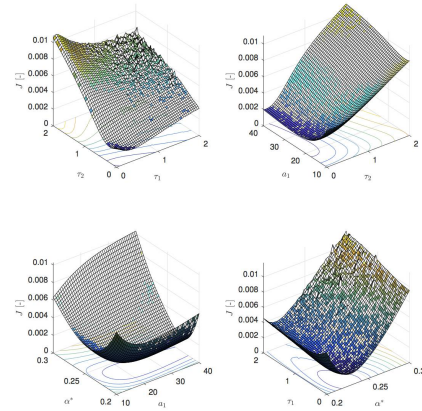


Fig. 20 Example of the cost function surface for the optimization of X-parameters. Note the roughness of the cost function surface for values of τ_1 and τ_2 larger than 1 s.

Table 6 Parameters estimation results and statistical tests. For the KS-test a significance of $\alpha = 0.1$ was used. $h = N$ means that the distribution can be assumed to be approximately normal, $h = \times$ when that this is not possible. The t -test and signed rank test were run with $\alpha = 0.01$ and a Bonferroni correction. For these tests, $h = *$ if the null hypothesis is rejected, and the parameter value is expected to differ from zero, $h = \circ$ otherwise.

Parameter		Results		KS-test		t -test		Signed rank	
Name	Unit	$\hat{\theta}$	$\sigma(\hat{\theta})$	p	h	p	h	p	h
C_{L0}	[-]	0.1758	0.0423	0.9135	N	0.0000	*	0.0000	*
$C_{L\alpha}$	[-]	4.6605	0.3965	0.9399	N	0.0000	*	0.0000	*
$C_{L\alpha^2}$	[-]	10.7753	3.8895	0.7665	N	0.0000	*	0.0000	*
C_{D0}	[-]	0.0046	0.0223	0.2648	N	0.2801	$\circ$	0.0837	$\circ$
$C_{D\alpha}$	[-]	0.2372	0.1443	0.5802	N	0.0000	*	0.0000	*
$C_{D\delta_e}$	[-]	-0.1857	0.0781	0.4845	N	0.0000	*	0.0000	*
$C_{D(1-X)}$	[-]	0.0732	0.0317	0.1101	N	0.0000	*	0.0000	*
C_{DC_T}	[-]	0.3788	0.0852	0.2042	N	0.0000	*	0.0000	*
C_{Y0}	[-]	0.0032	0.0035	0.8693	N	0.0000	*	0.0003	*
$C_{Y\beta}$	[-]	-0.5222	0.0682	0.9894	N	0.0000	*	0.0000	*
C_{Yp}	[-]	-0.5000	0.2244	0.9694	N	0.0000	*	0.0000	*
C_{Yr}	[-]	0.8971	0.4794	0.8368	N	0.0000	*	0.0000	*
$C_{Y\delta_a}$	[-]	-0.2932	0.0685	0.8481	N	0.0000	*	0.0000	*
C_{l0}	[-]	-0.0017	0.0001	0.8529	N	0.0000	*	0.0000	*
$C_{l\beta}$	[-]	-0.0454	0.0167	0.9067	N	0.0000	*	0.0000	*
C_{lp}	[-]	-0.1340	0.0620	0.7478	N	0.0000	*	0.0000	*
C_{lr}	[-]	0.1412	0.1287	0.5916	N	0.0000	*	0.0000	*
$C_{l\delta_a}$	[-]	-0.0853	0.0248	0.6586	N	0.0000	*	0.0000	*
C_{m0}	[-]	0.0183	0.0138	0.9811	N	0.0000	*	0.0000	*
$C_{m\alpha}$	[-]	-0.5683	0.1329	0.8496	N	0.0000	*	0.0000	*
$C_{m\delta_e X}$	[-]	-1.0230	0.1542	0.0736	$\times$	0.0000	*	0.0000	*
C_{mC_T}	[-]	0.1443	0.0498	0.8103	N	0.0000	*	0.0000	*
C_{n0}	[-]	0.0013	0.0000	0.3076	N	0.0000	*	0.0010	*
$C_{n\beta}$	[-]	0.0804	0.0093	0.5920	N	0.0000	*	0.0010	*
C_{nr}	[-]	-0.0496	0.0308	0.9029	N	0.0000	*	0.0010	*
$C_{n\delta_r}$	[-]	0.0492	0.0040	0.9162	N	0.0000	*	0.0010	*

for a large part α -related effects and bias would be interchangeable. However, removing either of these parameters in the C_L , C_D , or C_m -models was found to lead to a strong increase in modeling error, so both terms were still included in the final models.

D. Model Validation

Of the 34 recorded flight data sets, 7 were kept aside for model validation (80-20 split). Table 7 compares the MSE and coefficient of determination (R^2) for both the training and validation data. Time histories comparing model output

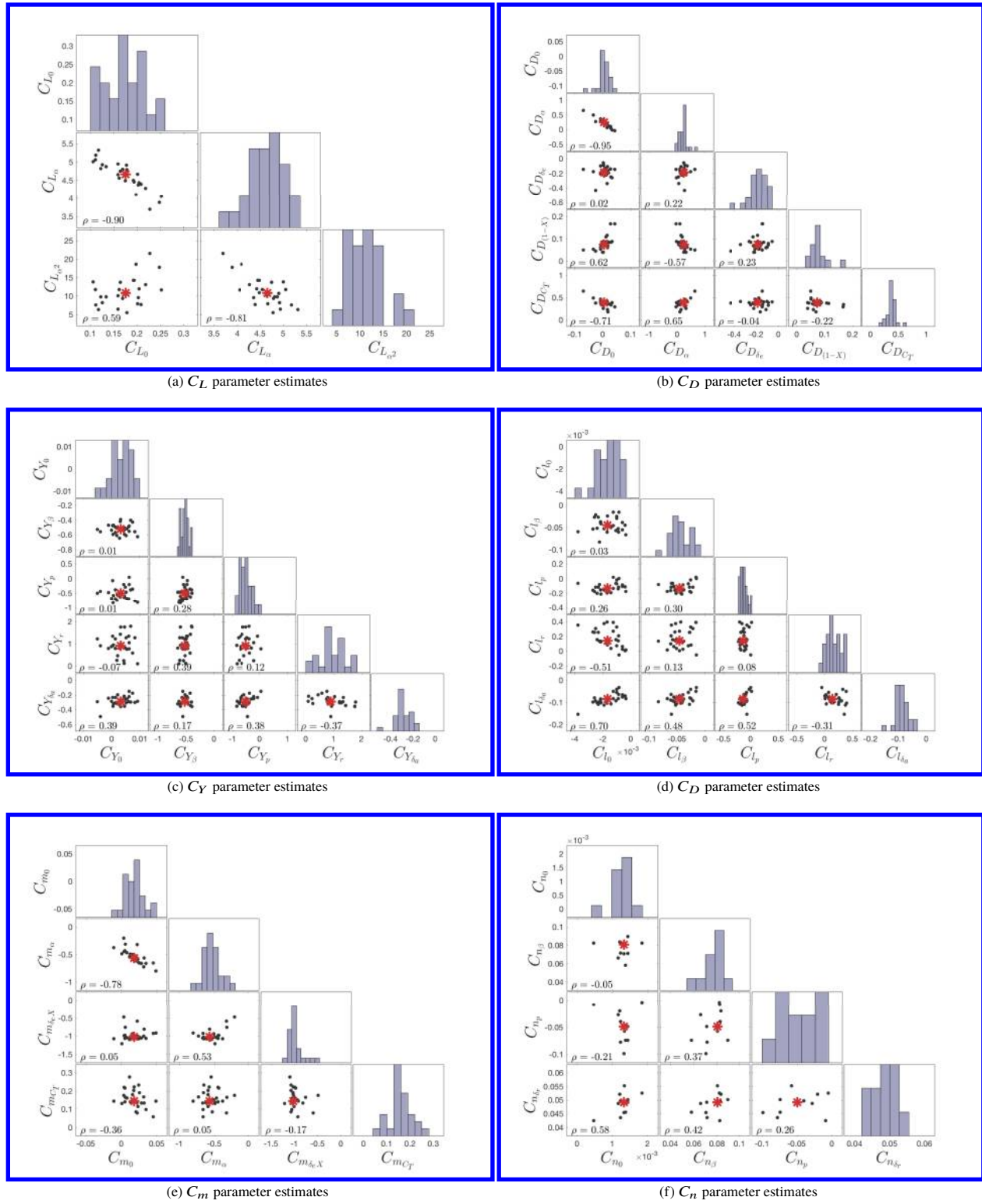


Fig. 21 Matrix plots of the parameters estimates. The diagonal plots are histograms that visualize an approximation of the parameter distributions. The off-diagonals are scatter plots of each of the parameter combinations. The red star marks the median value, and the values of the correlations are shown (ρ)

to measured data are presented for three selected validation data sets in Fig. 22. It was chosen to show the best, worst, and an average quality fit, measured by average normalized MSE over all coefficients.

Table 7 Properties of the estimated parameters of the aerodynamic model equations. MSE and R^2 are averages over the sets.

Coefficient	Training data (27 sets)				Validation data (7 sets)			
	MSE	R^2	$\min(R^2)$	$\max(R^2)$	MSE	R^2	$\min(R^2)$	$\max(R^2)$
C_L	1.65×10^{-3}	0.92	0.71	0.98	1.45×10^{-3}	0.91	0.77	0.96
C_D	1.01×10^{-4}	0.74	-1.47	0.97	6.72×10^{-5}	0.89	0.84	0.94
C_Y	4.68×10^{-5}	0.66	-0.67	0.91	4.55×10^{-5}	0.57	0.29	0.82
C_l	2.40×10^{-6}	0.54	-0.60	0.85	1.97×10^{-6}	0.47	0.08	0.92
C_m	9.93×10^{-5}	0.68	-0.39	0.92	9.87×10^{-5}	0.73	0.26	0.92
C_n	8.21×10^{-7}	0.49	-0.66	0.96	8.66×10^{-7}	0.12	-0.43	0.80

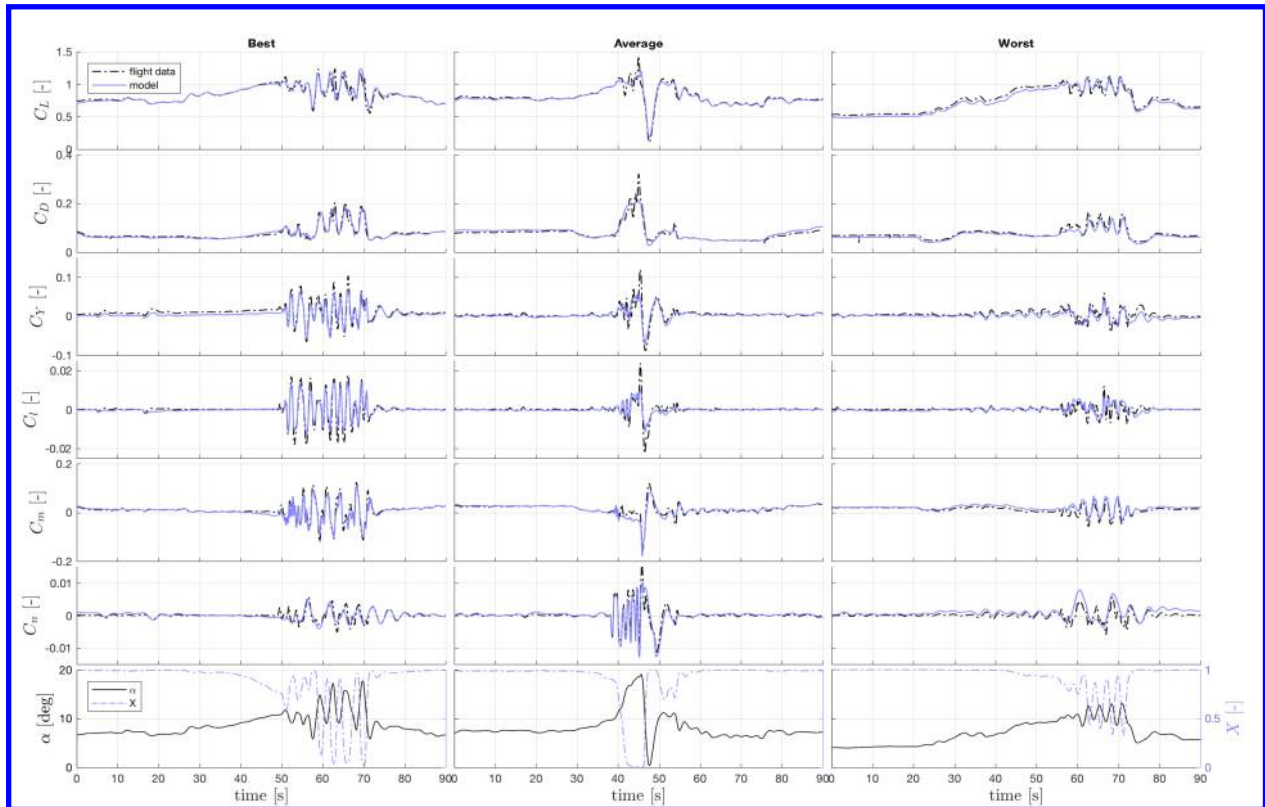


Fig. 22 Selected plots from the 7 validation data sets, showing model output versus measured force and moment coefficients. Left column contains data set 12, which has the lowest MSE, middle is set 5, and right is set 22. In the bottom plot, angle of attack and the calculated flow separation point variable are shown.

Fig. 22 shows good agreement of the fitted models with the flight test data, also during the stall. There are some stall-related effects that are not reflected by the model, for instance in the C_L and C_D time histories of the “average” plot in Fig. 22. However, the model re-aligns with the data once the aircraft recovers from the stall. An attempt was made to explain (part of) these features by considering X on left and right wings separately, as some authors in literature have proposed [14], but this did not lead to improved results for our data.

As Table 7 shows, the model quality is approximately the same on the training and validation sets, indicating that the model is neither under- nor overfitting. In some cases, the validation fit is even slightly better. This is caused by the random allocation of flight data maneuvers; some of them led to much lower errors. Because there are only 7 validation data sets, their average is influenced more significantly by the random allocation than is the case for the training data. It can be seen that for some sets, the worst R^2 scores are even negative. This holds often for the asymmetric models (lateral force, roll, yaw), which on average also show a lower quality fit than for the longitudinal model equations (lift,

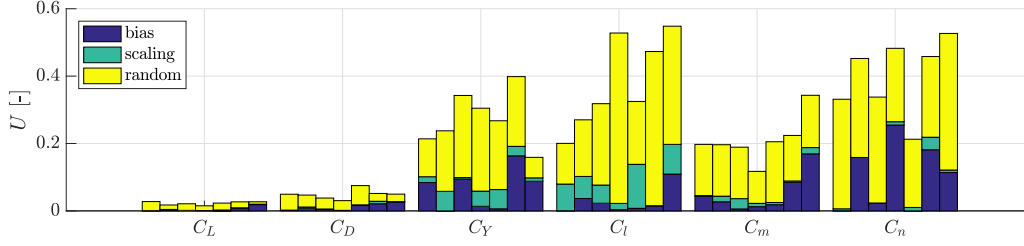


Fig. 23 Theil statistics of the 7 validation sets, grouped by coefficient model. Lower means better fit. Groups are sorted by average U -score over all coefficients, normalized by average score over that coefficient (from left to right: 12, 23, 25, 16, 5, 22, 7).

drag, pitch). In part, this can be explained by the definition of R^2 :

$$R^2 = 1 - \frac{\sum(\mathbf{y} - \hat{\mathbf{y}})^2}{\sum(\mathbf{y} - \mu_y)^2}. \quad (16)$$

In the extreme case that the measurement is (nearly) a constant, i.e., $\mathbf{y} \approx \mu_y$, then the denominator is (close to) zero. The more a time history is like a constant, the more any present errors will lead to a reduction in R^2 , which is the case for the asymmetric models, as can be seen in Fig. 22.

To gain more insight into the types of errors that occur in the model output, use is made of Theil statistics [28]. Theil's U -coefficient is a normalized metric for model quality, and ranges from 1 (worst case) to 0 (perfect fit). The error U can further be divided into a bias error U_{bias} , a variance error U_{var} , and a covariance error U_{cov} . These can be calculated as [28]:

$$U = \frac{\sqrt{\frac{1}{N} \sum(\mathbf{y} - \hat{\mathbf{y}})^2}}{\sqrt{\frac{1}{N} \sum \mathbf{y}^2} + \sqrt{\frac{1}{N} \sum(\hat{\mathbf{y}})^2}}, \quad (17)$$

$$U_{\text{bias}} = \frac{(\bar{\mathbf{y}} - \bar{\hat{\mathbf{y}}})^2}{\frac{1}{N} \sum(\mathbf{y} - \bar{\mathbf{y}})^2}, \quad (18)$$

$$U_{\text{var}} = \frac{(\sigma_y - \sigma_{\hat{y}})^2}{\frac{1}{N} \sum(\mathbf{y} - \bar{\mathbf{y}})^2}, \quad (19)$$

$$U_{\text{cov}} = \frac{2(1 - \rho_{y\hat{y}})\sigma_y\sigma_{\hat{y}}}{\frac{1}{N} \sum(\mathbf{y} - \bar{\mathbf{y}})^2}. \quad (20)$$

Here, σ and ρ indicate the standard deviation and the cross-correlation, respectively. Bias, variance, and covariance errors always add up to one due to their definition, so they can be used to divide U into fractions. U_{var} relates to scaling errors, while U_{cov} is caused by any errors that are uncorrelated to the signals. These statistics were computed for all the validation data sets, and the result is shown in Fig. 23. Within each coefficient group, the order of data sets is the same. They are ordered by average, normalized U over the six coefficients.

This figure confirms the previous observation that the longitudinal models have a higher quality fit. Furthermore, what can be seen is that generally speaking, the same trend can be seen for all coefficients; there is a relation between the fit quality for all models. In other words, some data sets (e.g., set 7) have large errors for all aerodynamic force coefficients, while for others the fit is overall much better.

Furthermore, it can be seen that the largest part of the error can be attributed to U_{cov} . This indicates that generally, the model parameters have been properly estimated. However, some maneuvers show that large parts of the error are caused by bias. The contribution of scaling errors to the total model error is small.

V. Discussion

This paper has presented a methodology for aerodynamic modeling of a Cessna Citation II, including the aerodynamic stall dynamics. It was applied to flight data, consisting of a series of quasi-steady wings-level and accelerated stalls,

Table 8 Comparison of the newly developed stall model, the previous stall model, and the normal flight envelope model of the Citation II in terms of the average mean squared error on the 7 validation data sets gathered for this research.

Model	Current MSE	Previous stall model[21] MSE	Difference	Normal flight envelope[26] MSE	Difference
C_L	1.45×10^{-3}	3.02×10^{-3}	+ 208%	3.35×10^{-2}	+2313%
C_D	6.72×10^{-5}	3.16×10^{-4}	+ 470%	1.38×10^{-3}	+2050%
C_Y	4.55×10^{-5}	n/a	n/a	2.47×10^{-4}	+ 543%
C_l	1.97×10^{-6}	n/a	n/a	1.30×10^{-5}	+ 662%
C_m	9.87×10^{-5}	2.89×10^{-3}	+2925%	7.84×10^{-4}	+ 795%
C_n	8.66×10^{-7}	n/a	n/a	1.61×10^{-5}	+1864%

that was specifically collected for this purpose at an altitude of 5,500 meters. Overall, the obtained stall model shows very good agreement with the validation data. The discussion of these results is presented in this section, and is split into three parts. First, the resulting model itself will be discussed. Several potential model issues are covered, and the simulated model outputs are compared to those of other Citation II models. After that, the inclusion of dynamic effects in the stall model is considered. Finally, an exploration of the required effort to expand the current model to the complete flight envelope is presented.

A. Stall Model

The model structure that was selected in this research is well-supported by the flight data. The multivariate orthogonal function modeling algorithm provided good guidance in identifying and selecting effective model terms. However, there are a few noteworthy results that are discussed here.

First of all is the fact that no terms relating to pitch rate q are selected, while these are commonly found in models for C_L , C_D , and especially C_m . The explanation for this is that the contribution of q to the lift, drag, or pitch moment is small in our (limited) flight test data set. The measured values of q in our maneuvers are close to zero when the aircraft is not in stall. Hence, identifying these relatively small effects due to q during stalled flight conditions – where strong nonlinearities and (random) disturbances are present – is challenging. A potential remedy for this is to gather new flight tests, where a dynamic maneuver such as a 3-2-1-1 on the elevator is added just before the aircraft stalls.

Another aspect of the model structure that is noteworthy is that no changes in control effectiveness for the aileron and rudder were identified. In qualitative comments, the test pilots stated that the control response of the aircraft did not change too much during the stall. However, a similar effect as for the elevator was expected. Perhaps the current approach of using X to add effects to control effectiveness terms is unsuitable for the aileron or rudder. Alternatively, it could be that not enough data is available to identify these effects. Dynamic maneuvers such as a 3-2-1-1 using aileron or rudder, performed right before the stall will give a stronger reference for comparison with the control performance during stall.

Next, the fact that the thrust coefficient is used as a regressor for the models for C_m and C_D is uncommon. As is discussed in Appendix C, these terms are corrections for mismatches in the engine model.

A final note on the model structure is regarding the use of X as a regressor. In this research, it has been shown that it is beneficial to use X as (part of) a term in the lift, drag, and pitch moment coefficient stall models. Not only was the well-known quadratic X -term in the Kirchhoff lift coefficient model [12, 15, 21] objectively selected by our model structure selection algorithm, but the explicit effect of X found for modeling the reduction in elevator effectiveness is a novel contribution. Hence, given new flight data and/or better mathematical transformations, other uses of X are thought to be possible. For instance, the static pitch stability term C_{m_α} is expected to change during stall. However, it must be kept in mind that X is only a single variable used to describe all the flow separation phenomena around the entire aircraft. This obviously is a simplification of reality. From a system identification perspective, a simple model is actually desired, as long as it still leads to accurate results. Nevertheless, it might not be possible to explain all relevant stall-related effects on the flight dynamics using X .

The identified stall model is significant improvement compared to the previous Citation II stall model as well as its normal flight envelope model when simulating stall maneuvers. Table 8 shows the difference in average MSE over the 7 gathered validation data sets for all three of these models. Lateral model terms were not included in the previous stall model due to a lack of the side slip measurement. The new stall model achieves a much lower error for all model outputs. It is interesting to see that the normal flight envelope model actually achieved a lower error on the pitch moment coefficient fit than the previous stall model.

Figure 24 visually compares the longitudinal model outputs of both these models, as well as the normal flight envelope Citation II model [26]. Fig. 25 shows the lateral model outputs of just the new stall model alongside the regular flight envelope model. The improvement in model fit is illustrated clearly, especially for the drag and pitch moments. The previous stall model can be seen to experience a significant bias during most part of the pitch moment time trace, which explains the large difference in MSE for the previous stall model. Also, the normal flight envelope model can be seen lack the reduction in lift during stall.

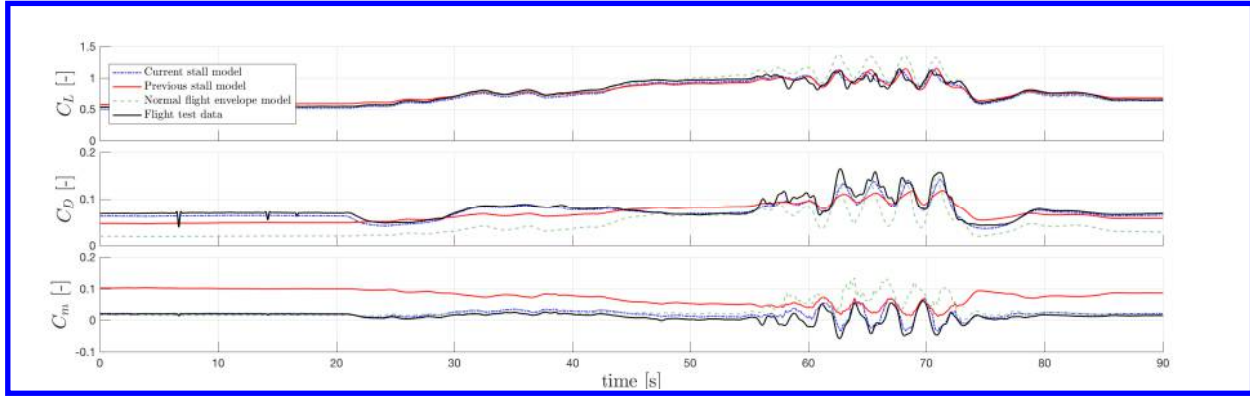


Fig. 24 Comparison of longitudinal outputs of the newly-obtained stall model to the previous version's. Additionally, outputs of the normal flight envelope model are shown, which is a validated dynamics model for the Citation II. Data set 22 (validation) is shown.

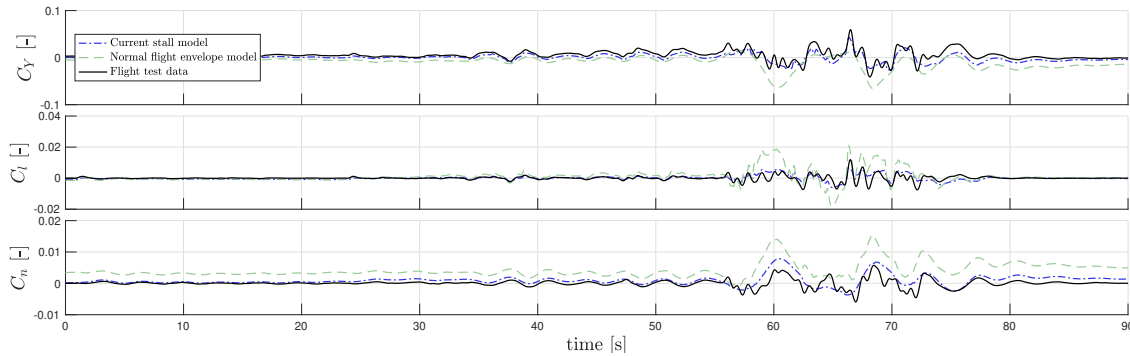


Fig. 25 Comparison of lateral outputs of the obtained stall model with the outputs of the validated normal flight envelope model for the Citation II. Data set 22 (validation) is shown.

B. Dynamic Stall Effects

One of the goals of this research was to include dynamic effects of stall into the model. These effects are parametrized by τ_1 and τ_2 . Including these terms in the model complicates both system identification as well as model evaluation, since an ODE needs to be solved to obtain X . Hence, it is sensible to review them: does the extra effort actually influence the model output in any significant way?

A simple test for this is to investigate what happens when either of these parameters is set equal to zero. The effect on the modeled lift coefficient of setting τ_1 equal to zero is shown in Fig. 26. It is shown that due to the time delay parametrized by τ_1 , the minimum value of X is larger, and therefore the maximum lift coefficient is higher than when τ_1 is neglected. As a result, the model provides a better match with the flight data. For τ_2 , the change in model output is negligible, with a less than 1% difference in MSE. This was to be expected, as the estimated value of τ_2 was already close to zero. This result shows that for the relatively benign quasi-steady stall maneuvers considered in this paper – i.e., as opposed to deep stall maneuvers – τ_2 can be neglected in modeling the stall dynamics of a Citation II type aircraft in the flight condition that was considered. Overall, these results show that while the hysteresis effect, parametrized by τ_2 ,

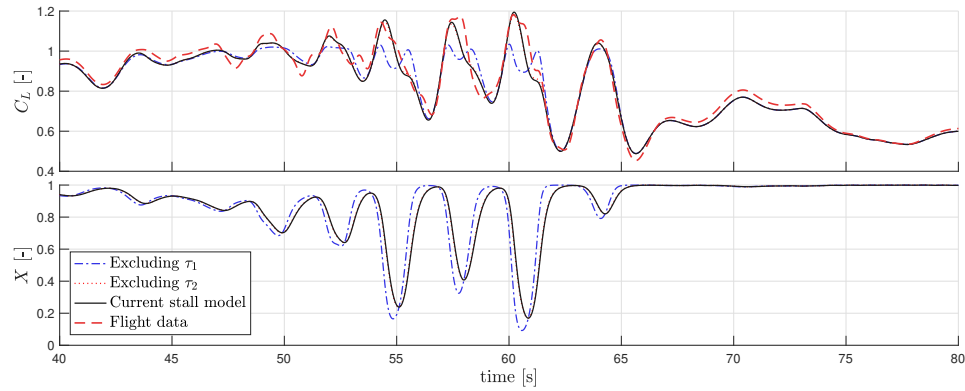


Fig. 26 Illustration of the effect on model output by setting either τ_1 or τ_2 equal to zero. It is shown that the maximum achieved lift coefficient is decreased when $\tau_1 = 0$. The effect of $\tau_2 = 0$ is so small that it is indistinguishable from the current stall model.

might not have a strong influence on model output for quasi-steady stall modeling, the dynamic time-delay effect by τ_1 is essential to ensure a realistic model output.

C. Towards a Full-Envelope Stall Model

The end goal of the research effort is to obtain a stall model that provides the most realistic simulator experience reasonably possible, for all flight conditions that are relevant for training scenarios. This model will then serve further research into what parts of such a stall model are relevant for effective training, and which are not. Before that is achieved, there are still many steps to take. The current stall model is only validated at an altitude of 5,500 meters and has so far only been used for a single pilot-in-the-loop experiment [19]. Regarding the research efforts that still are to be done, three major considerations are discussed here.

First of all, it is believed that the stall modeling methodology that was developed for this research will be a suitable tool for expanding the validated stall model envelope. It has been shown that Kirchhoff's methodology can capture the relevant nonlinearities in a stall maneuver, and that both the model structure and parameters can be identified. The same procedure as was described should be repeated for other selected altitudes, and other flap and gear configurations. The model parameters can then be made functions of flight condition. This is a common practice in aerodynamic modeling, and was also applied for the normal flight envelope model of the Citation II from [26]. The stall model structures should ideally be the same at every flight condition. If terms need to be added or removed for different flight conditions, this will pose an additional challenge.

Secondly, new flight tests will have to be conducted. These should be planned based on what flight conditions are to be required during training scenarios. If only few and specific scenarios are foreseen, one could perhaps do with fewer (costly) flight tests than what would be required for a full flight envelope model. For this research, about 30 maneuvers were done at the same flight condition. For model identification, this was concluded to be a good number. Also the piloting technique that was used for this research was generally found to be effective. A good addition to the flight test maneuvers would be the inclusion of dynamic excitation *before* the aircraft stalls. For example, a sequence of 3-2-1-1s or doublets on all three control axes could be performed. This should facilitate identification of changes in control responses. Important is that the aircraft is sufficiently dynamically excited both in- and outside the stall, and that high enough angles of attack are reached. No clear differences were found between wings-level and accelerated stall maneuvers in terms of identifiability. All control surfaces, including rudder, should be used during the stall. Without rudder inputs, yaw moment identification is not possible.

Finally, since the model's goal is to provide the best simulator experience possible, it will be vital to actually put a human in the loop fly it. This should be done even before expanding the stall model to the full flight envelope. Current model development was solely guided by quality of fit metrics such as MSE, and by visually inspecting plots. These metrics might overlook properties of the model that are vital for a good simulator experience, or conversely: they might severely penalize some errors that are irrelevant for the intended purpose. Hence, early simulator experiments as described in [19] will help in steering the research focus to the most relevant issues.

VI. Conclusion

Using a set of flight test data gathered specifically for this goal, a complete model of the quasi-steady stall dynamics of a Cessna Citation II is identified and validated around an altitude of 5,500 m, or 18,000 ft, in clean conditions. This flight data includes the use of control surfaces and dynamic excitation during recorded stalls, to also directly measure effects of degraded control effectiveness. Overall, the model output shows very good agreement with flight data, especially during the stalls. Kirchoff's theory of flow separation was used as modeling principle and was found to be effective in capturing the effects of stall on aircraft dynamics. Model structure selection was performed using multivariate orthogonal function modeling. Apart from modeling the effect on the lift, the flow separation point variable X was also selected by this model selection algorithm to capture the effect of stall on drag and elevator effectiveness. Especially the direct estimate of reduced elevator control effectiveness, supported by our flight test inputs during the stall, are novel result of this research. For our flight test data, X was not found to have an effect on the lateral dynamics (roll, yaw, lateral force) or aileron and rudder effectiveness. In conclusion, the model identification methodology that is developed in this research is considered to be well-suited for identifying stall models from flight test data. This methodology, as well as the obtained model for the Cessna Citation II, will serve as the basis for future research into improving stall models for training simulators as well as formalized guidelines for the required stall model fidelity.

References

- [1] Boeing Commercial Airplanes, "Report: Statistical Summary of Commercial Jet Airplane Accidents, Worldwide Operations, 1959-2015," Tech. rep., 2015. URL http://www.boeing.com/resources/boeingdotcom/company/about/_bca/pdf/statsum.pdf.
- [2] Lambregts, A. A., Nesemeier, G., Newman, R. L., and Wilborn, J. E., "Airplane Upsets: Old Problem, New Issues," *AIAA Modeling and Simulation Technologies Conference and Exhibit*, American Institute of Aeronautics and Astronautics, Reston, Virginia, 2008. <https://doi.org/10.2514/6.2008-6867>, URL <http://arc.aiaa.org/doi/10.2514/6.2008-6867>.
- [3] Crider, D. A., "The Need for Upset Recovery Training," *AIAA Modeling and Simulation Technologies Conference and Exhibit*, 2008. <https://doi.org/AIAA-2008-6864>.
- [4] Advani, S. K., and Field, J., "Upset Prevention and Recovery Training in Flight Simulators," *Proceedings of the AIAA Modeling and Simulation Technologies Conference, Portland (OR)*, 2011. <https://doi.org/10.2514/6.2011-6698>.
- [5] Advani, S. K., and Schroeder, J. A., "Global Implementation of Upset Prevention & Recovery Training," *Proceedings of the AIAA Modeling and Simulation Technologies Conference, San Diego (CA)*, 2016. <https://doi.org/10.2514/6.2016-1430>.
- [6] Federal Aviation Administration, "Advisory Circular 120-109A," , 2015. <https://doi.org/AFS-800AC91-97>, URL https://www.faa.gov/documentLibrary/media/Advisory/Circular/AC/_120-109A.pdf.
- [7] Chambers, J. R., and Grafton, S. B., "Aerodynamic Characteristics of Airplanes at High Angles of Attack," *Journal of Aircraft*, Vol. 31, No. 5, 1977, pp. 1109–1115.
- [8] Foster, J. V., Cunningham, K., and Fremaux, C. M., "Dynamics Modeling and Simulation of Large Transport Airplanes in Upset Conditions," *AIAA Guidance, Navigation, and Control Conference and Exhibit*, 2005. <https://doi.org/AIAA-2005-5933>, URL <http://arc.aiaa.org/doi/pdf/10.2514/6.2005-5933>.
- [9] Murch, A. M., and Foster, J. V., "Recent NASA Research on Aerodynamic Modeling of Post- Stall and Spin Dynamics of Large Transport Airplanes," *45th AIAA Aerospace Sciences Meeting & Exhibit*, 2007.
- [10] Abramov, N. B., Goman, M. G., and Khrabrov, A. N., "Aerodynamic Model of Transport Airplane in Extended Envelope for Simulation of Upset Recovery," *28th International Congress of the Aeronautical Sciences*, 2012. URL <http://www.icas-proceedings.net/ICAS2012/PAPERS/770.PDF>.
- [11] Groen, E. L., Ledegang, W., Field, J., Smaili, H., Roza, M., Fucke, L., Nooij, S. A. E., Goman, M. G., Mayrhofer, M., Zaichik, L., Grigoryev, M., and Biryukov, V., "SUPRA â€“ Enhanced Upset Recovery Simulation," *AIAA Modeling and Simulation Technologies Conference*, 2012.
- [12] Goman, M. G., and Khrabrov, A. N., "State-space representation of aerodynamic characteristics of an aircraft at high angles of attack," *Astrodynamics Conference*, 1992. <https://doi.org/10.2514/6.1992-4651>.
- [13] Fischenberg, D., "Identification of an Unsteady Aerodynamic Stall Model from Flight Test Data," *20th Atmospheric Flight Mechanics Conference*, 1995, pp. 138–146.

- [14] Fischenberg, D., and Jategaonkar, R. V., "Identification of Aircraft Stall Behavior from Flight Test Data," *RTO SCI Symposium on System Identification for Integrated Aircraft Development and Flight Testing*, 1998. URL <http://www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA361699{#}page=228>.
- [15] Dias, J. N., "Unsteady and Post-Stall Model Identification Using Dynamic Stall Maneuvers," *AIAA Atmospheric Flight Mechanics Conference*, 2015. <https://doi.org/10.2514/6.2015-2705>.
- [16] Da Ronch, A., Vallespin, D., Ghoreyshi, M., and Badcock, K. J., "Evaluation of Dynamic Derivatives Using Computational Fluid Dynamics," *AIAA Journal*, Vol. 50, No. 2, 2012, pp. 470–484. <https://doi.org/10.2514/1.J051304>.
- [17] Nie, Y., Van Kampen, E., Chu, Q. P., Kier, T. M., and Looye, G., "Geometry Based Quick Aircraft Modeling Method for Upset Recovery Applications," *AIAA Modeling and Simulation Technologies Conference*, , No. January, 2015, pp. 1–16. <https://doi.org/10.2514/6.2015-2034>, URL <http://arc.aiaa.org/doi/abs/10.2514/6.2015-2034>.
- [18] Teng, T. T., Zhang, T. S., Liu, S., and Grant, P. R., "Representative Post-Stall Modelling of T-tailed Regional Jets and Turboprops for Upset Recovery Training," *AIAA Modeling and Simulation Technologies Conference*, , No. January, 2015.
- [19] Smets, S. C. E., de Visser, C. C., and Pool, D. M., "Subjective Noticeability of Variations in Quasi-Steady Aerodynamic Stall Dynamics," *Proceedings of the AIAA Modeling and Simulation Technologies conference, San Diego (CA)*, 2019. <https://doi.org/10.2514/6.2019-1485>.
- [20] Schroeder, J. A., Bürki-Cohen, J., Shikany, D. A., Gingras, D. R., and Desrochers, P., "An Evaluation of Several Stall Models for Commercial Transport Training," *AIAA Modeling and Simulation Technology Conference*, 2014. <https://doi.org/doi:10.2514/6.2014-1002>.
- [21] van Horssen, L. J., Pool, D. M., and de Visser, C. C., "Aerodynamic Stall and Buffet Modeling for the Cessna Citation II based on Flight Test Data," *AIAA SciTech 2018, session: Modeling and Simulation of Air Vehicle Dynamics I*, 2018.
- [22] Morelli, E. A., Cunningham, K., and Hill, M. A., "Global Aerodynamic Modeling for Stall/Upset Recovery Training Using Efficient Piloted Flight Test Techniques," 2013.
- [23] Laban, M., "On-Line Aircraft Aerodynamic Model Identification," Ph.D. thesis, Delft University of Technology, 1994.
- [24] Norsett, S., Wanner, G., and Hairer, E., *Solving Ordinary Differential Equations I*, Vol. 8, 1993. <https://doi.org/10.1007/978-3-540-78862-1>, URL <http://www.springerlink.com/index/10.1007/978-3-540-78862-1>.
- [25] Ljung, L., *System Identification: Theory for the User (Second Edition)*, 1999. [https://doi.org/10.1016/S0005-1098\(01\)00214-X](https://doi.org/10.1016/S0005-1098(01)00214-X).
- [26] van den Hoek, M. A., de Visser, C. C., and Pool, D. M., "Identification of a Cessna Citation II Model Based on Flight Test Data," *4th CEAS Specialist Conference on Guidance, Navigation & Control*, 2017.
- [27] Dias, J. N., "High Angle of Attack Model Identification with Compressibility Effects," *AIAA Atmospheric Flight Mechanics Conference*, , No. January, 2015, pp. 1–15. <https://doi.org/10.2514/6.2015-1477>, URL <http://arc.aiaa.org/doi/10.2514/6.2015-1477>.
- [28] Bonner, M. S., and Gingras, D. R., "Evaluation of the Navy's F/A-18A-D powered-approach aerodynamics model," *Modeling and Simulation Technologies Conference*, 1997. <https://doi.org/10.2514/6.1997-3667>, URL <http://arc.aiaa.org/doi/10.2514/6.1997-3667>.
- [29] Leis, J. R., and Kramer, M. A., "The Simultaneous Solution and Sensitivity Analysis of Systems Described by Ordinary Differential Equations," *ACM Transactions on Mathematical Software (TOMS)*, Vol. 14, No. 1, 1988, pp. 45–60. <https://doi.org/10.1145/42288.46156>, URL <http://dl.acm.org/citation.cfm?id=46156>.

Appendices

A. Parameter Sensitivity of Solution of ODE

This Appendix presents the method used for computing the cost function gradient $\partial J(\theta, x)/\partial \theta$. Consider the example model structure from Eq. (7), which results in the following parameter vector: $\theta = [\tau_1 \ \tau_2 \ a_1 \ \alpha^* \ C_{L_0} \ C_{L_\alpha} \ C_{L_q}]^T$. The

relevant equations are repeated for clarity:

$$\begin{aligned} J(\theta, x) &= \frac{1}{N} \left(C_L - \hat{C}_L(\theta, x) \right)^\top \left(C_L - \hat{C}_L(\theta, x) \right) , \\ \hat{C}_L(\theta, x) &= C_{L_0} + C_{L_\alpha} \left(\frac{1+\sqrt{X}}{2} \right)^2 \alpha + C_{L_q} \frac{q\bar{c}}{V} , \\ \tau_1 \frac{dX}{dt} + X &= \frac{1}{2} (1 - \tanh [a_1 (\alpha - \tau_2 \dot{\alpha} - \alpha^*)]) . \end{aligned}$$

The first four parameters in θ influence the cost function via X , the rest does not. As a result, the application of the chain rule requires a separate treatment of X -parameters and others:

$$\frac{\partial J(\theta, x)}{\partial \theta_i} = \begin{cases} \frac{\partial J}{\partial \hat{C}_L} \frac{\partial \hat{C}_L}{\partial X} \frac{\partial X}{\partial \theta_i} & \text{when } \theta_i \in \{\tau_1 \ \tau_2 \ a_1 \ \alpha^*\} \\ \frac{\partial J}{\partial \hat{C}_L} \frac{\partial \hat{C}_L}{\partial \theta_i} & \text{when } \theta_i \in \{C_{L_0} \ C_{L_\alpha} \ C_{L_q}\} \end{cases} . \quad (21)$$

All the partial derivative terms in this equation are relatively straightforward to obtain using algebra, except for $\partial X / \partial \theta_i$. When using finite difference methods to compute the parameter sensitivity of a solution of an ODE (i.e., the flow separation point X), numerical issues can arise [24]. To avoid this, a different method is used instead [29]. First, Eq. (2) is rewritten to the following form:

$$\frac{d}{dt} X(t, x, \theta) = \frac{1}{\tau_1} \left(-X + \frac{1}{2} - \frac{1}{2} \tanh [a_1 (\alpha - \tau_2 \dot{\alpha} - \alpha^*)] \right) \quad \text{with} \quad X(0, x, \theta) = X_0(x, \theta) . \quad (22)$$

X_0 is the static mapping from α to X , i.e., without considering dynamic effects. Eq. (22) can be written as:

$$\frac{dX(t, x, \theta)}{dt} = G(X, t, x, \theta) . \quad (23)$$

Next, $\partial X / \partial \theta$, which is the partial derivative that we want to obtain, is defined as:

$$\frac{\partial X(t, x, \theta)}{\partial \theta} = S(t, x, \theta) . \quad (24)$$

When one takes the partial derivative of Eq. (23) with respect to θ , the following result is obtained:

$$\frac{\partial}{\partial \theta} \frac{d}{dt} X(t, x, \theta) = \frac{\partial G(X, t, x, \theta)}{\partial X} \frac{\partial X(t, x, \theta)}{\partial \theta} + \frac{\partial G(X, t, x, \theta)}{\partial \theta} . \quad (25)$$

Substituting Eq. (24) yields:

$$\frac{d}{dt} S(t, x, \theta) = \frac{\partial G(X, t, x, \theta)}{\partial X} S(t, x, \theta) + \frac{\partial G(X, t, x, \theta)}{\partial \theta} \quad \text{with} \quad S(0, x, \theta) = \mathbf{0} . \quad (26)$$

The terms $\partial G / \partial X$ and $\partial G / \partial \theta$ can be derived from Eq. (22). A second ODE has been obtained, which can be solved numerically. The solution can be substituted into Eq. (21) to complete the chain rule.

B. Results of Initial X -parameter Optimization

An important assumption is that the dependence of the optimal set of X -parameters on the selected C_L model structure is mild. When the optimization of the X -parameters is done using the initial model structure (as in Eq. (7)), then the results are as shown in Table 9. The parameters are clearly in the same order of magnitude, the assumption is concluded to hold.

Table 9 Results of estimating the X -parameters using the initial model structure for C_L .

Parameter		Results				
Name	Unit	$\hat{\theta}$	Diff. w. final	θ_{lb}	θ_{ub}	$\sigma(\hat{\theta})$
τ_1	[s]	0.4623	+81.5%	0.001	2.00	0.4227
τ_2	[s]	0.0212	+20.5%	0.000	2.00	0.1049
a_1	[-]	22.1161	-20.1%	15.000	40.00	4.7636
α^*	[rad]	0.2286	+9.7%	0.100	0.35	0.0200

C. Explanation for Including C_T As Regressor

The models for drag and pitch moment coefficient both contain the thrust coefficient C_T as a regressor term, since it was found that this has a strong beneficial effect on model quality. Since this might appear surprising, this Appendix presents the likely reason behind this.

First, recall that the aerodynamic force and moment coefficients, such as C_D and C_m , cannot be measured directly, but are determined based on other measurements. For this, it is required to have accurate knowledge of the engine thrust. However, also the thrust cannot be measured directly. An engine model is used for computing the thrust based on signals that *are* measured, such as thrust lever setting, and atmospheric conditions. The Citation II is equipped with two Pratt & Whitney JT15D turbofans. The models that were used for these are included in the Delft University Aircraft Simulation Model and Analysis Tool (DASMAT).

It is clear that any errors in this engine model propagate into errors into the “measured” aerodynamic force and moment coefficients, and thus lead to changes in the identified model parameters. If the error in thrust would have been constant, only the bias term of the model would change. However, the thrust was not constant during the recorded stall maneuvers, which also led to a changing error in the engine thrust. As a result, the bias term cannot correct for this error. Since the thrust coefficient C_T is computed based on the same engine model, it contains the same error. Hence, because it is correlated with a pattern in the data that cannot be otherwise explained, it is beneficial to the model quality to include C_T as a regressor.

It is recommended to do research and improve the accuracy of the engine model. If that is successful, these thrust-related model terms should no longer be necessary, which would be an improvement to model simplicity.

D. Individual Model Term Contributions of C_Y , C_l , and C_n

The contributions of the individual model terms were shown in the main part of the paper only for the longitudinal model outputs (i.e., C_L , C_D , C_m). The same plots for the lateral model outputs are presented here in the Appendices, in Figures 27, 28, and 29.

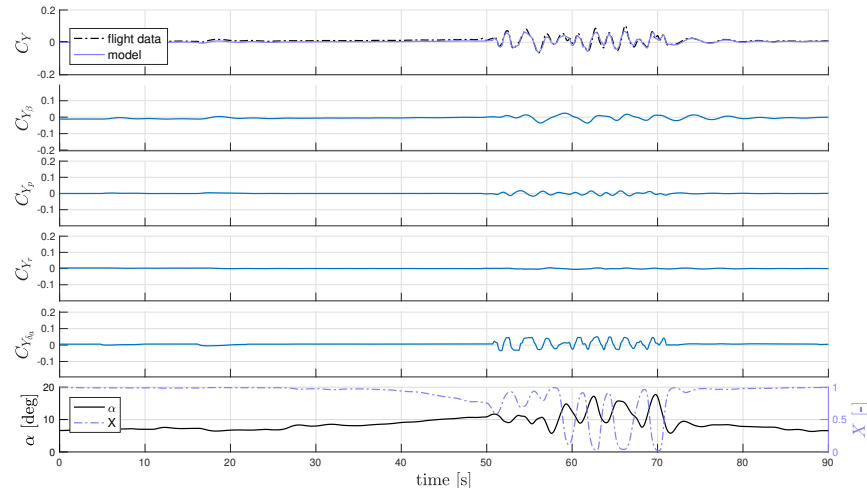


Fig. 27 Example of the model term contributions of the lateral force model, example shown is data set 12 (validation).

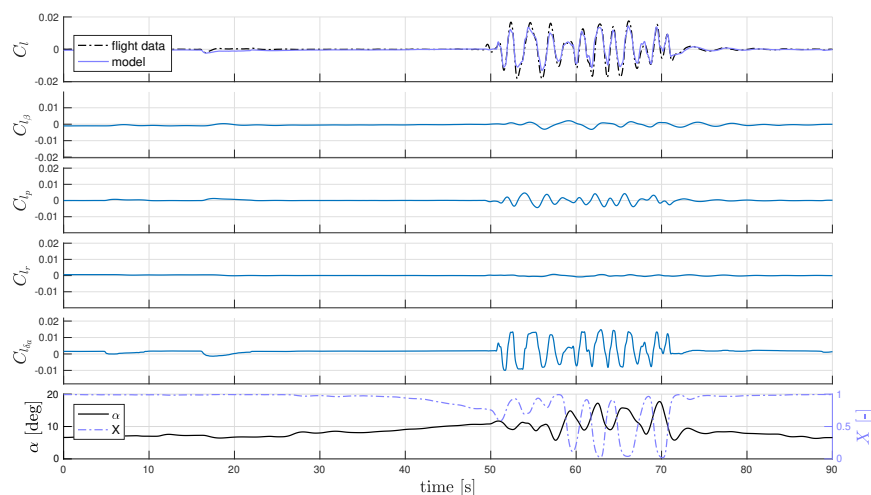


Fig. 28 Example of the model term contributions of the roll moment model, example shown is data set 12 (validation).

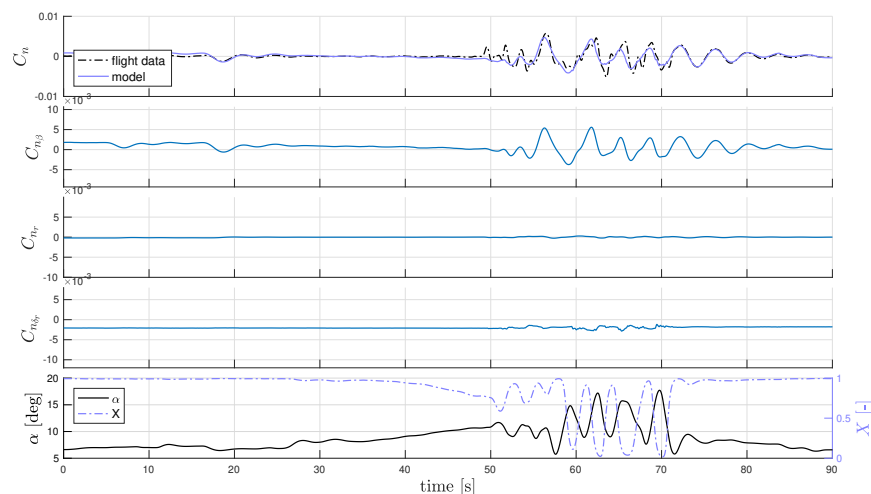


Fig. 29 Example of the model term contributions of the yaw moment model, example shown is data set 12 (validation).