


# RESIDUAL LEARNING FRAMEWORK FOR NONLINEAR AEROELASTIC–FLIGHT-DYNAMICS CONTROL-ORIENTED MODELING OF VERY FLEXIBLE AIRCRAFT

Divya Sanghi[1] & Carlos E. S. Cesnik[2]

[1, 2]Department of Aerospace Engineering, University of Michigan, Ann Arbor, MI 48109

## Abstract

Very flexible aircraft (VFA) exhibit strong aeroelastic coupling and geometric nonlinearities due to their high-aspect-ratio wings (HARWs) and lightweight structures. This paper presents a residual-learning extension of the bottom-to-top control-oriented modeling framework for VFA, providing a low-order, real-time-capable model that can capture the response across small and large deformations. The study considers a representative HARW transport aircraft model, and the University of Michigan's Nonlinear Aeroelastic Simulation Toolbox (UM/NAST) is used to conduct the nonlinear simulations that provide the higher-order reference. Three low-order modeling approaches are compared. The first is a semi-analytical flexible-aircraft model linearized about the trim condition. The second is a direct nonlinear VFA model, in which a neural network trained on the UM/NAST data predicts the complete nonlinear aerodynamic load. The third is the proposed residual model, in which the linearized aerodynamic prediction is retained as an explicit baseline and a neural network predicts only the nonlinear correction. The models are evaluated through aerodynamic-load prediction and maneuver-response simulations for single-surface and coupled doublet inputs. The direct nonlinear VFA model captures the large-amplitude aerodynamic behavior but can degrade the near-trim response. The residual model provides more balanced behavior, remaining close to the linearized baseline for small perturbations while capturing the nonlinear behavior at large deformations. Both data-driven models that capture the nonlinear aerodynamics can run faster than real time for the simulated maneuvers, supporting their use in control-oriented applications.

**Keywords:** Reduced-order modeling, residual learning, nonlinear aeroelasticity, very flexible aircraft, control-oriented modeling

## 1. Introduction

Next-generation aircraft designs are increasingly adopting longer and lighter wings to improve aerodynamic efficiency and reduce fuel consumption. However, these performance benefits come with challenges due to increased structural flexibility. Such configurations, referred to as very flexible aircraft (VFA), are characterized by low natural frequencies and undergo large deflections during maneuvers or gust encounters. These effects lead to pronounced geometric nonlinearities and strong coupling between nonlinear rigid-body motion and aeroelastic response, which creates additional challenges for both modeling and control [1].

Traditional six-degree-of-freedom (6-DOF) rigid-body flight-dynamics models, typically augmented with empirical corrections for structural flexibility [2], are inadequate for accurately capturing the nonlinear aeroelastic behavior of VFA. These methods fail to represent critical phenomena such as flutter, dynamic coupling, and large deformation effects [3, 4]. While high-fidelity models can capture these interactions, they often involve hundreds or thousands of states [5, 6, 7], making them impractical for real-time control applications.

Therefore, low-order flexible-aircraft models that capture the nonlinear aeroelastic behavior are needed for control-oriented modeling of VFA. One promising approach is to augment the 6-DOF rigid-body equations with nonlinear functions that represent the key structural and aerodynamic effects. This bottom-to-top (B2T) modeling approach [8, 9] preserves the physical meaning of states and parameters while allowing nonlinear corrections derived from high-fidelity simulations or experimental data to be embedded into reduced-order frameworks.

Early efforts included mean-axes formulations [3, 4], which decouple structural and rigid-body motion by defining an inertial reference at the deformed center of mass. More commonly, body-fixed reference frames have been used to accommodate large elastic deflections and varying mass properties [10, 11]. Different approaches have been developed, including the strain-based reduction approach by Su and Cesnik [11], Taylor-series expansions by Da Ronch et al. [12], and modal condensation techniques based on intrinsic beam theory by Palacios and Cea [13]. While these methods simplify the structural dynamics and preserve interpretability, the aerodynamic modeling is often either overly simplified or too computationally expensive for control-oriented use.

Duan et al. [8, 9] developed a semi-analytical, low-order modeling framework tailored for control-oriented applications involving very flexible aircraft. The framework comprises two formulations. The first is a linearized flexible-aircraft (FA) model evaluated about a trim condition [9], which is computationally efficient and physically interpretable but is not expected to remain accurate under the large deflections that motivate VFA modeling. The second is a fully nonlinear bottom-to-top model, referred to as B2T-VFA, which augments the 6-DOF rigid-body equations with nonlinear structural and aerodynamic terms approximated using feedforward neural networks trained on simulation data from a high-fidelity strain-based nonlinear aeroelastic framework implemented in the University of Michigan's Nonlinear Aeroelastic Simulation Toolbox (UM/NAST) [6, 7]. The resulting B2T-VFA model preserves geometric nonlinearities, elastic coupling, and inertial effects, while achieving substantial model-order reduction. Demonstrated on a high-aspect-ratio-wing (HARW) transport aircraft model, it showed good agreement with full-order UM/NAST results for coupled elevator–aileron input cases while requiring significantly fewer flexible states. Although the B2T-VFA framework performed well in modeling the nonlinear flight dynamics, the published results show reduced accuracy in the low-deflection regime and in the lateral-directional response, so challenges remain in improving the robustness and generalization of the neural-network predictions of the nonlinear terms.

To address these challenges, previous studies have explored residual learning as a strategy to improve the prediction of nonlinear terms in data-driven aerodynamic models, especially those based on neural networks. Kou and Zhang [14] proposed a layered reduced-order model that adds a nonlinear neural-network correction to a linear surrogate for aerodynamic-force prediction, improving the accuracy of limit-cycle-oscillation predictions and the agreement with high-fidelity simulations. He et al. [15] introduced a deep neural network that combines multi-fidelity aerodynamic data, using separate networks for the low-fidelity outputs and for the residual between low- and high-fidelity data, improving lift and drag predictions. Similarly, Zhang et al. [16] developed a multi-fidelity deep neural-network surrogate for aerodynamic shape optimization by learning linear and nonlinear correlations between coarse- and fine-grid simulations, achieving high accuracy at significantly reduced cost.

More recently, Lee et al. [17] used a residual neural-network architecture for hypersonic aerodynamic-coefficient increments, with separate blocks used to encode geometry and flow-condition effects before the final prediction. Demo et al. [18] applied DeepONet to learn corrections to projection-based reduced-order models, improving pressure-prediction fidelity. Yan et al. [19] employed neural-network residual Kriging for flight-load surrogate modeling with improved accuracy and efficiency. Feldwisch [20] augmented a linear unsteady-aerodynamic state-space model with a recurrent neural network for the nonlinear deviation from the linear response, improving load predictions in the transonic regime where shock motion and flow separation drive the aerodynamic nonlinearity. These studies highlight the potential of combining residual learning with physics-based models to improve the generalization of aerodynamic predictions.

Building on these insights and the framework of Duan et al. [8, 9], the objective of this study is to enhance the control-oriented modeling of very flexible aircraft through a residual-learning approach for the nonlinear aerodynamic terms. As introduced above, the linearized FA model is computationally

efficient and interpretable near trim but loses accuracy under large deflections, whereas the direct B2T-VFA model has nonlinear authority at larger amplitudes but does not reduce to the FA model in the small-amplitude limit. The residual model proposed in this study combines the interpretable FA baseline with a learned nonlinear correction.

Rather than training a neural network to represent the full nonlinear aerodynamic loads directly, the proposed approach retains the FA prediction as an explicit baseline and trains the network to learn only the nonlinear aerodynamic residual relative to that baseline. The terminology used throughout the paper is as follows. FA denotes the linearized flexible-aircraft model evaluated about trim. Direct VFA denotes the B2T-VFA model in which the neural network learns the complete nonlinear aerodynamic load vector. VFA-FA denotes the residual model in which the FA aerodynamic prediction is evaluated first and a neural network learns the nonlinear correction. The residual model preserves the near-trim accuracy and interpretability of the linear baseline while capturing the nonlinear behavior at large deflections. The three models are compared against full-order nonlinear UM/NAST simulation results for a representative HARW transport aircraft model. The study first presents the VFA-FA residual model and assesses its ability to remain accurate across both the linear regime, where the FA baseline is most reliable, and the nonlinear regime, where the learned aerodynamic correction is required. It then investigates model behavior across a broader set of maneuvers than previously considered [8], spanning single-surface and coupled inputs from low to high amplitude as well as cases outside the training set. The remainder of the paper is organized as follows. Section 2. summarizes the B2T-VFA equations and the residual-learning formulation. Section 3. describes the aircraft model and Sec. 4. presents the results. The final section summarizes the main conclusions and future work.

## 2. Modeling Framework

This section reviews the B2T modeling framework introduced by Duan et al. [9, 8] and details the aerodynamic residual formulation investigated in this study. The original B2T-VFA framework augments a low-order 6-DOF rigid-body dynamics model with nonlinear structural and aerodynamic terms approximated using neural networks, trained on high-fidelity aeroelastic simulations. This serves as the foundation for the residual-learning-based improvements introduced in Sec. 2.3. While the present work focuses on the aerodynamic terms since they are the main source of the nonlinearity, the framework can be extended to structural terms as well.

### 2.1 B2T-VFA Nonlinear Formulation Overview

The final form of the nonlinear equations of motion governing both the rigid-body and flexible states in the B2T-VFA model, adapted from Duan et al. [8], is given by

$$
E\frac{d}{dt}\begin{Bmatrix} V \\ \omega \\ \dot{\eta} \\ \eta \end{Bmatrix} = -\begin{Bmatrix} m\,\tilde{\omega}V + m\dot{\eta}^{\top}\dfrac{\partial^2 r_C(\eta)}{\partial \eta^2}\dot{\eta} + 2m\tilde{\omega}\dfrac{\partial r_C(\eta)}{\partial \eta}\dot{\eta} + m\tilde{\omega}^2 r_C(\eta) \\ m\tilde{r}_C(\eta)\,(\tilde{\omega}V) + \dot{\eta}^{\top}H_{\text{hess}}(\eta)\dot{\eta} + 2\,(J_0(\eta)\dot{\eta})\,\omega + \tilde{\omega}I_B(\eta)\omega \\ \left(m\dfrac{\partial r_C(\eta)}{\partial \eta}\right)^{\top}(\tilde{\omega}V) + \dfrac{1}{2}\dot{\eta}^{\top}\dfrac{\partial M}{\partial \eta}\dot{\eta} + 2\,(\nu(\eta)\dot{\eta})\,\omega - \omega^{\top}J_0(\eta)\omega + K\eta \\ -\dot{\eta} \end{Bmatrix} + \begin{Bmatrix} F_g + F_{\text{aero}} + F_{\text{thrust}} \\ M_g + M_{\text{aero}} + M_{\text{thrust}} \\ Q_g + Q_{\text{aero}} + Q_{\text{thrust}} \\ 0 \end{Bmatrix} \tag{1}
$$

where

$$
E = \begin{bmatrix} m & -m\tilde{r}_C(\eta) & m\frac{\partial r_C(\eta)}{\partial \eta} & 0 \\ m\tilde{r}_C(\eta) & I_B(\eta) & H_{\text{jac}}(\eta) & 0 \\ m\left(\frac{\partial r_C(\eta)}{\partial \eta}\right)^{\top} & H_{\text{jac}}^{\top}(\eta) & M(\eta) & 0 \\ 0 & 0 & 0 & I_n \end{bmatrix}. \tag{2}
$$

Equations (1)–(2), together with the translational and rotational kinematics equations for position and orientation (represented by Euler angles), form the complete system-level dynamics of the B2T-VFA framework. The detailed derivation of these equations, including the definitions of the auxiliary structural matrices such as $H_{\text{jac}}$, $H_{\text{hess}}$, $J_0$, and $\nu$, can be found in Duan et al. [8]. The state variables include the body-frame translational velocity $V = [u, v, w]^\top$, angular velocity $\omega = [p, q, r]^\top$, and the flexible modal amplitudes $\eta$ and their time derivatives $\dot{\eta}$.

The selected flexible state vector $\eta$ is designed to be a low-dimensional representation of the dominant strain-based modes, rather than the full set used in the full-order nonlinear model. Duan et al. [8] demonstrated that retaining a limited number of dominant modes is sufficient to accurately capture the key aeroelastic behavior, enabling substantial model-order reduction. The number of retained flexible modes, $n$, directly determines the total number of flexible states, $2n$, which includes both $\eta$ and $\dot{\eta}$. Additionally, $m$ denotes the total mass of the aircraft, and the remaining terms represent nonlinear structural, thrust, and aerodynamic contributions. In the direct B2T-VFA formulation, the nonlinear structural and aerodynamic terms are approximated using shallow feedforward neural networks trained on simulation data from the full-order nonlinear UM/NAST model. While this study uses UM/NAST as the higher-order nonlinear reference, the training data could in principle come from any high-order physics-based model or from experimental results.

Since the present work focuses only on the aerodynamic terms of the B2T-VFA formulation, the relevant terms are the translational force $F_{\text{aero}}$, the rotational moment $M_{\text{aero}}$, and the generalized flexible force $Q_{\text{aero}}$. These terms are treated as functions of all the states and the control-surface deflection inputs $\delta$. Therefore, the input dimension for the aerodynamic terms is $2n+6+n_\delta$, where $n$ is the number of retained flexible modes and $n_\delta$ is the number of control-surface inputs. The corresponding input and output dimensions are summarized in Table 1. This neural-network representation of the complete aerodynamic load vector is referred to in this paper as the direct approach, or Direct VFA. In the remainder of this paper, the structural and thrust terms are left in the original B2T-VFA form, while the aerodynamic terms in Table 1 are replaced by the residual formulation.

Table 1 – Summary of aerodynamic neural-network models for B2T-VFA nonlinear terms using the direct approach (adapted from [8]).

| Term | Description | Input dimension | Output dimension |
|---|---|---|---|
| $F_{\text{aero}}(\eta, \dot{\eta}, V, \omega, \delta)$ | Translational | $2n+6+n_\delta$ | 3 |
| $M_{\text{aero}}(\eta, \dot{\eta}, V, \omega, \delta)$ | Rotational | $2n+6+n_\delta$ | 3 |
| $Q_{\text{aero}}(\eta, \dot{\eta}, V, \omega, \delta)$ | Flexible | $2n+6+n_\delta$ | $n$ |

## 2.2 B2T-FA Linearized Aerodynamic Baseline

The bottom-to-top flexible-aircraft (B2T-FA) formulation introduced by Duan et al. [9] provides the linear aerodynamic baseline used in this work. Unlike the direct B2T-VFA approach described above, B2T-FA is a semi-analytical, control-oriented formulation based on a model linearized about a trimmed condition. This makes the model computationally efficient and physically interpretable, but also limits its validity to small perturbations around that trim condition. Therefore, B2T-FA is not expected to capture the nonlinear aerodynamic increments that arise when the aircraft undergoes large deformations. In this work, the B2T-FA aerodynamic prediction is used as the baseline for the residual-learning formulation introduced in the next section. The FA baseline is evaluated about a trimmed operating point and is given by

$$y_{\text{aero}}^{\text{FA}} = \bar{q}\left(C_0 + C_\beta \Delta\beta + C_\eta \Delta\eta + C_{\dot{\eta}} \dot{\eta} + C_\delta \Delta\delta\right), \tag{3}$$

where $\bar{q}$ is the trim dynamic pressure, $\Delta(\cdot)$ denotes perturbations from trim, $\beta = [V^\top,\ \omega^\top]^\top$ contains the rigid-body velocity vector $V$ and angular-rate vector $\omega$ entering the aerodynamic model, $\delta$ contains the control-surface deflections, and the $C$ matrices are aerodynamic sensitivity coefficients.

Equation (3) follows the structure of the FA control-oriented formulation [9]. Because these linearized sensitivities are evaluated about the trim condition, the FA prediction is the trusted aerodynamic refer-

ence in the small-deflection regime. The residual-learning formulation in the next section uses this FA prediction as the baseline and trains a neural network only on the remaining nonlinear aerodynamic contribution, allowing the residual model to retain the near-trim accuracy of the FA baseline.

## 2.3 Residual-Learning Formulation

Both data-driven aerodynamic models take as input the vector $x(t)$, which comprises the flexible states, the rigid-body flight-dynamics variables, and the control-surface commands rather than the full dynamic state of Eq. (1). This input vector is defined as

$$x(t) = \left[\eta^\top, \dot{\eta}^\top, V_\infty, \alpha_B, \beta_B, p, q, r, \delta_e, \delta_a, \delta_r\right]^\top, \tag{4}$$

where $V_\infty$, $\alpha_B$, and $\beta_B$ are the flight speed, body-axis angle of attack, and body-axis sideslip angle, and $\delta_e$, $\delta_a$, and $\delta_r$ are the grouped elevator, aileron, and rudder commands.
The Direct VFA aerodynamic model approximates the full nonlinear aerodynamic load with a neural network,

$$y_{\text{aero}}^{\text{direct}}(t) = y_{\text{aero}}^{\text{NN}}(x(t)). \tag{5}$$

While this direct approach can be effective given sufficient training data, it is purely data-driven with no physics-based anchoring and offers limited interpretability. In particular, Duan et al. [8] reported that the Direct VFA model does not reproduce the low-deflection response as accurately as the FA model. To address this, the present work adopts a residual-learning framework in which the network is trained to predict only the difference between the high-order nonlinear loads and the linearized FA baseline of Sec. 2.2, so that the physics-based baseline is preserved and only the unknown nonlinear contribution is learned. The aerodynamic loads are thus decomposed into the FA baseline and a learned residual,

$$y_{\text{aero}}^{\text{residual}}(t) = y_{\text{aero}}^{\text{FA}}(t) + \Delta y_{\text{aero}}^{\text{NN}}(x(t)). \tag{6}$$

This residual model, which combines the FA baseline with a learned nonlinear correction, is referred to in this paper as VFA-FA. The training residual is formed from the difference between the full-order nonlinear loads and the FA baseline. Compared with the direct approach, this formulation requires the linearized baseline to be evaluated before training, but in return it narrows the output range of each network and localizes learning to the regions where the nonlinearities are strongest. This localized structure also allows the training residual to be weighted according to its magnitude, so that small differences between the full-order and FA predictions near trim do not dominate the training of the nonlinear correction. The weighted VFA-FA training residual is defined as

$$\Delta y_{\text{aero}}^{\text{train}}(t) = \lambda_{\text{res}}(t)\left[y_{\text{aero}}^{\text{Full-order}}(t) - y_{\text{aero}}^{\text{FA}}(t)\right], \tag{7}$$

where $\lambda_{\text{res}} \in [0,1]$ is a smooth residual-amplitude weight that increases with the normalized difference between the full-order and FA aerodynamic predictions. Thus, the learned correction is encouraged to remain small in the near-trim regime, where the FA baseline is already accurate and the residual may be comparable to numerical or fitting noise, while approaching the full nonlinear residual where the FA model misses the nonlinear response. This weighting is applied only in defining the training residual. In deployment, the VFA-FA model evaluates the FA baseline and the learned correction according to Eq. (6). A further advantage of this decomposition is its modularity. Setting $\Delta y_{\text{aero}}^{\text{NN}} = 0$ in Eq. (6) recovers the pure FA model, so the physics-based baseline and the learned nonlinear correction stay separate and individually interpretable rather than fused into a single model. In both models, the output contains the six rigid-body and $n$ generalized modal aerodynamic loads. Table 2 summarizes the network inputs and outputs for the two models.

## 2.4 Network Training and Evaluation

Both data-driven aerodynamic models are trained on the same database of UM/NAST aerodynamic loads for the HARW aircraft test case, whose flight condition and excitation maneuvers are described in Sec. 3.. The database contains $16{,}335$ samples, split into training, validation, and test subsets.

Table 2 – Neural-network inputs and outputs for the Direct VFA and VFA-FA models.

| **Model** | **Neural-network input** | **Training output** | **Input dim.** | **Output dim.** |
|---|---|---|---|---|
| Direct VFA | $x(t)$ from Eq. (4). | $y_{\text{aero}}^{\text{Full-order}}$ | $2n+9$ | $n+6$ |
| VFA-FA | Same $x(t)$ as Direct VFA. | $\Delta y_{\text{aero}}^{\text{train}}$ from Eq. (7) | $2n+9$ | $n+6$ |

The Direct VFA and VFA-FA models share the same architecture and optimizer and differ only in their training output.
Each aerodynamic model is a feedforward network with a single hidden layer of $200$ neurons, hyperbolic-tangent activation, and a linear output. The layer sizes are $17 \to 200 \to 10$ for $n = 4$ and $29 \to 200 \to 16$ for $n = 10$. The networks are trained with the Levenberg–Marquardt algorithm, which converges efficiently at this size. The remaining settings are listed in Table 3.

Table 3 – Neural-network architecture and training settings for the Direct VFA and VFA-FA models.

| **Setting** | **Value** |
|---|---|
| Network | One hidden layer, $200$ tanh neurons |
| Normalization | Min–max scaling of inputs and outputs to $[-1, 1]$ |
| Optimizer | Levenberg–Marquardt, $30$ iterations |
| Samples (train/val/test) | $16{,}335$ $(11{,}435/2{,}450/2{,}450)$ |

The trained models are evaluated in two steps. In open-loop testing, each aerodynamic model infers the aerodynamic loads along a fixed UM/NAST state and control trajectory, and the predictions are compared with the nonlinear reference. This isolates the accuracy of the data-driven aerodynamics without feedback. In closed-loop testing, the reduced-order B2T model is integrated forward in time under the prescribed control doublets with the data-driven aerodynamics in the loop, and the resulting aircraft response is compared with the corresponding full-order nonlinear UM/NAST simulation. Both steps are needed because high open-loop accuracy is necessary but not sufficient. Small load errors can accumulate through the coupled rigid–elastic dynamics and perturb the trimmed equilibrium, so the closed-loop response is a more challenging test. The error in both steps is quantified by the root-mean-square error (RMSE) against the UM/NAST reference, defined in Sec. 4.. Both the training maneuvers and a set of held-out maneuvers are evaluated to distinguish in-sample accuracy from generalization.

## 3. Aircraft Model and Simulation Cases

This section describes the commercial transport high-aspect-ratio-wing (HARW) aircraft model, the flight condition, and the simulation cases used in this study. The nonlinear reference model of this test case was developed in Refs. [21, 22]. The configuration is representative of a long-range wide-body commercial transport aircraft with a wingspan of 70 m. The full-order nonlinear model contains 1085 states, consisting of 13 rigid-body states and 1072 flexible states, namely the $n = 536$ strain-based modal amplitudes and their rates. This higher-order reference is denoted full-order nonlinear ($n = 536$) hereafter. The original aircraft model includes multiple control effectors. For the present B2T comparison, the control inputs are grouped so that the left and right elevator sections deflect symmetrically, the left and right ailerons deflect anti-symmetrically, and the rudder deflects independently. The resulting control-surface vector is

$$\delta = \begin{bmatrix} \delta_e & \delta_a & \delta_r \end{bmatrix}^\top, \tag{8}$$

where $\delta_e$, $\delta_a$, and $\delta_r$ denote the grouped elevator, aileron, and rudder deflections, respectively.
The comparisons use 10-s doublet simulations at the trimmed flight condition $h = 8484.11$ m and $V = 253.98$ m/s, corresponding to approximately $M_\infty = 0.83$. Following the investigations of Duan et al. [8], the B2T models are evaluated with either $n = 4$ or $n = 10$ retained flexible modes. The $n = 4$ model retains the dominant low-frequency flexible modes, including the first out-of-plane bending

modes of the left and right wings. The $n = 10$ model additionally retains higher-frequency flexible modes, including additional bending and torsional content. The left and right wing-root out-of-plane bending curvatures are used later as structural response channels in the closed-loop comparisons. The neural-network aerodynamic models of Sec. 2.4 are trained on a database of UM/NAST simulations at this flight condition, generated from control-surface doublet inputs with amplitudes ranging from 1 to 5 degrees. The doublet maneuvers considered in the closed-loop comparison are listed in Table 4. Each maneuver is defined by the grouped elevator, aileron, and rudder doublet amplitudes applied about the trimmed control values. Figure 1 shows an example of the doublet profile used for control-surface inputs. The commands are introduced at $t = 1$ s, reversed at $t = 4$ s, and returned to trim at $t = 7$ s. The 1-deg maneuvers are used to assess the near-trim linear regime, while the 3- to 5-deg maneuvers are used to assess nonlinear response recovery.

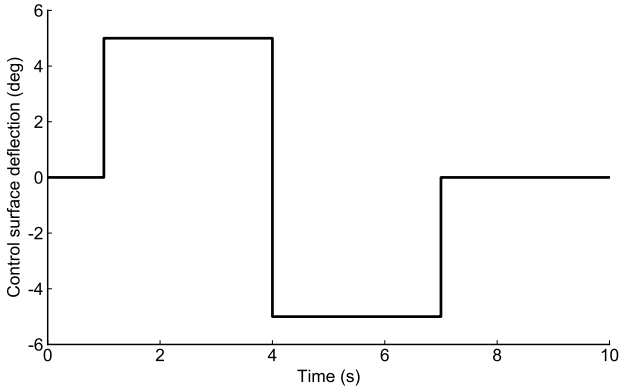



Figure 1 – Representative doublet control input applied about the trimmed state.

Table 4 – Doublet maneuvers considered for testing.

| Maneuver | Elevator | Aileron | Rudder |
|---|---|---|---|
| 1-deg elevator | $1^\circ$ | $0^\circ$ | $0^\circ$ |
| 1-deg aileron | $0^\circ$ | $1^\circ$ | $0^\circ$ |
| 1-deg coupled | $1^\circ$ | $1^\circ$ | $1^\circ$ |
| 3-deg elevator | $3^\circ$ | $0^\circ$ | $0^\circ$ |
| 3-deg coupled | $3^\circ$ | $3^\circ$ | $3^\circ$ |
| 4-deg coupled | $4^\circ$ | $4^\circ$ | $4^\circ$ |
| 5-deg coupled | $5^\circ$ | $5^\circ$ | $5^\circ$ |

To emphasize the degree of nonlinearity and how it varies with the control-surface amplitude, Table 5 reports the peak vertical wingtip deflection extracted from the full-order nonlinear reference histories, expressed as a percentage of the semispan and measured relative to the trimmed shape. Figure 2 presents the corresponding incremental wingtip deflection time histories for representative amplitudes, showing the geometrically nonlinear response as the command amplitude increases.

Table 5 – Maximum wingtip deflection relative to trim.

| Maneuver | Max. tip deflection [% semispan] |
|---|---|
| 1-deg elevator | 2.59 |
| 1-deg coupled | 2.78 |
| 3-deg elevator | 12.34 |
| 3-deg coupled | 12.57 |
| 4-deg coupled | 18.27 |
| 5-deg coupled | 23.35 |

The 3-deg maneuvers already exceed 12% semispan, and the coupled 5-deg maneuver reaches

23.4% semispan. These deformation levels are large enough for geometric nonlinearities to become important, which explains why the linear FA load error grows rapidly in the high-amplitude cases.

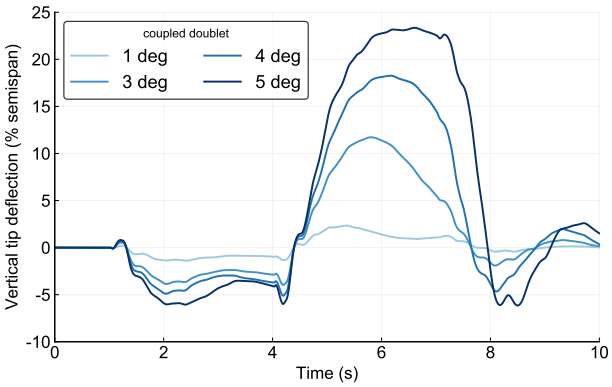



Figure 2 – Wingtip deflection for varying control-surface amplitudes.

## 4. Results

This section compares the FA baseline, Direct VFA, and VFA-FA residual models with the full-order nonlinear ($n = 536$) reference responses. The comparison is performed in two steps. The open-loop aerodynamic-load prediction evaluates whether the model reproduces the nonlinear aerodynamic coefficients along a fixed full-order nonlinear trajectory, while the maneuver-response simulations evaluate whether the resulting reduced-order model reproduces the coupled aircraft response when integrated forward in time. The $n = 4$ model is treated as the primary control-oriented configuration, where it runs faster than real time, while the $n = 10$ model is used to examine the mode-count sensitivity of the comparison. The closed-loop error metric aggregates the error over the 11 response channels ($u$, $v$, $w$, $p$, $q$, $r$, $\phi$, $\theta$, $\psi$, and the left and right wing-root curvature channels $\kappa_L$ and $\kappa_R$) and over all time samples of the 10-s maneuver. For clarity, the response figures show a representative subset of these channels, namely the three body velocities, the three Euler angles, and the two wing curvatures.

### 4.1 Aerodynamic Load Prediction

The Direct VFA and VFA-FA models differ first in the aerodynamic quantity assigned to the neural network. Direct VFA learns the full nonlinear aerodynamic load. VFA-FA keeps the FA prediction explicit and learns only the residual that remains after the FA baseline is removed from the full nonlinear load. This distinction is important because the direct output remains large even near trim, whereas the residual output is small in the linear regime and becomes significant only at higher deflections.

The Direct VFA and VFA-FA networks both fit their assigned coefficient outputs closely over the 16,335-sample aerodynamic database. On the held-out test split, both networks reach a mean channel-wise $R^2$ above 0.9999.

Figures 3 and 4 show the six rigid-body aerodynamic load histories for the 1- and 5-deg coupled maneuvers, comparing the full-order nonlinear reference loads with the Direct VFA, VFA-FA, and FA predictions. The loads are plotted as increments about the trimmed values. For the 1-deg maneuver the FA baseline stays close to the nonlinear loads, confirming that the linear model is adequate near trim, and the Direct VFA and VFA-FA predictions both reproduce the nonlinear loads. As the amplitude increases, the gap between the FA baseline and the nonlinear loads widens and becomes dominant for the 5-deg maneuver, where the FA baseline predicts a large rolling moment that the nonlinear reference does not develop, while the Direct VFA and VFA-FA predictions continue to track the reference. This gap between the nonlinear loads and the FA baseline is the nonlinear aerodynamic residual, which increases with maneuver amplitude and is what the residual-learning formulation is intended to capture.

The VFA-FA prediction in Figs. 3 and 4 reflects the size of this residual. Near trim the residual is small, so VFA-FA stays close to the FA baseline in the regime where the linear model is already adequate.

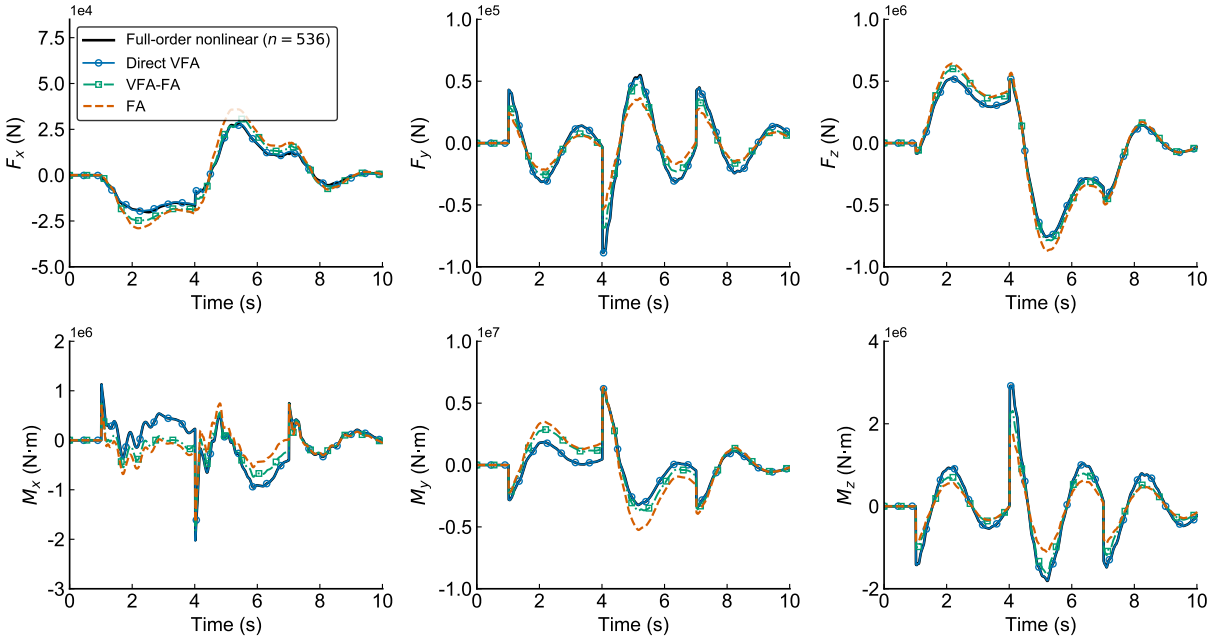



Figure 3 – Open-loop rigid-body aerodynamic load increments relative to trim for the 1-deg coupled maneuver ($n = 4$).

As the geometric nonlinearity develops, the residual increases until it becomes a dominant part of the load in the high-amplitude coupled maneuver, where VFA-FA recovers the nonlinear loads that the FA baseline misses. Learning only this residual keeps the nonlinear aerodynamic content as a distinct, physically meaningful quantity, in contrast with the Direct VFA model, which embeds the linear and nonlinear contributions together.

Because VFA-FA forms its prediction by adding the learned correction to the FA baseline, its accuracy is tied to the quality of that baseline. Near trim, where the FA prediction already represents almost the entire load, the residual is small and the model inherits the baseline accuracy directly. As the amplitude increases, the FA baseline omits a larger share of the load and the network must supply a correspondingly larger correction. A more accurate linearized baseline would therefore leave a smaller residual for the network to learn, which makes the learning problem better conditioned without altering the structure of the formulation.

## 4.2 Maneuver Response Prediction

This subsection evaluates the maneuver response in closed loop, with responses of FA, Direct VFA, and VFA-FA compared against the full-order nonlinear reference. The closed-loop comparison separates the three models by maneuver amplitude. Table 6 summarizes the aggregate RMSE for all selected maneuvers. Because this aggregate combines channels of different physical type into a single number dominated by the largest-magnitude channels, small differences between the two data-driven models are interpreted together with the response time histories rather than from the aggregate value alone. The small-amplitude maneuvers confirm that the FA baseline is the most accurate near trim. For $n = 4$, the mean RMSE over the three 1-deg maneuvers is 0.079 for FA, 0.107 for VFA-FA, and 0.240 for Direct VFA. This result is consistent with the purpose of the residual formulation, in which the model should remain close to FA when the nonlinear corrections are weak.

At larger amplitudes, the FA error increases rapidly, as expected. For the high-amplitude maneuvers from 3 to 5 deg, the $n = 4$ mean RMSE is reduced from 4.928 for FA to 2.310 for Direct VFA and 1.646 for VFA-FA, so the residual model is the most accurate $n = 4$ model in exactly the regime where the linear baseline fails. With $n = 10$, the two data-driven models are comparable at high amplitude and both reduce the FA error by a large factor. Direct VFA is lower on the 4- and 5-deg maneuvers, while VFA-FA gives the lowest error on the two 3-deg maneuvers. The advantage of the residual model over the FA baseline therefore holds at the higher amplitudes for both mode counts. At $n = 4$

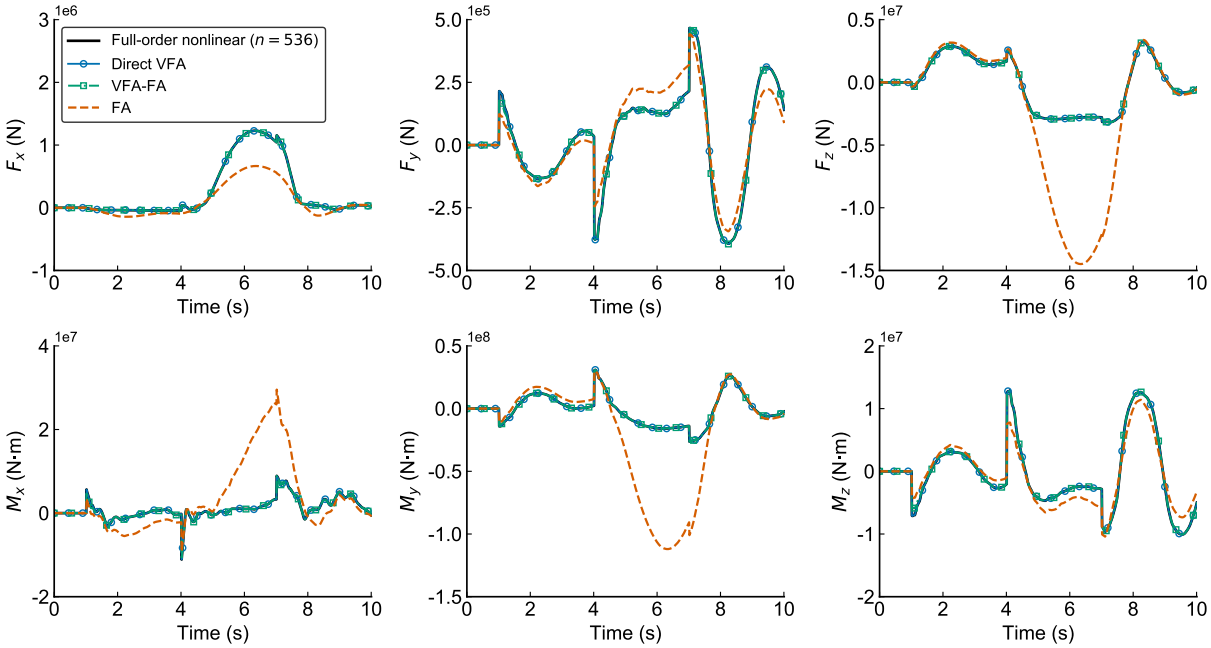



Figure 4 – Open-loop rigid-body aerodynamic load increments relative to trim for the 5-deg coupled maneuver ($n = 4$).

the residual model is also the more accurate of the two learned models, whereas at $n = 10$ Direct VFA is the more accurate at the highest amplitudes (4 and 5 deg). In both cases the residual model additionally retains the near-trim linear limit that Direct VFA does not capture well.

Table 6 – Closed-loop aggregate RMSE against the full-order nonlinear ($n = 536$) reference.

| **Maneuver** | $n = 4$ | | | $n = 10$ | | |
|---|---|---|---|---|---|---|
| | **FA** | **Direct VFA** | **VFA-FA** | **FA** | **Direct VFA** | **VFA-FA** |
| 1-deg elevator | 0.081 | 0.288 | 0.109 | 0.112 | 0.125 | 0.101 |
| 1-deg aileron | 0.021 | 0.102 | 0.021 | 0.023 | 0.087 | 0.027 |
| 1-deg coupled | 0.136 | 0.330 | 0.191 | 0.152 | 0.168 | 0.165 |
| 3-deg elevator | 2.121 | 1.511 | 0.628 | 2.258 | 0.492 | 0.392 |
| 3-deg coupled | 2.188 | 0.982 | 1.226 | 2.317 | 0.839 | 0.738 |
| 4-deg coupled | 4.954 | 2.577 | 2.065 | 5.110 | 0.705 | 1.458 |
| 5-deg coupled | 10.451 | 4.170 | 2.663 | 10.622 | 1.609 | 2.618 |

Figures 5 and 6 show the response channels for the 1-deg aileron and coupled 1-deg maneuvers. For the 1-deg aileron maneuver, the $n = 4$ VFA-FA RMSE is 0.021, essentially identical to FA and substantially lower than the Direct VFA value of 0.102. For the coupled 1-deg maneuver, the FA baseline remains the best $n = 4$ model, but VFA-FA is closer to FA than Direct VFA. These trends are consistent with the residual output, which is small near trim so that the model relies primarily on the FA component. The degradation of Direct VFA in this regime arises because the near-trim aerodynamic load is dominated by the self-consistent linearized response, and the direct network must reproduce this entire load from a finite-accuracy fit. The resulting fitting errors, although negligible relative to the full-amplitude loads, are comparable in magnitude to the small perturbation loads that drive the near-trim response, so they perturb the trimmed closed-loop dynamics. The residual model avoids this by retaining the analytic FA baseline exactly and adding only a learned correction that is near zero in the linear regime, which is why its small-amplitude error stays close to the FA value rather than inheriting the network fitting error.

Figure 7 shows the 1-deg elevator maneuver, the small-amplitude case in which the Direct VFA degradation is most pronounced. Its $n = 4$ RMSE of 0.288 is more than three times the FA value of 0.081, whereas the VFA-FA residual stays close to the baseline at 0.109. This maneuver most

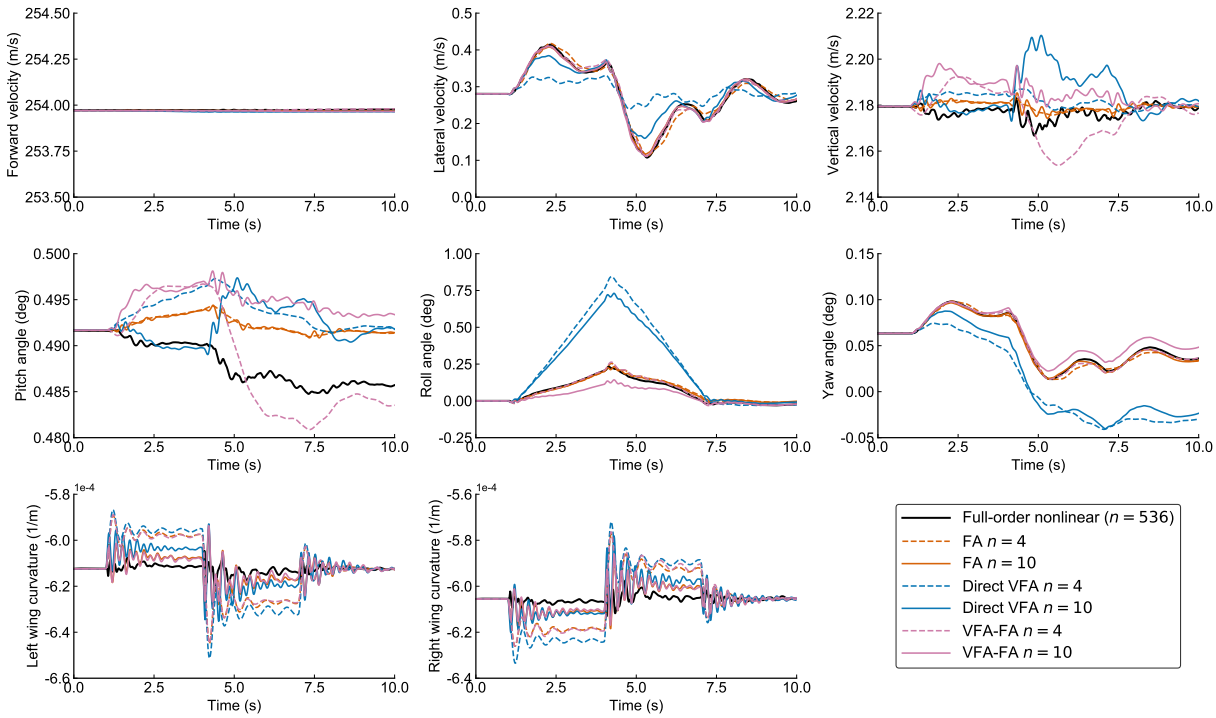



Figure 5 – Dynamic response to a 1-deg aileron doublet excitation.

directly shows that direct learning degrades the near-trim regime while residual learning preserves it. The high-amplitude maneuvers expose the limitation of the FA baseline and the value of the learned aerodynamic correction. Figures 8–10 show the 3-deg elevator, coupled 3-deg, and coupled 5-deg maneuvers. In the 3-deg elevator maneuver, the FA $n = 4$ RMSE is 2.121, whereas Direct VFA and VFA-FA reduce the error to 1.511 and 0.628, respectively. With $n = 10$, VFA-FA gives the lowest error for this maneuver, 0.392. The coupled 3-deg maneuver shows the same $n = 10$ behavior, with VFA-FA giving the lowest error.

At 5 deg, the nonlinear correction becomes more demanding. As Fig. 10 shows, the FA response stays near its small-amplitude shape while the full-order nonlinear reference develops large coupled deviations, so the FA baseline no longer tracks the reference and its RMSE exceeds 10 for both mode counts. Both data-driven aerodynamic models recover a large part of the missing nonlinear response. Relative to the earlier Direct VFA results of Duan et al. [8], the present Direct VFA implementation shows improved lateral-directional tracking in the high-amplitude coupled maneuver, especially in the lateral velocity and Euler-angle channels shown here. Within the present comparison, the improvement from the residual model is most evident in roll. For the 5-deg coupled maneuver with $n = 4$, VFA-FA tracks the roll response more closely than Direct VFA and gives the lowest coupled 5-deg error, 2.663 compared with 4.170 for Direct VFA, while the yaw response remains comparatively more difficult to improve. For $n = 10$, the Direct VFA model gives a lower coupled 5-deg error, 1.609 compared with 2.618 for VFA-FA, but the two data-driven models remain comparable and both improve markedly over the FA baseline value above 10. The residual model therefore demonstrates the desired nonlinear authority at both mode counts. It is the most accurate model at the largest amplitudes for $n = 4$ and is comparable to Direct VFA at the largest amplitudes for $n = 10$ while remaining the most accurate model on the 3-deg maneuvers.

To test generalization, the models are also evaluated on additional maneuvers that were not included in the training set. These maneuvers are not a full amplitude sweep, but they provide an initial check that the residual model is not simply memorizing the selected integer-amplitude closed-loop maneuvers. Table 7 summarizes the aggregate RMSE. For $n = 10$, Direct VFA has the lowest mean error over these four maneuvers, 0.172, followed by VFA-FA at 0.197 and FA at 0.248. For $n = 4$, VFA-FA reduces the mean error relative to Direct VFA and is close to the FA baseline. The mixed 2-deg elevator, 1-deg aileron, and 1-deg rudder maneuver is particularly interesting because it lies

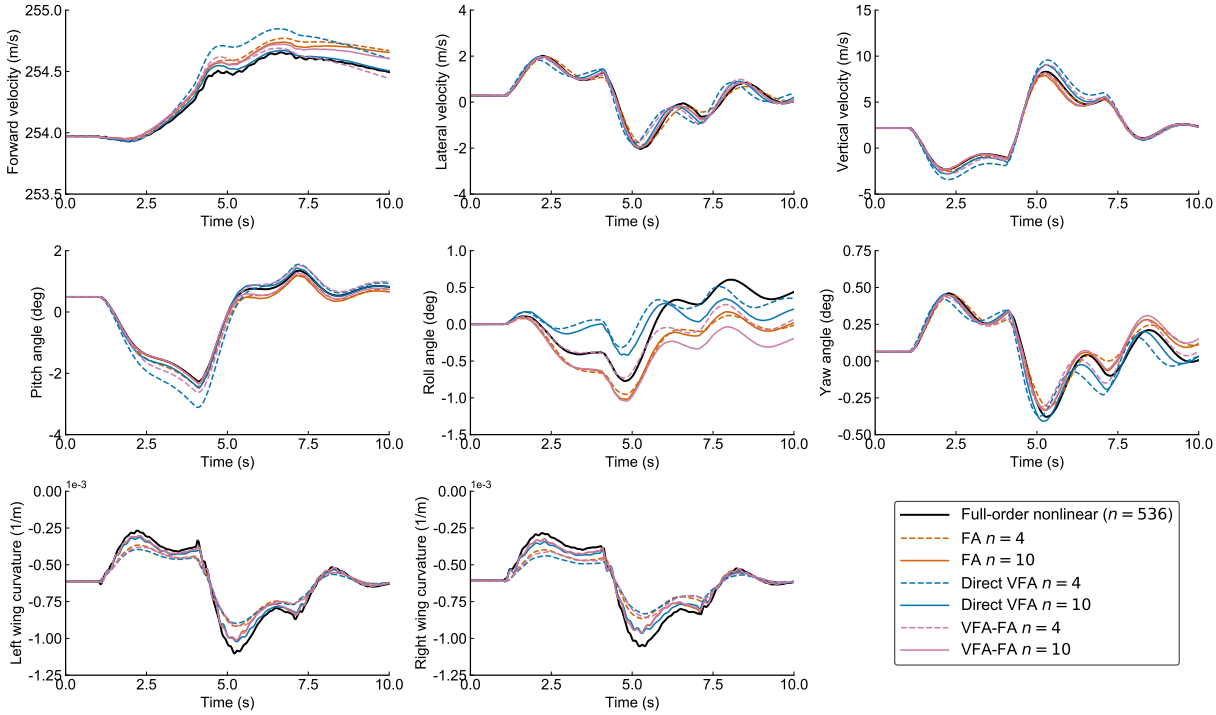



Figure 6 – Dynamic response to a coupled 1-deg elevator, aileron, and rudder doublet excitation.

between the small- and high-amplitude regimes. There, the $n = 4$ VFA-FA error is 0.374 compared with 0.604 for FA and 0.557 for Direct VFA.

Table 7 – Closed-loop aggregate RMSE for maneuvers not included in training, against the full-order nonlinear ($n = 536$) reference.

| **Maneuver** | $n = 4$ | | | $n = 10$ | | |
|---|---|---|---|---|---|---|
| | **FA** | **Direct VFA** | **VFA-FA** | **FA** | **Direct VFA** | **VFA-FA** |
| 1-deg elevator, 1-deg rudder | 0.103 | 0.316 | 0.214 | 0.122 | 0.143 | 0.113 |
| 0.5-deg coupled | 0.053 | 0.164 | 0.137 | 0.047 | 0.084 | 0.051 |
| 1-deg elevator, 0.5-deg aileron and rudder | 0.095 | 0.297 | 0.164 | 0.121 | 0.139 | 0.120 |
| 2-deg elevator, 1-deg aileron and rudder | 0.604 | 0.557 | 0.374 | 0.701 | 0.324 | 0.502 |

Table 8 shows the average simulation time and the corresponding speed-up relative to real time, measured as simulation wall time. These values can vary with the computing platform used to evaluate the models, so they are reported as relative comparisons rather than absolute benchmarks and provide a baseline check of how the mode count affects the runtime of the Direct VFA and VFA-FA models that capture the nonlinear behavior. All models ran faster than real time. The FA baseline is an analytical low-order model and is therefore expected to run substantially faster than real time, which the measurements confirm at roughly $9\times$ for $n = 4$. The Direct VFA and VFA-FA models add a neural-network evaluation at each step and run at comparable speed to each other, as expected because both infer a network of the same size. The mode count is the dominant cost driver, with the learned $n = 4$ models running at approximately $7\times$ real-time speed and the learned $n = 10$ models at approximately $1.5\times$. These measurements highlight a balance between response accuracy and computational cost, where the higher mode count improves accuracy at the expense of runtime.

## 5. Conclusions

This paper presented a residual-learning extension to a low-order B2T-VFA formulation, using a representative HARW aircraft test case, and compared it against the FA linearized baseline and the Direct

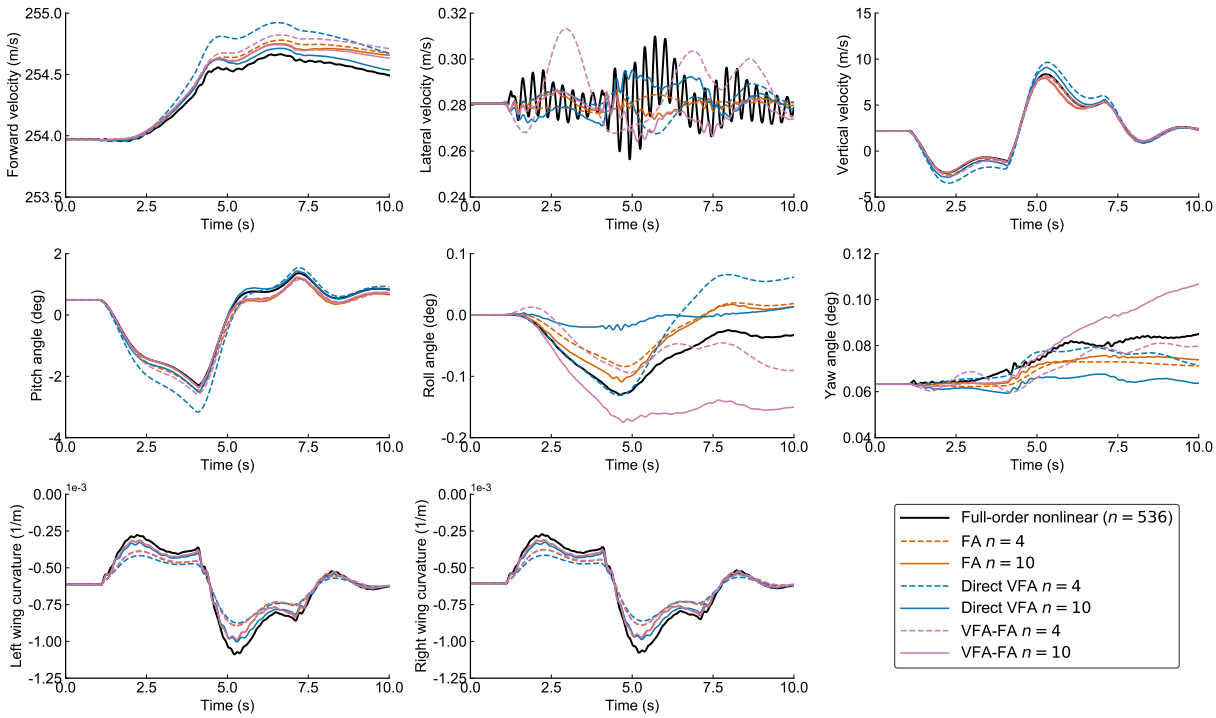



Figure 7 – Dynamic response to a 1-deg elevator doublet excitation.

Table 8 – Average computational cost from the 10-s closed-loop runs.

| **Model** | **Average simulation time [s]** | **Approximate speed-up ($\times$ real time)** |
|---|---|---|
| FA $n = 4$ | 1.10 | 9.1 |
| FA $n = 10$ | 5.45 | 1.8 |
| Direct VFA $n = 4$ | 1.48 | 6.8 |
| Direct VFA $n = 10$ | 6.26 | 1.6 |
| VFA-FA $n = 4$ | 1.45 | 6.9 |
| VFA-FA $n = 10$ | 6.52 | 1.5 |

VFA model using full-order nonlinear ($n = 536$) simulations. Both networks reproduced their assigned aerodynamic outputs accurately, and the load comparisons confirmed that nonlinear corrections become necessary as the FA error increases with amplitude. The maneuver-response results then showed distinct behavior among the three models with amplitude. Near trim, the FA baseline was the most accurate and direct learning instead degraded the response, while learned corrections were needed at high amplitude, where the linear baseline loses accuracy. The residual model combined both behaviors. For the $n = 4$ model, it preserved the near-trim accuracy of the FA baseline and gave the lowest mean error at high amplitude among the three models. With more retained modes ($n = 10$), both data-driven models improved substantially over the FA baseline at high amplitude and remained comparable to each other at the largest amplitudes, while the residual model stayed closest to FA near trim.

The residual model retains the linearized baseline explicitly and adds only a learned nonlinear correction, so it captures the nonlinear behavior without sacrificing the interpretability or near-trim reliability of the linear model. Since the correction is defined relative to this baseline, the accuracy of the residual model is tied to that of the baseline, and a more accurate FA model would further aid the residual learning. The data-driven models also ran faster than real time for the simulated maneuvers, with the $n = 4$ models the fastest, supporting their use in control-oriented studies. Future work will extend the residual formulation to the structural and inertial terms and to longer-duration maneuvers, and will further examine the mode-count dependence of the high-amplitude residual identified in the $n = 4$ and $n = 10$ comparison. It will also explore extensive hyperparameter tuning and residual-improvement

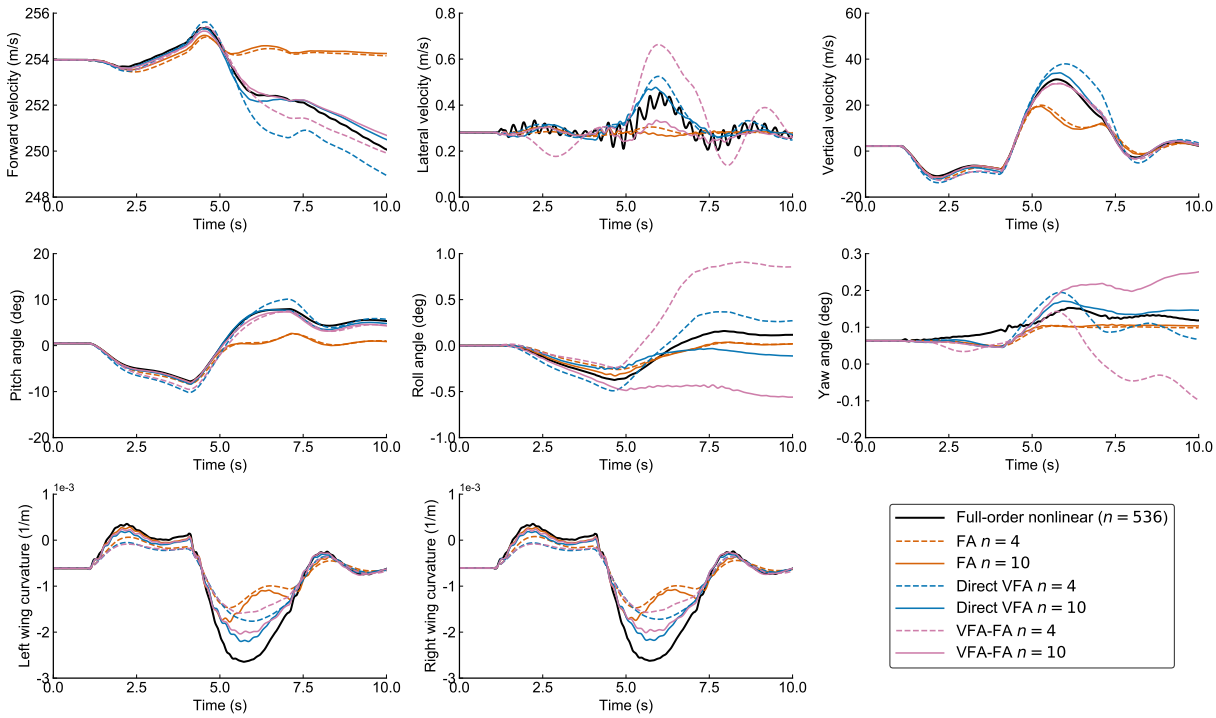



Figure 8 – Dynamic response to a 3-deg elevator doublet excitation.

techniques, a broader set of maneuvers, and the sensitivity of the model to training-data coverage. A further direction is to distribute the learned correction across the individual terms of the linearized aerodynamic operator in Eq. (3), learning state-dependent adjustments to the rigid-body, flexible-state, and control sensitivities rather than a single lumped load correction, making it more suitable for control-oriented modeling.

## Acknowledgments

This work was supported in part by the Center for Autonomous Air Mobility & Sensing (CAAMS), a National Science Foundation Industry–University Cooperative Research Center.

## 6. Contact Author Email Address

cesnik@umich.edu

## 7. Copyright Statement

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

## References

[1] C. E. S. Cesnik, R. Palacios, and E. Y. Reichenbach. Reexamined structural design procedures for very flexible aircraft. *Journal of Aircraft*, 51(5):1580–1591, 2014.

[2] David K. Schmidt and David L. Raney. Modeling and simulation of flexible flight vehicles. *Journal of Guidance, Control, and Dynamics*, 24(3):539–546, 2001.

[3] Kenneth Milne. Nonlinear flight equations for flexible aircraft using mean axes. In *AIAA Atmospheric Flight Mechanics Conference*, 1971.

[4] J R Canavin and P W Likins. Flexible aircraft equations of motion using mean axes. In *AIAA Flight Mechanics Specialist Conference*, 1972.

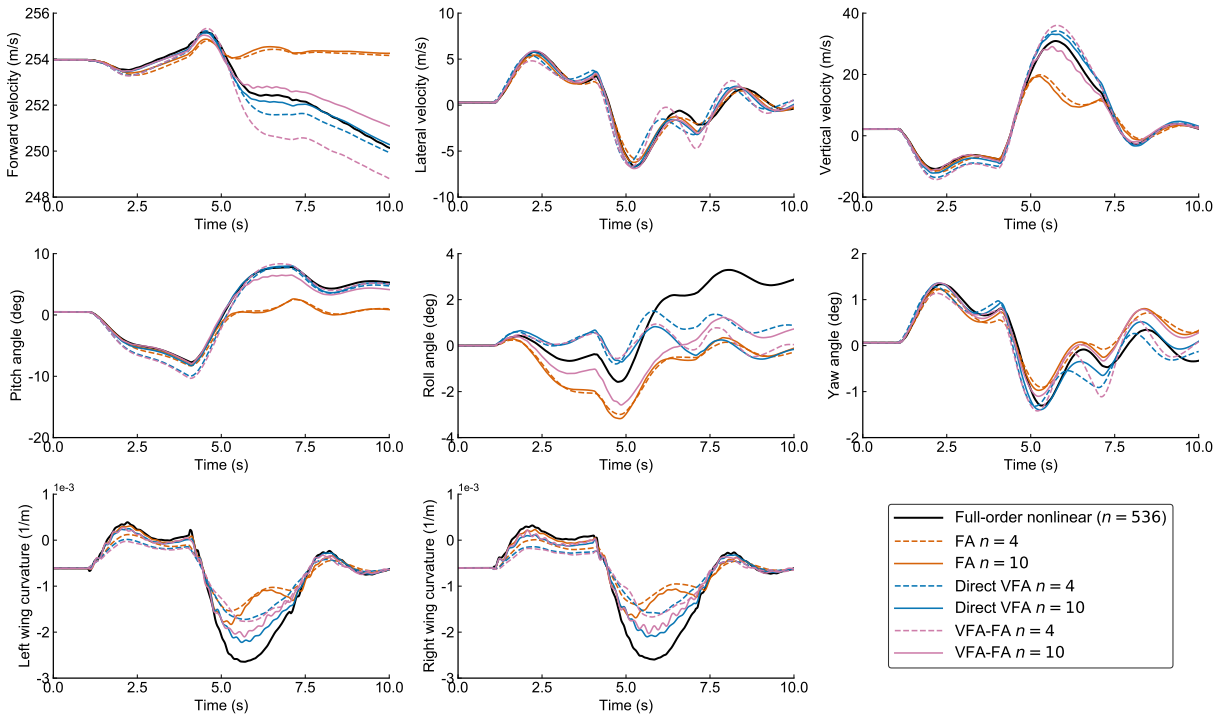



Figure 9 – Dynamic response to a coupled 3-deg elevator, aileron, and rudder doublet excitation.


[5] C. M. Shearer and C. E. S. Cesnik. Nonlinear flight dynamics of very flexible aircraft. *Journal of Aircraft*, 44(5):1528–1545, 2007.

[6] W. Su and C. E. S. Cesnik. Nonlinear aeroelasticity of a very flexible blended-wing-body aircraft. *Journal of Aircraft*, 47(5):1539–1553, 2010.

[7] Weihua Su and Carlos E. S. Cesnik. Dynamic response of highly flexible flying wings. *AIAA Journal*, 49(2):324–339, 2011.

[8] M. Duan, C. E. S. Cesnik, I. V. Kolmanovsky, and F. Vetrano. Nonlinear control-oriented modeling for very flexible aircraft. In *AIAA Scitech 2021 Virtual Forum*, 2021.

[9] M. Duan, C. E. S. Cesnik, I. V. Kolmanovsky, and F. Vetrano. Control-oriented modeling for flexible aircraft. *Journal of Aircraft*, 61(1), 2024.

[10] Giulio Avanzini, Augusto Rossoni, and Salvatore Pezzuto. Maneuvering response of very flexible aircraft by multibody dynamics. *Journal of Aircraft*, 47(4):1150–1160, 2010.

[11] W. Su and C. E. S. Cesnik. Strain-based geometrically nonlinear beam formulation for modeling very flexible aircraft. *International Journal of Solids and Structures*, 48(16-17):2349–2360, 2011.

[12] A. Da Ronch, K. Badcock, Y. Wang, A. Wynn, and R. Palacios. Nonlinear model reduction for flexible aircraft control design. In *AIAA Atmospheric Flight Mechanics Conference*, 2012.

[13] R. Palacios and A. Cea. Nonlinear modal condensation of large finite element models: Application of hodges's intrinsic theory. *AIAA Journal*, 57(10):4255–4268, 2019.

[14] J. Kou and W. Zhang. Layered reduced-order models for nonlinear aerodynamics and aeroelasticity. *Journal of Fluids and Structures*, 68:174–193, 2017.

[15] Lei He, Weiqi Qian, Tun Zhao, and Qing Wang. Multi-fidelity aerodynamic data fusion with a deep neural network modeling method. *Entropy*, 22(9), 2020.

[16] X. Zhang, F. Xie, T. Ji, Z. Zhu, and Y. Zheng. Multi-fidelity deep neural network surrogate model for aerodynamic shape optimization. *Computer Methods in Applied Mechanics and Engineering*, 373:113485, 2021.

[17] D. H. Lee, D. Lee, S. Han, S. Seo, B. J. Lee, and J. Ahn. Deep residual neural network for predicting aerodynamic coefficient changes with ablation. *Aerospace Science and Technology*, 136:108207, 2023.

[18] N. Demo, M. Tezzele, and G. Rozza. A deeponet multi-fidelity approach for residual learning in reduced order modeling. *Advances in Modeling and Simulation in Engineering Sciences*, 10(1):12, 2023.

[19] Q. Yan, Z. Wan, and C. Yang. Flight load calculation using neural network residual kriging. *Aerospace*, 10(7), 2023.

[20] Johan M. Feldwisch. Increased-order model for unsteady aerodynamic nonlinearities using neural net-

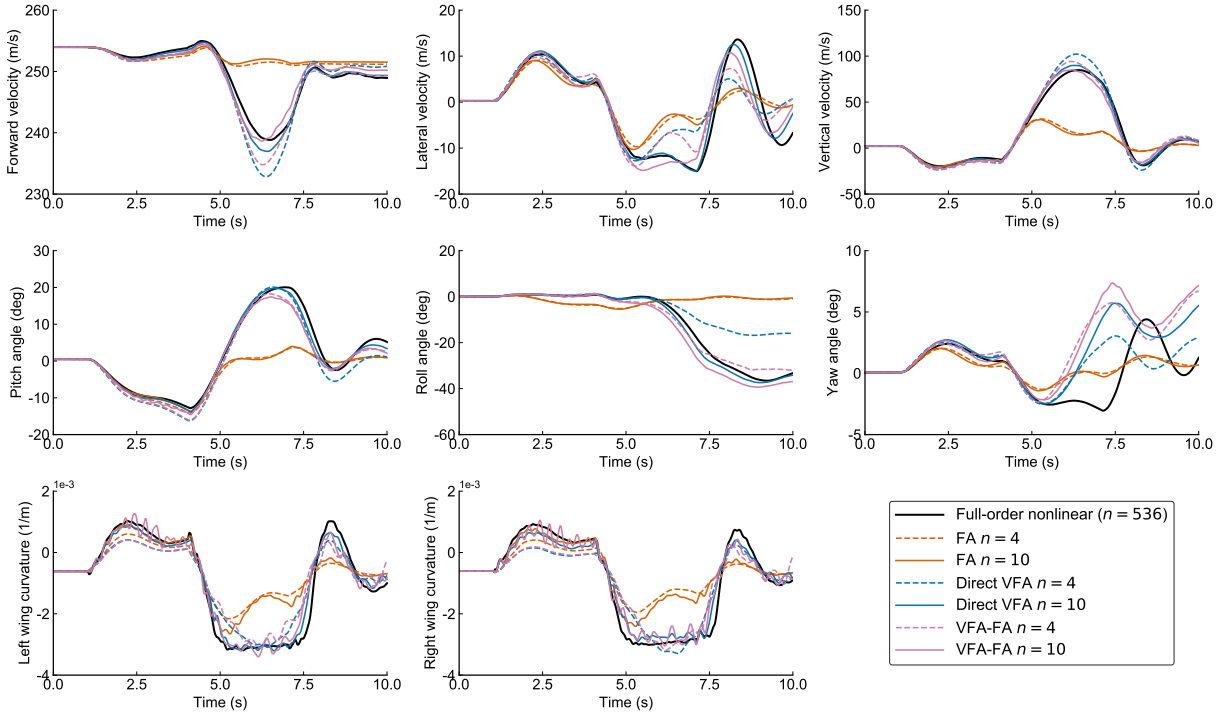



Figure 10 – Dynamic response to a coupled 5-deg elevator, aileron, and rudder doublet excitation.

works. *Journal of Aircraft*, 2025.

[21] C. Riso, D. Sanghi, C. E. S. Cesnik, F. Vetrano, and P. Teufel. Parametric roll maneuverability analysis of a high-aspect-ratio-wing civil transport aircraft. In *AIAA SciTech 2020 Forum*, Orlando, FL, January 2020.

[22] D. Sanghi, C. E. S. Cesnik, and C. Riso. Roll maneuverability of transonic high-aspect-ratio-wing aircraft with flared folding wingtips. *Journal of Aircraft*, 2024.