Statistical Wind-Tunnel Experimentation Advancements for eVTOL Aircraft Aero-Propulsive Model Development

Benjamin M. Simmons* and Ronald C. Busan†

NASA Langley Research Center, Hampton, Virginia 23681

Compared to many conventional aircraft, electric vertical takeoff and landing (eVTOL) vehicle designs present greater aero-propulsive complexity and many interacting factors requiring development of novel testing and model development strategies. This paper describes several advances in statistically-rigorous experiment design methods for wind-tunnel testing of eVTOL aircraft enabling accurate aero-propulsive model development. The techniques are applied to a subscale tilt-rotor eVTOL aircraft configuration built at NASA Langley Research Center that has 24 independent control effectors. The general testing approach, leveraging design of experiments and response surface methodology, is outlined and followed by a detailed description of the experiment design and modeling approach employed for different phases of the wind-tunnel testing. A new approach for gravitational tare modeling is developed and validated, which reduces the required wind-tunnel test time by nearly 50%. A mostly automated trim envelope determination approach is formulated to rapidly determine the aircraft flight envelope to inform transition testing for aero-propulsive model development. Efficient powered-airframe characterization tests are conducted using a nested I-optimal design approach formulated in previous simulation-based testing. The collective wind-tunnel testing approach facilitates rapid characterization of eVTOL aircraft configurations and produces accurate aero-propulsive models that can be used in flight dynamics simulations.

I. Introduction

MANY novel aircraft designs are currently under investigation to support a future Urban Air Mobility (UAM) transportation system [1, 2]. Vehicles supporting UAM operations require precise hover and efficient cruise capabilities as well as the ability to safely transition between flight regimes. Hybrid and electric vertical takeoff and landing (eVTOL) aircraft are a clear fit for this new transportation model. In general, eVTOL aircraft are a combination of traditional fixed-wing and rotary-wing aircraft including certain attributes from each type of vehicle. Rotary-wing aircraft features provide the ability to takeoff and land vertically, hover, and precisely maneuver in confined areas. Longer endurance, better efficiency, and the ability to operate at high speeds is derived from fixed-wing aircraft attributes. Distributed hybrid and electric propulsion technology used in many eVTOL aircraft concepts has further broadened the traditional aeronautical vehicle design space and resulted in many unique vehicle designs [3–10].

Although the operational utility of eVTOL aircraft has great potential, there are many research areas to be addressed prior to introduction into a UAM environment. One essential research area is accurate eVTOL vehicle aero-propulsive modeling enabling flight dynamics simulation development to support flight control system design and certification, research on practical flight operations, and many other areas. Efficient and accurate vehicle aero-propulsive model development, however, is challenged by several vehicle attributes, including: many control surfaces and propulsors, propulsion-airframe interactions, propulsor-propulsor interactions, high-incidence angle propulsor aerodynamics, vehicle instability, rapidly changing aerodynamics through transition, and large flight envelopes that need to be characterized by a global aero-propulsive model. Previous research has investigated methods for efficient eVTOL aircraft aero-propulsive modeling across their wide flight envelopes using computational predictions [11, 12] and wind-tunnel testing [13–18]. The present work builds on previous eVTOL aircraft modeling research to further refine an effective aero-propulsive modeling strategy for eVTOL aircraft using wind-tunnel experiments leveraging design of experiments (DOE) and response surface methodology (RSM) theory [19, 20].

The paper is organized as follows: Section II introduces the experimental aircraft. Section III gives an overview of the wind-tunnel testing effort and model identification approach. Subsequent sections then describe the experiment design and modeling strategy for each phase of testing, including: gravitational tare characterization in Sec. IV, transition

---

*Research Engineer, Flight Dynamics Branch, Member AIAA.

†Research Engineer, Flight Dynamics Branch, Member AIAA.

trim envelope determination in Sec. V, and transition aero-propulsive characterization in Sec. VI. Overall conclusions are summarized in Sec. VII.

II. Aircraft

The wind-tunnel testing and modeling approaches demonstrated in this paper are applied to the Research Aircraft for eVTOL Enabling techNologies (RAVEN) Subscale Wind-Tunnel and Flight Test (SWFT) model built at NASA Langley Research Center (LaRC). The RAVEN-SWFT is a 28.625% scale version of the RAVEN 1000-lb eVTOL aircraft concept [21], which has been designed in a collaborative effort between NASA LaRC and the Georgia Institute of Technology (Georgia Tech). The RAVEN aircraft is a tilt-rotor eVTOL aircraft configuration with six variable-pitch proprotors. The front four proprotors tilt forward and are operational throughout the entire flight envelope. The rear two proprotors do not tilt and serve as lifting proprotors in hover and transition. The rear proprotors are inactive in forward flight. The aircraft control surfaces included six flaperons, an all moving stabilator, and a rudder. In total, the vehicle has 24 independent control effectors:

- Six proprotor rotational speeds ($n_1, n_2, ..., n_6$)
- Six proprotor collective pitch angles ($\delta_{c_1}, \delta_{c_2}, ..., \delta_{c_6}$)
- Four nacelle tilt angles ($\delta_{t_1}, \delta_{t_2}, \delta_{t_3}, \delta_{t_4}$)
- Six flaperon deflection angles ($\delta_{f_1}, \delta_{f_2}, ..., \delta_{f_6}$)
- One stabilator deflection angle (δ_s)
- One rudder deflection angle (δ_r)

A schematic of the RAVEN aircraft with annotations showing the vehicle propulsor and control surface definitions is shown in Fig. 1. Flaperon and stabilator deflections are defined as positive trailing edge downward. Rudder deflection is defined as positive trailing edge left. As currently configured, proprotor 1, 3, and 5 rotate counterclockwise and proprotor 2, 4, and 6 rotate clockwise, as viewed from the perspective of each respective electric motor. The tested deflection angle range for each control effector for the version of the RAVEN-SWFT vehicle described in this paper is shown in Table 1.
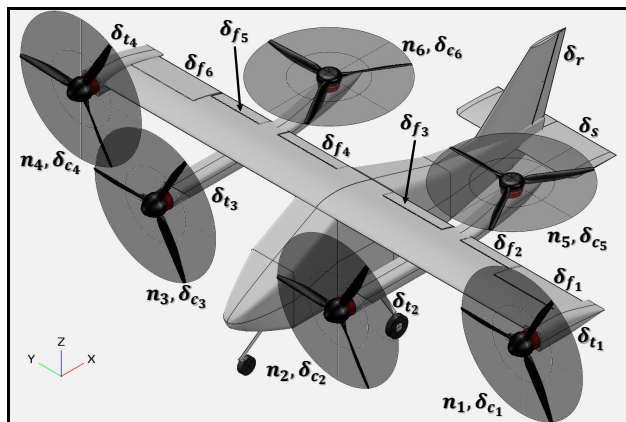

Fig. 1   RAVEN control effector definitions.

Table 1   RAVEN-SWFT control effector limits

Control Effector(s)	Units	Minimum	Maximum
$\delta_{c_1}, \delta_{c_2}, \delta_{c_3}, \delta_{c_4}$	deg	−2.5	+12
$\delta_{c_5}, \delta_{c_6}$	deg	−7	+7
$\delta_{t_1}, \delta_{t_4}$	deg	0	+96
$\delta_{t_2}, \delta_{t_3}$	deg	0	+88.4*
$\delta_{f_1}, \delta_{f_6}$	deg	−30	+65
$\delta_{f_2}, \delta_{f_3}, \delta_{f_4}, \delta_{f_5}$	deg	−35	+35
δ_s	deg	−22	+22
δ_r	deg	−30	+30

*The inboard tilt mechanisms were intended to have a maximum angle of 90 deg, but were restricted to 88.4 deg for the testing described in this paper due to a physical model limitation that has since been resolved.

The RAVEN-SWFT, pictured in the NASA Langley 12-Foot Low-Speed Tunnel (LST) [22] in Fig. 2, is a subscale aircraft configuration designed as a flight dynamics and controls testbed to advance eVTOL aircraft technology. As its name suggests, the vehicle is designed for use in both wind-tunnel and flight-test experiments. The aircraft has a target flight weight of 37 pounds, a wingspan of 5.7 ft (excluding the outboard nacelles), and a proprotor diameter of 19.5 inches. The RAVEN-SWFT was developed at NASA LaRC as one of several subscale eVTOL research aircraft intended to explore their unique flight characteristics and resolve implementation challenges to help bring similar full-scale vehicles into mainstream operation. Previous aircraft have included the LA-8 [10, 23] and GL-10 [7], which have enabled research in computational aerodynamic predictions [24, 25], wind-tunnel testing [13, 16, 26], high incidence angle proprotor aerodynamics [27–29], aero-propulsive modeling [7, 14, 17, 18, 30], flight controls [31, 32], and flight-test strategies [33–35].
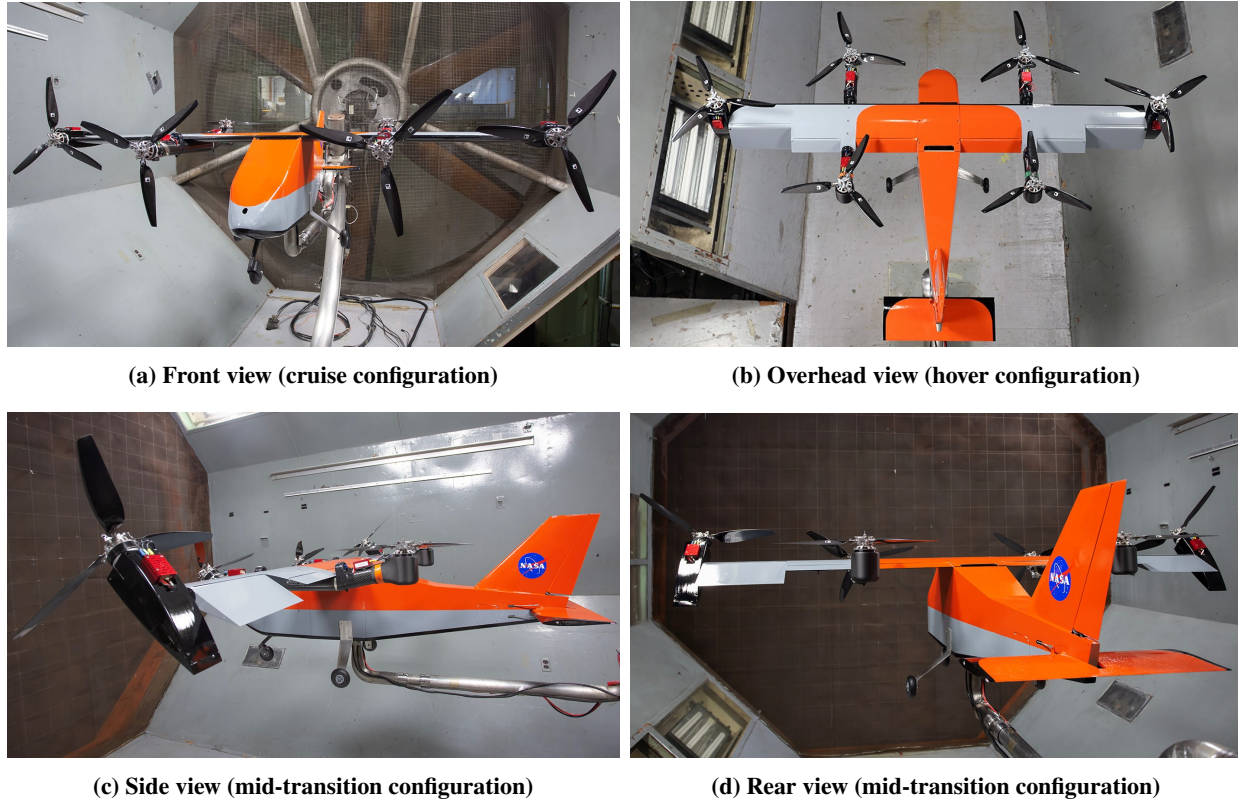
(a) Front view (cruise configuration)

(b) Overhead view (hover configuration)

(c) Side view (mid-transition configuration)

(d) Rear view (mid-transition configuration)

Fig. 2   RAVEN-SWFT mounted in the NASA Langley 12-Foot Low-Speed Tunnel.
(Credit: Lee Pollard, NASA Langley Research Center)

III. Wind-Tunnel Experiment and Model Identification Overview

RAVEN-SWFT wind-tunnel tests have been performed in the 12-Foot LST for the isolated proprotor [36], the isolated airframe (without proprotors), and powered airframe (with proprotors operating). This paper focuses on the initial static full-vehicle wind-tunnel test completed in December 2022, which acquired data to accomplish multiple objectives:

- integration and validation of a gravitational tare modeling approach to substantially reduce the number of required tare points,
- characterization of the motor and control surface servo-actuator dynamics,
- development of an isolated-airframe aerodynamic model,
- experimental determination of the RAVEN-SWFT trim envelope through transition to inform powered-airframe characterization experiments, and
- development of a powered-airframe aero-propulsive model throughout the transition envelope up to a dynamic pressure setting of $\bar{q} = 5$ psf (the practical limit for testing eVTOL vehicles in 12-Foot LST).

The test phases used to identify the powered-airframe aero-propulsive model are described through the remainder of the paper.

A. Experiment Design Overview

For each test phase described in this paper, the design matrix was formed using cuboidal, completely randomized, I-optimal response surface designs [19, 20, 37, 38] created based on the purpose, desired model precision, and number of test factors (n_f). I-optimal designs minimize the integrated prediction variance across the design space for an assumed model structure, which is a good design criterion choice for this test because the primary goal was to develop models with low prediction error. Test matrix design was performed using Design-Expert®, a commercially available statistical software package [39].

The design matrices were created using insight gained from previous related studies [17, 30] and were evaluated/refined

using design evaluation metrics prior to conducting the experiments. The primary benchmark used to assess and tune the experiment designs was the prediction variance, which provides insight into the precision of prediction for a given model structure. The variance of the predicted response is

$$\text{Var}[\hat{y}(\boldsymbol{x}_0)] = \sigma^2 \boldsymbol{x}_0^T \left(\boldsymbol{X}^T \boldsymbol{X} \right)^{-1} \boldsymbol{x}_0 \tag{1}$$

where $\hat{y}(\boldsymbol{x}_0)$ is the predicted response evaluated at the design space location $\boldsymbol{x}_0$ expanded to the form of the model structure, $\boldsymbol{X}$ is a matrix composed of the designed test points in the form of the model structure, and σ^2 is the measurement error variance [20]. From Eq. (1), the prediction variance is a function of the experiment design, the model structure, the location in the design space, and the measurement facility error variance. The unscaled prediction variance (UPV), defined as

$$\text{UPV} = \frac{\text{Var}[\hat{y}(\boldsymbol{x}_0)]}{\sigma^2} = \boldsymbol{x}_0^T \left(\boldsymbol{X}^T \boldsymbol{X} \right)^{-1} \boldsymbol{x}_0 \tag{2}$$

removes the dependence on σ^2 and, thus, can be used to compare experiment designs when σ^2 is unknown prior to experimentation.

Graphical presentation of the distribution of prediction variance throughout the design space is an effective way to assess experiment designs. Fraction of design space (FDS) plots, introduced by Ref. [40], depict the prediction variance distribution over the design space in a concise manner, where the prediction variance metrics are plotted against the FDS encompassing a prediction variance less than or equal to a particular value. It is also useful to consider the FDS including a particular model precision, quantified by the confidence interval half-width δ [38, 41, 42]. The model precision δ normalized by the response standard deviation σ plotted against FDS provides further insight into the prediction capability of the model developed from a particular experiment design, prior to conducting the experiment. For this study, a design was deemed to be adequate for fitting a particular model complexity if δ/σ was less than two for greater than 95% of the design space. The prediction variance threshold PV^* used to determine the FDS within a given model precision level is

$$\text{PV}^* = \left(\frac{\delta/\sigma}{t_{\alpha_p/2, N-p}} \right)^2 \tag{3}$$

where N is the number of test points, p is the number of parameters in the model, and α_p is the significance level chosen as $\alpha_p = 0.05$.

To create each I-optimal design matrix for the different wind-tunnel test stages, the number of free design points was selected to obtain an adequate FDS (FDS ≥ 0.95) with a normalized model precision $\delta/\sigma \leq 2$. The design optimization was conducted with one fixed center point. After the optimization was completed, additional center points were added within the design. This was done because it has been noted from previous experience that including multiple replicate points in the center of a design during the optimization process can repel the optimized points away from the interior of the design space. Multiple center points were included to each design to provide additional stabilization of the prediction variance within the experimental region and allow for estimation of pure error [19]. Furthermore, axial points were added to some test matrices, as will be described in future sections. Axial points, where each test factor is varied at a specified distance from the center of the design space while all other test factors are held at the center of the design space, aid in interpretation of the model fit for main effects because the axial points and predicted response can be directly compared in two-dimensional plots for each factor. Axial points also improve the design metrics if they are included in the points used for model identification. In addition to the points used for modeling, test points selected using a random number generator were used as validation data withheld from model identification. Randomized validation data are agnostic to the model development experiment design and the model complexity it was designed for, making this a good validation point selection strategy for designed experiments. The number of validation data points was selected to provide a good independent estimate of prediction error while remaining a modest number of test points. Using this overall approach as a foundation, the experiment design for each phase of testing will be described in more detail in Secs. IV-VI.

B. Model Identification Overview

Model identification for this effort focused on developing polynomial response surface equation (RSE) representations of the response variables as a function of vehicle state and control variables. Factors under test, or close variants (as discussed in the following sections), are defined as explanatory variables, and a model is identified from the data collected in wind-tunnel testing. The model structure identification and parameter estimation methods used for this

work were adapted from the System IDentification Programs for AirCraft (SIDPAC) software toolbox [43, 44]. The model structure was developed using the stepwise regression algorithm described in Ref. [45] and the model parameters were estimated using ordinary least-squares regression.

The stepwise regression algorithm is a combination of forward selection and backwards elimination of candidate regressors where a single regressor is either added to or removed from the model at each iteration. If the partial F-statistic for a term included in the model falls below a cutoff threshold $F(\alpha_p, 1, N - p)$, the term is removed from the model. Otherwise, the excluded candidate model terms with the highest correlation to the unmodeled portion of the response variable is added into the model. The algorithm was run automatically until the remaining excluded model terms did not surpass a partial F-statistic cutoff value when added to the model. The automated algorithm is effective in predicting dominant terms that should be included in the model, but is more challenged to determine which borderline terms with similar statistical modeling metrics are worthy of inclusion in the model structure (a task that is more obvious to a subject matter expert based on physical insight). This results in both model terms excluded from the model that should be included and other model terms included in the final model that should be excluded, based on analyst judgment. For example, if an interaction term for a control surface and a nearby propulsor is excluded from the model, consideration might be given to including the term. Conversely, if the model includes an interaction term for control effectors far apart on the vehicle lacking physical justification, consideration should be given to removing the term from the model. Most often these borderline model terms are either just above or below the statistical cutoff thresholds. Therefore, after automatic model structure selection is completed, it is beneficial to have a subject matter expert adjust the final model structure based on physical insight combined with statistical metrics.

After identifying the model structure and parameter estimates, model adequacy was examined using data withheld from the model development process. Model fit metrics and modeling residuals alone do not provide information about the model predictive capability. Assessment of model performance using validation data not used for modeling provides a more reliable estimate of model prediction accuracy. Validation assessment can be performed by analyzing the prediction residuals ($e = z - \hat{y}$) between the measured response z and predicted response $\hat{y}$. Comparison of modeling and prediction residuals is useful because a significant increase in the spread of prediction residuals compared to modeling residuals is a way of diagnosing an inadequate model. Residuals and their statistical properties can be given further interpretability by normalization. A good error normalization metric is the range of response variable measurements used to develop the model, $\text{range}(z_m) = \max(z_m) - \min(z_m)$. Range normalization provides a fair comparison between prediction error metrics for different response variables used for aircraft modeling because longitudinal responses are generally biased above or below zero and lateral-directional responses are generally centered about zero. The normalized residual vector is defined as:

$$ e^* = \frac{z - \hat{y}}{\text{range}(z_m)} \tag{4} $$

Similarly, the normalized root-mean-square modeling error (NRMSE) is defined as:

$$ \text{NRMSE} = \frac{1}{\text{range}(z_m)} \sqrt{\frac{(z - \hat{y})^T (z - \hat{y})}{N}} \tag{5} $$

The range-normalized NRMSE is a good metric for model validation because it:
  1) succinctly presents a quantitative assessment and comparison of the model fit and prediction performance,
  2) allows comparisons of the prediction capability for responses with different character and units, and
  3) is straightforward to interpret as a percent error quantity.

Another quantitative measure of the model adequacy is critical binomial analysis of residuals prediction error metric, e^*_{cv}. The process of computing e^*_{cv} is shown in Ref. [17]. Further explanation of critical binomial analysis of residuals and justification for using this metric to assess prediction error are given in Ref. [46].

IV. Gravitational Tare Characterization

For the wind-tunnel testing described in this paper, the forces and moments exerted on the aircraft are measured by an internal strain-gage balance. These include contributions from aero-propulsive forces and moments, as well as from gravitational loads (i.e., aircraft weight). To isolate the aero-propulsive forces and moments, the desired measurement from wind-tunnel testing, the contributions from gravitational forces must be removed. To determine the gravitational loads, the forces and moments are measured without wind-tunnel airflow, requiring additional test time, referred to as a tare run. For conventional airplane wind-tunnel testing, the variation of the gravitational tare forces and

moments generally only depends on the angle of attack; therefore, the additional testing required to remove gravitational contributions can be limited. However, for eVTOL aircraft wind-tunnel testing, vehicle components such as tilting wings or tilting rotors, change the mass distribution and, thus, must be accounted for in the gravitational tare. Further complications for DOE/RSM characterization experiments result from vehicle component positions not having the symmetry observed in flight operations. In previous eVTOL aircraft wind-tunnel testing, tare runs have been completed by running the same test matrix twice—first with the aircraft unpowered and without tunnel airflow, and then with the aircraft powered and airflow on—in essence, doubling the required test time. To resolve this inefficiency, a gravitational tare modeling approach was developed to obviate most of the required tare test time using DOE/RSM techniques.

A. Experiment Design

For the powered RAVEN-SWFT aircraft, it was postulated that eight factors could significantly influence the gravitational tare: the angle of attack (α), the four nacelle tilt angles (δ_{t_1}, δ_{t_2}, δ_{t_3}, δ_{t_4}), the large outboard flaps (δ_{f_1}, δ_{f_6}), and the all-moving stabilator (δ_s). For the isolated-airframe test, the tilt nacelle angles were fixed to their forward flight setting, resulting in only four test factors for the tare experiments. Because each of these test factors is moved independently for aero-propulsive model development, the effect of each tare test factor must be independently characterized.

I-optimal test matrices were developed for the four-factor and eight-factor tare modeling experiments. Each test matrix supported identification of a full quadratic design model (all possible linear, pure quadratic, and two-factor interaction model terms), pure cubic terms for each factor, and cubic interaction terms including angle of attack and one other test factor (e.g., for δ_{f_1}, the cubic interaction terms included $\alpha^2 \delta_{f_1}$ and $\alpha \delta_{f_1}^2$); the four-factor test matrix also supported up to sixth-order pure α terms. The test factor ranges for each tare experiment are shown in Table 2, which were designed to cover the full range of operability of the factors for the wind-tunnel test.

Table 2   Gravitational tare characterization experiment test factor ranges

(a) Isolated airframe (four test factors)			
Factor(s)	Units	Minimum	Maximum
α	deg	−8	+24
$\delta_{f_1}, \delta_{f_6}$	deg	−30	+30
δ_s	deg	−22	+22

(b) Powered airframe (eight test factors)			
Factor(s)	Units	Minimum	Maximum
α	deg	−8	+16
$\delta_{t_1}, \delta_{t_4}$	deg	0	+96
$\delta_{t_2}, \delta_{t_3}$	deg	0	+88.4
$\delta_{f_1}, \delta_{f_6}$	deg	−30	+65
δ_s	deg	−22	+22

The four-factor and eight-factor I-optimal gravitational tare experiment designs included 40 and 76 free design points, respectively, and one center point. After design optimization was completed, the design was supplemented with axial points for each test factor with a coded distance from the center of the design space of −1, −0.5, +0.5, and +1, resulting in $4 \times n_f$ additional points. Furthermore, 5 additional center points were added to the design. The final test matrix was re-randomized before conducting the experiment. In addition to the points used for modeling, 12 ($3 \times n_f$) and 16 ($2 \times n_f$) validation data points, respectively, were added using a random number generator. Table 3 summarizes the tare experiment designs, including the design model and the breakdown of test points; the FDS with $\delta/\sigma \leq 2$ for the points used in design optimization and the points used in model identification (the optimized model points, center points, and axial points) with the respective model form used to create the design. Figure 3 shows a two-dimensional slice of each tare experiment design.

B. Modeling Results

Gravitational tare models were developed as RSEs for the axial, normal, side, roll, pitch, and yaw balance voltage measurements as a function of the test factors. Gravitational tares predominantly influence the longitudinal responses (axial, normal, and pitch voltage), but the lateral-directional responses (side, roll, and yaw voltage) were also modeled to capture small offsets and deterministic variation where applicable. For the longitudinal responses, the R^2 values were 99.8% or higher, indicating that a large majority of the variation of the longitudinal voltage response was characterized by the models.

Table 3 Gravitational tare modeling experiment design summary

Factors	Design Model	Model Points	Center Points	Axial Points	Validation Points	Total Points	Optimization FDS with $\delta/\sigma \leq 2$	Model FDS with $\delta/\sigma \leq 2$
4	quadratic + pure cubic + cubic α interactions + α^4 + α^5 + α^6	40	6	16	12	74	0.994	0.999
8	quadratic + pure cubic + cubic α interactions	76	6	32	16	130	0.938	0.992

(a) Isolated airframe (four factors)

(b) Powered airframe (eight factors)
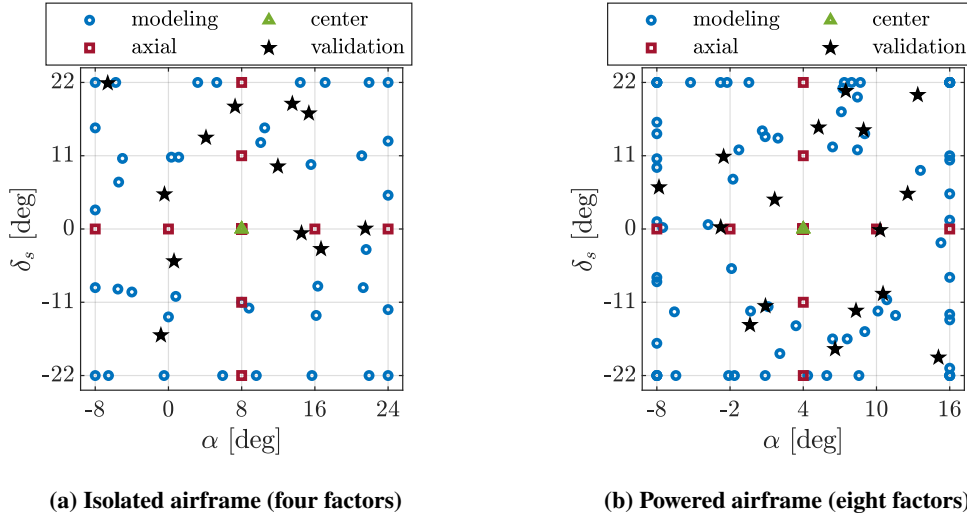
Fig. 3 Two-dimensional slices of the gravitational tare characterization experiment designs.

The gravitational tare models were rapidly identified and analyzed in the 12-Foot LST control room immediately following a tare run to promptly proceed with wind-on testing. To allow for quick assessment of modeling results for each response, the MATLAB®-based tare modeling software created for this test displayed plots to analyze the adherence to regression modeling assumptions and model fit adequacy [20, 47], including:
- normal probability plots of externally studentized modeling residuals,
- externally studentized modeling residuals against data point number,
- externally studentized modeling residuals against the predicted response,
- externally studentized modeling residuals against each test factor,
- model mean response and prediction intervals compared to measured data for the axial sweep of each test factor, and
- bar charts comparing the ordinary, adjusted, and predicted R^2.

Furthermore, normalized modeling and validation residual histories were compared to assess model prediction capability.

The selected model terms for the longitudinal tare models are listed in Table 4. The final gravitational tare models were used to make predictions of the tare voltage values for the test matrices executed throughout the RAVEN-SWFT wind-tunnel testing, removing the need for additional tare runs. To integrate the gravitational tare predictions into the 12-Foot LST software, the predicted values for each run schedule were stored in a data file replacing the traditional tare data files.

To help provide confidence in the gravitational tare modeling approach and verify successful integration of predicted tare files into the 12-Foot LST computer system, randomized isolated-airframe angle of attack sweeps were made with a modeled tare and traditional tare file. Note that the randomized α-sweep data collection strategy is intended to help expose and characterize the uncertainty in the measured data and, consequently, makes the raw data quality look noisier

Table 4   **Final model terms selected for the isolated-airframe and powered-airframe gravitational tare models**

Response	Isolated-Airframe Model Terms	Powered-Airframe Model Terms
Normal Voltage	α, α^2	α, α^2
Axial Voltage	$\alpha, \alpha^2, \alpha^5$	$\alpha, \alpha^2, \alpha^3$
Pitch Voltage	α, α^2	$\alpha, \delta_{t_1}, \delta_{t_4}, \delta_{f_1}, \delta_{f_6}, \alpha\delta_{t_1}, \alpha\delta_{t_2}, \alpha\delta_{t_3}, \alpha\delta_{t_4}, \alpha^2, \delta_{t_1}^2, \delta_{t_2}^2, \delta_{t_3}^2, \delta_{t_4}^2, \alpha\delta_{t_1}^2, \alpha\delta_{t_4}^2, \delta_{t_1}^3, \delta_{t_4}^3$

than a traditional one-factor-at-a-time (OFAT) sweep where a single test factor is varied by sequentially increasing or decreasing its value. Figure 4a shows the axial, normal, and pitch voltage values measured in a traditional tare run compared to the mean response and 95% prediction interval (PI) on the mean response for the identified tare model. The tare voltage responses predicted by the tare model are observed to accurately predict the measured tare voltage values and the measured data remain mostly within the 95% prediction interval. Figure 4b shows two separate angle of attack sweeps at $\bar{q} = 3.5$ psf, with one using a traditional tare and one using a modeled tare to remove the gravitational loads. The data processed using a traditional tare is plotted with error bars reflecting the $\pm 2\sigma$ measurement uncertainty bounds. The resulting variation of lift coefficient C_L, drag coefficient C_D, and pitching moment coefficient C_m with α is practically identical (within the measurement error bounds), as expected. These results suggest that the tare modeling procedures developed for this test, and expected to be used in future tests, were successful.
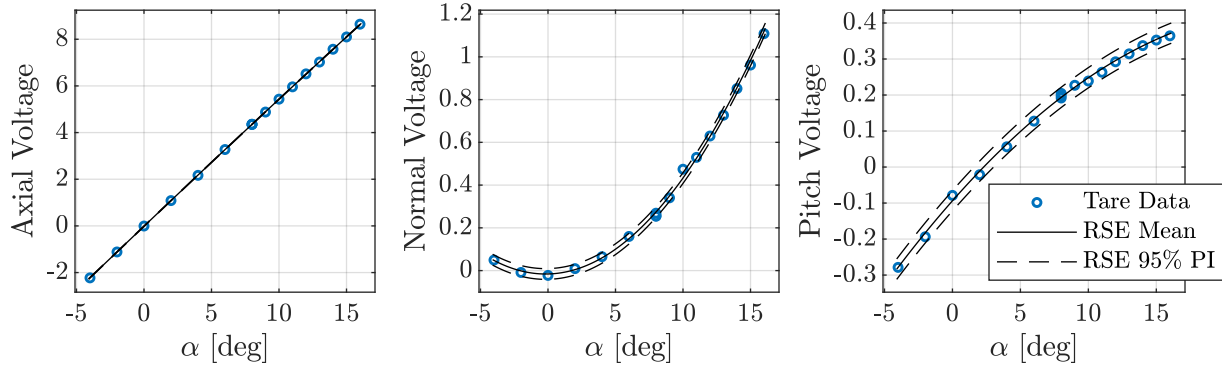
(a) Gravitational tare voltage data and model predictions
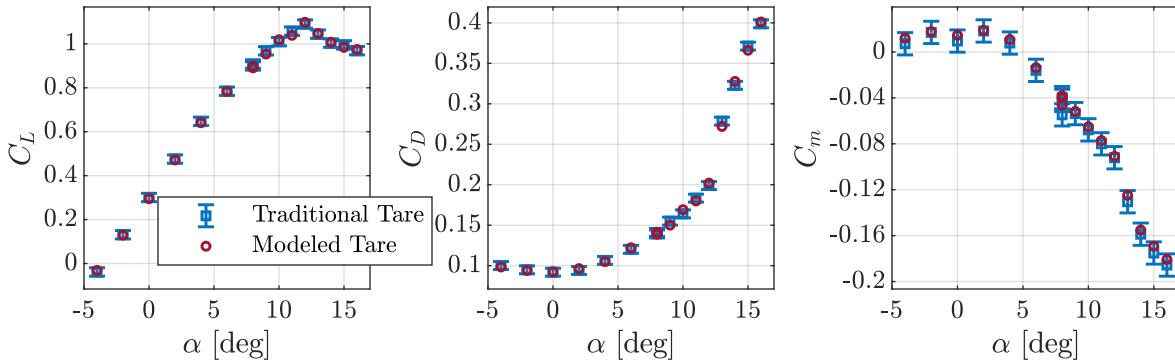
(b) Lift, drag, and pitching moment coefficients recorded using a traditional and modeled tare

Fig. 4   **Gravitational tare model validation using an isolated-airframe angle of attack sweep at $\bar{q} = 3.5$ psf.**

V. Transition Trim Envelope Determination

As described in Refs. [16, 17], determination of the trim envelope for an eVTOL aircraft is crucial for specifying the test factor ranges for powered-airframe wind-tunnel experiments to facilitate collection of useful data throughout transition. In previous work, initial exploratory OFAT testing was used to define the ranges of factors for subsequent DOE/RSM testing used for aero-propulsive model development. This was accomplished by finding trim points where the longitudinal forces were balanced and the pitching moment was zero by manually varying groups of control effectors with subject matter expertise [16, 26]. Although this approach was effective for determining the trimmable flight envelope through transition, it was time consuming and conducted in an ad hoc manner.

To improve the efficiency and accuracy of the trim envelope determination process for RAVEN-SWFT, a new approach was formulated using DOE/RSM applied from the perspective of process optimization (as opposed to characterization—the perspective of the testing conducted for the other stages of testing described in this paper). The trim envelope was determined by strategically linking certain control effectors together for the experiment, and then identifying a set of RSEs that allow a root-finding algorithm to find trim control effector settings for a level, unaccelerated transition at different airspeed conditions. Informed by the number of different airspeed settings used in previous eVTOL aircraft wind-tunnel testing [16, 17] and the capabilities of the 12-Foot LST, eight dynamic pressure settings were selected for testing:

$$\bar{q} = [0 \quad 0.25 \quad 0.62 \quad 1.16 \quad 1.87 \quad 2.75 \quad 3.79 \quad 5.00] \text{ psf}$$

which, at standard sea level conditions, corresponds to freestream velocity settings of:

$$V = [0 \quad 14.5 \quad 22.9 \quad 31.3 \quad 39.7 \quad 48.1 \quad 56.5 \quad 64.9] \text{ ft/s}$$

These airspeed settings include a hover condition with no tunnel airflow, as well as seven non-zero airspeed settings chosen to linearly vary airspeed between the lowest and highest practical settings for longer-duration testing in the 12-Foot LST. The experiment design discussed next was executed at each airspeed setting and was followed by identification of a set of RSEs for each airspeed setting.

A. Experiment Design

The main purpose of the trim envelope determination test phase was to find an operational transition flight envelope to focus on in subsequent, more precise powered-airframe characterization testing. Consequently, compared to the powered-airframe characterization experiments that will be described in Sec. VI, the experiment design for trim envelope determination testing included a reduced number of test factors and a more frugal test matrix design strategy. The seven test factors were:

1) Front motor command, $\eta_{m_{\text{front}}} = \eta_{m_1} = \eta_{m_2} = \eta_{m_3} = \eta_{m_4}$
2) Rear motor command, $\eta_{m_{\text{rear}}} = \eta_{m_5} = \eta_{m_6}$
3) Front collective pitch angle, $\delta_{c_{\text{front}}} = \delta_{c_1} = \delta_{c_2} = \delta_{c_3} = \delta_{c_4}$
4) Rear collective pitch angle, $\delta_{c_{\text{rear}}} = \delta_{c_5} = \delta_{c_6}$
5) Nacelle tilt angle, $\delta_t = \delta_{t_1} = \delta_{t_2} = \delta_{t_3} = \delta_{t_4}$
6) Outboard flaperon deflection angle, $\delta_{f_{\text{out}}} = \delta_{f_1} = \delta_{f_6}$
7) Inboard and midboard flaperon deflection angle, $\delta_{f_{\text{in}}} = \delta_{f_2} = \delta_{f_3} = \delta_{f_4} = \delta_{f_5}$

Although linked motor commands were used as test factors for operational convenience, averaged measured front and rear proprotor rotational speed (denoted n_{front} and n_{rear}, respectively) were used as the explanatory variables for modeling. Angle of attack (α), angle of sideslip (β), stabilator deflection (δ_s), and rudder deflection (δ_r) settings were held at a constant value of 0 deg for the trim envelope determination experiments described in this paper.

An I-optimal response surface design was created for a design model complexity of a full quadratic model with pure cubic terms, including 51 optimized points and 5 center points. The design does not include axial points nor validation data because of the preparative and parsimonious nature of this stage of testing. The FDS with $\delta/\sigma \leq 2$ for the points used in the design optimization and the points used in model identification (the optimized model points and all center points) for the aforementioned design model were 0.964 and 0.985, respectively. A sample two-dimensional test matrix slice in coded units is shown in Fig. 5. Although "Coded Variable 1" and "Coded Variable 2" are shown in the figure, similar plots would be obtained for other test variables and the variables would be converted into engineering units before executing testing.

Table 5 shows the ranges of factors tested at each dynamic pressure setting. The trim envelope determination test was conducted starting in hover and the dynamic pressure setting was progressively increased so that the previous regions could help to inform the next region based on the trim envelope determined at the previous dynamic pressure setting.
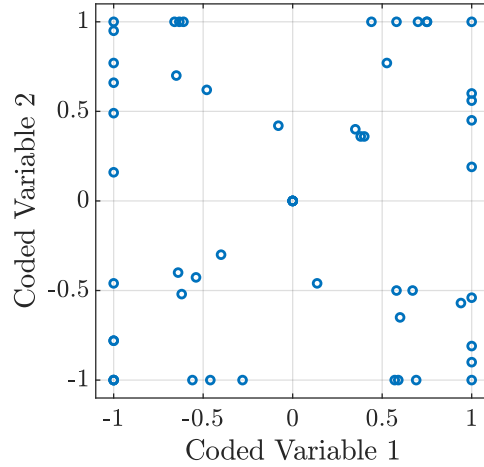
Fig. 5 Two-dimensional slice of the 7-factor trim envelope determination experiment design in coded units.

Table 5 Trim envelope determination test factor ranges at each dynamic pressure setting

$\bar{q}$	ηm_{front}	ηm_{rear}	$\delta_{c_{\text{front}}}$	$\delta_{c_{\text{rear}}}$	δ_t	$\delta_{f_{\text{out}}}$	$\delta_{f_{\text{in}}}$
psf	command	command	deg	deg	deg	deg	deg
0	[2500, 5600]	[2500, 5200]	[-2.5, 9]	[-4, 7]	88.4	[0, 65]	[0, 35]
0.25	[2500, 6000]	[2500, 5700]	[-2.5, 9]	[-4, 7]	[65, 88.4]	[0, 65]	[0, 35]
0.62	[2500, 5700]	[2500, 5200]	[-2.5, 9]	[-4, 7]	[60, 88.4]	[0, 65]	[0, 35]
1.16	[2500, 5200]	[2500, 5000]	[-2.5, 9]	[-4, 7]	[60, 88.4]	[0, 65]	[0, 35]
1.87	[2500, 5000]	[2500, 5000]	[-2.5, 9]	[-4, 7]	[50, 85]	[0, 65]	[0, 35]
2.75	[2500, 5000]	[2500, 5000]	[-2.5, 9]	[-4, 7]	[45, 80]	[0, 65]	[0, 35]
3.79	[2500, 5000]	[2500, 4800]	[-2.5, 10]	[-7, 7]	[40, 75]	[0, 65]	[0, 35]
5	[2500, 5000]	[2500, 4400]	[-2.5, 10]	[-7, 7]	[30, 70]	[0, 65]	[0, 35]

B. Modeling Results

After conducting trim envelope determination experiments at each dynamic pressure setting, response surface models were identified using similar near-real-time analysis capabilities to those described in Sec. IV. The identified RSEs were then used to determine the RAVEN-SWFT transition trim envelope. A trim envelope determination strategy was crafted to find control effector settings where the lift was equal to aircraft weight, the horizontal component of thrust balanced drag, and the pitching moment was zero throughout the RAVEN-SWFT transition envelope. The trim approach was developed to determine the trimmable flight envelope for the next stage of wind-tunnel testing in an expedited manner by the same researchers conducting the wind-tunnel test. Consequently, the following approach does not represent an optimal trim strategy for the aircraft; however, it was still successful in its intended purpose of finding a trimmable envelope to inform the next stage of wind-tunnel testing. Future analysis after completion of wind-tunnel testing will be applied to find optimal trim solutions, whereby objectives such as maximizing the available control authority or minimizing the power consumption can be incorporated into the trim strategy.

Trim analysis was performed at all tested dynamic pressure settings with $\alpha = 0$ deg, representing a level, unaccelerated transition. The control variables $\delta_{c_{\text{front}}}$, $\delta_{c_{\text{rear}}}$, and δ_t were specified to be free parameters used to find trim solutions using a modified Newton-Raphson solver available in SIDPAC [44]. The remaining control variables n_{front}, n_{rear}, $\delta_{f_{\text{out}}}$, and $\delta_{f_{\text{in}}}$ were fixed at a constant value while finding an individual trim solution. To find a set of trim solutions forming a trimmable flight envelope, the values of $n_{\text{front}} = n_{\text{rear}}$, $\delta_{f_{\text{out}}}$, and $\delta_{f_{\text{in}}}$ were swept throughout their tested range. The experimentally-determined trim envelope is shown in Fig. 6, where proprotor rotational speed, nacelle tilt angle, and front/rear collective pitch angle are plotted against freestream velocity. Note that the high-speed transition and forward

flight airspeed for the RAVEN-SWFT vehicle are higher than were able to be tested in the 12-Foot LST, which is why the full transition envelope is not shown in Fig. 6.
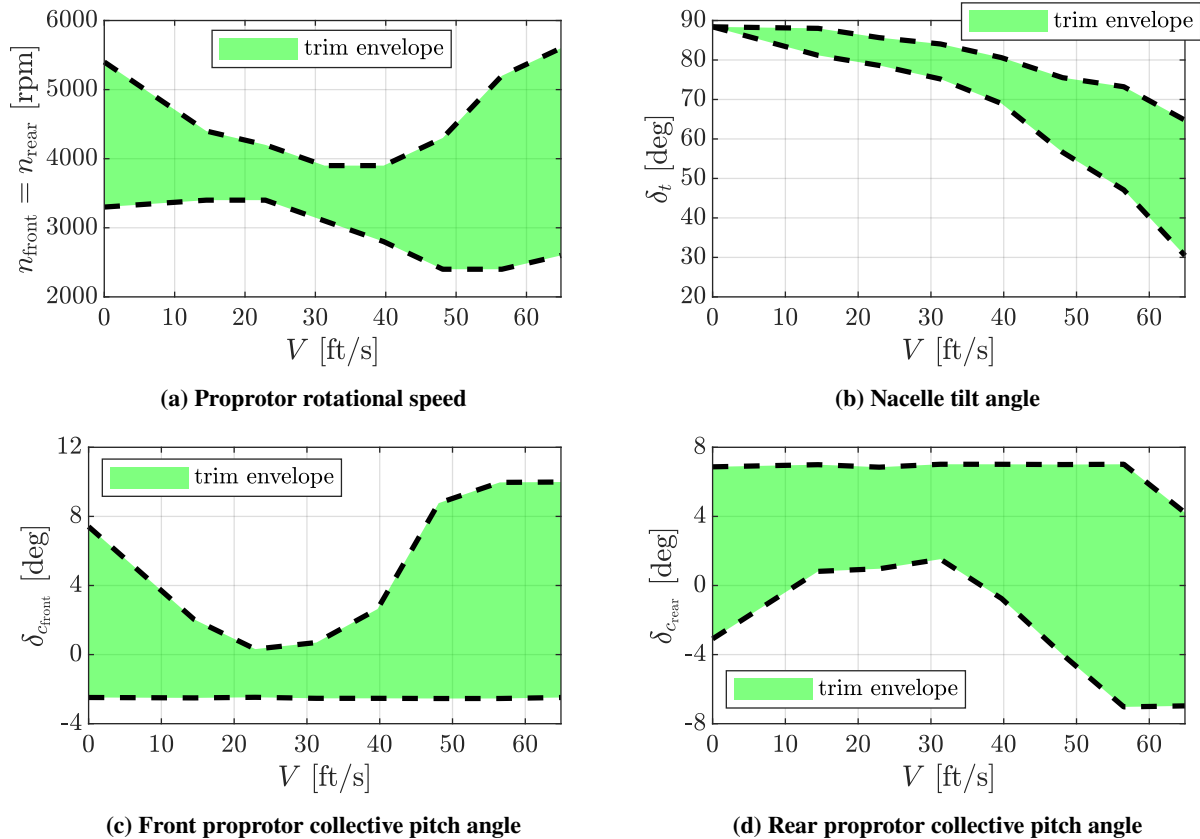

(a) Proprotor rotational speed

(b) Nacelle tilt angle

(c) Front proprotor collective pitch angle

(d) Rear proprotor collective pitch angle

Fig. 6   Experimentally-determined RAVEN-SWFT trimmable transition envelope with α = 0 deg.

VI. Powered-Airframe Characterization

Informed by the results from the trim envelope determination testing, static powered-airframe characterization experiments were conducted at each of the eight specified dynamic pressure settings. For the powered RAVEN-SWFT airframe, there are 26 independent factors at each dynamic pressure setting:

- Angle of attack, α
- Angle of sideslip, β
- Motor speed commands, $\eta_{m_1}, \eta_{m_2}, \eta_{m_3}, \eta_{m_4}, \eta_{m_5}, \eta_{m_6}$
- Collective pitch angles, $\delta_{c_1}, \delta_{c_2}, \delta_{c_3}, \delta_{c_4}, \delta_{c_5}, \delta_{c_6}$
- Nacelle tilt angles, $\delta_{t_1}, \delta_{t_2}, \delta_{t_3}, \delta_{t_4}$
- Flaperon deflection angles, $\delta_{f_1}, \delta_{f_2}, \delta_{f_3}, \delta_{f_4}, \delta_{f_5}, \delta_{f_6}$
- Stabilator deflection angle, δ_s
- Rudder deflection angle, δ_r

Although individual motor commands and airflow angles were used as test factors for operational convenience, measured proprotor rotational speeds for each propulsor ($n_1, n_2,..., n_6$) and body-axis velocity components v and w were used as the explanatory variables for modeling (see Ref. [17] for additional justification). The test factor ranges for each dynamic pressure setting are listed in Table 6. Table 7 lists the ranges for explanatory variables that are not test factors included in Table 6. As mentioned previously, the high-speed transition and forward flight airspeed for the RAVEN-SWFT vehicle are higher than were able to be tested in the 12-Foot LST. Future computational and/or flight-test model identification is expected to be conducted to provide aero-propulsive models at higher speeds.

Table 6   Powered-airframe characterization test factor ranges

$\bar{q}$	α,β	$\eta_{m_1},...,\eta_{m_4}$	η_{m_5},η_{m_6}	$\delta_{c_1},...,\delta_{c_4}$	$\delta_{c_5},\delta_{c_6}$	$\delta_{t_1},\delta_{t_4}$	$\delta_{t_2},\delta_{t_3}$	$\delta_{f_1},\delta_{f_6}$	$\delta_{f_2},...,\delta_{f_5}$	δ_s,δ_r
psf	deg	command	command	deg	deg	deg	deg	deg	deg	deg
0	0	[2500, 5400]	[2500, 5000]	[-2.5, 9]	[-4, 7]	[70, 96]	[70, 88.4]	[20, 65]	[-20, 20]	0
0.25	[-8, 8]	[2500, 5200]	[2500, 5000]	[-2.5, 9]	[-4, 7]	[65, 96]	[65, 88.4]	[20, 65]	[-20, 20]	[-20, 20]
0.62	[-8, 8]	[2500, 5100]	[2500, 4800]	[-2.5, 9]	[-4, 7]	[60, 90]	[60, 88.4]	[20, 65]	[-20, 20]	[-20, 20]
1.16	[-8, 8]	[2500, 5000]	[2500, 4700]	[-2.5, 9]	[-4, 7]	[60, 90]	[60, 88.4]	[20, 65]	[-20, 20]	[-20, 20]
1.87	[-8, 8]	[2800, 5000]	[2500, 4700]	[-2.5, 9]	[-4, 7]	[55, 85]	[55, 85]	[20, 65]	[-20, 20]	[-20, 20]
2.75	[-8, 8]	[2800, 5000]	[2500, 4200]	[-2.5, 10]	[-6, 7]	[50, 80]	[50, 80]	[15, 65]	[-20, 20]	[-20, 20]
3.79	[-8, 8]	[2900, 4900]	[2500, 4000]	[-2.5, 11]	[-7, 7]	[40, 75]	[40, 75]	[10, 60]	[-20, 20]	[-20, 20]
5	[-8, 8]	[3000, 4400]	[2500, 3800]	[-2, 12]	[-7, 7]	[30, 70]	[30, 70]	[0, 50]	[-20, 20]	[-20, 20]

Table 7   Powered-airframe characterization explanatory variable ranges (for variables not included in Table 6)

$\bar{q}$	v,w	$n_1,...,n_4$	n_5,n_6
psf	ft/s	rpm	rpm
0	0	[1900, 5900]	[2100, 6200]
0.25	[-2.0, 2.0]	[1600, 6100]	[2000, 6300]
0.62	[-3.2, 3.2]	[1600, 6100]	[2000, 6500]
1.16	[-4.4, 4.4]	[1500, 6200]	[2000, 6300]
1.87	[-5.5, 5.5]	[1600, 6200]	[2000, 6300]
2.75	[-6.7, 6.7]	[1500, 6100]	[1800, 6500]
3.79	[-7.9, 7.9]	[1400, 6200]	[1800, 6600]
5	[-9.0, 9.0]	[1400, 6200]	[1700, 6400]

A. Experiment Design

The combination of a regular and nested *I*-optimal design was selected as the response surface design type for the prediction accuracy and efficiency benefits explained in Ref. [30]; a similar design strategy was employed for this study. A set of five sequential test blocks was designed to acquire data for aero-propulsive model identification at each dynamic pressure setting. Blocks 1 and 2 were created as a two-block *I*-optimal design supporting identification of a full quadratic design model (all possible linear, pure quadratic, and two-factor interaction model terms). The design included 405 free design points and 6 center points (3 center points per block). Blocks 3 and 4 were created as a two-block nested *I*-optimal design for a quadratic design model, originally inspired by the concept and benefits of the nested face-centered central composite design proposed in Ref. [48]. The nested *I*-optimal design provides additional coverage of the interior portion of the design space, which aids in reducing model bias [20], while also allowing estimation of higher order model terms, such as pure cubic terms. The nested *I*-optimal design emulates the same design process as the regular *I*-optimal design, except that the design points are scaled to be at half the distance from the center of the design space. The nested *I*-optimal design was created separately from Blocks 1 and 2 (i.e., knowledge of the design points in Blocks 1 and 2 does not influence the design for Blocks 3 and 4). Block 5 included $2 \times n_f = 52$ validation test points selected using a random number generator; $4 \times n_f = 104$ axial points for each test factor with a coded distance from the center of the design space of -1, -0.5, $+0.5$, and $+1$; and 6 center points. Figure 7 shows a two-dimensional slice of the 26-factor experiment design in coded units. Sets of blocks are plotted sequentially with points from the previous blocks to show how design points progressively fill the design space. Table 8 summarizes the experiment design composition.

Design evaluation metrics were used to assess the design quality prior to conducting the wind-tunnel experiments. The evaluation approach and general results were similar to those presented in Ref. [30] for the sequence of a regular and nested *I*-optimal design, including properties of high statistical power, low correlation among candidate regressors, low prediction variance, and precise parameter estimates for the design model. Prediction variance metrics for the sequential block design are shown in Fig. 8 and Table 9, including results for Blocks 1-2 with a full quadratic evaluation

model, Blocks 1-4 with a full quadratic evaluation model, and Blocks 1-4 with a full quadratic model as well as pure cubic terms. Figure 8 shows the UPV, and δ/σ threshold values against FDS. Table 9 lists the FDS with $\delta/\sigma \leq 1.5$ and $\delta/\sigma \leq 2$ for each case. An adequate FDS (FDS ≥ 0.95) is obtained for all cases for $\delta/\sigma \leq 2$, as well as for Blocks 1-4 with $\delta/\sigma \leq 1.5$.
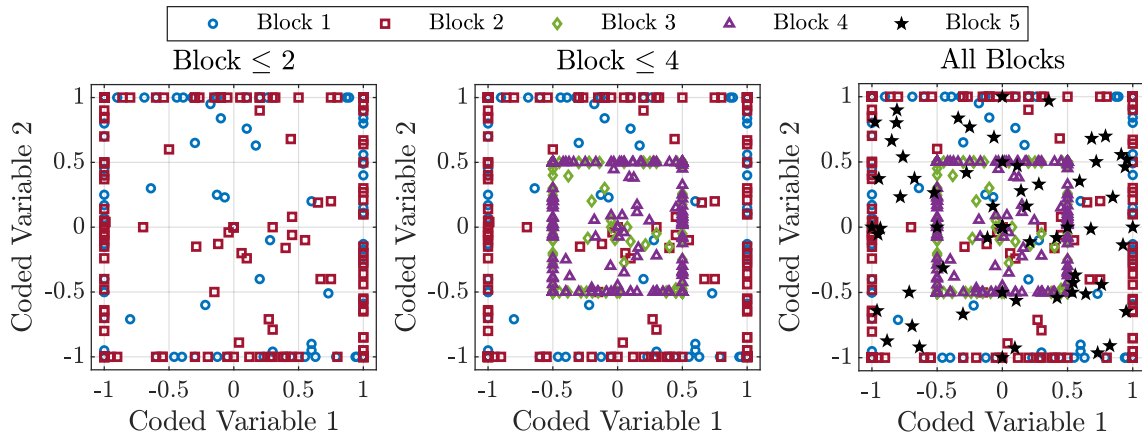

Fig. 7   Sequential two-dimensional slices of the 26-factor powered-airframe characterization experiment design.

Table 8   Powered-airframe experiment design summary

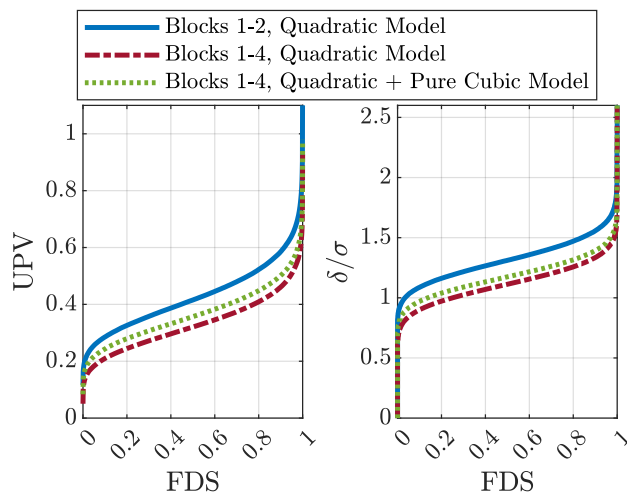| Block Number	Design Type	Design Model	Model Points	Center Points	Axial Points	Validation Points	Total Points	Cumulative Points
1	*I*-optimal	quadratic (1/2)	202	3	0	0	205	205
2	*I*-optimal	quadratic (2/2)	203	3	0	0	206	411
3	nested *I*-optimal	quadratic (1/2)	203	3	0	0	206	617
4	nested *I*-optimal	quadratic (2/2)	202	3	0	0	205	822
5	validation	N/A	0	6	104	52	162	984

Fig. 8   Prediction variance plots for the 26-factor powered-airframe experiment design.

Table 9   Prediction variance threshold FDS values for the powered-airframe experiment design

Blocks	Evaluation Model	FDS with $\delta/\sigma \leq 1.5$	FDS with $\delta/\sigma \leq 2$
1-2	Quadratic	0.837	0.999
1-4	Quadratic	0.986	1.000
1-4	Quadratic + Pure Cubic	0.970	1.000

13

B. Modeling Results

Separate aero-propulsive models were developed at each dynamic pressure setting to produce a transition model for the RAVEN-SWFT vehicle up to 5 psf (64.9 ft/s). The modeled responses are the dimensional body-axis aero-propulsive forces X, Y, Z in lbf and moments L, M, N in ft·lbf. The explanatory variables are defined as the body-axis velocity components v, w in ft/s, proprotor rotational speeds $n_1, n_2, ..., n_6$ in revolutions per second, proprotor collective pitch angles $\delta_{c_1}, \delta_{c_2},..., \delta_{c_6}$ in radians, nacelle tilt angles $\delta_{t_1}, \delta_{t_2}, \delta_{t_3}, \delta_{t_4}$ in radians, flaperon deflection angles $\delta_{f_1}, \delta_{f_2},..., \delta_{f_6}$ in radians, stabilator deflection angle δ_s in radians, and rudder deflection angle δ_r in radians. The choice of response variables and explanatory variables follows the eVTOL aircraft modeling convention presented and justified in Ref. [17].

1. Sample Local Modeling Results

To analyze the adherence to regression modeling assumptions and analyze model fit adequacy, sample residual diagnostics plots are shown in Fig. 9 for each response at $\bar{q} = 2.75$ psf. Figure 9a shows the normalized modeling and validation residuals, e_i^*, against data point number. Figure 9b shows normal probability plots for the externally studentized modeling residuals, t_i [20, 47]. Figure 9c shows the externally studentized modeling residuals against the predicted response, as well as 95% Bonferroni limits to help identify possible outliers. These plots show that the residuals for each response are reasonably independent (Fig. 9a) and normally distributed (Fig. 9b) with constant variance (Fig. 9c), satisfying the regression modeling assumptions and suggesting that the identified model parameters are meaningful. Furthermore, Fig. 9a also shows that modeling and prediction residuals appear to have similar magnitude, supporting the assertion that good predictive models have been identified. Residuals with similar character were obtained at the other tested dynamic pressure settings.

Figure 10 shows the modeled Z response compared to center and axial wind-tunnel data points for a subset of pertinent explanatory variables: w, n_4, δ_{t_4}, and δ_s. The model response is shown as RSE predictions of the mean response along with a 95% confidence interval (CI) on the mean response and a 95% prediction interval (PI) characterizing the uncertainty in the response predictions for individual data points [20]. The RSE predictions show good overall agreement with the measured wind-tunnel data and the measured data are all well within the 95% prediction interval. Similar predictive capability was observed for the dominant effects in other response variables and at other dynamic pressure settings.

2. Global Modeling Results

Figures 11-14 show modeling and prediction error metrics for models identified at each dynamic pressure setting. Tables 10-13 at the end of the paper list the numerical values shown in the figures. Figure 11 (Table 10) shows the coefficient of determination (R^2) value for each local model. The R^2 metric quantifies the amount of variation in the response variable about its mean value that is described by the model. Most local models have an R^2 value near or above 90% indicating that a significant portion of the variation in the response is described by the models. An abnormally low R^2 value is observed for Y at $\bar{q} = 0$ psf because there is no direct means of side force generation in hover. Figure 12 (Table 11) shows the number of model terms (n_p) identified for each local model. The number of parameters in the models is greater than many conventional aircraft modeling problems, but is consistent with previous eVTOL aircraft modeling efforts [12, 17]. Figure 13 (Table 12) shows the NRMSE computed using modeling data and validation data. The NRMSE values calculated using modeling and validation data for each respective local model are similar and low-valued indicating that a quality model has been identified. The binomial analysis of residuals prediction error metric (e_{cv}^*) value for each model is shown in Figure 14 (Table 13). Most local models have a e_{cv}^* value of less than 5% indicating that models with good prediction capability have been identified.

After developing local models at each dynamic pressure setting, a continuous transition model that can be used for flight dynamics simulations was formed using shape-preserving piecewise cubic interpolation [49, 50] between local model predictions. This interpolation method is preferred over cubic spline interpolation because of its tendency to avoid overshoot and oscillatory behavior.
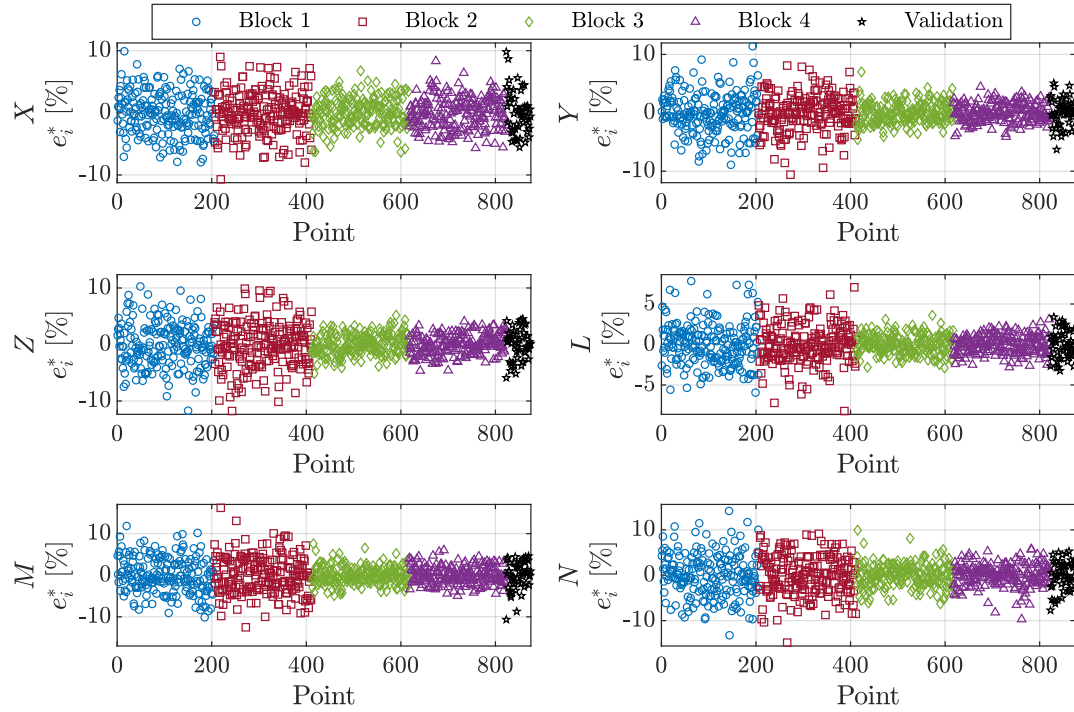
(a) Residual history
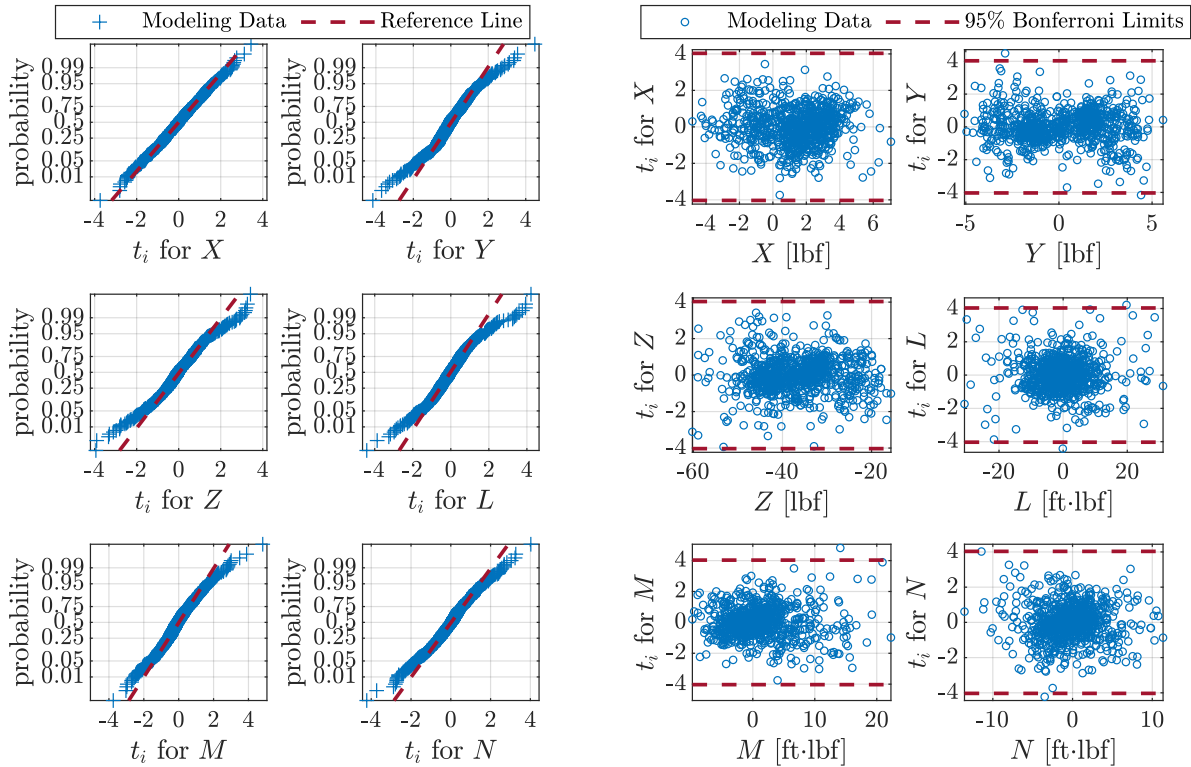
(b) Normal probability plot

(c) Residuals against predicted response

Fig. 9   Residual diagnostics for the powered-airframe model at $\bar{q}$ = 2.75 psf.
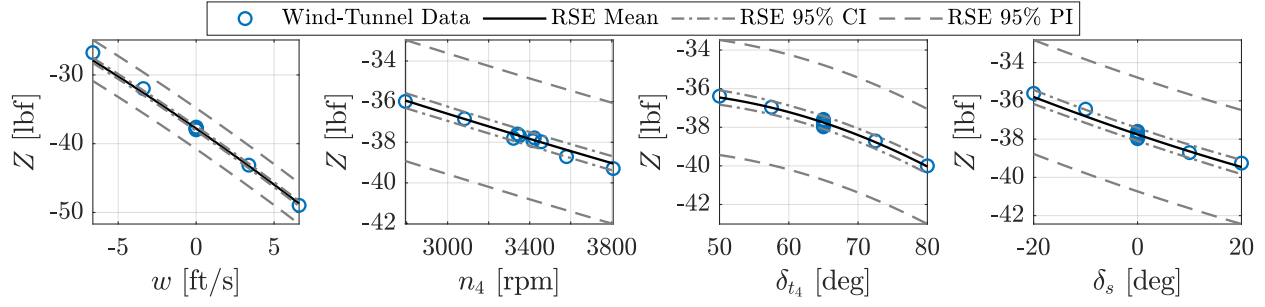
Fig. 10 Comparison of model predictions and wind-tunnel data for explanatory variable sweeps at $\bar{q}$ = 2.75 psf.
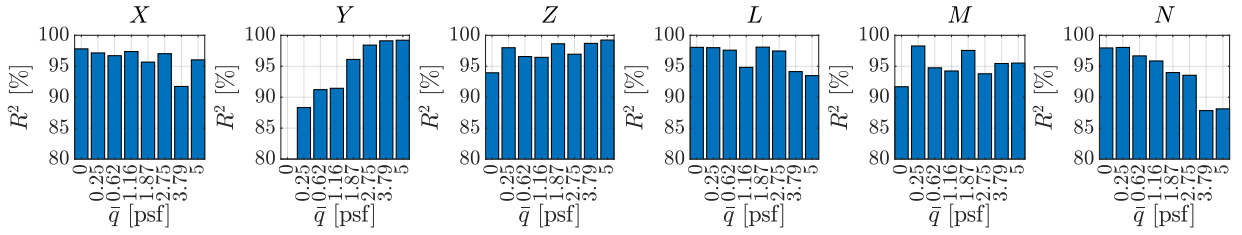

Fig. 11 Coefficient of determination, R^2, for each local model response.
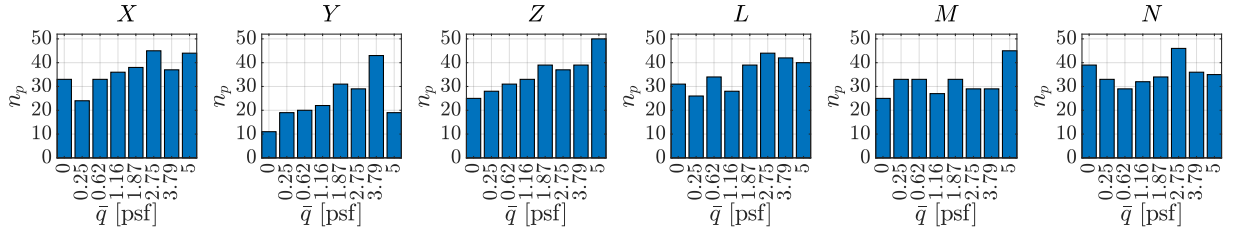

Fig. 12 Number of model parameters, n_p, identified for each local model response.
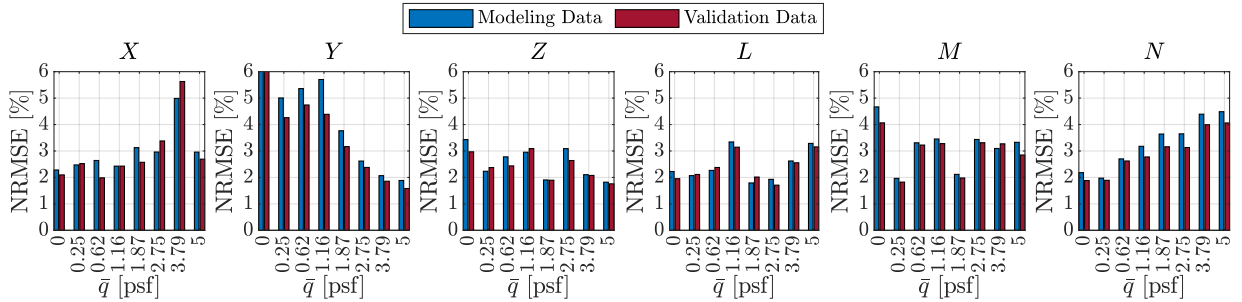

Fig. 13 Modeling and validation NRMSE for each local model response.
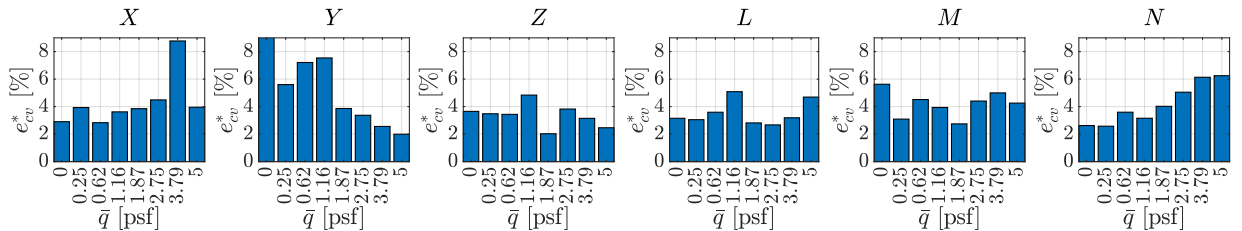

Fig. 14 Binomial analysis of residuals prediction error metric, e_{cv}^*, for each local model response.

VII. Conclusions

Electric VTOL vehicles present new challenges for aircraft modeling and are currently an important area of research. The mathematical models developed to characterize aero-propulsive phenomena for these complex aircraft must readily describe many control effectors and interactions, while also being amenable to drastically changing aerodynamics at numerous different flight conditions across a wide flight envelope. The result is a large modeling problem with significant nonlinearities and many interaction effects that requires an efficient experiment design within cost and time constraints. Design of experiments and response surface methodology theory provide efficient, statistically-rigorous experiment designs that enable accurate characterization of complex distributed propulsion aircraft.

This paper has presented multiple advances in eVTOL aircraft wind-tunnel testing using statistically-designed experiments enabling more efficient and accurate model development. The new wind-tunnel testing approaches were described in the context of the RAVEN-SWFT eVTOL vehicle that has recently been built and wind-tunnel tested at NASA Langley Research Center. A gravitational tare experiment design and modeling approach was developed to reduce test time by nearly 50% by eliminating the need for most gravitational tare runs. An efficient trim envelope determination strategy was applied across the testable transition envelope to quickly determine test factor ranges for subsequent aero-propulsive characterization experiments. Using test factor ranges reflecting the trimmable transition envelope, powered-airframe testing was conducted using a regular and nested *I*-optimal design at several dynamic pressure settings. This paper built on previous simulation work to provide the first experimental application of the nested *I*-optimal experiment design approach for complex aircraft characterization. The final aero-propulsive models created from the testing strategies discussed throughout the paper were shown to have good predictive capability across the tested flight envelope. Based on the successful modeling results, the core statistical testing advancements—gravitational tare modeling, trim envelope determination, and application of the nested *I*-optimal experiment design—are recommended for future wind-tunnel testing for complex aircraft configurations.

Acknowledgments

Powered-Airframe Testing Modeling and Validation Metric Tables

Table 10 Coefficient of determination, R^2, for each local model response (expressed as a percentage)

$\bar{q}$ [psf]	X	Y	Z	L	M	N
0	97.8	38.9	93.9	98.1	91.7	98.0
0.25	97.2	88.3	98.0	98.0	98.3	98.0
0.62	96.7	91.2	96.6	97.6	94.7	96.7
1.16	97.4	91.4	96.4	94.8	94.2	95.9
1.87	95.7	96.1	98.6	98.1	97.6	94.0
2.75	97.0	98.4	96.9	97.5	93.8	93.5
3.79	91.7	99.1	98.7	94.1	95.4	87.8
5	96.0	99.2	99.2	93.5	95.5	88.1

Table 11 Number of model parameters identified for each local model response

$\bar{q}$ [psf]	X	Y	Z	L	M	N
0	33	11	25	31	25	39
0.25	24	19	28	26	33	33
0.62	33	20	31	34	33	29
1.16	36	22	33	28	27	32
1.87	38	31	39	39	33	34
2.75	45	29	37	44	29	46
3.79	37	43	39	42	29	36
5	44	19	50	40	45	35

Table 12   NRMSE for each local model response (expressed as a percentage)

(a) Modeling data

$\bar{q}$ [psf]	X	Y	Z	L	M	N
0	2.28	13.11	3.43	2.22	4.67	2.18
0.25	2.47	5.01	2.23	2.07	1.96	1.97
0.62	2.64	5.36	2.78	2.27	3.31	2.70
1.16	2.43	5.70	2.95	3.34	3.45	3.18
1.87	3.13	3.77	1.90	1.79	2.12	3.64
2.75	2.96	2.62	3.09	1.93	3.44	3.65
3.79	4.99	2.07	2.10	2.62	3.09	4.39
5	2.96	1.88	1.82	3.29	3.33	4.49

(b) Validation data

$\bar{q}$ [psf]	X	Y	Z	L	M	N
0	2.10	10.27	2.97	1.95	4.06	1.88
0.25	2.52	4.26	2.37	2.11	1.82	1.89
0.62	1.98	4.74	2.44	2.38	3.22	2.62
1.16	2.43	4.39	3.09	3.14	3.28	2.78
1.87	2.57	3.17	1.89	2.01	1.98	3.16
2.75	3.38	2.38	2.64	1.71	3.31	3.13
3.79	5.63	1.86	2.08	2.55	3.27	3.99
5	2.69	1.58	1.76	3.16	2.85	4.06

Table 13   Binomial analysis of residuals prediction error metric, e_{cv}^*, for each local model response (expressed as a percentage)

$\bar{q}$ [psf]	X	Y	Z	L	M	N
0	2.90	13.88	3.65	3.15	5.63	2.62
0.25	3.93	5.60	3.48	3.05	3.09	2.57
0.62	2.83	7.21	3.44	3.59	4.51	3.59
1.16	3.62	7.54	4.84	5.08	3.94	3.15
1.87	3.85	3.86	2.03	2.81	2.74	4.02
2.75	4.49	3.37	3.82	2.67	4.40	5.05
3.79	8.77	2.56	3.15	3.19	5.00	6.13
5	3.95	1.99	2.46	4.69	4.26	6.25

References

[1] Antcliff, K. R., Whiteside, S. K. S., Kohlman, L. W., and Silva, C., "Baseline Assumptions and Future Research Areas for Urban Air Mobility Vehicles," *AIAA SciTech 2019 Forum*, AIAA Paper 2019-0528, Jan. 2019. https://doi.org/10.2514/6.2019-0528.

[2] Patterson, M. D., Antcliff, K. R., and Kohlman, L. W., "A Proposed Approach to Studying Urban Air Mobility Missions Including an Initial Exploration of Mission Requirements," *AHS International 74th Annual Forum & Technology Display*, May 2018.

[3] "eVTOL Aircraft Directory," *Electric VTOL News™*, https://evtol.news/aircraft, Accessed 01 November 2023.

[4] Johnson, W., Silva, C., and Solis, E., "Concept Vehicles for VTOL Air Taxi Operations," *AHS Technical Conference on Aeromechanics Design for Transformative Vertical Flight*, Jan. 2018.

[5] Silva, C., Johnson, W., Antcliff, K. R., and Patterson, M. D., "VTOL Urban Air Mobility Concept Vehicles for Technology Development," *2018 Aviation Technology, Integration, and Operations Conference*, AIAA Paper 2018-3847, Jun. 2018. https://doi.org/10.2514/6.2018-3847.

[6] Johnson, W., and Silva, C., "NASA Concept Vehicles and the Engineering of Advanced Air Mobility Aircraft," *The Aeronautical Journal*, Vol. 126, No. 1295, 2022, pp. 59–91. https://doi.org/10.1017/aer.2021.92.

[7] Rothhaar, P. M., Murphy, P. C., Bacon, B. J., Gregory, I. M., Grauer, J. A., Busan, R. C., and Croom, M. A., "NASA Langley Distributed Propulsion VTOL Tilt-Wing Aircraft Testing, Modeling, Simulation, Control, and Flight Test Development," *14th AIAA Aviation Technology, Integration, and Operations Conference*, AIAA Paper 2014-2999, Jun. 2014. https://doi.org/10.2514/6.2014-2999.

[8] Saeed, A. S., Younes, A. B., Cai, C., and Cai, G., "A Survey of Hybrid Unmanned Aerial Vehicles," *Progress in Aerospace Sciences*, Vol. 98, 2018, pp. 91–105. https://doi.org/10.1016/j.paerosci.2018.03.007.

[9] Kim, H. D., Perry, A. T., and Ansell, P. J., "A Review of Distributed Electric Propulsion Concepts for Air Vehicle Technology," *2018 AIAA/IEEE Electric Aircraft Technologies Symposium*, AIAA Paper 2018-4998, Jul. 2018. https://doi.org/10.2514/6.2018-4998.

[10] North, D. D., Busan, R. C., and Howland, G., "Design and Fabrication of the Langley Aerodrome No. 8 Distributed Electric Propulsion VTOL Testbed," *AIAA SciTech 2021 Forum*, AIAA Paper 2021-1188, Jan. 2021. https://doi.org/10.2514/6.2021-1188.

[11] Murphy, P. C., Buning, P. G., and Simmons, B. M., "Rapid Aero Modeling for Urban Air Mobility Aircraft in Computational Experiments," *AIAA SciTech 2021 Forum*, AIAA Paper 2021-1002, Jan. 2021. https://doi.org/10.2514/6.2021-1002.

[12] Simmons, B. M., Buning, P. G., and Murphy, P. C., "Full-Envelope Aero-Propulsive Model Identification for Lift+Cruise Aircraft Using Computational Experiments," *AIAA AVIATION 2021 Forum*, AIAA Paper 2021-3170, Aug. 2021. https://doi.org/10.2514/6.2021-3170.

[13] Busan, R. C., Rothhaar, P. M., Croom, M. A., Murphy, P. C., Grafton, S. B., and O'Neal, A. W., "Enabling Advanced Wind-Tunnel Research Methods Using the NASA Langley 12-Foot Low Speed Tunnel," *14th AIAA Aviation Technology, Integration, and Operations Conference*, AIAA Paper 2014-3000, Jun. 2014. https://doi.org/10.2514/6.2014-3000.

[14] Murphy, P. C., and Landman, D., "Experiment Design for Complex VTOL Aircraft with Distributed Propulsion and Tilt Wing," *AIAA Atmospheric Flight Mechanics Conference*, AIAA Paper 2015-0017, Jan. 2015. https://doi.org/10.2514/6.2015-0017.

[15] Murphy, P. C., Simmons, B. M., Hatke, D. B., and Busan, R. C., "Rapid Aero Modeling for Urban Air Mobility Aircraft in Wind-Tunnel Tests," *AIAA SciTech 2021 Forum*, AIAA Paper 2021-1644, Jan. 2021. https://doi.org/10.2514/6.2021-1644.

[16] Busan, R. C., Murphy, P. C., Hatke, D. B., and Simmons, B. M., "Wind Tunnel Testing Techniques for a Tandem Tilt-Wing, Distributed Electric Propulsion VTOL Aircraft," *AIAA SciTech 2021 Forum*, AIAA Paper 2021-1189, Jan. 2021. https://doi.org/10.2514/6.2021-1189.

[17] Simmons, B. M., and Murphy, P. C., "Aero-Propulsive Modeling for Tilt-Wing, Distributed Propulsion Aircraft Using Wind Tunnel Data," *Journal of Aircraft*, Vol. 59, No. 5, 2022, pp. 1162–1178. https://doi.org/10.2514/1.C036351.

[18] Simmons, B. M., Morelli, E. A., Busan, R. C., Hatke, D. B., and O'Neal, A. W., "Aero-Propulsive Modeling for eVTOL Aircraft Using Wind Tunnel Testing with Multisine Inputs," *AIAA AVIATION 2022 Forum*, AIAA Paper 2022-3603, Jun. 2022. https://doi.org/10.2514/6.2022-3603.

[19] Montgomery, D. C., *Design And Analysis of Experiments*, 8th ed., John Wiley & Sons, Inc., Hoboken, NJ, 2013.

[20] Myers, R. H., Montgomery, D. C., and Anderson-Cook, C. M., *Response Surface Methodology: Process and Product Optimization Using Designed Experiments*, 4th ed., John Wiley & Sons, 2016.

[21] German, B. J., Jha, A., Whiteside, S. K. S., and Welstead, J. R., "Overview of the Research Aircraft for eVTOL Enabling techNologies (RAVEN) Activity," *AIAA AVIATION 2023 Forum*, AIAA Paper 2023-3924, June 2023. https://doi.org/10.2514/6.2023-3924.

[22] "NASA Langley 12-Foot Low-Speed Tunnel," https://researchdirectorate.larc.nasa.gov/12-foot-low-speed-tunnel-12-ft-lst/, Accessed 18 April 2023.

[23] McSwain, R. G., Geuther, S. C., Howland, G., Patterson, M. D., Whiteside, S. K., and North, D. D., "An Experimental Approach to a Rapid Propulsion and Aeronautics Concepts Testbed," NASA TM–2020-220437, Jan. 2020.

[24] Geuther, S. C., and Fei, X., "LA-8 Computational Analysis and Validation Studies Using FlightStream," *AIAA SciTech 2021 Forum*, AIAA Paper 2021-1191, Jan. 2021. https://doi.org/10.2514/6.2021-1191.

[25] Ahuja, V., and Hartfield, R. J., "Reduced-Order Aerodynamics with the Method of Integrated Circulation," *AIAA SciTech 2022 Forum*, AIAA Paper 2022-0005, Jan. 2022. https://doi.org/10.2514/6.2022-0005.

[26] Geuther, S. C., North, D. D., and Busan, R. C., "Investigation of a Tandem Tilt-wing VTOL Aircraft in the NASA Langley 12-Foot Low-Speed Tunnel," NASA TM–2020–5003178, Jun. 2020.

[27] Simmons, B. M., "System Identification for Propellers at High Incidence Angles," *Journal of Aircraft*, Vol. 58, No. 6, 2021, pp. 1336–1350. https://doi.org/10.2514/1.C036329.

[28] Simmons, B. M., and Hatke, D. B., "Investigation of High Incidence Angle Propeller Aerodynamics for Subscale eVTOL Aircraft," NASA TM–20210014010, May 2021.

[29] Stratton, M., and Landman, D., "Wind Tunnel Test and Empirical Modeling of Tilt-Rotor Performance for eVTOL Applications," *AIAA SciTech 2021 Forum*, AIAA Paper 2021-0834, Jan. 2021. https://doi.org/10.2514/6.2021-0834.

[30] Simmons, B. M., "Evaluation of Response Surface Experiment Designs for Distributed Propulsion Aircraft Aero-Propulsive Modeling," *AIAA SciTech 2023 Forum*, AIAA Paper 2023-2251, Jan. 2023. https://doi.org/10.2514/6.2023-2251.

[31] Cooper, J. R., Ackerman, K. A., Rothhaar, P. M., and Gregory, I. M., "Autonomous Path-Following for a Tilt-Wing, Distributed Electric Propulsion, Vertical Take-Off and Landing Unmanned Aerial System in Hover Mode," NASA TM–2018–220109, Nov. 2018.

[32] Cook, J., and Gregory, I., "A Robust Uniform Control Approach for VTOL Aircraft," *VFS Autonomous VTOL Technical Meeting and Electric VTOL Symposium*, Jan. 2021.

[33] McSwain, R. G., Glaab, L. J., and Theodore, C. R., "Greased Lightning (GL-10) Performance Flight Research – Flight Data Report," NASA TM–2017–219794, Nov. 2017.

[34] Fredericks, W. J., McSwain, R. G., Beaton, B. F., Klassman, D. W., and Theodore, C. R., "Greased Lightning (GL-10) Flight Testing Campaign," NASA TM–2017–219643, Jul. 2017.

[35] Simmons, B. M., "System Identification Approach for eVTOL Aircraft Demonstrated Using Simulated Flight Data," *Journal of Aircraft*, Vol. 60, No. 4, 2023, pp. 1078–1093. https://doi.org/10.2514/1.C036896.

[36] Simmons, B. M., "Efficient Variable-Pitch Propeller Aerodynamic Model Development for Vectored-Thrust eVTOL Aircraft," *AIAA AVIATION 2022 Forum*, AIAA Paper 2022-3817, Jun. 2022. https://doi.org/10.2514/6.2022-3817.

[37] Goos, P., and Jones, B., *Optimal Design of Experiments*, John Wiley & Sons, West Sussex, United Kingdom, 2011.

[38] Anderson, M. J., and Whitcomb, P. J., *RSM Simplified: Optimizing Processes Using Response Surface Methods for Design of Experiments*, 2nd ed., CRC Press, Boca Raton, FL, 2017. https://doi.org/10.1201/9781315382326.

[39] "Design-Expert®," Version 13, Stat-Ease, Inc., https://www.statease.com/software/design-expert/, Accessed 11 May 2022.

[40] Zahran, A., Anderson-Cook, C. M., and Myers, R. H., "Fraction of Design Space to Assess Prediction Capability of Response Surface Designs," *Journal of Quality Technology*, Vol. 35, No. 4, 2003, pp. 377–386. https://doi.org/10.1080/00224065.2003.11980235.

[41] Anderson, M. J., Adams, W. F., and Whitcomb, P. J., "How to Properly Size Response Surface Method Experiment (RSM) Designs for System Optimization," 2016. https://cdn.statease.com/media/public/documents/how_to_properly_size_experiments_aimed_at_system_optimization.pdf, Accessed 17 October 2022.

[42] Whitcomb, P., "FDS—A Power Tool for Designers of Optimization Experiments," *Stat-Teaser Newsletter*, Stat-Ease, Inc., Sep. 2008, pp. 1–3. https://cdn.statease.com/media/public/documents/statteaser-0908.pdf, Accessed 17 October 2022.

[43] Morelli, E. A., and Klein, V., *Aircraft System Identification: Theory and Practice*, 2nd ed., Sunflyte Enterprises, Williamsburg, VA, 2016.

[44] "System IDentification Programs for AirCraft (SIDPAC)," *NASA Technology Transfer Program*, https://software.nasa.gov/software/LAR-16100-1, Accessed 03 September 2022.

[45] Klein, V., Batterson, J. G., and Murphy, P. C., "Determination of Airplane Model Structure from Flight Data by Using Modified Stepwise Regression," NASA TP-1916, Oct. 1981.

[46] DeLoach, R., "Assessment of Response Surface Models Using Independent Confirmation Point Analysis," *48th AIAA Aerospace Sciences Meeting Including the New Horizons Forum and Aerospace Exposition*, AIAA Paper 2010-741, Jan. 2010. https://doi.org/10.2514/6.2010-741.

[47] Montgomery, D. C., Peck, E. A., and Vining, G. G., *Introduction to Linear Regression Analysis*, 5th ed., John Wiley & Sons, Hoboken, New Jersey, 2012.

[48] Landman, D., Simpson, J., Mariani, R., Ortiz, F., and Britcher, C., "Hybrid Design for Aircraft Wind-Tunnel Testing Using Response Surface Methodologies," *Journal of Aircraft*, Vol. 44, No. 4, 2007, pp. 1214–1221. https://doi.org/10.2514/1.25914.

[49] Fritsch, F. N., and Carlson, R. E., "Monotone Piecewise Cubic Interpolation," *SIAM Journal on Numerical Analysis*, Vol. 17, No. 2, 1980, pp. 238–246. https://doi.org/10.1137/0717021.

[50] Moler, C. B., *Numerical Computing with MATLAB*, Society for Industrial and Applied Mathematics, Philadelphia, PA, 2004. https://doi.org/10.1137/1.9780898717952.