Design of Experiments Development for Rotorcraft Acoustic Flight Testing

James H. Stephenson
U.S. Army Combat Capabilities Development Command
Aviation & Missile Center
Hampton, VA, USA

Kyle A. Pascioni
NASA Langley Research Center
Hampton, VA, USA

ABSTRACT

A cooperative acoustics flight test campaign between the US Army and NASA was performed in which design of experiments (DOE) approaches were used to plan the flight test conditions. Three DOE designs were used, a face centered central composite design, circumscribed central composite design, and a hexagonal design. A traditional one-factor-at-a-time approach was also used, and interpolation points were planned to test for the strength of the DOE approaches. This paper documents the design methodology, discusses how response surface models were fit to the data, evaluates the overall response of the models, and evaluates the individual DOE designs. The response surface models were also used to design new test conditions of interest during the experiment, and that process is also documented. For the first time, DOE was shown to be an exceptionally useful tool for rotorcraft acoustics flight test planning, while the full power of the approach has yet to be reached.

INTRODUCTION

A comprehensive acoustics research flight test was conducted by the US Army's DEVCOM Aviation & Missile Center (AvMC) and the NASA Langley Research Center (LaRC) using an MD530F vehicle (Ref. 1). One of the primary objectives of the flight test was to investigate the use of design of experiments (DOE) for selecting flight test conditions. Traditional flight test design investigates one-factor-at-a-time (OFAT) and seeks to build a database of measurements that reasonably spans the entirety of the flight envelope that can be maintained steadily across a microphone array. This approach suffers from several limitations, but the major concern is that it requires a relatively large number of test conditions be measured. The required number of test conditions grows with increasing complexity of the physics under investigation. Design of experiments has been shown to be effective at reducing the measurement scope of experimental data sets in rotorcraft aerodynamics as well as other complex research fields (Refs. 2–6).

Let us assume that the speed and flight path angle (FPA) of a vehicle are all that are important for characterizing the acoustic emissions of a particular rotorcraft. This is what land-use modeling software, such as the Advanced Acoustic Model (Ref. 7), assumes today. In that case, measuring the vehicle at five flight speeds and five different flight path angles would result in a test matrix that has 25 individual conditions to measure. However, it is known that weight of the vehicle can influence the acoustics. In the OFAT approach, the researcher would repeat all 25 of those speed and FPA conditions at each vehicle weight of interest. Even simplifying and assuming that the vehicle flies in a 'heavy' and 'light' condition, that approach doubles the required measurement points. This problem is further exacerbated as the number of measurement parameters increases. We know that traditional helicopter noise can be nondimensionalized by four unique parameters (Ref. 8), and so the required test matrix should be quite large (Ref. 9). When Urban Air Mobility (UAM) vehicles are considered, with multiple propulsors and multiple potential configurations, then the measurement space can quickly become intractable. When considering meeting the needs of Future Vertical Lift (FVL) or UAM vehicles, OFAT testing cannot continue, and a new system for choosing adequate test conditions must be developed.

This paper describes several potential design of experiments approaches, provides analysis of the individual designs, and discusses in-field optimizations. The intent of this experiment is to provide more guidance for future rotorcraft acoustic flight test design development.

FLIGHT TEST DESCRIPTION

This data set comes from a comprehensive flight test that was conducted to address multiple unique research goals simultaneously. The flight test is described completely in Ref. 1, with specifics provided in Refs. 10–12. Only the relevant data necessary to discuss the DOE data set will be presented here, and the reader is referred to the other references for more complete details.

The test utilized a fully instrumented MD530F helicopter, shown in Figure 1 (Refs. 1, 13, 14). This is a light-utility

The Vertical Flight Society's 81st Annual Forum and Technology Display, Virginia Beach, VA, May 20-22, 2025.

Figure 1: MD530F vehicle as configured during flight test.

civilian vehicle with a single 350 HP, turboshaft, Rolls-Royce 250-C30 engine; additional MD530F aircraft characteristics are shown in Table 1. Aircraft positional and inertial state data were measured with NASA's Aircraft Navigation and Tracking System (ANTS), and are described in Ref. 1. This vehicle was outfitted with a significant suite of on-board instrumentation that is also described in Ref. 1.

Table 1: Aircraft specifications.

No. Main Rotor Blades	5
Main Rotor RPM, BPF	477, 39.75 Hz
No. Tail Rotor Blades	2
Tail Rotor RPM, BPF	2,848, 94.9 Hz

Microphone Instrumentation

This flight test had two primary microphone arrays (a snapshot array and a phased array). The snapshot array is the focus of the DOE work, and so is described here briefly. The snapshot array was comprised entirely of Wireless Acoustic Measurement System version II (WAMS II) systems, which are remotely controlled with local data storage. Each unit consists of a microphone, ground board, radio antenna, GPS receiver, and on-board SD card for data recording. Microphones 1 through 75 use the standard WAMS II setup for this test, which consisted of a GRAS 67AX microphone embedded in a 400 mm diameter round ground board. However, microphones 76 through 79 used an inverted B&K 4964 microphone above a 381 mm diameter ground board. In these configurations, microphones are offset from the center of the ground board to minimize edge effects, per SAE Aerospace Recommended Practice (Ref. 15). All microphones are sampled simultaneously and uninterrupted throughout a run at 50 kHz with 24-bit resolution, and are time synchronized with vehicle instrumentation using Coordinated Universal Time.

The snapshot array was designed to create hemispheres, i.e., to capture the acoustic levels at all emission angles below the aircraft, sufficient for land-use planning models such as the Advanced Acoustic Model and the Aircraft Noise Prediction Program 2 (Refs. 7, 16). The snapshot array was designed using a modified 8th order Lebedev distribution (Ref. 17) in a manner to provide significant coverage of a hemisphere as the vehicle passes over the center of the array. This array layout design is described in more details in Ref. 10. Microphone locations on the ground are shown in Figure 2. Microphones

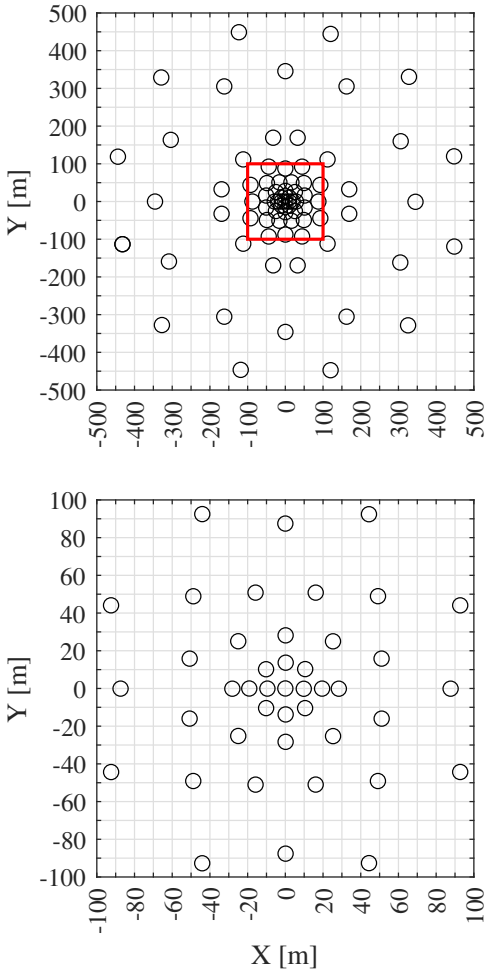


Figure 2: Snapshot array microphone locations shown on the ground.

were generally numbered sequentially counter clockwise and radially from the farthest out to the innermost microphones. A Cartesian coordinate system is used with microphone 79 defined as the center of the coordinate system for the snapshot array ('X' = 'Y' = 'Z' = 0). The coordinate system is defined such that 'X' is along the flight track and is positive in the primary flight direction; 'Y' is defined perpendicular to the flight track and is positive to the aircraft port (left) side; 'Z' is positive up.

Figure 3 shows the microphone locations on the hemisphere using a Lambert projection. The Lambert projection is defined such that azimuth is represented by counter-clockwise rotation around the figure, in the direction of the main rotor

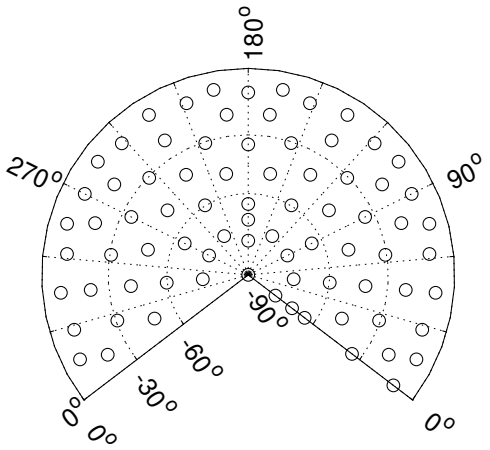


Figure 3: Microphone positions projected on a hemisphere, using a Lambert projection, for the designed flight condition.

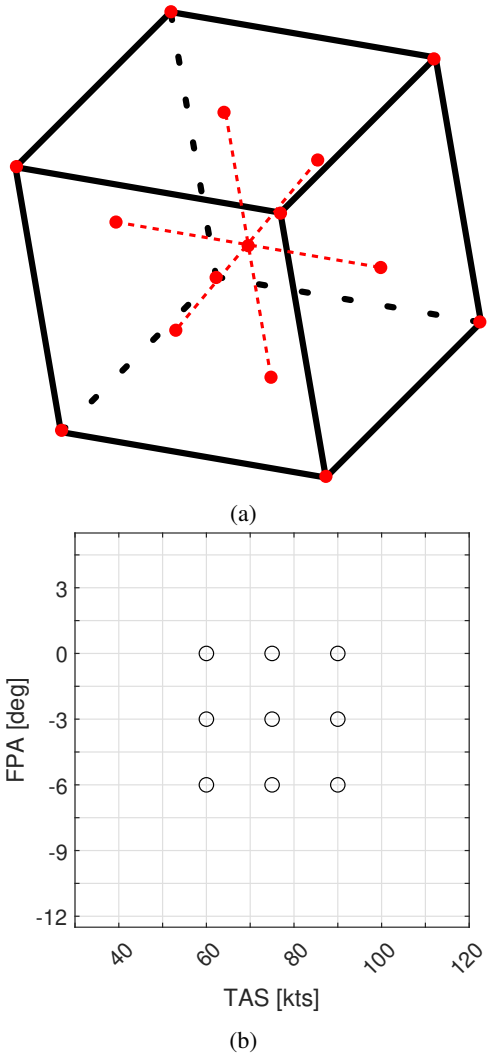
as seen from above. Elevation in the Lambert projection is represented in the radial direction, such that the center of the figure is directly beneath the rotor, while the edge is the horizon plane.

Flight Test Condition Design

A balanced approach to meet as many objectives simultaneously as possible was taken for the design of the flight test conditions, while also remaining well within the flight envelope of the vehicle and constrained by available flight time. In the end, three design of experiment techniques were deployed along with a more traditional one-factor-at-a-time (OFAT) approach to designing test conditions. Only two flight condition parameters were varied, true airspeed measured in knots (TAS) and flight path angle (FPA) measured in degrees. The center point of the design of experiments (75 KTAS, $-3°$ FPA) was chosen to be colocated with all designs, providing a stable location from which to deviate.

The first design of experiments approach used a face centered central composite design (FCCD), shown in Fig. 4. The FCCD places test conditions at the corners of a cube (box in two dimensions) with additional points at the center of the cube and the center of each face of the cube. In two dimensions, as tested here, that results in 9 unique test conditions in a three-by-three pattern, which appears to collapse to a form of OFAT in this case. The size of the box was chosen such that test conditions were located in level flight (where rotorcraft spend much of their time), and the highest speed descent condition comfortably achievable was chosen as the bottom right corner. These two conditions determined the center of the design as well as the extents of the box.

The circumscribed central composite design (CCCD) is similar, and shown in Fig. 5b. The CCCD again places test conditions at the center and on the corners of a cube, but instead of placing the remaining points on the center of the face of the



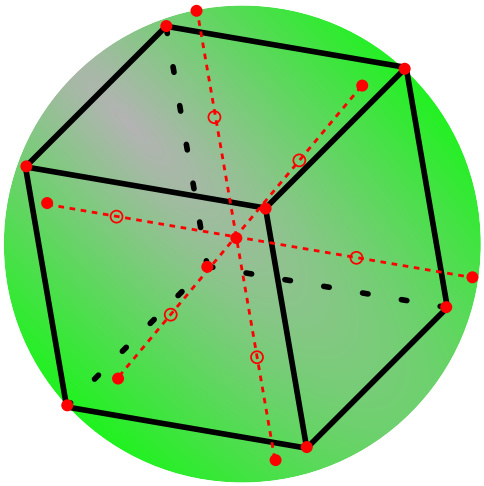
(a)

(b)

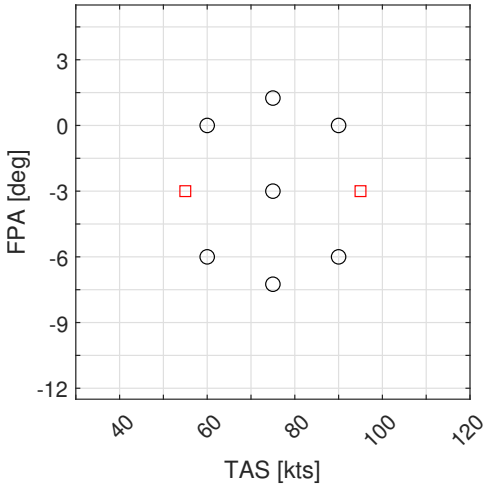
Figure 4: (a) Generic face centered central composite design where red circles show the test prescribed conditions. (b) a 2-D implementation of FCCD used for this flight test.

cube, it circumscribes a sphere around the cube and places the test conditions on the surface of that sphere normal to the center of the cube's faces (Fig. 5a). In the CCCD design, two test conditions were removed at the $-3°$ FPA level, as they were only 5 knots different than the FCCD test conditions shown in Fig. 4b, and it was judged that this difference was likely not enough (given experimental errors in piloting and weather) to justify the additional flight test hours to measure. So, for the CCCD conditions used later, the FCCD test points located at (60 KTAS, $-3°$ FPA) and (90 KTAS, $-3°$ FPA) are employed instead of the ideal conditions at 55 and 95 KTAS. The CCCD approach is expected to yield slightly better predictive values in climb conditions than the FCCD design, as the physics of climb will be represented in the final model.

The final DOE design used was a hexagonal approach. While the central composite designs provide effectively three levels



(a)



(b)

Figure 5: (a) Generic circumscribed central composite design where solid red circles show the prescribed test conditions. Open red circles are left to show the slight change over the design in the FCCD approach from Fig. 4a. (b) a 2-D implementation of CCCD used for this flight test, where the red squares are neglected in the final design.

in each direction (true airspeed, flight path angle), the hexagonal approach (shown in Fig. 6) has 5 levels in true airspeed and 3 levels in flight path angle. While the central point was maintained, this design was chosen to have a larger radius and to cover more of the flight envelope of the vehicle. The obvious drawback to this design is that it placed no test conditions in level flight, where vehicles spend the majority of their time. The hexagonal approach also does not extend to multi-dimensional conditions and so has limited use when considering applications to future vehicle concepts.

Finally, all of the test conditions were brought together and additional conditions chosen to complete the traditional OFAT approach (shown in Fig. 7). Here, two additional test conditions were chosen to fill out a reasonable spread of level flight speeds, two conditions were chosen at slow speed (40 KTAS) but steeper descents ($-6°$ and $-9°$), and a third condition at a higher speed climb condition (75 KTAS, 3° FPA). Then, a higher speed descent condition (75 KTAS, $-9°$ FPA) was chosen to round out the OFAT conditions.

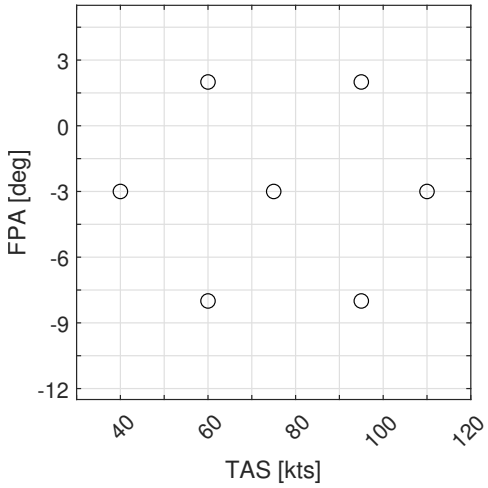


Figure 6: Hexagonal design of experiments approach used during flight test.

The purpose of DOE implementation is to enhance test efficiency, and so a response surface modeling (RSM) methodology was used. In order to test the accuracy of the RSM, it must be checked against extrapolation and interpolation data points. The OFAT test conditions serve as adequate extrapolation points for all DOE designs, but no interpolation points were available interior to the central composite designs. Thus, two conditions were selected to be inside of each of the DOE designs to test for the ability to interpolate on the interior portion of any response surface model that might be developed. In total, 25 unique test conditions were developed for the flight test that spanned all of the myriad requirements for the snapshot array flight days.

At this stage, all test conditions have been identified and prioritized, with DOE conditions given priority zero, meaning they were mission critical. Test conditions were assigned condition codes ('C' for climb, 'L' for level flight, 'D' for descent) and flown during the experiment with respect to their pre-determined priority levels. Table 2 identifies the test condition codes, airspeed and flight path angle specifications, as well as the number of times flown throughout the first flight test day.

GENERALIZED LINEAR MIXED MODELING

Acoustic response surface models can now be built following the process previously documented in Ref. 18. From the work shown in Ref. 18 and follow-on analysis, each individual mi-

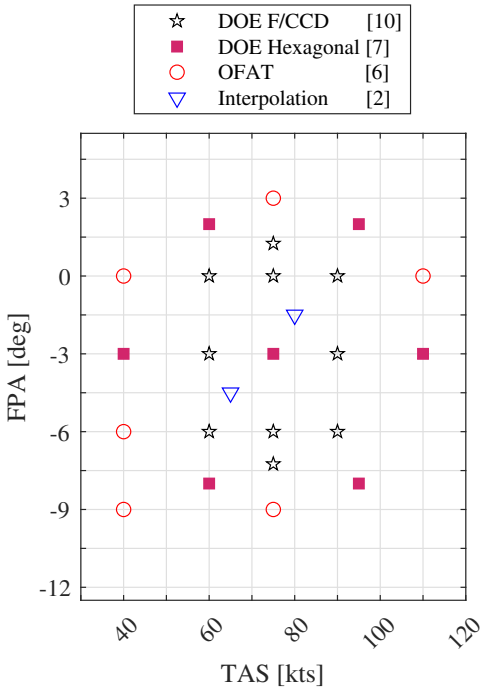


Figure 7: Complete flight test condition design showing central composite designs, hexagonal design, interpolation points, as well as the remaining traditional one-factor-at-a-time conditions needed to complete the matrix.

crophone measurement can be reasonably described using a quadratic response surface model based on Equation 1.

$$y \backsim V + \gamma + V^2 + \gamma^2 + V * \gamma \tag{1}$$

Here, 'y' is the acoustic metric of interest, and (V, γ) are TAS and FPA, respectively. The dimensional models (V, γ) were shown to be just as powerful as, if not more accurate than, the non-dimensional models in Ref. 18. As these parameters are significantly simpler to implement during the flight test, they were the primary focus. The metrics investigated here are the same used in Ref. 18, namely:

- overall sound pressure level (OASPL)
- A-weighted overall sound pressure level (OASPLA)
- blade-vortex interaction sound pressure level (BVISPL)
- perceived noise level (PNL)
- tone-corrected perceived noise level (PNLT).
- effective perceived noise level (EPNL)
- sound exposure level (SEL)

BVISPL was defined as a frequency integration from 200 Hz to 800 Hz, while OASPL was limited to frequencies from 20 Hz to 11 kHz. The remaining metrics are defined in their traditional sense.

Table 2: Flight condition names, prescribed flight conditions, and number of measured runs at that condition.

Condition Code	TAS [kts]	FPA [°]	Runs
C1	75	3	5
C2	60	2	5
C3	95	2	5
C4	75	1.25	5
L1	40	0	4
L2	60	0	5
L3	75	0	5
L4	90	0	9
L5	110	0	4
D1	80	-1.5	3
D2	40	-3	5
D3	60	-3	5
D4	75	-3	5
D5	90	-3	5
D6	110	-3	5
D7	65	-4.5	3
D8	40	-6	3
D9	60	-6	5
D10	75	-6	5
D11	90	-6	5
D12	75	-7.25	5
D13	60	-8	5
D14	95	-8	6
D15	40	-9	3
D16	75	-9	3

The RSMs using Eq. 1 are built unique for each metric (6) and each microphone (79). During the flight test this was done in a unique manner for only one half-second spectrum, while the vehicle crossed the center of the microphone array. The crossover point was determined not through GPS tracking data, as those were not available until after the test was complete, but by using the time when the center microphone recorded its maximum sound pressure level. This method was verified to be within one-half of a second of the actual crossover time using the experiment from Ref. 18 and so was deemed sufficiently accurate. Verification post-test showed on average this assumption was accurate within one-quarter of a second for the current experiment. Each metric from each microphone was investigated without correcting for spherical spreading as well as correcting for spherical spreading to the assumed crossover altitude for each run. Thus, the response surface models were based on a ground plane of data, and an approximate hemisphere of data, respectively.

The RSMs were built using normalized TAS and FPA as well

as dimensional quantities. Normalization of the variables can be used to reduce the influence of one variable over the other if a small range is used for one variable. Normalization was done each day, such that the minimum value prescribed for that day was set to 0, and maximum value was set to 1 and internal numbers were linearly mapped accordingly. Fortunately, the normalized variables resulted in identical results as the dimensional variables, and so only dimensional results will be used here.

The RSMs were built using an identity link function assuming a normal distribution of the metrics, as well as assuming a gamma distribution of the metrics. Ref. 18 showed that a normally distributed metric assumption was accurate within approximately 1.5 dB, while subsequent research that was presented at the conference showed that the gamma distribution resulted in an accuracy within 0.015 dB. For the majority of the analysis in this paper, dimensional predictor variables are used with an assumption of a gamma distribution in the response variables.

FLIGHT TEST APPLICATION

The flight test conditions were defined and the groundwork for building of the response surface models was in place prior to the start of the flight test. This meant that as soon as the first day of flight testing was accomplished, the RSMs were built and interrogated to begin defining new test conditions for subsequent days. The first day of testing for the conventional MD530F measured the conditions found in Table 3. Here, it should be noted that 'L6' and 'H1' (a hover condition) were added to assist in analysis with the phased array experiment documented in Ref. 11. The 'H1' case was not used in this present work.

During the flight test, the RSMs were built using all data that were available at the end of the day. There was insufficient time to evaluate one DOE design over another and still execute an optimization technique to determine extra data points necessary to measure on the following flight test day. Thus, 79 RSMs were built – one for each microphone – using the overflight data from the runs captured in Table 3, with the exception of 'H1' condition.

RSMs were built from the metric data collected from each microphone during the overpass, as well as by averaging metric data across the repeated flights of the same condition. Presently, investigation of only the averaged conditions will be discussed.

Several microphones were chosen for investigation of their response surface models. Figure 8 shows the OASPL metric modeled for microphone 13, which was ahead of the vehicle and slightly below the horizon plane. The RSM shows, in this direction, for a non-distance corrected signal, that the maximum noise is radiated at steep descent angles and very high speeds. Minimum noise for this directivity is located near level flight at very low speeds.

Table 3: Flight conditions measured during the first day of conventional MD530F testing.

Condition Code	TAS [kts]	FPA [°]	Runs
C1	75	3	3
C2	60	2	3
C3	95	2	3
C4	75	1.25	3
L6	20	0	3
L1	40	0	3
L2	60	0	3
L3	75	0	3
L4	90	0	4
L5	110	0	3
D1	80	-2	3
D2	40	-3	3
D3	60	-3	3
D4	75	-3	3
D5	90	-3	3
D6	110	-3	3
D7	65	-4	3
D8	40	-6	3
D9	60	-6	3
D10	75	-6	3
D11	90	-6	3
D12	75	-7.25	3
D13	60	-8	3
D14	95	-8	4
D15	40	-9	3
D16	75	-9	3
H1	0	0	1

The magnitude of the gradient of the RSM shown in Fig. 8 is provided in Fig. 9. From this figure, it appears that the highest rates of change relative to each axis are at the edges of the interrogation window (lowest and highest possible speeds) and could have been predicted from the quadratic assumption of the generalized linear model (Eq. 1). The use of the gradient method must be employed in the future with a higher order response model where gradients would provide more useful information.

Similar data can be found in Figure 10, which provides the BVISPL and the gradient of BVISPL from microphone 13. Here it is reinforced that peak BVI occurs for this vehicle in shallow descents and slow speed flights, while the highest rates of change occur near the edges of the interrogation window at the extremes in flight path angles.

This process of fitting RSMs and evaluating the gradient was automated for all microphones and all metrics. The maximum value of the RSM and the maximum value of the gradient of the RSM was identified for all runs in Table 3. This

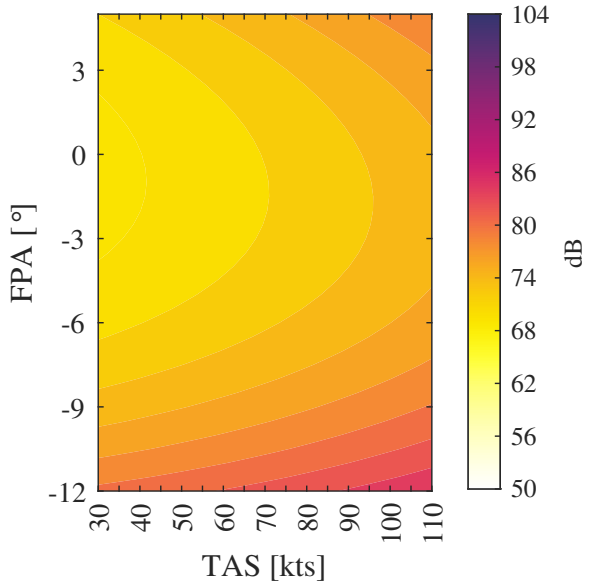


Figure 8: RSM using a generalized linear model with gamma distribution for a microphone ahead of the vehicle (microphone 13) and OASPL metric.

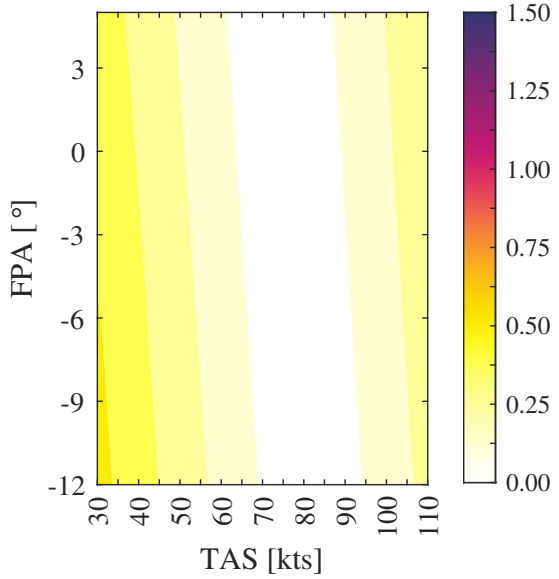


Figure 9: The magnitude of the gradient of the RSM from Fig. 8.

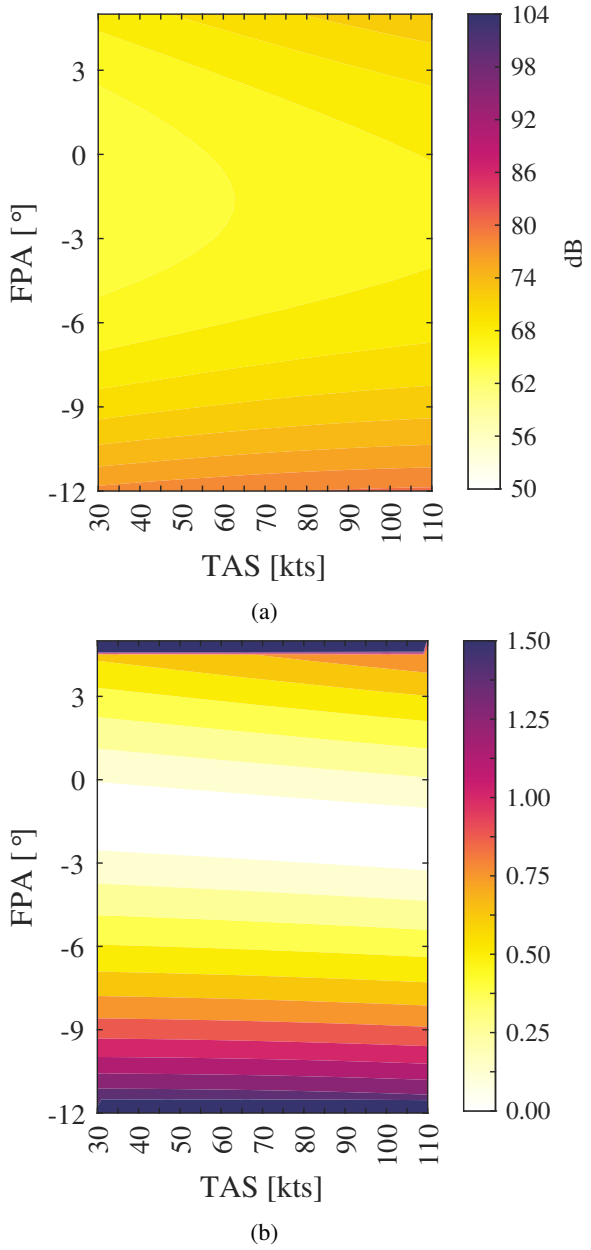


Figure 10: RSM using a generalized linear model with gamma distribution for a microphone ahead of the vehicle and (a) BVISPL metric and (b) gradient thereof.

resulted in multiple figures, such as Figure 11, which shows the flight condition resulting in maximum BVISPL predicted for each of the 79 microphones. Not all microphones experience strong BVI, and so predictions from some of these microphones will not yield useful information. In a holistic sense, however, this figure makes it clear that the -3 degree descent conditions are primary drivers of BVISPL above 60 knots, with slower speeds requiring higher descent angles. It should be noted that the range of frequencies investigated for BVISPL includes significant tail rotor contributions, and so the tail rotor may be contributing to this conclusion.

Figure 12 shows the location of maximum derivative of the BVISPL RSM, based on a gamma distribution of the response variable. It is clear that the maxima are stacked near the extremes of the interrogation window and are likely not providing significant useful information, but it does suggest low speed climbs and descents are of potential interest. However, Figure 13, shows the location of maximum derivative of the BVISPL RSM, based on a gamma distribution of the response variable and non-dimensionalized predictor variables. Here,

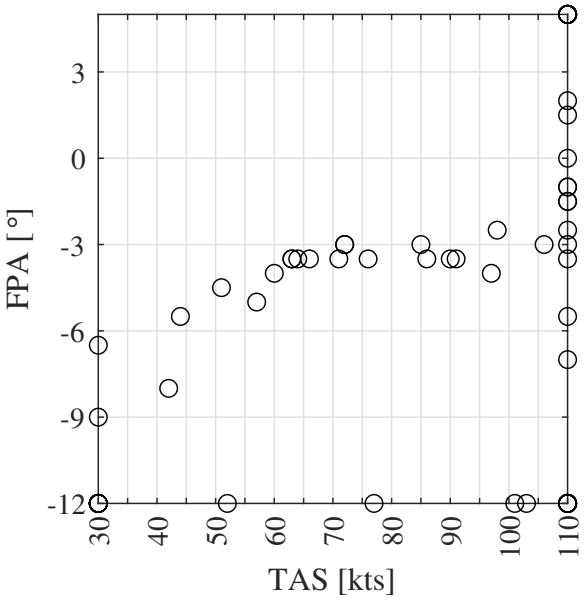


Figure 11: Maximum BVISPL predicted by the RSM for each microphone, trained on all data from Table 3, assuming a gamma distribution of the response variable.

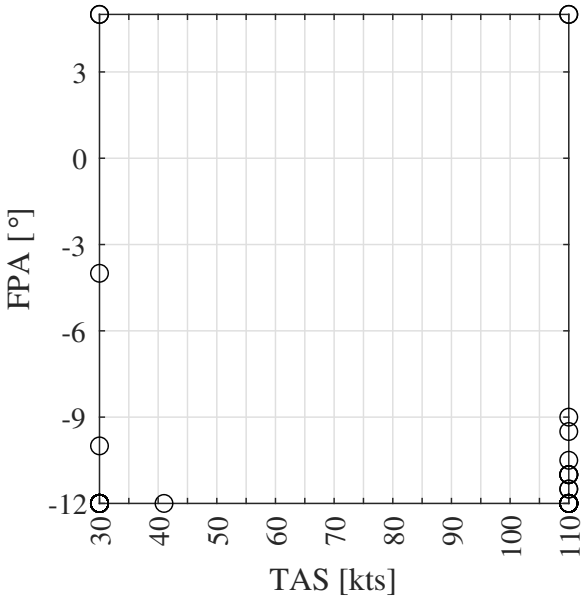


Figure 13: Maximum BVISPL gradient of each RSM for each microphone, trained on all data from Table 3, assuming a gamma distribution of the response variable and non-dimensionalized predictor variables.

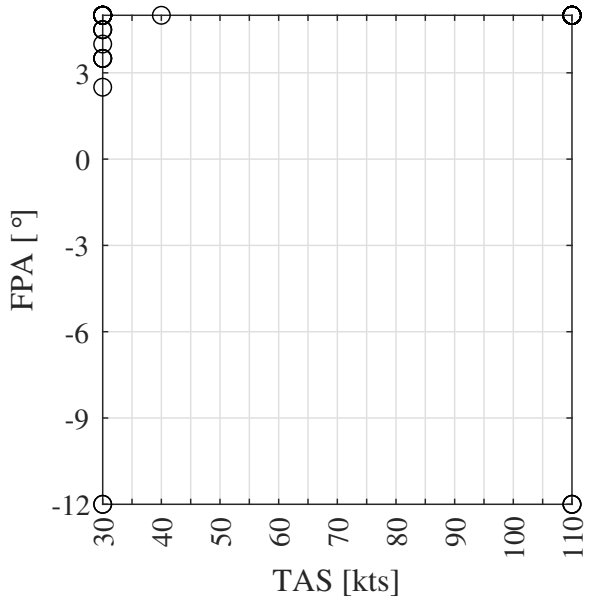


Figure 12: Maximum BVISPL gradient of each RSM for each microphone, trained on all data from Table 3, assuming a gamma distribution of the response variable.

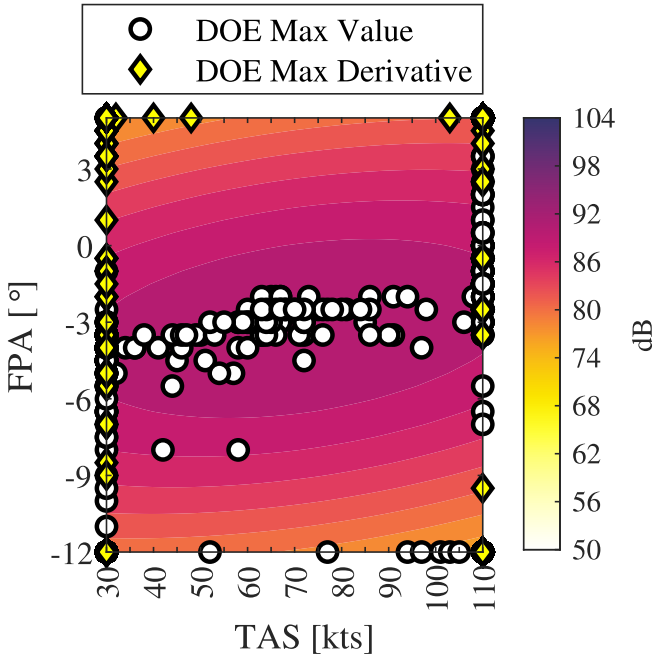


Figure 14: Contour shows BVISPL for the center microphone, while each circle represents a maximum value of a RSM for each microphone and each metric. Diamonds represent the maximum gradient of the RSM for each microphone and each metric, all assuming a gamma distribution of the response variable.

steep descents at slow and high speeds are potential locations of interest, but again these are at the edge of the interrogation window and are potentially a numerical fitting issue.

Combining all microphones and all metrics into a single image results in Figure 14. The contour level is from BVISPL metric for the center microphone (microphone 79), while the markers indicate locations of maximum value or gradient of the RSM, for each microphone, from every metric of interest.

Overall takeaways are that the -3 degree descent conditions are of primary interest for maximum metric values, while

slow speed and high speed conditions may be important for both maximum values and maximum gradients. Using non-dimensionalized predictor variables results in a largely similar distribution, shown in Figure 15. The non-dimensionalized predictor variables do show some preference for collecting −12 degree descent conditions for multiple flight speeds.

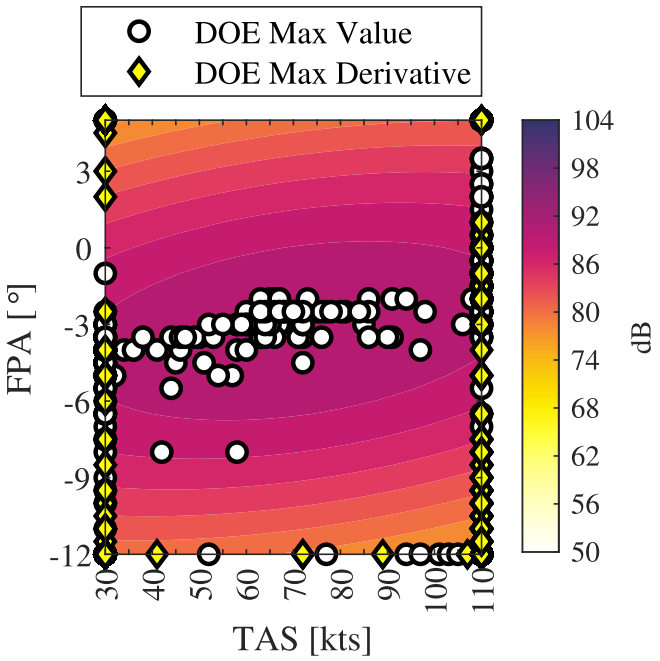


Figure 15: Contour shows BVISPL for the center microphone, while each circle represents a maximum value of a RSM for each microphone and each metric. Diamonds represent the maximum gradient of the RSM for each microphone and each metric, all assuming a gamma distribution of the response variable and non-dimensionalized predictor variables.

All of the previous data can be evaluated against the planned test points and used to suggest new test points worth investigating. A secondary data set also exists alongside this experiment, which provided slightly different results. Those results, not shown here, contributed to the identification of new test conditions as well. The new test conditions, which are unique and far enough away from pre-planned test conditions, are provided in Table 4. A second day of data was collected, in which these conditions were flown multiple times along with some of the originally planned flight conditions. The second day of testing measured the conditions found in Table 5.

ANALYSIS OF DESIGN OF EXPERIMENT APPROACHES

At this stage, a DOE method was developed for flight testing through the use of standard approaches. Analysis of the acquired data set was made using generalized linear modeling with non-dimensional and dimensionalized predictor variables, and six acoustic metrics as response variables. The RSMs were then evaluated for their maximum values and maximum gradients, and new test points were created and subsequently tested. Further evaluation is necessary to determine the value in the original DOE designs, as all acquired data were used so far to train the RSMs.

Table 4: Flight conditions developed during the flight test, using the RSM values and gradients previously described. The final column indicates that the point was developed based on the RSM models for maximum (M) metric and derivative (D) of the metric.

Condition Code	TAS [kts]	FPA [°]	Reason
C5	40	3	D
C6	60	3	D
C7	95	3	D
C8	40	5	D
C9	60	5	D
C10	75	5	M
C11	90	5	M
D18	40	-12	D
D19	60	-12	M
D20	75	-12	D/M

Optimization with DOE

The evaluation of the DOE designs will begin with the previously described development of new test conditions, this time repeated using only the test conditions from each DOE design. The intent here is to determine if the above designs were implemented, what would the impact be on providing new test conditions.

Figure 16 provides a similar approach as Figure 14, but this time created using only the FCCD design. Several distinct differences appear in this image, as compared to Figure 14. First, the identification of maxima values for RSM and maximal gradients are dispersed over a larger portion of the figure, and not grouped in the corners of the domain. Second, the maximal values would encourage more testing at the -3 degree descents (suggested before but test points already existed there), and now include test conditions at steeper descent angles for 70 knot flight speeds. This area was covered well by the overall test design, but could have influenced and improved the test plan if this area wasn't measured. It is interesting to note that the BVISPL contour is significantly different for this model then was previously shown in Figure 14. The errors in the RSM modeling using DOE will be analyzed below.

This analysis is continued using the CCCD design approach, as shown in Figure 17, and the Hexagonal design approach in Figure 18. The CCCD approach BVISPL contour for mi-

Table 5: Flight conditions measured during the second day of conventional MD530F testing.

Condition Code	TAS [kts]	FPA [°]	Runs
C1	75	3.0	2
C2	60	2.0	2
C3	95	2.0	2
C4	75	1.25	2
C5	40	3.0	3
C6	60	3.0	3
C7	95	3.0	3
C8	40	5.0	3
C9	60	5.0	3
C10	75	5.0	3
C11	90	5.0	3
L1	40	0.0	1
L2	60	0.0	2
L3	75	0.0	2
L4	90	0.0	5
L5	110	0.0	1
D2	40	-3.0	2
D3	60	-3.0	2
D4	75	-3.0	2
D5	90	-3.0	2
D6	110	-3.0	2
D9	60	-6.0	2
D10	75	-6.0	2
D11	90	-6.0	2
D12	75	-7.25	2
D13	60	-8.0	2
D14	95	-8.0	2
D17	75	-10.5	5
D18	40	-12.0	3
D19	60	-12.0	3
D20	75	-12.0	4

crophone 79 agrees better with the full data set RSM contour, and provides some unique test conditions as suggested by the maximal values. For the CCCD design, -6 degree slow speed conditions become important to measure for the maximal value analysis, while the gradient method suggests low speed conditions are of particular importance.

The hexagonal approach, shown in Figure 18, provides a contour map similar to that seen in Figure 16, although with less steep vertical gradients. Again, the -3 degree descent condition is called out as a particular need for additional testing, as well as some steeper descents around -6 degrees at around 40, 75, 85, and 100 knots.

Overall, the method for fitting RSMs to flight test data and investigating the maximum value and maximum gradient has been shown valuable in identifying new potential test conditions to investigate. It is clear that the maximum value method results in more obvious flight test conditions, with the gradient method tending to prefer locations on the edge of the interrogation window. Further analysis should be done to determine other viable methods for determining new test points 'on-the-fly', such as training a semi-empirical model and identifying areas where high errors are expected with the model (Ref. 19).

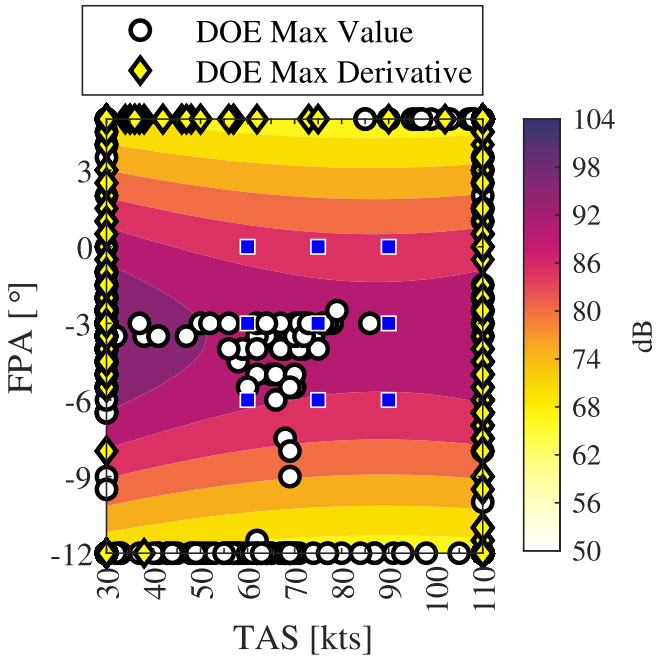


Figure 16: Contour shows BVISPL for the center microphone, while each circle represents a maximum value of a RSM for each microphone and each metric as developed by the FCCD DOE approach. FCCD conditions shown in squares. Diamonds represent the maximum gradient of the RSM for each microphone and each metric, all assuming a gamma distribution of the response variable.

RSM Errors with DOE

A principle element of the DOE process has yet to be investigated. That is, the ability for the RSMs to accurately predict locations of interest outside of the test envelope (extrapolation) as well as within the test envelope (interpolation). In order to conduct the analysis, RSMs are made using each experimental design. The models were then evaluated at the interpolation points and 'OFAT' points previously identified in Figure 7. The 'OFAT' conditions are extrapolation points for each of the three designs, and so provides an excellent chance to evaluate the RSM predictions.

As each of the interpolation and extrapolation points were measured, then the error can be determined as the difference between the measured response and predicted value for each metric at every microphone location. Figure 19 is a Lambert projection of the error predicted at each microphone location

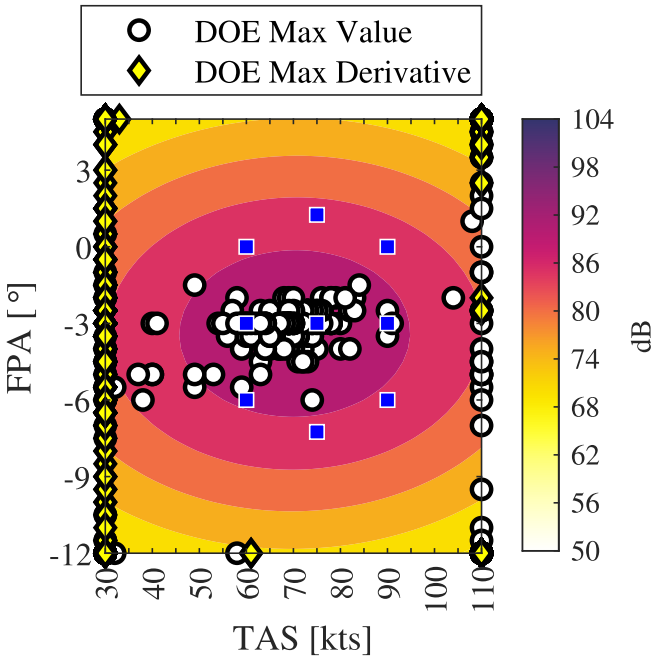


Figure 17: Contour shows BVISPL for the center microphone, while each circle represents a maximum value of a RSM for each microphone and each metric as developed by the CCCD DOE approach. CCCD conditions shown in squares. Diamonds represent the maximum gradient of the RSM for each microphone and each metric, all assuming a gamma distribution of the response variable.

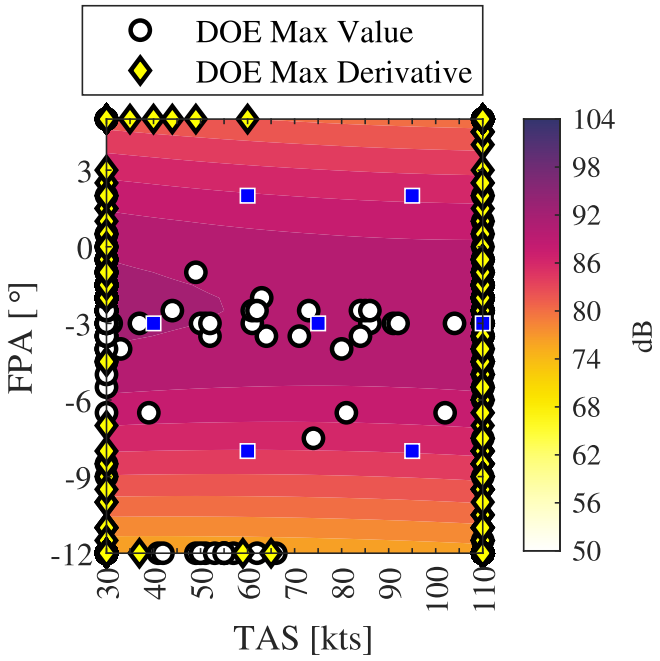


Figure 18: Contour shows BVISPL for the center microphone, while each circle represents a maximum value of a RSM for each microphone and each metric as developed by the Hexagonal DOE approach. Hexagonal conditions shown in squares. Diamonds represent the maximum gradient of the RSM for each microphone and each metric, all assuming a gamma distribution of the response variable.

for the OASPL metric using FCCD RSMs. From this figure, it is obvious that the errors in the RSM peak beneath the rotor and near the horizon plane and rear of the vehicle for the OASPL metric.

Similar figures can be produced, such as Figure 20, which shows the error distribution for OASPL from the CCCD design for an extrapolation condition 'D16'. Here, primary errors occur forward and slightly below the rotor suggesting that the RSMs do not adequately predict the BVI related noise as that is the primary radiation direction for BVI.

Figures such as these have been produced for all metrics, for each DOE approach and can continue to be qualitatively analyzed. However, a more robust statistical analysis has been conducted to quantify errors associated with DOE. Each of the errors for each microphone can be captured in a single box plot, as shown in Figure 21. These are standard box plots, where the red lines indicate median values, with the box ranging from 25th to 75th percentiles, and the whiskers extend to a maximum of 1.5 times the interquartile range. 'plus' symbols represent outliers, while the solid black line differentiates between interpolation and extrapolation points. Figure 21 is a bit of a misnomer as the 'interpolated' and 'extrapolated' points were used in the building of the RSM's. Instead, this figure acts as the anchor to compare the following images to. This is the 'ideal' condition, where the RSM's have performed to the

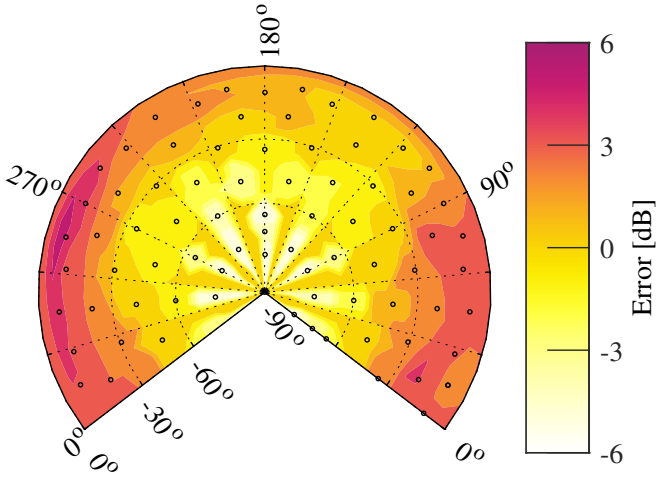


Figure 19: Lambert projection of the error predicted by the FCCD RSMs for the OASPL metric at interpolation condition 'D1'.

best of their ability as they are trained on the data set in question. It is useful to note that the median errors for all points in Figure 21 are within 1.5 dB, which is well within the expectation for repeated flight conditions. Overall this shows that the RSMs, when built on the data they are trying to predict, do an excellent job of recreating those data. This analysis

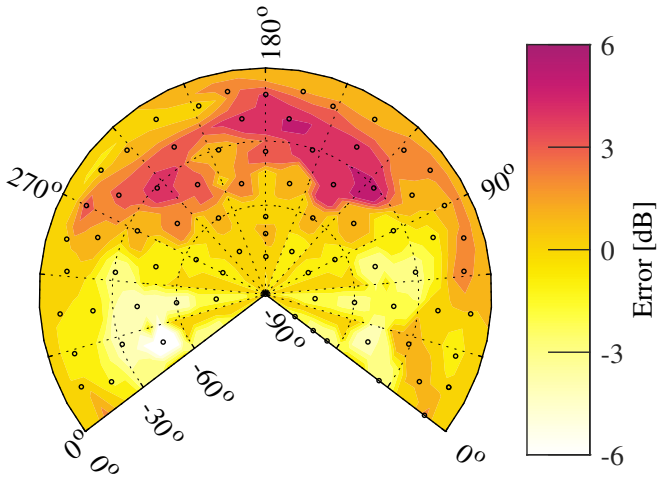


Figure 20: Lambert projection of the error predicted by the CCCD RSMs for the OASPL metric at extrapolation condition 'D16'.

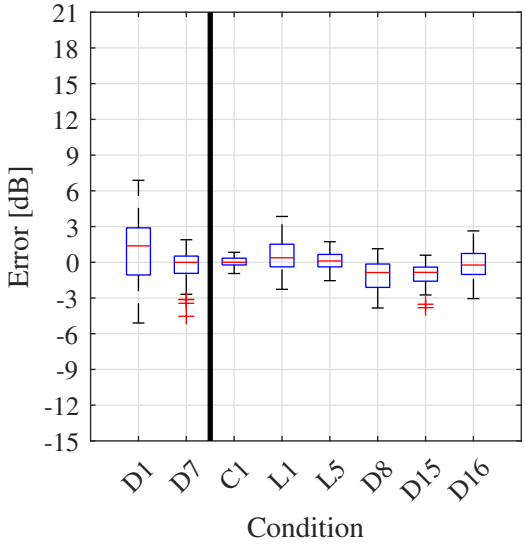


Figure 21: Box plot of RSM OASPL prediction errors for the complete design for interpolated conditions ('D1' and 'D7'), as well as extrapolation conditions ('C1', 'L1', 'L5', 'D8', 'D15', and 'D16').

does have the drawback of losing directionality information that was available in the Lambert error projections.

Figure 22 shows the OASPL prediction error distribution for FCCD design for both interpolated and extrapolated conditions. It is clear that the RSM for OASPL produces significantly lower median errors and a smaller distribution of errors for points located interior to the test conditions (interpolation). However, while the extrapolation points have higher distribution of errors, their median error is still quite small within 3 dB of the actual value. The same analysis can be done for other metrics, such as BVISPL (Figure 23) and SEL (Figure 24). The BVISPL analysis shows increased spread across the con-

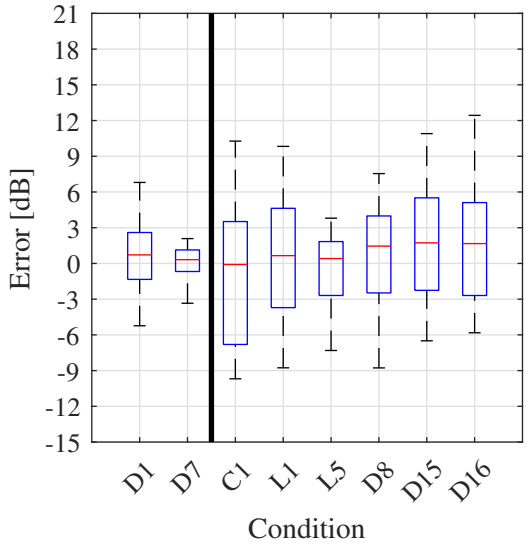


Figure 22: Box plot of RSM OASPL prediction errors for the FCCD design for interpolated conditions ('D1' and 'D7'), as well as extrapolation conditions ('C1', 'L1', 'L5', 'D8', 'D15', and 'D16').

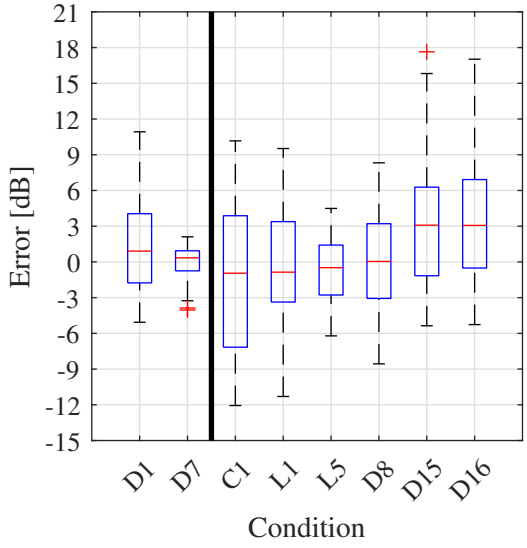


Figure 23: Box plot of RSM BVISPL prediction errors for the FCCD design for interpolated conditions ('D1' and 'D7'), as well as extrapolation conditions ('C1', 'L1', 'L5', 'D8', 'D15', and 'D16').

ditions, but the median errors stay within (or very close) to 3 dB from truth data. The SEL condition, however, shows significantly smaller errors, with the exception of the climb prediction condition ('C1'). This is to be expected as sound exposure level is an integrated quantity that is less sensitive to dynamic variations in vehicle state, compared to BVISPL.

The CCCD condition can be similarly investigated, and is shown in Figures 25 and 26. Examination of Fig. 25 compared to Fig. 22 shows similar levels of OASPL prediction for interpolation and extrapolation points. The CCCD design, in

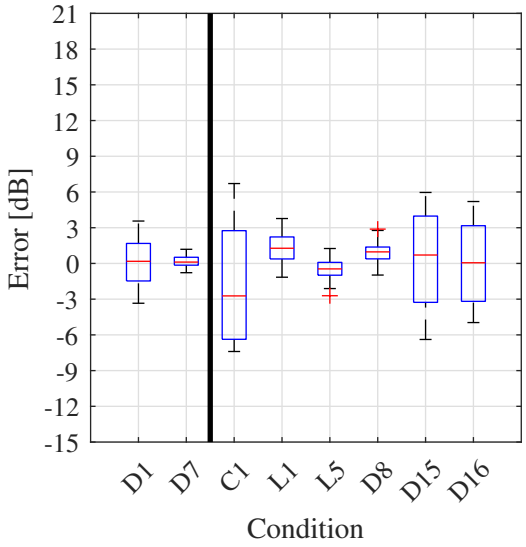


Figure 24: Box plot of RSM SEL prediction errors for the FCCD design for interpolated conditions ('D1' and 'D7'), as well as extrapolation conditions ('C1', 'L1', 'L5', 'D8', 'D15', and 'D16').

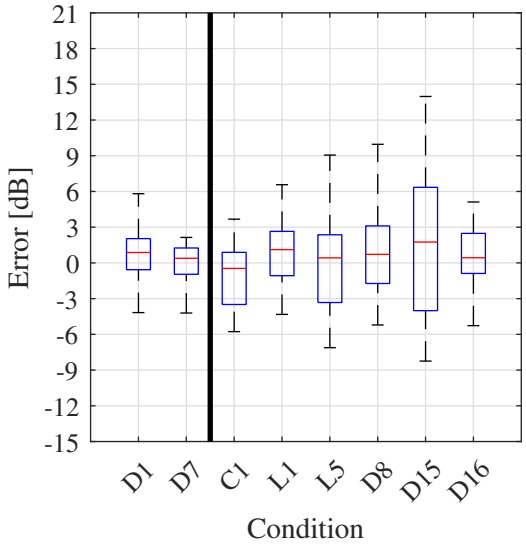


Figure 25: Box plots of OASPL prediction errors for the CCCD design for interpolated conditions ('D1' and 'D7'), as well as extrapolation conditions ('C1', 'L1', 'L5', 'D8', 'D15', and 'D16').

general, represents a lower spread of error levels with the exception of 'D15' extrapolation condition. However, the FCCD design in general has lower median errors across all microphones for OASPL predictions.

Sound exposure level for the CCCD design was also provided in Figure 26. Comparing to Figure 24, it can be seen that the CCCD has a higher spread in extrapolation errors, with the exception of condition 'C1'. Predicting 'C1' better is expected for the CCCD design as climb conditions are represented in the design, while FCCD conditions do not include climb conditions. The interpolation conditions yielded similar error distributions.

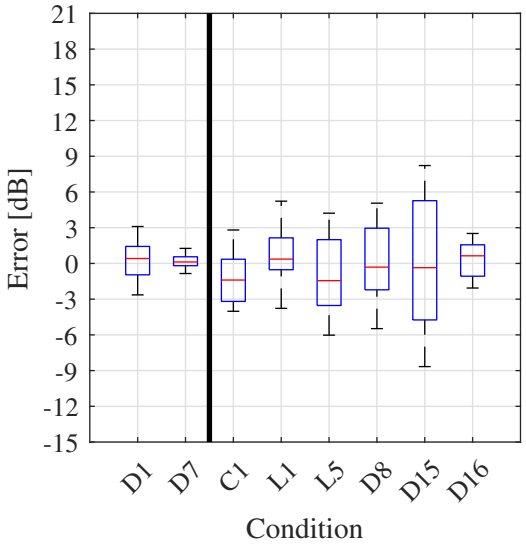


Figure 26: Box plots of SEL prediction errors for the CCCD design for interpolated conditions ('D1' and 'D7'), as well as extrapolation conditions ('C1', 'L1', 'L5', 'D8', 'D15', and 'D16').

Finally, the Hexagonal DOE design can be analyzed in Figures 27-29. The OASPL RSMs for the Hexagonal design

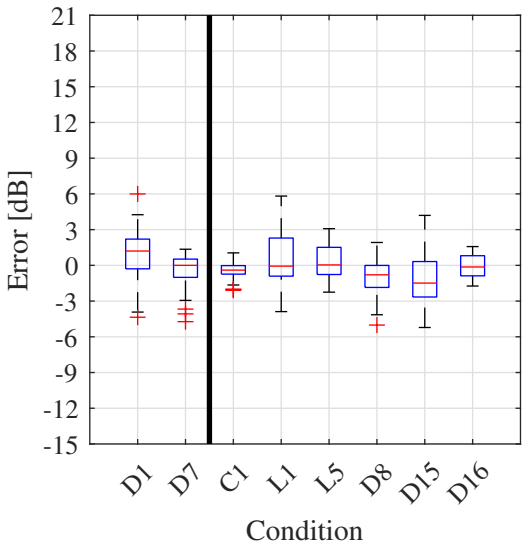


Figure 27: Box plots of OASPL prediction errors for the Hexagonal design for interpolated conditions ('D1' and 'D7'), as well as extrapolation conditions ('C1', 'L1', 'L5', 'D8', 'D15', and 'D16').

(Fig. 27) shows a significantly lower spread and median errors than either of the FCCD (Fig. 22) or CCCD (Fig. 25) designs for the same metric in both interpolation and extrapolation conditions. The errors for the Hexagonal design are in

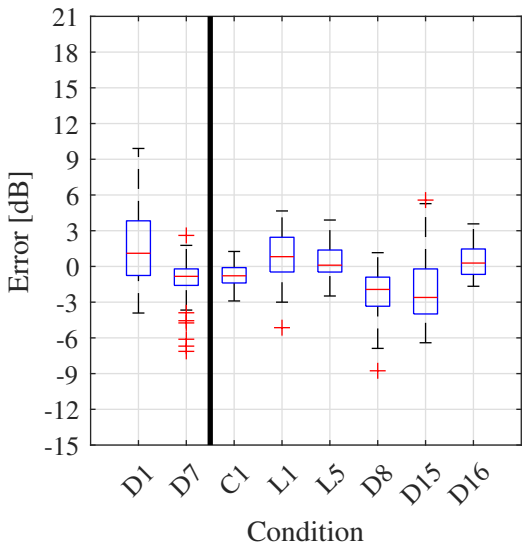


Figure 28: Box plots of BVISPL prediction errors for the Hexagonal design for interpolated conditions ('D1' and 'D7'), as well as extrapolation conditions ('C1', 'L1', 'L5', 'D8', 'D15', and 'D16').

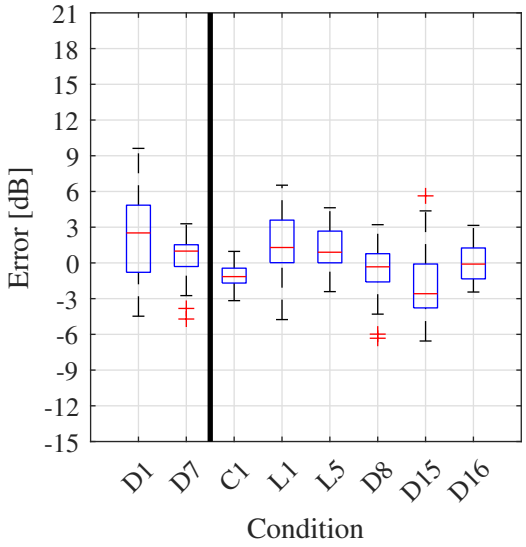


Figure 29: Box plots of PNLT prediction errors for the Hexagonal design for interpolated conditions ('D1' and 'D7'), as well as extrapolation conditions ('C1', 'L1', 'L5', 'D8', 'D15', and 'D16').

line with some of the best errors expected from Fig. 21, and this is likely because the Hexagonal design spans a significantly larger portion of the flight envelope than the FCCD or CCCD designs.

The Hexagonal design did not adequately perform with the interpolation condition 'D1' for the BVISPL metric (seen in Fig. 28), but overall performed excellently across all conditions. Also provided was the PNLT prediction for this condition, which shows small median errors and small spread of errors across all conditions and microphones.

CONCLUSIONS

A comprehensive acoustics research flight test was described, in which a design of experiments approach was undertaken to design the initial test conditions. Three unique approaches were considered for the two flight test parameters available (flight path angle and flight speed). A face centered central composite design, a circumscribed central composite design, and a hexagonal design were all employed. A fourth design was used, based on traditional one-factor-at-a-time, to complete the entire flight envelope.

A response surface methodology was used to fit the measured data. Each of the 79 available microphones was modeled using a quadratic generalized linear model assuming a gamma distribution of the response variables with respect to the predictor variables (flight path angle and flight speed). These models were built for each of the six acoustic metrics of interest (OASPL, OASPLA, BVISPL, PNL, PNLT, and EPNL).

The RSMs were first built using all available data at the end of the first day of testing. The RSMs were investigated for their maximum predicted value as well as their maximum gradient value. This was to highlight areas of interest where significant acoustic phenomena might be occurring and worth further investigation. From this analysis, 10 new test conditions were designed and measured on a subsequent day. It was highlighted that refinement of the gradient method is necessary through the use of a higher order model, as the assumed quadratic model used here always results in gradient maxima near the boundaries of the interrogation window.

Post-test analysis was conducted to investigate the effectiveness of the individual DOE designs and to determine if the above identification of new test conditions would have worked for the smaller DOE designs. It was demonstrated that the designs could have yielded new test conditions, and that the Hexagonal approach (which covered more of the flight envelope) produced fewer new test conditions interior to the design envelope.

The RSMs of the DOE designs were then investigated for their error distributions compared to interpolation conditions as well as extrapolation conditions. Overall the Hexagonal design condition appeared to provide the lowest error distribution, and this is again likely because it spanned a larger portion of the test envelope and so had a larger span of data to train upon. Of the two central composite designs, the circumscribed design provided slightly lower errors and was expected because this design has improved fitting characteristics with its flight conditions not perfectly symmetrical allowing for greater analysis of covariates.

This experiment pioneered the use of design of experiments for rotorcraft acoustic flight testing. It sets up the basis for developing such approaches for future vehicles, such as the urban air mobility community. However, this is neither the

best nor only approach to design of experiments for field testing. One particularly powerful method, not employed here, is to dynamically update a prediction model during testing. If a semi-empirical model is built prior to the test, then the errors of such model can be minimized through strategically designing test conditions around locations of maximal error (Ref. 19). This method is particularly powerful, but requires substantially more understanding of the vehicle prior to the flight test, and substantially more computational power than was available in the current methods. The approach taken here is a reasonable first step when the vehicle design or response is unknown prior to testing.

ACKNOWLEDGMENTS

The authors would like to sincerely thank Nathan Cruze (NASA LaRC) for his advice on generalized linear modeling, and Peter Parker (NASA LaRC) for his advice on design of experiments that helped guide the planning of the flight test. Their expert advice was crucial to the success of this work.

REFERENCES

1. Stephenson, J. H., Pascioni, K. A., Houston, M. L., Stutz, C. M., and Martin, P. B., "Overview of a Comprehensive MD530F Acoustic Flight Test," Proceedings of the 81st Annual Forum of the Vertical Flight Society, Virginia Beach, VA, May 2025.

2. Overmeyer, A. D., Tanner, P. E., Martin, P. B., and Commo, S. A., "Case Studies for the Statistical Design of Experiments Applied to Powered Rotor Wind Tunnel Tests," 2015 AIAA Aviation Forum, Dallas, TX, June 2015.

3. Lundstedt, T., Seifert, E., Abramo, L., Thelin, B., Nyström, Å., Pettersen, J., and Bergman, R., "Experimental design and optimization," *Chemometrics and Intelligent Laboratory Systems*, Vol. 42, (1-2), 1998, pp. 3–40.

4. Keane, A., "Wing optimization using design of experiment, response surface, and data fusion methods," *Journal of Aircraft*, Vol. 40, (4), 2003, pp. 741–750.

5. Mandenius, C.-F. and Brundin, A., "Bioprocess optimization using design-of-experiments methodology," *Biotechnology progress*, Vol. 24, (6), 2008, pp. 1191–1203.

6. Thurman, C. S. and Baeder, J. D., "Blade-Wake Interaction Noise for Small Hovering Rotors, Part I: Characterization Study," *AIAA Journal*, Vol. 61, (6), 2023, pp. 2552–2569.

7. Page, J. A., Wilmer, C., Schultz, T., Plotkin, K. J., and Czech, J., "Advanced Acoustic Model Technical Reference and User Manual," Technical Report Project WP-1304, SERDP, May 2009.

8. Greenwood, E., *Fundamental Rotorcraft Acoustic Modeling from Experiments (FRAME)*, Ph.D. thesis, University of Maryland, 2011.

9. Watts, M. E., Greenwood, E., Sim, B., Stephenson, J. H., and Smith, C. D., "Helicopter Acoustic Flight Test with Altitude Variation and Maneuvers," Technical Memorandum NASA/TM-2016-219354, NASA Langley Research Center, Hampton, VA 23681, USA, December 2016.

10. Houston, M. L., Stephenson, J. H., Pascioni, K. A., and Stutz, C. M., "Snapshot Array Design Considerations for Rotorcraft Noise Characterization," Proceedings of the 81st Annual Forum of the Vertical Flight Society, Virginia Beach, VA, May 2025.

11. Pascioni, K. A., Stutz, C. M., Houston, M. L., and Stephenson, J. H., "Phased Array Measurements on a Full-Scale Helicopter," Proceedings of the 81st Annual Forum of the Vertical Flight Society, Virginia Beach, VA, May 2025.

12. Stutz, C. M., Stephenson, J. H., Pascioni, K. A., and Houston, M. L., "Acoustic Assessment of an MD530F Helicopter in Maneuvering Flight," Proceedings of the 81st Annual Forum of the Vertical Flight Society, Virginia Beach, VA, May 2025.

13. "N20AT FAA Registry Information," Technical report, Federal Aviation Authority, Available at: https://registry.faa.gov/AircraftInquiry/Search/NNumberResult?nNumberTxt=20AT, Accessed 30 Jan 2025.

14. Pascioni, K. A., Stephenson, J. H., Houston, M. L., and Stutz, C. M., "Overview of the 2024 Army/NASA Acoustic Research Flight Test," NASA Acoustics Technical Working Group Meeting, Fall 2024, Cleveland, Ohio, October 2024.

15. "Ground-Plane Microphone Configuration for Propeller-Driven Light-Aircraft Noise Measurement," ARP 4055, SAE, November 2007.

16. Lopes, L. and Burley, C., "Design of the next generation aircraft noise prediction program: ANOPP2," 17th AIAA/CEAS Aeroacoustics Conference, Portland, OR, June 2011.

17. Lebedev, V. and Laikov, D., "A Quadrature Formula for the Sphere of the 131st Algebraic Order of Accuracy," *Doklady Mathematics*, Vol. 59, (3), 1999, pp. 477–481.

18. Stephenson, J. H., "Generalized Linear Modeling of Rotorcraft Acoustic Flight Test Data," Proceedings of the 80^{th} Annual Forum of the Vertical Flight Society, Montréal, Québec, Canada, May 2024.

19. Greenwood, E., personal communication, June 2022.