Project 042 Acoustical Model of Mach Cut-off Flight

The Pennsylvania State University

Project Lead Investigator
Victor W. Sparrow
Director and Professor of Acoustics
Graduate Program in Acoustics
The Pennsylvania State University
201 Applied Science Bldg.
University Park, PA 16802
+1 (814) 865-6364
vws1@psu.edu

University Participants
Pennsylvania State University
- P.I.(s): Dr. Victor W. Sparrow (PI), Dr. Michelle C. Vigeant (Co-PI)
- FAA Award Number: 13-C-AJFE-PSU, Amendments 20, 33, and 42
- Period of Performance: June 28, 2016 - December 31, 2019
- Task(s):
    1. Assess and extend modeling capability for Mach Cut-off events (a.k.a. Task 1A)
    2. Study human perception of Mach Cut-off sounds

Project Funding Level
$150K for 2017-2018 and $170K for 2018-2019, The Pennsylvania State University

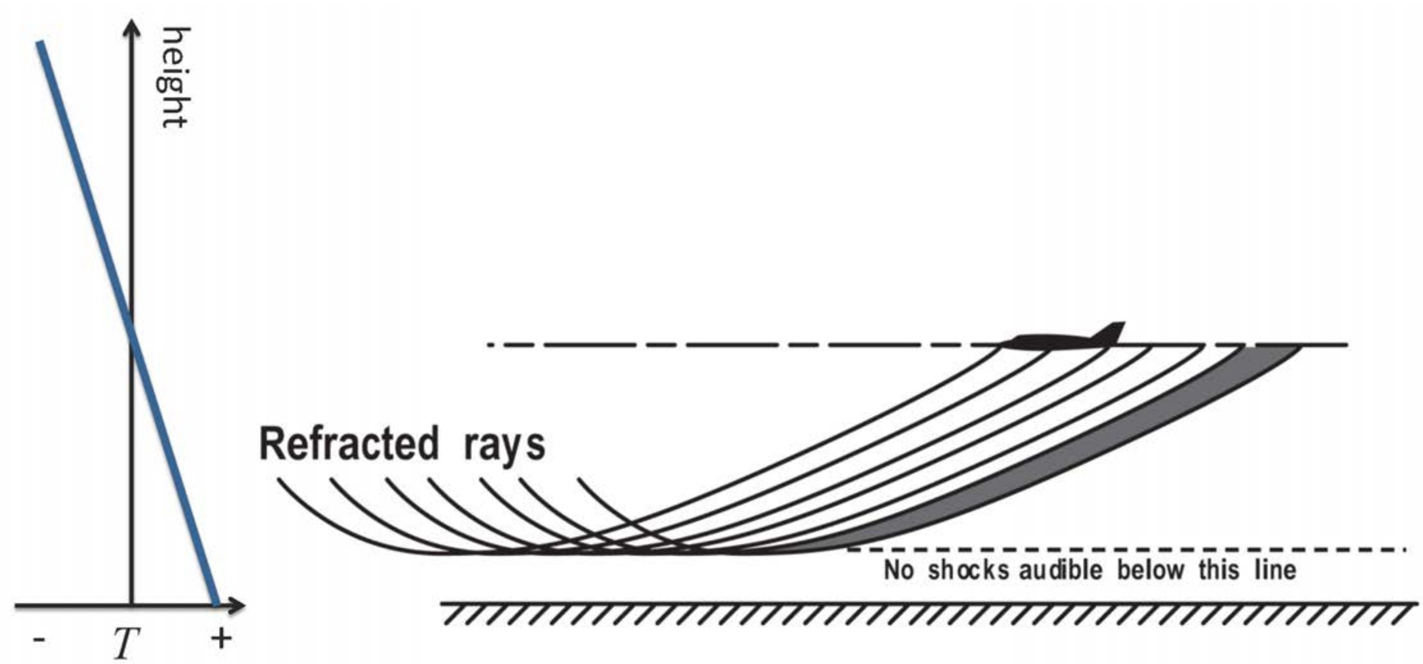
Aerion Corporation is providing cost-share in-kind matching funds to Penn State. Our point of contact at Aerion is Jason Matishek, jrmatisheck@aerioncorp.com. Aerion is providing the necessary near-field CFD data and other relevant information to help guide the project team make accurate predictions of the Mach cut-off sonic boom signatures that may be produced by Aerion's future supersonic aircraft.

Investigation Team
Pennsylvania State University
Principal Investigator: Victor W. Sparrow
Co-Investigator: Michelle C. Vigeant
Graduate Research Assistant Zhendong Huang (assessment and extension of Mach cut-off models)
Graduate Research Assistant Nick Ortega (human perception of Mach cut-off sounds)

Project Overview
ASCENT Project 42 brings together resources to provide preliminary information to the FAA regarding the noise exposure of supersonic aircraft flying under Mach cut-off conditions. Studies in the 1970s showed that Mach cut-off supersonic flight was possible, but there is currently no data establishing the frequency and extent of noise exposures and no guidelines for managing such exposures. Penn State is shedding light on the Mach cut-off phenomena.

Aerion Corporation and many others believe that Mach cut-off supersonic flight is both viable [Plotkin, et al., 2008] and very likely to be acceptable to the public. But there is a lack of data to back up this assertion. Thus, research needs to be conducted to provide a technical basis for rulemaking regarding Mach cut-off operations.

The basic concept of Mach cut-off relies on the fact that the ambient temperature is substantially colder at flight altitudes than on the ground. Hence, the speed of sound is substantially slower at flight altitudes than at the ground. As illustrated in the simplified graphic below, it is possible to fly in a range of Mach numbers (perhaps between Mach 1.0 and Mach 1.15) while having the sonic boom noise refract (bend) upwards such that the rays never reach the ground. However, the reader should be aware that this picture is over-simplified since the temperature profile in the atmosphere is never a smooth, linear function as depicted here. For higher Mach numbers, the sonic boom will impact the ground before refracting upward.



Simplified view of Mach cut-off where sonic boom noise does not reach the ground surface. Left: ambient temperature versus height. [Sparrow] Right: aircraft and ray diagram showing refraction of sonic boom [NASA].

Little is known about the noise impact of Mach cut-off operations for future supersonic aircraft. The concept of Mach cut-off was introduced by Lockheed engineers in the mid-1960s [Shurcliff, 1970]. NASA conducted some field experiments in the early 1970s, focusing on other speed regimes of flight, validating some of the Mach cut-off theory for some of the sound field. This research was conducted in Nevada with a 466 m (1,529 ft) tower [Haglund and Kane, 1973]. Then to more directly address the Mach cut-off issue, a theoretical and experimental study was conducted in the mid-1970s with FAA support. The studies estimated altitudes and Mach number regimes to ensure the focus boom does not reach the ground. That field campaign used fighter jets flying out of Langley AFB to a test area in the Atlantic Ocean off Wallops Island, Virginia [Perley, 1977]. Using the available instrumentation, the study concluded that Mach cut-off flight was feasible.

In none of those studies were any recordings made of sufficient quality to assess human response to the Mach cut-off noise. The theoretical studies estimating the altitude and Mach number restrictions for focus boom avoidance assumed a simple atmospheric model (linear sound speed profile), and did not include real-world atmospheric effects. Hence the 1960s-1970s work was very good, but is only a start to determining appropriate flight conditions for routine Mach cut-off supersonic flights over the continental United States.

Task 1A- Propagation Modeling with Enhanced Ray-tracing Capabilities

The Pennsylvania State University

Objective(s)

Research will be conducted to understand Mach cut-off operations and how often people would hear the unique Mach cut-off sounds. This includes estimating the flight altitude and Mach number restrictions for focus boom avoidance including real-world atmospheric effects and providing assessments of practical Mach cut-off flight.

Research Approach

Methodology

Mach cut-off depends on the upward refraction of the sound in the atmosphere due to the temperature and wind speed gradients. To examine this phenomenon, the original propagation theory [Nichols, 1971] was retraced for extensibility. Nichols' theory was used back in 1970s to assess Mach cut-off. In that theory, the atmosphere is assumed to have only vertical variations of temperature and horizontal wind. Based on one form of the refraction law specifying the normal direction of a wavefront, it was argued that for upward sound refraction, a direct sonic boom noise would not reach the ground as long as the acoustic wavefront normal becomes parallel to the ground aloft. The "safe" flight altitude and distance between the sonic boom caustic and the ground can then be determined.

One limitation of earlier Mach cut-off theories (including Nicholls') and existing tools (e.g. PCBoom) is that vertical winds are not included. In a realistic atmosphere, however, a noticeable vertical wind can sometime exist. A sea breeze is one example in which a wind is blowing from sea to land, which normally occurs along coasts during daytime which involves with ascending and descending motion of the air. As the vertical wind becomes non-negligible, both the launch angle of sound rays away from the aircraft and the altitude of the caustic will be affected, as shown in Figure 1. Thus, it's important to clarify the difference between the refraction law for a sound ray and that for the normal to a wavefront in a moving atmosphere [Ostashev, 2001].

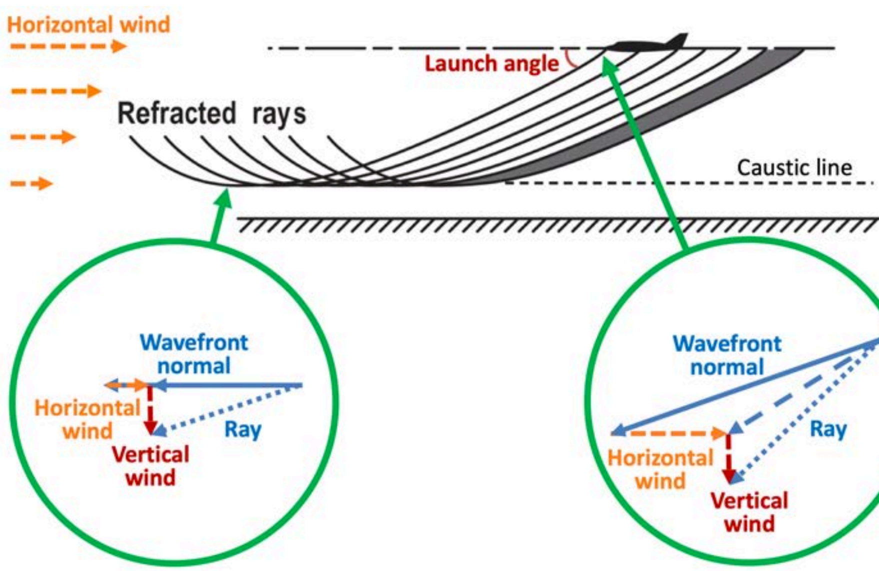


Figure 1. Contributions of a vertical wind to the difference between the wavefront normal and ray directions.

Another limitation in Nicholls' theory is, by calculating the cut-off Mach number based on the atmospheric conditions only at the flight and ground levels, the impact of realistic vertical atmospheric profiles in between those two levels on the sound propagation hasn't been included. For sonic boom propagation over a long horizontal distance, horizontal variations of the atmosphere can also be important for Mach cut-off.

In order to have an acoustical model that can lead to more accurate estimates of the safe cut-off altitude, flight Mach number and ground speed for Mach cut-off flight, 3-D ray tracing equations have been examined [Pierce, 1989]. Based on which, a 3-D 4th order Runge-Kutta integration ray tracing scheme has been developed, which can read-in realistic atmospheric data that includes arbitrary speed of sound variations as well as arbitrary three-dimensional winds including longitudinal, lateral and vertical wind components. It can be used to simulate sonic boom propagation through the 3-D atmosphere.

The data from Climate Forecast System version 2 (CFSv2) was used for the atmosphere [Saha, 2014]. CFSv2 provides an averaged and smoothed atmosphere globally at a horizontal resolution of 0.5 degree in both longitude and latitude, with 37 pressure levels up to an altitude of 40 km. The data is provided every 6 hours since January 1st, 2011 and updated on a daily basis that includes temperature, pressure, eastward and northward winds, vertical wind, and ground elevation that were used among other variables.

As shown in Figure 2, the horizontal resolution of the CFSv2 grid points corresponds to a distance of approximately 55 km over the ground. Since typical horizontal distances the sonic boom can travel before it reaches a cut-off point would be mostly in the range of 80 – 100 km, the CFSv2 grid point resolution is not good enough to provide information about the horizontal variation of the atmosphere in a local ray-tracing scale. Thus, the atmosphere is assumed to be vertically stratified in the computational domain of the simulation for each single case. On the other hand, in a global scale, the horizontal change of the atmosphere is included in our model. For each simulation, after the instantaneous location of the cruising aircraft along the flight path had been specified, this can be achieved by choosing the atmospheric data at the nearest CFSv2 grid point to represent the local atmospheric condition at the aircraft location.

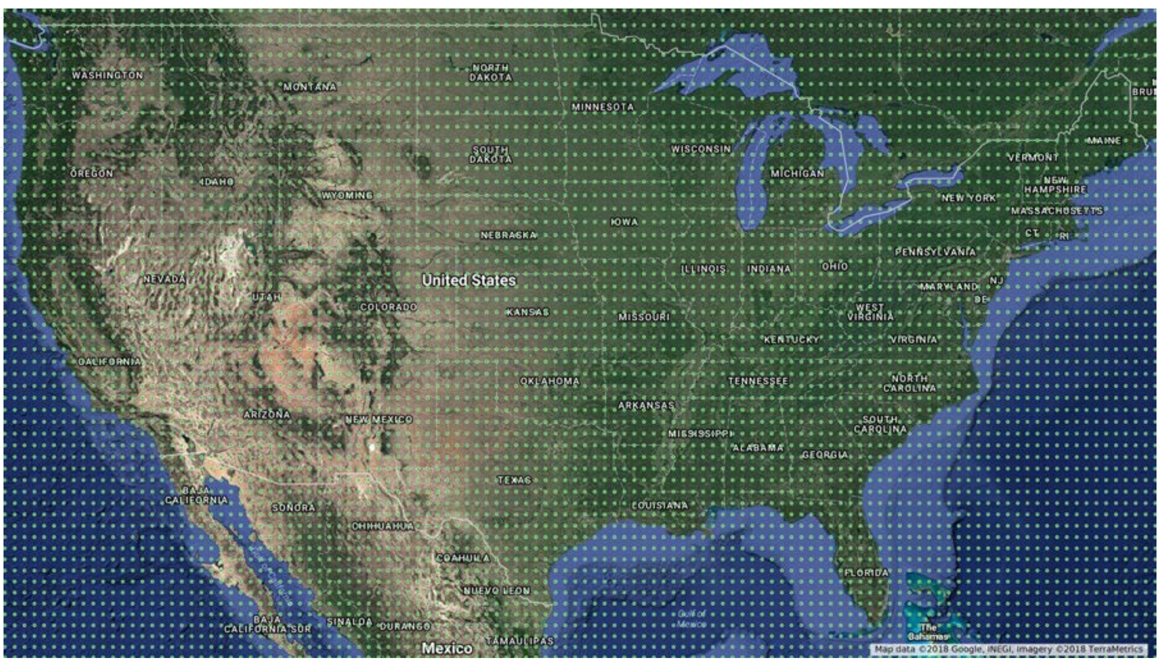


Figure 2. Contiguous United States map with grid points of the CFSv2 data.

With the input atmospheric data, the next step is to incorporate the Mach cut-off operational parameters proposed by Aerion Corporation. For this study, both a lower flight altitude of 12.5 km (41,010 ft) and a higher one of 15.24 km (50,000 ft) were adopted. The flight direction was inferred based on the location of the aircraft along the flight path.

When an aircraft is flying under the Mach cut-off condition, a focused boom is formed at the cut-off altitude, which corresponds to loud sound energy. Below that cut-off altitude, there is an evanescent wave decaying rapidly towards the ground, which typically covers a distance of a few hundred feet. To ensure the caustic never intercepts the ground, an adequate safety margin needs to be allowed [Plotkin, 2008]. So far, the cut-off Mach number was determined as the largest Mach number possible with at least a 500 m clearance of a caustic on the ground. The corresponding true air speed, aircraft ground speed, and vertical distance between the caustic and ground were then calculated. This method was used to assess the viability of Mach cut-off operations for one of the busiest air routes in the U.S. at a particular time of a year.

Results

As an example, a preliminary 3-D ray-tracing result using CFSv2 data is given for an aircraft flying over Los Angeles under cruise condition at an altitude of 12.5 km (41,010 ft) at 12 PM UTC on Jan 1, 2017, with a flight direction of 21.6 degree north of east, as shown in Figure 3. The cut-off Mach number for this case is 1.09, and the corresponding aircraft ground speed is 747.83 mph.

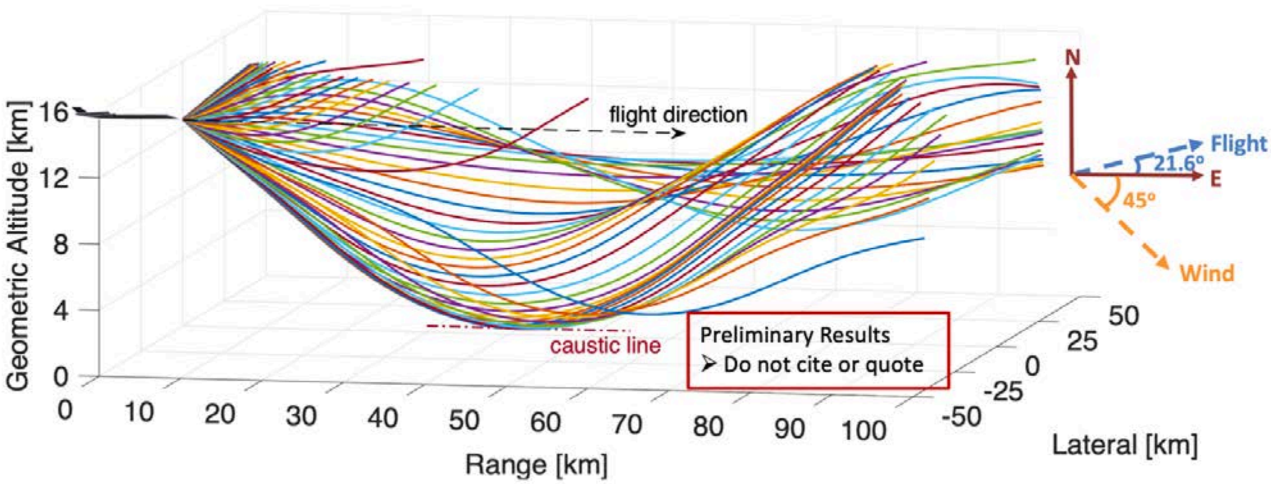


Figure 3. 3-D ray-tracing diagram, isometric view.

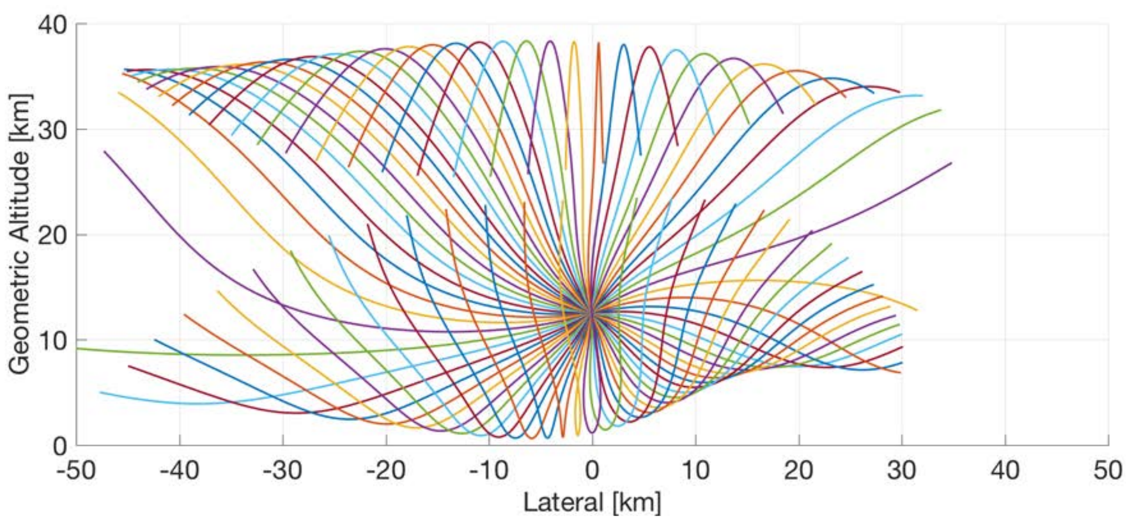


Figure 4. A side view of the 3-D diagram.

Figure 4 is a sideview of the 3-D diagram shown in Figure 3. In Figure 4, the flight direction is pointing perpendicularly out of the paper. It shows that the crosswind bends the sound rays toward the left side, and it's not necessarily the bottom ray (which has the deepest launch angle) that's the most likely one to reach the ground. Thus, it can be important to trace not only the bottom ray, but also the neighboring side rays.

Another example given below is to examine the impact of the vertical wind on the Mach cut-off altitude. In this example, the aircraft is flying at an altitude of 12.5 km northward. The CFSv2 atmospheric data for 12 PM UTC on September 10, 2017 in Miami, Florida is used. The atmospheric profiles are given in Figure 5.

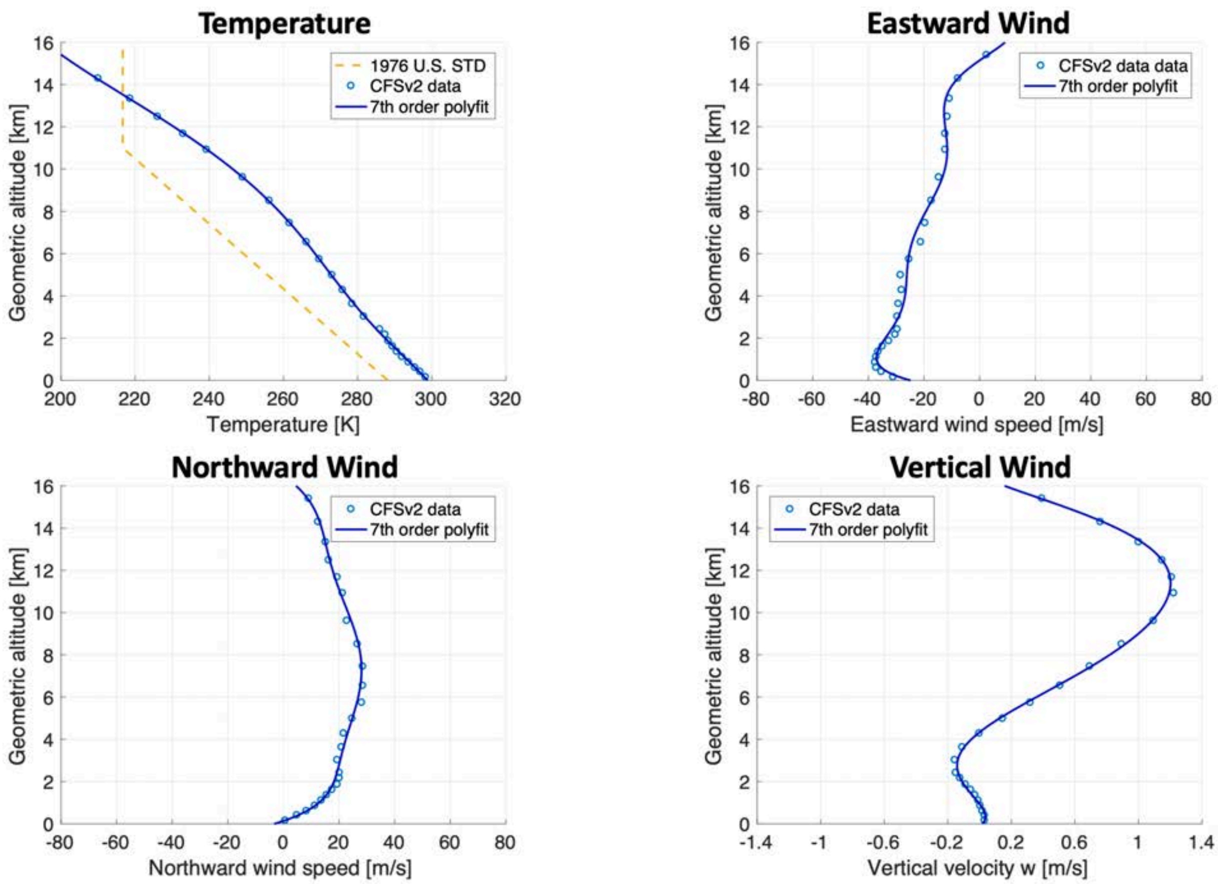


Figure 5. Atmospheric condition at 12 PM UTC on Sept. 10, 2017 in Miami, FL.

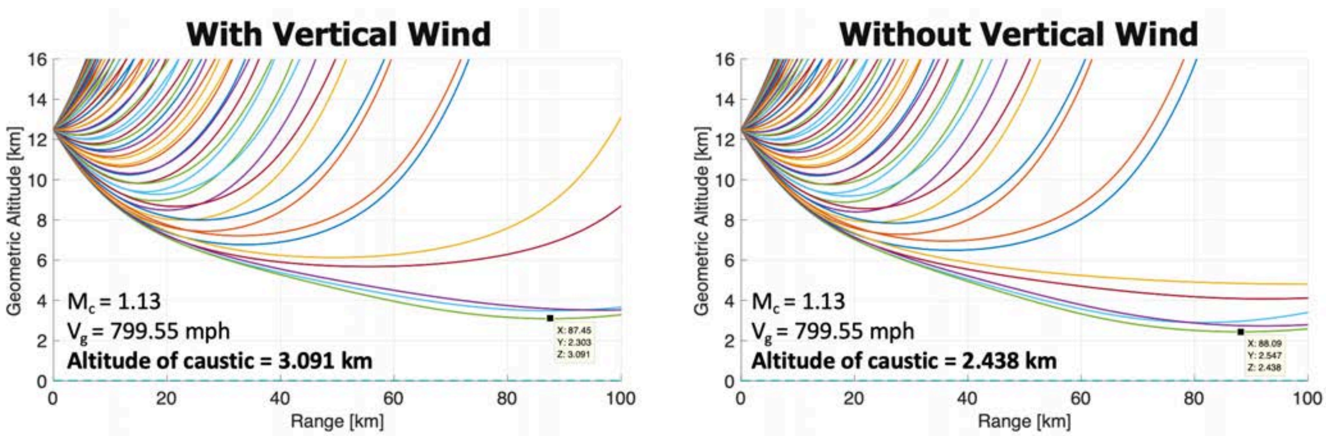


Figure 6. The impact of vertical wind on the Mach cut-off altitude.

In Figure 6, both of the two pictures are side views of the ray diagram. The only difference is, the vertical wind was included for the calculation shown in the left subfigure, while it was ignored for that shown on the right. A simple comparison shows that including a vertical wind while keeping other atmospheric variables the same can contribute to a 653 m difference in the Mach cut-off altitude. This means, sometimes the vertical wind is an important factor. It should also be stated that this particular atmospheric event is an unusual one, corresponding to when hurricane Irma hit Florida.

In a last example, the safe altitudes and speeds for Mach cut-off flight for a hypothetical round trip between Los Angeles and New York (as shown is Figure 7) were investigated, in which two different flight altitudes were used. The cut-off Mach number was determined as the largest Mach number possible with at least a 500 m clearance of a caustic above the ground. The ray-tracing calculation results for the westbound and eastbound cases are given in Figures 8 and 9, respectively.

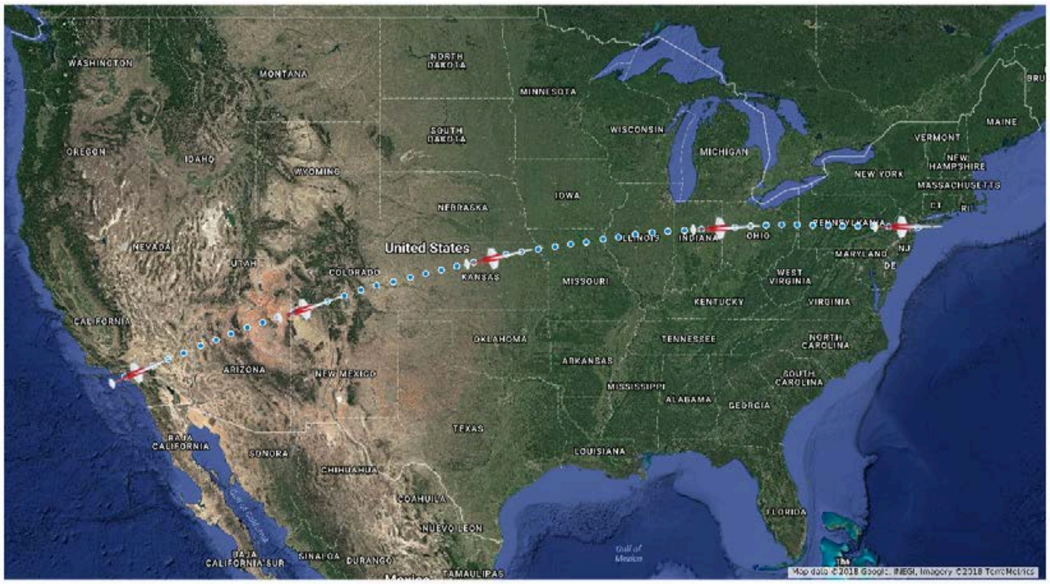


Figure 7. Air Route from LAX to JFK, Los Angeles to New York.

The top picture in Figure 8 provides the results for the westbound case. It shows when the aircraft was leaving the east coast, flying at a higher altitude didn't necessarily contribute to a higher cut-off Mach number. This is primarily because, the dominant eastward wind speed is decreasing with increasing altitude at altitudes above 10 km so that in the same region the effective sound speed (thermodynamic speed of sound plus wind speed) increases with increasing altitude. Thus, a higher cut-off Mach number can be achieved by flying at an altitude where the local effective sound speed is a minimum. However, a higher Mach number does not necessarily correspond to a higher aircraft ground speed unless we are talking about the same flight altitude, because the Mach number only represents a relative speed. In this case, we are comparing two different flight altitudes, since the temperature and wind speed at those two different altitudes are also different. Although sometimes a higher Mach number is achieved at a higher altitude, adjusting the flight altitude between 12.5 km and 15.24 km didn't gain significant benefit in the maximum aircraft ground speed.

Comparing Figures 8 and 9, it can be seen that, it's easier to implement Mach cut-off for a westbound flight than for an eastbound flight whenever the wind is mainly blowing eastward, which seems to be relatively common. In other words, it's easier to implement Mach cut-off when flying into a head wind than with a tail wind. When it's inevitable to fly with a tail wind, one should try to fly at an altitude where the local longitudinal wind speed is smaller. This is so that the local effective sound speed can be smaller, and this may help to achieve a higher Mach number that's well above Mach 1.0 if the temperature is not varying significantly in that region, even though by doing so one may not be able to increase the aircraft ground speed.

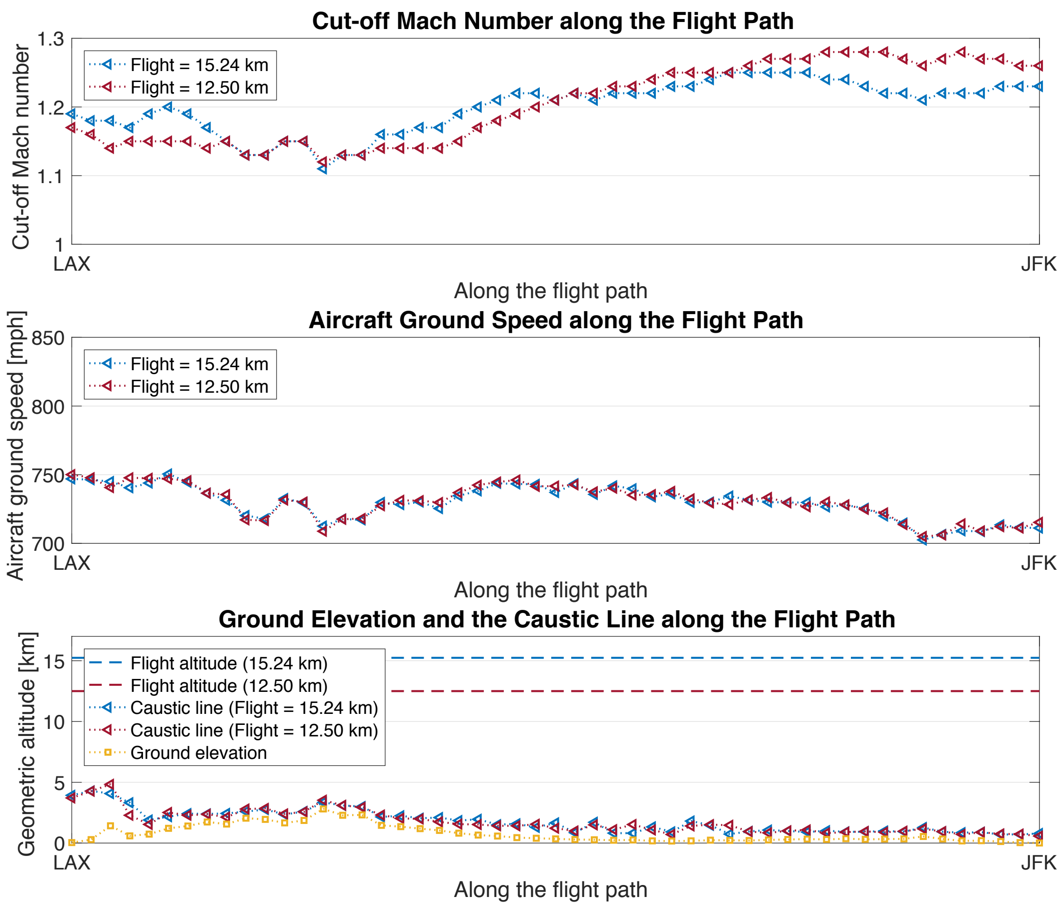


Figure 8. JFK to LAX air route (westbound into a headwind) on Jan 1 2017 at 12 PM UTC.

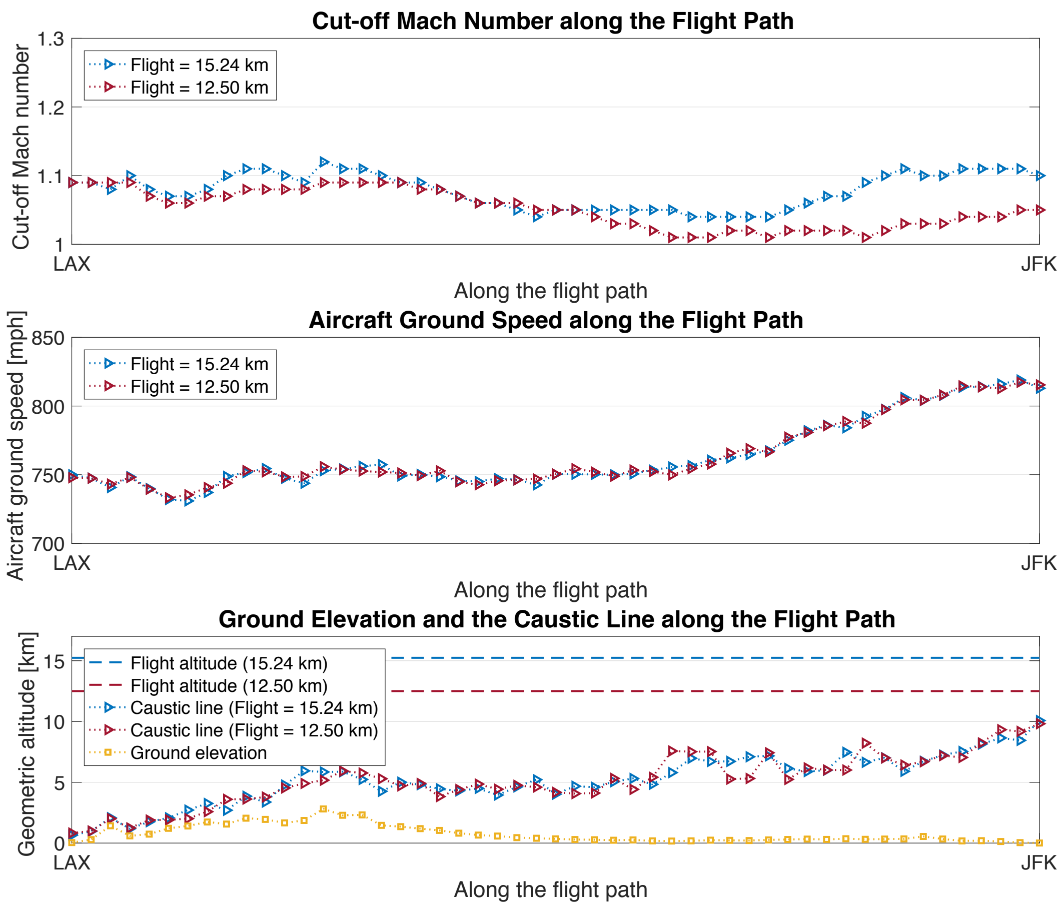


Figure 9. LAX to JFK air route (eastbound with a tailwind) on Jan 1 2017 at 12 PM UTC.

Milestone(s)

Milestone	Date
A 3-D ray-tracing algorithm had been developed, which showed vertical wind speed can be an important factor in the prediction of safe altitude.	January 31, 2018
The safe altitudes and speeds for Mach cut-off flight for a hypothetical trip from Los Angeles to New York was investigated.	March 31, 2018
Examine other atmospheric model datasets, which have better time and spatial resolution than the CFSv2 datasets	September 30, 2018

Major Accomplishments

In this research, a 3-D ray-tracing algorithm has been developed for the acoustical model of Mach cut-off flight, which can read-in realistic atmospheric data that includes arbitrary speed of sound variations as well as arbitrary three-dimensional winds. It is shown that vertical wind can be an important factor in the prediction of safe altitude. For certain wind profiles and aircraft headings a change in Mach number of 0.01 can send a focus boom down to the ground, even at a low Mach number such as 1.04. The safe altitudes and speeds for Mach cut-off flight for a hypothetical trip from Los Angeles to New York was investigated. Because of the persistent tailwind for such an eastward flight, Mach cut-off is much more challenging than for a westward flight.

Publications

Z. Huang and V. W. Sparrow, "An improved Mach cut-off model based on a 3-D ray tracing method and realistic atmospheric data," *J. Acoust. Soc. Am.*, **143**(3), 1913 (2018), invited presentation at 175th Meeting of the Acoustical Society of America, Minneapolis, MN, USA, May 07-11, 2018.

Outreach Efforts

Z. Huang and V. W. Sparrow, " An Improved Mach Cut-off Model Based on a 3-D Ray Tracing Method and Realistic Atmospheric Data," Poster for the Penn State Center for Acoustics and Vibration (CAV) Spring Workshop, University Park, PA, April 24-25, 2018.

Awards

None

Student Involvement

Zhendong Huang is the graduate research assistant supported by Project 42 at Penn State on this task. He is pursuing his Ph.D. in the Penn State Graduate Program in Acoustics.

Plans for Next Period

The project team is examining other atmospheric models with much better temporal and spatial resolutions than the CFSv2 datasets, so that it would be possible to include impact of temporal variations of the atmosphere along a flight path, as well as the contribution of horizontal variation of the atmosphere in a local scale, on the safe altitude and cut-off Mach number. This approach is based on excellent inputs from the Penn State Department of Meteorology on appropriate atmospheric datasets. A statistical analysis is underway which will help to answer the question that how often people could hear sounds from Mach cut-off operations.

References

G. Haglund and E. Kane, "Flight test measurements and analysis of sonic boom phenomena near the shock wave extremity," NASA Report CR-2167 (1973).

Z. Huang and V. W. Sparrow, "An improved Mach cut-off model based on a 3-D ray tracing method and realistic atmospheric data," *J. Acoust. Soc. Am.*, **143**(3), 1913 (2018), invited presentation at 175th Meeting of the Acoustical Society of America, Minneapolis, MN, USA, May 07-11, 2018.

J. Nicholls, "A note on the calculation of 'cut-off' Mach number, " Meteorological Mag. **100** 33-46 (1971).

V. E. Ostashev, D. Hohenwarter, K. Attenborough, P. Blanc-Benon, D. Juvé, & G. H. Goedecke, "On the refraction law for a sound ray in a moving medium, " Acta Acustica united with Acustica, **87**(3), 303-306 (2001).

A. Pierce, "Acoustics: An Introduction to its Physical Principles and Applications, " Acoustical Soc. Am., New York (1989).

K. Plotkin, J. Matisheck, and R. Tracy, "Sonic Boom Cutoff Across the United States," in 14th AIAA/CEAS Aeroacoustics Conference (29th AIAA Aeroacoustics Conference), AIAA Paper 3033 (2008).

S. Saha et al., "The NCEP Climate Forecast System Version 2," J. Clim., **27**(6), 2185–2208, Mar. (2014).

Task 2- Subjective Study on Annoyance, Metrics, and Descriptors

The Pennsylvania State University

Objective(s)

- Identify the key perceptual attributes of Mach-cutoff ground signatures.
- Determine how these attributes are correlated with annoyance ratings of these signals.
- Identify a metric appropriate for predicting annoyance due to Mach-cutoff ground signatures.

Research Approach

Introduction

This research will be the first major contribution to the body of literature on the perception of Mach cut-off. The ultimate goal of this work is to identify a metric that can be used to predicting annoyance due to Mach-cutoff flights. Subjective testing data will inform metric selection and corresponding values for possible use in certification. Mach-cutoff ground signatures are unique sounds and have not been part of the public's day-to-day experience, so the impact on community annoyance is difficult to predict. In addition, these signatures are perceptually different from N-shaped sonic booms. As such, new subjective tests are necessary to assess perception of these sounds and ultimately predict annoyance in communities.

The task is subdivided into three studies: (1) descriptor study, (2) annoyance factor study, and (3) absolute annoyance study. The descriptor study was nearing completion at the end of the last period and was used to identify a set of perceptual attributes that could be used to describe Mach-cutoff ground signatures for the second study, which is the annoyance factor study. This second study is nearing completion and will identify correlation between annoyance and various noise metrics. A secondary goal of this study is to evaluate any possible correlations between annoyance and the attributes identified in the first study. The third study, the absolute annoyance study, is still in the development stages. The purpose of this study is obtain annoyance ratings of the signatures relative to other common traffic noises, i.e. road, rail, and subsonic aircraft noise, and to further evaluate the proposed metric(s) from the second study to use for predicting community annoyance response to Mach-cutoff ground signatures.

Study 1 – Descriptor study

Data Collection

Study 1 was designed as an exploratory study to identify terms used to describe perceptual attributes associated with these signatures since no prior work exists on this topic. Several different subjective testing methods were considered to obtain this set of terms as described in the 2016-2017 annual report. The method of Free Choice Profiling (FCP) was selected since it allows participants complete freedom to identify attributes they associate with this type of sound without any influence from the test structure or administrator. Free Choice Profiling was developed by Williams and Langron [21,2] to study perception of foods, but it has been applied in the audio field [3-5]. The general procedure includes three steps: vocabulary generation, vocabulary refinement, and stimulus ratings.

A set of 24 stimuli was used in this study, which where a subset of Mach cutoff recordings obtained from NASA's Farfield Investigation of No-boom Thresholds (FaINT) recordings [6]. (The selection process was detailed in the 2016-2017 annual report and involved careful listening and categorization of the stimuli.) For the vocabulary generation step, the participants were required to listen to the stimuli in two groups of 12 to generate ideas of appropriate descriptors. Subjects were then instructed to enter up to 10 separate terms that they felt emphasized perceptual differences between the stimuli (see Figure 1 for the testing interface). Participants could listen to each individual stimulus as many times as they chose. The process was then repeated with the other half of the stimuli.

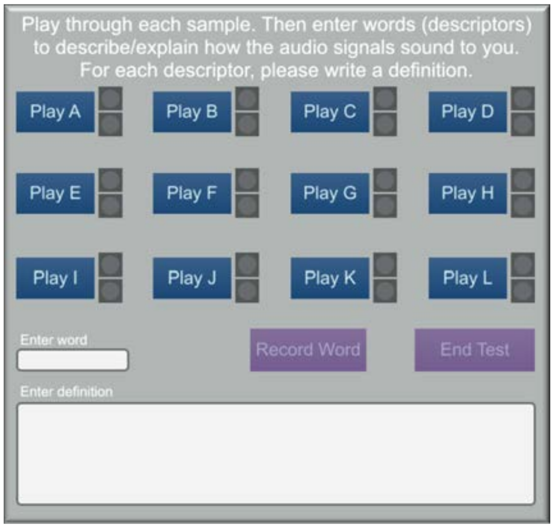


Figure 1. Study 1user interface for descriptor entry. This interface allowed test subjects to enter their own descriptors for Mach-cutoff stimuli along with accompanying definitions.

Once a subject had generated descriptors, the list was reduced to a maximum of 5 though an interview with the test administrator (Ortega). Interview questions included "Are any of the descriptors describing the same aspect of the sound as another descriptor?" and "Would it be difficult to rate stimuli on any of these descriptors?". Once the list was narrowed down to 3-5 attributes, subjects rated this set of attributes for each of the 24 stimuli using the interface shown in Figure 2. (The screen pictured was repeated for each stimulus, with each of the subject's descriptors appearing.) This process served as the final step of the FCP procedure. Note that the stimulus order was randomized for all sets for both listening tests in steps 1 and 3.

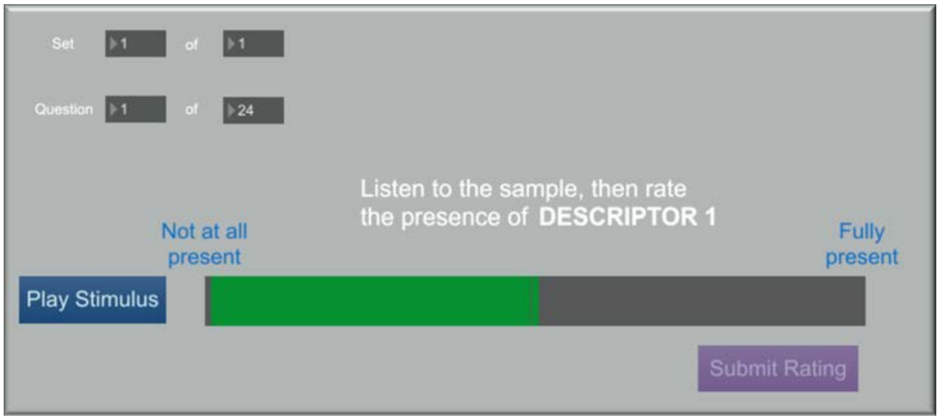


Figure 2. Study 1 user interface for stimulus rating. This interface allowed subjects to rate stimuli on their own descriptors. During testing, a subject's descriptor replaced the label "DESCRIPTOR 1".

Test participants were recruited from the State College community. Inclusion criteria were minimum age of 18 years old and normal hearing thresholds of 25 dB HL or lower for the 125 -4k Hz octave bands. In addition, since musicians have been shown to provide more reliable data in these types of tests with much less training than non-musicians [7], a target was set to have about 2/3rds of the participants be musicians and the remainder non-musicians. In this manner, it was hoped that any noise in the data would be reduced. Non-musicians, however, were explicitly also recruited so that it would be possible to evaluate how representative the terms used by the musicians were of the general population.

A total of 28 subjects participated in the study, where their ages ranged from 18 to 38 years of age, median age 22, with 14 males and 14 females. While this age range is low relative to the general population, this study was intended to be a preliminary study only, and a broader age range was used in the second study and will also be used in the third study. The target of having approximately 2/3rds of the participants was roughly achieved with a total of 19 musicians participating in the study.

Data Analysis

The first step of the analysis of Free Choice Profiling data is the Generalized Procrustes Analysis (GPA) method [8]. This analysis method unifies the rating sets of each subject in order to pool together the results of all participants. This step is necessary since the subjects rated the stimuli using their own individual terms with (potentially) different definitions even for the same terms. The first step of GPA is to scale each set of ratings by subtracting the mean ratings and multiplying each rating on all terms for one subject by one value. This process normalizes each subjects' ratings, adjusting for subjects that use more of the scale than others. Second, the normalized rating sets are rotated in the multi-dimensional space to align the ratings of each stimulus. If each subject "places" the stimuli in their own subjective space, with axes defined by their own descriptors, then GPA rotates all spaces until the stimuli in every space lie on top of each other, as much as possible. Third, GPA defines new axes through Principal Component Analysis (PCA). These axes define a "consensus space".

The average stimulus ratings plotted against the first two principal components in the consensus space are shown in Figure 3. Immediately apparent is that stimulus 23 appears to be an outlier and a separate analysis, described below, supported the decision to exclude the data from this stimulus. A clustering algorithm was run on the rest of the data to find groups of similar stimuli. Overall, the clustering analysis revealed that a single cluster best characterizes all of the stimuli. However, since this finding didn't allow for more specific groupings, the next possible number of clusters resulting from the analysis was three clusters, which led to interpretable results. The three clusters labeled 1, 2, and 3 in Figure 3 are characterized by softer sounds, more impulsive sounds, and more high frequency sounds, respectively.

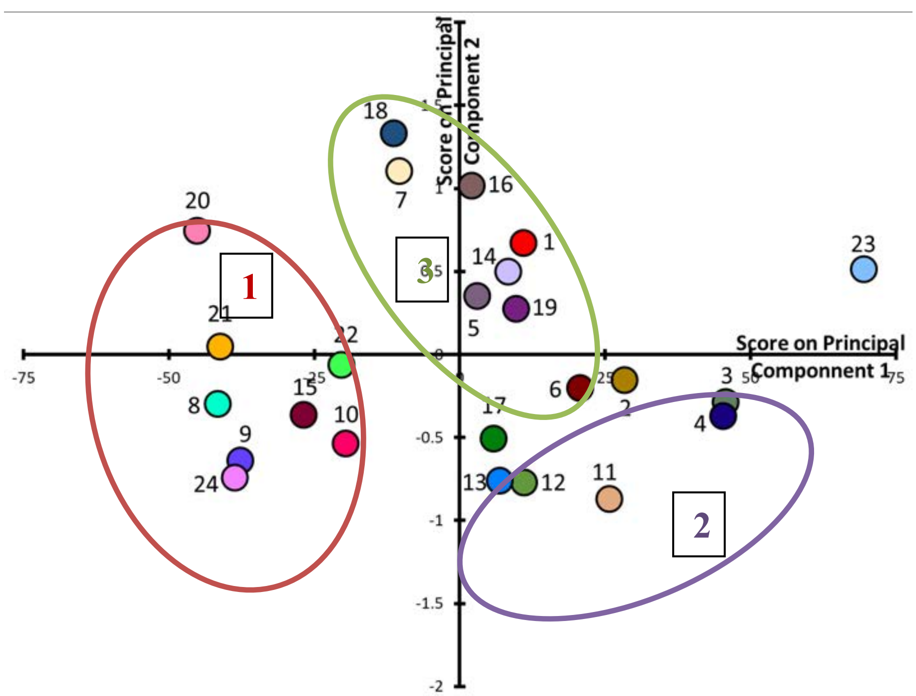


Figure 3. Average stimulus principal component scores after Generalized Procrustes Analysis. Circles show three main clusters. Note that the ratings of stimulus 23 were excluded from the remainder of the analysis as it was determined to be an outlier. Cluster 1 is characterized by softer sounds, 2 by more impulsive sounds, and 3 by more high frequency sounds

To further investigate if stimulus 23 was an outlier, the variance in the principal component scores for all stimuli were calculated. (Traditional statistical analyses can be run once the data has been transformed as described previously into the consensus space). As shown in Figure 4, the variance in the scores of stimulus 23 is nearly twice the average variance for the other stimuli. Therefore, based on both the visual inspection of the results shown in Figure 3 of stimulus 23 scores and the computation of the variance, stimulus 23 was omitted from all data analyses.

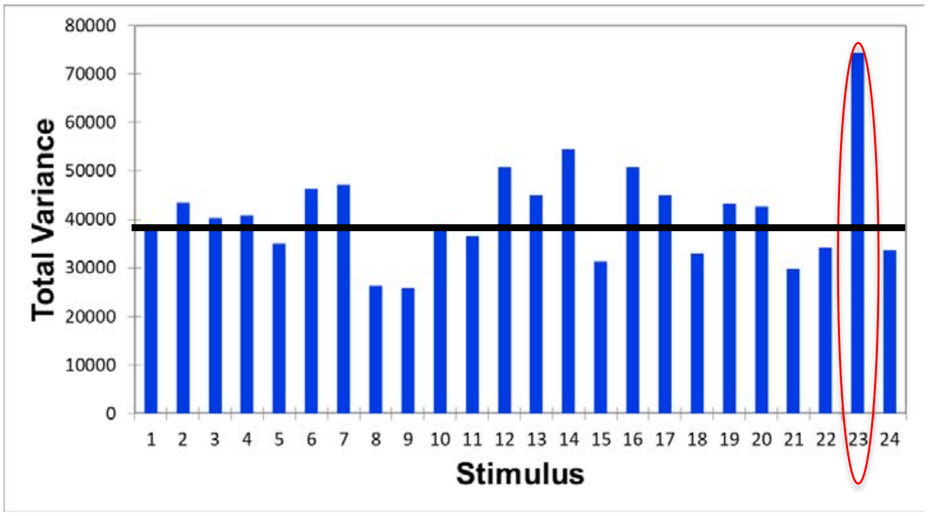


Figure 4. Total variance by stimulus. Note that the variance for stimulus 23 is nearly twice that of the average indicated by the black horizontal line.

In a similar manner to obtaining a set of scores for each principal component for each stimulus, a set of scores can be obtained for each attribute (descriptor) rated by each participant. These results can then be analyzed to determine the number of components (factors) that can explain the results for each individual attribute and the associated factor loading for each descriptor can be obtained. The results of this analysis for all rated terms are shown in Figure 6, where the results are plotted as a function of where each individual term falls on the scales (axes) of the two factors (components). This set of scores for each term can then be used to carry out a clustering analysis to determine how the terms may be grouped to ultimately identify a set of 2 to 4 perceptual attributes associated with Mach-cutoff ground signatures, which can then be used in follow-on studies to identify a metric to predict annoyance.

With the factor loadings obtained in the previous step, meaning can be given to the principal components. First, the terms were clustered using k-means clustering on the factor loadings. In Figure 6, these clusters are highlighted by the four circles. The largest cluster is on the positive end of the first axis and contains terms related to loud, impulsive sounds (example terms are “thunderous” and “powerful”). On the negative end is the next largest cluster, with terms related to soft sounds (e.g. “distance” and “soft”). These clusters show that the first principal component, Factor 1, is likely related to a sense of loudness. The second and third clusters place low frequency terms such as “rumble” on the negative end of the second axis and high frequency terms, such as “whistling” and “white noise”, on the positive end, showing that the second principal component, Factor 2, is likely related to pitch perception (frequency). This information was used to group the terms into clusters as shown. These clusters along with a fifth are broken down into the most common terms from each cluster in Table 1. The most commonly used terms were related to “Thunder” in the first cluster and the next most common term was “Rumble” in the third cluster. No obvious representative terms could be identified for the remaining three clusters. For the second and fourth clusters categorize as “soft” and “high frequency”, respectively, it was deemed unnecessary to identify representative terms, since those clusters appeared to have opposite meaning to “Thunder” and “Rumble”, respectively. For the final cluster, while no clear term emerged, it was decided that this perceptual feature may be worth further investigation. After a close review of the provided terms and definitions in this group, the word “Swoosh” was selected to represent this group, which was a term provided by a non-musician.

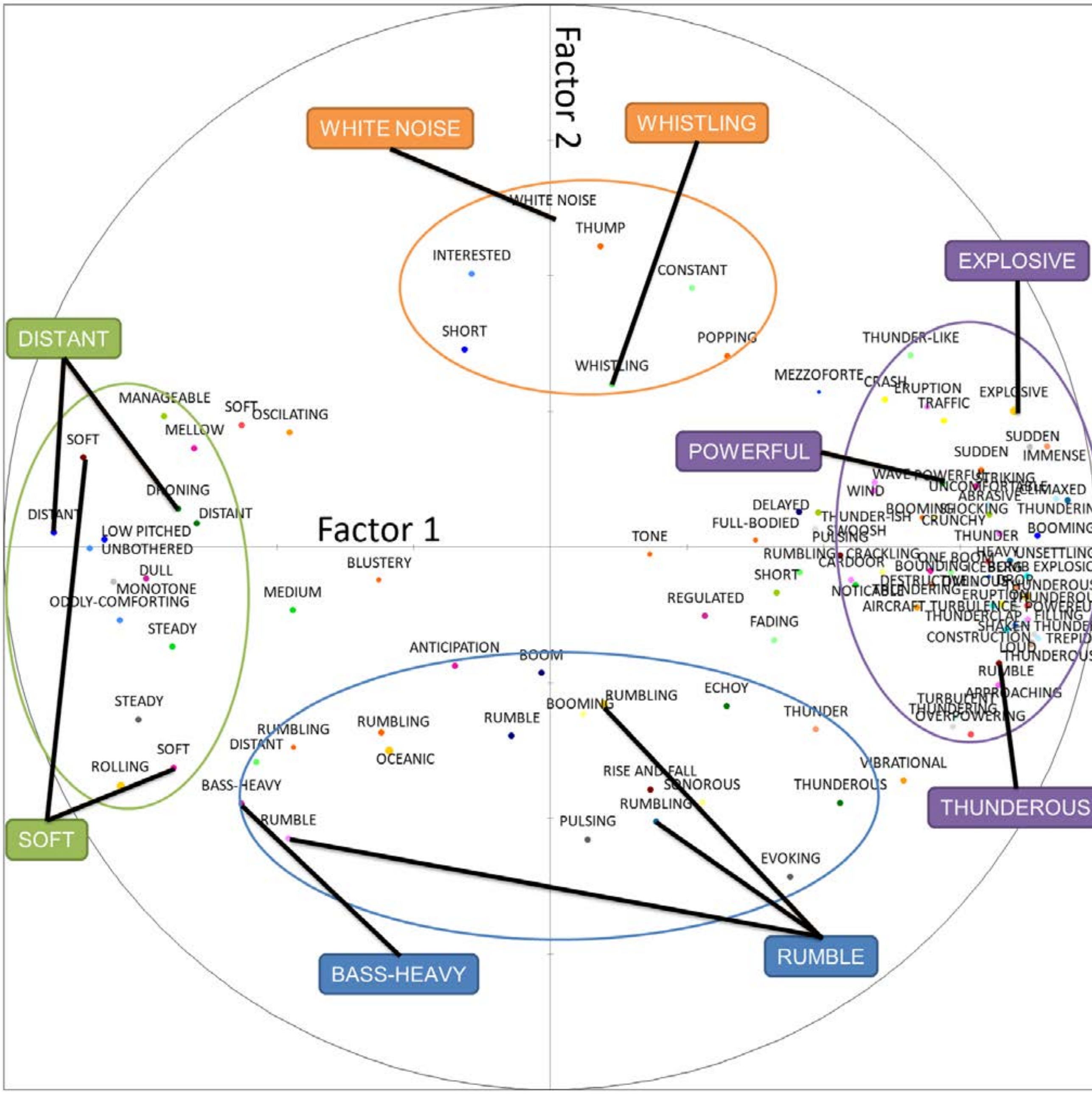


Figure 6. Descriptors plotted by loadings on first two principal components (Factor 1 and Factor 2). Key terms have been highlighted. Ovals show first four clusters of words.

Table 1. Descriptors grouped into clusters. "Thunder" was the most common term. Rumble was the most common term not related to thunder.

Loud		Soft		Low-f		High-f		Swell/Swoosh	
Thunder	13	Soft	3	Rumble	8	White noise	1	Ocean terms	3
Explosion/boom	10	Distant	3	Heavy bass	2	Whistling	1	Pulsing	2
Sudden	5	Steady	2			Short	1	Swoosh	1

Study 2: Annoyance factor study

Study Design

Having established terms linked to the perception of Mach-cutoff ground signatures in Study 1, the second study was designed to link these perceptual aspects to perceived annoyance and noise metrics. The three terms that were selected from the results of the first study were "thunderous", "rumbly" and "swooshing", where the first two each represented one of the two primary factors, and the third was selected as a possible third factor, although not strongly supported by the PCA results. The terms were defined as follows for the participants: "Thunderous indicates the degree to which the sound resembles the crack of nearby thunder (booming, explosive, etc.)"; "Rumbly is defined by an excess of low-frequency sound"; "Swooshing is defined as a sound that transitions from lower to higher pitch and from louder to quieter volume"; and finally for "annoyance", "Annoying describes how annoyed you feel by the sound".

In terms of the experimental design, it was now possible to carry out a more conventional subjective study using a common set of attributes for all participants to evaluate. The experimental method chosen was a combination of the paired comparison and rating scale methods, which has been shown to be an easier task for participants and yield more consistent responses over other test designs. However, this approach, relative to the one used in the first study where participants rated each stimulus individually, is much more time consuming to complete since paired comparison requires that participants evaluate all stimuli relative to all of the other stimuli in pairs. As a result, the number of test questions grows relatively quickly for a small number of stimuli, e.g. for a set of only 6 stimuli, there are 15 possible combinations for stimuli grouped in pairs. Due to testing time constraints to minimize participant fatigue, it was not possible to use all 24 of the stimuli from the first test. As a result, a subset of 6 stimuli was used in this study, where two stimuli from each of the three groups were selected as representative from each group respectively. An additional 3 simulated stimuli were included in the annoyance set, resulting in a total of 36 pairs for that set (further details are below). The overall test was subdivided into four sets, one for annoying and one for each of the three attributes.

The participants were presented a pair of stimuli and asked to select which stimulus had more of a given attribute, i.e. annoying, thunderous, rumbly, or swooshing. Subjects also had to rate the degree with which the one stimulus had more of a given attribute by using the scale shown in Figure 7. (Note that the different attributes ("THUNDEROUS", "RUMBLY", "SWOOSHING", or "ANNOYING") would replace the label "ATTRIBUTE" depending on which term the subject was rating. First, subjects would listen to both stimuli.) More details about the test procedure are provided in the next section.

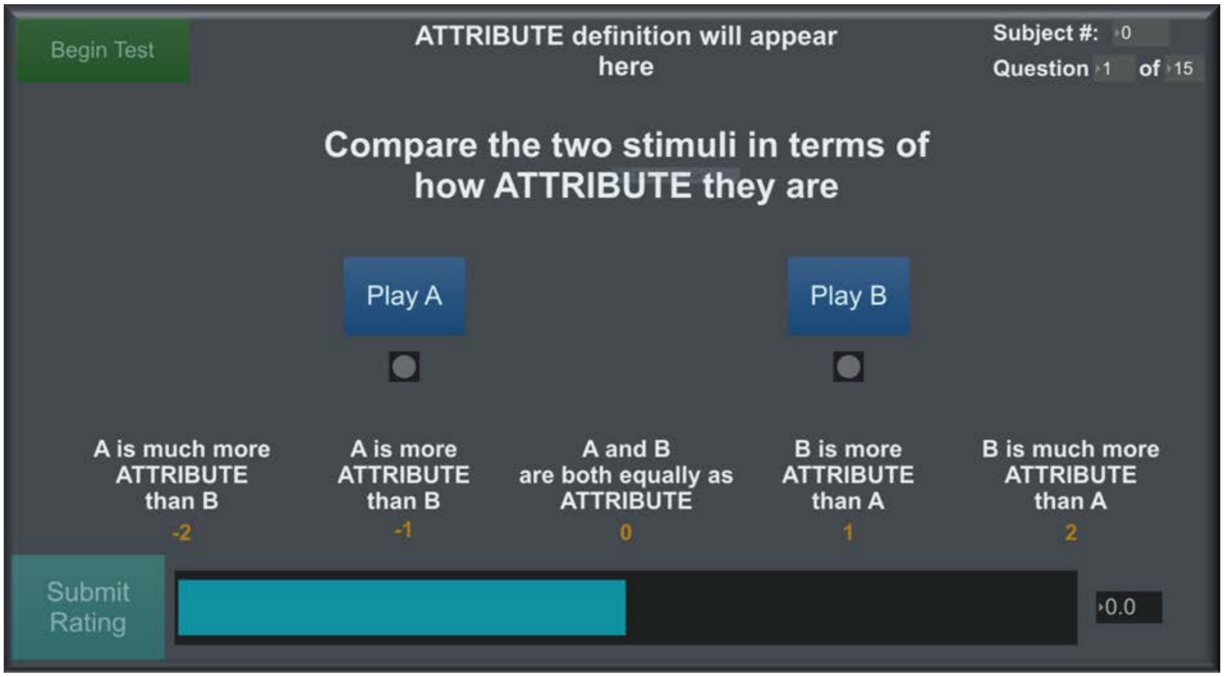


Figure 7. Study 2 rating interface. Subjects would click on the "Play A" and "Play B" buttons, decide which stimulus had more of the given attribute, i.e. annoying, thunderous, rumbly or swooshing, and rate how large that difference was using the slider bar.

The primary independent variable in Study 2 was different Mach cutoff signatures, the same as in the first study. However, for this second study, an additional independent variable was also of interest: the effect of the initial Mach-cutoff evanescent wave on perception, where this initial part of the signature was simulated from the work in Task 1. Three such signatures were synthesized for different elevations above the ground by modifying two input N-waves from glider recordings using a Hilbert Transform method. However, these signals lacked realism due to the lack of post-boom noise. As such, the simulation outputs were combined with a Mach-cutoff recording that had no audible evanescent shock noise and only post-boom noise, labelled "original" in Figure 8. This recording was selected from the set of 6 since it contained minimal background noise. The research team listened to the resulting waveforms and determined that they sounded realistic. The 3 resulting waveforms along with the original recording are plotted in Figure 8. Note that the signals only differed in the early part of the waveform, where the simulation was layered in. Due to testing time constraints, these stimuli were only included in the annoyance set.

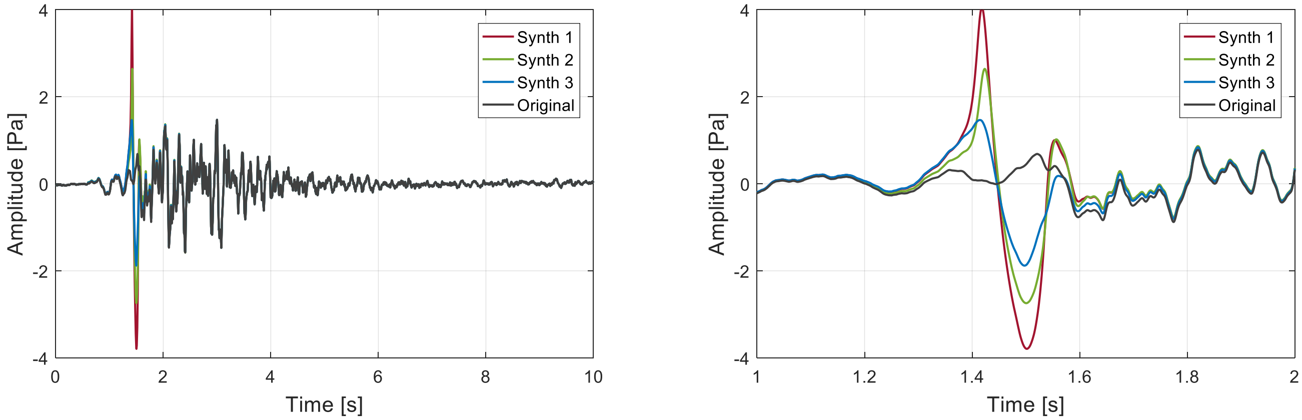


Figure 8. The three additional stimuli included in the annoyance rating set. The three simulated wave forms contain only the initial Mach-cutoff evanescent wave and not the remaining portion of a typical signature. As a result, the synthesized stimuli were combined with one of the six selected recordings to add post-boom noise. The full waveform is shown on left, while a detailed view is shown on right.

A final design element for this second study was to broaden the age of the participants to be more representative of the general population and not only college-aged students. Targets were set to collect 10 participants from each of the four age ranges: 18-29, 30-39, 40-49, and 50-60, with a goal of including 40 participants. To remove the possibility of gender effects on the results, a second target was to recruit an equal number of males and females.

Data Collection

The testing procedure for each subject was as follows. First, subjects would read and sign the informed consent form. Then, a hearing screening would be administered to ensure subjects had normal hearing in the range of interest (1000 Hz and below). Next, subjects would rate the stimuli on a paired comparison scale for each of the terms in four sets, one for each term. The first three sets were in a randomized order and used the terms "thunderous", "rumbly", and "swooshing". As 6 stimuli were used for these sets, there were a total of 15 comparisons made for each of these terms. For the last set, subjects rated the 9 stimuli on the term "annoying" (36 comparisons). Before each set, tutorial slides were presented defining the attribute to be rated in that set (as defined above). Breaks were provided between each set, with a break in the middle of the "annoying" set. Afterwards, demographic and other information were collected through questionnaires. Average testing time was around 80 minutes per subject, with only 2 subjects taking longer than 90 minutes.

As of 5 November 2018, 42 subjects had participated in the study. Age and gender demographics are summarized in Figure 9. In general, there were fewer subjects in the highest age group, but overall the subjects' ages were well distributed across the target age range. The gender ratio is skewed slightly towards female, with a 4:3 female to male ratio. It should also be noted that 7 of the 42 subjects met our criteria to be classified as a musician, though this was not a requirement for participation. Future analyses may evaluate the effect of musical training on the results.

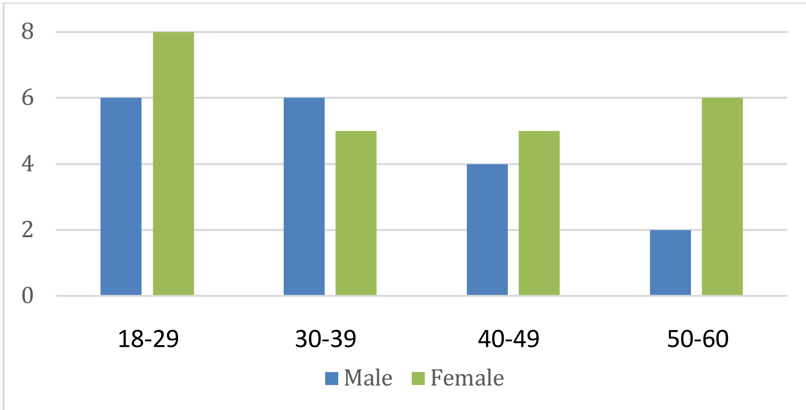


Figure 9. Demographic breakdown of Study 2 participants. Participants dropped off with age, and more females participated than males.

Data Analysis (Preliminary: 39 subjects)
The following analyses are **preliminary** and are based on the results from 39 out of the 42 subjects run (those run before 30 September 2018).

For all analysis, ratings on the paired comparison scale must be converted into ratings for individual stimuli on a relative scale. This conversion is achieved by summing across all pairs for a given stimulus. First, all stimuli are given an initial "rating" of zero. Then, for each pair, the stimulus rated as having more of the given attribute has the magnitude of the comparison rating added to the total rating for that stimulus. This process results in a single rating value for each stimulus for each participant. All the following analyses were run using these ratings.

The first statistical analysis looked at the effect of demographics on these ratings. A multivariate analysis of variance (MANOVA) was run with the ratings on each of the four attributes as the output variables and age category, gender, noise sensitivity, and Mach-cutoff signature (stimuli). A significant effect of ratings on individual stimuli was found ($p < 0.05$), but no statistically significant effects of age category or gender. The average annoyance ratings separated by demographic groups are shown in Figure 10. While there are some visible differences, these are not statistically significant and the overall shapes of the individual curves are similar, which means that the age of a participant did not influence their annoyance ratings. In other words, these findings suggest that future studies may not need to recruit as broad an age range which was done for this study. Finally, the effect of an individual's noise sensitivity was also not found to have a significant effect on the ratings.

The effect of the simulated boom on annoyance ratings was also investigated. For this analysis, ratings were calculated using only pairs including the three simulated booms and the original recording (4 stimuli, 6 pairs). This was done to directly compare the effect of the simulated boom. ANOVA was run with annoyance as the output variable and subject number and stimulus (now only 4 stimuli) as the input variables. While the effect of subject number was significant (at $p < 0.05$), the effect of stimulus was not significant (at $p < 0.05$). This result indicates that the simulated boom portion of the stimulus did not significantly affect the perception of annoyance and invites investigation into whether annoyance from Mach-cutoff ground signatures is due to the post-boom noise only.

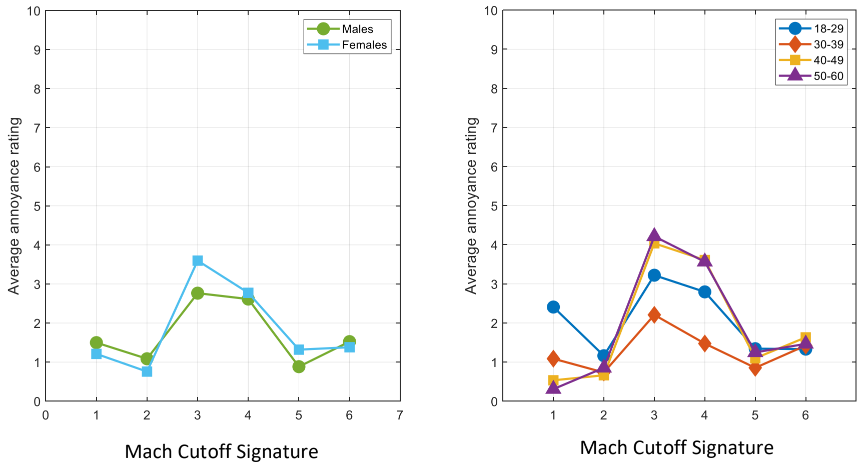


Figure 10. Annoyance ratings of the six Mach cutoff recorded signatures separated by gender (left) and age group (right). No statistically significant effects of either gender or age on ratings were found.

Annoyance ratings were also modeled using a linear regression of ratings of "thunderous", "rumbly", and "swooshing". To run this analysis, ratings on the three attributes were converted into factor scores through a varimax rotation. This process leads to orthogonal regressors, which is assumed in linear regression. The regression revealed a strong dependence of "annoying" ratings on "thunderous" ratings as shown in Table 2. While the dependence was significant for each factor, the slope estimate was strongest for the factor related to "thunderous" ratings, and metrics related to this factor will likely be the strongest predictors of annoyance. This finding will be further investigated in the next period.

Table 2. Linear regression results for dependence of annoyance on perceptual factors. While all factors had a significant effect on annoyance, the thunderous factor had the largest effect (slope).

Factor	p-value (slope $\neq$ 0)	Slope estimate	Standard error
Rumbly	<.0001	0.41	0.10
Swooshing	.0029	0.30	0.10
Thunderous	<.0001	0.79	0.10

Multiple metrics were calculated for each stimulus in order to investigate possible quantities that could be used to predict annoyance and the three perceptual attributes evaluated. Calculation procedures were written in MATLAB. Procedures were written to calculate weighted *sound exposure levels* (SEL) using A-, B-, C-, D-, and E- weighting (SEL_A, etc.). In addition, *perceived sound exposure level* (PL_SEL) was calculated according to the FaINT report [6], as was the sound quality metric *sharpness*.

The calculated metrics were used to evaluate the degree of correlation of these metrics with the rating data, as shown in Table 3. The highest correlations, which are indicated in **bold**, suggest that loudness-based metrics are somewhat correlated with the ratings on "annoying" and "thunderous". These metrics were also often correlated with the other perceptions of "rumbly" and "swooshing" as well, though to a lesser extent. It is apparent that loudness influences all the ratings, so a more complete analysis will be conducted in the next period to parse out which portion of the variance is explained by each metric.

Table 3. Correlations between calculated metrics and average ratings.

	SEL_A	SEL_B	SEL_C	SEL_D	SEL_E	PL_SEL	Sharpness	RC-based
Thunder	**0.92**	**0.95**	0.83	0.89	**0.97**	0. 88	0.61	0.39
Rumble	0.28	0.68	0.81	0.67	0.56	0. 30	0.06	0.62
Swoosh	0.86	**0.93**	0.76	0.82	**0.94**	0. 83	0.43	0.53
Annoy	0.79	**0.97**	**0.91**	**0.91**	**0.94**	0. 77	0.41	0.46

This preliminary analysis indicated that none of these metrics were highly correlated with “rumbly” ratings without also being correlated with thunderous ratings. A non-traditional aircraft noise annoyance metric was explored as a possible metric to predict this attributed. A new metric was developed based on Room Criteria Mark II (RC Mark-II). RC Mark-II was designed to analyze background noise in rooms due to heating, ventilation, and air conditioning (HVAC) systems. This metric was chosen since it includes an indicator for the perception of “rumble”. In particular, the numerical quantity used in the RC Mark-II metric that predicts the perception of “rumble” was selected as the basis for this proposed “RC-based” metric. However, due to the lack of higher frequency content in these signatures, which would typically be present in HVAC noise, an early step in the calculation has been changed to be based on the mid- and low-frequency octave bands. The correlation analysis of this proposed RC-based metric and the perceptual ratings are shown in the last column of Table 3. As shown, the correlation values are relatively low for this proposed metric, however, further analyses are planned to evaluate if the variance due to loudness is reducing the predictive power of this metric.

Milestone(s)

Milestone	Planned Due Date	Status
Report on Mach cut-off descriptor subjective test	November 1, 2017	Complete
Report on experimental design of metric and annoyance study	February 1, 2017	Complete
Report on possible metrics that could be used to predict annoyance due to Mach cut-off sounds	July 31, 2018	In progress

Major Accomplishments

- Subjective testing for descriptor study (Study 1) was completed – 28 subjects participated in the study
- Study 1 data were analyzed – stimuli were characterized and the most significant terms were determined to be “Thunderous” and “Rumbly”.
- Annoyance factor study (Study 2) was designed – paired comparison test designed to look at Annoyance along with perception of “Thunder”, “Rumble”, and “Swoosh”.
- Study 2 data collection completed – 42 subjects participated in the full study (by Nov. 5).
- Multiple metric calculation procedures written for Study 2 data analysis – metric calculation procedures written for sharpness, sound exposure level (A-, B-, C-, D-, and E- weighting), Perceived Loudness, Perceived Sound Exposure Level, and an RC-based metric to possibly predict “Rumbly” ratings.
- Study 2 preliminary annoyance data analysis completed – annoyance deemed to be related to perception of thunder; ratings not likely affected by demographics of subjects.
- Further statistical analyses planned for Study 2 – in particular, a stepwise regression will be used to determine which metrics to include in the full analysis.

Publications

Acoustical Society of America Spring 2018 meeting abstract [9].

Outreach Efforts

CAV Workshop 2018 – Poster presentation: This consisted of one poster including the results from the descriptor study.

Awards

CAV Workshop 2018 student poster competition – 3rd place.

Student Involvement

Nicholas Ortega was primarily responsible for test design, stimulus selection, test preparations, statistical analyses, and presentation preparations for the first two studies. He also presented the poster at the CAV workshop and the talk at the ASA Conference.

Jonathan Broyles started in September 2018 and has been responsible for test design for Study 3. He will complete work on Study 3 during the following period.

Plans for Next Period

The major goals for the next period are to (1) complete the data analysis for Study 2 to propose a set of possible metrics to predict annoyance due to Mach cutoff ground signatures by January 2019 [Ortega] and (2) design and conduct a third subjective study to investigate the degree of annoyance due to Mach cutoff signatures relative to typical traffic noise [Broyles].

Based on the results of the second study wherein the inclusion of simulated stimuli did not have a significant effect on the annoyance ratings, it was decided that the third study will not include any simulated stimuli. The aims of the third study were further adjusted after meeting with AERION representatives in September 2018. From the discussions, it was determined that utilizing everyday traffic stimuli in context to the FaINT Mach-cutoff signatures based on annoyance would be beneficial in developing an annoyance metric. The everyday traffic stimuli of interest include road, rail, and subsonic air traffic noise. The proper authorities for each stimulus will be contacted to obtain permission to obtain recordings. Recording the stimuli will commence late November and be complete by December 2018. The experimental design will be developed in early 2019, data collection to take place in March – April 2019, and data analysis and final report by July 2019.

References

1. A. A. Williams; S. P. Langron, "The Use of Free-choice Profiling for the Evaluation of Commercial Ports", J. Sci. Food Agri. **35** 558-568 (1984).
2. V. Dairou; J.-M. Sieffermann, "A Comparison of 14 Jams Characterized by Conventional Profile and a Quick Original Method, the Flash Profile", J. of Food Sci. **67**(2) 826-834 (2002).
3. L. Gaëtan, ""Individual Vocabulary Profiling of Spatial Enhancement Systems for Stereo Headphone Reproduction", Audio Engineering Society Convention 119, paper 6629 (2005).
4. T. Lokki; J. Pätynen; Antti Kuusinen, *et al.,* "Concert hall acoustics assessment with individually elicited attributes", J. Acousti. Soc. Am. **130**(2) 835-849 (2011).
5. M. Gerzon, "Periphony: With-Height Sound Reproduction," J. Audio Eng. Soc. **21**(1) 2-10 (1973).
6. L. J. Cliatt II; M. A. Hill; E.A. Haering Jr., "Mach cutoff analysis and results from NASA's Farfield Investigation of No-boom Thresholds", 22nd AIAA/CEAS Aeroacoustics Conference, AIAA 2016-3011 (2016).
7. M. S. Lawless; M. C. Vigeant, "Effects of test method and participant musical training on preference ratings of stimuli with different reverberation times." J. Acoust. Soc. Am. **142**(4) 2258-2272 (2017).
8. J. C. Gower, "Generalized Procrustes Analysis," Psychometrika **40**(1) 33-51 (1974).
9. N. D. Ortega; M. C. Vigeant; V. W. Sparrow, "Perceptual characterization of Mach cut-off Sonic Booms," J. Acoust. Soc. Am. **143** (3, Pt. 2) 1936 (2018). Presentation at ASA Spring 2018 in Minneapolis, MN, USA in May 2018.