



Partnership for Air Transportation
Noise and Emissions Reduction
An FAA/NASA/Transport Canada-
sponsored Center of Excellence



Low Frequency Noise Study

prepared by
Kathleen K. Hodgdon, Anthony A. Atchley,
Robert J. Bernhard

April 2007

Final Report
PARTNER Low-frequency Noise Study

Kathleen K. Hodgdon, Anthony A. Atchley, Robert J. Bernhard

PARTNER-COE-2007-001

Opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the FAA, NASA or Transport Canada.

The Partnership for Air Transportation Noise and Emissions Reduction — PARTNER — is a cooperative aviation research organization, and an FAA/NASA/Transport Canada-sponsored Center of Excellence. PARTNER fosters breakthrough technological, operational, policy, and workforce advances for the betterment of mobility, economy, national security, and the environment. The organization's operational headquarters is at the Massachusetts Institute of Technology.

The Partnership for Air Transportation Noise and Emissions Reduction
Massachusetts Institute of Technology, 77 Massachusetts Avenue, 37-395
Cambridge, MA 02139 USA
<http://www.partner.aero>
info@partner.aero

The PARTNER Low Frequency Noise (LFN) team consisted of university members from The Pennsylvania State University, Purdue University, and The University of Central Florida, with industry and government partners from Boeing, G. E. Aircraft Engines, The Metropolitan Washington Airports Authority, The Volpe National Transportation Systems Center, and Wyle Laboratories during the LFN field study conducted in 2004. Members from Bombardier, Harris Miller Miller and Hanson, and Pratt and Whitney later joined the team in an advisory capacity during the analysis of data and drafting of this report.

This work was funded by the Office of Environment and Energy, U.S. Federal Aviation Administration under Contract FAA 03-C-NE-PSU. The project was managed by Dr. Mehmet Marsan.

The content of this report draws heavily from the masters theses of graduate students at Penn State and Purdue University. They are denoted with "*" below.

University Team Members:

Penn State

Anthony Atchley

Xiao Di

Bradley Dunkin*

Thomas Gabrielson

Remy Gutierrez*

Thomas Hettmansperger

Kathleen Hodgdon

Erin Horan*

Matthew Nickerson*

Peter Shapiro*

Purdue University

Robert Bernhard

Patricia Davies

Luc Mongeau

Daniel Robinson*

Univ. of Central Florida

Roger Wayson

Sponsors:

FAA: Lourdes Maurice, Mehmet Marsan, Tom Connor

NASA: Kevin Shepherd, David McCurdy

Advisory Board Members:

Boeing: Mahendra Joshi, Belur Shivashankara

Bombardier: Mark Huising

G.E. Aircraft Engines: Philip Gliebe

HMMH: Nick Miller

Metropolitan Washington Airport Authority: Neal Phillips, James Rushing

Pratt and Whitney: Ebad Jahangir

Volpe National Transportation Systems Center: Gregg Fleming, Eric Boeker, Mike Lau, Chris Roof

Wyle Laboratories: Ben Sharp, Kenneth J. Plotkin

Table of Contents

List of Acronyms and Symbols.....	1
1. Executive Summary	3
2. Overview of the Study and Organization of the Report.....	9
3. Background	10
3.1 Subjective Perception, Weighting Networks and Low Frequency Annoyance.....	10
3.2 Findings of Previous Airport Studies of Low-Frequency Noise	11
3.3 Objective Indicators of Low Frequency Noise Annoyance.....	14
4. Addressing FICAN Findings on MSP Expert Panel Report	19
5. Low-Frequency Noise Field Study Design	21
5.1 Airport Selection Process.....	21
5.2 Source Noise Measurements.....	24
5.3 Measurements of Low-Frequency Noise Impact Residential Structures.....	26
5.4 Meteorological Data.....	32
6. Source Noise.....	33
6.1 Sideline Noise During Start-of-Takeoff Roll and Acceleration	33
6.2 Thrust Reverser Noise.....	36
6.3 Low Frequency Propagation Modeling	39
7. Noise and Vibration Impact on Residential Structures	43
7.1 Metrics for Vibration Impact and Subjective Perception of Annoyance.....	43
7.2 Vibration Impact on Residential Structures.....	44
7.2.1 Window Vibration	44
7.2.2 Wall and Floor Vibration	48
7.2.3 Hubbard's Sound Pressure Level Threshold Criteria.....	50
7.2.4 Outdoor Metric Indicators of Indoor Audible Window Rattle	51
7.2.5 Correlation of Outdoor Sound Pressure Metrics with Interior Vibration Levels..	52
8. Design and Analysis Subjective Jury Trials.....	56
8.1 Spectral Balance and Rattle Subjective Study Design.....	56
8.1.1 Noise Signatures Used in the Spectral Balance and Rattle Subjective Tests	57
8.1.2 Subjective Test Methodology	59
8.1.3 Analysis Methodology	60
8.1.4 Results.....	62

8.1.4.1	Spectral Balance Study.....	62
8.1.5	Audible Rattle Study.....	65
8.1.6	Objective Metrics of Signatures Used in the Subjective Studies.....	68
8.1.7	Correlation of Objective and Subjective Rankings.....	69
8.1.7.1	Findings from the Spectral Balance Study.....	70
8.1.7.2	Findings from the Audible Rattle Study.....	71
8.1.8	Comparison of Objective and Subjective Grouping of Signatures.....	72
8.1.8.1	Spectral Balance Study.....	72
8.1.8.2	Audible Rattle Study.....	73
8.1.9	Comparison of Spectral Balance Signatures to Signatures obtained near Florida Airport.....	74
8.1.10	Correlation of Results with the Spectrum of the Signature.....	75
8.2	Ability of Tokita Nakamura Threshold to Prediction Aviation Noise Impact.....	77
8.2.1	Experimental Design.....	80
8.2.2	Data Analysis and Results.....	80
9.	Rattle and Low Frequency Sound Insulation Studies.....	84
9.1	Laboratory Study of Rattle Onset.....	84
9.1.1	Rattle Results.....	85
9.1.2	Theoretical Models of Rattle.....	90
9.2	Low Frequency Sound Insulation Study.....	94
10.	Summary, Findings and Recommendations.....	98
	References.....	102

List of Acronyms and Symbols

Acronyms

ANSI	American National Standard Institute
BTL	Bradley-Terry-Luce statistical method
BWI	Baltimore-Washington International Airport
CFR	Code of Federal Regulations
DNL	Day Night Level
FAA	Federal Aviation Administration
FICAN	Federal Interagency Commission on Aviation Noise
FICON	Federal Interagency Committee on Noise
HA	Highly Annoyed
HF	High frequency
HP	High-performance window
HVAC	Heating Ventilating and Air-Conditioning
IAD	Washington Dulles International Airport
ICP	Integrated Circuit Piezoelectric
IMAGINE	Improved Methods for the Assessment of the Generic Impact of Noise in the Environment
ISO	International Standards Organization
LF	Low-Frequency
LFN	Low-Frequency Noise
MAC	Metropolitan Airports Commission
MSP	Minneapolis St. Paul International Airport
NASA	National Aeronautic Space Administration
OAG	Official Airline Guide
OITC	Outdoor-Indoor Transmission Class
PARTNER	Partnership for AiR Transportation Noise and Emissions Reduction
PE	Parabolic Equation
SAE	Society of Automotive Engineers
SASS II	Gulfstream Supersonic Acoustic Signature Simulator II
SPL	Sound Pressure Level
STOR	Start-of-Takeoff-Roll
STC	Sound Transmission Class
TC	Transport Canada

TR	Thrust Reversal
UHP	Ultra-High-Performance Window
USGS	United States Geological Survey

Symbols

L, L_A, L_C	Sound Pressure Level (Flat-, A- and C-Weighted)
$L_{A,LF}$	A-weighted Low-Frequency Sound Pressure Level
L_{dn}	Day-Night Average Sound Level
L_E, L_{AE}, L_{CE}	Sound Exposure Level (Flat-, A- and C-Weighted)
L_{eq}, L_{Aeq}, L_{Ceq}	Equivalent Sound Pressure Level (Flat-, A- and C-Weighted)
$LFNR$	Low Frequency Noise Rating
$LFSL$	Low Frequency Sound Level
L_{LF}	Low-Frequency Sound Level
LL	Steven's Loudness (ISO 532A)
$LLSEL$	Loudness Level Sound Exposure Level (ISO 226-1987)
$L_{max}, L_{Amax}, L_{Cmax}$	Maximum Sound Pressure Level (Flat-, A- and C-Weighted)
PNL	Perceived Noise Level
Q	Disturbance Index
$Q_A (\alpha=3/40)$	Disturbance Index, A-weighting
$Q_{PNL} (\alpha=1/10)$	Disturbance Index, Perceived Noise Level
$Q_{PNL} (\alpha=3/40)$	Disturbance Index, Perceived Noise Level
R_s	Spearman Rank Correlation Coefficient

1. Executive Summary

This report documents a study to investigate human response to the low-frequency content of aviation noise, or low-frequency noise (LFN). The study comprised field measurements and laboratory studies. The major findings were:

1. Start-of-takeoff-roll, acceleration down the runway, and thrust reversal generate high levels of LFN (below 200 Hz) at critical distances from runways (around 3000 ft in the study) which can be annoying to people living around airports.
2. Hubbard exterior sound level criteria works well as a first level assessment tool for vibration/rattle due to LFN.
3. A-weighted Sound Pressure Level (L_{Amax}) and C-weighted Sound Pressure Level (L_{Cmax}) metrics correlate well with laboratory based subjective response to indoor aircraft noise when LFN levels are low to moderate. The same holds for rattle annoyance (again for low to moderate level LFN). Also, multiple low level LFN events may cause rattle (i.e. simultaneous multiple runway operations).
4. When high levels of LFN are present, Tokita & Nakamura thresholds with C-Weighted Sound Exposure Level (L_{CE}) metric should be used as an indicator of potential for LFN annoyance. The low-frequency noise based metrics did not perform as well as L_{CE} . Data lower than 50 Hz is needed to assess vibration/rattle annoyance.
5. The risk of window rattle is lowered with preload and avoiding resonance response in the design. Outdoor-Indoor Transmission Class (OITC) is a better rating for rattle prone applications than Sound Transmission Class (STC) commonly used in rating windows for transmission loss.

The City of Richfield and the Metropolitan Airports Commission established a Low-Frequency Noise Expert Panel in 1998 to study existing and potential impacts of low-frequency aircraft noise in communities around Minneapolis St. Paul International Airport (MSP). The MSP Expert Panel issued their findings in 2000 in a three volume report.¹ After reviewing the report and meeting with a subset of the MSP Expert Panel in 2001, the Federal Interagency Commission on Aviation Noise (FICAN) issued a report in 2002 containing its response to the MSP Expert Panel's findings and recommendations for future research in low-frequency noise around airports.² FICAN's response was critical of the MSP Expert Panel report in several respects. As a result of this dissent, Congress (H1200, House Congressional Record 12 February, 2003) directed the Federal Aviation Administration (FAA) to conduct a further study

of low-frequency aircraft noise to address the issues raised by FICAN. The follow-on study was conducted on behalf of the FAA by the Partnership for AiR Transportation Noise and Emissions Reduction (PARTNER), an FAA/NASA/Transport Canada-sponsored Center of Excellence.

Reasons for focusing on the low-frequency components of aircraft noise are that

- 1) low-frequency sound encounters less absorption as it travels through the air than higher frequency sound, so it persists for longer distances from the airport,
- 2) the amount of sound transmitted from the outside to the inside of buildings is greater at low frequencies sound than at higher frequencies,
- 3) A-weighting metrics commonly used in assessing the impact of aircraft noise deemphasizes low frequencies,
- 4) standard noise models used for assessing airport noise neglect source noise below 50 Hz,³ and
- 5) prior research indicates that frequencies in the 20–80 Hz range have an influence on the perception of low-frequency noise.^{1,4,5}

FICAN recommended that further research consider the following:

1) That measurements be conducted in houses within critical distances from runways identified in previous studies of low-frequency aircraft noise, in particular one conducted at Baltimore-Washington International Airport (BWI).⁶ Measurements should include exterior noise and window, wall, and floor vibration with a frequency range extending down to a few hertz to capture the low-frequency impact. The vibration measurements should be based on the recommendations by the American National Standard Institute (ANSI) Standard S3.29-1983 (R1996).⁷ In addition, the measured noise and vibration levels should be compared to thresholds for tactile perception of vibration, known as the "Hubbard criteria,"⁸ used to establish the extent of the effect of low-frequency noise at BWI.

2) Have panels of subjects rate the annoyance of individual aircraft events in the houses. Conduct statistical analysis to establish what combination of physical measures gave the best prediction of annoyance ratings. Assess the ANSI Standard [S12.9, Part 4]⁹ Low-Frequency Level (L_{LF}) as a descriptor of low-frequency noise.

3) Study the efficacy of sound insulation in a stepwise fashion, beginning with the most rattle-prone features of houses, the windows and doors. FICAN's idea was to use the same subjects as in Recommendation 2 to assess the impact of insulation.

In addition to considering FICAN's recommendations, PARTNER investigators reviewed prior studies of low-frequency noise in the vicinity of airports^{1,4,6,10,11} and published archival literature, met with representatives of the City of Richfield, members of the Metropolitan Airports Commission, FICAN, and the MSP Expert Panel, consulted with other experts and researchers, and established a set of airport selection criteria to design the follow-on low-frequency noise study.

The airport selection criteria included physical attributes such as a representative aircraft fleet mix, houses within appropriately 4000 ft of the runways, flat terrain with minimal obstructions between the runway and houses, and minimal interference of other sources of transportation

noise. Because the initial plans included subjective evaluations of noise, it was important to avoid communities with prior histories of low-frequency noise problems which could bias the findings. The cooperation of an airport was also necessary. Airports with no history of low-frequency noise problems were reluctant to cooperate out of concern that the study would create a problem where none had existed. The investigators judged the best compromise was to conduct the field study at an airport that met the necessary physical attributes to permit implementation of FICAN Recommendation 1, but abandon the in-residence subjective evaluations component. As a result, FICAN Recommendation 2 was implemented through laboratory-based subjective evaluations substituted in their place.

One of the topics discussed during the meeting between the investigators and FICAN was the practicality of implementing Recommendation 3 regarding the sound insulation study. The investigators determined that to obtain statistically valid results and account for different housing construction, a large number of houses would have to be evaluated. Doing so would require resources well beyond those available for the current study. However, a laboratory-based insulation study was incorporated into the study.

The final low-frequency noise study included the following three components:

- Field measurements conducted at Washington Dulles International Airport (IAD), which included measurements of aircraft noise alongside active runways and noise and vibration impact at two residential structures located approximately 3000 ft from the runways.
- Three subjective tests conducted in an aircraft noise simulator to determine which metrics correlate best with human response to the types of aircraft noises that might be experienced in houses located near active runways.
- Laboratory studies of window rattle and low-frequency sound insulation to develop correlations between low-frequency noise, building design, and the vibration, interior noise and rattle response of houses in communities surrounding airports, with the goal of making recommendations for reducing or eliminating rattle emissions.

The principal findings and recommendations of this study are as follows:

1) Field Measurements

Source Noise

Finding: The highest levels of noise near the runway during start-of-takeoff-roll, acceleration down the runway, and thrust reversal are at frequencies below 200 Hz.

Measurements of sideline noise at start-of-takeoff-roll show that the larger the aircraft, the higher the noise levels, and the levels steadily decrease as the aircraft moves down the runway. Measurements of noise levels during thrust reversal do not show the same trend with aircraft size. The largest aircraft and highest thrust-rating category do not have the highest noise levels.

Recommendation: The Integrated Noise Model uses forward-thrust noise data to model thrust

reverser noise. Thrust reverser noise was identified by the MSP Expert Panel as a potential significant contributor to low frequency noise annoyance and was shown in this study to have significant levels of low-frequency noise. Both the levels and directivity of thrust reverser noise should be investigated further to determine if modifications to noise models are warranted.

Noise and Vibration Impact at Residential Structures

Finding: Measured vibration levels of windows in houses located within 3000 ft of runways can exceed the Hubbard threshold criteria, indicating the potential for vibration to be perceived by occupants. The thresholds were exceeded to a greater degree on a rattle-prone window. For the most part, the vibration levels of secure windows fell below the Hubbard thresholds. The level of wall vibrations for takeoff or landing events having the highest exterior peak C-weighted sound levels can exceed the Hubbard threshold. The vibration levels of the floors did not rise significantly above the background level.

Hubbard's exterior sound pressure level criteria, cited in FICAN's response to the MSP Expert Panel, are consistent with the direct vibration measurements and proved to be good indicators of the onset of window rattle.

Sound Exposure Level L_E , Low-Frequency Sound Level L_{FSL} and Low-Frequency Level L_{LF} , and the Maximum Sound Pressure Level L_{max} correlated well with vibration levels of a rattle-prone window. The A-weighted metrics correlated poorly with the acceleration levels of a rattle prone window. .

Recommendation: The Hubbard exterior sound pressure level threshold criteria should be used as a first assessment of the potential for low-frequency noise impact.

Recommendation: Modern windows have optional plastic grid inserts. The rattle thresholds for these types of windows should be assessed.

2) Subjective Assessment of Low-Frequency Noise

Spectral Balance Study:

Finding: The spectral balance study included only single event signatures. Several level-, loudness-, and perceived noisiness-based metrics correlated well with subjective evaluations of indoor aircraft noise, in particular, L_{Amax} and L_{Cmax} .

Recommendation: Because L_{Amax} and L_{Cmax} are simple metrics to implement, they should be used to predict subjective response to indoor aircraft noise when the levels are appropriate for A- and C-weightings and there are not high levels of low-frequency noise.

Audible Rattle Study:

Finding: The rattle trial included four signatures that resulted from noise impacts from events on

multiple runways simultaneously. Audible window rattle was more likely to be observed for a combined event than for a single event.

Signatures that contained audible rattle were not ranked as the most annoying, most likely because the rattle content was audible, but not loud, relative to the overall noise content of the signature. This result is consistent with other studies of noise containing audible rattle.¹² The subjective rankings of the rattle signatures were grouped together and in the same order relative to the non-rattle signatures, regardless of analysis method. A-weighted and perceived noisiness-based metrics correlated well with the subjective rankings.

Recommendation: Similar to the recommendation from the spectral balance study, L_{Amax} should be used to predict subjective response to indoor aircraft noise when the levels are appropriate for A-weightings and there are no high levels of audible low-frequency noise. Assessment of rattle impact should include both single and multiple events in areas where noise from multiple runways can impact a neighborhood simultaneously. The combined events may create sufficient inaudible low frequencies to induce an audible window rattle.

Assessment of Tokita & Nakamura Threshold for predicting perception of LFN:

Finding: The Tokita & Nakamura annoyance thresholds were validated as predictors of annoyance due to low-frequency aircraft noise. They were found to relate favorably to the subjective annoyance assessments. Linear regression analysis showed that the C-weighted sound exposure level L_{CE} was the best single-metric predictor of subjective annoyance response, explaining over 90% of the variability of the data set. L_{CE} correlated better with the subjective data than metrics specifically designed to quantify low-frequency noise impact.

Recommendation: The Tokita & Nakamura thresholds should be used as indicators of the potential for annoyance due to low-frequency aircraft noise. L_{CE} should be used as a single-number metric for assessing the potential for annoyance when high levels of low-frequency aircraft noise are present.

Finding Valid Across All Subjective Studies: For interior levels without a strong low frequency component the A-weighting captured the perception. For interior levels with strong low frequencies, C-weighting correlated better than A-weighting. Loudness based metrics that included the full frequency range of interest also correlated well. The low frequency based metrics did not correlate better than the level and loudness based metrics. Level influenced perception more than rattle content when assessed in the laboratory.

Overall, the findings suggest that people are responding to the broad spectral content and any predictive metric should quantify the full broadband noise. Loudness algorithms should include frequency content below 50 Hz to optimally correlate with the perception of low frequency noise.

3) Laboratory Rattle and Sound Insulation Studies

Finding: The rattle study explained why rattle can occur at acceleration levels below 1g where previous models had predicted onset. Resonant systems tend to have a rattle frequency band around resonance. This result was verified experimentally. The models developed during the study give the capability to identify mitigation strategies.

Recommendation: A general strategy for eliminating rattle in resonant systems is to increase preload and design the systems so that excitation does not coincide with system resonance.

Finding: The window transmission loss study showed that transmission loss performance is degraded at low frequency by resonance. These resonances are either due to panel vibration or from mass-air-mass interactions of the windows and air gaps between them. Where low frequency excitation occurs, our studies show that the Outdoor-Indoor Transmission Class (OITC) rating is a better than the Sound Transmission Class (STC) rating for identifying the performance of windows.

Recommendation: The Outdoor-Indoor Transmission Class should be used for rating window performance.

2. Overview of the Study and Organization of the Report

This study was mandated by the United States Congress (H1200, House Congressional Record 12 February, 2003) to address the issues raised by FICAN² concerning the report of the MSP Expert Panel.¹ In addition to addressing FICAN's recommendations, this study is intended to contribute to a better understanding of the impact of low-frequency aircraft noise on communities, and assess which metrics are most effective in predicting it. This follow-on study was conducted on behalf of the FAA by the Partnership for AiR Transportation Noise and Emissions Reduction (PARTNER), an FAA/NASA/Transport Canada-sponsored Center of Excellence. In addition to considering FICAN's recommendations, PARTNER investigators reviewed prior studies of low-frequency noise in the vicinity of airports^{1,4,6,10,11} and published archival literature, meet with representatives of the City of Richfield, members of the Metropolitan Airports Commission, FICAN, and the MSP Expert Panel, consulted with other experts and researchers, and established a set of airport selection criteria to design the follow-on low-frequency noise study.

This study included both field measurements at Washington Dulles International Airport, laboratory-based subjective jury trials, and laboratory-based rattle and low frequency sound insulation studies. Key aspects of the study included:

- Measurement of aircraft source noise close to runways
- Measurement of noise and vibration impact at residential structures close to runways
- Three subjective jury trials to assess impact of low frequency noise
- Correlation of jury test results with metrics
- Identification of metric(s) that correlate best with subjective responses
- Investigation of quantitative measurements to assess the potential for annoyance due to low-frequency noise
- Laboratory-based study of rattle and low-frequency sound insulation
- Investigation of low-frequency sound propagation models

This document is the final report of the PARTNER low-frequency noise study. Section 1 is the Executive Summary. Section 3 contains background including a summary of previous studies of low-frequency noise around airports and objective metrics used to predict low-frequency noise annoyance. Section 4 discusses how this study addresses FICAN's recommendations. The process used to select an airport for the field measurement part of this study, along with the design of the field measurement, are described in Section 5. Section 6 presents the results of the source noise measurements. The noise and vibration impact on residential structures are discussed in Section 7. Section 8 describes the design and results of the subjective tests, while Section 9 details the rattle and low-frequency sound insulation parts of the study. Section 10 summarizes the study and lists the principal findings and recommendations.

3. Background

Much of the research on annoyance due to low frequency noise has been devoted to studying the response of subjects, either at home or at work, to long-term exposure to LFN from noise sources such as industrial plants, wind turbines, HVAC systems, high-speed trains, pile drivers, and loud music. A literature review revealed a shortage of information based on subjective tests of perception of realistic aviation noise signatures. A brief review of perception of low frequency noise, the findings of previous airport noise studies and objective indicators of low frequency noise are presented in this section.

3.1 Subjective Perception, Weighting Networks and Low Frequency Annoyance

Loudness has been measured subjectively, resulting in the equal loudness contours established for pure tones¹³ and bands of noise.¹⁴ The reference frequency is chosen as 1 kHz, and a loudness of 40 phons is represented by 40dB SPL at 1 kHz. Subjectively, there is 10 dB per doubling of loudness at 1kHz, so each increment of 10 phons corresponds to twice as loud. The phon contours are sloped to present sounds that are equally loud along each contour as a function of frequency. Above 100 Hz, an increase of 6-10 dB in sound pressure level is perceived as a doubling of loudness, depending on the slope of the curves at a given frequency. At the low frequencies the contours are very close together indicating that we experience a more rapid growth of loudness in the lower frequencies. Less sound energy is required to double our perception of loudness in the low-frequency region relative to the energy required to double our perception above 100 Hz. This phenomenon of our hearing system represents a reduced dynamic range in the low frequency region. This contributes to the potential for a low-frequency sound that is perceived as too loud to be closer to our threshold of detectability than would occur in the higher frequency region, increasing the potential for a rapid growth in annoyance with a minimal growth in loudness.

Research has shown that annoyance due to low frequencies increases rapidly with level. Subjective jury trials conducted with low frequency noise containing spectral peaks below 50 Hz, found that the signatures were judged as more annoying than loud even with an 11-12-phon decrease in loudness level. For low frequencies at low sensation levels loudness habituation occurred more rapidly than annoyance habituation, contributing to the difference between loudness and annoyance perception in this region. Longer noise exposure resulted in an increase in annoyance relative to loudness.¹⁵

The A-weighting curve is roughly the inverse of the 40-phon equal loudness contour, and was designed to mimic the human ear's response to sound of that loudness. The 40-phon curve represents the level of a tone that is necessary at each frequency to be equally as loud as a 1 kHz tone at 40 dB SPL. It was not designed to evaluate loudness significantly greater than 40 phon, and does not accurately characterize noise perception above that level. The A-weighting function is also not designed to evaluate noise that contains significant low-frequency content,¹⁶ as it has a sharp roll-off at low frequencies. Aircraft noise spectra can be heavily tilted towards low frequencies and may not be appropriately assessed by regulations based on the A-weighting network^{3,11, 12,17,18}

There are estimates that state that approximately 2.5% of the population may have a low frequency threshold that is at least 12dB more sensitive than the average threshold. This means that an increased sensitivity would affect nearly 1, 000,000 people in the EU-15 countries in the 50-59 year old age group, a group that generates many complaints.¹⁹

3.2 Findings of Previous Airport Studies of Low-Frequency Noise

Previous research efforts addressing noise, vibration and human perception of aviation noise, with specific relevance to low-frequency noise impact, have been conducted at several airports, including San Francisco International Airport (SFO), Baltimore-Washington International Airport (BWI), Boston Logan International Airport (BOS), Minneapolis-St. Paul International Airport (MSP) dating back to the mid-1980s, and most recently at Amsterdam Schiphol (AMS). A brief overview of some of these related investigations and their pertinent findings follow. Additional information can be found in a Wyle Laboratories report entitled "Status of Low-Frequency Noise Research and Mitigation."¹¹

San Francisco International Airport (SFO)

In 1986 and 1987, studies were conducted at San Francisco International Airport (SFO) concerning directivity patterns for low-frequency noise²⁰ and the differences in low-frequency noise exposure between backblast noise experienced by communities located behind aircraft taking off and aircraft overflight noise.²¹ The studies showed that communities at an angle of 40° to 50° from the jet exhaust axis experienced maximum low-frequency noise levels and that backblast noise had both more low frequency noise and longer duration than overflight noise. The findings also indicated that C-weighting worked best to describe low-frequency departure noise.

Another study²² conducted in 2000 reported field measurements of low-frequency noise, as well as spectral and temporal signatures of low-frequency backblast noise as compared to overflight and sideline noise. However, the main focus of the study was the effect of low-frequency noise on people. The report contains a summary of current knowledge on start of takeoff roll (SOTR) backblast noise, and its predominantly low-frequency spectral content, dealing with source noise characteristics, propagation of low-frequency noise, and mitigation efforts. It also discusses the effects of weather on low-frequency noise.

Baltimore Washington International Airport (BWI)

An analysis of start of takeoff roll (SOTR) noise was conducted at Baltimore Washington International Airport (BWI) over a period of 34 consecutive days in January-February 1990.¹⁰ A noise monitor was placed at a house 4000 feet behind and about 45° to the side of the start of runway 15R. Accordingly, the house was exposed to noise emissions from the rear of aircraft engaging in SOTR operations. The noise monitor reported hourly L_{eq} values, from which an overall L_{dn} value of 67.9 dB was calculated.

The data analysis indicated that there were three significant contributions to the overall L_{dn}

other than SOTR operations: 1) engine maintenance run-ups (59.8 dB), (2) non-airport background noise (55.3 dB), and 3) spurious instrumental readings (59.8 dB). When these levels were subtracted from the overall calculation, the remaining contribution from SOTR operations is 65.9 dB.

The results from the data obtained on site test were then compared to data from permanent noise monitor sites around BWI. The report suggests that models may not adequately predict impact at sites near SOTR operations. When compared with the data, it appeared that the model underestimated the noise from Stage 3 aircraft more than Stage 2 aircraft. The report notes that modeling ground operations is more challenging than modeling over-flight events due to the greater significance of difficult-to-model conditions such as foliage, barriers, wind, and temperature gradients.

In 1996 and 1997, two reports^{23,24} were published based on prior studies at BWI that dealt with modifications to existing houses to insulate them from low-frequency noise. Noise, measured in both dBA and dBC, was reduced significantly in several instances, although the cost to insulate each of the homes from low-frequency noise was in the \$40,000 to \$50,000 range; significantly higher than the cost for traditional sound insulation.

A low-frequency noise study measuring sound levels, vibration levels, and human assessment of the noise was conducted at BWI over a period of one week in August 1997.⁶ The study was conducted, monitoring the objective and subjective evaluation in a home located 3200 feet behind and to the side of Runway 28. The noise monitors recorded equivalent sound pressure level L_{eq} and sound exposure level L_E , both A- and C-weighted, inside and outside the house, for individual takeoff events. The vibration measurement was made on the wall of the house facing the runway. For subjective evaluation, the resident of the house was asked to rate how objectionable takeoff events were. The report concluded that human reaction does not depend solely upon the low frequency content of the event. The study found that C-weighted levels had a closer correspondence with both annoyance ratings and vibration levels than did A-weighted levels.

Boston Logan International Airport (BOS)

A two-pronged study, released in 1996, was undertaken at Boston Logan International Airport (BOS), examining the change in community noise levels after the switch from Stage 2 to Stage 3 aircraft as well as effectiveness of standard noise insulation at reducing low-frequency noise in addition to higher frequencies. The first part of the study, which is more applicable to the issue of low frequency noise, found a significant decrease in overall noise as well as a decrease specifically at frequencies below 100 Hz in areas that are normally affected by backblast and sideline noise.¹¹

Minneapolis-St. Paul International Airport (MSP)

In 1998 the Metropolitan Airports Commission and the City of Richfield commissioned an Expert Panel to examine current and future impacts of low-frequency aircraft noise in the communities surrounding Minneapolis-St. Paul International Airport (MSP). The scope of the project was broad, and in 2000 the Expert Panel released its findings in three volumes.¹ Included in the Expert Panel's report were, among other topics, sources of low frequency noise, studies of proper choice of a low-frequency noise descriptor, metrics, measurements of low-frequency noise reduction, and options for low-frequency noise mitigation.

The Expert Panel found that rattle-related annoyance was an effect of low-frequency aircraft noise for residents living within a mile of two different runways. In addition, they determined that noise from the reverse thrust during an aircraft's landing was an area needing more research.

While the 2000 Expert Panel report was a comprehensive study of low frequency airport noise, there is some disagreement over its conclusions, notably by the Federal Interagency Committee on Aviation Noise (FICAN).²

Amsterdam Schiphol (AMS)

The project at Amsterdam Schiphol Airport (AMS) included modeling, measurements, and recommendations regarding low-frequency noise from the Polderbaan runway.⁴ Proximity of the runway end to houses had resulted in complaints, and the authors were engaged to characterize the situation and recommend mitigation strategies. Community complaints exceeded both pre-construction predictions and numbers expected from noise-monitoring data (Phase I). Since the predictions and noise-monitoring data are based on A-weighted levels, the hypothesis was that structural vibration due to low-frequency noise was causing the complaints, and vibration measurements at houses in affected neighborhoods supported this (Phase II).

In Phase III of the project, parabolic equation modeling was used to predict noise levels for various atmospheric and ground conditions. The model was validated using acoustic measurements taken near the runway, at a house in the affected neighborhood, and at intermediate points. In addition, vibration data was obtained to determine if the ground was a vibration path. The conclusion was that vibration at the house was due exclusively to airborne noise. Complaints were not received when there was no wind, but modeling predicted a noise level increase of up to 10 dB for a northeast wind over a rigid ground surface. Thus attenuation of 10 dB would be desirable, with the frequency range around 31.5 Hz being of the greatest concern.

The proposed mitigation measures include barriers, ground absorption, modified operations, insulation of residences, and the more speculative strategies of active sound cancellation and wind generation. Barriers would need to be 10-15 meters high to provide a reduction of 6 dB, and barriers near the runway would affect aircraft safety. Modifying the ground cover with gravel beds or thick vegetation would likely provide the needed attenuation, although the necessary level of vegetation would take years to develop and the gravel bed approach is unproven on this scale. Insulation would be impractical at the frequencies concerned.

Airport operations could be modified such that the Polderbaan runway is not used by the noisiest aircraft, or under adverse wind conditions. This would require significant regulatory

changes, however, and the impact on communities near other runways would need to be evaluated. The authors mention the unproven strategies of active sound cancellation and wind generation, pointing out that the fans required to generate a favorable wind would themselves be a significant source of noise. The most feasible and effective options seem to be ground cover modification or airport operations modification.

3.3 Objective Indicators of Low Frequency Noise Annoyance

Complaints about aviation noise are primarily driven by human perception. In other words, when someone asks "Is there a low frequency noise problem?" they usually mean, are people annoyed by it? To answer this question, and to be able to predict and assess the impact of aviation noise, the objective metrics need to be related to the subjective perception of the impact by people.

Due to the attenuation of low frequencies by A-weighting filter networks, it is generally accepted that A-weighted sound pressure levels underestimate annoyance to low frequency noise. A considerable amount of effort has been devoted to finding reliable, objective metrics for predicting low-frequency noise annoyance. Leventhall provides an excellent review.²⁵

Much of the research has been devoted to studying the response of subjects, either at home or at work, to long-term exposure low-frequency noise from noise sources such as industrial plants, wind turbines, HVAC systems, high-speed trains, pile drivers, and loud music. (See Ref. 25 and the references therein.) Subjective test designs typically fall into one of three categories: 1) exposing subjects to continuous, steady levels of single tones or noise having various bandwidths and spectral shapes; 2) exposing subjects to modulated noise; or 3) exposing subjects to realistic recordings of low frequency noise sources. A number of factors have been found to affect subjective perception of low-frequency noise, including level, spectral shape, tonal content, and modulation rate.

One common conclusion across studies is that as the frequency decreases, the degree of annoyance increases more rapidly with sound pressure level. This phenomenon is referred to as narrowing of dynamic range. Equal loudness contours are used to relate objective sound pressure levels to subjective perception of loudness. These contours are compressed at low frequencies, meaning that the difference in sound pressure level of a low-frequency signal that is just audible and one that is considered annoying is much less than at higher frequencies. In other words, a low-frequency signal can go from being audible to annoying with a relatively small change in level.

Regardless of the nature of test signals, two themes are common to the conclusions. The first is the need to measure low-frequency noise indoors. Correlating indoor response to an outdoor measure of low-frequency noise is problematic because of variability in low-frequency transmission loss through structures, and the possibility of exciting low-frequency resonances in structures and rooms that lead to levels of low-frequency noise larger than might be expected based on outdoor levels.

Poulsen²⁶ reviews a number of acceptability criteria that have been proposed for LFN in Europe. The solid black line in Figure 1 shows an aggregate criteria curve developed from data in that article. Exactly how this type of curve is used varies from country to country. Generally

speaking, if one or more levels in a 1/3-octave band exceed the criteria curve, low frequency noise is considered to be a nuisance. Below 40 Hz, the criteria curve roughly approximates the threshold for audibility. The inference is that at very low frequencies, noise that is audible is unacceptable. This statement does not say that the noise is annoying, however, narrowing of dynamic range at low frequencies does suggest that the threshold of annoyance would track the criterion curve, offset above it. The question is by how much?

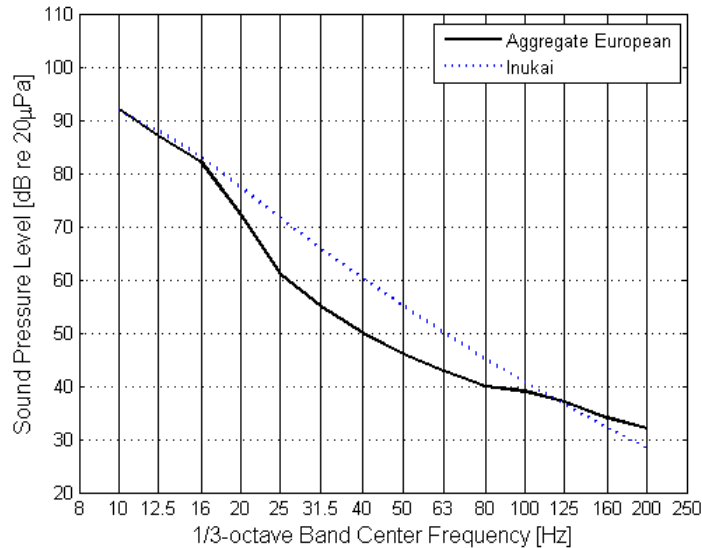


Figure 1. Graph of sound pressure level in dB re 20µPa versus 1/3-octave band center frequency showing criteria for annoyance due to Low Frequency Noise.

One class of low-frequency noise metric involves an A-weighted, low frequency sound pressure level. Investigators recognize that A-weighting does a good job of representing subjective perception of noise. However, the A-weighted level of a broadband signal is dominated by the levels in the mid to high frequencies. As a result, a standard A-weighted level does not capture annoyance due to low-frequency noise. The Danish method of assessing annoyance due to low-frequency noise is to apply the A-weighting filter corrections to 1/3 octave bands from 10 to 160 Hz. The energy sum over these low frequency bands gives the $L_{A,LF}$. Recommended limits of $L_{A,LF}$ have been established. For example, the recommended limit for a dwelling during the day is 25 dB $L_{A,LF}$, averaged over 10 minutes.

Poulsen also conducted a listening test to determine the best method of assessing annoyance due to low-frequency noise. The signatures were realistic recordings of noises containing low frequency noise, including highway traffic noise, drop forge, gas turbine, fast ferry, distant steel factory, generator, air compressor, and music transmitted through a building. Results indicated that the Danish method described above gives the best objective metric of subjective response.

Inukai, et al²⁷ also conducted subjective tests to determine acceptability limits for ordinary adults. This limit is shown by the blue dotted line in Figure 1, and is in reasonable agreement

with the aggregate criteria curve.

Because of the limited frequency band used in calculating $L_{A,LF}$, it cannot be measured with a standard sound level meter. These criteria are for long-term exposure and measurements are averaged over the period of minutes. It is not clear how applicable these limits are to aviation noise.

Other low-frequency noise metrics that focus on the noise levels in particular 1/3-octave bands have also been proposed. One of these is the Low-Frequency Level (L_{LF}) described in ANSI Standard S12.9, Part 4⁹ and referred to by FICAN.² L_{LF} is the level calculated from the time-average, mean-squared pressure in the 16, 31.5, and 63-Hz octave bands. The MSP Expert Panel was critical of using L_{LF} , because it had not previously been used to characterize aircraft noise and its reliance on data in the 16 Hz band. Aircraft certification data does not extend down to 16 Hz, and so L_{LF} cannot be calculated with the Integrated Noise Model (INM).

The MSP Expert Panel used the Low-Frequency Sound Level (L_{FSL}). L_{FSL} is defined as "a single-event metric that sums the maximum one-third octave band sound levels from 25 to 80 Hz, inclusive, that occur during the course of an individual aircraft passby."²⁸ They also proposed a dose based on L_{FSL} . FICAN was critical of the use of L_{FSL} primarily because it uses one-third octave band data that may be measured at different times during an event. They considered it to be constructed artificially.²

Work on the subjective perception of low frequency noise by Tokita and Nakamura in the early 1980's has been applied to airport noise. They conducted a series of studies to address the question of how and why infrasound and low frequency noise affects humans differently than other noises.^{29,30} Two similar experiments were conducted in the low and high frequency range, investigating perception above and below 100 Hz. The subjective responses were catalogued and the researchers were able to identify thresholds where the response turned from "No" to "Yes" and from "Yes" to "Very Much." These thresholds were identified for the subject's response to the noise when prompted about their ability to detect a sound, feel annoyance, feel displeasure, have an oppressive feeling, or feel a vibration. Responses that are peculiar to the low frequency stimuli are feelings of oppression and vibration. The usual methods of quantifying noise discount the low-frequency and infra sound bands and so do not properly evaluate the impact due to low frequency noise^{29,30}.

Two versions of their thresholds are shown in Figure 2. The left-hand plot is a redrawn version of a graph presented in the MSP Expert Panel report.¹ It presents Tokita and Nakamura's thresholds (i.e. sound pressure levels) at which subjective response to direct exposure to low frequency sounds changes as functions of frequency from one characterization to another. (It should be noted that the threshold of audibility is not shown on the graph.) The right-hand plot is a similar graph redrawn from a Wyle Labs report on a Low-Frequency Noise study conducted at Amsterdam's Schiphol airport.⁴ The thresholds in this figure are those of Tokita and Nakamura for direct exposure, compensated for a nominal transmission loss model for a house from a 1982 NASA report by Stephens, Shepherd, Hubbard, and Grosveld.³¹ The right-hand plot, the "Schiphol plot," represents the nominal levels of outdoor sound that would elicit a particular response from occupants inside a house. An important feature of these figures is the downward-pointing wedge-like region characterized by oppressive/chest vibration response. In this frequency range there is relative little difference between the levels at which a

person detects the noise and at which that noise is perceived as being oppressive. This illustrates narrowing of dynamic range mentioned previously.

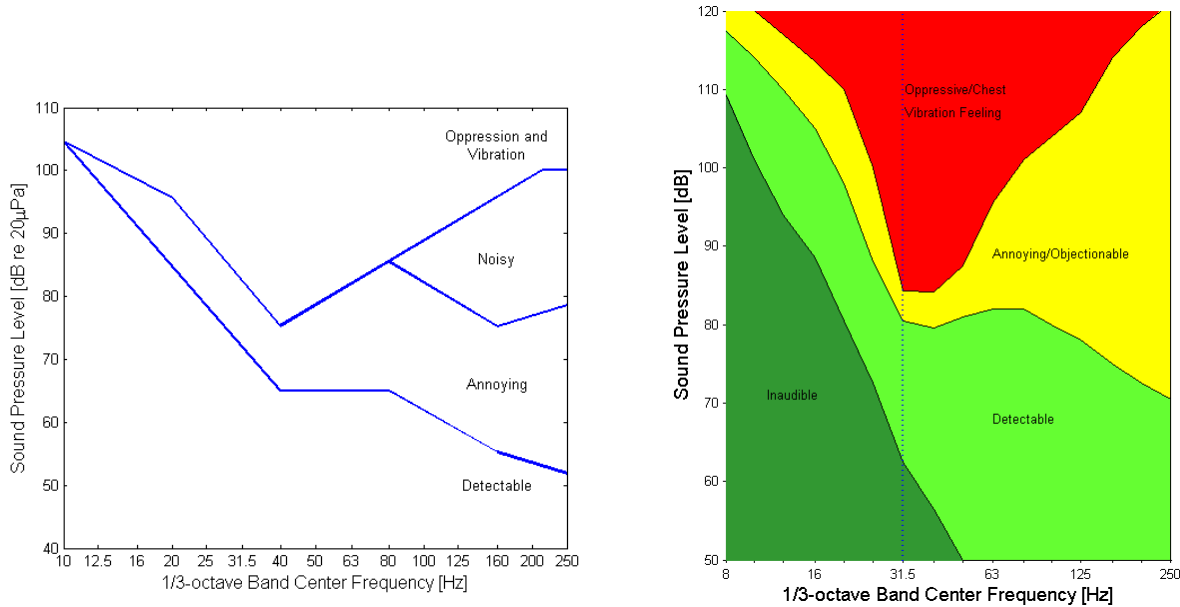


Figure 2. Two different versions of Tokita and Nakamura's threshold for the perception of low-frequency noise. The left-hand plot is for direct exposure. The right-hand plot is for nominal levels of outdoor sound that would elicit a particular response from occupants inside a house.

Figures 3 and 4 illustrate the utility of the criteria curves or thresholds in assessing the potential for annoyance due to low-frequency noise. Figure 3 shows the results of a parabolic-equation model of sound propagation in the vicinity of Schiphol from Ref. 4. The colors represent values of sound pressure levels. The figure on the left is a no-wind condition, the one on the right a wind blowing from the northeast. One of the test houses at Schiphol was located approximately 2 km to the southwest of the runway. Comparing the two figures, the northeast wind results in an increased sound level at the location of the house. The reason for this is that sound is refracted downward in the downwind direction.

Figure 4 is the "Schiphol plot" described previously with data recorded outside the houses as described in the Schiphol report.⁴ The solid and dashed lines represent the highest 1/3-octave band sound pressure levels recorded outside the house on two different days. The solid line corresponds to a day when the wind was not from the northeast. The dashed line represents the levels on a day with a northeast wind. As expected, a northeast wind results in higher levels. On this day, the levels exceed the annoying threshold in some of the 1/3 octave bands. There is also anecdotal evidence from Schiphol of increased complaints on the "northeast wind day." The results of Figure 4 were used in two ways. First, they provided anecdotal support for the validity of the thresholds. Second, they provided an indication of how much mitigation would be necessary to reduce annoyance.

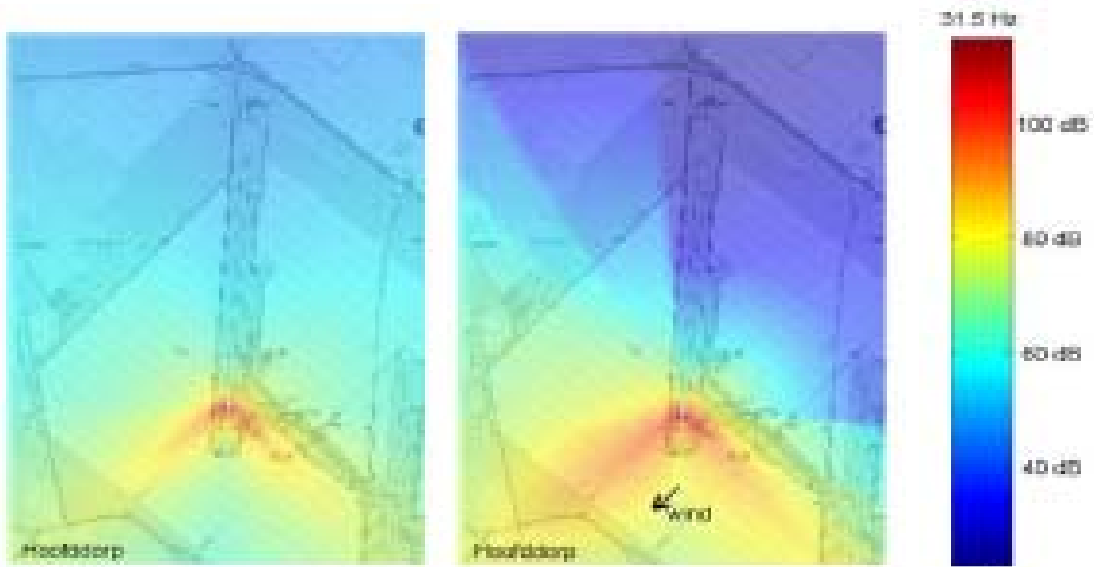


Figure 3. Results of a parabolic equation model of low-frequency sound propagation showing the influence of a northeast wind on the sound levels to the southwest of the runway. (From Ref. 4.)

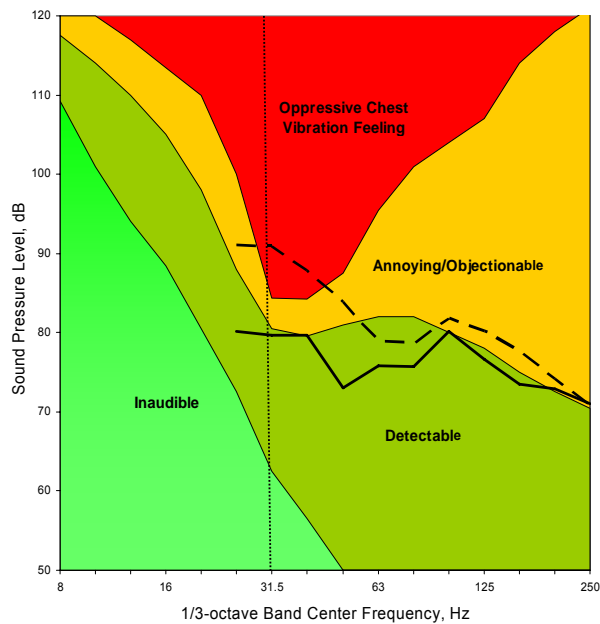


Figure 4. The Tokita and Nakamura's thresholds from along with data recorded outside the houses as described in Ref. 4. The solid and dashed lines represent the highest 1/3-octave band sound pressure levels recorded outside the house on two different days. The solid line corresponds to a day when the wind was not from the northeast. The dashed line represents the levels on a day with a northeast wind. (From Ref. 4.)

4. Addressing FICAN Findings on MSP Expert Panel Report

The City of Richfield and the Metropolitan Airports Commission established a Low-Frequency Noise Expert Panel in 1998 to study existing and potential impacts of low-frequency aircraft noise in communities around Minneapolis St. Paul International Airport (MSP). The MSP Expert Panel issued their findings in 2000 in a three volume report.¹ After reviewing the report and meeting with a subset of the MSP Expert Panel in 2001, the Federal Interagency Commission on Aviation Noise (FICAN) issued a report in 2002 containing its response to the MSP Expert Panel's findings and recommendations for future research in low-frequency noise around airports.²

FICAN recommended that further research consider the following:

1) That measurements be conducted in houses within critical distances from runways identified in previous studies of low-frequency aircraft noise, in particular one conducted at Baltimore-Washington International Airport (BWI).⁶ Measurements should include exterior noise and window, wall, and floor vibration with a frequency range extending down to a few hertz to capture the low-frequency impact. The vibration measurements should be based on the recommendations by the American National Standard Institute (ANSI) Standard S3.29-1983 (R1996).⁷ In addition, the measured noise and vibration levels should be compared to thresholds for tactile perception of vibration, known as the "Hubbard criteria,"⁸ used to establish the extent of the effect of low-frequency noise at BWI.

2) Have panels of subjects rate the annoyance of individual aircraft events in the houses. Conduct statistical analysis to establish what combination of physical measures gave the best prediction of annoyance ratings. Assess the ANSI Standard [S12.9, Part 4]⁹ Low-Frequency Level (L_{LF}) as a descriptor of low-frequency noise.

3) Study the efficacy of sound insulation in a stepwise fashion, beginning with the most rattle-prone features of houses, the windows and doors. FICAN's idea was to use the same subjects as in Recommendation 2 to assess the impact of insulation.

In addition to these recommendations, FICAN provided responses to the MSP Expert Panel findings in four major areas, namely the effects of low frequency aircraft noise, identifying an appropriate descriptor for LFN and LFN dose, investigating the relationship between LFN and annoyance, and defining an acceptability criteria for LFN. Specific actions were incorporated in the PARTNER low-frequency noise study to address both the FICAN recommendations and responses.

Assessment of the Effects of Low-Frequency Noise

The Low-Frequency Noise Study conducted research in houses located within critical distances (3000 ft) from runways, as identified in the BWI⁶ and MSP Expert Panel reports.¹ Measurements included interior and exterior noise levels down to 8 Hz, and vibration measurements on window, wall and floor as recommended. The experimental design included root mean square (rms) acceleration measurements of structural vibration, as well as measurements of audible rattle. The locations of the accelerometers were in accordance with FICAN's suggestion to measure acceleration as suggested in Hubbard⁸ and ANSI S3-29-1983 (R

1966).⁷ The acceleration data obtained was evaluated and compared to values given in Hubbard and ANSI S3-29(1983) as well to outdoor metrics, to address FICAN's suggestion to examine measurements against the Hubbard Criteria as cited in the BWI study.⁶ The measurement of rms acceleration provided a measure of primary structural vibration, as well as rattle due to secondary vibrations.

Identify a Descriptor for Low-Frequency Noise and Low-Frequency Noise Dose

A time-averaged dose descriptor such as day-night average noise level (L_{dn}), is standard for aviation noise impact. FICAN questioned the validity of the MSP Expert Panel's use of the Low-Frequency Sound Level (L_{FSL}) as a descriptor for low-frequency noise and low-frequency noise dose. L_{FSL} is a low frequency single-event noise metric. FICAN also recommended that the Low-Frequency level (L_{LF}) defined in ANSI S12.9, Part 4, be considered as a potential metric. L_{LF} was not specifically designed for aviation noise, but was intended to apply to all sources of low-frequency energy, and included sound pressure levels as low as the 16 Hz one-third octave band. To address these suggestions, this PARTNER study obtained noise measurements that included interior and exterior noise levels down to 8 Hz. The time series recordings were made for numerous single noise events, and combined events (events on multiple runways simultaneously). The recordings were used for subjective jury evaluations and metrics were calculated to identify the metrics that best correlated with subjective perception of annoyance. Assessment of a Low-Frequency Noise dose was not addressed in this study, however, low-frequency single-event metrics were assessed.

Evaluate the Relationship between Low-Frequency Noise and Annoyance

Three subjective jury evaluations were conducted to assess the relationship between low-frequency noise and annoyance. The trials assessed the influence of the spectral balance of the aviation noise signature, the presence of rattle in aviation noise and applicability of the Tokita and Nakamura thresholds to low-frequency aircraft noise.^{29,30} These thresholds have been useful in other low-frequency noise studies.⁴ A statistical analysis of the subjective response data was conducted to establish what combination of physical measures gave the best prediction of annoyance judgments. Metrics were correlated with the results of subjective jury evaluations conducted in an aircraft noise simulator.

Investigate the Acceptability Criteria for Low-Frequency Noise

A dosage-response relationship for noise exposure and high annoyance was recommended in 1992 by The Federal Interagency Committee On Noise FICON (as cited by FICAN) identifying the 65 dB DNL contour as representing 12.3% Highly Annoyed (HA) and the 75 dB DNL as representing 36.5% HA. Based on a social survey conducted at MSP, the Expert Panel recommended the use of L_{FSL} , a single event metric to determine the annoyance dose response. The Expert Panel recommended that a 70dB L_{FSL} represented the 12.3 % HA and that 87 dB represented the 36.5 % HA. Assessment of a Low-Frequency Noise dose was not addressed in this study. However, jury trials were conducted to assess metrics that best correlated with subjective perception of annoyance. Investigators also conducted a laboratory investigation of rattle-prone features of house structure elements (windows and doors) and a low-frequency transmission loss study to address FICAN's recommendation for a sound insulation study.

5. Low-Frequency Noise Field Study Design

5.1 Airport Selection Process

The investigators identified a set of criteria that guided the selection of an airport for the field study. These criteria were:

1. The airport should have a representative commercial aircraft fleet mix.
2. The airport should permit measurement of source noise along runways including side-line noise on start of takeoff roll and acceleration, and thrust reverser noise.
3. The local terrain should provide a direct sound propagation path from runway to homes with no large structures between runway and community.
4. Houses should be located within 3000-4000 feet of runways based on the MSP and the low-frequency study at BWI studies and allow both exterior noise and interior noise and vibration measurements.
5. Houses should be of various age and construction type.
6. There should be minimal interference from other noise sources such as highways, trains, construction etc.
7. The airport should have no previous history of a low-frequency noise problem to reduce the potential for bias in the surrounding community.

Table 1 provides a list of candidate airports having houses within 4000 feet of the runways considered by the investigator. No single airport satisfied all six criteria. One of the most difficult conflicts to resolve was that airports with no history of low-frequency noise problems were reluctant to cooperate out of concern that the study could create a problem where none had existed. The investigators judged the best compromise to conducting the field study at an airport that met the necessary physical attributes to permit implementation of as many of the FICAN recommendation as possible. The most significant compromise was that the in-residence subjective evaluation component had to be abandoned and laboratory-based subjective evaluations substituted in their place.

Washington Dulles International Airport (IAD) was found to meet the majority of the criteria. The airport had a well distributed fleet mix and a surrounding terrain that did not vary in elevation greatly. The area surrounding the runways allowed measurements of source noise along the sideline, and did not contain any structures that would influence the path of noise propagation. In addition to these factors, the airport has two residential structures within 3000-4000 feet of the runways. The Metropolitan Washington Airports Authority agreed to host the low-frequency noise study under the condition that the research team was not permitted to interview community members during the time period that the field measurements were conducted.

The field measurements were conducted October 5-22, 2004 to coincide with study of noise levels and directivity behind aircraft during start of takeoff roll by NASA and Volpe National Transportation Systems Center. In preparation for their study, the NASA/Volpe investigators conducted an extensive search of airports using selection criteria that were compatible in a number of respects with those of the PARTNER study. Their selection of IAD served as confirmation that IAD was appropriate for the low-frequency noise study.

Table 1: List of candidate airports having houses within 4000 feet of runways.

Distance from homes and measurement sites to runways					Other Considerations
BWI (3 sites) <i>Baltimore-Washington International</i>	3200 feet* (site 7)	4600 feet (site 6)	7800 feet (site 3)	Note: all sites were to the East or Northeast of runways (distance is measured from the end of Runway 28)	Train Tracks to the West
MSP (mult. sites) <i>Minneapolis-St. Paul International</i>	Measurements were made from 2600 feet to 5000 feet			Note: all sites were to the North, Northwest, or Northeast of runways	Waterway to the East
ATL <i>William B Hartsfield International</i>	Neighborhoods begin 2280 feet away in the North				Highways to the East, West and South
CLT <i>Charlotte Douglas International</i>	Neighborhoods begin 3630 feet away to the North, 2310 feet away to the West, and 2310 feet away to the East				Highway to the East
CVG <i>Cincinnati/Northern Kentucky International</i>	Neighborhoods begin 3300 feet away to the West, 1650 feet away to the Northeast, and 2970 feet away to the Southeast				Highways to the North, and West
IAD <i>Washington Dulles International</i>	Neighborhoods begin 3700 feet away to the East (Measured from sideline of 19L/1R)				Highway to the Northeast
IND <i>Indianapolis International</i>	Neighborhoods begin 1980 feet away to the Northwest, and 2700 feet away to the Southwest				Highways to the East and South FEDEX HUB
MCO <i>Orlando International</i>	Neighborhoods begin 3630 feet away to the North, and 4620 feet away to the East				Highway to the North Lakes to the North and South
ONT <i>Ontario International</i>	Neighborhoods begin 2640 feet away to the North, and 1650 feet away to the West				Highways to the North and Southwest UPS HUB
PDX <i>Portland International</i>	Neighborhoods begin 4620 feet away to the Southwest				Waterway to the North
PHL <i>Philadelphia International</i>	Neighborhoods begin 3960 feet away to the North, and 2475 feet away to the West				Highway to the North Waterway to the South UPS HUB
SDF <i>Louisville International- Standiford Field</i>	Neighborhoods begin 1500 feet away to the East, 3600 feet away to the West, 2000 feet away to the North, and 4000 feet away to the Southwest				Highways to the North and East UPS HUB
SEA <i>Seattle-Tacoma International</i>	Neighborhoods begin 2500 feet away to the North, and 2640 feet away to the West				Highway to the North

*Vibration measurements were completed at BWI (site 7, 3200 feet from runway)

BWI and MSP were used as reference sites for the other candidate sites

The NASA/Volpe team provided significant logistical support to the PARTNER investigators both prior to and during the field measurements.

The IAD field study included measurements of source noise along side Runways 19R, 30, and 1R, and noise impact at two unoccupied residential structures, a brick-facade house and a stone-facade house located on airport property within approximately 3000 ft of runways. The measurements at the houses consisted of outdoor and indoor noise as well as measurements of wall, window and floor vibration. Figure 5 shows an overhead photograph of IAD, indicating the runway layout, locations of the residential structures, and their proximities to the runways.



Figure 5. Overhead photograph of Washington Dulles International Airport (IAD) showing the runway layout, locations of the residential structures and their proximity to the runways. The arrows indicate the predominant direction of runway traffic during the measurements. (Image courtesy of the USGS.)

5.2 Source Noise Measurements

Microphone arrays were set up along three runways to record sideline noise during start of takeoff roll (SOTR) and acceleration down the runway on Runways 30 and 1R, and sideline noise during thrust reverser (TR) deployment during landing on Runways 19R and 1R. For the three weeks of the study, the wind generally dictated Runway 30 would be used for takeoffs and 19R and 1R would be used for landings. Air traffic patterns during these three weeks followed a general trend. In the early morning hours a large number of aircraft departed from Runway 30 while the two other runways were used primarily for landings. These early departures and arrivals consisted of both commercial and general aviation aircraft. In the mid-morning hours there was a lull in the traffic. In the early afternoon to evening hours there was an increase in both arrivals and departures, with the arrivals including the larger aircraft used for international flights.

The layout of the microphone arrays was determined based on the length of the runway, available equipment, and the purpose of the measurements for a particular runway. Runway 19R was used primarily to record thrust reverser signatures. The goal was to capture the full deployment of the thrust reverser. After observing representative landing operations, investigators determined that a 3000 ft (914 m) long array centered on the 3500 ft (1066 m) mark from the landing threshold was sufficient to capture the thrust reverser event, considering the variability in the duration and where the thrust reverser would be engaged.

Figure 6 illustrates the instrumentation setup to measure thrust reverser noise along Runway 19R. Microphones (prepolarized G.R.A.S. 1/2 inch 40AE microphones connected to G.R.A.S. 1/2 inch ICP 26CA preamplifiers with 90 mm diameter open-cell foam windscreens) were mounted on tripods 200 ft from the runway centerline and 4 ft. (1.2 m) above the ground. (The only exception was Microphone 5 which was mounted at 15 ft. (4.6 m)). Microphones 4 and 5 were co-located at the array center, at the 3500 ft point of the runway. The other microphones were spaced 500 ft apart. Distances were measured with a laser range finder. The ± 1 dB bandwidth of the microphones is 3.15 Hz - 20 kHz. The ± 0.2 dB bandwidth of the preamplifiers is 2 Hz-100 kHz.

The signal from each microphone was digitized at 50 k samples/sec per channel and recorded on a 16-bit TEAC GX-1 data recorder. The recorder parameters were set using a laptop computer running TEAC GX software. The digital recorder was connected to each microphone/pre-amp unit using RG-62 coaxial cable. At the beginning and end of each measurement session an end-to-end (i.e. from microphone to recorder) calibration was performed for each microphone channel using a G.R.A.S. model 42AB sound calibrator, which generates a 1 kHz tone at 114 dB re 20 μ Pa.

Observers stationed at the digital recorder identified each landing event by aircraft type, airline, tail number, and time of the event. Landing events were also recorded with a video camera, set up 509 ft (155 m) from the center line of the runway and 3440 ft (1048 m) from the landing threshold, to visually track the aircraft down the runway. The video recording was used to identify tail numbers and where along the runway the thrust reverser was deployed and stowed.

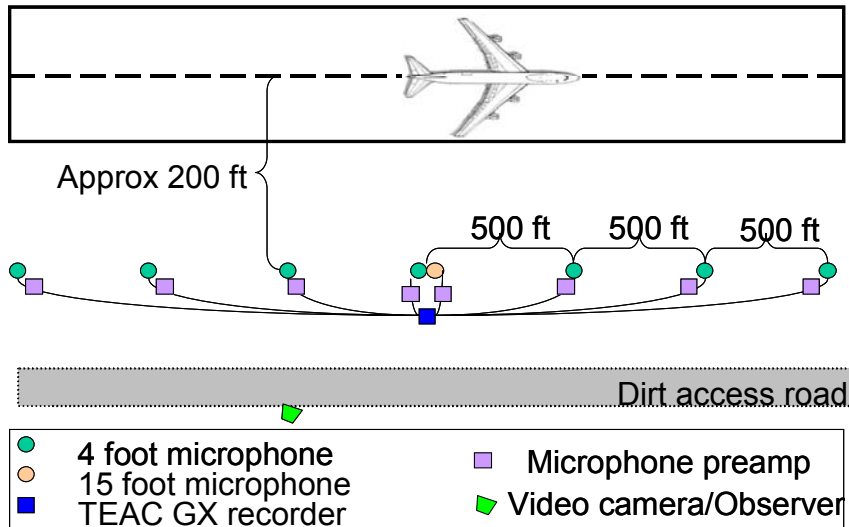


Figure 6. Diagram of instrumentation setup to measure thrust reverser noise along Runway 19R.

During the period of the field measurements Runway 30 was used for takeoff. A microphone array similar to that used along Runway 19R was set up along the sideline of Runway 30 to capture start-of-takeoff-roll signatures. The microphone locations are indicated in Table 3. Two different configurations were used. The first ("original") array spanned a total distance of 1500 ft (457 m). Airport construction and standing water during the time of measurements limited this distance. The microphones were located 330 ft (100 m) from the center line of the runway with the center of the array 1000 ft (305 m) from the start of the runway. Subsequent microphones were spaced 250 ft (76 m) apart. As with the setup along Runway 19R, the microphones were mounted 4 ft (1.2 m) above the ground, with the exception of Microphone 5, which was 15 ft (4.6 m) above the ground.

As the study progressed, Investigators later determined that to monitor and record on two runways simultaneously, they would have to reduce the array size. The locations of the microphones for the "abbreviated" array are indicated in the bottom row of Table 3.

Table 2. Positions of microphones from the start of the runway for Runway 19R.

Mic 1	Mic 2	Mic 3	Mic 4/5	Mic 6	Mic 7	Mic 8
All values are distance from landing threshold in feet (meters).						
2000 (610)	2500 (762)	3000 (914)	3500 (1066)	4000 (1218)	4500 (1370)	5000 (1524)

Table 3. Positions of microphones from the start of the runway for Runway 30. The microphones were located 330 ft (100 m) from the center line of the runway.

	Mic 1	Mic 2	Mic 3	Mic 4	Mic 5	Mic 6	Mic 7	Mic 8
Array	All values are distance from start of runway in feet (meters)							
Original	250 (76)	500 (152)	750 (229)	1000 (305)	1000 (305)	1250 (381)	1500 (457)	1750 (533)
Abbreviated	250 (76)	625 (190)		1000 (305)	1000 (305)	1375 (419)	1750 (533)	

Runway 1R was not part of the original experimental design. However, investigators observed during the first week of the field study that the traffic pattern on this runway would provide valuable data and so the design was changed to include it. This runway was the only one of the three that serviced both departures and arrivals at the same time. The microphone locations are indicated in Table 4. The start of the array was closer to the landing threshold than on 19R so that it could capture the start-of-takeoff-roll noise and it also covered a span of 1143 meters so as to be sure to capture most of the thrust reverser noise. This array had to be abbreviated later on in the study to facilitate recording on multiple runways. The microphones were placed at a distance of 375 ft (114 m) from the center line.

Table 4. Positions of microphones from the start of the runway for Runway 1R.

	Mic 1	Mic 2	Mic 3	Mic 4	Mic 5	Mic 6	Mic 7
Setup	All values are distance from start of runway in feet (meters)						
Original	1500 (457)	2250 (686)	3000 (914)	3000 (914)	3750 (1143)	4500 (1370)	5250 (1600)
Abbreviated			3000 (914)		3750 (1143)	4500 (1370)	5250 (1600)

5.3 Measurements of Low-Frequency Noise Impact Residential Structures

In addition to noise measurements alongside runways, noise and vibration measurements were made at two abandoned residential structures located on airport property. The first structure was a single story framed brick facade house with a basement located approximately 2800 ft (850 m) to the west of Runway 19R. This structure is labeled Brick House in the upper left portion of Figure 5. The kitchen, located in the southeast corner of this house and having one wall roughly parallel to Runway 19R, was chosen as the room to be instrumented for the indoor part of the study. An overhead view of the Brick House is shown in Figure 7, indicating the kitchen, garage and approximate location of the outdoor microphone. The exterior of the Brick House is shown in Figure 8.

The second structure, labeled Stone House in the lower right portion of Figure 5, was a stone facade, single story dwelling with a basement, located approximately 2700 ft (820 m) from runway 1R and 3100 ft (960 m) from Runway 30. A bedroom located in the northwest corner of



Figure 7. Overhead photograph of the Brick House indicating the kitchen, garage, and approximate location of the outdoor microphone. Runway 19R is approximately 2800 ft (850 m) to the right. (North is up in this photograph.) (Image courtesy of the USGS.)



Figure 8. Northeast view of the exterior of the brick house.



Figure 9. Overhead photograph of the Stone House indicating the instrumented bedroom and approximate location of the two outdoor microphones. Runway 1R is approximately 2700 ft (820 m) to the right, Runway 30 is 3100 ft (960 m) to the left. North is up in this photograph. (The two small structures that appear in this image to the right of the Stone House were not present during the field study.) (Image courtesy of the USGS.)



Figure 10. Exterior view of Stone House looking from the east. Runway 30 is approximately 3100 ft (960 m) to the west (towards the background). One of the outdoor microphones is shown in the foreground.

the house was instrumented. As shown in Figure 9, this room has one wall facing the departure end of Runway 30. The other wall was roughly perpendicular to Runway 1R. Although stone houses are not typical construction, the exterior noise and window vibration measurements provided useful information.

The instrumentation setup at each house was similar. An outdoor microphone placed at a height of 4 ft (1.2 m), approximately 50 ft (15 m) from the corner of the house in the direction of the runway recorded exterior noise. One outdoor microphone was used at the Brick House (see Figure 7), two were used at the Stone House (see Figures 9 and 10). The microphone/preamplifier/windscreen types and calibration procedure were the same as that used along the runways.

Figures 11 and 12 show the instrumentation setup in each house. The test rooms were similar in size, approximately 10 ft × 15 ft (3 m × 5 m), and instrumented with an indoor microphone, a HEAD binaural analysis system, and a sound level meter to record the indoor noise levels, and with accelerometers to record wall, window and floor vibrations. The microphone and HEAD system were placed in the room at a location to capture noise events similar to what someone sitting in the room might experience. The accelerometer locations were chosen based on previous work by Hubbard,⁸ Mayes, et al,³² Schomer and Sias,³³ and ANSI Standard S3.29-1983.⁷ The accelerometer locations in the Brick and Stone Houses were slightly different. In the Brick House, accelerometers were mounted on the east and south walls and on windows located in the east and south walls. In the Stone House, one accelerometer was mounted on the west window, one on the north wall, and one each on two different windows in the north wall. One window was securely mounted. The other window rattled during certain noise events. An accelerometer was placed on this window as well as on the window next to it, to compare the vibration characteristics of a rattling versus non-rattling window. All of the wall and window accelerometers measured vibrations perpendicular to the wall or window. A fifth accelerometer was used in each house to monitor vertical vibrations of the floor.

The signal paths are shown schematically in Figure 13. The accelerometers were Vibrametrics Model 1000A with a nominal sensitivity of approximately 10 mV/g. Each was powered through a PCB signal conditioning box model 480E09 with the gain set to 100. These accelerometers were mounted using wax. The floor in the Brick House was covered with linoleum tile, while that in the Stone House was covered with thin carpet. The floor accelerometers were mounted to a nail driven through the floor covering into the subfloor.

Due to its location in the relatively undeveloped area to the west of Runway 19R, the Brick House was only slightly impacted by events other than aircraft operations on the runway. The Brick House had one outdoor microphone placed approximately 50 ft (15 m) from the southeast corner of the house closest to Runway 19R. The indoor microphone was positioned about 3 ft (1 m) from the noise-impacted east wall. The dominant outdoor background noise when aircraft were not present tended to be the sound of crickets in the fields surrounding the house. For the purpose of this study, the Brick House accelerometer response measured perpendicular to the east wall and window (parallel to Runway 19R) was defined to be the X-axis, while the responses perpendicular to the south wall and window were identified as the Y-axis.

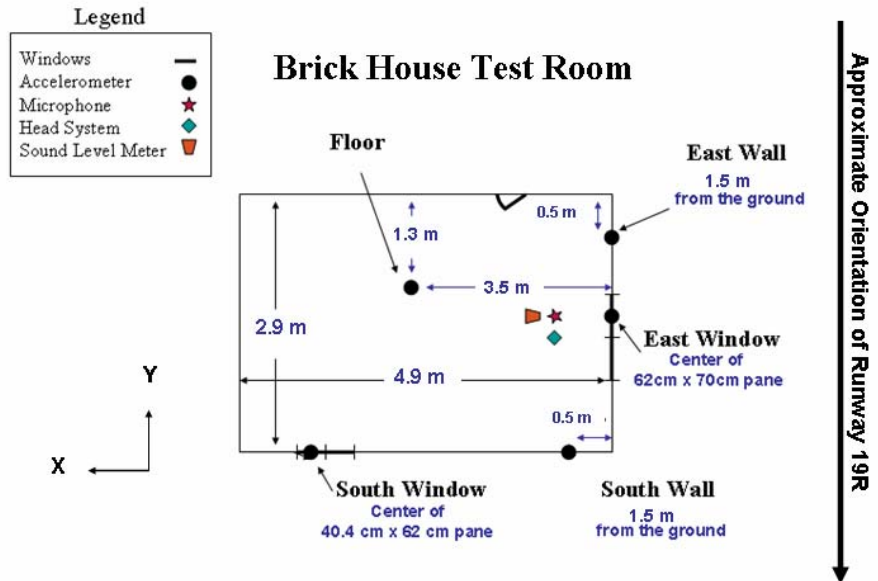


Figure 11. Diagram of instrumented room in the Brick House. (Not to scale.)

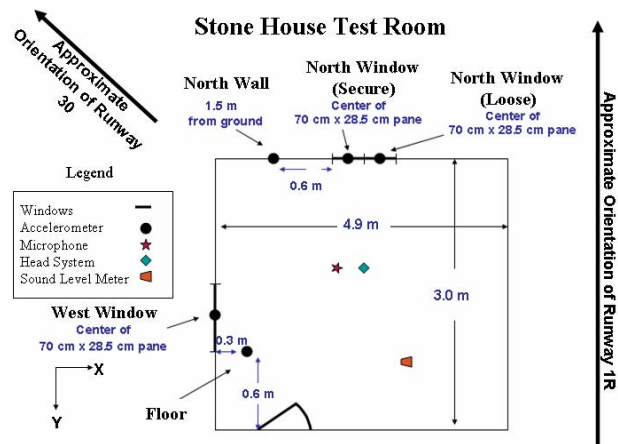


Figure 12. Diagram of instrumented room in the Stone House. (Not to scale.)

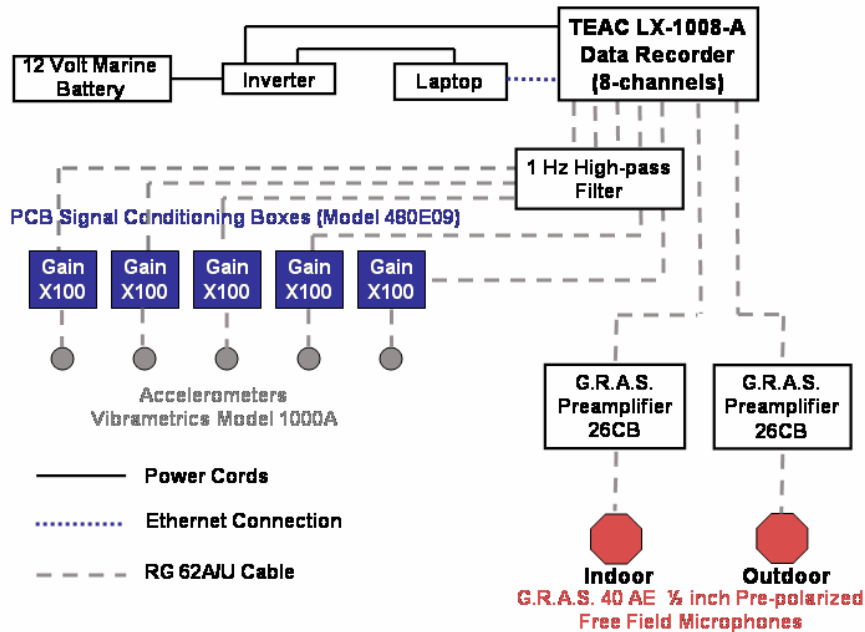


Figure 13. Schematic diagram of signal paths used at the two residential structures.

The Stone House, located south of the passenger terminals and between Runways 30 and 1R, was impacted by more airport noise events due to its more central location than the isolated Brick House. In addition to aircraft taking off and landing on runways, taxiways to and from these runways were near enough that noise from planes idling and throttling up engines also impacted the house. Onsite observations at the Stone House indicated that there was also large impact from thrust reverser and sideline acceleration due to operations on 1R. Outdoor microphones were positioned approximately 50 ft (15 m) from the two corners of the stone house closest to each of the runways. This allowed the opportunity to study the effects of individual events, and combined events due to events occurring simultaneously on both runways. Often, two noise events would overlap at the Stone House, increasing the impact and making it difficult to pinpoint the source. At the Stone House, the indoor microphone was located about 3 ft (1 m) from the north wall. The accelerometer on the west window measured vibrations in what was defined to be the X direction, and measurements made on the north wall defined the Y-axis.

In an effort to record rattle similar to what residents might experience, investigators installed shelving with cups, dishes, etc. on the north and west walls of the Stone House. However, these items did not rattle. In an attempt to artificially generate audible rattle, a plate was hung from a string and rested slightly against the center of windows in both houses. This rattle augmentation was fashioned after one used previously by Schomer and Averbuch.³⁴

All the data was recorded as time series data using an 8 channel, 24 bit TEAC LX-10 recorder at 48 ksamples/second per channel. The accelerometer data was passed through a 1 Hz high-pass filter to attenuate some very low frequency interference that was observed, and believed to be the result of construction that was being conducted at IAD during the time of the field study.

The data was recorded using two methods to allow for the calculation of metrics for single events or metrics assessing the noise impact averaged over an hour. Recording for single events was initiated by observers in the house who were notified of an approaching aircraft by observers at the runway. Recording stopped after the aircraft could no longer be heard at the house. The second method was a continuous recording that involved starting the recording and having the observers leave the premises for periods of one to two hours. This method limited noise contamination from the observers, and provided a means to conduct hour-long averages of overall impact at a given location. Recording continuously over a 24-hour period was not feasible; the hour-long averages provided the opportunity to compare the noise impact in hour-long increments throughout the day.

5.4 Meteorological Data

Noise levels in the vicinity of airports are affected by many environmental factors, including atmospheric attenuation of sound, temperature and how it changes with height, wind speed and direction, and how they change with height, and atmospheric turbulence. Ground effects (e.g., plowed field, vegetation, snow covered field, rain soaked field, desert, etc.) and irregular terrain also affect sound propagation significantly. Ground and terrain effects were minimized to the extent possible by site selection. Investigators measured relevant meteorological data, not available from typical airport weather reports, during the period of the field study.

Meteorological data was gathered at three sites for the duration of the study. One site used a 2D sonic anemometer to measure wind speed, a second site used two conventional 3D anemometers to measure wind speed and how it changed with height above the ground, and the third site used 3D sonic anemometers and temperature probes to measure changes in wind with height as well as the temperature gradient.

6. Source Noise

The source noise investigation is intended to characterize the low-frequency content of aircraft noise near the runway. The information is intended to supplement the database of source noise characteristics.

6.1 Sideline Noise During Start-of-Takeoff Roll and Acceleration

The aim of the analysis of start-of-takeoff-roll (SOTR) signatures is to determine the frequency content and levels of sideline noise, recorded from positions parallel to and along the runway, and to determine how these change with aircraft type. Figure 14 shows a spectrogram of a SOTR event for a Boeing 777. The horizontal axis is time, the vertical frequency on a logarithmic scale. The numbers along the vertical axis correspond to 10, 100, 1000, 10 000 Hz. The upper frequency range is 20 kHz. Color indicates the sound pressure level in dB re 20 μ Pa. The aircraft passage takes place from about the 5 second mark to the 30 second mark. The near-parallel striations in the upper left portion of the graph, beginning at around 5 sec, represent tonal signals characteristic of inlet fan engine noise. The highest sound pressure levels, most noticeable by the darker colors in the 15 – 20 second period, occur at frequencies below 200 Hz.

It can be difficult to differentiate between signal and background noise at the lowest frequencies. This is typical of outdoor noise measurements and care must be taken to assure that the signal-to-noise ratio is sufficient for subsequent analysis. The typical signal-to-noise ratio for the measurements reported here is illustrated in Figure 15. This figure shows a graph of the frequency spectrum of the unweighted, energy-averaged equivalent sound level (L_{eq}) for all the start-of-takeoff-roll events recorded on Runway 30 at IAD, regardless of aircraft types. (Energy-averaged means that the mean-square pressures are averaged first and then that average is converted to dB. The 10 dB down points are used to determine the duration of the events.) The different colors indicate recordings from different microphones (refer to Table 3). The signal typically exceeds the background level by about 20 dB down to 16 Hz. The dip in level in the vicinity of 300 Hz is due to ground interference. The highest levels are in the frequencies below 200 Hz, which supports previous work indicating that behind the aircraft, significant sound energy was found below 200 Hz.¹¹

The lower graph in Figure 15 shows the A-weighted L_{eq} (L_{Aeq}). The de-emphasis of low frequencies by A-weighting is evident when comparing the upper and lower graphs.

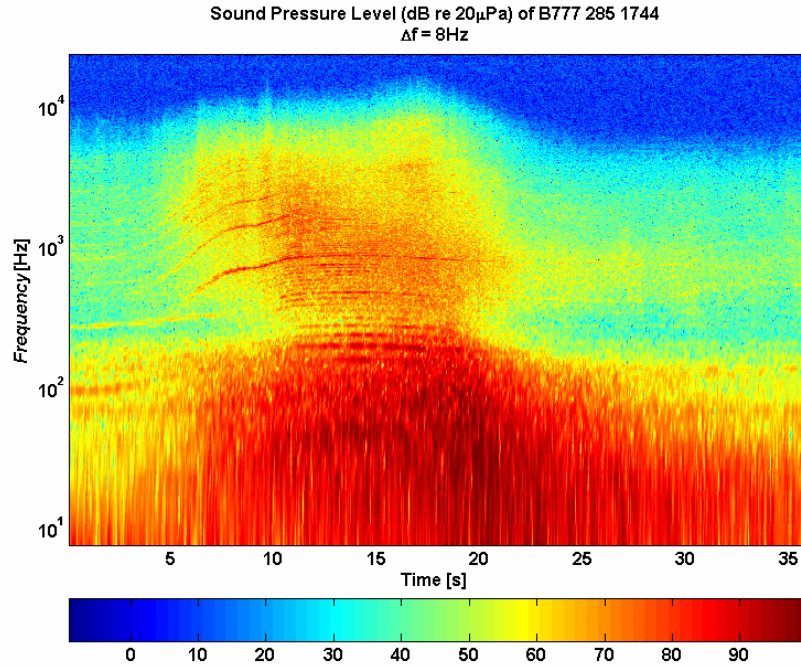


Figure 14. Spectrogram of sideline noise during start-of-takeoff roll (event B777 285 1744) as measured by Microphone 1 (see Table 2) 330 ft (100 m) from the centerline and 250 ft (76 m) from the start of Runway 30.

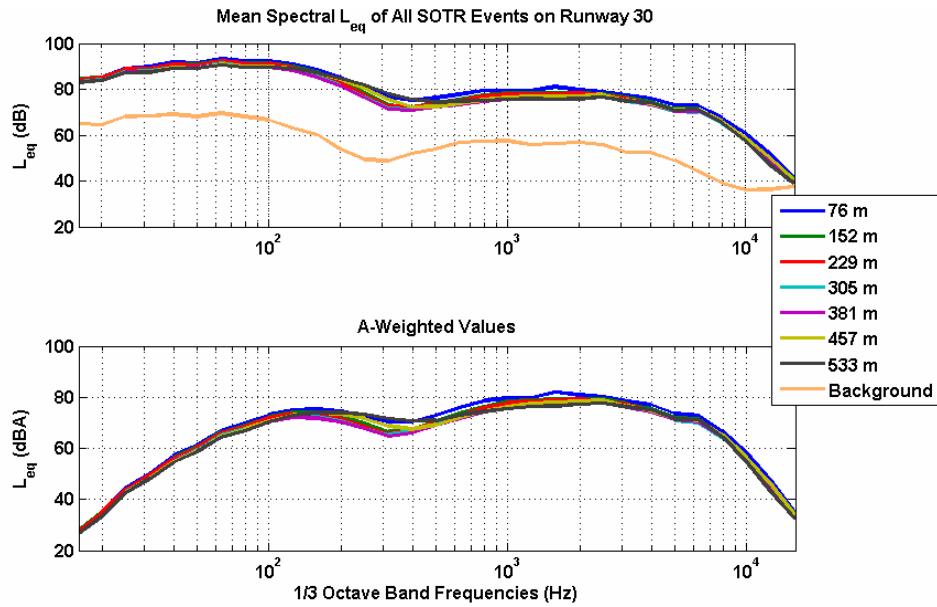


Figure 15. Mean Spectral L_{eq} for all SOTR events for different microphone positions. The upper graph shows the unweighted L_{eq} as well as the typical background level. The lower graph shows the A-weighted L_{eq} .

A common metric used to assess aircraft noise is the Sound Exposure Level (L_E). L_E can be thought of as the decibel level of a noise event of 1 second duration that contains the same energy as a different event of different duration. For this reason, L_E is useful in comparing events of different durations. The Sound Exposure Levels of noise recorded on the sideline of the runway during start-of-takeoff-roll (STOR) and acceleration down the runway are shown in Figure 16, categorized according to aircraft model and single engine maximum thrust. The vertical scale is L_E . The horizontal axis is categories of aircraft grouped by single-engine thrust ratings. The aircraft falling into a given category are also indicated on the graph. The thrust ranges for the categories were chosen such that there are roughly equal numbers of aircraft in each category. Within a grouping, the different data points indicate different microphones moving farther down the runway from left to right. The error bars indicate one standard deviation. In general, the heavier the aircraft, the higher the noise levels. During SOTR the highest levels were found to be at the beginning of the runway and decrease as the aircraft moved down the runway. The levels vary by 3-4 dB over the 1500 ft length of the microphone array.

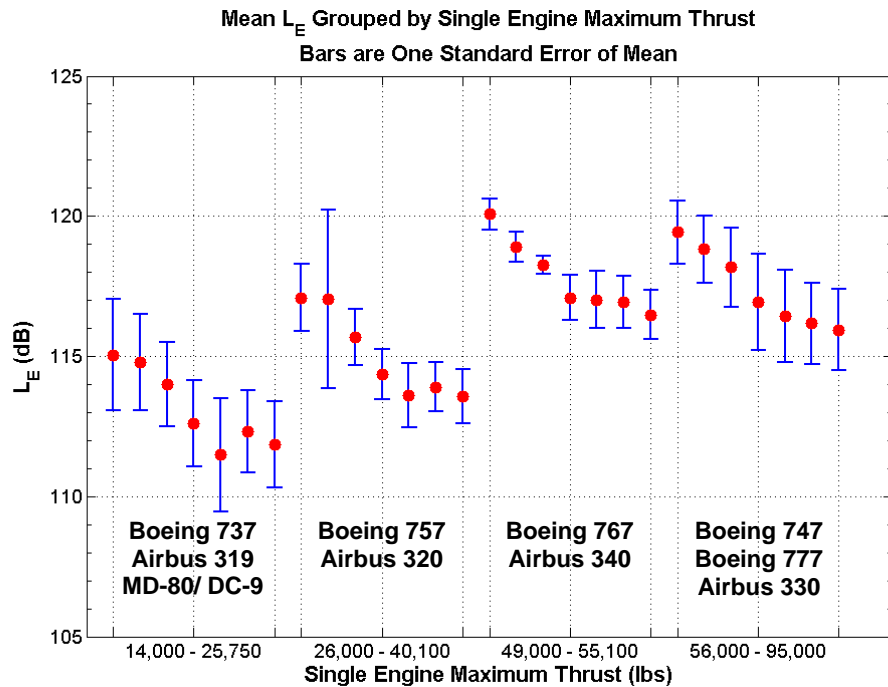


Figure 16. Sound Exposure Level (L_E) for sideline noise during start-of-takeoff-roll and acceleration down the runway grouped by single-engine maximum thrust. Within each of the four groups, data points indicate L_E measured at successive locations along the runway.

6.2 Thrust Reverser Noise

The thrust reverser noise data was analyzed in a similar way as the start-of-takeoff-roll data. A spectrogram of a thrust reverser event is shown in Figure 17. The event takes place from about the 4 second mark to the 16 second mark. The highest sound pressure levels, indicated by the darker red colors are most noticeable in the 5 – 12 sec period and occur at frequencies below 200 Hz. The typical signal to noise ratio is illustrated in Figure 18 which shows the mean spectral L_{eq} for all thrust reverser events on Runway 19R. The signal to noise ratio is equivalent to that for the start-of-takeoff-roll measurements shown in Figure 15. The highest levels are at the lowest frequencies, below 200 Hz. The MSP Expert Panel identified this noise as a potential significant contributor to low-frequency noise annoyance.

The thrust reverser Sound Exposure Levels (L_E) for different single engine thrust ratings are shown in Figure 19. The aircraft in the third thrust category have the highest levels. This trend is also seen in Figure 20 which is a plot of the average L_E as a function of aircraft type. The B757 and B767 have the highest levels, although B747, B777, and A330 aircraft have highest maximum takeoff weight

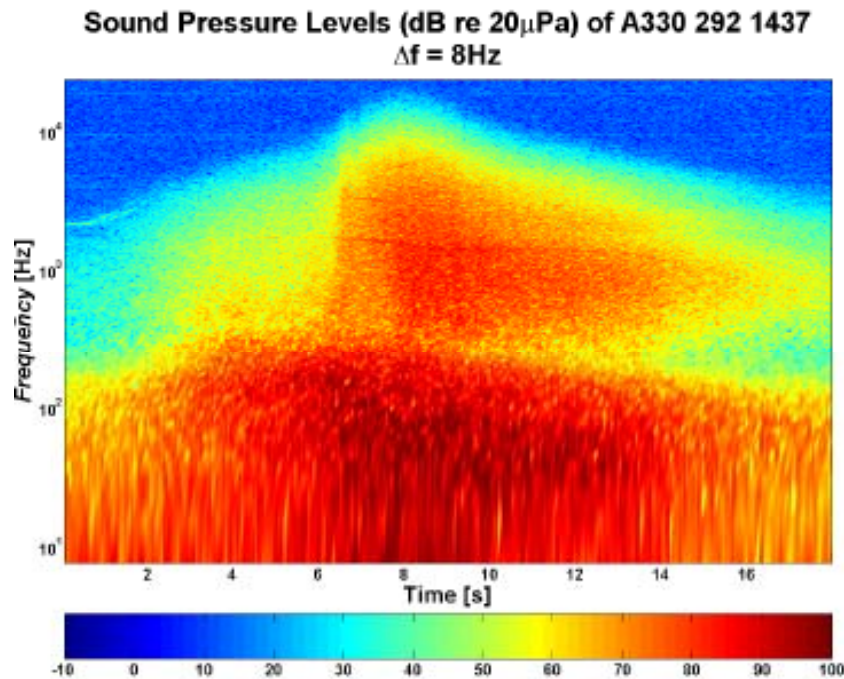


Figure 17. Spectrogram for thrust reverser event A330 292 1437 measured with Microphone 3 (see Table 1) 200 ft (61 m) from the centerline and 3000 ft (914 m) from the start of Runway 19R.

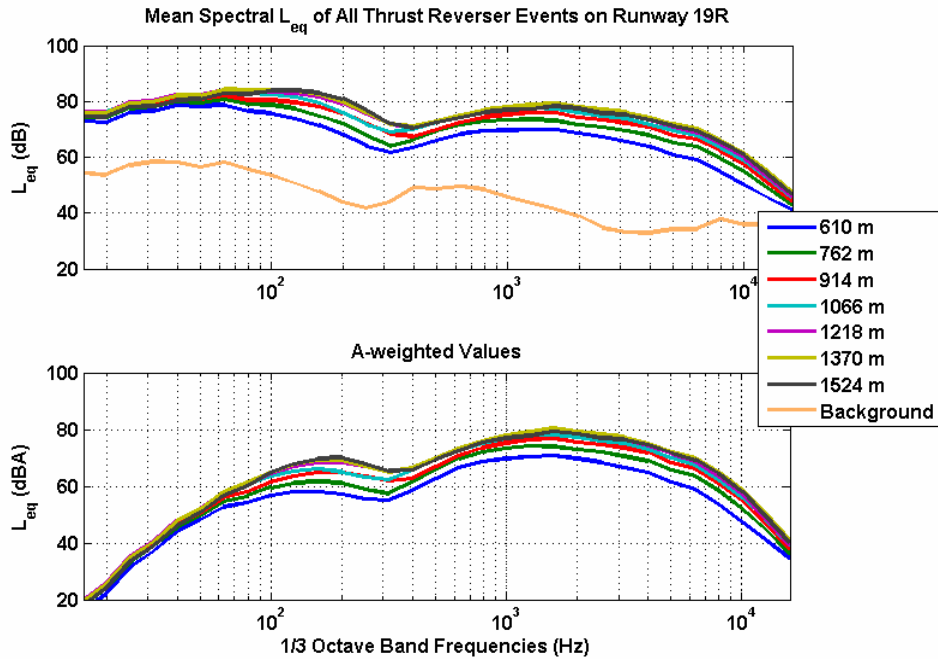


Figure 18. Mean Spectral L_{eq} for all TR events on Runway 19R for different microphone positions. The upper graph shows the unweighted L_{eq} as well as the typical background level. The lower graph shows the A-weighted L_{eq} .

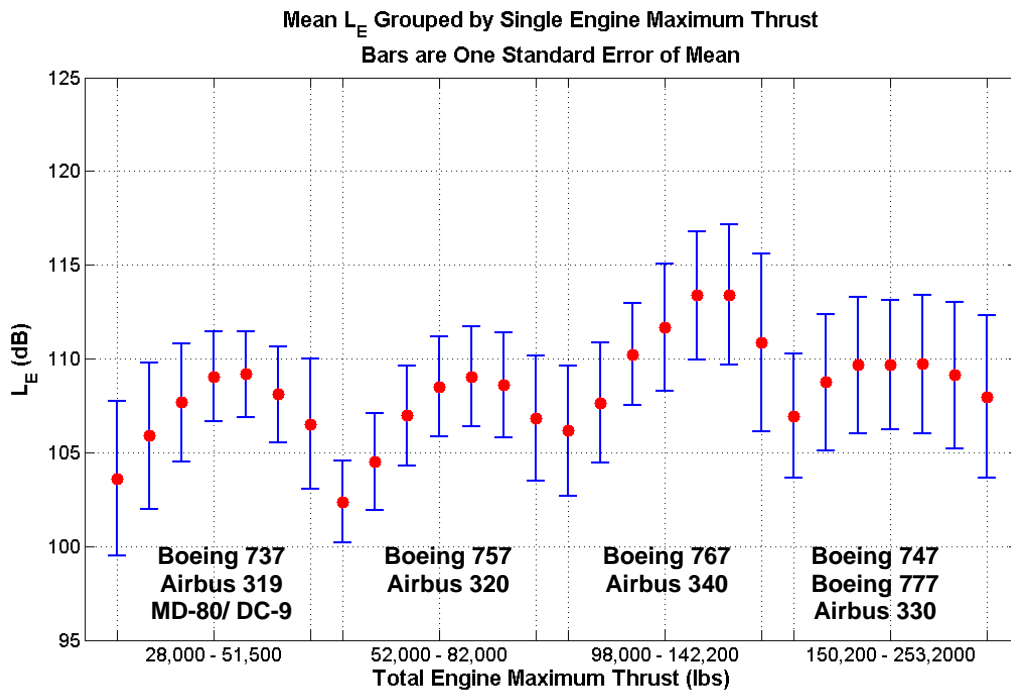


Figure 19. L_E for thrust reverser events grouped by single-engine maximum thrust. Within each of the four groups, data points indicate L_E measured at successive locations along the runway.

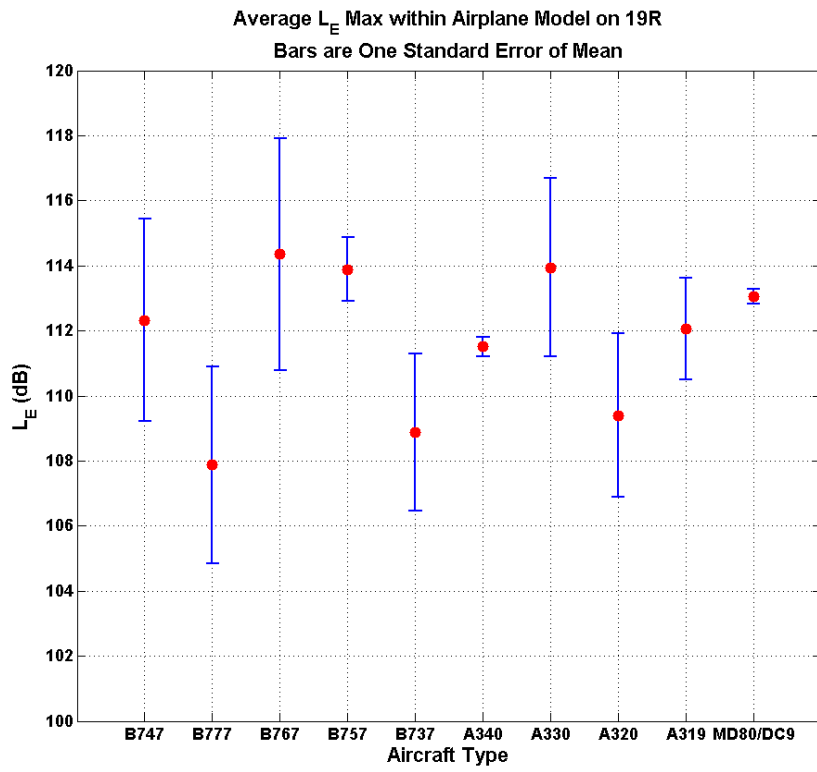


Figure 20. Average thrust reverser L_E by aircraft type.

6.3 Low Frequency Propagation Modeling

The low-frequency noise level in the vicinity of an airport is affected by many factors, including the noise source level and its spectrum, environmental temperature profile, wind direction and wind speed profile, ground effects (plowed field, vegetation, snow covered field, rain soaked field, desert, etc.), irregular terrain effects, atmospheric turbulence, and atmospheric attenuation. Standard modeling practice for assessing airport noise^{3,35,36} assumes spherical spreading, does not take meteorological conditions or terrain into account, and neglects source noise below 50 Hz. Therefore, it is important to determine what enhancements to standard practice are necessary to account for low frequency noise both in terms of the source noise database and propagation algorithm.

As a small first step, the deviation from spherical spreading due to variations in meteorological conditions was investigated using a parabolic equation model. The spherical spreading assumption, while very useful, is an approximation, and is not valid when meteorological conditions result in significant levels of atmospheric refraction. Atmospheric refraction is caused by variations in winds and temperature with height above the ground. Ray-tracing models are able to take atmospheric refraction into account, but are valid only at high frequencies. Parabolic-equation (PE) models can account for atmospheric refraction and are valid at low frequencies. State of the art PE models can also take terrain, ground impedance, and atmospheric turbulence into account.

Ideally, one would use the runway recordings as input to propagation models to predict the noise levels at the residential structures used as part of the IAD field study. However, because the characteristics of the source change as the aircraft moves down the runway, it is difficult to do a one-to-one comparison of runway noise with residential noise. Instead, a simpler approach was taken as a proof-of-concept. The meteorological data recorded at IAD was analyzed to determine the range of meteorological conditions experienced during the field study. The best and worse case conditions were used to determine the sensitivity of the PE noise predictions to this variability.

Table 5 summarizes the relevant meteorological parameters used for parabolic equation predictions. Three conditions are used. The "neutral" parameters represent a neutral atmosphere, i.e., no wind and no temperature gradient. The upwind and downwind conditions involve both wind and temperature gradients. Upwind (downwind) means that the wind is blowing from the receiver (source) toward the source (receiver). Wind raises the noise background at low frequencies, so data was recorded only in low wind conditions. Therefore, the range of conditions experienced at IAD was relative small.

Table 5. Typical wind speed and temperature at heights of 2 and 10 m for three wind directions used for parabolic-equation prediction of propagation loss.

	Wind (2m)	Wind (10m)	Temp. (2m)	Temp. (10m)
Neutral	0.0 m/s	0.0 m/s	15.0 °C	15.0 °C
Upwind	1.0 m/s	3.0 m/s	15.4 °C	14.7 °C
Downwind	1.0 m/s	3.0 m/s	15.4 °C	14.7 °C

Figure 21 shows the effect of the three different meteorological conditions on the propagation of 25, 50, and 200 Hz sound. Each figure is a graph of sound pressure level in dB relative to that of an omnidirectional source as a function of height and range. The source is located 2 m above the ground, assumed to be a rigid, flat surface. The graphs on the top row show the neutral conditions. The propagation from source to receiver obeys spherical spreading under these conditions. The middle and bottom rows correspond to the upwind and downwind conditions listed in Table 5, respectively. The most important conclusions are drawn by focusing attention on lower heights and longer ranges. Comparing the three conditions, it can be seen that levels may differ by 10 - 20 dB, even for the relatively small differences in meteorological conditions used here.

The model used for this study does not include atmospheric turbulence. As a result the levels predicted in the "shadow zone" in the upwind condition are under predicted. Sound scattered from inhomogeneities caused by atmospheric turbulence tends to fill in the shadow zone. As a result, a more realistic prediction of the levels would be about 20 dB lower than those predicted for the neutral condition. Turbulence can be included in parabolic-equation models, however, to do so the wind speed would need to be sampled at a rate of 10 – 20 Hz.

Figure 22 shows the relative sound pressure level for a receiver at a height of 2 m as a function of range for 50 Hz and the three meteorological conditions considered in Figure 21. The solid black line is for the crosswind condition, the dash-dot red line for downwind, and the dashed blue line for upwind. These curves are essentially horizontal slices through the three 50 Hz graphs in Figure 21 at a height of 2 m. The graph indicates the differences in level that an observer standing on the ground might experience under different weather conditions. The dotted blue line extending from the upwind case at ranges beyond 700 m is intended to indicate the approximate effect of including atmospheric turbulence in the prediction model. This line is based on the assumption that scattering of sound from turbulence into the shadow zone results in a level approximately 20 dB lower than the neutral case. As mentioned above, turbulence is not included in the model. The dotted curve is only included to indicate its importance.

While some argue that meteorological conditions get averaged out in cumulative metrics, such as the annual L_{dn} making more sophisticated propagation models unnecessary, weather effects do not entirely average out over a year.^{37,38} Three observations are warranted: 1) Meteorology conditions can have a significant effect on single-event levels and can affect noise contours; 2) If predictions of low-frequency noise (below 50 Hz) are necessary, current models are insufficient; 3) If overall improved accuracy is desirable, the potential improvement afforded by more sophisticated models should be assessed. European nations are developing a PE-based model called IMAGINE as part of their transportation noise modeling effort.³⁹

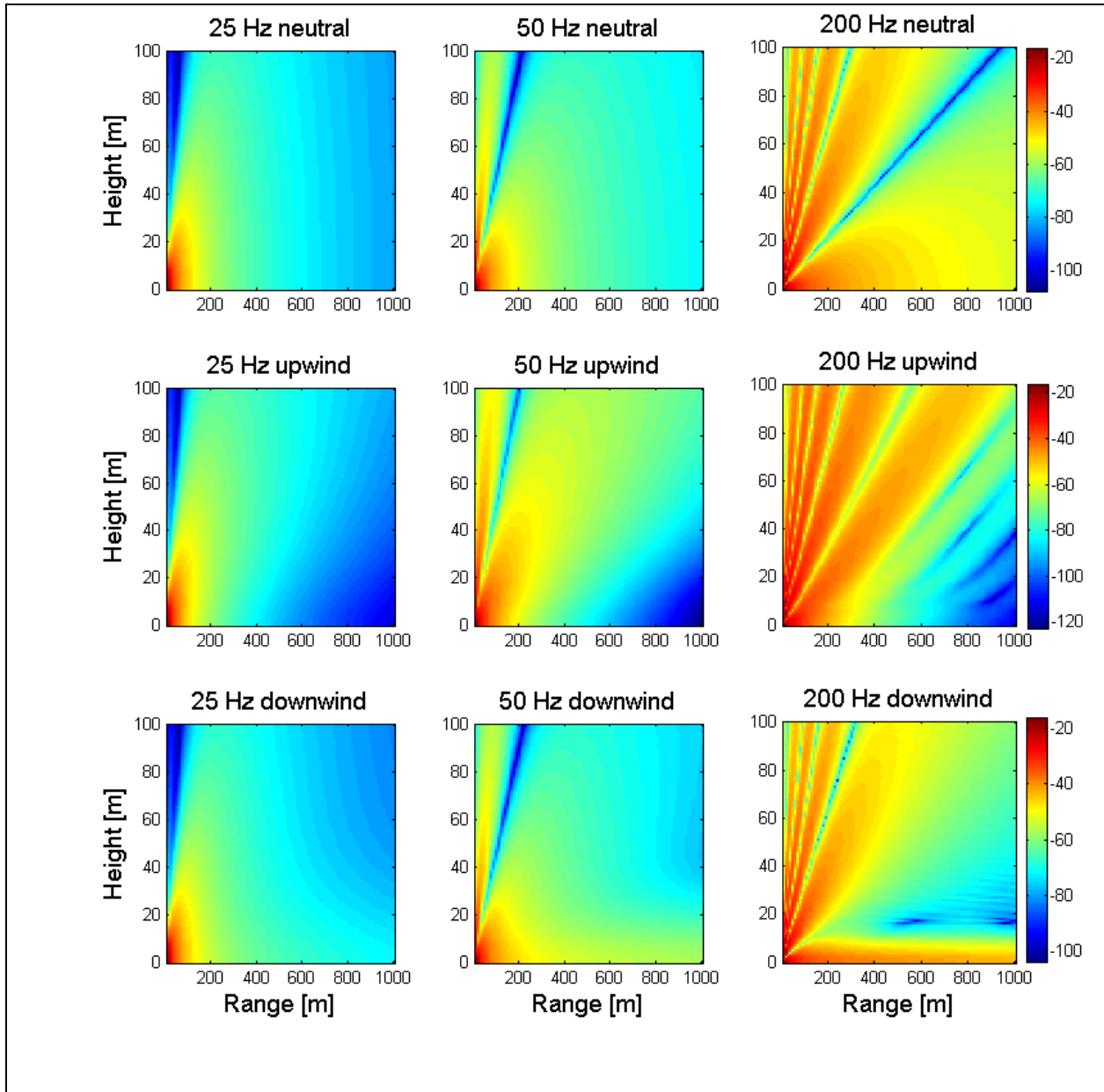


Figure 21. Parabolic equation predictions of the propagation loss of 25, 50, and 200 Hz sound from an omnidirectional source located 2 m above a rigid, flat surface for the three different meteorological conditions listed in Table 5. Each figure is a graph of sound pressure level in dB relative to the source as a function of height and range. The levels at low heights and large ranges can vary by more than 10 dB depending on the weather conditions. The model does not include atmospheric turbulence.

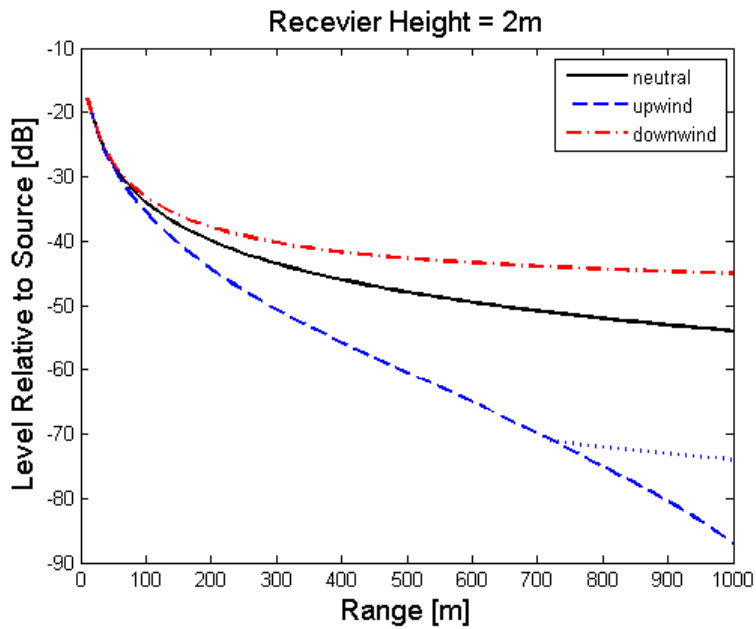


Figure 22. Graph of the sound pressure level relative to the source for a receiver at 2 m height as a function of range for 50 Hz sound and the three meteorological conditions considered in Figure 17. The solid black line is for the neutral condition, the dash-dot red line for downwind, and the dashed blue line for upwind. The dotted blue line extending from the upwind case at ranges beyond 700 m is intended to indicate the approximate effect of including atmospheric turbulence in the prediction model.

7. Noise and Vibration Impact on Residential Structures

The noise and vibration measurements at the two residential structures are intended to 1) assess the level and frequency content of noise-induced vibrations, 2) compare the levels with established thresholds for perception of vibration, and 3) determine which of the objective metrics of the outdoor sound listed best correlate with the vibration level.

7.1 Metrics for Vibration Impact and Subjective Perception of Annoyance

One of the objectives of this study is to investigate a range of objective metrics for their ability to predict the vibration impact based on the outdoor sound levels. Four categories of objective metrics were used for this purpose: level-based metrics, loudness-based metrics, metrics designed specifically for low-frequency noise, and perceived noisiness. The metrics used in this study are listed below.

Level

- L_{max} , L_{Amax} , L_{Cmax} Unweighted, A- and C-weighted Maximum Sound Pressure Level
- L_E , L_{AE} , L_{CE} Unweighted, A- and C-weighted Sound Exposure Level
- Q_A ($\alpha=3/40$) A-weighted Disturbance Index

Loudness

- LL Steven's Loudness (ISO 532A) and Zwicker Loudness (ISO 532B)
- $LLSEL$ Loudness Level Sound Exposure Level (ISO 226-1987)

Low-Frequency

- $LFNR$ Low Frequency Noise Rating
- $LFSL$ Low Frequency Sound Level
- L_{LF} Low Frequency Level

Perceived Noisiness

- PNL Perceived Noise Level
- Q_{PNL} ($\alpha=3/40$) Perceived Noise Level Disturbance Index
- Q_{PNL} ($\alpha=1/10$) Perceived Noise Level Disturbance Index

These community noise metrics are used to assess level, loudness and perceived noisiness of noise signatures. The Maximum Sound Pressure Level (L_{max}) reported here, is the maximum level occurring during any 0.5 second period of an event. Sound Exposure Level (L_E) is the decibel level of a noise event of duration 1 second that contains the same energy as a different event of different duration. Loudness Level (LL) is a single number rating in phons and correlates with the human assessment of loudness (in sones). Zwicker (ISO 532B) and Stevens Loudness Level (ISO 532A) were both developed and standardized to calculate loudness for complex noises. Loudness-Level Sound Exposure Level ($LLSEL$)^{40, 41} is the overall time-integrated phon level used to assess annoyance from environmental noise.

The low-frequency objective metrics are included to assess the low-frequency component.

These metrics include Low-Frequency Noise Rating (*LFNR*)⁴², Low-Frequency Sound Level (*LFSL*)⁴³ developed as a predictor for aircraft noise-induced rattle, and Low-Frequency Level (*L_{LF}*)⁴⁴.

The Perceived Noise Level (*PNL*) is a single number rating of the noisiness of a sound.⁴⁵ Disturbance index (*Q*) is “an average of the time-varying weighted noise level, measured over a specified time period to form an equivalent continuous noise level”.⁴⁶ It provides a single number quantitative measure of a noise event that is an empirically derived variation of the Equivalent Sound Pressure Level (*L_{eq}*). It has been used in Germany for rating aircraft noise and in Austria to assess road traffic noise.⁴⁶ The Disturbance Index was designed to more closely correlate with subjective data, and is based on two values, the type of level one seeks to measure, and a free parameter α . When α is set to 1/10, *Q* is equal to *L_{eq}*. In this study α is set to 3/40 as this variant of *Q* was found to correlate well with subjective rankings in previous studies.

Berglund et. al. have defined loudness by relating it to the magnitude of the sound, noisiness to the quality of the sound, and annoyance to the nuisance aspect of the sound within a given context. For both aircraft noise and community noise, the researchers observed that noisiness and annoyance increased proportionally with increased loudness. The effect was more pronounced for annoyance.^{47,48} Kryter distinguishes between loudness and noisiness based on the concept that the descriptor loudness does not infer an emotional based response (a low level of loudness may be desirable), but the descriptor noisiness always implies a negative connotation. He attributes noisiness to the combination of one or more of the following: loudness, tonality, duration, and the aspects of duration, impulsiveness, and variability.⁴⁹

There are more recently developed time-varying loudness models that incorporate many more aspects of human hearing than those incorporated in weighted sound pressure level and stationary loudness ISO532B. In several studies where the loudness exceeded 5% of the time (percentile loudness) estimated from these models correlate well with human perception of loudness, which is a primary driver in noise annoyance.

7.2 Vibration Impact on Residential Structures

7.2.1 Window Vibration

Figure 23 displays graphs of the root-mean-square (rms) acceleration in dB re $1\mu\text{g}$ of two windows in the kitchen of the Brick House as functions of frequency for thrust reverser events from different types of aircraft. Referring back to Figure 11, the east window is in the wall roughly parallel with Runway 19R, while the south window is in the wall roughly perpendicular to the runway. The general features of the responses are the same for both windows. There is a low frequency peak at around 40 Hz, and then numerous peaks at higher frequencies above 100 Hz. These higher-frequency peaks are likely modal resonances of the window panes. The levels of these peaks are about the same in both windows. However, the lower-frequency peak, likely due to a resonance in a larger structure such as the wall, is significantly higher in the east window than in the south. The east and south walls windows were of similar construction, therefore, the increased level on the east window is indicative either of a dependence on wall construction and/or window installation, or wall/window orientation with respect to the runway.

Figure 24 shows a comparison of the measured one-third octave band acceleration levels of the window in the east wall of the Brick House to the Hubbard threshold for perception of vibration.⁸ The aircraft type is identified by the first few letters of the filename given the legend. For example, B747_292_1612_BH refers to a Boeing 747. All of the events are landings on Runway 19R. This figure demonstrates that the acceleration levels of this particular window were below the threshold for tactile perception in most of the 1/3-octave bands. However, the level in the 40 Hz band approached the threshold in some cases. The data used by Hubbard to establish the threshold for perception are based on whole body vibration. As he points out that it is not clear how the concept of whole-body vibration applies to window or wall vibrations. Nevertheless, the thresholds do indicate perceptible vibration levels of various structures.

The Stone House was impacted strongly by thrust reverser and sideline acceleration due to operations on Runway 1R, and also impacted by start-of-takeoff-roll on Runway 30. Figure 25 shows a comparison of the Hubbard threshold to measured one-third octave band acceleration levels of the window in the west wall of the Stone House for various types of events. Most of the curves correspond to multiple events occurring simultaneously on more than one runway from various types of aircraft, indicated as "Mult" in the legend. The Hubbard threshold is reached or exceeded in one of the one-third octave band during two of the events.

The response of the north window of the Stone House is shown in Figure 26. In contrast to the measurements on the west window (Figure 25), there is a greater tendency for the thresholds to be exceeded on this window. This difference could be partially explained by the differences in orientation of the windows with respect to the runway. In particular one might expect that the west wall would be impacted more from operations on Runway 30 than from those on 1R. However, the location of the west wall puts it in the vicinity of the quiet zone of aircraft taking off from Runway 30 and it is shielded from operations on 1R.

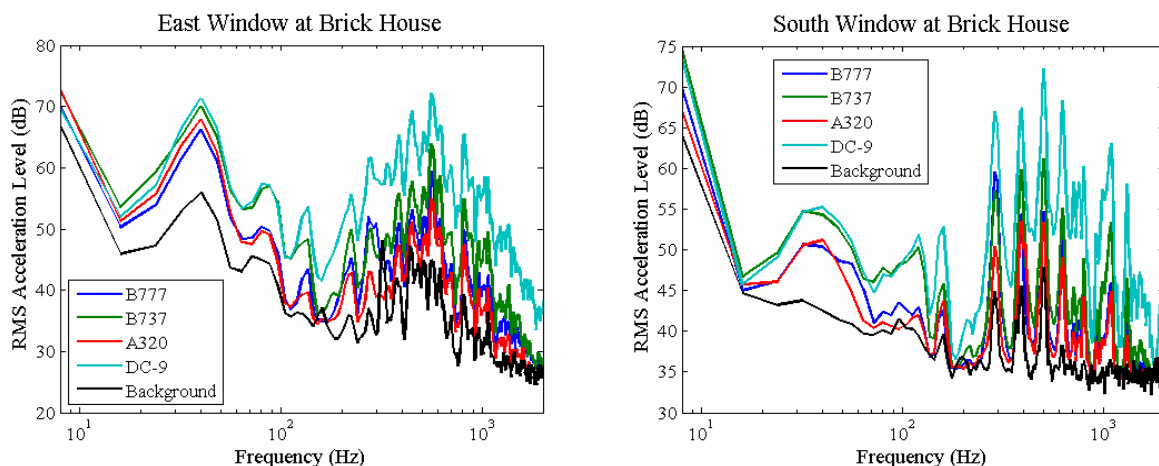


Figure 23. Acceleration levels (in dB re 1µg) vs. frequency of the east and south windows in the brick house due to thrust reverser events for different aircraft types on Runway 19R. The differences in the levels of the low-frequency peaks of the two windows are indicative of a dependence on orientation with respect to the runway.

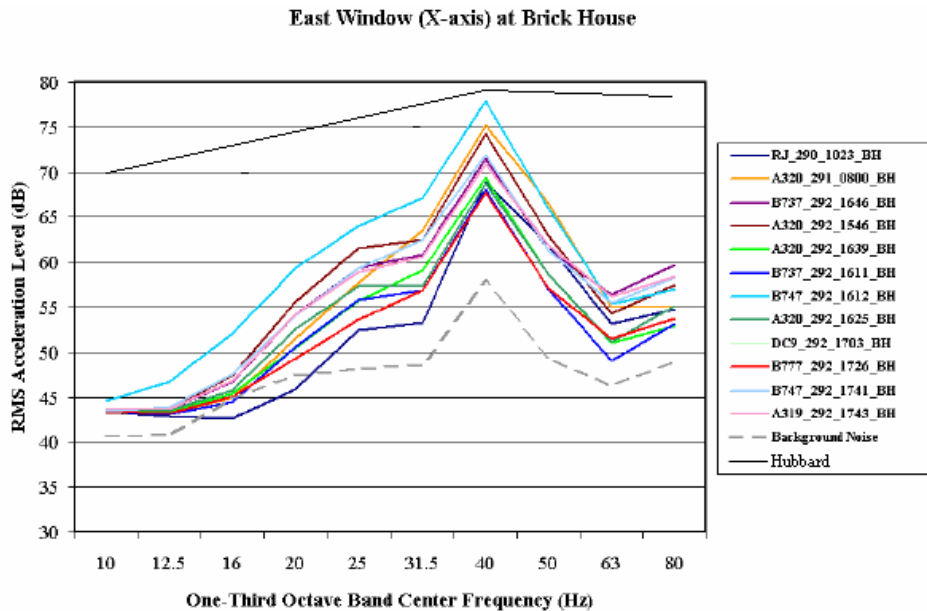


Figure 24. Comparison of the acceleration levels (dB re 1 μ grms) measured on the window on the east wall of the Brick House to the Hubbard criteria threshold for tactile perception of vibration for landings of various types of aircraft.

The vibration levels of windows in the houses located within about 3000 ft of runways can exceed the Hubbard criteria. Based on the two houses used in IAD, the vibration level is greater in houses that are impacted by events on two runways simultaneously. There is also evidence of a dependence on window/wall orientation with respect to the runways. These results demonstrate the need to give careful consideration to both the orientation of the house with respect to runways and also the nature of operations on that runway. This indicates that the potential for window vibration is generally more of a concern because of the possible production of secondary rattle rather than the tactile perception of vibration.

Investigators noticed that a loose window pane at the Stone House sometimes produced audible rattle during noise events, yet this window pane was located in the same frame as a well-installed window that did not rattle. An accelerometer was placed on both the loose and the secure windows in order to compare the rattle and non-rattle characteristics within the same window frame. Figure 27 shows a comparison of acceleration response of the secure (left) and loose (right) windows to the same event, a Boeing 777 landing. The peak levels are higher for the loose window and extend over a broader range of frequencies for a longer duration of time than for the secure window. These characteristics are indicative of rattle.

Figure 28 shows a comparison of the rms acceleration levels of the secure (left) and loose (right) windows for several types of events. (The left-hand figure is the same as Figure 22). As would be expected, the levels of the loose window are higher than those of the secure window. Moreover, the levels are well above the Hubbard threshold over a broad range of frequencies. The events plotted are those when audible rattle events occurred, therefore, not only are the vibration levels great enough to be perceived, they also correspond to audible rattle.

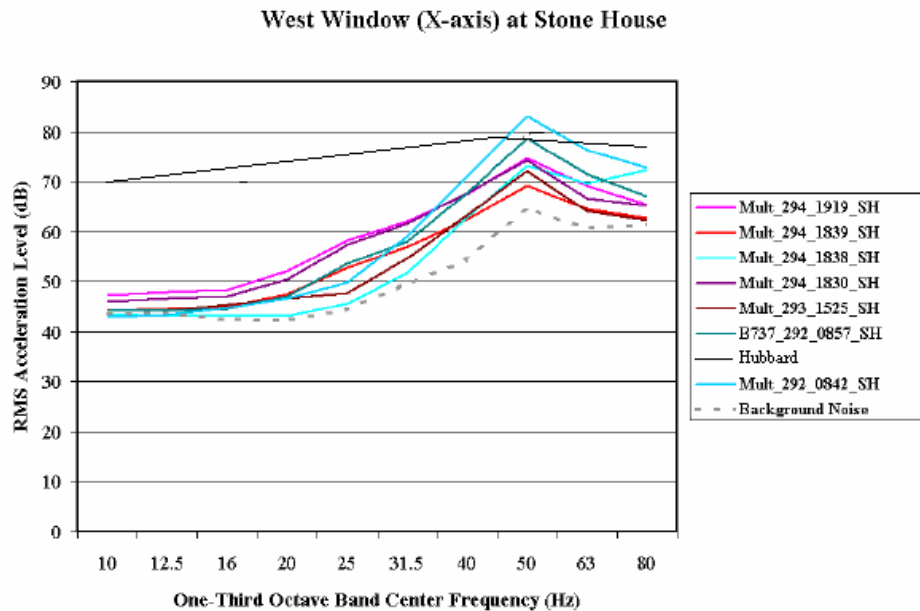


Figure 25. Comparison of the acceleration levels (dB re 1 μ grms) measured on the west window of the Stone House to the Hubbard criteria threshold for tactile perception of vibration. Most of the curves correspond to multiple events occurring simultaneously on more than one runway from various types of aircraft, indicated as "Mult" in the legend.

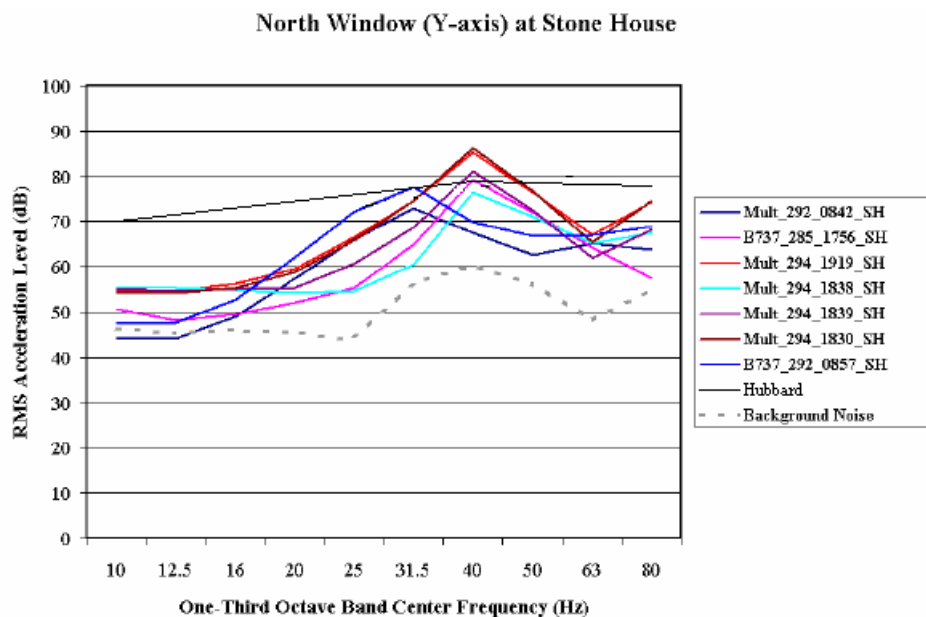


Figure 26. Comparison of the acceleration levels (dB re 1 μ grms) measured on the north window of the Stone House to the Hubbard criteria threshold for tactile perception of vibration. Most of the curves correspond to multiple events occurring simultaneously on more than one runway from various types of aircraft, indicated as "Mult" in the legend.

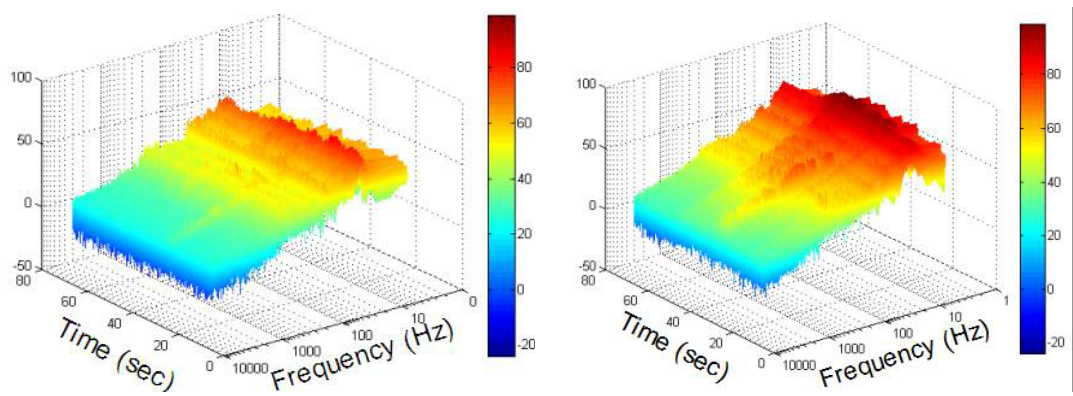


Figure 27. Window acceleration levels impact from Boeing 777 landing. The response of the secure window is on the left, that of the loosely coupled window. Increased acceleration levels can be seen in red. The scale at the right indicates the color associated with the dB level (re 1 μ grms). The pattern seen in the right-hand figure is characteristic of rattle.

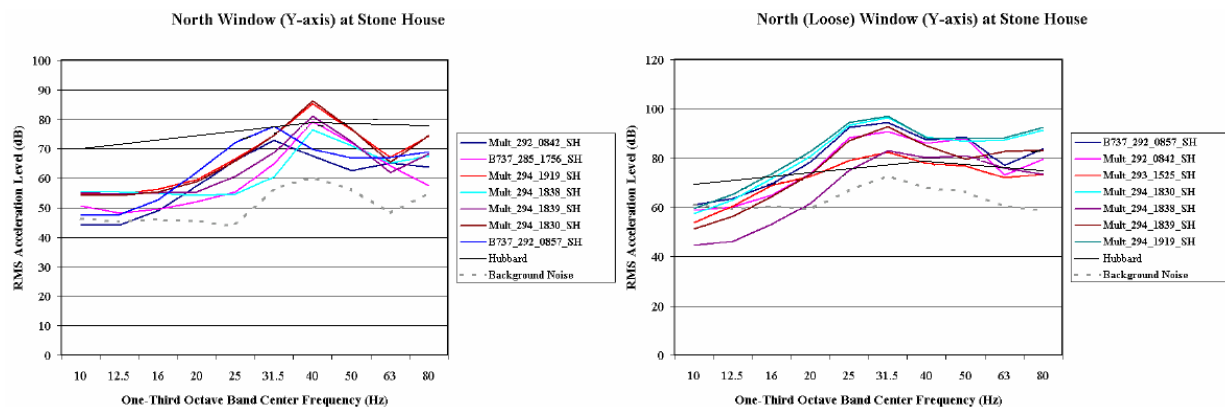


Figure 28. Comparison of the acceleration levels (dB re 1 μ grms) measured on the north secure window of the Stone House to those measured on the loose north window.

7.2.2 Wall and Floor Vibration

The wall and floor vibration levels are much lower than those of the windows. As observed previously by Hubbard,⁸ the wall acceleration levels were approximately 10 dB lower than the window levels for the same event. A previous study at BWI found the maximum C-weighted sound pressure level L_{Cmax} to strongly correlate with maximum wall acceleration levels.⁶ Figure 25 shows the maximum rms acceleration level (in dB re 1 μ g) for the east wall of the Brick

House (red squares) and the north wall of the Stone House (open circles) as functions of L_{Cmax} . The shaded area indicates the range of Hubbard's threshold for tactile perception of vibration for frequencies ranging from 10 to 100 Hz. As might be expected, the vibration level of the wall in the Stone House is only slightly dependent on outdoor sound level and well below the thresholds. The vibration levels of the Brick House wall are higher than those of the Stone House and generally increase with L_{Cmax} . This result is consistent with the BWI study.⁶ The vibration levels for the events with the highest outdoor peak C-weighted sound levels are in the range where the vibrations might be perceived.

The acceleration levels of the floors measured in either house (not shown) did not rise significantly above background levels for any of the events recorded. Consequently, we concluded that the floor vibrations levels are significantly below the thresholds for perception. The floors in the instrumented rooms were located above basements in both houses rather than on cement slabs.

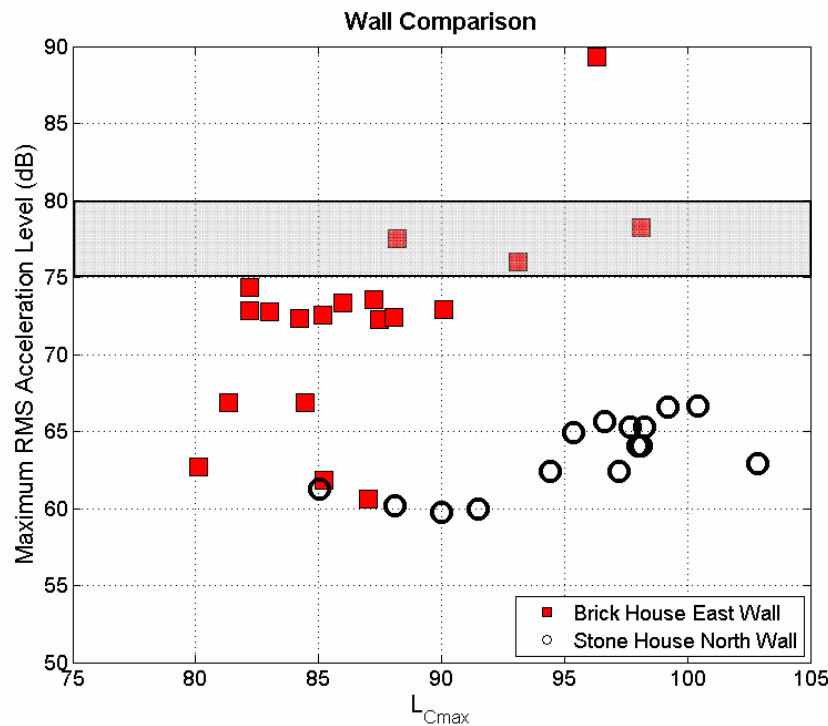


Figure 29. Graph of maximum rms acceleration level (in dB re $1\mu g$) for the east wall of the Brick House (blue diamonds) and the north wall of the Stone House (purple squares) vs. the maximum outdoor C-weighted sound pressure level (in dBC re $20\mu Pa$.) The shaded areas indicate the range of Hubbard's threshold for tactile perception of vibration for frequencies ranging from 10 to 100 Hz.

7.2.3 Hubbard's Sound Pressure Level Threshold Criteria

In the previous sections, rms acceleration levels were compared to the acceleration threshold criteria established by Hubbard in Figure 10 of Ref. 8. However, Hubbard also developed a set of criteria for outdoor sound pressure level (SPL) sufficient to cause perceptible vibration of house structure elements (windows, walls, and floors). (These SPL criteria are shown in Figure 9 of Ref. 8.) Figure 30 shows the Hubbard SPL thresholds for window, wall, and floor vibrations in the 10 – 100 Hz frequency range. The data points represent the 1/3-octave band levels during the 0.5 second interval corresponding to the peak sound pressure level for the event. The upward-pointing triangles are for event Mult_294_1919_SH that caused the largest acceleration levels in the north windows of the Stone House shown in Figure 28, as this event also caused the loose window to rattle. The levels exceed both the window and wall Hubbard SPL threshold criteria. The downward-pointing triangles are for event Mult_293_1522_SH, an event that did not cause the loose window to rattle. The levels for this event do not exceed the Hubbard criteria to any significant degree. This comparison suggests that the Hubbard outdoor SPL threshold criteria serve as a useful guideline for determining when perceptible vibration of house elements might occur.

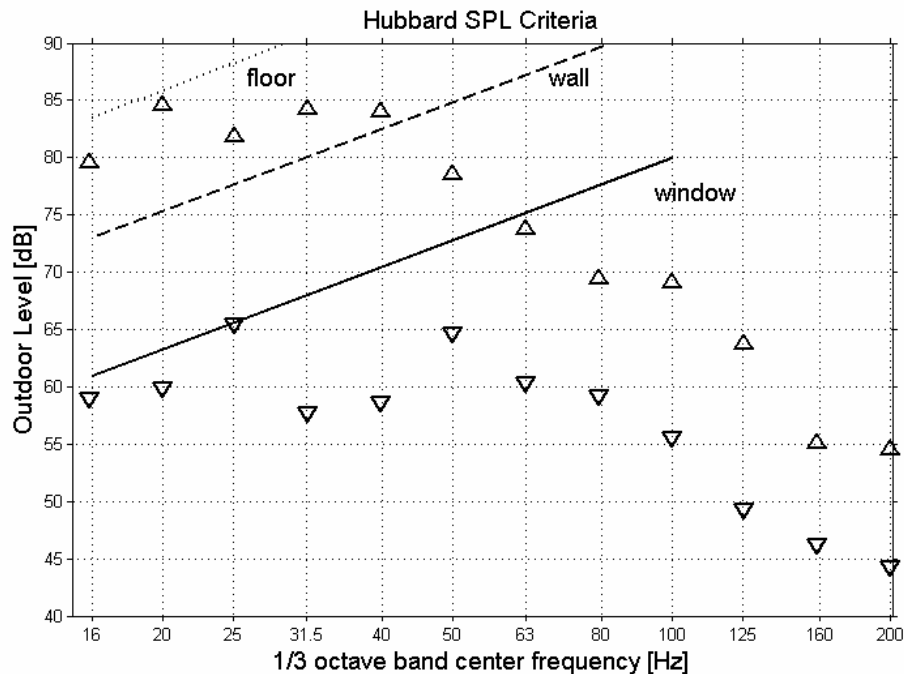


Figure 30. Peak outdoor sound level at the Stone House compared to Hubbard's SPL thresholds for perceptible vibration of house structure elements. The upward-pointing triangles are for the event Mult_294_1919_SH that caused the largest acceleration levels in the north windows of the Stone House shown in Figure 28. This event caused the loose window to rattle. The downward-pointing triangles are for event Mult_293_1522_SH, an event that did not cause the loose window to rattle.

7.2.4 Outdoor Metric Indicators of Indoor Audible Window Rattle

The previous sections discussed the utility of the Hubbard criteria for predicting perceptible vibrations of house structure elements. In this section, the metrics discussed in Section 7.1 are correlated with the rms acceleration level of the rattle-prone window in the Stone House, to identify which correlate best with window rattle.

Seventeen events (13 with rattle, and 4 without) recorded at the Stone House were analyzed to search for such a metric. The results are shown in Figure 31. The vertical axis indicates whether rattle was or was not audible on the indoor microphone recording of these events. "No audible rattle" is indicated by black open circles, "yes, rattle is audible" by the red Xs. The figure on the left compares the unweighted, C-, and A-weighted maximum outdoor sound pressure level L_{max} , while that on the right compares L_{Cmax} , L_E , and L_{FSL} . There is a range of overlap in which rattle both is and is not audible. However for each metric, there is a level above which only rattle always occur. For instance, referring to the upper left panel, the data indicates that when L_{max} exceeds 100 dB, this particular window always rattles. The thresholds for L_{Cmax} and L_{Amax} are approximately 98 and 95 dB, respectively. Notice however, that for both L_{max} and L_{Cmax} 9 of the 13 rattle events are above the "rattle only" threshold, compared to only 3 of 13 for L_{Amax} . Therefore, L_{max} and L_{Cmax} provide more meaningful thresholds than does L_{Amax} .

The figure on the right compares L_{Cmax} , L_E , and L_{FSL} . These three metrics perform more-or-less equally well, in that rattle always occurs above a threshold of roughly 97 dB.

These results are for a particular window and other windows will respond differently. Nevertheless, the results can be used as a guideline for which metrics might prove useful in assessing the potential for rattle. Modern windows have optional plastic grid inserts. Anecdotal observations indicate that these inserts are prone to rattle.

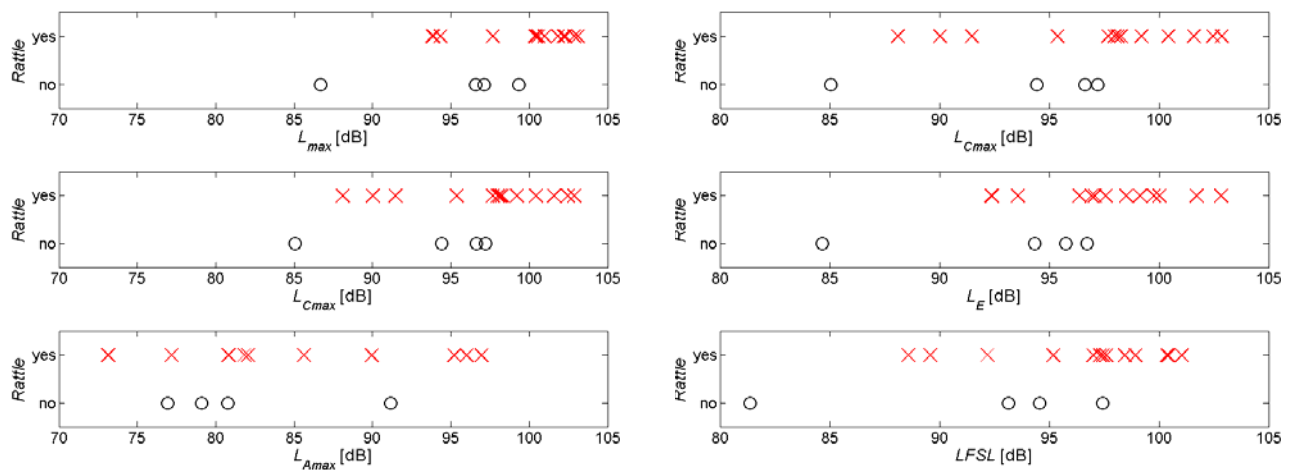


Figure 31. Correlation of the occurrence of audible rattle from the rattle-prone window in the Stone House heard on the indoor recording and six metrics calculated from the outdoor microphone.

7.2.5 Correlation of Outdoor Sound Pressure Metrics with Interior Vibration Levels

The outdoor sound pressure metrics were compared to the vibration levels of the windows and walls to determine which metrics are the best predictors of vibration levels. The results are shown in Tables 6-10.

Table 6 shows the results for the rattle-prone window in the Stone House. In each case, the metrics were compared to the maximum acceleration levels of the window. The Pearson Correlation Coefficient is an indicator of the degree of correlation between the acceleration level and the metric. A correlation coefficient of 1.0 indicates a perfect linear correlation. The P-value obtained is a measure of the statistical significance of the correlation. For example, a P-value of 0.01 means that the probability that the correlation could be due to chance is 1%. Smaller P-values correspond to higher levels of statistical significance.

The metrics appearing above the heavy line in Table 6 all correlated with the maximum acceleration level at a significance level of greater than 99%, i.e., the P-value is less than 0.01. The metrics that correlated best were, in order, the unweighted Sound Exposure Level L_E , the Low-Frequency Sound Level L_{FSL} , the Low-Frequency-Level L_{LF} , and the maximum sound pressure level L_{max} . C-weighted Sound Exposure Level L_{CE} and Maximum Sound Pressure Level L_{Cmax} also correlated well, although to a lesser degree. The A-weighted metrics correlated poorly with the acceleration levels.

Table 6. Statistical Correlation between metrics from the outdoor microphone and the maximum acceleration level of the rattle-prone window at the Stone House. Metrics in the rows above the heavy line show linear correlation with the maximum acceleration level at a significance level exceeding of 0.01 meaning that the probability that the correlation could be due to chance is less than 1%.

Rattle-Prone Window at Stone House		
Metric	Pearson Correlation	P-value
L_E	0.950	0.000
L_{FSL}	0.947	0.000
L_{LF}	0.930	0.000
L_{max}	0.913	0.000
L_{CE}	0.898	0.000
L_{Cmax}	0.838	0.000
$LLSEL$	0.548	0.011
$Q_{PNL \alpha = 3/40}$	0.547	0.011
$Q_{PNL \alpha = 1/10}$	0.541	0.012
PNL	0.533	0.014
$Q_A \alpha = 3/40$	0.445	0.036
L_{Amax}	0.431	0.042
L_{AE}	0.422	0.046

In contrast, the correlation of the outdoor metrics with the maximum acceleration levels of the windows that did not rattle does not present as clear a picture. The results shown in Tables 7 for the windows in the north and west walls of the Stone House, and Table 8 for the windows in the east and south walls of the Brick House shown in Table 8. Generally, the metrics that correlated well for the rattle-prone window in Table 6 correlate the poorest for the secure windows, and vice versa.

Overall, the correlations for the Stone House west window are only fair at best, the maximum value being 0.786, and the reason for this poor correlation is as of yet not fully understood. One possible explanation has to do with the orientation of the windows with respect to the runways. The instrumented room in the northwest corner of the Stone House is oriented to the rear of takeoff operations on Runway 30. Therefore, the room may be shielded somewhat by the quiet zone behind the aircraft. However, the Stone House was impacted by operations on Runway 1R. The north wall is roughly perpendicular to 1R, while the west wall is around the corner. Therefore, the west window could be shielded from impact, whereas the north window is not.

A similar trend is seen for the Brick House. While the correlation coefficients are low for both windows, they are higher for the east window than for the south window. A dependence on orientation with respect to the runway was also noted previously in this report based on analysis of the vibration data for the windows.

The correlations between the metric values and maximum acceleration levels for the walls in the Brick and Stone Houses are shown in Tables 9 and 10, respectively. The correlation coefficients are generally low. Previously in Figure 29, the maximum acceleration levels for the east wall of the Brick House and north wall of the Stone House were plotted against L_{Cmax} . Examining Tables 9 and 10, this metric is nearly best correlated for the Brick House east wall and moderately correlated for the Stone House north wall.

Table 7. Statistical Correlation between metrics from the outdoor microphone and the maximum acceleration level of the secure windows at the Stone House.

Stone House Windows

North Secure Window ($n = 17$)		
Metric	Pearson	
	Correlation	P-value
$Q_{A \alpha = 3/40}$	0.947	0.000
$Q_{PNL \alpha = 3/40}$	0.946	0.000
$Q_{PNL \alpha = 1/10}$	0.946	0.000
PNL	0.943	0.000
LLSEL	0.925	0.000
L_{Amax}	0.904	0.000
L_{Cmax}	0.837	0.002
L_{CE}	0.778	0.000
L_{max}	0.716	0.015
L_{AE}	0.671	0.003
L_E	0.480	0.096
LFSL	0.124	0.370

West Secure Window ($n = 17$)		
Metric	Pearson	
	Correlation	P-value
LLSEL	0.701	0.001
$Q_{PNL \alpha = 3/40}$	0.675	0.002
$Q_{A \alpha = 3/40}$	0.672	0.002
PNL	0.665	0.002
L_{AE}	0.662	0.002
$Q_{PNL \alpha = 1/10}$	0.659	0.002
L_{Cmax}	0.569	0.008
L_{CE}	0.543	0.012
L_{max}	0.510	0.018
L_{Amax}	0.476	0.026
L_E	0.401	0.056
LFSL	0.351	0.084

Table 8. Statistical Correlation between metrics from the outdoor microphone and the maximum acceleration level of the windows in the east and south walls of the Brick House.

Brick House Windows

East Window		
Metric	Pearson	
	Correlation	P-value
LFSL	0.764	0.000
L_{eq}	0.758	0.000
L_{Ceq}	0.745	0.000
L_{CE}	0.706	0.000
L_{max}	0.687	0.001
L_E	0.679	0.001
$Q_{PNL \alpha = 3/40}$	0.626	0.002
$Q_{PNL \alpha = 1/10}$	0.623	0.003
$Q_{A \alpha = 3/40}$	0.62	0.003
PNL	0.619	0.003
L_{Amax}	0.618	0.003
L_{Cmax}	0.616	0.004
L_{Aeq}	0.615	0.004
L_{AE}	0.615	0.004
LLSEL	0.604	0.004

South Window		
Metric	Pearson	
	Correlation	P-value
$Q_{PNL \alpha = 3/40}$	0.561	0.008
L_{Aeq}	0.559	0.008
$Q_{A \alpha = 3/40}$	0.547	0.006
$Q_{PNL \alpha = 1/10}$	0.547	0.009
PNL	0.536	0.011
LLSEL	0.505	0.016
L_{Ceq}	0.494	0.018
L_{Amax}	0.458	0.028
L_{EA}	0.454	0.029
LFSL	0.377	0.062
L_{eq}	0.364	0.069
L_{Cmax}	0.351	0.076
L_{max}	0.331	0.09
L_{CE}	0.328	0.091
L_E	0.177	0.241

Table 9. Statistical Correlation between metrics from the outdoor microphone and the maximum acceleration level of the east and south walls of the Brick House.

Brick House Walls

East Wall			South Wall		
Metric	Pearson		Metric	Pearson	
	Correlation	P-value		Correlation	P-value
L_E	0.641	0.002	L_{CE}	0.616	0.003
L_{Cmax}	0.639	0.002	L_{max}	0.581	0.006
L_{max}	0.597	0.004	LFSL	0.572	0.006
L_{CE}	0.594	0.004	L_E	0.564	0.007
LFSL	0.529	0.012	PNL	0.543	0.010
L_{EA}	0.514	0.014	$Q_{PNL \alpha = 3/40}$	0.528	0.012
PNL	0.432	0.036	$Q_{PNL \alpha = 1/10}$	0.51	0.015
LLSEL	0.418	0.042	L_{AE}	0.509	0.015
$Q_{PNL \alpha = 3/40}$	0.405	0.047	LLSEL	0.495	0.019
$Q_{PNL \alpha = 1/10}$	0.402	0.049	L_{Cmax}	0.492	0.019
$Q_{A \alpha = 3/40}$	0.396	0.052	$Q_{A \alpha = 3/40}$	0.484	0.021
L_{Amax}	0.362	0.069	L_{Amax}	0.468	0.025

Table 10. Statistical Correlation between metrics from the outdoor microphone and the maximum acceleration level of the north wall of the Stone House.

North Wall at Stone House ($n = 15$)		
Metric	Pearson	
	Correlation	P-value
LLSEL	0.784	0.000
L_{CE}	0.778	0.000
$Q_{PNL \alpha = 1/10}$	0.773	0.000
$Q_{PNL \alpha = 3/40}$	0.771	0.000
PNL	0.755	0.000
L_{Cmax}	0.739	0.001
L_E	0.737	0.001
LFSL	0.696	0.002
L_{max}	0.695	0.002
$Q_{A \alpha = 3/40}$	0.694	0.002
L_{AE}	0.671	0.003
L_{Amax}	0.651	0.004

8. Design and Analysis Subjective Jury Trials

8.1 Spectral Balance and Rattle Subjective Study Design

Initially two separate jury trials, one focused on spectral balance of noise signatures and another focused on the perception of audible rattle embedded in noise, were conducted. The goals of the tests were to assess the subjective response to aviation noise signatures recorded in the residential structures, and to correlate objective metrics to those subjective ratings. A third, follow-on study was conducted to assess thresholds for low-frequency aviation noise annoyance. The two initial studies are discussed first, followed by the third.

The signatures were presented to the subjects in Gulfstream's aircraft noise simulator known as Supersonic Acoustic Signature Simulator II (SASSII). (Gulfstream designed SASSII to simulate noise from supersonic aircraft. The present low-frequency noise study has nothing to do with supersonic flight. We used the simulator because of the high fidelity of low frequency sound reproduction.) The simulator is designed to present frequencies as low as 7 Hz. As such, the low frequency content of the signatures can be perceived both audibly and through tactile sensations. The spectral balance and rattle subjective studies presented signatures at a level that was in the audible range of perception. The low frequency threshold study included some signatures that may have induced a vibro-tactile sensation due to the high-level low frequency content.

The simulator, housed in a mobile 32-foot (9.8 m) long trailer, is shown schematically in Figure 32. Subjects were seated two at a time in the 7 by 11 foot (2.1 by 3.4 m) audio booth, facing the sound reproduction system located behind an acoustically-transparent screen. SASS-II provides a laboratory presentation format that is more realistic than listening under headphones. The participants are seated to listen in a space, rather than wearing headphones. The audio system was balanced so that signals experienced by the two subjects were matched to the extent possible.

Use of a laboratory setting allows for greater control of the test signatures and increased validity in the statistical findings, though a laboratory environment is not as natural as listening in one's home. Annoyance ratings obtained in the laboratory may not include the effect of intrusiveness the noise has on daily activities. A potential follow-on test of intrusiveness could be conducted in which participants are engaged in activity such as reading while listening to low-level background recordings of a news show. The aviation signatures could be presented over this background noise to conduct a test of intrusiveness.

Subjects for the study were recruited from a variety of different sources on Penn State's campus and, to recruit older subjects, advertisements were placed at a local senior center and on various community bulletin boards. The spectral balance study included 36 subjects and the rattle study included 41 subjects. All subjects had hearing within normal limits (less than 20 dB hearing loss) in the range from 250 Hz to 8 kHz when screened using a Maico MA 19 portable audiometer. Subjects were aware of what types of noises they would be rating, and what the goals of the study were, based on the following:

- The informed consent signed by each participant stated that the study addressed community impact of noise.

- The recruiting poster had a picture of an airplane and referred to aviation noise.
- The directions and response sheet used annoyance as the measure of the signature.

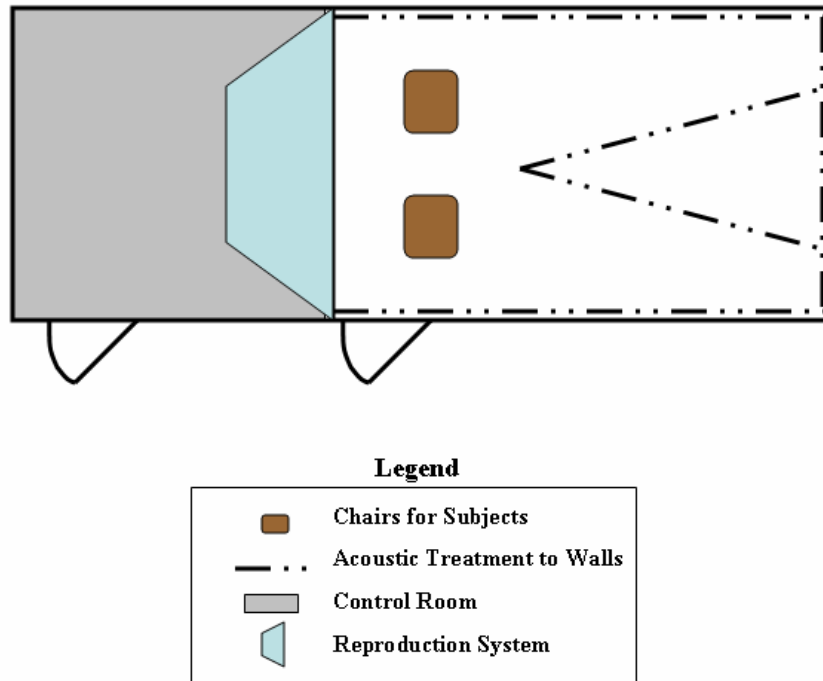


Figure 32. Diagram of the Gulfstream Supersonic Signature Simulator II used for subjective presentations.

8.1.1 Noise Signatures Used in the Spectral Balance and Rattle Subjective Tests

Six signatures derived from indoor recordings were used for each of the two trials. Tables 11 and 12 contain information about the signatures used in the spectral balance study and the rattle study, respectively. The left-hand column contains the filename used to designate the various signatures. The first part of the filename indicates the aircraft type, when known. "U" indicates that the aircraft type is unknown. "Mult" indicates that the signature consisted of multiple events. The signature designation includes the year-date in 2004, time of day, and location, where SH indicates Stone House, and BH indicates Brick House. The table also includes the average sound pressure level L_{eq} , and the C- and A-weighted sound pressure levels for the duration of the signature. The annoyance rank column indicates the relative annoyance of the signatures. A ranking of 1 is the most annoying and the ranking of 6 indicates the least annoying signature. The right-hand column provides a brief description of the signature. The signatures varied in both loudness and spectral content over the duration of the recording.

Table 11. Listing of the six signatures used in the spectral balance phase of the subjective study. The left-hand column indicated the file name of the signature and the letter designator used in subsequent analysis. The right-hand column provides a brief description of the signature. Other columns include the unweighted and C- and A-weighted sound pressure levels, all in dB re 20 μ Pa, and the results of relative annoyance of the signatures as rated by the subjects. (1 being the most annoying).

Spectral Balance Signatures (Approx. 30 secs in length)	L_p	L_C	L_A	Annoyance Rank	Description Indoor Microphone
B747 289 1607 BH indoor→ A	75.3	68.0	60.3	2	High LF, moderate HF, window open
A319 292 1557 BH indoor→ B	74.8	54.4	36.4	6	Moderate LF, low HF, low levels, window closed
DC9 292 1703 BH indoor → C	74.3	67.1	63.8	1	High LF, high HF, high levels, window closed
U 294 1814 SH indoor → D	68.5	57.7	47.8	4	Long event, moderate LF and HF, window closed
U 295 0720 SH indoor → E	68.8	52.6	39.6	5	Shorter event, lower LF and HF window closed
U 295 1043 SH indoor → F	68.9	60.3	51.4	3	Long event, moderate LF and HF window closed

Table 12. Listing of the six signatures used in the rattle content phase of the subjective study. The content of the various columns of the table are the same as described in the caption for Table 12.

Rattle/Non-Rattle Signatures (Approx. 45 secs in length)	L_p	L_C	L_A	Annoyance Rank	Description Indoor Microphone
Mult 292 0847 SH indoor→ A	72.6	67.7	47.3	4	LF rumble, rattle at start. Overflight noise at end, window closed
A320 289 0845 SH indoor→ B	70.6	66.0	50.3	3	Quiet start. LF rumble rattle landing at end, window closed
Mult 292 0841 SH indoor → C	76.9	67.9	47.8	5	Quiet rattle start. LF rumble duration, window closed.
Mult 294 1738 SH indoor → D	72.9	66.0	56.0	2	Non-rattle. HF buzz saw tones, window closed
Mult 285 1756 SH indoor → E	66.8	61.3	40.7	6	Non-rattle. Low levels like “C”, window closed.
B747 2891607 BH indoor → F	75.1	71.8	63.4	1	Non-rattle, high LF, moderate HF present, window open.

For the spectral balance study, three of the selected signatures were recorded at the Brick House and three at the Stone House. Care was also taken to ensure that these six signatures exhibited a range of low and high frequency content and level. The design included a selection of signatures that had different qualitative perceptive identifiers (e.g. Signature B sounded “soft”; Signature D sounded “hollow”). Values of several objective metrics calculated from the signatures were used as a selection tool to identify signatures that had values that varied within and across the metrics. The signatures were approximately 30 seconds in duration.

For the rattle jury trial, three of the signatures contained audible rattle and three did not. For each rattle signature there was an attempt to choose a non-rattle signature with an unweighted overall sound pressure level within about ± 2 dB of that of one of the rattle signatures. An open

window signature was also presented to include the condition in which homeowners may have their windows open during a noise event. The signatures were approximately 45 seconds in duration.

One signature was common to both jury trials to provide overlap. Signature F in the audible rattle trial is the same as Signature A of the spectral balance trial, although the playback levels were slightly different. The spectral balance trial included all single event signatures. The rattle trial included four signatures that resulted from “Mult” or combined events. The combined events occurred when the house received noise impact from multiple runways simultaneously. Audible window rattle was more likely to be observed for combined events than for single events.

8.1.2 Subjective Test Methodology

The spectral balance study used signatures with variations in the balance between the low and high frequency content. The rattle study presented signatures that were similar when measured by various noise metrics, but varied as to the presence, or absence, of rattle content. The rattle content was just detectable above the background of aircraft noise. For both tests, some of the differences between signatures were small.

As a result, the subjective test methodology consisted of two judgments, a paired comparison and an assessment of relative annoyance. The method of paired comparison has advantages if fine judgments are required. Subjects generally find choosing one signature over another easier than assigning a rating value to individual signature, thus paired comparisons are a good testing methodology for naïve, untrained subjects.⁵⁰ The test design provides for close comparison between pairs of signatures with the rank sum and the Bradley-Terry-Luce analyses, while affording a relative comparison of the strength of the annoyance in each pair by including the rating scale.

Signatures were presented to subjects to compare in pairs, using the statistical method of paired comparisons. Each jury member was presented with two signatures (A then B) and asked to choose which one was more annoying and then to indicate the degree to which they found it more annoying on a response sheet (see Figure 33). To rate the relative annoyance of the signatures, subjects were asked to place an X on the line to indicate the degree of referenced to the "No Preference" mark. The rating scale used adjectives that anchor points along the line such as “strong preference for A,” “moderate preference for A,” “no preference,” “moderate preference for B,” and “strong preference for B”. The distance of the X from the no preference mark was used in subsequent analysis to determine relative annoyance.

The design was a two alternative forced choice. Subjects were presented with two options and had to make a selection. They were not allowed to mark directly over the "No Preference" line, but could mark as close to it as they chose, as long as they made a choice.

Mark an "X" along the scale to denote which noise is more annoying.



Figure 33. Response sheet used by participants in the subjective trials to indicate which one of a pair of signatures was more annoying and the degree to which it was more annoying. Subjects indicated their rating by placing an X on the line.

Subjects were presented the pairs of sounds in both orders: e.g., Signature A followed by Signature B (A then B), and at some other time in the test, Signature B followed by Signature A (B then A). Six signatures (A through F) were presented in all distinct pairs. Subjects were not asked to compare a signature to itself, i.e., an (A - A) pair was not presented. Because there were 6 signatures in each test, subjects evaluated 30 pairs, hearing each signature 10 times, 5 times as the first signature in the pair and 5 times as the second signature in the presented pair.

The subjective study was counterbalanced, meaning that subjects evaluated each possible pair in both orders. The presentation of the reverse ordered pair for each combination affords an evaluation of ordering effects. If a subject, particularly when comparisons are difficult, always chooses the first signature heard, an order effect will be observed when results are analyzed. Typically results from the (A then B) comparison are averaged with those from the (B then A) comparison to remove this ordering effect. The ordering of the pairs presented was also randomized for each subject, again to reduce any ordering effects. Order effect was evaluated by comparing the findings for Order 1 to the findings for Order 2.

The spectral balance study included 36 subjects and the rattle study included 41 subjects.

8.1.3 Analysis Methodology

The subjects' responses to pair-wise comparison of the signatures were analyzed using three different statistical methods. These analysis methods are summarized in this section. Each method of analysis results in a ranking of the signatures from most to least annoying. The rankings are then correlated with objective metrics calculated for the signatures to determine which metrics are the best predictors of annoyance.

Paired Comparison Rank Sum Method (Based on Signature Selection)

The subjects were forced to choose one of the signatures presented as a pair as more annoying than the other. All possible pairs of the six signatures were presented in both possible orders. The signatures can be ranked based on the number of times a particular subject picked a

particular signature as being the more annoying of a pair. (In fact, three rankings can be made. One each based on the two orders of presentation and a third based on the sum of the rankings for each order.)

The rankings were determined as follows:

1. If the signature was chosen as the more annoying of the pair it was given a score of 2.
2. The remaining signature that was chosen as less annoying was given a score of 1.
3. The scores for each signature were summed across all subjects, resulting in rank sums.
4. The rank sums for the six signatures were then ranked in order. The highest sum was given a rank of 1 (most annoying) and lowest sum given a rank of 6, (least annoying).

This method was applied separately for each of the presentation orders and from the average of the two orders.

Paired Comparison Bradley-Terry-Luce Method (Based on Signature Selection)

The Bradley-Terry-Luce (BTL) method of paired comparisons uses the probability that a particular signature was chosen in a paired comparison to place that signature on an annoyance scale. The Bradley-Terry-Luce method is commonly used in psychoacoustic testing, and assumes that the signatures being compared exhibit the same features.⁵¹ That is, the features used to compare Signature A with Signature B must be the same as those used when comparing Signature A with Signature C. Subjects may change their focus during the test session, altering the features that they use to evaluate the signatures. For instance, if subjects begin to perceive similarities in the signatures, they may choose to disregard subtle differences in those similar features (such as the rattle content), and shift their focus to more distinguishing features (such as loudness).

Relative Strength of Annoyance Rating (Based on Location of Mark on Response Sheet)

Having the subjects indicate their perception of a chosen signature by placing a mark along the line from no preference to strongly annoying, provides a measure of the relative strength of subject's annoyance. Consider Signature A. The relative strength of annoyance of a Signature A by a particular subject is determined by first measuring the distance from the no preference mark to the mark the subject placed on the response sheet each time Signature A was judged to be the more annoying of a pair. Summing up the distances, from the subject's response for each time they evaluated Signature A results in the relative strength annoyance of Signature A, for that particular subject. The process is repeated for each of the remaining five signatures. Summing up the relative strengths of annoyance across all subjects results in a total relative strength of annoyance.

The following steps were used to determine the degree to which a subject found a particular signature annoying, called the strength of annoyance.

1. For each pair, the subjects marked the response scale to identify the degree to which a signature was more annoying.

2. The distance along the line from ‘no preference’ was measured with a digital caliper.
3. The distance was converted to a percentage of total distance along the line from no preference to strongly, for each pair.
4. The percentages were summed for each subject for each signature in each order
5. The sum for each signature was divided by 5 (the number of times that signature was presented in each order) to get an average value for each signature for each subject
6. The average values were summed across all subjects and divided by the number of subjects to get the average across all subjects.

8.1.4 Results

8.1.4.1 Spectral Balance Study

The results of these analyses for the spectral balance study are shown in Table 13. The upper section of the table contains the values calculated for the three measures of annoyance. The lower section contains the rankings based on each value. A rank of 1 means the signature was found most annoying. A rank of 6 means the signature is least annoying. For each trial, the rankings were the same for each of the measures of annoyance.

The values in the strength of annoyance column are the averages for a particular signature across all subjects, normalized by the length of the line on the response sheet. The rank sum is a measure of the number of times a particular signature was chosen as the more annoying of a pair, summed across all subjects. The Bradley-Terry-Luce (BTL) values are calculated according to the algorithm discussed in Ref. 52.

The values of the relative strength of annoyance for the six signatures are plotted in Figure 34. Results for the two presentation orders are shown in the left and middle panels. The panel on the right is the average of the two order values. The letters next to the symbols refer to the designations of the six signatures used in Table 11. The most noticeable difference between the results for the two orders of presentation is for Signatures C and A. However, each presentation order results in the same ranking C, A, F, D, E, B from most to least annoying. The annoyance rating of the signatures is relative to one another, not absolute.

The results of the rank Sum analysis are shown in Figure 35. There is less of an order effect than with the relative strength of annoyance analysis. Both methods result in the same order.

The average values of the relative strength on annoyance and rank sum from the right hand panels of Figures 34 and 35 are compared to the results of the Bradley-Terry-Luce analysis from Table 13. The BTL analysis results in the same ordering of the signatures as the other two methods of analysis and the relative spacing of the values compare favorably.

Table 13. Results of analysis of subjective responses to the spectral balance jury trial. The upper section contains the values of the different measures of annoyance. The lower section contains the rankings that result from those values. A rank of 1 (6) means most (least) annoying.

Signature	Order One		Order Two		BTL
	Strength of Annoyance	Rank Sum	Strength of Annoyance	Rank Sum	

A	0.265	301	0.486	334	1.543
B	0.004	195	0.034	195	-2.829
C	0.353	322	0.635	356	1.914
D	0.101	247	0.186	285	0.052
E	0.034	210	0.087	239	-1.193
F	0.257	261	0.264	304	0.513

Signature	Order One		Order Two		BTL
	Strength of Annoyance	Rank Sum	Strength of Annoyance	Rank Sum	
A	2	2	2	2	2
B	6	6	6	6	6
C	1	1	1	1	1
D	4	4	4	4	4
E	5	5	5	5	5
F	3	3	3	3	3

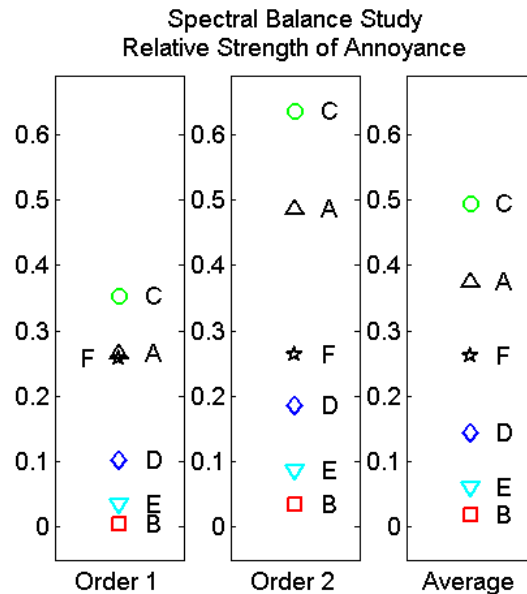


Figure 34. Values of relative strength of annoyance for the spectral balance study from Table 13. Results for the two presentation orders are shown in the left and middle panels. The panel on the right is the average of the two order values. The letters next to the symbols refer to the designations of the six signatures used in Table 11.

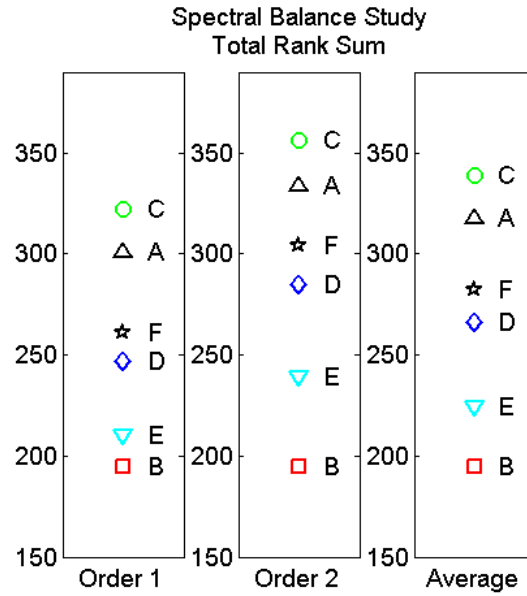


Figure 35. Values of rank sum for the spectral balance study from Table 13. Results for the two presentation orders are shown in the left and middle panels. The panel on the right is the average of the two order values. The letters next to the symbols refer to the designations of the six signatures used in Table 11.

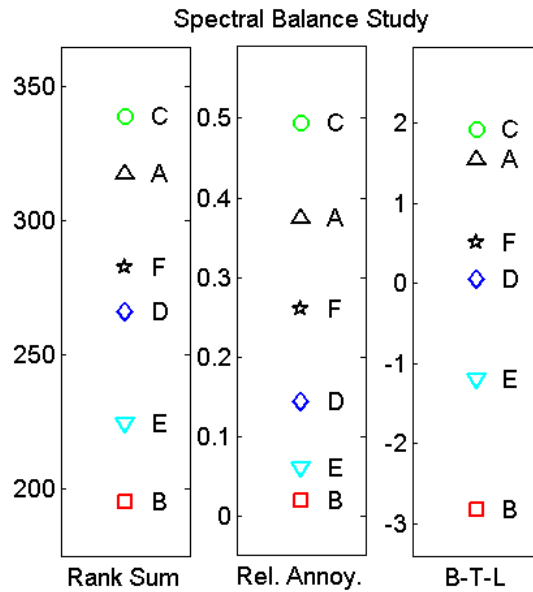


Figure 36. Comparison of the average values of relative strength of annoyance and rank sum to the values of the Bradley-Terry-Luce analysis for the spectral balance study.

8.1.5 Audible Rattle Study

The results of the analyses for the audible rattle study are shown in Table 14. Comparisons of the three methods of analysis are shown in Figures 37-39. The conclusions with regards to order effect and order are similar to those found in the spectral balance study. Recall that Signature F in this trial is the same signature as A used in the spectral balance study. The ordering of the signatures is the same for all three methods of analysis and for both orders of presentation. The order is F, D, B, A, C, E from most to least annoying.

The signatures that contained a rattle event (A, B, and C) were not ranked as the most annoying. This finding may, perhaps, be attributed to the fact that the rattle while audible is not particularly loud. While rattle is clearly detectable when listening to the signatures, the energy contributed by the rattle is a small percentage of the energy in the whole aircraft event. Broadband energy-based metrics would then be relatively insensitive to this rattle presence. The subjects' perceptions of the signatures were insensitive to rattle, indicating that their judgments were primarily based on the overall level of the sound. In a home situation, sensitivity to rattle may increase, as it could be associated with potential property damage. Powell and Shepherd cited a similar result in a laboratory test.¹²

It would have been possible to boost the relative levels of rattle to determine at what point rattle becomes the determining factor in perceived relative annoyance. However, this study used the relative levels of the rattle and the aircraft noise as recorded in the houses. The houses used in this study were constructed of brick and stone. Less substantial construction types than the houses in this study could experience a higher level of rattle than was recorded in this study. A boosted rattle level study could be useful in predicting a threshold of annoyance for rattle in a background of noise.

The three signatures with rattle were also grouped together and in the same order relative to the other three signatures, regardless of analysis method.

Table 14. Results of analysis of subjective responses to the audible rattle jury trial. The upper section contains the values of the different measures of annoyance. The lower section contains the rankings that results from those values. A rank of 1 (6) means most (least) annoying.

Signature	Order One		Order Two		BTL
	Strength of Annoyance	Rank Sum	Strength of Annoyance	Rank Sum	
A	0.071	278	0.134	304	-0.393
B	0.125	300	0.176	323	0.002
C	0.064	277	0.063	264	-1.064
D	0.356	370	0.254	351	1.470
E	0.004	213	0.007	216	-2.732
F	0.635	407	0.421	387	2.716

Signature	Order One		Order Two		BTL
	Strength of Annoyance	Rank Sum	Strength of Annoyance	Rank Sum	
A	4	4	4	4	4
B	3	3	3	3	3
C	5	5	5	5	5
D	2	2	2	2	2
E	6	6	6	6	6
F	1	1	1	1	1

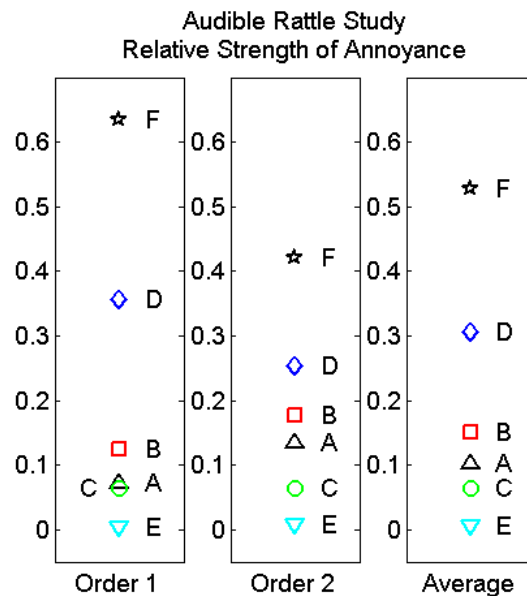


Figure 37. Values of relative strength of annoyance for the audible rattle study from Table 14. Results for the two presentation orders are shown in the left and middle panels. The panel on the right is the average of the two order values. The letters next to the symbols refer to the designations of the six signatures used in Table 12.

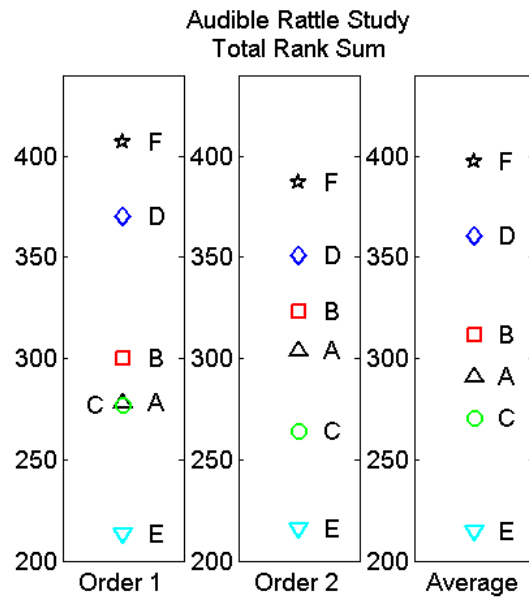


Figure 38. Values of rank sum for the audible rattle study from Table 14. Results for the two presentation orders are shown in the left and middle panels. The panel on the right is the average of the two order values. The letters next to the symbols refer to the designations of the six signatures used in Table 12.

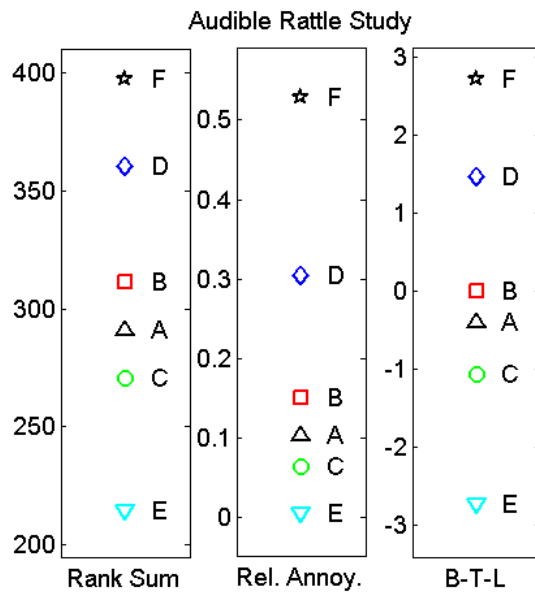


Figure 39. Comparison of the average values of relative strength of annoyance and rank sum to the values of the Bradley-Terry-Luce analysis for the audible rattle study.

8.1.6 Objective Metrics of Signatures Used in the Subjective Studies

The objective metrics listed in Section 4.4 were calculated from recordings made inside the simulator for each of the signatures used in the subjective trials. The results are shown in Tables 15 and 16 for the spectral balance and audible rattle studies, respectively. In some cases a metric was used in only one of the subjective trials. The * listed under Signatures B and E for *PNL* indicates that the levels of these signals were too low to calculate *PNL* according to ISO 3891.

Table 15. Values of metrics for spectral balance signatures as recorded upon playback in simulator. Levels are expressed in units of dB (or dBA or dBC) re: 20 μ Pa. The unit for Loudness is Sones. (All metrics calculated using MATLAB except for Loudness (Zwicker), which was calculated by using the HEAD Acoustics system.) * Indicates that levels were too low to calculate *PNL* for these signatures.

	A	B	C	D	E	F
<i>L_{max}</i>	89.2	79.3	91.4	79.0	76.8	80.6
<i>L_{Amax}</i>	80.5	58.6	84.8	67.1	61.1	70.4
<i>L_{Cmax}</i>	86.9	75.8	89.8	77.5	74.1	79.3
<i>L_{eq}</i>	71.8	61.9	68.9	61.4	59.8	63.2
<i>L_{Aeq}</i>	60.0	37.5	63.4	46.8	40.1	50.5
<i>L_{Ceq}</i>	68.8	58.3	67.4	59.7	56.8	61.9
<i>LFNR</i>	67	42	69	53	47	57
<i>LFSL</i>	77.5	68.4	72.7	65.9	65.2	67.9
<i>L_{LF}</i>	64.6	55.3	61.6	59.1	55.4	61.6
<i>Loudness (Zwicker)</i>	7.5	2.3	7.7	4.2	2.9	4.9
<i>LLSEL</i>	63.9	42.8	68.6	54.8	45.7	59.0
<i>Q_A ($\alpha = 1/10$)</i>	58.6	37.0	61.4	46.1	39.6	49.5
<i>Q_{PNL} ($\alpha = 3/40$)</i>	65.2	26.1	68.2	50.2	33.6	55.7
<i>Q_{PNL} ($\alpha = 1/10$)</i>	66.6	31.0	70.2	51.7	37.0	57.3
<i>PNL</i>	66.4	*	70.0	48.8	*	57.1

Table 16. Values of metrics for audible rattle study signatures as recorded upon playback in simulator. Levels are expressed in units of dB (or dBA or dBC) re: 20 μ Pa. The unit for Loudness is Sones. (All metrics calculated using MATLAB except for Loudness (Zwicker), which was calculated by using the HEAD Acoustics system.) * Indicates that levels were too low to calculate PNL for these signatures.

	A	B	C	D	E	F
L_{max}	89.2	84.3	86.6	84.0	80.1	91.3
L_{Amax}	65.7	66.7	65.7	74.0	56.7	83.0
L_{Cmax}	84.1	81.2	84.1	82.8	75.8	89.0
L_{eq}	71.8	68.4	70.1	67.0	65.6	74.9
L_{Aeq}	47.4	50.3	47.8	56.0	40.8	63.4
L_{Ceq}	67.7	66.0	67.7	65.8	61.3	71.8
L_E	88.3	83.9	85.0	82.2	80.9	90.5
L_{AE}	63.9	65.8	62.6	71.1	56.1	79.0
L_{CE}	84.3	81.5	82.5	80.9	76.7	87.4
L_{FSL}	84.4	82.0	85.4	80.2	77.1	87.8
$LLSEL$	54.0	58.5	56.2	64.9	44.8	67.5
$Q_A (\alpha = 1/10)$	47.1	49.9	47.5	55.4	40.7	62.6
$Q_{PNL} (\alpha = 3/40)$	49.7	53.4	49.5	61.8	*	66.8
$Q_{PNL} (\alpha = 1/10)$	50.6	53.9	50.6	62.8	*	68.0
PNL	49.0	56.8	52.5	66.1	*	69.9

8.1.7 Correlation of Objective and Subjective Rankings

A ranking (e.g. from loudest to quietest) of each signature was formed based on the values of the objective metrics presented in Tables 15 and 16. The hypothesis, that the objective and subjective rankings will be correlated, was tested to determine which objective metrics are the best predictors of the order of the subjective rankings. Only the rank order is used in this analysis, not the values of objective metrics or the measures of annoyance of the individual signatures.

The Spearman Rank Correlation Coefficient (R_s) was used to determine the strength of the correlation. The Spearman Rank Correlation Coefficient (R_s) is a non-parametric measure of association that is used in the psychoacoustic community as well as across other research disciplines.^{53,54} Like the Pearson Correlation Coefficient, it is a statistical measure used to assess the amount of discrepancy, or agreement, between two sets of rankings. The Spearman Rank Coefficient measures whether or not the two rankings are the same. The confidence level of 95% was selected for the correlation, indicating that the probability that the resulting values of R_s could have occurred by chance are 5% or less. The 95% confidence level is applied by setting a parameter known as the critical value of $\alpha = 0.05$, the significance level, and is chosen prior to

conducting the test. We selected a fixed significance of $\alpha = 0.05$. (Note that α is the nomenclature typically used to represent the significance level in tables providing critical values of test statistics, and is not to be confused with the free parameter defined for the calculation of the Disturbance Index Q .) With six signatures ($N = 6$) the value of R_s must equal or exceed 0.829, to attain a 95% confidence level. P-values, or the attained significance level, are the level of significance achieved during the test. Smaller P-values correspond to test results having higher levels of statistical significance.

8.1.7.1 Findings from the Spectral Balance Study

Table 17 shows the metrics that had the best correlations with the subjective rankings for the Spectral Balance trial. The Spearman Rank Correlation Coefficient R_s for the twelve metrics included in the table equaled or exceeded the critical value of 0.829 for a 95% confidence level. Seven metrics produced perfect rank correlation, meaning that the order of the objective and subjective rankings were identical.

The rank-order correlation study of the Spectral Balance test results showed that:

1. A variety of metrics showed good correlation between the orders based on the objective metrics and those resulting from the subjective jury trials.
2. Only three of the metrics listed in Table 16 failed to achieve the critical value of R_s , specifically L_{eq} , L_{FSL} , and PNL .
3. L_{FSL} , a metric that is specifically designed for low-frequency noise, fell below the specified R_s critical value, while several other low-frequency-specific metrics rated above the critical value.

Table 17. Correlation analysis between rankings based on subjective test results and objective metrics for the Spectral Balance trial. Metrics having Spearman Rank Correlation Coefficients above critical value of 0.829 are included (critical value for $\alpha = 0.05$ and $N = 6$).

Metrics with Perfect Correlation		Metrics with High Correlation	
Metric	R_s Value	Metric	R_s Value
L_{Amax}	1.00	L_{Cmax}	0.943
$LFNR$	1.00	L_{Aeq}	0.943
$LLSEL$	1.00	L_{LF}	0.899
QA	1.00	L_{Ceq}	0.886
$QPNL (\alpha = 1/10)$	1.00	L_{max}	0.829
$QPNL (\alpha = 3/40)$	1.00		
$Loudness$	1.00		

8.1.7.2 Findings from the Audible Rattle Study

Table 18 shows the Spearman Rank Correlation Coefficients for all the metrics listed in Table 16 for the audible rattle trial. The metrics for which the correlation between the objective and subjective rankings reaches the 95% confidence level, corresponding to $R_s \geq 0.829$, are grouped above the heavy line. (In the case of two of the metrics (L_{maxC} and L_{eqC}), there were ties in the rankings of some of the signatures. In these cases the tied Spearman Rank Correlation Coefficient was used.)

Table 18. Correlation analysis between rankings based on subjective test results and objective metrics for the audible rattle trial. The value of the Spearman Rank Correlation Coefficients necessary to achieve a 95% confidence level is 0.829.

Metric	R_s
$Q_{PNL} (\alpha = 3/40)$	1.00
L_{AE}	1.00
L_{Amax}	0.99
$Q_{PNL} (\alpha = 1/10)$	0.97
LLSEL	0.94
$Q_A (\alpha = 3/40)$	0.94
L_{Aeq}	0.94
PNL	0.90
L_{maxC}	0.52
L_E	0.49
L_{CE}	0.49
L_{eq}	0.49
L_{max}	0.49
L_{Ceq}	0.46
LFSL	0.43

The Rank-Order correlation studies in the rattle test showed that:

1. Rankings based on $Q_{PNL} (\alpha = 3/40)$ and L_{AE} were perfectly correlated with the rankings derived from the subjective test results.
2. The rankings based on A-weighted metrics, Disturbance Indices and LLSEL were highly correlated with the ranking from the subjective tests.
3. The order based on unweighted and C-weighted metrics failed to attain the 95% confidence level.

In addition to the Spearman Rank ordinal data analysis, the interval data obtained in the strength of annoyance trials was analyzed using linear regression analysis. Some values obtained were as follows: Q_{PNL} resulted in a correlation of 0.958, L_{EA} was 0.923, $LLSEL$ was 0.822, L_{EC} was 0.445, and L_{FSL} had a correlation of 0.266. The Linear regression correlation values followed the same general pattern as the Spearman Rank correlations, although the correlations were not as strong across the metric set, due to the increased analysis resolution provided by the use of interval data.

8.1.8 Comparison of Objective and Subjective Grouping of Signatures

The focus of the previous section was on the correlation of the orders the signatures were ranked by the objective metrics and subjective ratings. Several objective metrics correlated with the subjective ranking. This type of analysis is useful as a first pass for determining which metrics warrant further investigation. A more stringent test for which objective metrics best predict subjective perception is based on the values of the metrics and the relative groupings of the signatures based on them.

8.1.8.1 Spectral Balance Study

The results of the order correlation analysis for the spectral balance study were presented previously in Table 17. The metrics L_{Amax} , $LFNR$, $LLSEL$, Q_A ($\alpha=3/40$), and Q_{PNL} ($\alpha=3/40$) all had perfect correlations (i.e., a Spearman Rank Correlation Coefficient $R_S = 1.00$) with the order as ranked by the subjects evaluating the signatures. The values of these five metrics are plotted in Figure 40, along with the rank sum and Bradley-Terry-Luce values. Several other metrics had R_S values which also satisfied the 95% confidence level. One of these metrics, L_{Ceq} , is also shown in Figure 40. The L_{Ceq} findings were statistically significant but the correlation was not as strong ($R_S = 0.886$).

The purpose of this figure is to assess which of the objective metrics best predicts, not only the order of the subjective ratings (as expressed by the rank sum and B-T-L values), but also the grouping and spacing. In order to facilitate this assessment, the ranges of the vertical scales have been adjusted so that the data spans about the same vertical distance in the plots.

The first observation to make is that the order of the signatures is the same in all the plots, with the exception of L_{Ceq} , the only metric in the figure that did not have a perfect correlation in Table 17. Another observation is, after excluding L_{Ceq} , the other objective metrics displayed similar groupings of the signatures: C with A, F with D, and E with B. The rank sum shows the same groupings. The BTL analysis grouped C with A, and F with D, but the E with B grouping is not as tight with the other metrics. If one concentrates on the four most annoying signatures (C, A, F, and D), both the groupings and spacings of the two subjective ratings and the five objective agree well.

These findings indicate that the objective metrics with perfect correlations in Table 17 also do a good job of predicting the relative spacing of the subjective ratings. Metrics from each of the four categories of candidate metrics discussed in Section 7.1, level-based, loudness-based, low frequency, and nosiness, showed perfect correlation and compare well with the subjective

ratings.

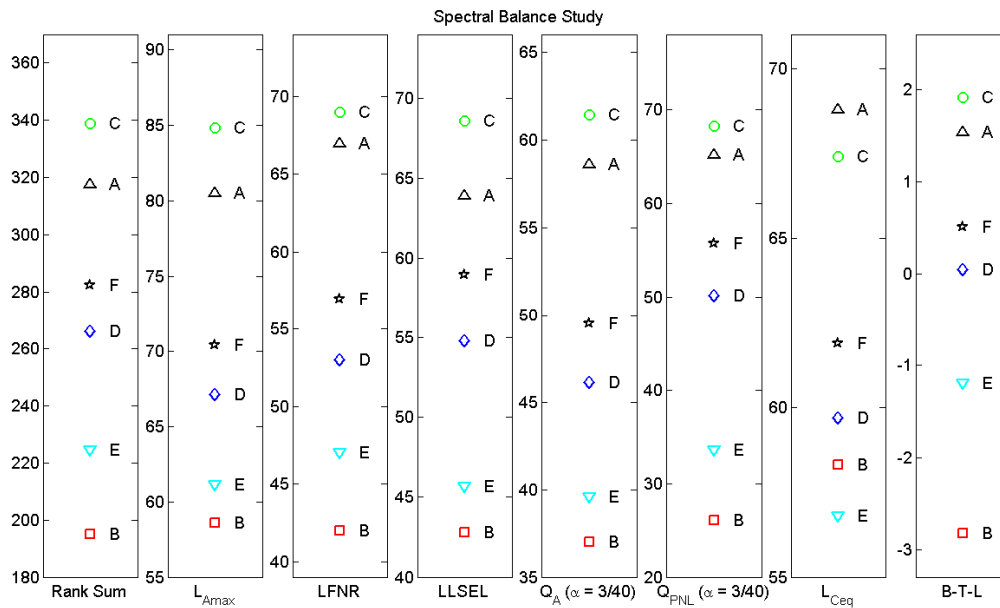


Figure 40. Comparison of objective metrics that correlated well with the subjective ordering for the spectral balance study. The scales have been adjusted so that the range of values of each metric fills the vertical scale. The rank sum and Bradley-Terry-Luce subjective values are shown in the left and right panels, respectively.

8.1.8.2 Audible Rattle Study

The results of the order correlation analysis for the audible rattle study were presented previously in Table 18. The metrics Q_{PNL} ($\alpha=3/40$) and L_{EA} showed perfect correlation. Other highly correlated metrics include L_{Amax} , L_{Aeq} and $LLSEL$. The order based on L_{Ceq} was poorly correlated with the subjective ordering. The values of these six metrics are plotted in Figure 41, along with the rank sum and Bradley-Terry-Luce values. It should be noted that because of the low signal level, a value could not be calculated for Q_{PNL} for Signature E.

As was the case with Figure 40, the purpose of Figure 41 is to assess which of the objective metrics best predicts, not only the order of the subjective ratings (as expressed by the rank sum and BTL values), but also the grouping and spacing. The ranges of the vertical scales have been adjusted so that the data spans about the same vertical distance in the plots.

All the metrics have signatures F and E at the far ends of the scale. Signature F was recorded with an open window and was the loudest signature. Also, with the exception of L_{Ceq} , which had a very low correlation coefficient, all the metrics produced the same order as the subjective ratings. The purpose of including L_{Ceq} in the figure is to demonstrate how a poorly correlated metric compares to a more highly correlated one. The statements in the remainder of this section

will assume that L_{Ceq} is excluded from consideration.

Signatures A, B, and C all contained audible rattle and were described as having a low frequency rumble characteristic, but they were not found to be the most annoying. Nevertheless, all the metrics, objective as well as subjective, grouped these three signatures together. Groupings of signatures with similar characteristics such as rattle helps one to understand the ability of particular metrics to not only correlate with the subjective rankings, but also may aid in identifying characteristics within signatures that lead to such ranking.

Inspection of Figure 41 indicates that the rank sum, BTL, L_{AE} , L_{Aeq} , Q_{PNL} ($\alpha=3/40$), and $LLSEL$ agree well with one another with respect to the relative spacing of the signatures.

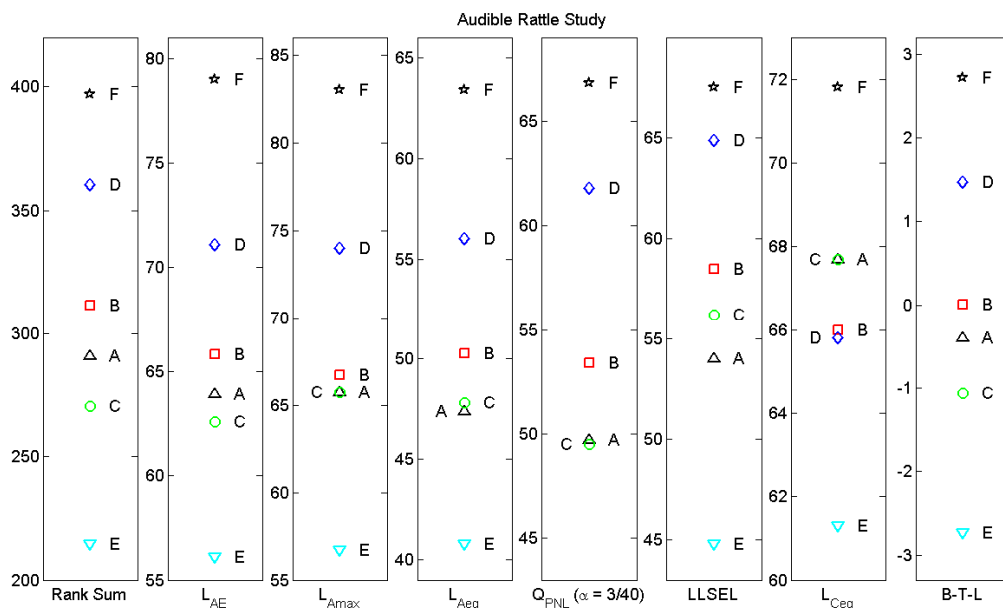


Figure 41. Comparison of objective metrics that correlated well with the subjective ordering for the audible rattle study. The scales have been adjusted so that the range of values of each metric fills the vertical scale. The rank sum and Bradley-Terry-Luce subjective values are shown in the left and right panels, respectively.

8.1.9 Comparison of Spectral Balance Signatures to Signatures obtained near Florida Airport

The spectral balance signatures were also used in a test conducted at Purdue's Herrick Laboratory where sounds were played over Etymotics ER-2 in-ear phones while subjects sat in a sound proof booth. These high fidelity in-ear phones are used in testing because of their relatively flat frequency response. The sounds were presented along with 15 modified recordings of two aircraft measured at a Florida airport that were at the same level or louder than the Dulles recordings. Subjects rated sounds directly rather than choosing between two sounds; they were

asked to place a mark on a five point scale labeled: not at all (2.0), slightly (3.5), moderately (5), very (6.5), and extremely annoying (8.0). (The numbers in parentheses are those assigned to the position of the words on the line.) The subjects in this test were asked to imagine that they were outside, rather than inside a house, because the other sounds were based on outside recordings. Signature C was the most annoying, closely followed by A. The average annoyance ratings for these sounds were below “moderately annoying” (around 4.4). Signatures F and D were rated below “slightly annoying” (around 3), while Signatures E and B were rated, on average, as very close to “not at all annoying.” If these had been judged as interior sounds, these may well have been found to be more annoying. The Florida sounds were 30 seconds in duration and were calibrated to be either around 59 dB (A) or 68 dB (A). This test thus resulted in the same annoyance rankings and similar relative annoyance ratings spacing as in the spectral balance study (C, A, F, D, E, B) described above, with the additional information on where the ratings lie on the scale from “not at all annoying” to “extremely annoying”. The additional information obtained in this test is very similar to that provided in the relative strength of annoyance analysis, which also used a five point scale to mark annoyance combined with a paired comparison selection.

8.1.10 Correlation of Results with the Spectrum of the Signature

The general conclusions drawn from the two subjective studies were that the order derived from the annoyance ratings of the signatures correlates with the orders predicted from metrics falling into different classes, namely level-, loudness-, low frequency-, and noisiness-based metrics. Of the level based metrics, A-weighted metrics are generally better than C-weighted metrics. Insight into the reasons for this finding can be found by examining the 1/3 octave band spectra of the signatures used in the subjective tests. These spectra are shown in Figures 42 – 44.

The unweighted spectra are shown in Figure 42. The shapes of the spectra are bimodal, peaking in the low frequencies in the 20 – 40 Hz range and in the high frequencies in the 500 – 1000 Hz range. The reason for the dip between the two peaks is ground reflection. The highest levels are found at the lowest frequencies.

The A- and C-weighted spectra for the spectral balance signatures are shown in Figure 43. The differences between the two weightings are evident. Limiting ones attention to the region around 1 kHz in the A-weighted spectrum (on the left), the order of annoyance indicated by the number in the parenthesis in the legend follows the level of the signatures. The same behavior is found in Figure 44 which shows the A-and C-weighted spectra for the audible rattle signatures.

The correlation between the A-weighted spectra and the rank order of the signatures does not explain why A-weighting works well. However, the finding is consistent with those of prior studies^{47,48} that, in the absence of a dominant feature of a noise signature, loudness of the event inside the house, is the most important characteristic. The signatures used in these studies were broadband with no dominant tonal content. The results do indicate the complicated nature of determining how the low frequency content of broadband aviation noise signatures influences perception.

If one compares the levels in the 20 – 250 Hz range in Figures 42-44 with the Tokita and Nakamura thresholds in the left panel of Figure 2, the low-frequency levels of the signatures

rated as being most annoying fall into the annoying region in some of the 1/3-octave bands. This comparison implies that the Tokita Nakamura thresholds may apply to low frequency aviation signatures. However, the thresholds are based on exposure to single tones or narrow bands of noise. A useful follow-on study would be to conduct a subjective study in which the high frequency part of the spectrum is the same for all the signatures, but the low frequency part is varied to determine thresholds, such as the Tokita and Nakamura thresholds, for aviation noise.

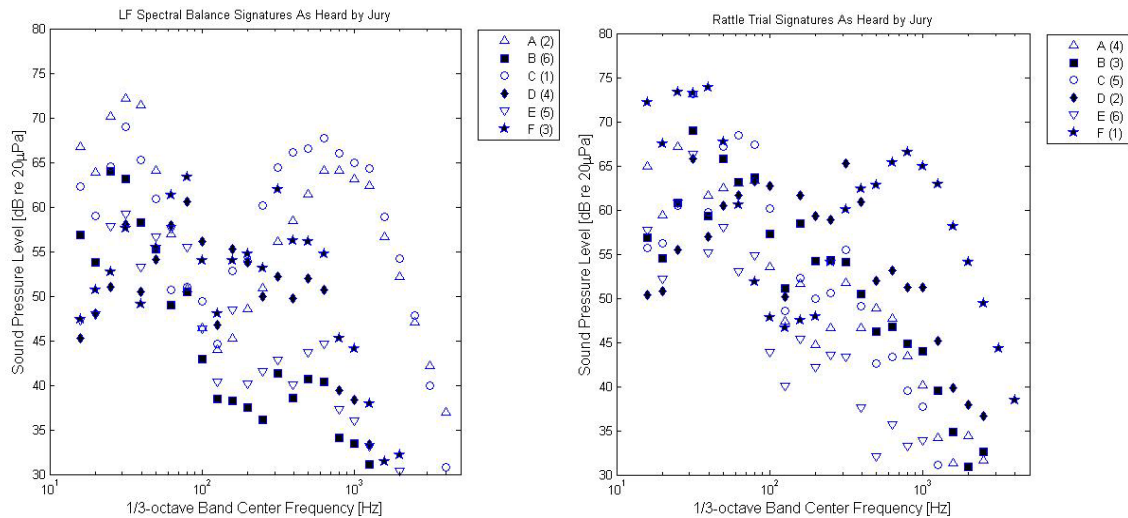


Figure 42. 1/3-Octave Band Levels of the signatures used in the spectral balance (left) and audible rattle (right) studies. Letters in the legend denote which signature, (numbers) indicate order of annoyance (1 = most annoying).

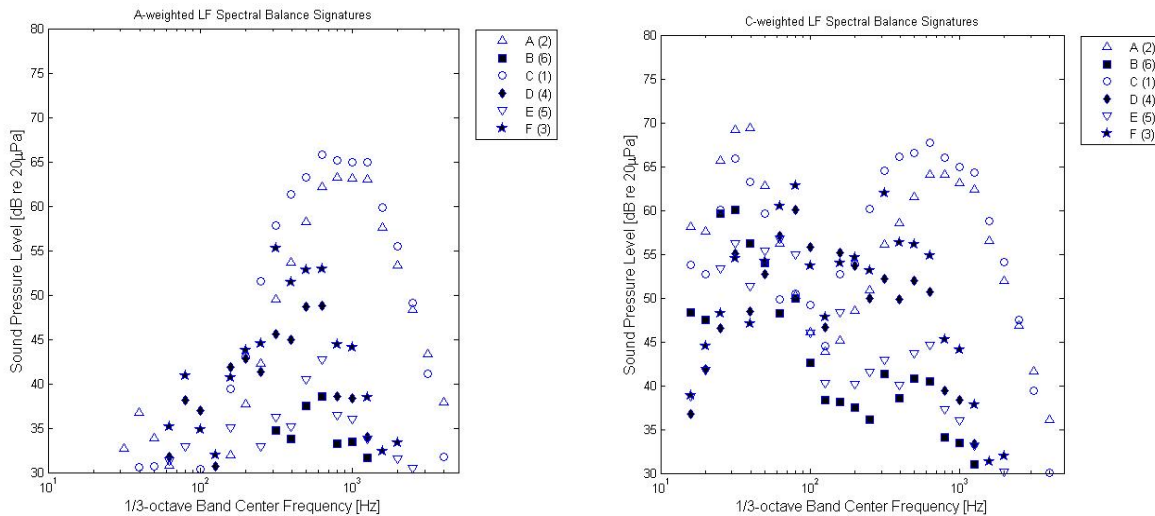


Figure 43. 1/3-Octave Band Levels A-weighted (left) and C-weighted (right) spectra of the signatures used in the spectral balance study. Letters in the legend denote which signature, (numbers) indicate order of annoyance (1 = most annoying).

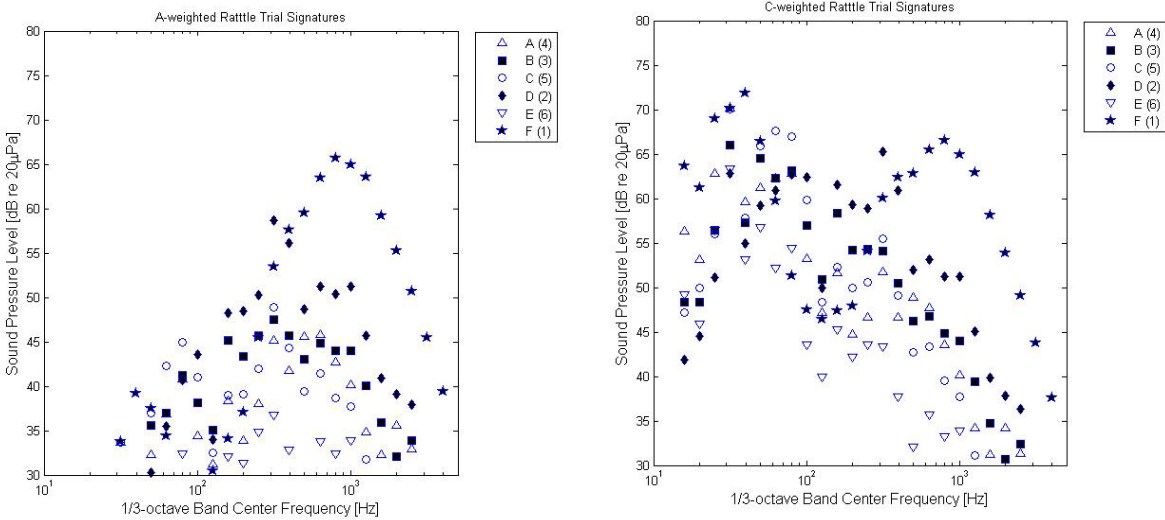


Figure 44. 1/3-Octave Band Levels A-weighted (left) and C-weighted (right) spectra of the signatures used in the audible rattle study. Letters in the legend denote which signature, (numbers) indicate order of annoyance (1 = most annoying).

8.2 Ability of Tokita Nakamura Threshold to Prediction Aviation Noise Impact

This study was conducted to assess the ability of the Tokita Nakamura thresholds^{29,30} to predict subjective perception of aviation noise signatures. Twenty-three signatures recorded at IAD were evaluated based on flat-weighted level, quality of recording, and low frequency noise content. Three signatures were selected as the basis for this study. The original signatures were modified to create additional signature variants that differed in low-frequency noise content. The three baseline signatures were chosen based on the presence of four criteria: a consistent pressure versus time envelope, free of overt extraneous stimuli such as rattle, the presence of a significant difference between their A- and C- weighted levels, and their levels needed to closely approach or cross the Tokita and Nakamura low-frequency annoyance thresholds.

Four variants of the three signatures, resulting in a total of 15 signatures, were used in the jury trial. The variations of the signatures were designed to exceed the Tokita and Nakamura thresholds in specific regions. Figure 45 shows a more detailed version of the Tokita and Nakamura thresholds than when these thresholds were first presented in Section 3 of this report. The signature synthesis process utilized a set of low-pass, band-pass, and high-pass filters implemented in the time domain, which isolated the different target frequency ranges. The high-pass filter was implemented to remove the target low-frequency range. This provided a variant of each of the signatures, which could be used as a counterpoint in the jury trial, testing whether or not low-frequency content adds to or detracts from annoyance. An additional signature was filtered through the High – Pass filter that resulted in a signature with the low-frequency range removed. This High – Pass signature was designed such that the A-weighted level would match the original signature’s A-weighted level within ± 1 dB. These 15 signatures were then presented to subjects in the Gulfstream Supersonic Acoustic Signature Simulator SASS-II. The signatures

as recorded in the simulator are presented in Figures 46-48 along with the Tokita and Nakamura thresholds.

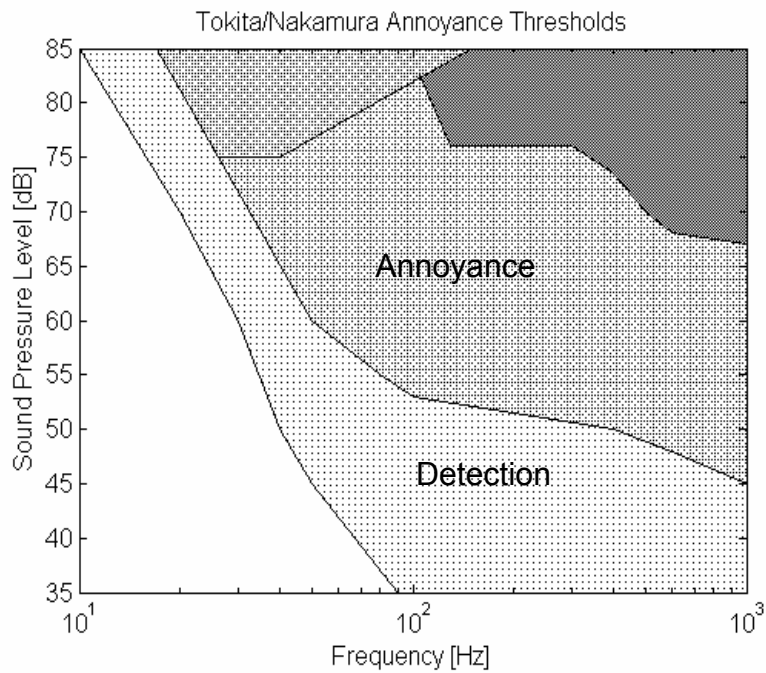


Figure 45. Tokita and Nakamura annoyance thresholds.

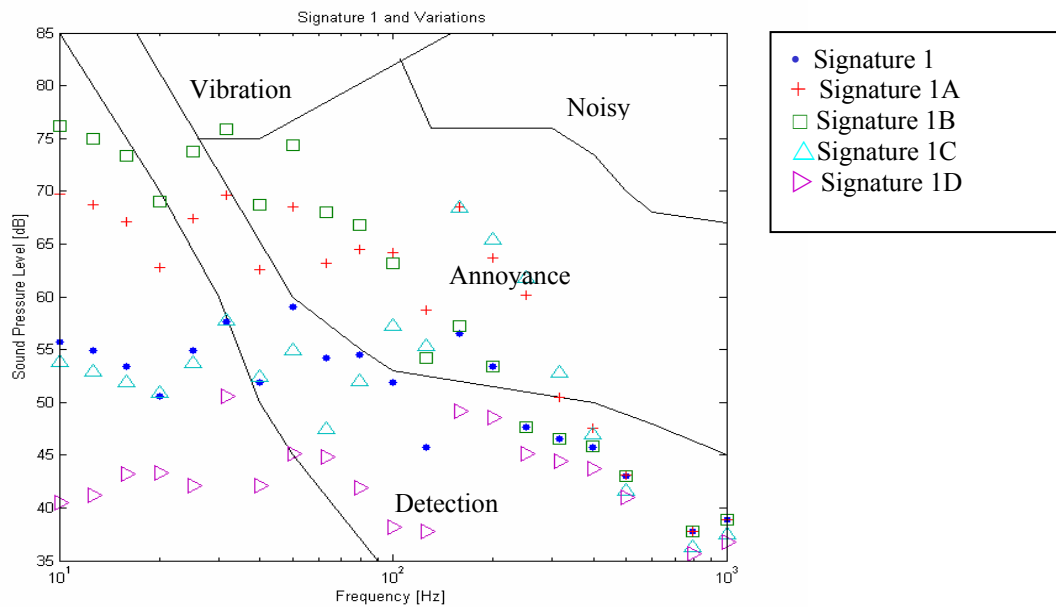


Figure 46. Signature Set 1. Original signature based on IAD dataset U_292_1808_SH.

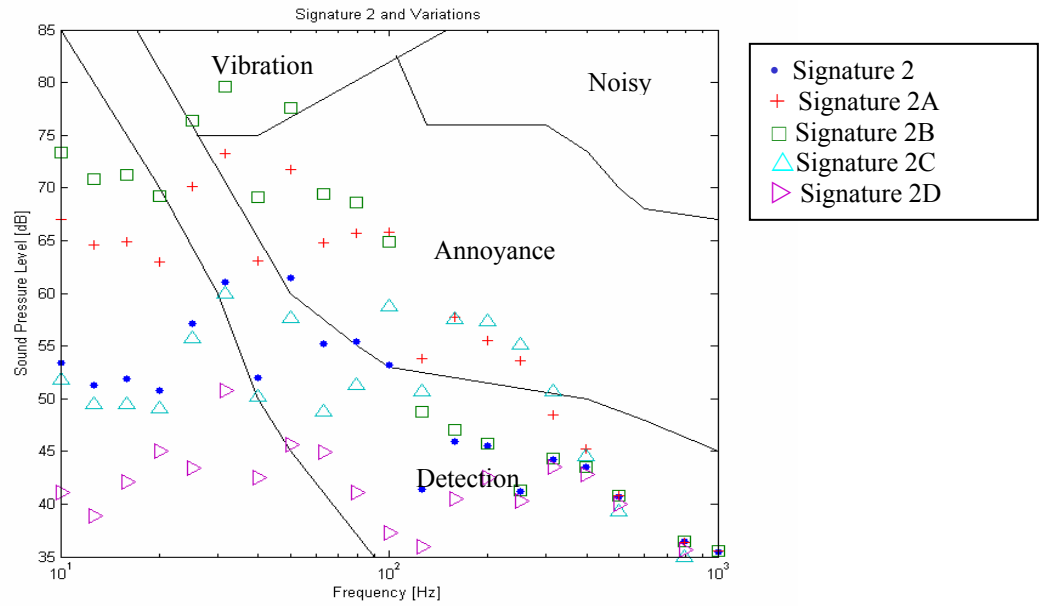


Figure 47. Signature Set 2. Original signature based on IAD dataset A320_289_0844_SH.

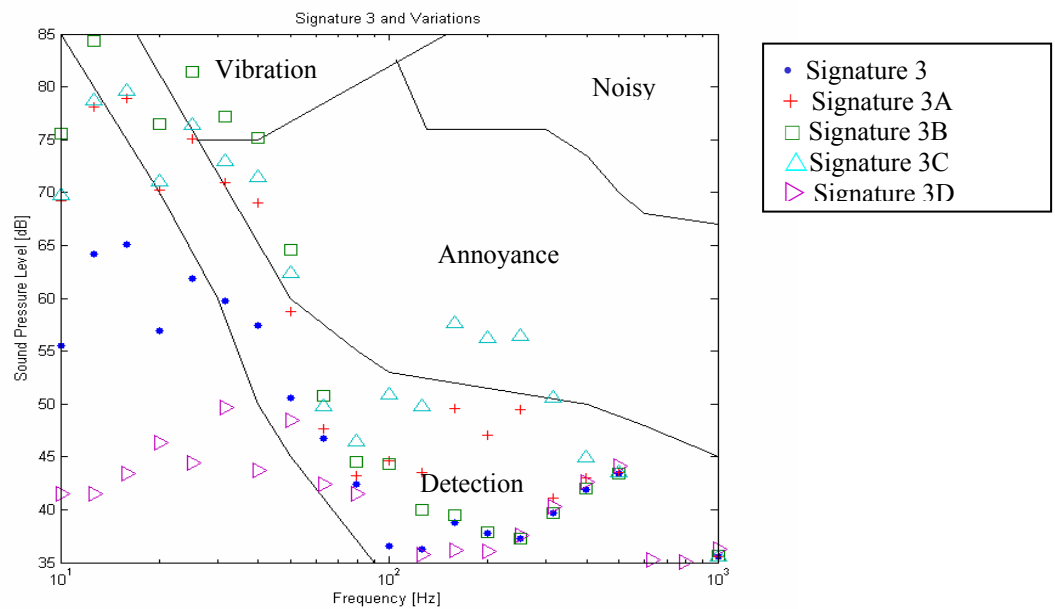


Figure 48. Signature Set 3. Original signature based on IAD dataset B747_292_1612_BH.

8.2.1 Experimental Design

The 15 signatures previously described were presented to 34 subjects in the Gulfstream simulator. The subjects were a diverse mixture of students and professionals solicited from the community around State College, PA. There were 23 male and 11 female participants in this study with an age range of 19 to 58 years of age with an average age of 31 years.

Participants in the jury trial completed informed consents and a demographic survey. The survey also asked questions concerning their attitudes toward and tolerance of low-frequency noise and aircraft noise among other topics. The participants were instructed to describe a time they felt relaxed at home. They were then instructed to listen to the noises in this mood as if they are relaxing at home. Pre-trial hearing test were conducted to assess the hearing sensitivity of participants (hearing threshold level < 25dB in frequency range of interest).

The signatures were presented to the subjects counterbalanced and replicated to detect ordering effects. This resulted in subjects judging a total of 60 pairs of signatures. Subjects were broken into four groups; each group received a different random presentation of the signatures. Subjects were directed to listen to each noise and rate it for annoyance on the given line scale labeled with the five annoyance descriptors recommended by the International Commission on the Biological Effects of Noise (ICBEN)⁵⁵ as verbal anchors for the range of response. Subjects were instructed to then select the signature of each pair that was more annoying. Subjects were given a break between each signature set to prevent fatigue. Each trial lasted about 90 minutes, including breaks. After the test subjects were asked to verbally answer a few questions concerning if they would predict a shift in their annoyance threshold due to repetitive exposure.

8.2.2 Data Analysis and Results

Three methods were used to analyze the data, the Bradley-Terry-Luce method, the rank sum method, and strength of annoyance. Only the results of the strength of annoyance analysis are presented here. The strength of annoyance means the degree to which a subject found a particular signature annoying. Subjects were asked to indicate the degree to which they found each signature annoying by marking an X along the line scale. The distance of each mark from the "not at all annoying" label was measured using a digital calibrator.

An example of the results of the strength of annoyance analysis is shown in Figure 49 for Signature Set 1. The vertical scale represents strength of annoyance with the ICBEN⁵⁵ descriptors displayed in place of the numerical values. The data for each of the signature variants is presented along the horizontal scale. The line in the center of each box represents the median of the data, the value shown is the mean of the data and the distance between the two show the skew of the data. The boxes represent the interquartile region; 75% of the data points lie below the top of each box and 25% of the data points lie below the bottom line. The whiskers then show the upper and lower 25% distributions of the data. The asterisks mark outliers in the data set.

This figure shows that Signature variants A and B are assessed as the most annoying. The original signature and the C variant were rated as less annoying. Variant D was rated the least

annoying. These results are consistent with the degree to which the signatures exceed the Tokita and Nakamura thresholds. Referring back to Figure 46, Signature 1B exceeded the Tokita and Nakamura thresholds to the greatest extent, followed by Signature 1A. Signatures 1 (the original) and 1C exceeded the detection threshold but generally fell below the annoying threshold. Signature 1D fell well below the annoying threshold.

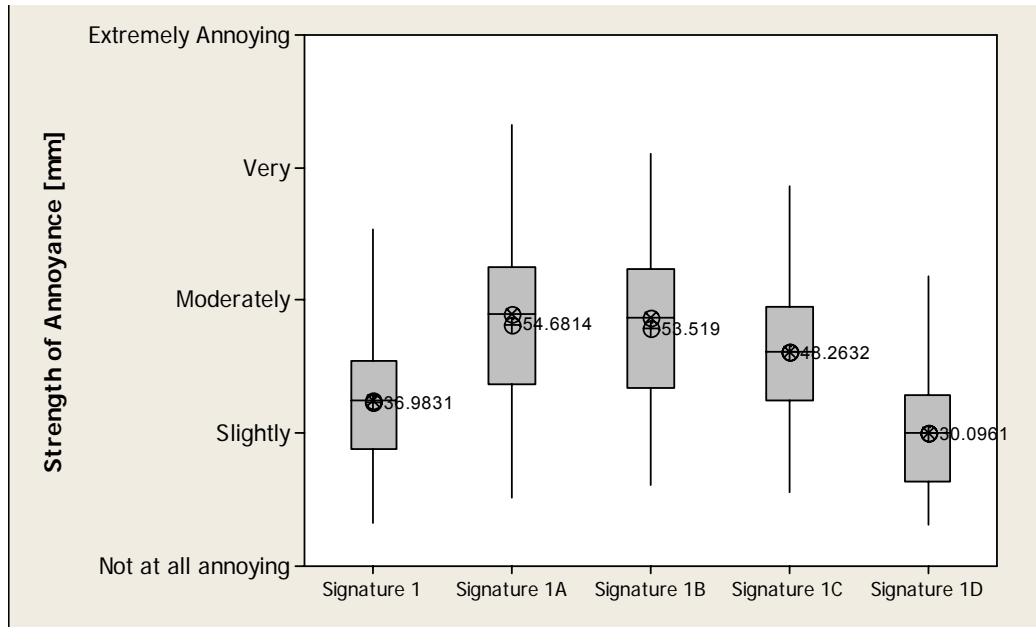


Figure 49. Example of a box plot analysis of the strength of annoyance on Signature Set 1.

Linear regression analysis was used relate the objective metrics to the subjective strength of annoyance assessments. The results are shown in Figures 50 – 52 for L_{CE} , L_{FSL} , and L_{LF} , respectively. The dashed red lines show the 95% confidence intervals for the linear fit. Several conclusions may be drawn. The Tokita and Nakamura thresholds can be used to assess the potential for annoyance to low-frequency aircraft noise. L_{CE} provided the best correlation to the perceived level of annoyance. L_{FSL} and L_{LF} correlate less well than L_{CE} but about the same as one another. The findings suggest that people are responding to the broad spectral content and any predictive metric should quantify the full broadband noise.

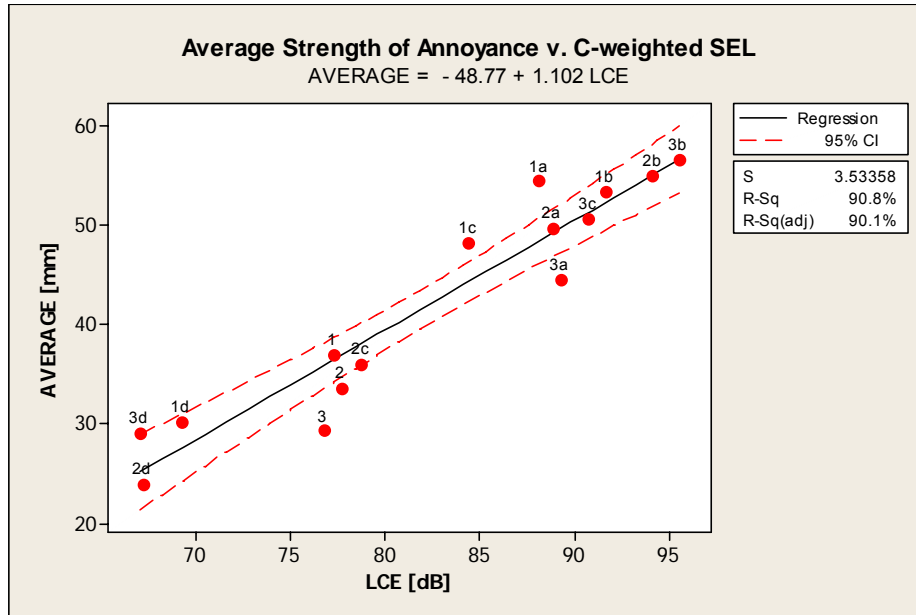


Figure 50. Linear regression analysis of the correlation of L_{CE} with the average strength of annoyance.

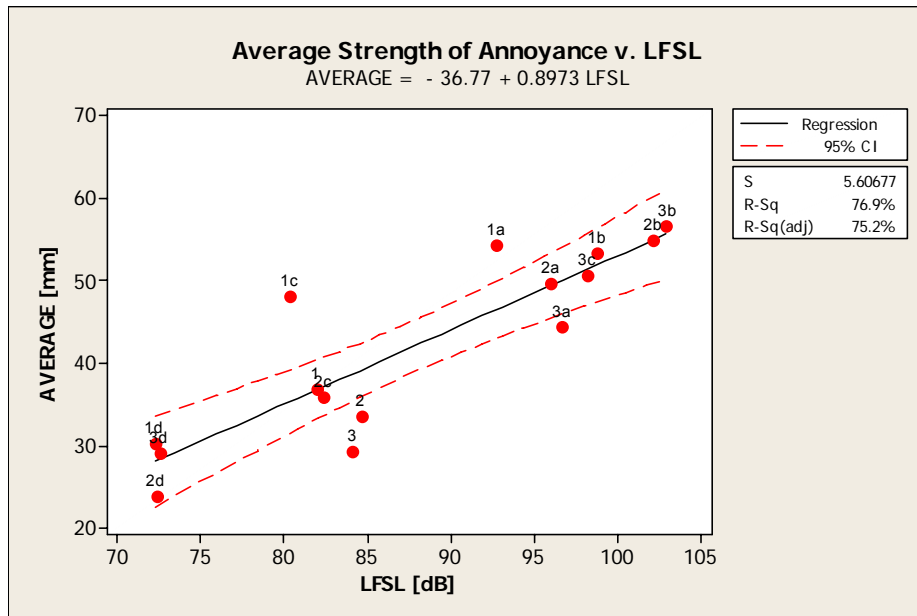


Figure 51. Linear regression analysis of the correlation of LFSL with the average strength of annoyance.

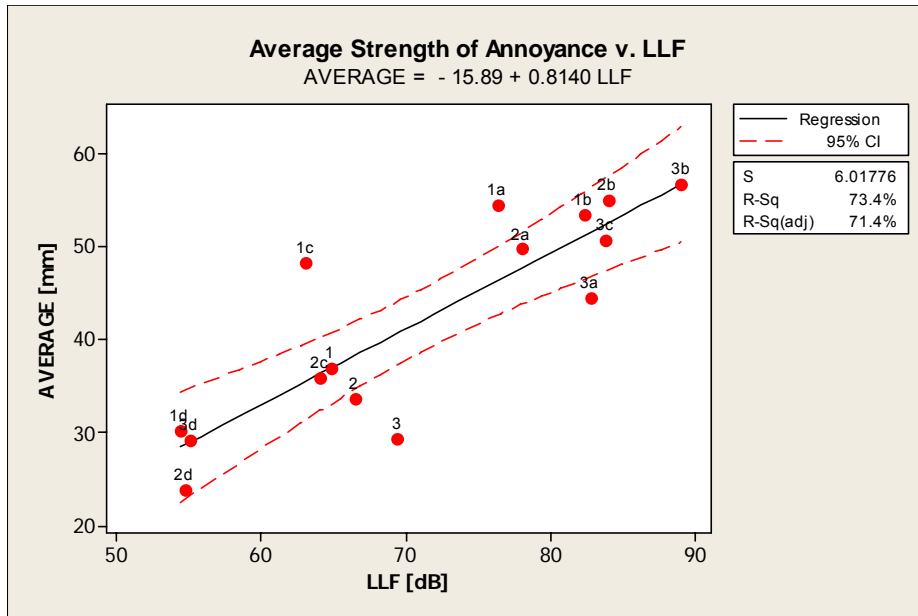


Figure 52. Linear regression analysis of the correlation of L_{LF} with the average strength of annoyance.

9. Rattle and Low Frequency Sound Insulation Studies

9.1 Laboratory Study of Rattle Onset

The objectives of the rattle and sound insulation studies are to develop correlations between low-frequency sound pressure levels, building design, and the frequency and amplitude of vibration, interior noise and rattle in communities surrounding airports. The vibration response of panels insonified with large amplitude, low-frequency acoustic waves was investigated. Similar models were developed for doors and windows. The models allow a better understanding of the causes of rattle and sound transmission and help identify suitable criteria. Mitigation efforts will be developed with this understanding.

High levels of low frequency noise are created by aircraft during takeoff and landing. A by-product of low frequency noise is the excitation of many structures into vibrations. The acoustically induced structural vibration results in rattle and low frequency interior sound that contributes to both the perception and the annoyance of the noise source. Investigation of the mechanisms leading to rattle and the development of rattle mitigation strategies are needed to reduce rattle emissions and thus complaints due to annoyance. In addition to the control of rattle mechanisms, passive sound insulation provides a mitigation strategy to attenuate sound energy from transmitting into homes. Current sound insulation methods are designed to reduce noise in the audible frequency range, typically above 250 Hz. Because sound transmission at frequencies less than 250 Hz is due to different mechanisms it is possible that transmission of high-performance, retrofit doors and windows typical of current sound insulation programs near airports will be less than doors and windows with lower rating. A series of highly rated windows were tested at the low frequency transmission loss facilities at the NASA Langley Research Center.

Rattle is the rapid loss and re-establishment of contact between loosely connected objects.⁵⁶ This study was conducted to investigate the mechanisms contributing to rattle by development of theoretical models and experimental testing of windows known to rattle. The theoretical models are lumped-parameter, single-degree-of-freedom models of elements typically found in homes. The models are divided into two classes: resonant and non-resonant systems. Previous research conducted by others^{57,58,59} have developed non-resonant models to describe rattle. However, many rattle observations do not behave as predicted by such models. No previous investigation has modeled rattle onset for resonant systems. Rattle criterion are determined for various excitation sources including random and harmonic base motion and forced excitation. These criteria include the rattle threshold and also the rattle bandwidth, a feature of resonant systems.

An *in-situ* experiment was conducted at Purdue University's Ray W. Herrick Laboratories. Four windows known to be susceptible to rattle were excited via high-fidelity playback of three large-amplitude, low-frequency noise signals. These windows are more prone than most windows to rattle but provided a controllable case for which to study rattle of resonant systems and to qualitatively verify the simple rattle models. The signals included pre-recorded aircraft take-off, swept sine signal, and pink noise. The vibration and acoustic response of each of the windows was measured to determine the relationship between frequency and acceleration level for onset of rattle. Modal response of the windows corresponded to rattle onset near resonant frequencies. This behavior is consistent with predicted response of the theoretical resonant models.

The rattle criteria developed by the models and validated by the laboratory study provide mitigation strategies for the reduction of rattle in current housing structures and design criteria for future structures. By reducing and possibly eliminating rattle emissions, a major source of annoyance is minimized. This research is intended to provide immediate solutions to reduce rattle using preload strategies and stiffness adjustments.

9.1.1 Rattle Results

Measurements were collected for four windows. The results for all four windows were similar. Thus, discussion of one window is sufficient. Measurement locations include outdoor microphone measuring the noise source, indoor room microphone, and window-mounted accelerometer.

The acceleration response of one window to the highest level swept sine signal is shown in Figure 53. The measured sound level at the surface of the window to the highest level swept sine signal is shown in Figure 54. The responses shown in were recorded simultaneously in time. A spectrogram plot is used to display frequency versus time in the top subplot of both Figures 53 and 54 with the reference scale above the spectrogram. The time history of the (1/8) second time-averaged equivalent level, L_{eq} , is shown in the bottom subplot of the same figures. Unweighted, A-weighted, and C-weighted equivalent sound levels, L_{eq} , are shown in Figure 54. For the highest signal strength the resulting overall sound level was 100 dBC. The swept sine signal sweeps down from 700 Hz to 20 Hz repeating after 78 seconds. The bottom-most diagonal line shown in both figures is the response to the input frequency from the frequency generator, while the parallel lines are harmonic distortion caused from driving the loudspeakers at high levels. On the spectrogram plot of the swept sine signal shown in Figure 53, rattle events are indicated by vertical highlights, indicating non-linear, broadband window response to sinusoidal excitation.

Overall acceleration levels, L_{eq} , of the lower east window to the four signal amplitudes of the swept sine are shown in Figure 55. Four black, vertical lines at 35, 44, 52, 66 Hz indicate resonances of the window determined by modal analysis. The four signal strengths corresponded to a series of overall outside sound levels measured at the surface of the window, which are indicated by the different line colors (blue = 70 dBC overall outside sound level, green = 80 dBC, orange = 90 dBC, and black = 100 dBC). The range of frequencies over which the window rattled is shown in bolded linestyle. Rattle onset was determined via visual inspection of the window response spectrogram plots when the window was excited into broadband response. It can be seen that acceleration level increase at frequencies near the resonance frequencies indicating resonant response.

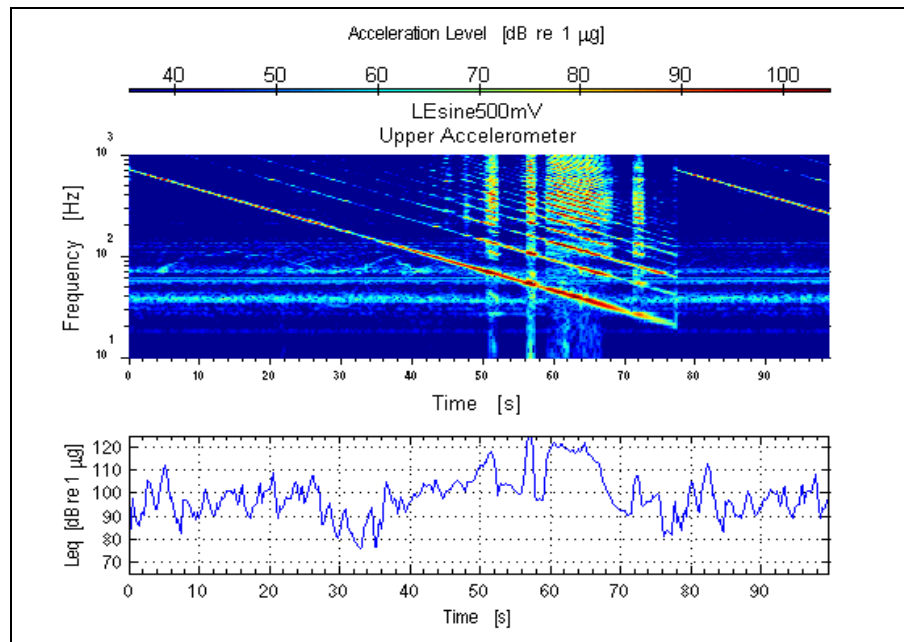


Figure 53. Vibration response of the window to highest swept sine signal with average sound pressure level of 100 dBC measured at the surface of the window.

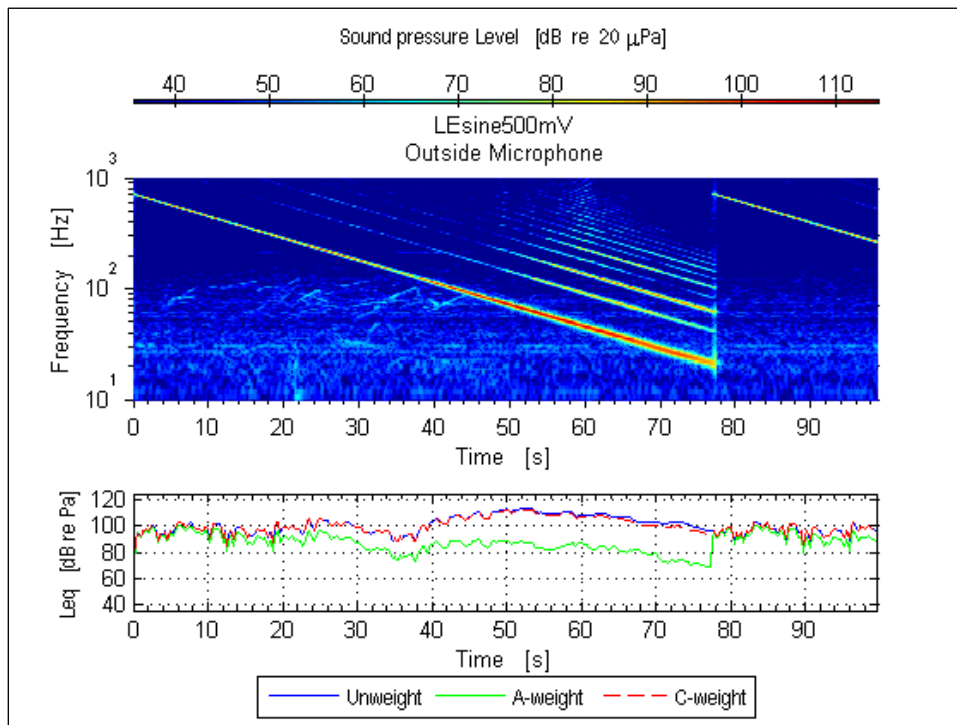


Figure 54. Sound pressure level measured at the exterior surface of the window of swept sine signal.

The spectrogram in Figure 56 is identical to Figure 53 except the frequency scale is only over the low frequency range (10-200 Hz). The bottom-most diagonal trace is the swept sine (700-20 Hz) from the function generator and the parallel traces are harmonic distortion from driving the loudspeaker at high levels. The tick marks on the frequency axis indicate resonances of the window which were determined by modal analysis. The first rattle event (seen as a vertical highlight) occurs when the swept sine passes through 65 Hz and corresponds to the resonance of the window at 65 Hz. The second rattle event also corresponds to modal resonance of the window at 53 Hz. Again rattle occurs at resonance of 44 Hz and carries through to the 35 Hz resonance. The final rattle event occurs in concurrence when the swept sine passes through 30 Hz also a resonance of the window.

Several observations can be made based on the results shown in Figure 55. First, the rattle did not occur at any frequency for the lowest signal amplitude indicated by the C-weighted 70 dB line, even though resonant behavior is evident from the increased acceleration level near the natural frequencies. Thus, for the right combination of parameters (excitation amplitude, preload, material stiffness), rattle can be mitigated. Secondly, for this window, rattle occurred only at acceleration levels greater than 100 (dB re 1 μg) or 0.1 g_{rms} . Thirdly, the rattle bandwidth increases for increasing excitation amplitude. Fourthly, the rattle onset threshold (acceleration amplitude) for the window is essentially the same regardless of the mode shape

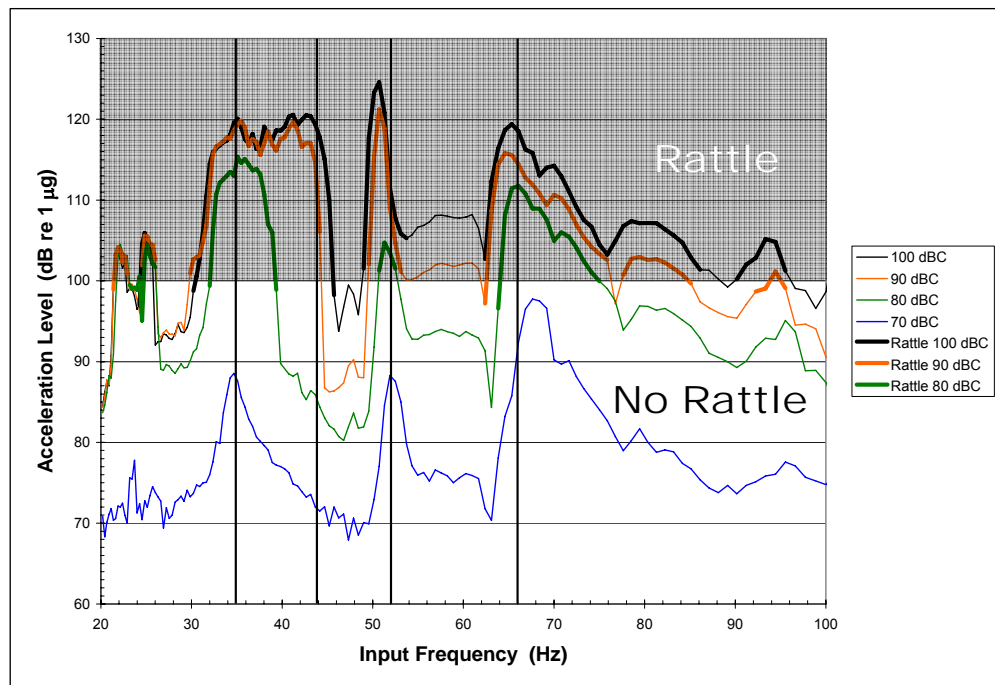


Figure 55. Response of the window to swept sine.

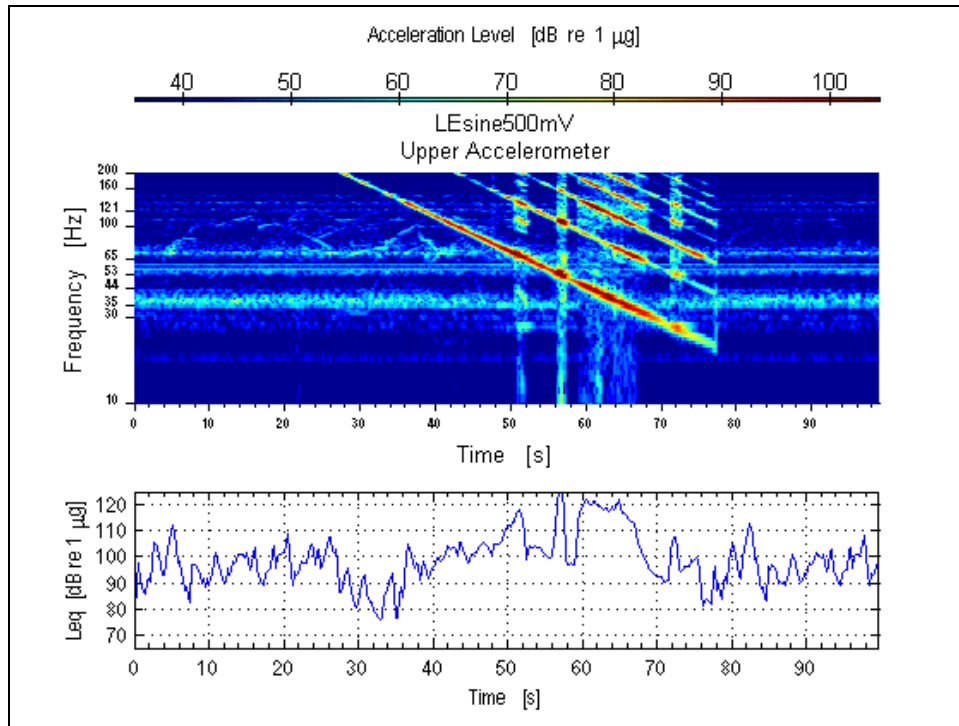


Figure 56. Same plot as Figure 53 but with smaller frequency scale (10-200 Hz).

The rattle behavior of the window corresponds well to the behavior of the Case 6 resonant rattle model discussed in the next section. The Case 6 rattle system model predicts an upper and lower rattle onset threshold centered about the natural frequency of the system. In the rattle experiment upper and lower rattle onset thresholds were seen centered about the resonances of the window. The Case 6 rattle system model predicted that for increasing excitation acceleration amplitude the rattle bandwidth increased. This same phenomenon is seen in the rattle experiment.

The aircraft take-off signal can be treated as a random signal in time. Figure 57 is the spectrogram plot of the outside microphone measuring the aircraft take-off signal with overall equivalent sound pressure level, $L_{eq} = 95$ dB at the window. The aircraft signal used in this study was a recording of a typical aircraft take-off from a single location. Aircraft take-off noise contains significant low frequency content, as seen by the nearly 20 dB difference between the A-weighted and C-weighted L_{eq} at the beginning of the measurement. The low frequency content is consistent throughout the signal.

The vibration response of the window is shown in Figure 58. The tick marks on the vertical axis (frequency) denote modal resonances of the window. The window responds readily at those frequencies. The vibration response of the window is shown in Figure 58. The tick marks on the vertical axis (frequency) denote modal resonances of the window. The window responds readily at those frequencies. Rattle events occur at the vertical highlights, indicating non-linear,

broadband response. Several rattle events occur between the time interval 20-27 seconds. This cluster is explained by the higher amplitude signal frequency content in the range of the resonant modes (50-150 Hz) over that time interval. The higher amplitude signal frequency content is shown in Figure 57 as orange (90 dBC) and red (100 dBC) speckles surrounded by a predominately yellow (85 dBC) region. The response to both the swept sine and aircraft/random signal show that rattle occurs for exposure to sound levels greater than 100 dBC.

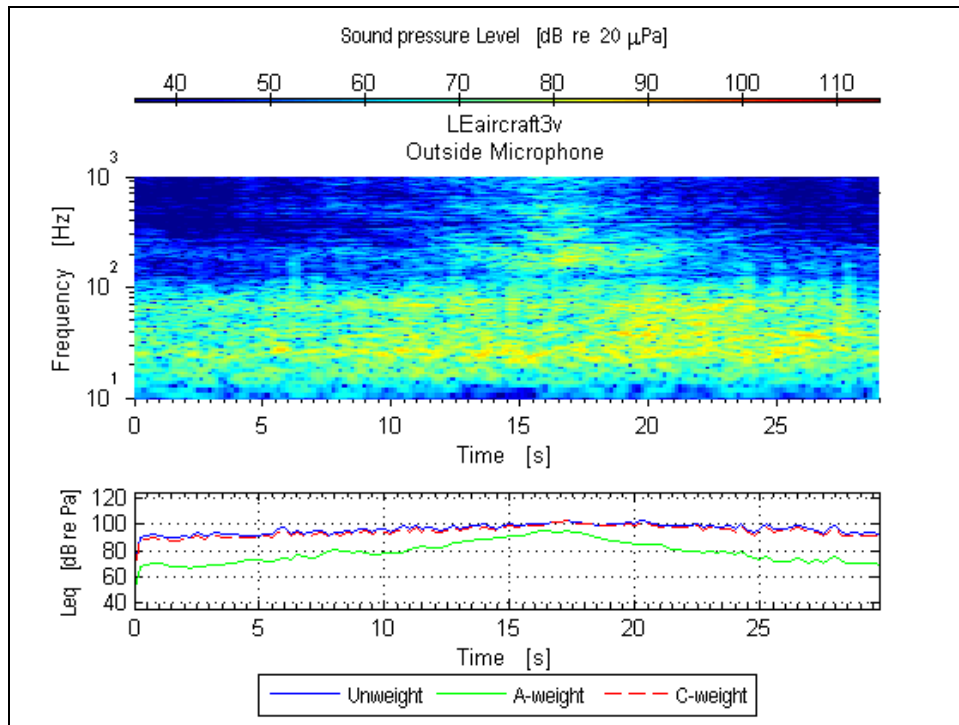


Figure 57. Sound pressure level measured at the surface of the lower east window of aircraft take-off signal.

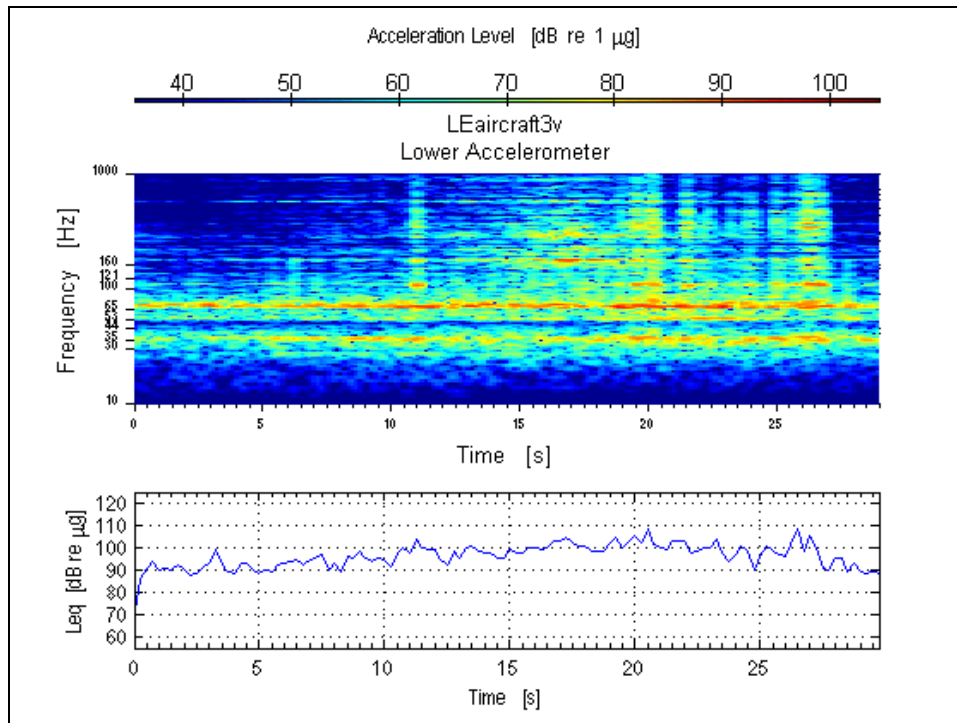


Figure 49. Vibration response of the lower east window to aircraft take-off signal.

9.1.2 Theoretical Models of Rattle

To explain the behavior of the window rattle, we developed analytical models of rattle for resonant and non-resonant systems. Three non-resonant rattle models are shown in Table 19 and three resonant rattle models are shown in Table 20. Steady-state harmonic base acceleration excitation of amplitude, A_b , was considered for Cases 1 through 4, and a harmonic base displacement excitation of amplitude, Y_b , was considered for Cases 5 and 6. Case 1 is in agreement with the findings of Hubbard⁸ that objects in normal contact with a vibrating surface rattle for input acceleration levels greater than the acceleration of gravity, $1g$. The resonant models demonstrate that rattle is possible for acceleration levels less than $1g$. For resonant systems rattle will occur over a range of frequencies, termed the rattle band, centered about the resonance of the system if the amplitude of vibration is large enough. Rattle bands can be minimized in general by using significant preloads, mitigating rattle and also by decreasing the natural frequency of the system. In the stiffness controlled region damping is not effective in mitigating rattle. However, in the damping controlled region, which is centered about the natural frequency, damping is effective in mitigating rattle. Unfortunately, for most typical systems the rattle band is greater than the damping controlled region indicating that damping is not a significant mitigation strategy.

In Table 19, the rattle criterion are summarize for six models excited by various vibration

inputs. Rattle mitigation strategies are summarized in Table 21.

Resonant behavior was observed in the rattle experiment of four windows. The rattle characteristics of the window most closely conformed to the Case 6 model. Rattle occurred in a band centered on the natural frequencies of the windows. Rattle occurred only when acceleration levels exceeded a particular level. For each window the rattle onset threshold was less than the acceleration of gravity, 1g.

Table 19. Rattle criterion for non-resonant models.

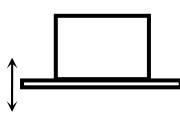
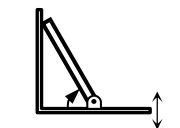
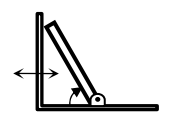
		Harmonic Motion Excitation
Case 1		$A_b > g$
Case 2		$A_b > g$
Case 3		$A_b > \frac{g}{2 \tan \theta}$

Table 20. Rattle criterion for resonant models.

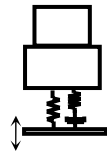
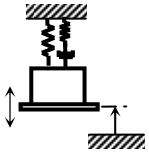
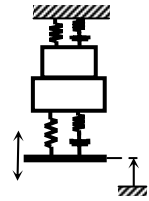
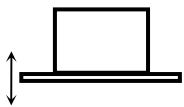
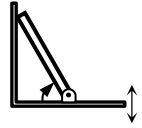
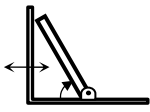
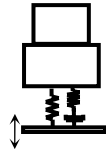
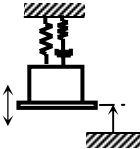
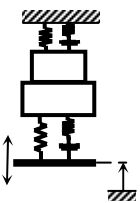
		Harmonic Motion Excitation
Case 4		$\sqrt{1 - \sqrt{(A_b/g)^2 (1 + \gamma^2) - \gamma^2}} < (\omega/\omega_n) < \sqrt{1 + \sqrt{(A_b/g)^2 (1 + \gamma^2) - \gamma^2}}$
Case 5		$(\omega/\omega_n) > \sqrt{1 + \sqrt{((y_p - y_{p,\min})/Y_b - 1)^2 - \gamma^2}}$
Case 6		$\omega_1 > \omega_n > \omega_2$ $\sqrt{\frac{y_p - y_{p,\min} + Y_b}{y_p - y_{p,\min} + (\omega_n/\omega_2)^2 Y_b}} < \left(\frac{\omega}{\omega_n}\right) < \sqrt{\frac{y_p - y_{p,\min} - Y_b}{y_p - y_{p,\min} - (\omega_n/\omega_2)^2 Y_b}}$ $\omega_1 < \omega_n < \omega_2$ $\sqrt{\frac{y_p - y_{p,\min} - Y_b}{y_p - y_{p,\min} - (\omega_n/\omega_2)^2 Y_b}} < \left(\frac{\omega}{\omega_n}\right) < \sqrt{\frac{y_p - y_{p,\min} + Y_b}{y_p - y_{p,\min} + (\omega_n/\omega_2)^2 Y_b}}$

Table 21. Rattle mitigation strategies.

		Harmonic Motion Excitation
Case 1		$A_b > g$
Case 2		$A_b > g$
Case 3		$A_b > \frac{g}{2 \tan \theta}$
Case 4		Increase pre-load. Decrease natural frequency (reduce stiffness or increase mass)
Case 5		Increase pre-load Decrease natural frequency (reduce stiffness or increase mass)
Case 6		Increase pre-load. Soften/stiffen one element relative to the other.

9.2 Low Frequency Sound Insulation Study

The low frequency sound reducing potential of window designs and construction methods with improved low frequency sound transmission loss, such as double and triple pane windows, was assessed. Transmission loss measurements were made via the sound intensity method for five high STC-rated residential windows. The vibration response of the windows to stationary, random excitation noise during the transmission loss test was measured. In addition, a separate mobility test was conducted using the impact hammer method. The objective of this part of the investigation was to assess the impact of improved acoustic insulation of windows and to evaluate their low frequency sound transmission properties.

The windows in this study are classified into two general groups: high-performance (HP) and ultra-high-performance (UHP). The high-performance windows were selected from the manufacturer specifications based on a mid-thirties STC rating. The ultra-high-performance windows were selected from the manufacturer specifications based on a mid-forties STC rating.

Transmission loss tests were done at the Structural Acoustics Loads and Transmission (SALT) facility at the NASA Langley Research Center. The one third octave band transmission loss measurements of all five windows studied are shown in Figure 59.

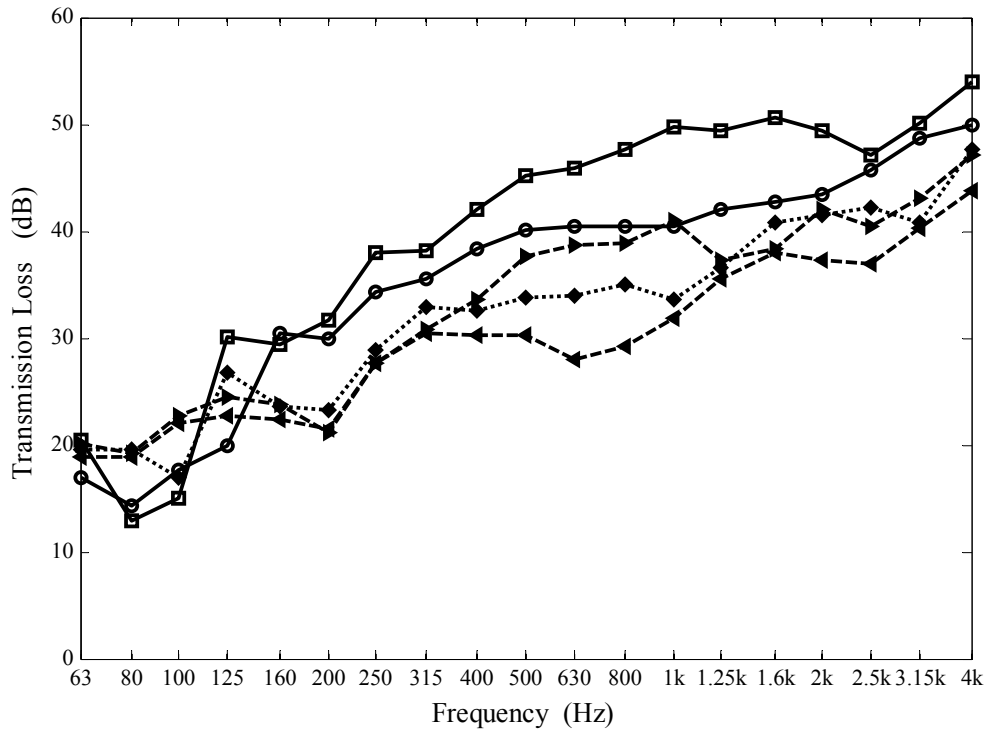


Figure 59. Transmission loss of five high-performance windows. Window A (o-), Window B (□-), Window C (◆-), Window D (◀-), Window E (▶-).

The transmission loss of the ultra-high-performance group (Windows A and B) is consistently higher than the high-performance group (Windows C, D, and E) at frequencies above 125 Hz. Windows A and B have higher STC rating, because the STC calculation includes transmission loss data in the frequency range 125-4000 Hz. However, below 125 Hz the transmission loss of the ultra-high-performance group was as much as 7 dB less than the transmission loss of the high-performance group. For the five windows tested in this investigation, the ultra-high-performance windows were consistently poorer than the high-performance windows in the 80 and 100 Hz bands.

A summary of the Sound Transmission Class (STC) and Outdoor-Indoor Transmission Class (OITC) for all five windows is shown in Table 22. Both rating schemes are a single-number metrics intended to characterize the transmission loss properties of a panel or structure. STC rating, as defined by ASTM standard E413⁶⁰, is intended to be applied panels or structures between adjoining interior spaces, though industry has often applied it to external elements such as doors and windows. For the calculation of the STC rating one third octave band transmission loss data in the frequency range 125-4000 Hz is used. The OITC rating, as defined by ASTM standard E1332⁶¹, is intended to quantify outdoor-indoor transmission loss. The OITC rating is a single number metric that is based on measured transmission loss of the sound insulating element in the frequency range from 80 to 4000 Hz. The OITC rating is determined from the difference of measured transmission loss from idealized, A-weighted transportation noise spectrum. The idealized transportation noise spectrum is representative of highway, rail, and aircraft noise. OITC is intended to be applied to exterior elements such as doors and windows.

Table 22. Rattle mitigation strategies.

Class	Window	STC	OITC
UHP	A	41	30
	B	46	29
HP	C	37	29
	D	33	28
	E	38	30

The UHP windows were possibly designed with the STC rating in mind (which neglects transmission loss data below 125 Hz). The STC rating does not quantify the poor low frequency performance of the UHP windows, whereas the OITC rating does. Consideration of the OITC rating shows the five windows to have approximately the same performance.

Mass-air-mass resonance may fall in the 56-4000 Hz frequency range. The mass-air-mass resonance is a well-known phenomenon in multi-panel structures, where the air space acts as compliance between the masses of two of the glass panels. The mass-air-mass resonant frequency, f_m , is approximately determined by⁶²:

$$f_m = \left[\frac{\rho_0 c_0^2}{4\pi^2 d} \left(\frac{m_1 + m_2}{m_1 m_2} \right) \right]^{1/2}$$

where m_1 and m_2 are the masses of two of the glass panels, d is the depth of the air space between the glass panels, and ρ_0 and c_0 are the density and speed of sound, respectively. The estimated mass-air-mass resonances for each test window are tabulated in Table 23.

Table 23. Estimated mass-air-mass resonances for five test windows.

Mass-Air-Mass Resonant Frequencies	
<i>Window</i>	<i>Frequency (Hz)</i>
A	115, 125, 200
B	85, 95, 208
C	180
D	200
E	183

From Table 23 it can be seen that mass-air-mass resonances can exist in the low frequency range. Theory predicts that for the frequencies immediately following the mass-air-mass resonance the transmission loss increases at 18 dB/octave.⁶² Trend-lines following this 18 dB/octave increase have been plotted in Figure 60 to highlight the mass-air-mass resonances. Both flexural resonances and mass-air-mass resonances contribute to the decrease of transmission loss at low frequencies.

The transmission loss, acceleration level, and acceleration level results of one window, Window C, are shown in Figure 60. The findings of this window were typical of all five windows studied in this investigation. In Figure 60, one third octave band transmission loss is shown as bolded linestyle. The narrow band transmission loss is overlaid in black. The average acceleration level as determined by acoustic excitation during the transmission loss measurement is shown as gray “dash” linestyle. The acceleration level, as determined by a separate mobility measurement is overlaid in “dash-dot” linestyle. The mass-law transmission loss prediction and 18 dB/octave are overlaid as black “dash” and solid gray linestyles, respectively.

In Figure 60 the average acceleration level and the acceleration level are consistent, as expected. Also, regions of increased acceleration and acceleration response (“peaks”) correspond to a decrease in transmission loss (“dips”). For example, this phenomenon occurs for Window C in Figure 51 at 63, 85, 107, 180, and 270 Hz. The resonant response of the window corresponds to decreased transmission loss at the resonance frequencies.

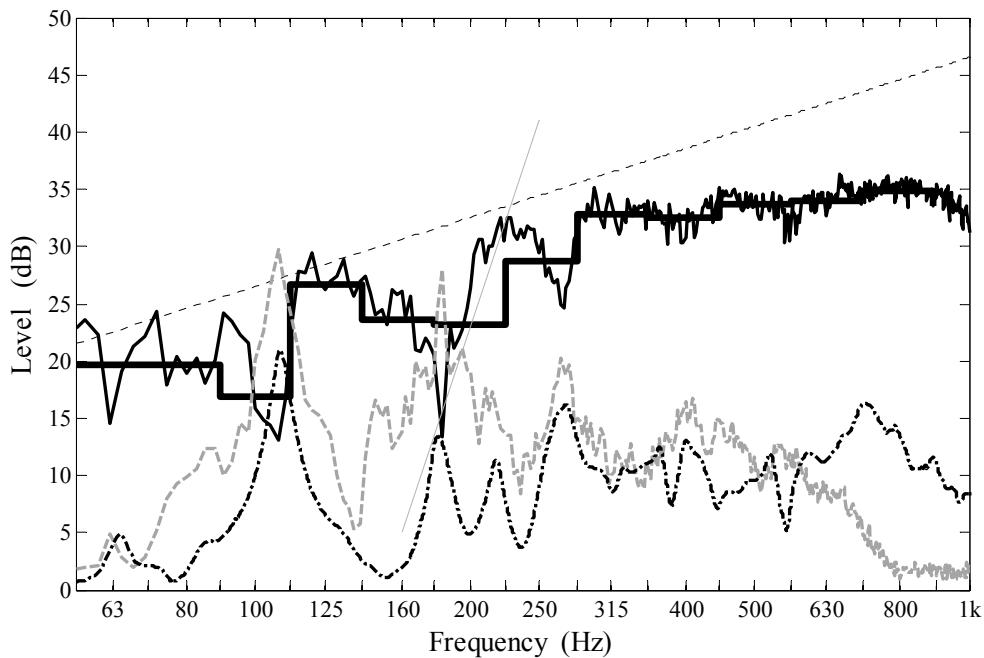


Figure 60. Window C sound transmission loss (TL) and acceleration level vs. frequency; narrow-band TL (—), one-third octave band TL (—), mass-law TL (..), average acceleration level (—), average acceleration level (-.-), 18 dB/octave trend-line (—).

The low frequency resonant response of the windows is the most likely cause for decreased transmission loss at low frequencies. The resonant response of the windows included flexural and mass-air-mass resonances. Flexural resonant response is damping-controlled and thus minimized by increasing damping of the glass panel in the window. This could be accomplished by constructing the window with laminate glass and/or increasing the effective damping of the seals between the glass panel and the window unit frame. The mass-air-mass resonance is primarily controlled by the mass of the individual glass panels and the distance between them. If possible, these resonances should be shifted by adjusting mass of air gaps to optimize window performance.

For the five windows tested in this investigation the three windows with lower STC rating (C, D and E) performed better in the 80 and 100 Hz one-third octave bands than the two windows in the ultra-high-performance class (A and B). This is because two mass-air-mass resonances were in the 80 and 100 Hz one-third octave band for both ultra-high performance windows. This is a small sample and cannot be generalized to all such windows, but it does illustrate that high Sound Transmission Class (STC) rating does not ensure good low frequency transmission loss performance. Consumers may be misled that high STC rating windows equivocate to good low frequency performance. The OITC rating is recommended instead of the STC rating for exterior components such as doors and windows. The OITC rating is intended for exterior elements where as STC is not and the OITC rating includes frequency content down to 80 Hz thus providing a better single-number metric of low frequency transmission loss performance.

10. Summary, Findings and Recommendations

This document is the final report of the PARTNER Low-Frequency Noise Study mandated by the United States Congress (H1200, House Congressional Record 12 February, 2003) to address the issues raised by FICAN² concerning the report of the MSP Expert Panel.¹ In addition to addressing FICAN's recommendations, this study is intended to contribute to a better understanding of the impact of low-frequency aircraft noise on communities, and assess which metrics are most effective in predicting it.

PARTNER investigators reviewed prior studies of low-frequency noise in the vicinity of airports^{1,4,6,10,11} and published archival literature, meet with representatives of the City of Richfield, members of the Metropolitan Airports Commission, FICAN, and the MSP Expert Panel, consulted with other experts and researchers, and established a set of airport selection criteria to design the follow-on low-frequency noise study.

This study included both field measurements at Washington Dulles International Airport, laboratory-based subjective jury trials, and laboratory-based rattle and low frequency sound insulation studies. Key aspects of the study included:

- Measurement of aircraft source noise close to runways
- Measurement of noise and vibration impact at residential structures close to runways
- Three subjective jury trials to assess impact of low frequency noise
- Correlation of jury test results with metrics
- Identification of metric(s) that correlate best with subjective responses
- Investigation of quantitative measurements to assess the potential for annoyance due to low-frequency noise
- Laboratory-based study of rattle and low-frequency sound insulation

The principal findings and recommendations of this study are as follows:

1) Field Measurements

Source Noise

Finding: The highest levels of noise near the runway during start-of-takeoff-roll and acceleration down the runway and during thrust reversal are at frequencies below 200 Hz.

Measurements of sideline noise at start-of-takeoff-roll show that the larger the aircraft, the higher the noise levels, and the levels steadily decrease as the aircraft moves down the runway. Measurements of noise levels during thrust reversal do not show the same trend with aircraft size. The largest aircraft and highest thrust-rating category do not have the highest noise levels.

Recommendation: The Integrated Noise Model uses forward-thrust noise data to model thrust reverser noise. Thrust reverser noise was identified by the MSP Expert Panel as a potential significant contributor to low frequency noise annoyance and was shown in this study to have significant levels of low-frequency noise. Both the levels and directivity of thrust reverser noise

should be investigated further to determine if modifications to noise models are warranted.

Noise and Vibration Impact at Residential Structures

Finding: Measured vibration levels of windows in houses located within 3000 ft of runways can exceed the Hubbard threshold criteria, indicating the potential for vibration to be perceived by occupants. The thresholds were exceeded to a greater degree on a rattle-prone window. For the most part, the vibration levels of secure windows fell below the Hubbard thresholds. The level of wall vibrations for takeoff or landing events having the highest exterior peak C-weighted sound levels can exceed the Hubbard threshold. The vibration levels of the floors did not rise significantly above the background level.

Hubbard's exterior sound pressure level criteria, cited in FICAN's response to the MSP Expert Panel, are consistent with the direct vibration measurements and proved good indicators of the onset of window rattle.

Sound Exposure Level L_E , Low-Frequency Sound Level L_{FSL} and Low-Frequency Level L_{LF} , and the Maximum Sound Pressure Level L_{max} correlated well with vibration levels of a rattle-prone window. The A-weighted metrics correlated poorly with the acceleration levels a rattle-prone window.

Recommendation: The Hubbard exterior sound pressure level threshold criteria be used as a first assessment of the potential for low-frequency noise impact.

Recommendation: Modern windows have optional plastic grid inserts. The rattle thresholds for these types of windows should be assessed.

2) Subjective Assessment of Low-Frequency Noise

Spectral Balance Study:

Finding: The spectral balance study included only single event signatures. Several level-, loudness-, and perceived noisiness-based metrics correlated well with subjective evaluations of indoor aircraft noise, in particular, L_{Amax} and L_{Cmax} .

Recommendation: Because L_{Amax} and L_{Cmax} are simple metrics to implement, they should be used to predict subjective response to indoor aircraft noise when the levels are appropriate for A- and C-weightings and there are not high levels of low-frequency noise.

Audible Rattle Study:

Finding: The rattle trial included four signatures that resulted from noise impacts from events on multiple runways simultaneously. Audible window rattle was more likely to be observed for a

combined event than for a single event.

Signatures that contained audible rattle were not ranked as the most annoying, most likely because the rattle content was audible, but not loud, relative to the overall noise content of the signature. This result is consistent with other studies of noise containing audible rattle.¹² The subjective rankings of the rattle signatures were grouped together and in the same order relative to the non-rattle signatures, regardless of analysis method. A-weighted and perceived noisiness-based metrics correlated well with the subjective rankings.

Recommendation: Similar to the recommendation from the spectral balance study, L_{Amax} should be used to predict subjective response to indoor aircraft noise when the levels are appropriate for A-weightings and there are not high levels of audible low-frequency noise. Assessment of rattle impact should include both single and multiple events in areas where noise from multiple runways can impact a neighborhood simultaneously. The combined events may create sufficient inaudible low frequencies to induce an audible window rattle.

Assessment of Tokita & Nakamura Threshold for predicting perception of LFN:

Finding: The Tokita & Nakamura annoyance thresholds were validated as predictors of annoyance due to low-frequency aircraft noise. They were found to relate favorably to the subjective annoyance assessments. Linear regression analysis showed that the C-weighted sound exposure level L_{CE} was the best single-metric predictor of subjective annoyance response, explaining over 90% of the variability of the data set. L_{CE} correlated better with the subjective data than metrics specifically designed to quantify low-frequency noise impact.

Recommendation: The Tokita & Nakamura thresholds should be used as indicators of the potential for annoyance due to low-frequency aircraft noise. L_{CE} should be used as a single-number metric for assessing the potential for annoyance when high levels of low-frequency aircraft noise are present.

Finding Valid Across All Subjective Studies: For interior levels without a strong low frequency component the A-weighting captured the perception. For interior levels with strong low frequencies C-weighting correlated better than A-weighting. Loudness based metrics that included the full frequency range of interest also correlated well. The low frequency based metrics did not correlate as well as the level and loudness based metrics. Level influenced perception more than rattle content when assessed in the laboratory.

Overall the findings suggest that people are responding to the broad spectral content and any predictive metric should quantify the full broadband noise. Loudness algorithms should include frequency content below 50 Hz to optimally correlate with the perception of low frequency noise.

3) Laboratory Rattle and Sound Insulation Studies

Finding: The rattle study explained why rattle can occur at acceleration levels below 1g where previous models had predicted onset. Resonant systems tend to have a rattle frequency band around resonance. This result was verified experimentally. The models developed during the study give the capability to identify mitigation strategies.

Recommendation: A general strategy for eliminating rattle in resonant systems is to increase preload and design the systems so that excitation does not coincide with system resonance.

Finding: The window transmission loss study showed that transmission loss performance is degraded at low frequency by resonance. These resonances are either due to panel vibration or from mass-air-mass interactions of the windows and air gaps between them. Where low frequency excitation occurs, our studies show that the Outdoor-Indoor Transmission Class (OITC) rating is a better than the Sound Transmission Class (STC) rating for identifying the performance of windows.

Recommendation: The Outdoor-Indoor Transmission Class should be used for rating window performance.

References

- ¹ Fidell, S., Harris, A. S., and Sutherland, L. "Findings of the Low-Frequency Noise Expert Panel of Richfield-MAC Noise Mitigation Agreement of 17 December 1998," Report to the City of Richfield, Minnesota and the Minneapolis Metropolitan Airports Commission, 2000.
- ² Federal Interagency Committee on Aviation Noise (FICAN), "FICAN on the Findings of the Minneapolis-St. Paul International Airport (MSP) Low-Frequency (LFN) Expert Panel," published on the FICAN website <http://www.fican.org>, 2002.
- ³ Society of Automotive Engineers, Committee A-21, Aircraft Noise, "Procedure for the Calculation of Airplane Noise in the Vicinity of Airports," Aerospace Information Report No. 1845, Warrendale, PA, March 1986.
- ⁴ Sharp, B. H., Beeks, T., and Veerbeek, H., "Groundnoise Polderbaan Overview of Results", A joint Wyle, TNO and NLR Report. Wyle Report 06-02, Wyle Laboratories, 2006.
- ⁵ Sutherland, L.C., Sharp, B. H., Mantey, R.A. "Preliminary Evaluation of Low Frequency Noise and Vibration Reduction Retrofit Concepts for Wood Frame Structures", Wyle Research Report 83-26, Wyle Laboratories, 1983.
- ⁶ Miller, N.P., Reindel, E.M., Senzig, D.A., and Horonjeff, R.D., "Study of Low Frequency Takeoff Noise at Baltimore-Washington International Airport," HMMH Report No. 294730.03/293100.09, Harris Miller Miller & Hanson, Inc., April 1998.
- ⁷ American National Standard ANSI S3.29-1983, "Guide to the Evaluation of Human Exposure to Vibration in Buildings."
- ⁸ Hubbard, H.H., "Noise Induced House Vibrations and Human Perception," *Noise Control Engineering J.*, **19**, 49-55 (1982).
- ⁹ American National Standard ANSI S12.9-2005, Part 4, "Quantities and Procedures for Description and Measurement of Environmental Sound - Part 4: Noise Assessment and Prediction of Long-term Community Response."
- ¹⁰ Harris, Miller, Miller, and Hanson. "Analysis of Start-of-Takeoff Roll Aircraft Noise Levels at Baltimore/Washington International Airport", HMMH Report No. 290600.1, Harris Miller Miller & Hanson, Inc., August 1990.
- ¹¹ Ben H. Sharp, Yuri A. Guovich, and William W. Albee. Status of Low-Frequency Aircraft Noise Research and Mitigation. Technical Report WR 01-21, Wyle Laboratory, September 2001.
- ¹² Powell, C.A, Shepherd, K.P. "Aircraft Noise Induced Building Vibration and Effects on Human Response," *Proceedings of InterNoise*, 567-572, 1989.
- ¹³ Fletcher H. and W.A. Munson, Relation between Loudness and Masking, *J. Acoust. Soc. Am*, 1937, 9, 1-10.
- ¹⁴ Stevens, S.S. Perceived level of noise by Mark VII and decibels (E). *J. Acoust. Soc. Am.*, 1972, 51, 575-601.
- ¹⁵ Broner, N. Low Frequency Noise Assessment-A New Insight? Leventhall, H. G., Ed. *Proceedings Low Frequency 2004 11th International Meeting on Low Frequency Noise and Vibration and its Control*, Aug. 30 –Sept. 1, 2004, Multiscience Publishing Co. pp. 27-34.
- ¹⁶ Leventhall, G. A Review of Published Research on Low Frequency Noise and its Effects. Defra Publications, 2003.
- ¹⁷ Fidell, S., Silvati, L., Pearsons, K., Lind, S., and Howe, R. Field study of the annoyance of low-frequency runway sideline noise. *J. Acoust. Soc. Am.*, 106:1408–1415, 1999.
- ¹⁸ Schomer, P. D. The importance of proper integration of and emphasis on the low-frequency sound energies for environmental noise assessment. *Noise Control Eng. J.*, 52(1):26–39, 2004.
- ¹⁹ Leventhall, H.G., Low Frequency Noise and Annoyance, *Noise and Health*, Vol 6, No. 23, April-June

- 2004, pp. 59-72.
- ²⁰ Kesterson, J.W., Vondemkamp, M.T., and Connor, W.K., "Investigation of Aircraft Departure Noise in Community Areas Behind Runways 1L and 1R at San Francisco International Airport," Tracor Applied Sciences Project 076-439, October 1986.
- ²¹ Connor, W.K., "Investigation of Aircraft Departure Noise in Community Areas Behind Runways 1L and 1R at San Francisco International Airport," Tracor Applied Sciences Project 076-439(-01), February 1987.
- ²² Pearsons, K. et al, "Study of Level, Annoyance and Potential Mitigation of Backblast Noise at San Francisco International Airport," BBN Technologies Report No. 8257, January 2000.
- ²³ Acoustical Design Collaborative, Ltd., "Low Frequency Residential Noise Insulation Study for BWI Airport," Acoustical Design Collaborative, Ltd. Project No. 95.08, August 1996.
- ²⁴ Acoustical Design Collaborative, Ltd., "BWI Low Frequency Noise Analysis for Allwood Neighborhood," Acoustical Design Collaborative, Ltd. Project No. 96.01, September 1997.
- ²⁵ Leventhall, H.G., "Low Frequency Noise and Annoyance," *Noise & Health* **6:23**, 59-72, 2004.
- ²⁶ Poulsen, T., "Comparison of Objective Methods for Assessment of Annoyance of Low Frequency Noise with the Results of a Laboratory Listening Test," *J. Low Freq. Noise, Vibration and Active Control*, **22**(3), 117-131, 2003.
- ²⁷ Inukai, Y., Yamada, S., Ochai, H., and Tokita Y., "Acceptable Limits and Their Percentiles for Low Frequency Noise in Ordinary Adults and Complaints," *Proceedings of Low Frequency 2004*, Maastricht The Netherlands, G. Leventhall editor, Multi-Science Publications, 117-127, 2004.
- ²⁸ Reference 1, Section 2.6.
- ²⁹ Nakamura, S., Tokita, Y., *Frequency Characteristics of Subjective Responses to Low Frequency Sound*. International Conference on Noise Control Engineering, Nederlands Akoetisch Genootschap, Delft, The Netherlands, 735-738, 1981.
- ³⁰ Tokita, Y, Nakamura, S. *Frequency Weighting Characteristics for Evaluation of Low Frequency Sound*, 1981 International Conference on Noise Control Engineering Nederlands Akoetisch Genootschap, Delft, The Netherlands, 39-742, 1981.
- ³¹ Stephens, D.G, Shepherd, K.P, Hubbard, H.H., and Grosveld F.W., "Guide to the Evaluation of Human Exposure to Noise from large Wind Turbines," NASA Tech. Memo 83288, March 1982.
- ³² Mayes, W.H., Stephens, D.G., Holmes, H.K., Lewis, R.B., Holliday, B.G., Ward, D.W., Deloach, R., Cawthorn, J. M., Finley, T. D., Lynch, J. W., et al., "Noise-Induced Building Vibrations Caused by Concorde and Conventional Aircraft Operations at Dulles and Kennedy International Airports," NASA Technical Report TM-78769, 1978.
- ³³ Schomer, P.D. and Sias, J.W., "On Spectral Weightings to Assess Human Response Indoors to Blast Noise and Sonic Booms," *Noise Control Engineering J.*, **46**(2), 57-71 (1998).
- ³⁴ Schomer, P.D. and Averbuch, A., "Indoor Human Response to Blast Sounds that Generate Rattles," *J. Acoust. Soc. Am.*, **86**(2), 665-673 (1998).
- ³⁵ International Civil Aviation Organization, "Recommended Method for Computing Noise Contours around Airports," ICAO Circular 205, 1987.
- ³⁶ European Civil Aviation Conference, "Standard Method of Computing Noise Contours around Civil Airports", ECAC Document 29R, 1986 (2nd edition 1997).
- ³⁷ Plotkin, K.J., Ikleheimer, B., and Huber, J., "The effects of atmospheric gradients on airport noise contours," Wyle Report WR 02-06, September 2003.
- ³⁸ Plotkin, K.J., "Continued Analysis of the effects of atmospheric gradients on airport noise contours,"

- Wyle Report WR 04-12, April 2004.
- ³⁹ M.G. Dittrich, IMAGINE Work Package 4—Aircraft Noise Sources, <http://www.imagine-project.org/artikel.php?ac=direct&id=55> [accessed June 2007].
- ⁴⁰ Schomer, P.D., “The importance of proper integration of and emphasis on the low-frequency sound energies for environmental noise assessment,” *Noise Control Eng. J.*, **52** (1), 26-39 (2004).
- ⁴¹ Schomer, P.D., Suzuki, Y., and Saito, F., “Evaluation of loudness-level weightings for assessing the annoyance of environmental noise,” *J. Acoust. Soc. Am.*, **110** (5), Pt. 1, 2390-2397 (2001).
- ⁴² Broner N. and Leventhall, H.G., “Low Frequency Noise Annoyance Assessment by Low Frequency Noise Rating (LFNR) Curves,” *Journal of Low Frequency Noise and Vibration*, **2** (1), 20-28, 1983.
- ⁴³ Fidell, S., Silvati, L., Pearsons, K., Lind, S., and Howe, R., “Field study of the annoyance of low-frequency runway sideline noise,” *J. Acoust. Soc. Am.*, **106**(3), 1408-1415 (1999).
- ⁴⁴ American National Standards Institute ANSI S12.9-1996-Part 4 "Quantities and Procedures for Description and Measurement of Environmental Sound—Part 4: Noise Assessment and Prediction of Long-Term Community Response," Acoustical Society of America, New York, 1996.
- ⁴⁵ International Standard ISO 3891, "Acoustics-Procedure for Describing Aircraft Noise Heard on the Ground," 1978.
- ⁴⁶ Schultz, T.J., *Community Noise Rating*, Applied Science Publishers LTD, second edition, 1982.
- ⁴⁷ Berglund, B., Berglund, U., Lindvall, T. (1975) Scaling Loudness, Noisiness and Annoyance of Community Noises. *J. of the Acoust. Soc. of Amer.* 60 (5), pp. 1119-1125.
- ⁴⁸ Berglund, B., Berglund, U., Lindvall, T. (1975) Scaling Loudness, Noisiness and Annoyance of Aircraft Noise. *J. of the Acoust. Soc. of Amer.* 57 (4), pp. 930-934.
- ⁴⁹ Kryter, K.D., *The Handbook of Hearing and the Effects of Noise*, p.52. San Diego, CA. Academic Press, 1994.
- ⁵⁰ David, H., *The Method of Paired Comparisons*, Charles Griffin and Company Limited, London, Great Britain, 1988.
- ⁵¹ Zimmer, K., Ellermeier, W., Schmid, C. “Using Probabilistic Choice Models to Investigate Auditory Unpleasantness”, *ACTA ACUSTICA united with ACUSTICA*, **90**, 1019-1028, 2004.
- ⁵² Krishnaiah, P.R., and Sen, P.K., eds., *Handbook of Statistics, Vol. 4*. Elsevier Science Publishers, 299-326, 1984.
- ⁵³ Cohen, B. H., *Explaining Psychological Statistics*, 2ndEd., John Wiley and Sons Inc., NY, NY 2001.
- ⁵⁴ Daniel, W.W., *Applied Nonparametric Statistics*, Houghton Mifflin Company, Boston, MA, 1978.
- ⁵⁵ Fields, et. al., 2001, “Standardized General-Purpose Noise Reaction Questions for Community Noise Surveys: Research and A Recommendation,” *J. Sound and Vib.*, 242(4), 641-769.
- ⁵⁶ Kavarana, F., and Rediers, B., "Squeak and Rattle – State of the Art and Beyond," SAE Paper 99NV273, Noise & Vibration Conf. Proceedings, 1999.
- ⁵⁷ Crandall, S.H. and Kurzweil, L., "On the Rattling of Windows by Sonic Booms," *J. Acoust. Soc. Am.*, **44**, 464-472, 1968.
- ⁵⁸ Clevenson S.A., Experimental Determination of the Rattle of Simple Models, NASA Langley Research Center, NASA-TM-78756, 1978.
- ⁵⁹ Sutherland, L.C., "Low Frequency Response of Structures," Wyle Research Report, WR-82-18, 1982.
- ⁶⁰ ASTM E413, Classification for Rating Sound Insulation, American Society for Testing and Materials (ASTM), West Conshohocken, PA, (2004).
- ⁶¹ ASTM E1332, Standard Classification for Determination of Outdoor-Indoor Transmission Class, American Society for Testing and Materials (ASTM), West Conshohocken, PA, (2003).

⁶² J.M. Mason and F.J. Fahy, The Use of Acoustically Tuned Resonators to Improve the Sound Transmission Loss of Double-Panel Partitions, *Journal of Sound and Vibration*, v. 124, n. 2, pp. 367-379, (1988).