A LABORATORY STUDY OF SUBJECTIVE RESPONSE TO SONIC BOOMS MEASURED AT WHITE SANDS MISSILE RANGE

by

Brenda M. Sullivan and Jack D. Leatherwood

SUMMARY

The Sonic Boom Simulator of the Langley Research Center was used to quantify subjective loudness response to boom signatures consisting of: (a) simulator reproductions of booms recently recorded at White Sands Missile Range; (b) idealized N-waves; and (c) idealized booms having intermediate shocks. The booms with intermediate shocks represented signatures derived from CFD predictions. The recorded booms represented those generated by F15 and T38 aircraft flyovers and represented a variety of waveforms reflecting the effects of propagation through a turbulent atmosphere. These waveforms included the following shape categories: N-waves, peaked, rounded, and U-shaped. Results showed that Perceived Level and Zwicker Loudness Level were good estimators of the loudness of turbulence modified sonic booms. No significant differences were observed between loudness responses for the several shape categories when expressed in terms of Perceived Level. Thus Perceived Level effectively accounted for waveform differences due to turbulence. Idealized booms with intermediate shocks, however, were rated as being approximately 2.7 dB(PL) less loud than the recorded signatures. This difference was not accounted for by PL.

INTRODUCTION

NASA Langley Research Center has conducted a series of laboratory studies (references 1-4) to investigate subjective loudness and annoyance response to simulated sonic booms. These studies, conducted using the Langley sonic boom simulator, were performed in support of NASA High-Speed Research Program efforts to develop a high-speed civil transport (HSCT) aircraft. The resulting data were used to: quantify the effects of boom shaping (minimization) on subjective loudness of outdoor booms (ref. 2), determine effects of boom waveform asymmetry (ref. 3) on loudness, define both loudness and annoyance response to simulated outdoor and indoor booms (ref. 4), and evaluate several metrics as estimators of boom loudness and/or annoyance.

All of the above studies used simple boom waveforms that represented idealized booms predicted by theory. Many of the studies included rise time as a variable and thus did address one of the principal effects of atmospheric propagation on sonic booms (that is, modifications of boom rise times). However, signatures representative of the many complex shapes that real booms can take upon propagation through the atmosphere have not been considered. For example, the shape of the front shock may be affected by molecular relaxation and viscosity in ways that are fairly well understood (ref. 5). Turbulence, however, affects the waveforms in ways that can only be predicted statistically (ref. 6). Obviously, it is impractical to define and synthesize the many signatures required to simulate such a process in the sonic boom simulator. A more reasonable approach is to use ground-measured signatures obtained from actual aircraft flyovers. Fortunately, an extensive collection of recorded boom signatures made during recent tests at White Sands Missile Range is available (ref. 7) for use.

The overall purpose of this paper is to quantify subjective loudness response to a set of signatures selected from those recorded at White Sands Missile Range. The selected signatures represented several shapes that sonic booms may take when modified by turbulence. These shapes have been categorized by previous investigators (see ref. 8) as : N-waves, Peaked waves (booms with one or more peaks on both front and rear shocks), Rounded waves (booms with rounded front and rear shocks), and "U-shaped" waves (booms having two strong positive-going peaks). An additional category, labeled "Intermediate" waves, which are characterized by three or more distinct shocks, was also included.

Several metrics were identified by the previous investigators (refs. 1-4) as being good loudness estimators for the idealized outdoor boom signatures. These metrics were: Steven’s Mark VII Perceived Level, Zwicker Loudness Level, and A-weighted sound exposure level. Each of these effectively predicted the loudness of a wide range N-wave and front-shock minimized boom shapes. An additional objective of this paper was to investigate the performance of these metrics for the White Sands booms.

EXPERIMENTAL METHOD

Sonic Boom Simulator

The experimental apparatus used in this study was the Langley Research Center’s Sonic Boom Simulator. Construction details, performance capabilities, and operating procedures of the simulator are given in reference 1. The simulator, shown in Figure 1, is a person-rated, airtight, loudspeaker-driven booth capable of accurately reproducing user-specified sonic boom waveforms at peak sound pressure levels up to approximately 138 dB. Input waveforms are "predistorted" to compensate for nonuniformities in the frequency response characteristics of the booth and sound reproduction system.

Test Subjects

Forty-eight test subjects (30 female, 18 male) obtained from a subject pool of local residents were used in this study. Ages of the test subjects ranged from 18 to 61 years with a median age of 31.5 years. All subjects were required to undergo audiometric screening prior to the test in order to insure normal hearing.

Experimental Design

Test Stimuli

The test stimuli consisted of simulator reproductions of recorded sonic booms from flyovers of F15 and T38 aircraft at White Sands Missile Range (ref. 7) and several computer-generated idealized booms. The White

Sands boom signatures were examined for the purpose of selecting booms representative of the four categories described earlier. Thirteen booms, having a range of rise times, were selected for inclusion in the experiment. These were: three N-waves (figure 2a), four peaked booms (figure 2b), three rounded booms (figure 2c), and three U-shaped booms (figure 2d).

The three booms within the U-shaped category were chosen on the basis of their being extremely unlike the classic N-wave shape. A U-shaped boom is characteristic of focussed booms which arise when booms travelling along two different paths converge at one location. These can arise from accelerating or manuevering aircraft or as a result of atmospheric propagation phenomena. Only one of the three U-shaped booms used in this test had a shape that likely resulted from focussing. The remaining two were extracted from recordings showing two boom events, which appeared to result from the boom travelling two different paths to the microphone location and arriving there at slightly different times. In each of these two cases, the second event was selected as most resembling a U-shaped signature.

Two idealized waveforms with intermediate shocks (figure 2e) were included in the study. These were based upon CFD predictions of the booms expected from possible HSCT designs. "Intermediate" booms have extra shocks between the front and rear shocks, which can be heard as separate events, and were included to determine whether these "multiple" booms would be judged differently. Three idealized N-waves (figure 2f) were also included in the stimuli set. These had rise times of .25, 3, and 8 milliseconds. The idealized N-waves were included to provide a basis for comparing subjective responses to real versus idealized booms. The five idealized booms were synthesized using the boom waveform generation capabilities of the sonic boom simulator.

Several additional comments pertaining to the test stimuli used in this study are warranted. The first concerns a practical problem encountered during the process of reproducing the recorded booms within the simulator. The field recordings, made at a sample rate of 8kHz, contained audible background noise. This noise had to be removed before the recorded booms could be compared with the noise-free computer-generated idealized booms. Since the noise was broadband, it could not be removed by filtering without adversely affecting the boom waveforms. To remove the noise, the recorded booms were "traced" to select the salient points of the time histories. Briefly, the procedure used was as follows: The sampled data for a waveform were examined and, where there were rapid fluctuations in pressure, all available data points would be selected. When the pressure fluctuations were small and slowly varying, intermediate data points were skipped. The selected points were then joined by straight-line segments and interpolation was used to create a time history at the sample rate (38.5 kHz) used by the simulator. An example of an original boom and its "traced" waveform is given in figure 3. The traced booms were then preprocessed to account for the nonuniform frequency response of the simulator booth (see ref. 1). When the preprocessed versions of the original and "traced" booms were played into the simulator, they sounded very similar, except the "traced" booms did not have the background noise.

A second point of interest relates to the durations of the sonic boom signatures used in this study. It is the intent of the NASA Langley subjective response studies to provide information on the effects of sonic booms created by a future HSCT, which will generate sonic booms having durations significantly longer than those made by the F15 and T38 aircraft. To more accurately simulate future HSCT signatures, the F15 and T38 signatures of the present study were "stretched" by increasing the time between the pressure shock at the front of the waveform and the similar shock at the tail (the intervening slow pressure change is not audible to the human ear, nor is it of sufficient amplitude to be felt). The shapes and rise times of the front and rear shocks were thus unchanged by this procedure. The duration selected for use was 300 milliseconds. The resulting boom signatures are assumed to represent HSCT booms after propagation through the atmosphere. It is known that most turbulence effects occur in the lower few thousand feet of the atmosphere and so the sonic booms created by the F15 and T38, and those created by an HSCT, would all pass through the same lower layer of the atmosphere and undergo similar modifications.

No attempt was made in this study to replicate the actual levels of the booms recorded at White Sands. Instead, a range of levels was included in order to provide data for evaluating the several metrics and for assessing the relative effects on subjective loudness of signature shape differences due to turbulence.

Scaling Method

The scaling method used was magnitude estimation. The ability of subjects to make reliable and accurate ratio judgments of sonic boom loudness was demonstrated in reference 9. The procedure used is summarized as follows: A sonic boom stimulus, designated as the standard, was presented to a subject. The standard was a N-wave with a 3 millisecond rise time and peak overpressure of 0.89 psf. This standard was assigned a loudness value of 100 by the experimenter. The standard was then followed by three comparison booms. The task of a subject was to rate the loudness of each comparison boom as compared to the loudness of the standard. For example, if a subject felt that a comparison boom was twice as loud as the standard, then the subject would assign it a value of 200. If the comparison boom was felt to be only one-fourth as loud as the standard, then the he/she would assign it a value of 25. After three comparison stimuli were judged, the standard was repeated and another three comparison booms were evaluated. This procedure was repeated until all booms within a session (and all sessions) were completed. The subjects were free to assign any number of their choosing (except negative numbers) to reflect their loudness opinions. The instructions explaining how to use the magnitude estimation procedure are given in Appendix A.

Test Structure

The test consisted of the 18 booms described earlier, each of which was presented at five levels, for a total of 90 presentations. These were randomly assigned to two sessions of 45 booms each. To reduce order effects, the booms within each session were presented in reverse sequence to one-half of the test subjects.

Test Procedure

Subjects were delivered to the laboratory in groups of three, with one group in the morning and one group in the afternoon on any given day. Upon arrival at the laboratory, each group was briefed on the overall purpose of the experiment, system safety features, and their rights as test subjects. A copy of these briefing remarks is given in Appendix B. The subjects were then given specific instructions related to the test procedure to be followed and in the use of the magnitude estimation procedure (Appendix A). At this point, the subjects were taken individually from the waiting room to the sonic boom simulator. At the simulator, the magnitude estimation scaling procedure was reviewed and the subject listened to several boom stimuli, played with the simulator door open, in order to become familiar with the type of sounds she/he would be asked to evaluate. The subject was then given a practice scoring sheet and seated in the simulator with the door closed. A practice session was then conducted in which the subject rated a set of stimuli similar to those used in the actual test sessions. Upon completion of the practice session, the scoring sheet was collected and any questions were answered. The first test session was then conducted. After all subjects completed the first session, they were then cycled through the remaining sessions. No further practice sessions were given.

Data Analysis

The boom pressure time histories measured within the simulator were computer processed to calculate sound exposure level in terms of three frequency weightings and to calculate two loudness metrics. The sound exposure level metrics were: unweighted sound exposure level (L_{UE}), C-weighted sound exposure level (L_{CE}), and A-weighted sound exposure level (L_{AE}). The loudness metrics were Stevens Mark VII Perceived Level (PL) and Zwicker Loudness Level (LLZ).

The central tendency parameter used to characterize the magnitude estimation scores was the geometric mean of the magnitude estimates for each stimulus. It is customary (see reference 10, for example) to use geometric averaging with magnitude estimation since the distribution of the logarithms of the magnitude estimates is approximately normal. Furthermore, subjective loudness is a power function of the physical intensity of a sound. Such a power function is linear when expressed in terms of the logarithms of the subjective loudness and sound pressure level.

DISCUSSION OF RESULTS

Metric Considerations

The overall performance of each metric as a loudness estimator was assessed by computing two sets of parameters using the obtained subjective data. The first set of parameters were the correlation coefficients between the subjective ratings and the levels of each metric. As noted earlier, the subjective ratings were characterized by the logarithm of the geometric means and metric levels were calculated from boom measurements made within the simulator. The correlation coefficients are measures of the degree of relationship between each metric and the obtained subjective ratings. The second set of parameters were the standard errors of estimate of the best-fit linear regression lines describing the relationship between subjective ratings and levels of each metric. These represent the prediction accuracies (or precision) of each metric. The smaller the standard error of estimate, the greater the prediction accuracy. Both of these parameters were calculated for the complete stimuli set (that is, all categories combined) and are displayed in Table 1. They were also calculated for each boom category and are presented in Table 2. Scatter plots showing the obtained subjective data for each metric are shown in figure 4.

Examination of Tables 1 and 2 indicates that PL and LLZ consistently correlated highest with subjective ratings and exhibited the lowest standard errors of estimate for all boom categories. (Analysis indicated that the differences in correlation coefficients and standard errors of estimate between PL and LLZ were not statistically significant.) The remaining metrics performed well for some boom categories, but not for others. These results indicate that both PL and LLZ are good loudness estimators for the range of turbulence modified outdoor sonic boom signatures of this study.

Loudness and Boom Category Considerations

Since each boom category represents a "shape" that a boom may assume after propagation through atmospheric turbulence, any differences in loudness ratings between the various categories, for a given metric and metric level, are indicative of an atmospheric effect that is not accounted for by that metric. If these differences are small and/or statistically insignificant, it can be concluded that atmospheric alterations of the boom shapes are accounted for by the metric. Since PL was identified in an earlier study (reference 4) as the metric of choice for general use in estimating subjective effects due to both indoor and outdoor sonic booms, it is the metric that is used to assess loudness of booms within each category.

The linear regression lines describing the relationship between the logarithm of the geometric means of the loudness magnitude estimates and PL for each boom category are presented in figure 5. [The scatter plot for this figure was given in figure 4(e)]. Inspection of the regression lines indicates that all are tightly grouped except for the line representing the two intermediate booms. This regression line, which has a slope similar to those for the other boom categories, indicates that the intermediate booms were rated as being less loud than the booms in the other categories for equivalent PL. Statistical analysis showed that the differences in slope of the regression lines were not statistically significant and that no significant differences existed between the loudness ratings for all categories except the intermediate boom category. The loudness scores for the intermediate booms were significantly lower than those for each of the other categories. Dummy variable analysis indicated that the average difference between the loudnesses of the intermediate booms and the loudnesses of the remaining booms was equivalent to approximately 2.7 dB(PL). The probability level for significant differences was 0.001.

The above results do not imply that booms of different shape, but comparable peak overpressure, were rated equally loud. In fact, this was not the case. What these results do show is that the PL metric effectively accounted for the turbulence-induced shape differences between the recorded White Sands signatures as well as the differences between the idealized N-waves and the White Sands booms. However, the PL metric did not fully account for the reduced loudness of the intermediate booms. This does not mean that PL is inapplicable to booms with intermediate shocks or that corrections to PL to account for these effects are required. The present results were based upon only two intermediate-shock signatures. Definitive conclusions must await the results of additional experiments to examine the effects of multiple shocks for a wider range of intermediate-shock boom parameters.

The reason for the reduced loudness of the intermediate booms is unclear. It is possible that temporal masking, due to the relatively closely spaced multiple shocks, may have played a role. Also loudness asymmetry between the front and rear portions of these signatures may have had a minor effect. It was demonstrated in an earlier study (ref. 3) that asymmetrical signatures were generally perceived by subjects as being quieter than symmetrical signatures of equivalent PL. This effect depended upon which part (front or rear) of a signature was loudest and upon the degree of asymmetry, defined as the difference between front and back loudnesses (measured in terms of PL). However, only one of the intermediate signatures in the present study had a significant degree of asymmetry. This boom is shown in the left part of figure 2(e) and had an asymmetry of about 17 db(PL). According to reference 3, this would result in an equivalent reduction in PL of less than 0.5 dB. This is not sufficient to explain the approximately 2 dB reduction in PL observed in figure 5.

CONCLUDING REMARKS

The sonic boom simulator of the Langley Research Center was used to quantify subjective loudness response to boom signatures consisting of: (a) simulator reproductions of booms recently recorded at White Sands Missile Range; (b) idealized (computer-generated) N-waves; and (c) idealized booms having intermediate shocks. The booms with intermediate shocks represented signatures derived from CFD predictions. The recorded booms represented those generated by F15 and T38 aircraft flyovers during the White Sands tests.

Results were used to assess the performance of several metrics as loudness estimators of the recorded and idealized booms. The recorded signatures consisted of a variety of waveforms reflecting the effects of propagation through a turbulent atmosphere. They were categorized according to the shape of the waveforms and comparisons of loudness judgments between the various categories were then made. Specific comments and findings of this study are summarized as follows:

1. Perceived Level (Steven’s Mark VII) and Zwicker Loudness Level were the best estimators of the loudness of turbulence modified sonic booms. These metrics correlated highest with obtained loudness ratings and had the lowest prediction errors. The results provided additional support for an earlier recommendation of Perceived Level as the metric of choice for assessing and/or predicting sonic boom subjective effects.
2. No significant differences in loudness were observed between sonic booms within the following categories: N-wave (recorded), peaked, rounded,

U-shaped, and N-waves (idealized). Booms having intermediate shocks (intermediate category), however, were rated as being less loud than those in the other categories. The reduced loudness of the intermediate booms was approximately equivalent to a 2.7 dB reduction in PL. Thus, Perceived Level did not fully account for the reduced loudness of the intermediate booms used in this study.

3. Reasons for the reduced loudness of the booms with intermediate shocks are unclear. Possible contributing factors may have been temporal masking and boom asymmetry. Additional tests to validate the loudness effects due to intermediate shocks and to explore possible explanatory mechanisms are desirable.

APPENDIX A

Subject Instructions

This test will consist of six test sessions. Prior to the first test session each of you will be taken individually to the simulator where you will listen to sounds that are similar to those you will be asked to rate. We will then place you in the simulator and a practice scoring session will be conducted. Upon completion of the practice session we will collect the practice rating sheets and answer any questions you may have concerning the test. At this point two test sessions will be conducted. You will then return to the waiting room while the other members of your group complete a similar test. You will return to the simulator two more times to complete the remaining test sessions.

During a test session we will play a series of sonic booms over the loudspeakers in the door of the simulator. The first sonic boom that you hear, and every fourth boom thereafter, will be a **REFERENCE** boom that you will use to judge how loud the other booms are. In order to help you keep track of which boom is the **REFERENCE** boom, it will always be preceded by a short beep. The **REFERENCE** boom will remain the same throughout the test. Your task will be to tell us how loud each of the other booms are as compared to the **REFERENCE** boom. You will be provided rating sheets for use in making your evaluations. The ratings sheets will indicate when a **REFERENCE** boom will be played and the sequence of **REFERENCE** and other booms will be organized as follows:

```
   <--------------------beep
R=100 <-----------------reference
1.________
2.________
3.________
   <--------------------beep
R=100 <-----------------reference
4.________
5.________
6.________
```

The scoring procedure will be as follows: The short beep will indicate to you that the boom which follows is the **REFERENCE** boom. Please listen to it carefully because you will compare the other booms to it. For this purpose the **REFERENCE** boom will be assigned a loudness value of 100. Thus you do not score the **REFERENCE** boom because it will always be equal to 100. You will then hear a sequence of three comparison booms. After listening to each comparison boom you should decide how loud you think it is relative to the **REFERENCE** boom and assign it a number accordingly. This number will be entered on the appropriate line of the scoring sheet. For example, if you feel the comparison boom is three times louder than the **REFERENCE** boom then you would give it a loudness score of 300. If you think the comparison boom is only one-fourth as loud as the **REFERENCE** boom you would give it a loudness score of 25. You may choose any number you wish as long as it faithfully represents your impression of the relative loudness of the comparison and **REFERENCE** booms. After evaluating three comparison booms in this manner you will hear the beep again, followed by the **REFERENCE** boom and three more comparison booms. This will be repeated within a test session until the test session is completed. Remember! There are no right or wrong answers. We are interested only in how loud the booms sound to you.

APPENDIX B

General Briefing Remarks

You have volunteered to participate in a research program designed to evaluate various sounds that may be produced by certain aircraft. Our purpose is to study people's impressions of these sounds. To do this we have built a simulator which can create sounds similar to those produced by some aircraft. The simulator provides no risk to participants. It meets stringent safety requirements and cannot produce noises which are harmful. It contains safety features that will automatically shut the system down if it does not perform properly.

You will enter the simulator, sit in the chair, and make yourself comfortable. The door will be closed and you will hear a series of sounds. These sounds represent those you could occasionally hear during your routine daily activities. Your task will be to evaluate these sounds using a method that we will explain later. Make yourself as comfortable and relaxed as possible while the test is being conducted. You will at all times be in two-way communication with the test conductor, and you will be monitored by the overhead TV camera. You may terminate the test at any time and for any reason in either of two ways: (1) by voice communication with the test conductor or (2) by exiting the simulator.

REFERENCES

1. Leatherwood, J.D.; Shepherd, K.P.; and Sullivan, B.M.: A New Simulator for Assessing Subjective Effects of Sonic Booms. NASA TM 104150, December, 1991.

2. Leatherwood, J.D.; and Sullivan, B.M.: Laboratory Study of Effects of Sonic Boom Shaping on Subjective Loudness and Acceptability. NASA TP 3269, October 1992.

3. Leatherwood, J.D.; and Sullivan, B.M.: Effect of Sonic Boom Asymmetry on Subjective Loudness. NASA TM 107708, December 1992.

4. Leatherwood, J.D.; and Sullivan, B.M: Loudness and Annoyance Response to Simulated Outdoor and Indoor Sonic Booms. NASA TM 107756, May 1993.

5. PLotkin, K.J.: The Effect of Turbulence on the Loudness of Minimized Sonic Boom Signatures. NASA CP 3172, vol. 1, 1992, pp. 77-95.

6. Proceedings of Conference on High-Speed Research: Sonic Boom, NASA CP 3172, February 25-27, 1992.

7. Willshire, W.L.; and DeVilbiss, D.W.: Preliminary Results From the White Sands Missile Range Sonic Boom Propagation Experiment. NASA CP 3172, vol. 1, 1992, pp. 137-149.

8. Maglieri, D.J.; and Plotkin, K.J.: Aeroacoustics of Flight Vehicles: Theory and Practice. Edited by H. H. Hubbard, NASA RP 1258, vol. 1, August 1991, pp. 519-561.

9. McDaniel,S.; Leatherwood, J.D.; and Sullivan, B.M.: Application of Magnitude Estimation Scaling to the Assessment of Subjective Loudness Response to Simulated Sonic Booms. NASA TM 107657, September 1992.

10. Stevens, S.S.: Psychophysics, John Wiley & Sons, 1975, pp. 269-270.

Table I. Correlation Coefficients and Standard Errors of Estimate Between the Mean Ratings and Each Metric. (Based on the Total Stimuli Set, n=90).

METRIC	CORRELATION COEFFICIENT	STANDARD ERROR OF ESTIMATE
PL	0.9544	0.0606
LLZ	0.9663	0.0522
L_{AE}	0.9168	0.0810
L_{CE}	0.8499	0.1069
L_{UE}	0.7700	0.1294

Table II. Correlation Coefficients and Standard Errors of Estimate For Each Metric and Boom Category. Standard Errors of Estimate are indicated by Parentheses.

CATEGORY	PL, dB	LLZ, dB	L_{AE}, dB	L_{CE}, dB	L_{UE}, dB
N-WAVE n=15	0.9758 (0.0477)	0.9805 (0.0429)	0.8989 (0.0956)	0.8873 (0.1007)	0.8431 (0.1174)
PEAKED n=20	0.9656 (0.0564)	0.9683 (0.0541)	0.9544 (0.0647)	0.9301 (0.0796)	0.9288 (0.0803)
ROUNDED n=15	0.9854 (0.0312)	0.9872 (0.0293)	0.9892 (0.0268)	0.9577 (0.0528)	0.9440 (0.0605)
INTERMEDIATE n=10	0.9629 (0.0626)	0.9680 (0.0583)	0.9625 (0.0630)	0.9596 (0.0653)	0.9280 (0.0865)
U-SHAPED n=15	0.9680 (0.0547)	0.9635 (0.0584)	0.8754 (0.1055)	0.9481 (0.0694)	0.8110 (0.1277)
IDEALIZED N-WAVES n=15	0.9536 (0.0633)	0.9680 (0.0528)	0.9364 (0.0738)	0.7897 (0.1290)	0.7560 (0.1376)

FIGURE 1

FIGURE 2A - N-WAVES

FIGURE 2B - PEAKED

FIGURE 2C - ROUNDED

FIGURE 2D - U-SHAPED

FIGURE 2E - INTERMEDIATE

FIGURE 2F - IDEALIZED N-WAVES

FIGURE 3

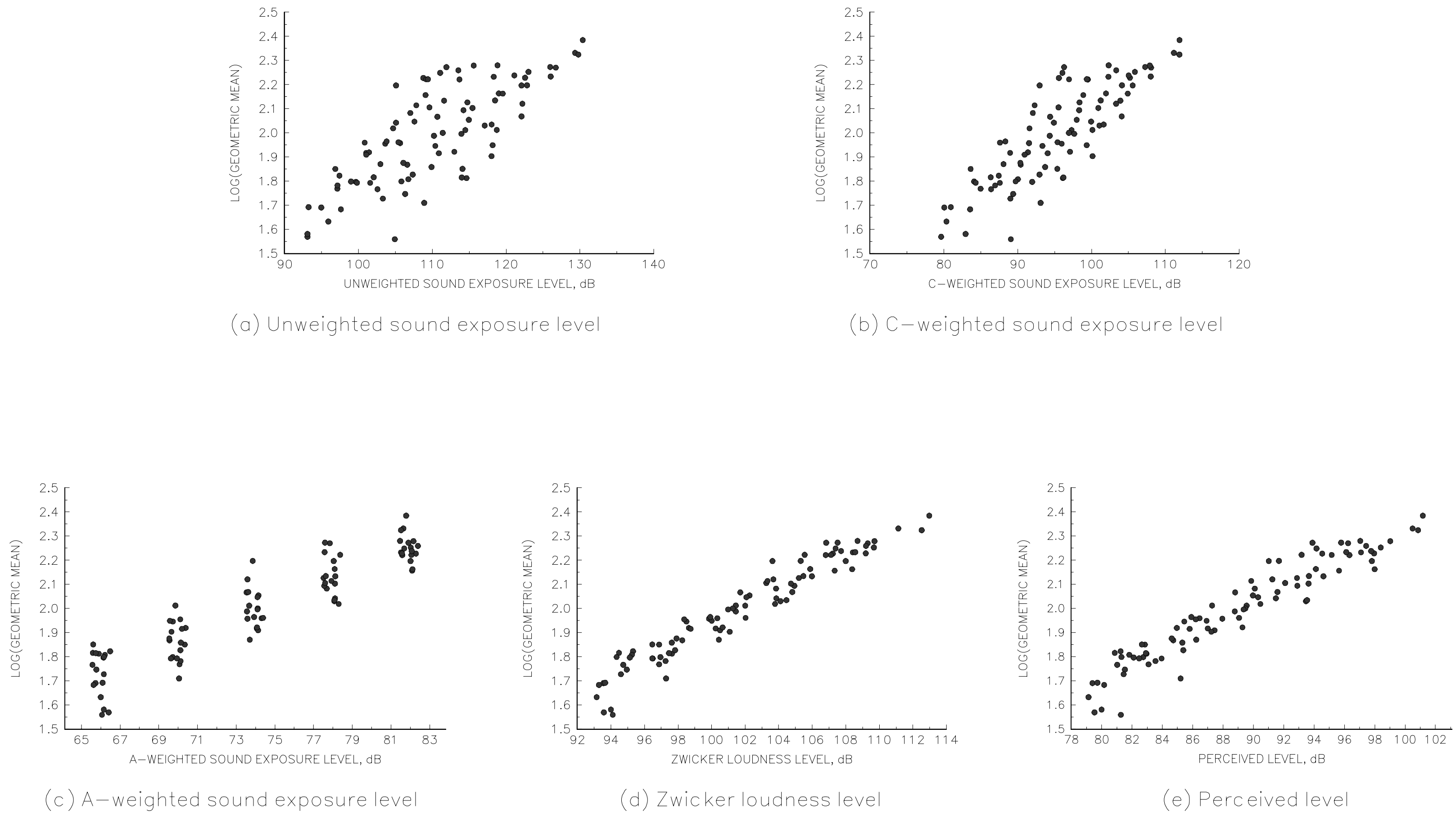

Figure 4.— Comparisons of judged loudness of sonic boom signatures with measured sound exposure levels and calculated loudness levels.

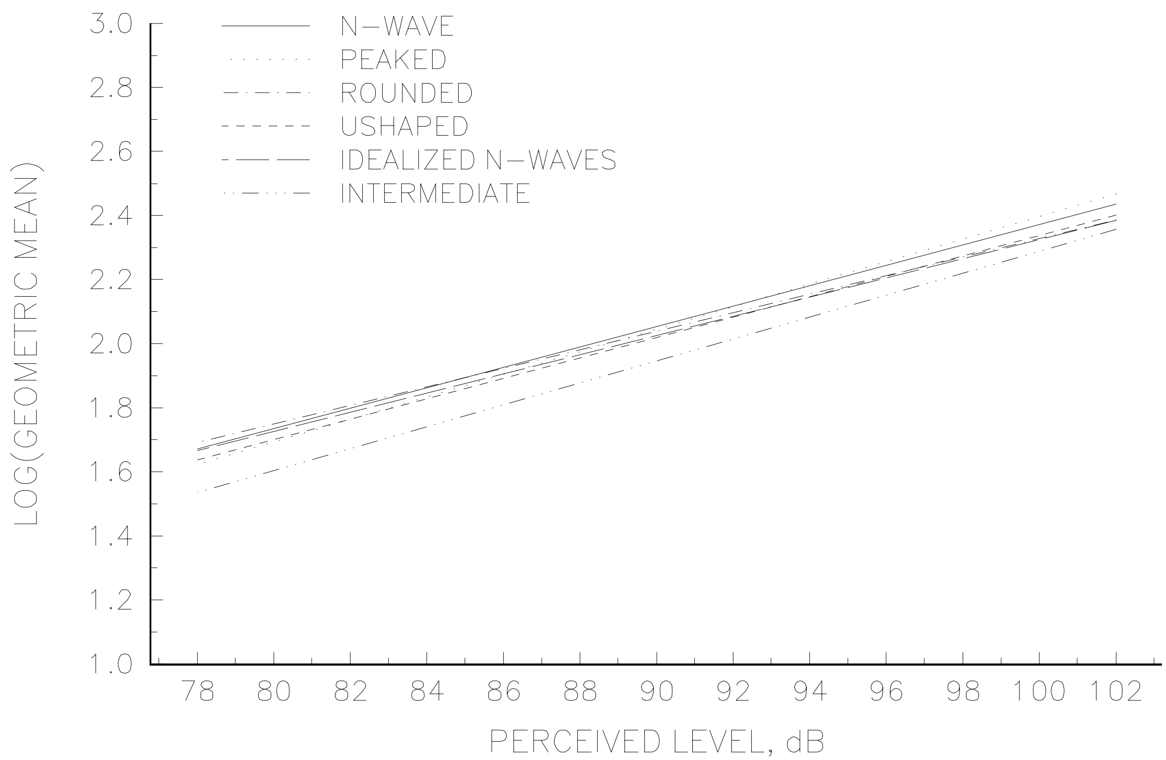

Figure 5.– Linear regression lines for each boom category in terms of the PI metric.

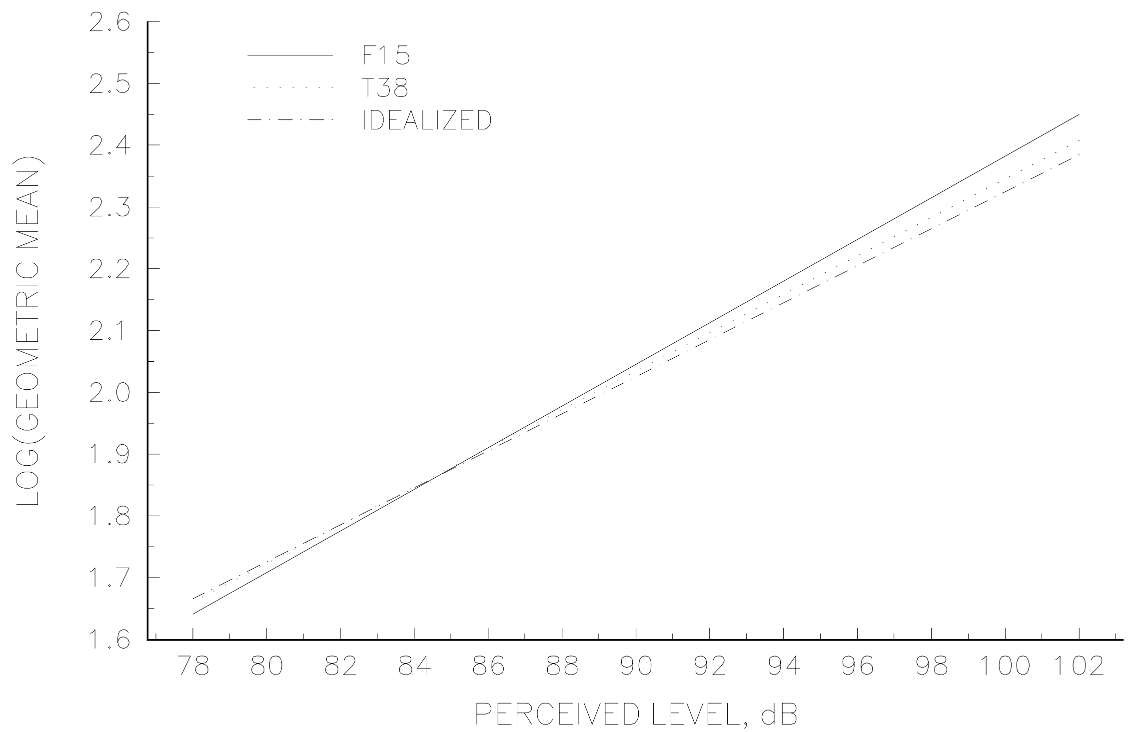

Figure 6.— Comparison of loudness ratings between the two aircraft and the idealized N–waves.