AN OVERVIEW OF METHODS TO QUANTIFY ANNOYANCE DUE TO NOISE WITH APPLICATION TO TIRE-ROAD NOISE

Sponsored by
American Concrete Pavement Association

HL 2008-2   Report 430180000 8000008325-1

Submitted by: Sarah McGuire, Graduate Research Assistant
Patricia Davies, Principal Investigator

Approved by: Patricia Davies, Director
Ray W. Herrick Laboratories

FEBRUARY 2008

Summary

Methods used to quantify annoyance due to environmental noise are described as well as the results of an analysis of vehicle passby recordings. Most methods currently used to quantify the impact of noise are based on average or maximum A-weighted sound pressure levels. It is known that the loudness of a sound is the primary driver in annoyance; however, it is also known that other sound characteristics, such as the presence of tones, fluctuations, spectral balance and impulsiveness, also impact annoyance. While A-weighted sound pressure levels may be strongly correlated to loudness levels for some sets of sounds, the relationship between the two can change if the characteristics of the sounds in the set change. Recently developed models of loudness incorporate many characteristics of the human hearing system including changes in response behavior with increased level, frequency masking and temporal masking. The A-weighting is derived from the 40 phon equal loudness contour and is only truly appropriate for analysis of relatively quiet sounds. In contrast, the loudness models are appropriate for sounds over a wide range of levels and incorporate the change in sensitivity to different frequencies of different sound levels, which is indicated in the equal loudness contours.

In recent laboratory tests statistics of time-varying loudness model predictions, particularly the Loudness exceeded 5% of the time (N_5), have been shown to be more highly correlated to subjects' annoyance ratings of transportation noise signatures than metrics based on the A-weighted sound pressure level. However, when analyzing community survey responses, researchers have typically used A-weighted metrics (e.g.,

L_{dn} or L_{Aeq}) in dose-response relationships and have not kept sufficient noise data to make loudness predictions. Thus it is difficult to examine how statistics of loudness predictions might be used to generate improved noise dose-response relationships. Well designed community surveys, including collection of the corresponding noise data, are difficult to design, as well as being time consuming and expensive to conduct. Before recommending that surveys be conducted to generate the data which would allow new annoyance models to be assessed, it is important to see: (1) whether there are sufficiently large variations in loudness for signals with the same L_{Aeq} level; large deviations for the same L_{Aeq} would indicate that L_{Aeq} is not a good measure of the loudness of a sound; and (2) if other sound attributes vary sufficiently across various road noise measurements to warrant inclusion in a community impact model.

Over 100 recordings of pass-by noise on four different types of pavements were made. For each recording metrics were calculated: those based on A-weighted sound pressure level, those from time-varying loudness models, and the so-called psychoacoustic/sound quality metrics: Roughness, Sharpness, and Fluctuation Strength which predict levels of sound attributes that, in addition to loudness, are commonly perceived by people listening to sounds. The relationship between the calculated metric values, particularly L_{Aeq} and N_5 was examined as was the variation of these metrics across all recordings made. The range of N_5 values was quite large at some L_{Aeq} levels. There were also significant variations in the sound quality metrics for the set of sounds analyzed. The report concludes with recommendations for future work.

Table of Contents

1. Introduction

A-weighted metrics are used to assess road noise and their impact on communities. The Federal Highway Administration recommends that interior noise levels for residences, schools, and buildings do not exceed an interior hourly L_{Aeq} level of 52 dB(A) (U.S. Department of Transportation, 1995). In addition standards for assessing road-noise recommend the measurement and calculation of A-weighted noise metrics. For example, when conducting statistical pass-by recordings, the maximum A-weighted Level is determined. There are, however, many arguments against the use of A-weighted metrics for the assessment of transportation noise. One common argument is that A-weighted metrics do not account for the different factors that affect response to noise, such as annoyance.

Annoyance is a complex phenomenon. Even when restricting oneself to annoyance caused by noise, non-acoustic factors can play a strong role. Memory, expectation and context play a part. Also in some cases fear caused by a sense of danger can intensify reactions. Visual input and the relationship between the person hearing the sound and those associated with or partly responsible for the noise generation, can also affect response. For example, good relationships between a community and an airport can increase acceptance of airport noise.

It is claimed that the prediction of annoyance using noise measurements alone only explains a small percentage of the variance of the responses in community surveys. This might lead one to conclude (wrongly) that there is no point in reducing noise because its effect is overwhelmed by other factors. Yet, from Schultz's (1976) data it is clear that as noise levels (as measured by using L_{dn}) go up the percent highly annoyed increases. Part of the problem is that when people model annoyance and consider non-acoustic factors they are often using relatively simple models that do not take into account the complex interactions that occur when people hear sounds – sounds evoke memories and spawn neural processing which leads to identification of the source and judgments of the sound which factor into the context in which the sound is heard. These highly complex mechanisms are difficult to understand and thus model.

Another problem with noise-response modeling could be the use of very simple measurements of the sound based on A-weighted sound pressure levels. Weightings of sound pressure are based on the information in the equal loudness contours; from the A-weighting based on the 40 phon curve to the C-weighting at very high levels (above 80 phons). Most passby sounds vary in level, so a single weighting is often not appropriate to apply across the whole duration of the sound. Also, these weightings do not account for other attributes of the human hearing system that have been studied and modeled in the psychoacoustics community. It should be noted that psychoacoustics research is a natural progression from the research that led to the equal loudness contours, and has resulted in continuous improvements in models that predict people's perception of sounds. There are now very accurate models that can be used to predict how people perceive the loudness of a sound through time (Zwicker and Fastl, 1998; Moore and Glasberg, 2002). These models have been shown to produce levels highly correlated to people's perception of the loudness of sounds in a variety of applications, yet they are typically not used when evaluating environmental noise or when trying to explain noise-annoyance dose-response relationships.

The question arises why not? Familiarity with A-weighted sound pressure level and ease of calculation, its use in standardized noise measurements, the availability of measurement devices that automatically calculate metrics derived from it, and the guidelines or legal requirements based around A-weighted sound pressure measurements all serve to impede consideration of new community noise impact models. Furthermore, when noise surveys have been conducted little detailed noise information is kept, the derived metrics (usually very few, and these are typically functions of A-weighted sound pressure) at respondees locations are kept but the ability to calculate newer noise metrics that are functions of time-varying loudness have been lost because the original time histories have been discarded. Thus there is insufficient data to validate proposed noise-annoyance models. It is expensive to conduct surveys where people's responses, detailed noise information, as well as information on the community and transportation operations are collected. For model validation a comprehensive set of situations should be captured making the scope of the project very large. While the newer loudness models are accurate, they do require significantly more computation. The question arises whether over the set of sounds that people are likely to hear are there differences in the information captured by the newer metrics that would not be captured in the much simpler A-weighted metrics. If all the sounds heard are very quiet there may be a very high correlation between the two types of metrics and one might conclude that there is no point in pursuing more complicated measurements. However, if you are interested in using a metric that would work for predicting human response to a variety of noise sources (road-tire noise, engine and exhaust noise, aircraft and train noise, power plant noise), there may well be a need for a more accurate model based on an up-to-date understanding of the behavior of the human auditory system, because for many of these sounds the A-weighed sound pressure level is no longer an accurate reflection of how people are hearing and processing the sound.

In most annoyance modeling for environmental sounds focus has been on metrics that quantify the noise levels which is often, but not always correctly, interpreted as measures of the loudness of the sound. However, it is known that other characteristics also influence how acceptable a sound is, for example, the degree and rate of fluctuation in loudness, the presence of tonal components in the sound, and the sound character as reflected in how sharp, rough or harsh it sounds. The gaps in concrete segments in roads create fluctuations in noise level; there is a shift in pitch of the road noise when vehicles move from one type of surface to another. For some vehicles and surfaces there is more low frequency noise generated and for others more high frequencies. Could these characteristics affect annoyance due to tire-road noise, and how do these effects compare with that of loudness alone?

In this research we limited the investigation to the evaluation (through calculation of metrics) of sounds arising from vehicle passbys on a number of different road surfaces and only those passbys that were dominated by tire-road noise. We were interested in answering the following questions: (1) What are the differences between the loudness-based metrics and those based on A-weighted sound pressure? (2) Do the results warrant further investigation of loudness metrics for evaluation of tire-noise dominated passbys? (3) Is there enough variation in other sound characteristics, to warrant further investigation?

In the following, a review of the literature related to annoyance, with emphasis on transportation noise is given. Following is a description of a series of passby measurements that were taken and the results of a metric analysis of those sounds. The report ends with some concluding comments and recommendations for future work.

2. Annoyance Models

Annoyance is the most evaluated impact of transportation noise on communities. It is often assessed using social surveys. Questionnaires or interviews are conducted in order to determine the level of annoyance for areas of a specific noise level. Respondents are asked to rate their degree of annoyance on a scale from not annoyed to highly annoyed. L_{dn}, Day-Night Average Sound Level, is used to determine the noise impact. This metric is based on average A-weighted sound pressure levels and has a 10 dB penalty for events occurring at night between 10 p.m. and 7 a.m. In 1978, Schultz analyzed the data from road, rail, and aircraft noise surveys in order to determine a relationship between the percent highly annoyed and L_{dn}. Figure 1 is the dose-response curve that was developed based on data from 11 noise surveys. This curve has been widely used for assessing the impact of transportation noise. However, there is a considerable amount of scatter of the survey data used in the analysis. For example, for an L_{dn} of 70 dB(A) the percent highly annoyed could be any value between 10 to 40 %.

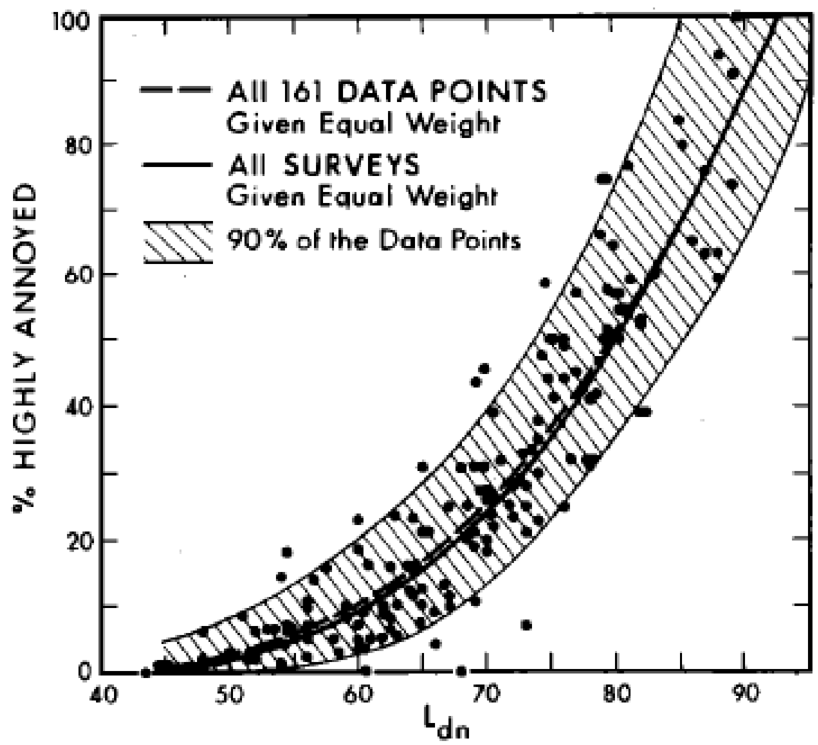

Figure 1: Dose-response curve for transportation noise (Schultz, 1978)

There have been several attempts to improve on Schultz's work. Additional curves have been developed by Finegold, Harris and von Gierke (1994), Fidell, Barber, and Schultz (1991), and Miedema and Vos (1998) using additional survey data sets. Miedema and Vos' is the most recently developed dose-response relationships. They reassessed the surveys used by Schultz, as well as additional noise survey data. In total 22 datasets were used in the analysis. Unlike Schultz, Miedema and Vos did not create one dose-response curve for all types of transportation noise. Rather they found that the percent highly annoyed depended on the noise source. Therefore, they developed three curves one for rail, road, and aircraft noise. The dose-response relationships are indicated in Figure 2. They determined that aircraft noise was more annoying, then road and rail noise.

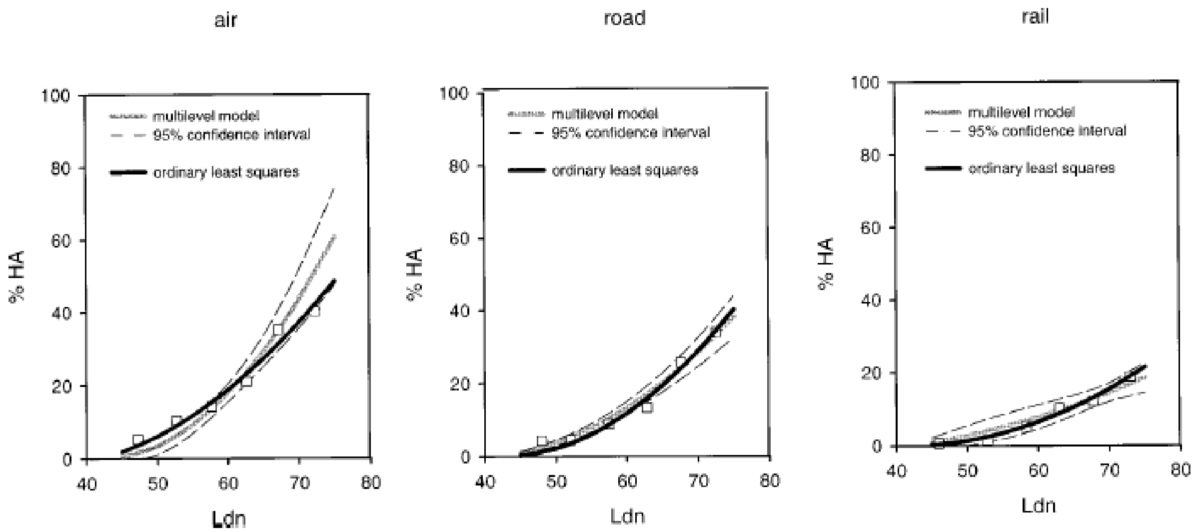

Figure 2: Dose-response relationships for aircraft, road, and rail noise (Miedema and Vos, 1998)

Results of more recent studies, have questioned the applicability of Miedema and Vos' curves. For example, annoyance studies conducted in Norway found that a higher percentage of people were annoyed then would be predicted by Miedema and Vos' dose-response relationships (Klaeboe, Amundsen, Fyhri, and Soldberg, 2004). In addition studies have been conducted in order to determine whether road noise is actually more annoying then rail noise. The results of a laboratory test conducted by Fastl, Fruhmann, and Ache (2003), indicated that sounds of the same Loudness, not dB(A) level, were judged equally annoying. Therefore the differences in annoyance found in the curves may be simply an artifact of the metric being used.

In addition to assessing additional survey data sets in order to determine an accurate dose-response relationship; attempts have also been made to increase the correlation between individual response data and noise level, by improving L_{dn}. In 1974, the Environmental Protection Agency (EPA) suggested several adjustments that should be made to L_{dn}. These included adjustments for different seasons, the amount of background noise, community attitudes and previous exposure to the noise, as well as adjustments for tonal and impulsive sounds. Schomer (2002) expanded on this idea of improving L_{dn} and came up with a revised list of adjustments. Schomer revised the levels of the EPA adjustments and also added additional adjustments in order to account for the increase in annoyance caused by noise-induced rattle, and also to account for differences in time duration over which L_{dn} is calculated.

2.1 Non-acoustic Factors and Annoyance

Schomer's adjusted L_{dn} contained adjustments for both qualities of the sound as well as for situational and attitudinal variables. This raises the question, of whether the variation in individual annoyance for the same L_{dn} is due to the use of a noise metric which does not account for the different qualities of the sound that affects annoyance, or is it due to other non-acoustic factors. While it is most likely that both components play an integral role, the majority of more recent research on environmental noise has focused on identifying non-acoustic factors that affect annoyance.

It has been suggested that the level of annoyance to transportation noise is dependent on the specific activities that are interrupted. The activities that can be affected include sleep. Transportation noise can cause a delay in sleep onset, an increase in awakenings and motility, a change in sleep structure, and a decrease in subject evaluated sleep quality (Griefahn, Marks, and Robens, 2006). In addition transportation noise can disturb activities such as reading, and interfere with communications. Taylor, Hall, and Birnie (1987) created a probabilistic model based on reported speech interference inside and outside, and awakenings, and found that these effects could be used to reasonably predict general annoyance to road traffic noise.

Table 1: Personal variables that affect annoyance.

Variable	**Change in Annoyance**	**Description**
Fear	Increases	Fear of the noise source
Capacity to cope with the noise	Decreases	Ability to adapt to the noise
Preventability	Increases	A belief that the level of noise can be prevented
Belief that source is unhealthy	Increases	A belief that the noise is impacting ones' health
General Annoyance to Source	Increases	Annoyance to the source in addition to the noise it produces
Noise Sensitivity	Increases	General sensitivity to noise
Personal Benefit	Decreases	e.g., frequent use of the highway or road
Recognition	Decreases	A belief that officials recognize the noise problem

Fields' (1993) has examined reports for 282 environmental noise surveys and found that several attitudinal variables also affect annoyance. These include noise sensitivity, a fear of the noise source, and a belief that the noise can be prevented. Miedema and Vos (1999) similarly found that the predominate factors affecting response were noise sensitivity and fear. Other factors that have been suggested include the predictability of the noise (whether the noise level is increasing over time), and whether there is a perceived benefit of the noise such as close access to the highway (Flindell and Stallen, 1999). Demographic variables such as age, gender, and education have not been found to have a strong affect on response. Table 1 is a list of personal variables that have been suggested as having an impact on annoyance.

2.2 Acoustic Factors and Annoyance

Acoustic factors also effect how a sound is perceived. One component of sound that is known to impact annoyance is low frequency noise. Low frequency noise is often considered to be sounds below 250 Hz. Several laboratory studies have been conducted to determine how annoyance is affected by low frequencies. Persson and Björkman (1988) conducted a study using broad band fan noise centered at 80, 250, 500, and 1000 Hz. They found that the noise centered at 80 Hz was found to be more annoying then the other three sounds. In terms of road noise Nilsson (2007) assessed the annoyance and perceived loudness of road traffic sounds with low, medium, and high levels of low frequency noise. Subjects were asked to assess the annoyance and loudness of the recordings on a magnitude scale. The perceived loudness and annoyance were compared to the A-weighted sound pressure level of the sounds. This comparison indicated that sounds with a high level of low frequency noise were considered both louder and more annoying then sounds of the same A-weighted sound pressure level with low and medium levels of low frequency noise.

Another factor that can impact annoyance is the tonality of the sound. Annoyance does increase when a tonal component is present. However, the exact nature of that increase is dependent not only on the frequency of the tonal component, but also on the level of the tone, the overall level of the entire sound, if other tones are present, and on masking effects (Hellman, 1984). Tonality is an important component of pavement noise. Transversely grooved concrete pavements in particular have been known to cause a "whine" due to prominent tones. A field study has been conducted in New Jersey, in which the tonality of different pavements, including transverse grooved, longitudinal grooved, and asphalt pavements were assessed. For this study, Billera, Schmidt, and Miller (1997) defined their own tonality metric; it was related to the ratio of the level of the tone to the surrounding noise. Only by using both their defined metric, and the average A-weighted sound pressure level, were they able to fully assess the variations in sound produced by the different pavements.

Studies on HVAC noise have indicated other acoustic factors that may affect response to noise. In addition to tonality, and low frequency noise, the spectral balance causes perception of the sound to change (Blazier Jr., 1981). Also, from results of laboratory tests it has been concluded that annoyance depends on the variations in the loudness of a sound. Annoyance increases for larger fluctuations in level, and also varies for different modulation rates (Bradley, 1994).

2.2.1 Loudness and Psychoacoustic Metrics

There are various psychoacoustic metrics, which quantify these specific qualities of sound. They include two metrics which account for how people perceive loudness.

While originally developed for stationary noise sources, these loudness models have been updated and improved so that they now predict loudness as a function of time, i.e., how humans perceive the loudness of a sound as its characteristics vary. There is a model by Moore and Glasberg (2002), a development of their earlier stationary model (Moore and Glasberg, 1996), and also a model by Zwicker (Zwicker and Fastl, 1999) that builds on the stationary loudness calculation described in Part B of (ISO532, 1975). They both incorporate nonlinear and frequency and temporal masking effects of the human hearing system. The primary difference between the two models is how they treat low frequency noise, Moore and Glasberg's model perhaps having more widespread support in the psychoacoustics community and Zwicker's model used more widely in the engineering community. When assessing the loudness of time-varying sounds, the statistics such as the maximum level or percentile loudness are calculated. For sounds whose level varies through time, N_5, the loudness exceeded 5% of the time, has been found to be a robust indicator of perceived overall loudness of the sound event. As sounds become more impulsive (quicker onset and rise time, and shorter duration), higher levels of the loudness, N_2 or N_{max}, may be more appropriate measures of the loudness of the event.

Refer to (Zwicker and Fastl, 1999) for information on the following sound attributes and the models used to predict the attribute strength. When people listen to a variety of sounds and are asked to describe or rate them, an analysis of the results will almost always show that people respond spectral balance (dull, sharp, metallic), fluctuation (unsteady, varying), tones (hum, screech), roughness (rough, smooth) and impulsiveness (startling, sudden). The sharpness of a sound is related to the spectral balance. As the amount of high (above 1000 Hz) versus low (below 1000 Hz) frequency energy increases the sharpness increases, and as energy above 3000 Hz increases the sharpness increases very rapidly. In the model used to predict sharpness, the frequency location of the centroid of the loudness spectrum after weighting frequencies above 3000 Hz more strongly, is determined. There are also two metrics which are related to the variation in loudness of a sound; they are Fluctuation Strength and Roughness. A modulation frequency of approximately 20 Hz serves as a transition between the two metrics. The Fluctuation Strength accounts for lower fluctuation rates where a listener can track the variations in loudness. It reaches a maximum for modulation frequencies of 4 Hz. An example of a sound with a high Fluctuation Strength would be a car alarm. Roughness accounts for faster variations in loudness. It reaches a maximum for modulation frequencies of approximately 70 Hz. Unlike Fluctuation Strength, the actual variation in loudness can generally not be tracked, but gives a rough character to the sounds. An example of a sound that would have a high Roughness value would be the sound of paper tearing.

In order to account for the impact of all four of these sound attributes on annoyance, a Psychoacoustic Annoyance model was developed (Zwicker and Fastl, 1999); this is a refinement of the Unbiased Annoyance model that Zwicker developed. The Psychoacoustic Annoyance (*PA*) model is a function of the Loudness exceeded 5% of the time (N_5), Sharpness (S), Fluctuation Strength (F) and the Roughness (R) of the sound:

$$PA = N_5\left(1 + \sqrt{w_S^2 + w_{FR}^2}\right) \qquad (1)$$

where

$$w_S = 0.25(S - 1.75)\log_{10}(N_5 + 10) \quad S > 1.75,$$
$$w_S = 0, \quad S < 1.75, \tag{2}$$

and

$$w_{FR} = \frac{2.18}{(N_5)^{0.4}}(0.4F + 0.6R). \tag{3}$$

This model is not widely used, but there are examples where it has been used to explain response behavior in annoyance studies for various types of transportation noise, see, e.g., (Zwicker and Fastl, 1999; More and Davies, 2007). If the fluctuations in the sound were insignificant ($R=F=0$) and the Sharpness (S) is below 1.87 then the output would be equal to N_5. The difference between Psychoacoustic Annoyance and N_5 is therefore an indication of how these other sound attributes might influence annoyance.

One sound attribute that this model does not account for, but is related to annoyance, is the tonality of the sound. There are several metrics that could be used to assess the tonalness of sounds. The Tone-to-Noise Ratio (ANSI, 1995), the Prominence Ratio (ANSI, 1995), and the Joint-Nordic method (Pedersen, Søndergaard, and Andersen, 1999) are different ways of comparing the level of a tone to the level of the surrounding noise. Aures' Tonality (von Aures, 1985) is a more sophisticated psychoacoustics-based model that takes into account the variation in tonalness with frequency, the bandwidth of the tonal feature, and the prominence of the feature above the surrounding noise floor after frequency masking has been taken into account. The ratio of the addition to loudness due to the tonal components, to the loudness of the total signal, is also part of Aures' model.

2.2.2 Road Noise and Psychoacoustic Metrics

Several studies have been conducted to assess the use of Loudness metrics when evaluating road traffic noise. Fastl (1989) conducted an experiment using road recordings of 17 minutes in duration. Subjects were asked to continuously assess the loudness of the recordings. Fastl determined that the Loudness exceeded 4% of the time, based on Zwicker's model, was able to account for the perceived loudness of road noise. He also determined that metrics such as L_{Aeq} underestimated the perceived loudness. The results of a more recent experiment by Fastl, Patsouras, Bayer, and Beckenbauer (2007) support those earlier results. An experiment was conducted using 4 second recordings of individual car passbys. These recordings were made for different sound absorbing road surfaces. Subjects were asked to evaluate the loudness of the recordings. There was close agreement between subjective evaluations of the loudness and of N_5. In addition to Fastl's experiments, Golebieweski, Makarewicz, Nowak, and Preis (2003) also assessed the use of Zwicker's Loudness. They conducted a subject test using individual pass-by recordings for dense and porous asphalt surfaces. Subjects were asked to rate their annoyance to each sound. Loudness was found to have a stronger correlation to annoyance then *SEL(A)*.

Few tests have been conducted which not only assessed the use of Loudness metrics but also the use of other psychoacoustic metrics as well. A subjective test was recently conducted by Raggam, Cik, Hoeldrick, Fallast, Gallasch, Fend, Lackner and Marth (2007) in which discomfort to road noise and the relationship to different psychoacoustic metrics were examined. They conducted a subjective test using road recordings from three different surfaces: concrete, asphalt-concrete and stone matrix asphalt. Three minute segments of recordings were presented to subjects. Subjects were asked to rate the level of discomfort they experienced when listening to the sounds. It was determined that Loudness was more correlated to subject ranked discomfort then the A-weighted sound level. Also it was determined that Roughness and Sharpness were reasonably correlated to subjects' discomfort level.

2.3 Concluding Comments on Annoyance Modeling and Metrics

In tests where people have compared responses to passby noise to both A-weighted and Loudness based metrics, the Loudness metrics typically are more highly correlated to people's perceptions. While loudness is clearly a dominant factor in annoyance, other sound attributes can also increase annoyance levels. Some of the normalization factors in Schomer's model (Schomer, 2002) are related to sound attributes other than loudness (presence of tones, impulses etc.) and when he accounts for them in his model, there is a reduction in the unexplained variance in survey responses. In the psychoacoustics and sound quality literature there are numerous examples supporting the use of measures of additional sound attributes in human response (e.g., annoyance, discomfort) models. In the few road noise studies where sound attributes other than loudness have been considered, there is evidence that these attributes may be playing a role in people's responses.

3. Road Noise Measurements

To assess the different qualities of tire pavement noise which may result in different annoyance levels, sound recordings were made of individual vehicle pass-bys on four different road surfaces. The four pavements were porous friction course asphalt (PFC), stone matrix asphalt (SMA), densely graded asphalt (DGA), and transversely grooved concrete. Two of the surfaces, SMA and PFC, were on I74 East near Indianapolis, the section of DGA tested was on US52 in Lafayette and the section of transversely grooved concrete was on I62 in Evansville. Pictures of three of the road surfaces are in Figure 3.

a) b) c)

Figure 3: Tested pavements: a) densely graded asphalt (DGA), b) transversely grooved concrete and c) porous friction course asphalt (PFC).

3.1 Passby Recording Methodology

The sound recordings were made using a Brüel & Kjær Sound Level Meter Type 2250, with a Brüel & Kjær 4189 ½ inch microphone. Calibration was performed before each set of measurements using a Brüel & Kjær 4228 pistonphone. For the recordings the microphone was positioned 7.5 m from the center of the nearest lane, and 1.2 m above the road surface. Recordings of both cars, and heavy trucks were made. The resulting set of recordings was evaluated for quality. Those recordings that contained high levels of background noise, due to traffic on the opposite lanes were not used. In addition, recordings that were dominated by high levels of engine or muffler noise, or sounds that resulted from other abnormal characteristics were not further analyzed.

3.2 Analysis of Recordings

For each recording several A-weighted noise metrics including L_{Amax}, *SEL(A)*, and L_{Aeq} were calculated, as well as various psychoacoustic metrics. In addition, several of the tonality metrics were calculated. However, the values of the tonality metrics were often found to be zero; for the pavements tested one surface was not found to be more tonal then the others and hence results are not shown in the following figures. However, the number of pavements tested was small, and it is known that tonality can be an important component of tire-pavement noise. Plots of the rest of the calculated metric values for each surface are in Figure 4.

There was large spread of the metrics calculated for each surface, due to differences in vehicles, tires and speed. However, despite the large spread of data, there are still apparent differences in metric values for the surfaces tested. For example, the Fluctuation Strength tended to be higher for PFC than the other three road surfaces, and the Psychoacoustic Annoyance showed a trend of being higher for SMA. However, these trends are based on vehicles traveling at different speeds, so it is not possible to make direct comparisons between surfaces where speeds were significantly different. For example the vehicles on the concrete surface were traveling approximately 45 mph which was significantly slower then the speed of vehicles on the other three surfaces.

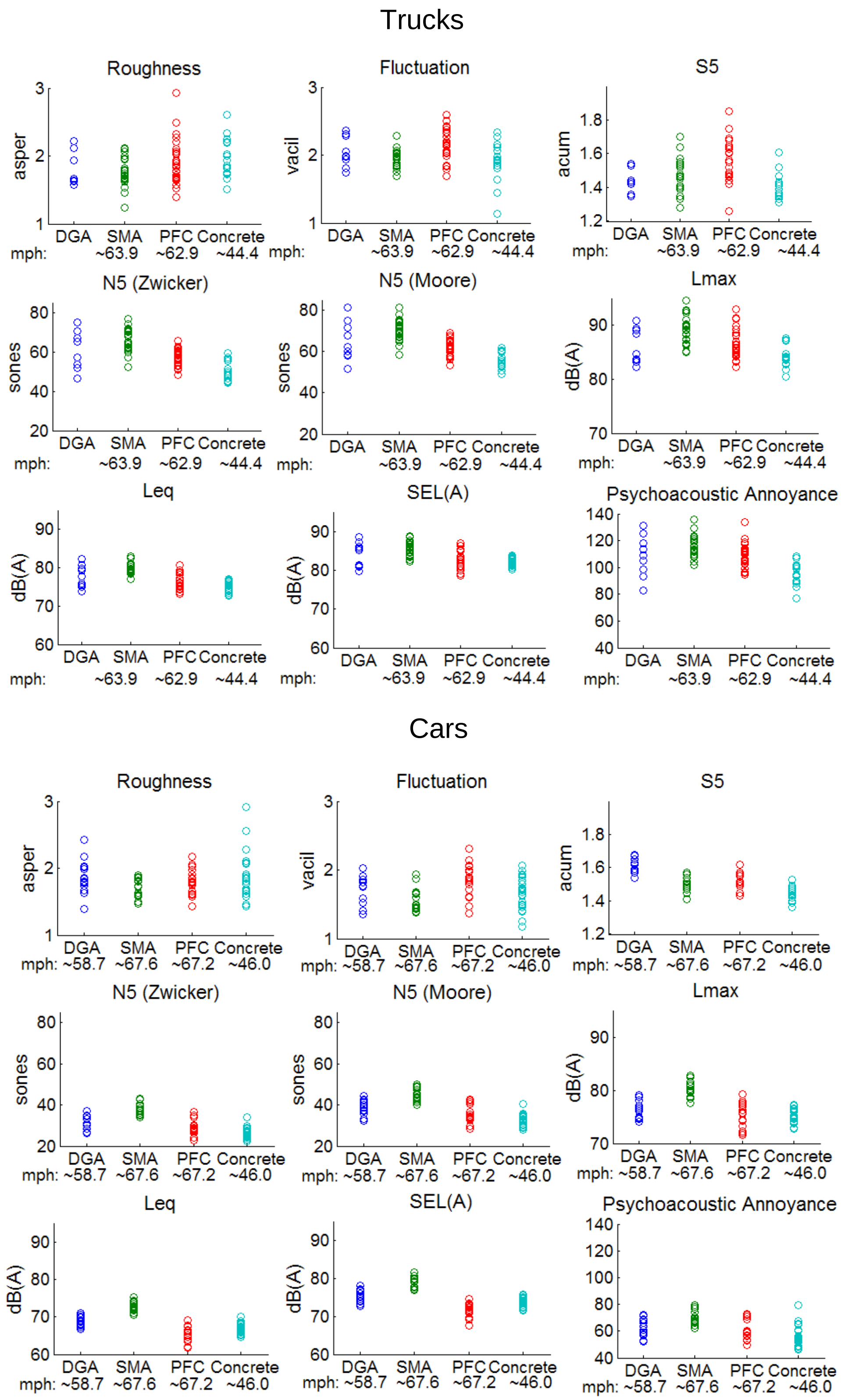

Figure 4: Noise metrics for individual truck and car pass-by events. Average speed of vehicles on each surface is indicated on the horizontal axes. The lower levels for passbys on concrete are probably due to the lower speed of the vehicles.

By comparing the Zwicker N_5 values and the Psychoacoustic Annoyance values, it can be seen that there is almost a doubling of values, from which it can be concluded that sound attributes other than loudness have a strong impact on Psychoacoustic Annoyance. As mentioned above, if these other attributes were not playing a role, Psychoacoustic Annoyance and N_5 would be the same. While this Psychoacoustic Annoyance model has, in the past, provided insight into people's responses in psychoacoustics tests, it should be emphasized that the model is very much a *model in development*. It may well be the case that the relative importance of the attributes may change depending on the range of sounds that are being evaluated. In situations where sounds are of similar loudness, these other attributes may be weighted more heavily than in situations where the sounds heard are of very different loudness levels. Figure 5 is a plot of Psychoacoustic Annoyance compared to *N5* values for the recordings. The Psychoacoustic Annoyance values are much higher then N5 due to the other sound attributes specifically Fluctuation and Roughness. Also for sounds of the same N5 value, the variation in *PA* can be as large as 10-20.

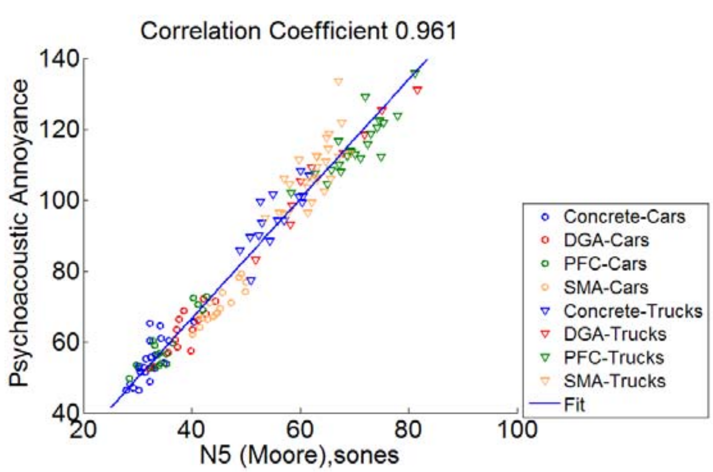

Figure 5: Correlation between Psychoacoustic Annoyance and N5

It is widely known that tire-pavement noise increases with vehicle speed. The increase in sound pressure level with speed is dependent on the particular road surface; however it approximately increases by 0.2 dB for every mile per hour increase in speed. (Bennert, Hanson, Maher, and Vitillo, 2005). In order to get an approximation of the variation in loudness that would occur if all of the vehicles were traveling the same speed, the relationship of 0.2 dB/mph was used in order to normalize the recordings. The vehicles were normalized to 60 mph. N_5 was re-calculated for both the car and truck recordings. The mean N_5 value and the standard deviation of the mean are shown in Figure 6. The noise produced by vehicles on the concrete surface was found to have a similar N_5 value to the noise produced by SMA. However, these results are just a rough approximation of the variation in loudness that may occur. The spectrum of the signal would change with speed and therefore affect the loudness estimates. Due to these expected changes, other psychoacoustic metrics were not calculated for these normalized recordings.

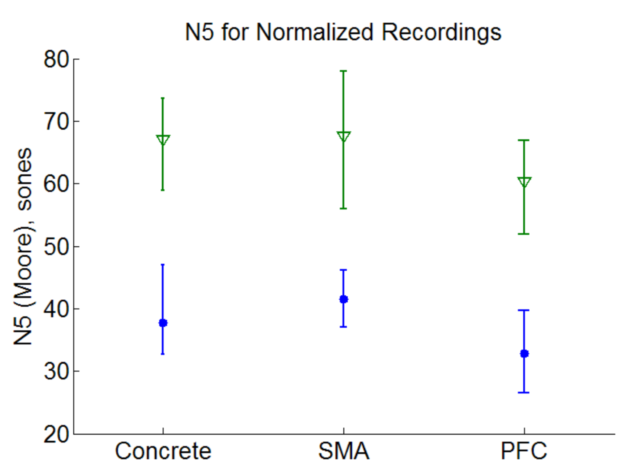

Figure 6: N_5 levels for recordings normalized to the same speed based on a 0.2dB /mph increase model. Green lines are for normalized truck recordings and blue lines are for normalized car recordings.

In addition to evaluating the variation in metric values for different road surfaces, the correlation between A-weighted metrics and Loudness metrics was also assessed. The results are shown in Figure 7. There is reasonable correlation between the Loudness and A-weighted metrics. However, there is also significant scatter in the Loudness versus A-weighted metric plots about the best fit line. Therefore, sounds of the same Loudness can often have a wide range of A-weighted sound pressure levels. In order to assess the amount of variation, the sounds were grouped together according to loudness. Signals within each grouping did not have a variation in Loudness of more then 1 sone. The variation in L_{Aeq} and L_{Amax} for each of these groups are plotted in Figure 8. Variations in L_{Aeq} and L_{Amax} as large as 7-8 dB(A) did occur for sounds of approximately the same N_5 and N_{max} level, respectively.

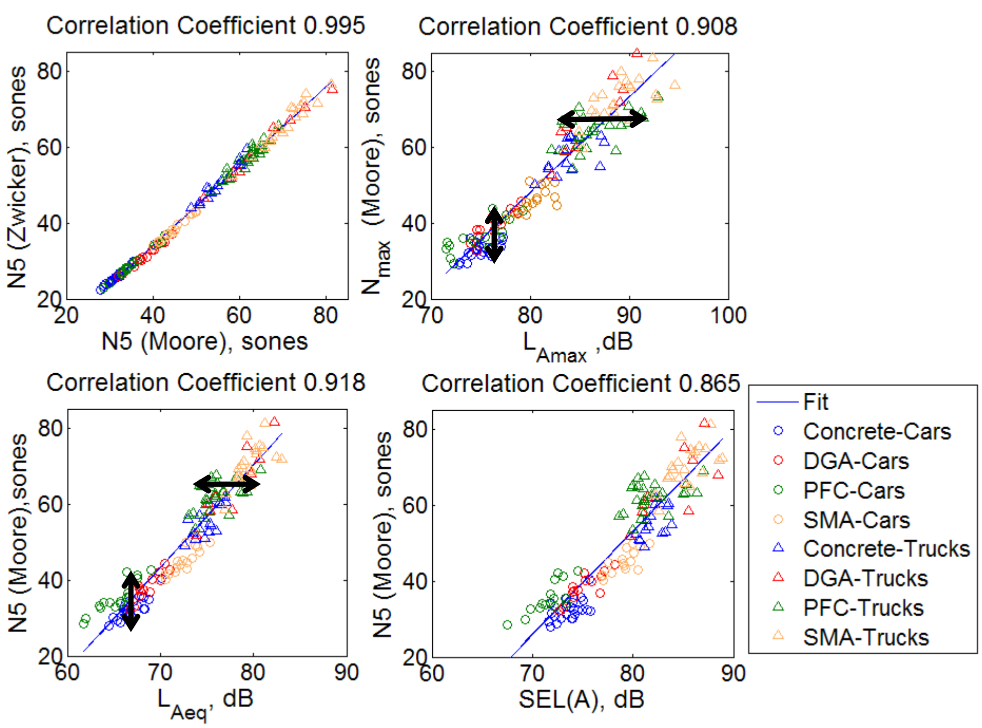

Figure 7: Correlation between Loudness metrics and A-weighted metrics. Note the significant variations in Loudness for the sounds with the same L_{Amax} and L_{Aeq}.

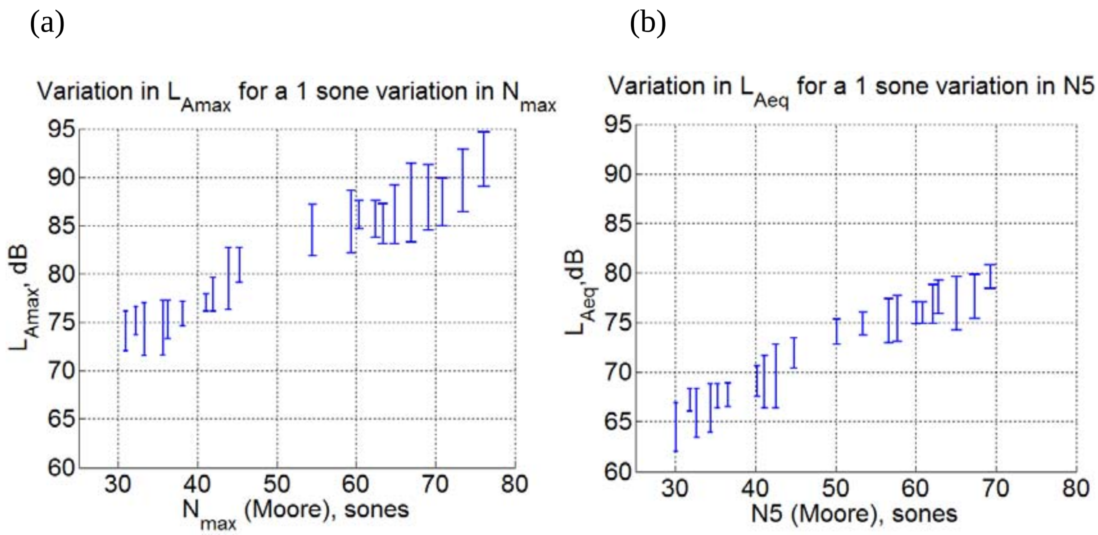

Figure 8: Variation in A-weighted metrics for a 1 sone variation in Loudness metrics. (a) L_{Amax} *vs.* N_{max} and (b) L_{Aeq} *vs.* N_5.

3.3 Twenty-Four Hour Simulations of Traffic

The metrics that were calculated and compared above were for single pass-by events. However when assessing the annoyance of communities due to road noise, $L_{dn,}$ or metrics based on the average daytime noise levels are typically used. Therefore, a 24 hour simulation of vehicle traffic was created in order to compare *N5* to L_{dn}. For the simulation, it was assumed that there were no overlapping events, only one vehicle passed by at a time. The recordings were 3 seconds in length, so the maximum number of events that could occur per hour was 1200. The pattern of vehicles per hour for a 24 hour time period was determined from data on hourly traffic flow obtained from the Illinois Department of Transportation. For the simulation the truck and car recordings for a specific surface were randomly selected. The number of each selected was based on the percentage of heavy vehicles chosen for the particular simulation. After randomly selecting the vehicle recordings, L_{dn} and the N_5 for the entire period were calculated. The simulation was then repeated 100 times to determine the variation in daytime noise levels due to the selection of vehicle recordings. Simulations were also conducted for different percentages of heavy vehicles varying from 5% to 50%. Furthermore, it was repeated using either only the loudest third or the quietest third of the truck recordings, therefore simulating a best and worst case noise scenario. The results of the simulation are in Figure 9. As indicated, when using L_{dn} the levels for passbys on SMA and concrete surfaces are found to be quite similar. However, there is more apparent variation between the levels of the two surfaces when N_5 is used and there is a strong linear relationship between *N5* and the fraction of cars on the road.

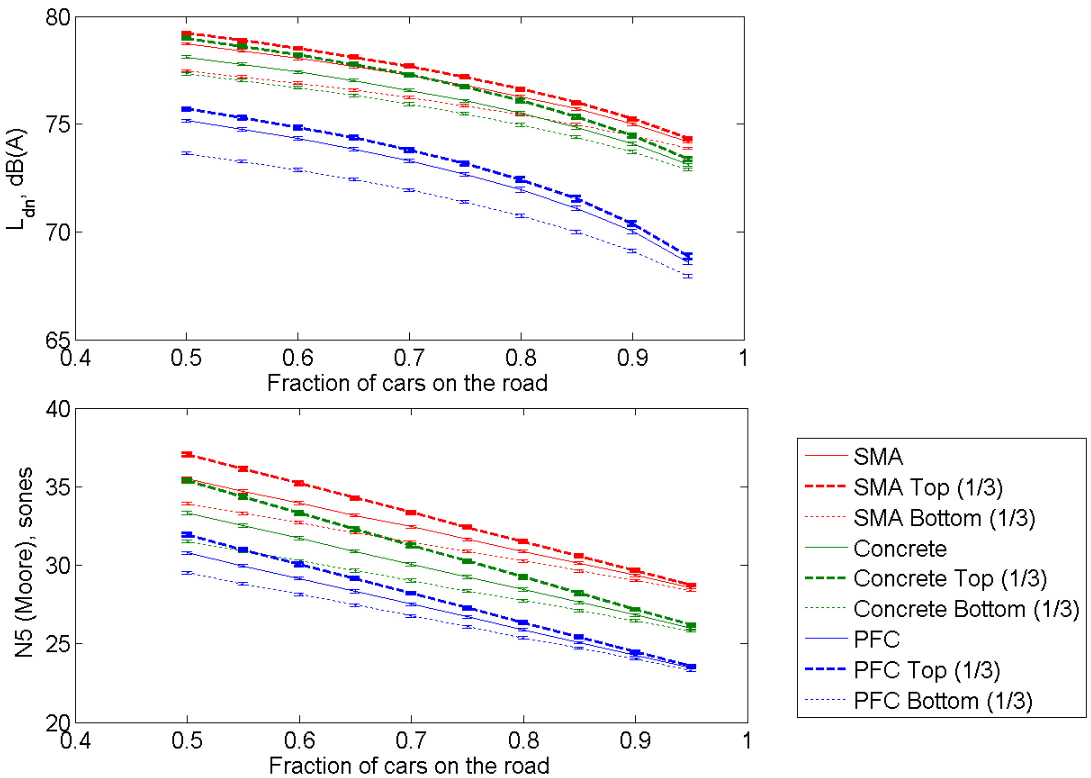

Figure 9: L_{dn} and N_5 for daytime simulations.

3.4 Discussion on Analysis of Pass-by Recordings

The metric analysis was conducted for a limited number of road surfaces. However, the results of the analysis of individual vehicle pass-by recordings indicate that different pavements do result in differences in the quality of the sounds. Even the very limited analysis of passby noise described in this report demonstrates that there are significant variations in sound attribute strengths for different vehicles and different road surfaces. The perceived strength of sound attributes, such as loudness, fluctuation, roughness and sharpness cannot be quantified by using A-weighted noise metrics, but can be assessed by using various psychoacoustic metrics. Assessing the strength of these sound attributes in addition to loudness is important when trying to determine the degree of annoyance that a community might experience. Results of past research that led to the development of the Psychoacoustic Annoyance model indicate that these variations could lead to very different levels of annoyance even when sounds are of the same loudness.

4. Concluding Comments and Recommendations for Future Work

There is sufficient evidence in the literature to warrant examination of the use of loudness metrics in noise-annoyance dose-response relationships related to road noise. When sounds contain high levels of low frequency energy, it is difficult to defend the use of A-weighted sound pressure levels in quantifying the noise level. Time-varying loudness models include the nonlinear behavior of the human hearing system and can predict the changes in frequency weighting (C- vs. A-) as sounds become louder and quieter. In vehicle passbys the sound can vary from very quiet to very loud and thus there is no single weighting that is appropriate to use throughout the sound event. Time-varying loudness models naturally account for these different auditory responses at different noise levels.

From an examination of a large set of vehicle passby recordings dominated by road-tire noise, it was found that the loudness of the passby (as measured by using Loudness exceeded 5% of the time – N_5) often varied considerably even when the sounds had the same average or peak A-weighted sound pressure level. There is evidence that these loudness models are very accurate predictors of the loudness of sounds and hence this variability is an indication that the A-weighted sound pressure level is not a robust predictor of loudness for this set of sounds.

In the analysis large variations in other sound attribute metrics were found and it is possible that these sound attributes may also strongly affect annoyance. One model in the literature – Psychoacoustic Annoyance (Zwicker and Fastl, 1998) – predicts that these attributes cause an almost doubling of the predicted annoyance level (compare the N_5 and Psychoacoustic Annoyance values in Figure 5).

4.1 Future Work

Ultimately the design and execution of a set of community surveys to populate a database would be desirable. Such a database would include community information, traffic and road surface information, noise and survey responses, and could be used by various researchers to validate and develop models to explain community response to noise. Use of a common database would allow for direct comparison of different models. However, this is a very large undertaking and several issues need to be addressed before surveys can be designed so that sufficient data is gathered for testing proposed models.

A traffic noise tool that enables prediction of sound time histories (that could be played back to people) at various locations in a community is desirable. These time histories could then be analyzed to derive whichever metrics are of interest. Current traffic noise prediction models are limited and it is not possible to calculate time-varying loudness or other psychoacoustic metrics with the amount of information retained in these models. This capability would be useful to people developing more accurate models of sound attributes and the impact that they have on annoyance, speech communication, learning and health, for example.

Building on the sound prediction capability, a survey simulation tool would be useful. This would involve predicting the sound at points on a grid; developing a comprehensive set of candidate models for community impact; choosing a sample of the population (based on noise, health, economic and demographic information) to survey and doing a virtual survey. An analysis of the data gathered from the virtual survey would determine whether it is possible to retrieve the dose-response relationships used in the simulation. Thus the survey simulation tool and subsequent analysis could be used to help in the design of a survey to ensure that cause-effect relationships can be identified with some degree of accuracy if a particular sampling strategy is adopted and traffic conditions are of a particular type.

There is also a need to have a time-history database of passby recordings on different surfaces for different vehicles. In particular, for concrete road surfaces a database would need to contain recordings for different concrete textures; so that the texture that creates not only the lowest noise level but causes the least amount of annoyance could be determined. In addition the recordings that have been analyzed in this paper were not tonal. However, passby noises from road surfaces such as transversely tined concrete do contain a well-defined pitch. A large database, which included recordings for these types of surfaces, would allow the tonality of pavements to be assessed.

Also passby recordings could be used to perform comparisons between road noise and other transportation noise. Miedema and Vos indicated that the level of annoyance depends on the source of the noise. They found, that aircraft noise was considered more annoying then road noise which was more annoying then rail noise. Laboratory studies comparing these three transportation noises could be conducted in order to determine whether the difference in annoyance is due to knowledge of the sound source or whether it is due to differences in sound characteristics.

Finally laboratory studies need to be conducted in order to determine if roughness and other sound attributes play a role in annoyance responses to road traffic noise, or whether the response is dominated by loudness because of the large range of levels produced by different vehicles. The results of such studies could be used to validate and/or improve the Psychoacoustic Annoyance model for use in predicting response to passby noise.

References

American National Standard Measurement of Sound Pressure Levels in Air, American National Standards Institute ANSI S1.113-1995.

T. Bennert, D. Hanson, A. Maher, and N. Vitillo, "Influence of pavement surface type on tire/pavement generated noise," *Journal of Testing and Evaluation* **33(2)** (2005).

D. Billera, B. Schmidt, and W. Miller, "Stop the Whine! Narrow band noise level measurements of three highway pavements," *Transportation Research Record* **1601**, 71-76 (1997).

W. E. Blazier Jr. "Revised noise criteria for application in the acoustical design and rating of HVAC systems," *Noise Control Eng. J.* **16(2)**, 64-73 (1981).

J. S. Bradley, "Annoyance caused by constant-amplitude and amplitude-modulated sounds containing rumble," *Noise Control Eng. J.* **42(6)**, 203-208 (1994).

H. Fastl, "Average loudness of road traffic noise," *Proceedings of Inter-noise 89*, Newport Beach California, 815-820 (1989).

H. Fastl, M. Fruhmann, and S. Ache, "Railway bonus for sounds without meaning?" *Acoustics Australia/WESPAC-VIII* (2003).

H. Fastl, C. Patsouras, S. Bayer, and T. Beckenbauer, "Psychoacoustic evaluation of sound absorbing road surfaces," *Proceedings of the 19th International Congress on Acoustics*, Madrid, Spain (2007).

S. Fidell, D.S. Barber, and T.J. Schultz, "Updating a dosage-effect relationship for the prevalence of annoyance due to general transportation noise," *J. Acoust. Soc. Am.* **89(1)**, 221-233 (1991).

J. M. Fields, "Effect of personal and situational variables on noise annoyance in residential areas," *J. Acoust. Soc. Am*. **93(5)**, 2753-2763 (1993).

L.S Finegold, C.S. Harris and H.E.von Gierke, “Community annoyance and sleep disturbance: updated criteria for assessing the impacts of general transportation noise on people,” *Noise Control Eng. J.* **42(1)**, 25-30 (1994).

I. H. Flindell and P.J. M. Stallen, “Non-acoustical factors in environmental noise,” *Noise & Health* **3,** 11-16 (1999).

B. R. Glasberg and B. C. J. Moore. “A model of loudness applicable to time-varying sounds,” *Journal of the Audio Engineering Society*, **50(5)**, 331–42 (May 2002).

R. Golebiewski, R. Makarewicz, M. Nowak, and A. Preis, “Traffic noise reduction due to the porous road surface,” *Applied Acoustics* **64,** 481-494 (2003).

B. Griefahn, A. Marks, and S. Robens, “Noise emitted from road, rail, and air traffic, and their effects on sleep,” *Journal of Sound and Vibration* **295(1-2)**, 129-140 (2006).

R.P. Hellman, “Growth rate of loudness, annoyance, and noisiness as a function of tone location within the noise spectrum,” *J. Acoust. Soc. Am.* **75(1)**, 209-218 (1984).

“Highway traffic noise analysis and abatement policy and guidance,” U. S. Department of Transportation, Federal Highway Administrations, Washington D.C. (1995).

ISO 532, “Acoustics – Method for calculating loudness level,”*International Standards Organization*, 1975. http://www.iso.org/iso/iso_catalogue.htm

R. Klaeboe, A.H. Amundsen, A. Fyhri, and S. Soldberg, “Road traffic noise- the relationship between noise exposure and noise annoyance in Norway,” *Applied Acoustics* **65**, 893-912 (2004).

H.M.E. Miedema, and H. Vos, “Exposure-response relationships for transportation noise,” *J. Acoust. Soc. Am.* **104(6)**, 3432-3445 (1998).

H.M.E. Miedema and H. Vos, “Demographic and attitudinal factors that modify annoyance from transportation noise,” *J. Acoust. Soc. Am.* **105(6)**, 336-3344 (1999).

B.C.J. Moore and B.R. Glasberg, “A revision of Zwicker’s loudness model,” *Acustica-Acta Acustica* **82(2),** 335-345 (1996).

S. More and P. Davies, “An examination of the influence of spectral balance on annoyance ratings of aircraft sounds,” *Proceedings of Inter-noise 2007,* Istanbul, Turkey, (Aug. 2007).

M. E. Nilsson, “A-weighted sound pressure level as an indicator of short-term loudness or annoyance of road-traffic sound,” *Journal of Sound and Vibration* **302(1-2)**, 197-207 (2007).

T.H. Pedersen, M. Søndergaard, and B. Anderson, “Objective method for assessing the audibility of tones in noise, Joint Nordic Method-Version 2,” Delta Acoustics and Vibrations (1999).

K. Persson and M. Björkman, “Annoyance due to low frequency noise and the use of the dB(A) scale,” *Journal of Sound and Vibration* **127(3),** 491-497 (1988).

R.B. Raggam, M Cik, R.R. Hoeldrick, K. Fallast, E. Gallasch, M. Fend, A. Lackner, and E. Marth, “Personal noise ranking of road traffic: subjective estimation versus physiological parameters under laboratory conditions,” *International Journal of Hygiene and Environmental Health* **210(2)**, 97-105 (2007).

P.D. Schomer, “On normalizing DNL to provide better correlation with response,” *Sound and Vibration* **36(12)**, 14-25 (2002).

T.J. Schultz, “Synthesis of social surveys on noise annoyance,” *J. Acoust. Soc. Am.* **64(2)**, 377-405 (1976).

S.M. Taylor, F.L. Hall and S.E. Birnie, “Transportation noise annoyance: testing of a probabilistic model,” *Journal of Sound and Vibration* **117(1)**, 95-113 (1987).

W. von Aures, “The sensory euphony as a function of auditory sensations,” *Acustica* **59,** 130-141 (1985).

E. Zwicker and H. Fastl, *Psychoacoustics, Facts and Models.* Springer, New York, 2nd Edition, (1999).