Bibliographical reference

Modelling Annoyance from Combined Traffic Noises: An Experimental Study, S. Kuhnt, C. Schürmann, M. Schütte, E. Wenning, B. Griefahn, M. Vormann and J. Hellbrück, Acta Acustica **vol. 94** (Number 3), 2008, pp. 393-400

DOI
https://doi.org/10.3813/AAA.918047

ACTA ACUSTICA UNITED WITH ACUSTICA
Vol. 94 (2008) 393 – 400

DOI 10.3813/AAA.918047

Modelling Annoyance from Combined Traffic Noises: An Experimental Study

S. Kuhnt
Eindhoven University of Technology, The Netherlands

C. Schürmann
Dortmund University of Technology, Germany

M. Schütte, E. Wenning, B. Griefahn
Institute for Occupational Physiology, Dortmund, Germany

M. Vormann, J. Hellbrück
Catholic University Eichstätt-Ingolstadt, Germany

Summary
Annoyance is one of the most studied reactions to noise. Nevertheless, little is known about the effect of the simultaneous occurrence of noise from different sources. Existing models which predict annoyance resulting from combined noise sources are derived from results for single sources and have not yet been validated. The present study empirically investigates actual annoyance as caused from different combinations of road and rail noise in a laboratory experiment. 72 volunteers were exposed to different noise scenarios consisting of combinations of road and rail traffic noise. During noise presentation, test persons had to carry out a task on a personal computer. After each noise scenario, they had to rate their subjective annoyance. A statistical model is derived from the resulting data set, which describes the relationship between noise exposure, task difficulty and annoyance.

PACS no. 43.50.Qp, 43.50.Rq, 43.50.Lj, 43.50.Ba

1. Introduction

Transportation noise is one of the most annoying environmental pollutants. In 1994, about 30% of the EU citizens were exposed to traffic noise levels greater than 55 dBA, 13% were exposed to levels even greater than 65 dBA [1]. In many cases, people are exposed to noise from different sources at the same time. For developing action plans for noise abatement an explanation of the functional relationship between noise and annoyance ratings is useful. A carefully performed meta analysis, which includes roughly 60 000 extended interviews [2] has shown that aircraft-noise annoys most and rail-noise least, while the differences increase with noise loads.

There are several approaches aiming at the prediction of annoyance resulting from two or more noise sources. Commonly known are, for example, the energy summation model [3, 4], the dominant source model [5, 4] and the annoyance equivalence model [4]. Whereas the first one assumes that the total annoyance can be solely predicted by the total noise exposure level, the dominant source model implies that only the loudest among several sound sources has an effect on annoyance. However, Miedema has shown that empirical data do not always support either of these models leading to the more sophisticated annoyance equivalence model [4]. There, the noise level of each single source is transformed into equally annoying sound levels of a pre-chosen reference source.

So far, all three models are derived from empirical studies regarding only single sources. The study presented here provides an empirical data-set for combined rail and road traffic sounds. Its experimental set-up resembles a working situation: Sound levels vary within a realistic range of indoor noise exposure and participants focus on solving tasks. Based on the obtained empirical data set, the functional relationship is analysed between the two types of traffic noise, interaction effects, task difficulty, further covariates and the resulting acute annoyance.

2. Experimental set-up

Traffic noises used in this study consist of combined railway and road noises with graded sound intensities and a duration of 5 minutes, cf. Table I. The noises of the individual sources, “rail” and “road” have been separately recorded using a stereo microphone and have been converted digitally into Dolby-Surround 5.1 signals. All signals are established and provided as 16 bit, 44.1 kHz interleaved wave files by SASS acoustic research & design

Received 31 October 2006, revised 15 february 2008,
accepted 20 February 2008.

Table I. Used sound configurations for the evaluation process. (The equivalent continuous sound levels are given in dBA SPL re 20 μPa.)

	Background 34 34.0	Railway 46 45.4	Railway 64 63.4
Backgr. 34 34.0	37.0	45.4	63.7
Road 46 46.6	47.0	50.0	64.0
Road 64 64.6	64.7	64.8	68.0

GmbH. The single sources, rail and road, were available with a level of 34, 46 and 64 dBA SPL (re 20 μPa). The 64 and 46 dBA versions are identical for both types of noise, except that the 64 dBA version is decreased digitally to get the 46 dBA sound. The 34 dB versions are represented by a noise with spectral properties like background noise in urban vicinity, without occurrence of specific perceptible acoustical events. This noise has been chosen instead of further decreasing the level of the recordings. So, there is no partial masking of the quiet parts of the road and rail signals by the background noise in the laboratories.

We have chosen 34, 46 and 64 dBA SPL (re 20 μPa) as graduations of the sound intensities so that an optimal evaluation of the dose response model has become possible. Besides, in all these noise combinations, the contributing sources are still clearly discernable. Table I presents the resulting combinations in the following way: level of railway / level of road. For example, the sound “64/46” combines 64 dBA rail with 46 dBA road noise. The levels in this table refer to the total sound levels as measured at the typical position of the subjects’ head during the listening tests.

Figure 1 shows the level over time for both sound sources. The railway noise consisted of 12 pass-bys of trains with freight trains, passenger trains, single engines, and high speed trains. Although unrealistic within a general environment, such situations can indeed be found in close proximity to main railway stations. The main reason for choosing this scenario was, however, to construct an indoor sound level without too much structure while adhering to the given noise level. The road noise scenario consists of noises from 74 passenger cars passing by at constant speed, no further sources are perceptible.

In Figure 2 the power spectral density of each single signal is presented. The spectra of the road and rail-noise are quite similar, despite of differences in the frequency region between 50 and 100 Hz and above 4 kHz.

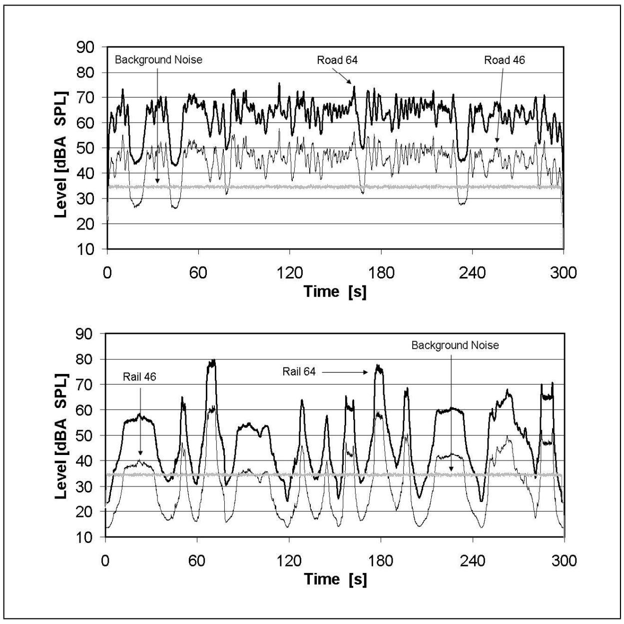

Figure 1. Level versus time for road- and rail-noise for the entire sound-sources at the head-position of the subjects during the tests.

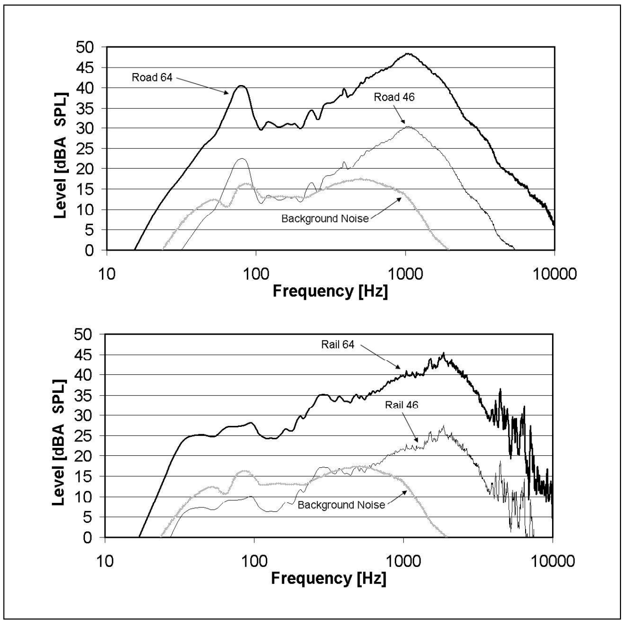

Figure 2. Power spectral densities at the head-position of the subjects during the tests.

Figure 3 displays the distribution of sound levels within each scenario. The levels of rail noise varied between 23-81 dBA and of road noise between 43-75 dBA. The deviation between the shapes of rail 64 and rail 46 is caused by the fact that the recording level was set for levels greater than ~20 dBA: Lower levels become indistinguishable and gather in the area close to 20 dBA. However, this is no problem for the hearing experiments, because the background noise in the laboratories was around 20 dBA.

The experiment itself was conducted in three laboratories (situated in the German cities Dortmund, Essen and Eichstätt). The acoustical properties of all three laboratories were adjusted to the recommendations for the operation of 5.1-dolby surround systems (ITU-R BS.1116-1). In every laboratory, the identical experimental plan was carried out. The procedure was controlled by a computer program: Sounds, questions and rating scales were auto-

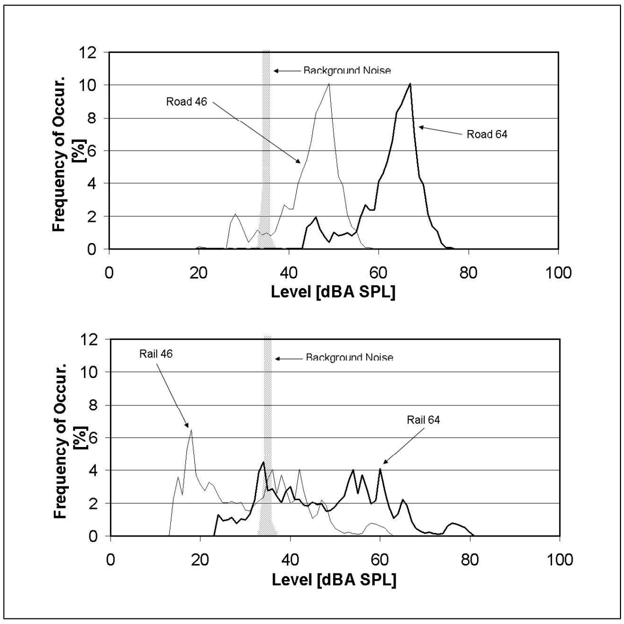

Figure 3. Level-distribution at the head-position of the subjects during the tests. For comparison the area of the background noise is shaded: almost 100% of the levels have values between 33 and 35 dBA.

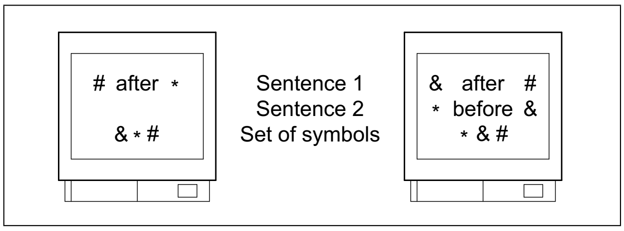

Figure 4. Easy (left side) and difficult (right side) version of the Grammatical-Reasoning-Test.

matically presented and all responses were entered by a keyboard.

During the sound presentation in the experiment, subjects had to carry out the Grammatical Reasoning Test (GRT) to simulate a work condition. The test is part of the ADARD-STRES battery [6]. The GRT requires logical thinking and additionally places demands on working memory. The stimuli of the GRT are sentences of varying syntactic structure and a set of symbols which is presented simultaneously. The test person has to analyze the particular sentence and must decide whether the sentence is concordant with the sequence of the symbols in the symbol set. The test was presented in an easy and difficult version [7]. In the easy version there is one sentence accompanied by a set of symbols (see Figure 4 left part). In the difficult version there are two sentences and one set of symbols (see Figure 4 right part). The test persons were instructed to correctly solve as many test problems as possible in the 5 minute experimental sections.

All 9 sound scenarios were presented on both task levels, resulting in 18 conditions in total. The sound scenarios were presented in random order with the only restriction of maximal two consecutive conditions with equal task level.

Table II. Covariates. N: number of factor levels.

Variable	N	levels
noise level	3	34, 46, 64 dBA
noise source	2	rail, road
gender	2	male, female
time of day	2	morning, afternoon
task difficulty	2	easy, difficult
laboratory	3	Eichstätt, Dortmund, Essen

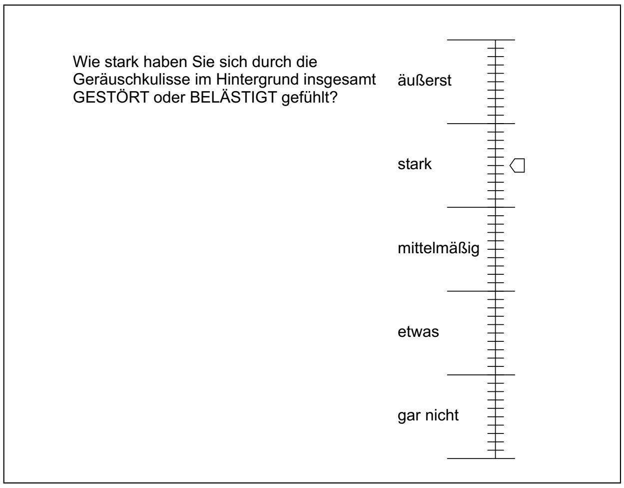

Figure 5. Annoyance question and rating scale (“How much did you feel annoyed or disturbed in total by the noise?” - “not at all”, “slightly”, “moderately”, “very”, “extremely”).

The experiment lasted 7.5 hours, including introduction, practice sections, rating sections and pauses. For each volunteer, the experiment was divided into two sessions, one taking place in the morning (starting between 8.30 h and 9.00 h), the other one in the afternoon (starting between 12.30 h and 13.30 h). The two sessions took place within 8 days. Table II summarises the independent variables of the experiment.

After the presentation of each scenario, a rating scale for assessing the level of annoyance and a rating scale for assessing the level of loudness were presented in random order. Thereafter, also task difficulty and the perception of strain had to be rated. The ratings were followed by a pause lasting between 2 and 15 minutes depending on individual preferences.

Here, we focus on annoyance ratings. Annoyance was recorded using an adapted version of the recommendation by the International Commission on the Biological Effects of Noise (ICBEN) with the reading “How much did you feel annoyed or disturbed in total by the noise?” [8, 9]. The ratings were measured by means of a vertically oriented rating scale consisting of five categories (as shown in Figure 5). The labels followed the original suggestion of ICBEN: “not at all” - “slightly” - “moderately” - “very” - “extremely”. Similar to the proceedings of hearing acoustics, each category was subdivided into 10 graduations resulting in a 50-point rating scale. Like in the category loudness scale (see [10]), the participants were instructed to choose first the appropriate category and then to make a finer choice. The chosen grade was shown as a number. The numbers from 1 to 10 refer to "not at all", the digits from 11 to 20 to "slightly" and so on, where 50 is the maximum.

The experiments have been approved by the Local Ethics Committee and the participants have given their written consent. Subjects were recruited following the same rules for each laboratory: 12 male and 12 female volunteering students between 20 and 30 years (overall mean 24.1 ± 2.9 years). Using a screening questionnaire, hearing capability was controlled. Diagnosis of a hearing disorder, differences in hearing of both ears, Tinnitus, long-term hearing phenomena and wearing a hearing aid were criteria of exclusion from the experiment. Test persons, who had had an inflammation of the middle ear in their infancy for several times underwent an audiometry to determine their hearing capability. Every participant obtained a financial compensation.

At the beginning of the experiment, individual noise sensitivity was measured by the Noise Sensitivity Questionnaire (NoiSeQ; [11, 12]) Every session started with abbreviated scales measuring the situation-specific mood (basing on MDBF, see [13]).

The experimental procedure was explained in the first session and the use of rating scales and tasks (see below) was practiced. During the introduction and practice section, no sound was presented. Afterwards test persons listened to anchor sounds representing the loudest and the quietest sound used in the whole experiment. These arrangements were taken to minimize the effect of interindividual differences in the use of scales.

3. Modelling annoyance

For the development of a dose-response model that explains the relationship between acute annoyance ratings and the covariates a regression model is adequate. Covariates are for example used to characterise noise level, task difficulty, gender of the proband et cetera. Because annoyance ratings as measurements of the response variable are measured on a discrete and bounded scale, the assumptions made in classical linear regression are violated.

We can find a more suitable model in the wider class of generalized linear models (see [14]).

A generalized linear model is based on the following assumptions: We denote the vector of covariates by x. The variable of interest Y^* follows for a given setting x of the covariates, denoted by $Y^* \mid x$, a distribution from an exponential family, which is further characterised by

- the expectation value $E(Y^* \mid x)$
- a linear predictor, summing up individual covariates multiplied with the corresponding elements of an unknown parameter vector β
- a continuous link function g that links the predictor with $E(Y^* \mid x)$.

Table III. Covariates and coded levels. DO: Dortmund, EI: Eichsätt, E: Essen(0).

Variable		values(coded value)
total noise level	x_{level}	37, …, 68 dB(see text)
dominant source	x_{dom}	road(-1), none(0), rail(1)
equal noise levels	x_{enl}	levels are not(0)/ are equal(1)
subject's gender	x_{sex}	female(-1), male(1)
time of day	x_{day}	morning(-1), afternoon(1)
task difficulty	x_{diff}	low(-1), high(1)
laboratory1	x_{lab1}	DO(-1), EI(1), E(0)
laboratory2	x_{lab2}	DO(-1), EI(0), E(1)

In our experiment, the response variable acute annoyance denoted by Y can take values between 1 and 50. We shift the values by one and define $Y^* = Y - 1$. The purpose of this is to get a variable, hence Y^*, which ranges from 0 up to some value n and therefore has the sample space of a binomial distribution. We now can assume that Y^* is binomially distributed $(IB(n, \pi))$ with parameters $n = 49$ and unknown π. The binomial distribution is part of the exponential family and can be used to build up the generalized linear model. The expectation value of Y^* is $E(Y^*) = 49\pi$, where π is unknown but depends on the vector of covariates x, thus $\pi(x)$. The variance, as determined by the binomial distribution, is $Var(Y^*) = n\pi(x)(1 - \pi(x))$.

The variance within the data set is often higher than suggested by the variance given by the binomial model. In such cases, a further parameter is introduced, the so-called overdispersion parameter ϕ, and it is assumed that $Var(Y^*) = \phi 49\pi(x)(1-\pi(x))$. So, the overdispersion parameter measures how much greater the actual variance of the data is compared to the model-based one.

From the experimental set-up we have a number of covariates, which we are coded for the statistical analysis as shown in Table III. Each presented noise scenario combines rail and road noise. We characterise the noise exposure by three covariates. Since the total noise level is not the arithmetic sum of the individual levels, it is not recommendable to add individual levels. Instead, we take the total noise level as shown in Table I (used sound-configuration). So, for instance, the (intended) 64/64 scenario has (used) $x_{noise\,level} = 68$, whereas the 64/34 and the 64/46 scenario both have $x_{noise\,level} = 64$ if decimal places are rounded. As it is no longer clear which source is louder, we use the dummy variables x_{dom} to indicate the dominating one. We set $x_{dom} = +1$ if rail noise is louder than road noise and $x_{dom} = -1$ if road noise is louder, $x_{dom} = 0$ if there is no difference. Since there might be another effect if equally loud noises are combined, we introduce another dummy variable x_{enl} that is set to $x_{enl} = 1$ if a scenario has equal noise levels and $x_{enl} = 0$ if the noise levels are different. So x_{dom} accounts for different kinds of noise whereas x_{enl} allows for a possible interacting effect if both noises are equally loud. Note that these effects are independent from each other. All the other dichotomous variables are coded with levels -1 and +1.

The predictor is constructed from the independent variables; these are the so-called main effects and the interactions, which are given by the product of the individual contributing factors' levels. Initially, one could think of a predictor consisting of all covariates and their two-way interactions. In this case

$$\begin{aligned} g\big(\pi(x)\big) \;=\; & \beta_0 + \beta_{noiselevel}\cdot x_{noiselevel} + \beta_{dom}\cdot x_{dom} \\ & + \beta_{enl}\cdot x_{enl} + \beta_{sex}\cdot x_{sex} + \beta_{day}\cdot x_{day} \\ & + \beta_{diff}\cdot x_{diff} + \beta_{lab1}\cdot x_{lab1} + \beta_{lab2}\cdot x_{lab2} \\ & + \beta_{noiselevel,dom}\cdot x_{noiselevel}\cdot x_{dom} \\ & + \beta_{noiselevel,enl}\cdot x_{noiselevel}\cdot x_{enl} \\ & + \ldots + \beta_{diff,lab2}\cdot x_{diff}\cdot x_{lab2}. \end{aligned} \tag{1}$$

Using a link function is necessary to transform the predictor, ranging within (-∞, ∞), onto the parameter space of the parameter π which is [0, 1]. Concerning the choice of a link function g the results from a pre-stage experiment with more noise levels recommend the use of the complementary loglog link function (cloglog) that is $g(x) = log(-log(1-x))$ [15].

The usual maximum likelihood approach for the estimation of the unknown parameters β_j is based on the assumption of independent observations. This assumption may be violated here as each subject gives more than one rating. An estimation method that incorporates the dependence structure of the observation is given by the so-called generalized estimating equations (GEE). These allow for a more efficient parameter estimation even if the correct correlation structure is unknown. Here, we assume that a subject's ratings' correlation follows an AR(1) process, i.e. the rating at position k depends on the one from position k-1. A detailed description of the AR(1) property can be found in [16]. Unless stated otherwise, all of the results in this article are based on GEE-approaches with AR(1) assumption.

The predictor equation (1) will likely contain factors which have negligible effects on annoyance ratings. So, a method is required which yields a parsimonious but still adequate predictor that includes all relevant factors. Such a method is performed by a stepwise model selection procedure which aims at finding the most significant for the model [17]. We employ the forward directed selection that in each step picks up the most significant factor if significant at the 1% level. If its p-value drops below this level in a later step, it is then removed again. It is common practise to remove main effects only if interactions involving that term have already been removed. The procedure stops if no more significant factors can be found.

Performing the stepwise model selection procedure gives the results as shown in Table IV. A closer look at the laboratory effects and their interactions with the noise level reveals that they only affect the ratings in the quieter scenarios where there is neither noise louder than 46 dBA. The difference is rather small, and its maximum of 2.4 in the estimated expected annoyance rating between laboratories occurs in the 34/34 scenario. Based on the fact that the inclusion of laboratory effects will not be helpful in the use of the model as prediction model, we further reduce the model by excluding them. This gives the final regression results as shown in Table V.

Table IV. Results from GEE-analysis (stepwise regression). NL×diff: Total noise level × task difficulty, lab1×NL: laboratory1 × total noise level, lab2×NL: laboratory2 × total noise level.

factor	parameter estimate	standard error	p-value
intercept (β_0)	−2.3251	0.0956	< 0.0001
total noise level	0.0318	0.0013	< 0.0001
task difficulty	−0.104	0.0490	0.0337
equal noise levels	0.0628	0.0165	0.0001
NL×diff	0.0027	0.0008	0.0011
laboratory1	0.3252	0.1302	0.0125
laboratory2	0.4998	0.1375	0.0003
lab1×NL	0.0040	0.0018	0.0232
lab2×NL	0.0072	0.0020	0.0004
Overdispersion	8.582		

Table V. Results from GEE-analysis (final model equation). NL×diff: Total noise level × task difficulty.

factor	parameter estimate	standard error	p-value
intercept (β_0)	−2.3121	0.0999	< 0.0001
total noise level	0.0316	0.0014	< 0.0001
task difficulty	−0.0980	0.0489	0.0451
equal noise levels	0.0628	0.0165	< 0.0001
NL×diff	0.0026	0.0008	0.0016
Overdispersion	8.697		

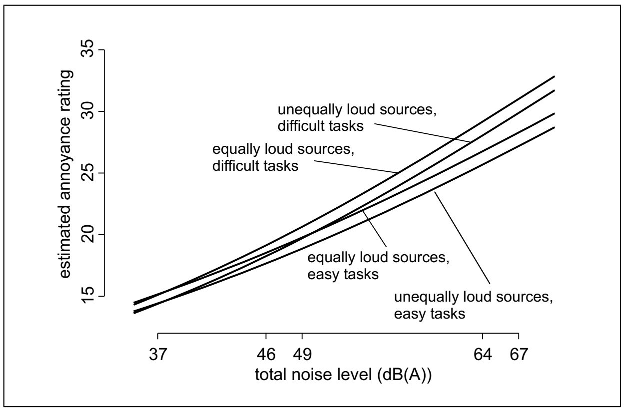

Figure 6. Estimated dose-response relationship for mean annoyance rating at different noise level.

So, the final predictor equation is

$$\begin{aligned} g\big(\pi(x)\big) \;=\; & \beta_0 + \beta_{level}x_{level} + \beta_{diff}x_{diff} + \beta_{enl}x_{enl} \\ & + \beta_{level,diff}x_{level}x_{diff}. \end{aligned} \tag{2}$$

and using the estimated parameter vector $\hat{\beta}$ we get

$$\begin{aligned} g\big(\hat{\pi}(x)\big) \;=\; & -2.3121 + 0.0316x_{level} - 0.098x_{diff} \\ & + 0.0628x_{enl} + 0.0026x_{level}x_{diff}. \end{aligned} \tag{3}$$

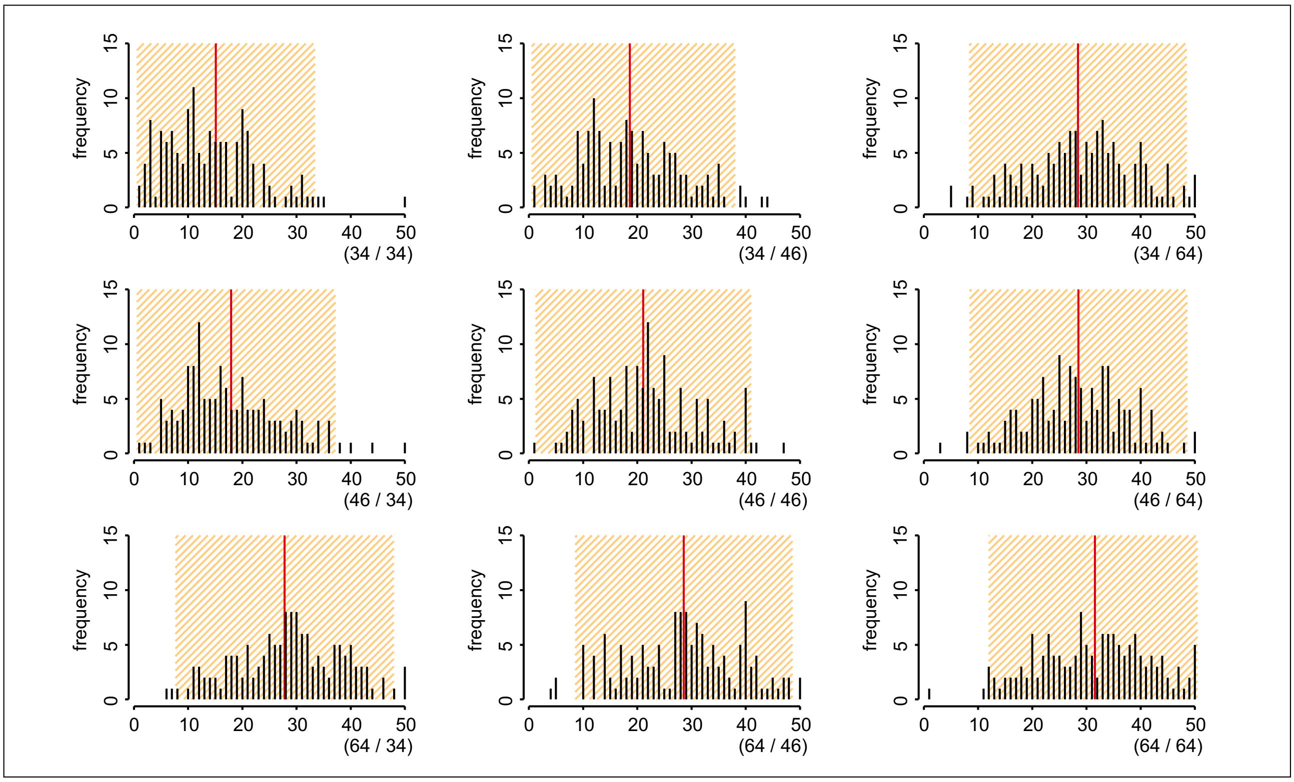

Figure 7. Density plot of the observed annoyance ratings for the nine combinations of rail - road noise scenarios while doing difficult tasks. Vertical lines show the model-estimated mean of the ratings, the shaded area marks the region in which 95% of the ratings are to be expected.

As we can clearly see, average annoyance rating increases if total noise level is raised. It is, unsurprisingly, the most important predictor, too, which augments the predictor by 0.98 if we move the sound level from 37 to 68 dB(A). Notably, this effect interacts with task difficulty: Considering the easier tasks the predictor increase is 0.90, whereas for the more difficult tasks it is 1.06. These findings strongly support the assumption that the subjective evaluation of noise is affected by simultaneous actions, to a degree that depends on the sound level itself.

Concerning the effects of combined sources we find that differences in annoyance ratings are not to be found if either road noise or rail noise represents the dominant source. Remarkably, there is small but discernible effect if both sound levels are equal. In this case, the expected annoyance rating will be greater compared to a situation with the same total noise level but unequally contributing sources. This indicates that there is a combination effect.

As an example of how to apply the model, assume that a subject has to do difficult tasks in the loudest sound setting (64 dBA of both rail and road noise). Then $x_{level} = 68$, $x_{diff} = 1$ and $x_{enl} = 1$. Consequently, we have

$$g\big(\hat{\pi}(x)\big) = -2.3121 + 0.0316 \cdot 68 - 0.0980 + 0.0628 + 0.0026 \cdot 68 = -0.0217$$

yielding

$$\hat{\pi}(x) = 1 - \exp\big(-\exp(-0.0217)\big) = 0.6241$$

and

$$\hat{E}(Y^* \mid x) = \hat{\pi} \cdot 49 = 30.58.$$

The estimated rating would then be

$$\hat{Y} = \hat{Y}^* + 1 = 31.58.$$

The standard deviation as estimated by the model, including the overdispersion parameter ϕ, is

$$\begin{aligned} \hat{sd}(\hat{Y} \mid x) &= \sqrt{\hat{\phi} \cdot \mathrm{Var}(\hat{Y} \mid x)} \\ &= \sqrt{8.697 \cdot 0.6241 \cdot 0.3759 \cdot 49} \\ &= 9.999. \end{aligned} \tag{4}$$

A visual interpretation of estimated annoyance ratings is also helpful. Figure 6 shows estimated dose-response curves for varying noise levels relative to the dichotomous variables of task difficulty and equal noise levels. The non-linear characteristic of the link function is discernable. It can also be seen that the differences in ratings between easy and difficult settings increase with raising sound levels.

However, estimated annoyance rating for an individual situation is only of limited use. As shown in Figure 7 by the histograms, the observed ratings exhibit a large variance causing wide intervals in which to expect a singular rating with 95% probability. For example, in the 64/64 scenario with an estimated rating of 34.3, we would expect 95% of all ratings to lie between 16 and 50.

We therefore look at the model from a different point of view. Instead of individual ratings we consider the proportion of certain annoyance ratings. Due to our use of a scale

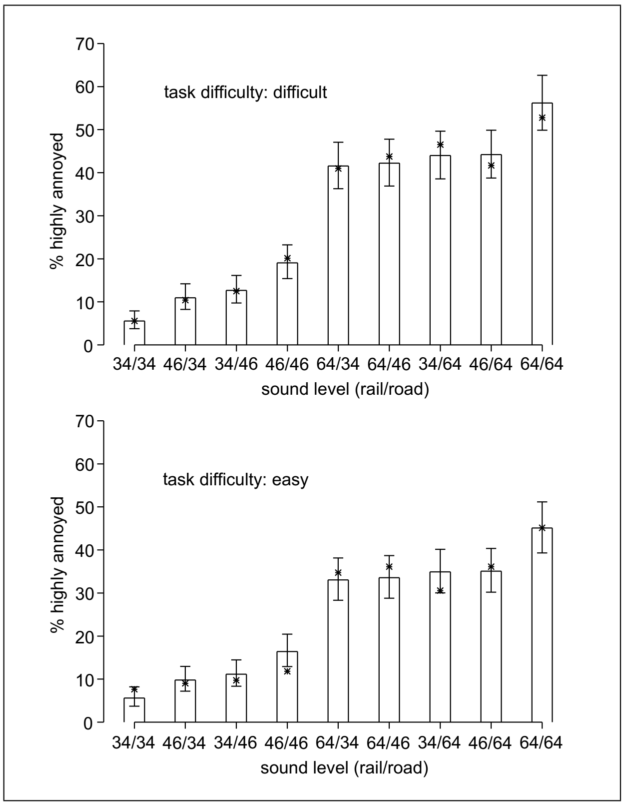

Figure 8. Observed percentages of highly annoyed subjects (asterisks) and estimated proportion of highly annoyed (box) together with 95% confidence intervals (vertical lines).

as in ISOTC 43/SC 1 2002, we classify a rating of 31 or more as one from a subject feeling "highly annoyed". This is consistent with [8], where also the 40% highest possible ratings are rated as highly annoyed, but different from Mediema's use of the top 28%. In terms of our model, the probability that an individual rating is "highly annoyed" is $P(Y > 30 \mid x)$ for a given scenario x.

Figure 8 presents the observed proportions of "highly annoyed" ratings for the two levels of difficulty and varying total noise levels. These lie almost always within the 95%-confidence interval for the percentage of "highly annoyed". For instance, if we consider a 46/34 scenario (which means that either source of noise is at 46 dBA and the other at 34 dBA) the proportion of "highly annoyed" subjects solving difficult tasks is expected to be within 9.2% and 15.3% (95% confidence interval), whereas the observed percentage is in fact 11.5%.

We conclude that although the model allows estimating individual ratings in a given situation, the prediction of a singular rating is inefficient due to the high distributional variance. But if we turn to the proportion of "highly annoyed" ratings the model gives useful results.

4. Discussion

The present study empirically investigated sound combinations from rail and road traffic. The selected dose-response model identifies the noise level as the most powerful predictor for annoyance reactions. Other included variables - as task difficulty and equality of noise levels - are of minor importance. Note that the effect of the equality of noise levels variable is essentially an interaction effect. The factors noise source, dominant source, time of day, and gender of subjects were also considered, but showed no significant influence on the annoyance rating. Using the results from the final model equation estimates of annoyance can be made not only for the noise settings used in the experiment but also for other scenarios with different noise level combinations.

As indicated by the large overdispersion parameter, there is a considerable variance in the annoyance ratings between subjects. Individual differences in weighting and integrating certain aspects of a given noise scenario into the final annoyance judgement could be a cause. The interindividual variation might be reduced by taking individual variables into account as for example noise sensitivity, mood, or critical tendencies [18]. However, knowledge about these will usually not be available in a prediction situation, which renders them useless for an applicable model.

The conducted study aimed at developing a quantitative model that allows for specific traffic situations the prediction of annoyance. The model turned out to be even more suiTable for the prediction of the proportion of "highly annoyed" subjects. In practise, the proportion of "highly annoyed" persons will be of higher importance, as they represent the target group for noise abatement (see [4]).

The prediction model presented here deviates from Miedema's linear approach, as we take the dose-response-relationship to be principally S-shaped, an assumption which is also backed up by earlier experiments [15]. Concerning the way of implementing noise and its sound levels into the prediction equation, the results from our model do not contradict neither the energy summation, nor the annoyance equivalence model. The dominant source model, however, is not suiTable for our data, as it would predict an increase in annoyance only if a louder source was added to a given source. A factor of equal noise level would not be needed within such a model. Besides, transforming the sound levels of noise sources into corresponding levels of a reference noise source as proposed in the equivalence model is also not necessary in our situation. In this case, the annoyance equivalence model collapses to the energy summation model. The energy summation model is closest related to our model, as the total noise level is the only relevant predictor of annoyance.

The developed model is applicable only for noise scenarios resulting from trains and cars passing at constant speed and predicts acute annoyance. Therefore, it seems to be worthwhile to conduct further experiments in order to clarify, whether the presented model can be applied to noises from vehicles in other motion states (accelerating, breaking, ...) and to a greater variety of vehicles (motor bikes, lorries, ...). Apart from the sound pressure level it seems promising to take different patterns of pass-by noise into account. In the conducted experiment subjects concentrated on a grammatical task with two difficulty levels. The difficulty of the activity was found to be significant, but with a limited effect, such that further examinations seem advisable. Tasks with a wider range of difficulty and with demands on different mental processing systems should be considered. As differences between laboratory and field studies [19, 20] are to be expected, the presented prediction model must be validated in or adapted to field situations. The structure of the developed model and the available data basis can easily be extended to further factors and factor levels.

Acknowledgements

The study was performed within the Research Network “Quiet Traffic” which was supported by the German Ministry of Education and Research. Rudolf Bisping (SASS acoustics) produced the sounds employed in the experiment and conducted experiments in the Essen laboratory.

References

[1] OECD: Synthesis report of the OECD project on environmentally sustainable transport EST. International est! Conference.

[2] H. M. E. Miedema, H. Vos: Exposure-response relationships for transportation noise. Journal of the Acoustical Society of America **104** (1998) 3432–3445.

[3] S. M. Taylor: A comparison of models to predict annoyance reactions to noise from mixed sources. Journal of Sound and Vibration **81** (1982) 123–138.

[4] H. M. E. Miedema: The relationship between exposure to multiple noise sources and noise annoyance. Journal of the Acoustical Society of America **116** (2004) 949–957.

[5] C. Rice, K. Izumi: Factors affecting the annoyance of combinations of noise sources. Proceedings of the Institute of Acoustics **8** (1986) 325–332.

[6] A. G. for Aerospace Research, D. (AGARD): Human performance assessment methods. AGAR-Dograph No. 308, Neuilly sur Seine, North Atlantic Treaty Organisation, 1989.

[7] C. A. Shingledecker: A task battery for applied human performance assessment research. Report AFAMRL-TR-84-071. Wright-Patterson Air Force Base, Dayton (OH), 1984.

[8] J. M. Fields, R. G. de Jong, T. Gjestland, I. H. Flindell, R. F. S. Job, S. Kurra, P. Lercher, M. Vallet, T. Yano: Standardized general-purpose noise reaction questions for community noise surveys: Research and a recommendation. Journal of Sound and Vibration **242** (2001) 641–679.

[9] ISO TC 43/SC 1 2002: Acoustics - assessment of noise annoyance by means of social and socio-acoustic surveys. Geneva, 2002.

[10] J. Hellbrück: Category subdivision scaling - A powerful tool in audiometry and noise assessment. – In: Recent trends in hearing research. Festschrift for Seiichiro Namba. H. Fastl, S. Kuwano, A. Schick (eds.). Bibliotheks- und Informationssystem der Universität Oldenburg, Oldenburg, 1996, 317–336.

[11] M. Schütte, A. Marks, E. Wenning, B. Griefahn: The development of the noise sensitivity questionnaire. Noise & Health **9** (2007) 15–24.

[12] M. Schütte, S. Sandrock, B. Griefahn: Factorial validity of the noise sensitivity questionnaire. Noise & Health **9** (2007) 96–100.

[13] R. Steyer, P. Schwenkmezger, P. Notz, M. Eid: Der mehrdimensionale befindlichkeitsfragebogen (MDBF). Hogrefe, Göttingen, 1997.

[14] P. McCullagh, J. A. Nelder: Generalized linear models. Chapman and Hall, London, 1989.

[15] S. Kuhnt, C. Schürmann, B. Griefahn: Annoyance from multiple transportation noise: Statistical models and outlier detection. Methods of Information in Medicine **5** (2004) 510–515.

[16] J. W. Hardin, J. M. Hilbe: Generalized estimating equations. Chapman and Hall/CRC, Boca Raton, 2003.

[17] A. Agresti: Categorical data analysis (2nd edition). Wiley-Interscience, New York, 2002.

[18] N. D. Weinstein: Individual differences in critical tendencies and noise annoyance. Journal of Sound and Vibration **68** (1980) 241–248.

[19] T. Ronnebaum, B. Schulte-Fortkamp, R. Weber: Evaluation of combined noise sources. – In: Contributions to psychological acoustics. A. Schick, M. Klatte (eds.). 1997.

[20] G. Hauck: Lästigkeitsunterschied zwischen den Geräuschen des Straßenverkehrs und des Schienenverkehrs. Zeitschrift für Lärmbekämpfung **38** (1991) 162–166.