Leq + X: Re-Assessment of exposure-response relationships for aircraft noise annoyance and disturbances to improve explained variance

Julia HAUBRICH1; Sarah BENZ1; Mark BRINK2; Rainer GUSKI3; Ullrich ISERMANN; Beat SCHÄFFER2; Rainer SCHMID4; Dirk SCHRECKENBERG1; Jean-Marc WUNDERLI5

1 ZEUS GmbH, Germany
2 Swiss Federal Office for the Environment, Switzerland
3 Ruhr University Bochum, Germany
4 German Aerospace Center, Germany
5 Empa, Swiss Federal Laboratories for Materials Science and Technology, Switzerland

ABSTRACT
There is evidence from the literature that an increased fraction of people is highly annoyed by aircraft noise by given average sound levels than in previous decades. Among possible reasons are changes of the air traffic over the last decades such as the increased number of movements along with a less noisy fleet mix that might be relevant for noise responses but are not adequately reflected by average sound metrics. In the Leq+X project, data of two German and two Swiss field studies comprising aircraft noise exposure and annoyance data of about 37'700 residents living around altogether seven airports are re-analyzed. Among others, the analyses include mixed-effects multilevel models of the percentage highly annoyed on the \(L_{den}\), on the \(L_{Aeq}\) (day/night), on the \(L_{Aeq}\) combined with either maximum sound level \(L_{p,AS,max}\) or number of events above threshold \(N_{ATT}\), and on the \(L_{p,AS,max}\) combined with \(N_{ATT}\) for different thresholds. Further models include the \(L_{Aeq}\) with non-energy-equivalent parameters \(k = 20\) or \(30\), which weight the number of events more strongly than the energy-equivalent \(L_{Aeq,24h,k=10}\). In this paper, the methodological approach and first results will be presented.

Keywords: aircraft noise annoyance, noise metrics, number of events

1. INTRODUCTION
Aircraft noise annoyance is one of the most common negative aspects of (living near) an airport. On the part of residents, noise annoyance is related to several health and wellbeing issues like sleep disturbances and cardiovascular diseases (for recent overviews, see, e.g., 1, 2). For airport operators, it can - at least indirectly, for example through greater and more effective opposition and counteractions to airport expansion - also have negative economic effects.

Noise annoyance is a complex psychological response to noise. According to Guski and colleagues (3), noise annoyance comprises three elements: (i) the experience of an often repeated noise-related disturbance and the behavioural response to it, (ii) an emotional/attitudinal response to the sound and its disturbing impact, (iii) the perception of control of the noise situation.

Over the past few years, there has been a growing number of publications showing that the proportion of people highly annoyed by aircraft noise (%Highly Annoyed, %HA) is higher than would be expected according to common exposure-response curves (such as the much-quoted one from Miedema and Oudshoorn, 4). At comparable aircraft sound levels, %HA is usually higher. Evidence
from a total of 62 studies was summarized in a systematic WHO review and meta-analysis of effects of environmental noise on annoyance by Guski, Schreckenberg, and Schuemer, (3). Some of their main results are depicted in Figure 1.

Figure 1: Relationship between $L_{den}$ and %Highly Annoyed for Aircraft Noise according to different studies and for different airports in different years; figure taken from the WHO Evidence Review for Environmental Noise Annoyance from Guski et al. (3). Black line: weighted (according to sample size) average of 12 samples in which the $L_{den}$ was calculated directly; red line: exposure-response curves according to Miedema & Oudshoorn (4); green line: estimates of Janssen & Vos (5) based on 7 recent aircraft noise examinations.

Figure 1 illustrates that %Highly Annoyed in the data based on the WHO full dataset (black line) and in the estimates of a more recent study by of Janssen & Vos (5, green line) increases earlier with increasing $L_{den}$ compared to the much-cited exposure-response curve according to Miedema & Oudshoorn (4, red line). There are also apparent differences between the different samples, each of which relates to a specific airport in a given year.

Among the possible causes for this finding, for example, methodological reasons were discussed. Janssen et al. (6), however, excluded various methodological aspects as explanatory factors for the increasing trend in an analysis of 34 studies from the years 1967 to 2005.

Also, the change context at an airport at the time of an investigation, that is, whether it could be classified as a ‘high rate change’-airport (HRC) or as a ‘low rate change’-airport (LRC) at a given time, is discussed as another possible cause for the trend recognizable in Figure 1. According to Janssen and Guski (7), a HRC-airport is characterized by the fact that within the last three years, planning intentions for a long-term increase in flight movements or a relocation of flight routes have been made public or still been publicly discussed after such notice. If this criterion does not apply at a given point in time, an airport will be considered as a LRC-airport at that time. Several studies (7, 8) show that the change context of an airport bears a significant effect on aircraft noise annoyance, but does not suffice as the sole explanation for the increase of %Highly Annoyed at comparable continuous sound levels.

As another possible reason for this increase, also an increase in flight movements is discussed. This is based on the finding that since the mid-1990s, the characteristics of aircraft noise have changed to the effect that, on average, airplanes have become quieter, while at the same time there are more flight movements. Examples of acoustic predictors based on event frequencies include $NAT_{LS}$ (‘number above threshold’; indicating the frequency of exceeding the threshold $LS$) and $L_{P,AS,max,air}$, indicating the number of overflights with maximum A-weighted sound pressure levels. For the $NAT_{LS}$, for example, in the NORAH study by Schreckenberg et al. (8) correlations with annoyance measures have been observed ranging from $r=\cdot 31$ to $r=\cdot 56$ at different airports.

Based on these considerations, the present study investigates whether the predictive power for %Highly Annoyed can be improved if this criterion is not only predicted by the equivalent continuous sound levels but when its underlying parameters - maximum level and frequency of noise
events are instead or additionally included into the prediction. Due to the differences between samples, i.e., between different airports in different years, which can be seen in Figure 1, it is also examined whether specific predictions for each sample or certain sub-samples can improve the predictive power further. To identify possible relevant sample-related factors, several sample and airport characteristics are considered as potential predictors.

2. PROCEDURE, METHODS

2.1 Data basis, samples and participants

For this study, existing data sets with individual annoyance measures and address-related acoustic exposure variables from these two Swiss and two German studies are used: ‘Laermstudie (LS) 2000’ (9), ‘SiRENE study’ (10), ‘RDF study’ (11), and ‘NORAH study’ (8).

In total, the data of 12 samples were available for the present study, where a sample is meant to be one survey at one airport at one specific point in time. The total number of participants is $N=34250$; 17909 (52.3%) female, aged 18-100 years ($M=57.5$, $SD=14.9$).

2.2 Acoustic predictors

As acoustic predictors, the parameters listed below were used. These were either extracted directly from the data sets of the above-mentioned samples ($L_{\text{Aeq}}$, $L_{\text{den}}$) on the basis of the data available from the studies, or - for the determination of values for $L_{\text{Aeq},24h,k=20}$ and $L_{\text{Aeq},24h,k=30}$ - in some cases approximations were used (for details see Benz et al., 12).

- $L_{\text{Aeq},24h,k=10}$, $L_{\text{Aeq},24h,k=20}$, $L_{\text{Aeq},24h,k=30}$: equivalent continuous sound level $L_{\text{Aeq},k}$ for the 24-hour day for different equivalence parameters $k$
- $L_{\text{den}}$: the day-night-night level as defined in the EU Environmental Noise Directive
- $\log(NAT_{24h,50})$, $\log(NAT_{24h,60})$, $\log(NAT_{24h,70})$, $\log(NAT_{24h,80})$: the logarithmic number of aircraft noise events with A-weighted maximum sound levels above a threshold of LS decibels ("Number above Threshold" $NAT_{LS}$)
- $L_{\text{AS,max,ari,50}}$, $L_{\text{AS,max,ari,60}}$, $L_{\text{AS,max,ari,70}}$, $L_{\text{AS,max,ari,80}}$: the arithmetically averaged A-weighted maximum sound levels ($L_{\text{AS,max,ari},k}$) above these thresholds

2.3 Sample and airport-specific predictors

The following sample and airport-specific predictors were considered: a) the number of flight movements on an average 24-hour day, b) night-flight rate, c) five year trend of flight movements (increasing or decreasing), d) fleet mix (for the starting movements over the 24 hours of the whole day; cf. 13, 14), e) change context (HRC vs. LRC, see above).

2.4 Assessment of high annoyance

Annoyance was assessed via self-report using the standard 5-point ICBEN noise annoyance scale (all samples except those from the LS2000 study) or the 11-point ICBEN noise annoyance scale.

For the determination of %Highly Annoyed, the upper two values of the 5-level ICBEN scale (4 and 5) were summarized into the category "highly annoyed" (HA); for analyses particularly referring to the LS2000 study, the upper three values on the numerical ICBEN scale from 0 to 10 (8, 9 and 10) were used instead.

2.5 Method

To investigate whether frequency-based parameters can improve the predictive power of the acoustic predictors, the models presented in Table 1 were examined. Each model includes one predictor (lines 1-4) or a combination of two predictors (line 5 and beyond). Multilevel analyses were carried out for each model using the package "lme4" of the statistical software R (15).
Table 1 - Summary of the predictors of the multi-level models

<table>
<thead>
<tr>
<th>Model group</th>
<th>Level 1-predictors (acoustic predictors)</th>
<th>Level 2-predictors (Sample and airport-specific predictors)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$L_{\text{den}}$</td>
<td>$L_{\text{den}}$</td>
<td>$L_{\text{Aeq},24h,k=10}$</td>
</tr>
<tr>
<td>$L_{\text{Aeq},\text{Aeq}}$</td>
<td>$L_{\text{Aeq},24h,k=20}$</td>
<td>(sample) flight movements</td>
</tr>
<tr>
<td>$L_{\text{den}} + \log(NAT_{LS})$</td>
<td>$L_{\text{den}, \log(NAT_{24h,50})}$</td>
<td>nightflight rate</td>
</tr>
<tr>
<td></td>
<td>$L_{\text{den}, \log(NAT_{24h,60})}$</td>
<td>trend (five year trend of flight movements)</td>
</tr>
<tr>
<td></td>
<td>$L_{\text{den}, \log(NAT_{24h,70})}$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$L_{\text{den}, \log(NAT_{24h,80})}$</td>
<td></td>
</tr>
<tr>
<td>$L_{\text{Aeq}} + \log(NAT_{LS})$</td>
<td>$L_{\text{Aeq},k=10, \log(NAT_{24h,50})}$</td>
<td>NoiseStarts24 (fleet mix)</td>
</tr>
<tr>
<td></td>
<td>$L_{\text{Aeq},k=10, \log(NAT_{24h,60})}$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$L_{\text{Aeq},k=10, \log(NAT_{24h,70})}$</td>
<td>change context (HRC vs. LRC)</td>
</tr>
<tr>
<td></td>
<td>$L_{\text{Aeq},k=10, \log(NAT_{24h,80})}$</td>
<td></td>
</tr>
</tbody>
</table>

In the first step, it was examined based on the intraclass correlation (ICC, 16), whether the second level predictors enhance the predictive power significantly. The sample variable, characterized by the airport and the survey year, served as the classification variable.

Next, each the best model with exactly one predictor and the best model with two predictors were identified. The criterion was the goodness of fit of the models (based on the Akaike Information Criterion, AIC, 17). Finally, these two models were compared statistically with a Chi-Square Differences Test.

Other outcome measures considered are the odds ratios, each separately for the intercept (the $\beta_0$ parameter) and the slopes (the $\beta_1$ parameter and, if a model has two acoustic predictors, additionally the $\beta_2$ parameter).

3. RESULTS

A summary of the results of the multi-level analyses is shown in Table 2.

For all models, the ICC is above .10 (range: .10 - .21), indicating that the 2nd level predictors should be kept. In the models with one acoustic predictor, the model with $L_{\text{den}}$ achieves better goodness of fit than the various $L_{\text{Aeq}}$ models.

For the models with two predictors, the models from the groups $L_{\text{den}}$ and $\log(NAT_{LS})$ as well as $L_{\text{Aeq},24h,k=10}$ and $\log(NAT_{LS})$ are better than those from the group $L_{\text{AS,max,ari,LS}}$ and $\log(NAT_{LS})$. The best model with two predictors has $L_{\text{den}}$ and $\log(NAT_{24h,70})$ as joint predictors.

A Chi-Square Differences Test comparing these two models yields a $p$-value of $p < .001$, which is highly significant, indicating that the more complex model with two predictors yields significantly better goodness of fit than the model with one predictor.

Overall, thus, the model associated with the best predictive power is the model with the combined predictors $L_{\text{den}}$ and $\log(NAT_{24h,70})$. 

In conclusion, the results of the multi-level analyses indicate that the model with the combined predictors $L_{\text{den}}$ and $\log(NAT_{24h,70})$ provides the best fit for the data.
Table 2 - Summary of the results of the multi-level analyses for the prediction of %HA aircraft noise

<table>
<thead>
<tr>
<th>Predictor 1</th>
<th>Predictor 2</th>
<th>Best Model</th>
<th>ICC</th>
<th>AIC</th>
<th>Marginal $R^2$</th>
<th>Conditional $R^2$</th>
<th>OR $\beta_0$ (Intercept)</th>
<th>OR $\beta_1$ (Slope)</th>
<th>OR $\beta_2$ (Slope)</th>
<th>Significant Term(s)</th>
<th>OR sig. Term(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$L_{den}$</td>
<td>/</td>
<td>AIM</td>
<td>.121</td>
<td>29417.1</td>
<td>.151</td>
<td>.246</td>
<td>.47***</td>
<td>2.25***</td>
<td>/</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>$L_{Aeq,24h,k=10}$</td>
<td>/</td>
<td>AIM</td>
<td>.144</td>
<td>29417.7</td>
<td>.146</td>
<td>.255</td>
<td>.45***</td>
<td>2.23***</td>
<td>/</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>$L_{Aeq,24h,k=20}$</td>
<td>/</td>
<td>AIM</td>
<td>.165</td>
<td>29985.5</td>
<td>.152</td>
<td>.299</td>
<td>.44***</td>
<td>2.33***</td>
<td>/</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>$L_{Aeq,24h,k=30}$</td>
<td>/</td>
<td>AIM</td>
<td>.209</td>
<td>30557.2</td>
<td>.154</td>
<td>.356</td>
<td>.44***</td>
<td>2.42***</td>
<td>/</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>$L_{den}$</td>
<td>log($NAT_{24h,50}$)</td>
<td>FM</td>
<td>.121</td>
<td>29406.9</td>
<td>.209</td>
<td>.264</td>
<td>.60**</td>
<td>2.73***</td>
<td>.96***</td>
<td>HRC</td>
<td>.51***</td>
</tr>
<tr>
<td>$L_{den}$</td>
<td>log($NAT_{24h,60}$)</td>
<td>FM</td>
<td>.119</td>
<td>29368.7</td>
<td>.213</td>
<td>.262</td>
<td>.58***</td>
<td>2.16***</td>
<td>1.27***</td>
<td>HRC</td>
<td>.49***</td>
</tr>
<tr>
<td>$L_{den}$</td>
<td>log($NAT_{24h,70}$)</td>
<td>FM</td>
<td>.119</td>
<td>29358.5</td>
<td>.214</td>
<td>.264</td>
<td>.60***</td>
<td>2.15***</td>
<td>1.26***</td>
<td>HRC</td>
<td>.55***</td>
</tr>
<tr>
<td>$L_{den}$</td>
<td>log($NAT_{24h,80}$)</td>
<td>FM</td>
<td>.121</td>
<td>29408.1</td>
<td>.209</td>
<td>.264</td>
<td>.60***</td>
<td>2.62***</td>
<td>ns</td>
<td>HRC</td>
<td>.51***</td>
</tr>
<tr>
<td>$L_{Aeq,24h,k=10}$</td>
<td>log($NAT_{24h,50}$)</td>
<td>AIM</td>
<td>.144</td>
<td>29418.5</td>
<td>.188</td>
<td>.236</td>
<td>.60**</td>
<td>2.27***</td>
<td>ns</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>$L_{Aeq,24h,k=10}$</td>
<td>log($NAT_{24h,60}$)</td>
<td>AIM</td>
<td>.138</td>
<td>29382.9</td>
<td>.216</td>
<td>.264</td>
<td>.57**</td>
<td>2.19***</td>
<td>1.25***</td>
<td>HRC</td>
<td>.52***</td>
</tr>
<tr>
<td>$L_{Aeq,24h,k=10}$</td>
<td>log($NAT_{24h,70}$)</td>
<td>AIM</td>
<td>.136</td>
<td>29387.0</td>
<td>.192</td>
<td>.237</td>
<td>.61**</td>
<td>1.87***</td>
<td>1.23***</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>$L_{Aeq,24h,k=10}$</td>
<td>log($NAT_{24h,80}$)</td>
<td>AIM</td>
<td>.146</td>
<td>29419.5</td>
<td>.190</td>
<td>.238</td>
<td>.60**</td>
<td>2.26***</td>
<td>ns</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>$L_{p,AS,max,ari,24h,50}$</td>
<td>log($NAT_{24h,50}$)</td>
<td>AIM</td>
<td>.133</td>
<td>29434.6</td>
<td>.216</td>
<td>.260</td>
<td>.62**</td>
<td>1.81***</td>
<td>1.70***</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>$L_{p,AS,max,ari,24h,60}$</td>
<td>log($NAT_{24h,60}$)</td>
<td>AIM</td>
<td>.125</td>
<td>29436.6</td>
<td>.230</td>
<td>.269</td>
<td>.60**</td>
<td>1.39***</td>
<td>2.08***</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>$L_{p,AS,max,ari,24h,70}$</td>
<td>log($NAT_{24h,70}$)</td>
<td>AIM</td>
<td>.112</td>
<td>29811.7</td>
<td>.208</td>
<td>.240</td>
<td>.64**</td>
<td>ns</td>
<td>2.34***</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>$L_{p,AS,max,ari,24h,80}$</td>
<td>log($NAT_{24h,80}$)</td>
<td>AIM</td>
<td>.104</td>
<td>30751.2</td>
<td>.158</td>
<td>.189</td>
<td>.67**</td>
<td>.86*</td>
<td>2.08***</td>
<td>--</td>
<td></td>
</tr>
</tbody>
</table>

Note. ICC: intraclass coefficient. AIM: “augmented intermediate model”, FM: “fullmodel”. AIC: Akaike information criterion. * p<.05, ** p<.01, *** p<.001; ns: not significant.
Table 2 shows that exactly one level 2-predictor is significant in this model, which is the ‘change context.’ A look at the odds ratios of the regression parameters shows that the $L_{den}$ exerts the largest effect on $\%HA$ - the odds ratios for $\beta_1$ are higher than those for $\beta_2$ for $\log(NAT_{24h,70})$ (2.15 vs. 1.26, respectively). The odds ratios for the significant level 2-predictor are about 4x smaller than those for $L_{den}$ and less than half as big as those for $\log(NAT_{24h,70})$ ($\beta_3=.55$). Thus, the ‘change context’ implies a significant additional contribution to the explained variance, but its effect is weaker than that of either of the two acoustic predictors.

Figure 2 shows the exposure-response curve for the relationship between $L_{den}$ and $\%HA$ for three classes of $\log(NAT_{24h,70})$, each separately for ‘high-rate change’ (HRC) and ‘low rate change’-airports (LRC), respectively. It can be seen that at the LRC-airports, the curves for the high, middle and lower $\log(NAT_{24h,70})$-groups differ mainly in their height, with the curve representing low $\log(NAT_{24h,70})$-values being above the curve representing middle values at higher $L_{den}$-values.

At the HRC-airports, however, the curve representing high $\log(NAT_{24h,70})$-values stands out, in that it has a much lower slope than the other two.

![High-rate change (HRC) airports](image1)

![Low-rate change (LRC) airports](image2)

Figure 2: Exposure-response curve for the relationship between $L_{den}$ and $\%HA$ for HRC airports (upper panel) and LRC airports (lower panel), separated for three classes of $\log(NAT_{24h,70})$ (see boxes on the right for detailed information). Dashed lines indicate the upper and lower confidence intervals to the respective curve of the same color. Grayed out are the extrapolated parts of the curves outside the actual range of $L_{den}$.
4. DISCUSSION AND CONCLUSIONS

This research project was based on the assumption that the "aircraft noise annoyance trend," showing an increased rate of highly annoyed persons at a comparable continuous sound level, could be associated with an increase in the frequency of flight movements. It has been postulated that the simultaneous consideration of a frequency-based predictor - in particular, log(NATLs) - in combination with an acoustic predictor - especially LAeq, Lden or Lp,AS,max,ari - yields a better predictive power of %Highly Annoyed than the respective acoustic predictor on its own.

For this purpose, the goodness of fit and predictive power of different models with single acoustic predictors and acoustic predictors in combination with frequency-based predictors were compared with each other using multi-level analyses. To take into account sample- and airport-specific effects, additional sample-specific characteristics were investigated as level 2-predictors.

The results show that a combination of Lden and log(NATLs) has better predictive power than models with a single predictor (Lden or LAeq,24h,k with k=10, k=20 or k=30) or a model with a combination of LAeq,k=10 and log(NATLs) or Lp,AS,max,ari and log(NATLs).

In addition, a sample-specific consideration of the effects turns out to be significant. Of the level 2-predictors examined here, the 'change context' proved to be significant ('high rate change'-airport, HRC, vs. 'low rate change'-airport, LRC), with %Highly Annoyed significantly higher for HRC.

ACKNOWLEDGEMENTS

The study was funded by the Swiss Federal Office for the Environment (‘Bundesamt für Umwelt, BAFU’), Bern, Switzerland; and by the Ruhr-University of Bochum, Germany.

REFERENCES


life around Frankfurt Airport. International journal of environmental research and public health, 7(9), 3382-3405.


