Estimation of field psychoacoustic indices and predictive annoyance models for road traffic noise combined with aircraft noise

Laure-Anne Gille, Catherine Marquis-Favre

Estimation of field psychoacoustic indices and predictive annoyance models for road traffic noise combined with aircraft noise

Laure-Anne Gille[a1] and Catherine Marquis-Favre[1]

Univ Lyon, ENTPE, Laboratoire Génie Civil et Bâtiment, 3 rue Maurice Audin, F-69518 Vaulx-en-Velin, France

Annoyance due to urban road traffic noise combined with aircraft noise was studied, using both laboratory and field survey data. Laboratory data were used to propose: i) partial annoyance models considering psychoacoustic indices and noise sensitivity, and ii) total annoyance models considering noise indices or partial annoyance models. To predict partial and total annoyance in field, a methodology was proposed to estimate the different psychoacoustic indices, involved in annoyance models, from L_{den} values. This methodology, interesting from a practical point of view, and the proposed annoyance models enable a good prediction of the field partial and total annoyance ratings compared to models based on L_{den} only. These results show that there is a need to develop the methodology of psychoacoustic index estimation from noise map L_{den} and also partial and total annoyance models.

[a] also at: CEREMA, Direction Territoriale Île-de-France, 21-23 rue Miollis, 75732 Paris Cedex 15, France

I. INTRODUCTION

In urban areas, noise annoyance is a major source of concern. Road traffic is the most annoying noise source in terms of numbers of those concerned, but aircraft is the most annoying in terms of percentages of those exposed to a given noise level (*e.g.*[1,2]). European cities of more than 100,000 inhabitants manage noise exposure using strategic noise maps[3]. These maps using L_{den}, the day-evening-night level, represent noise exposure to each transportation noise source. This index is also used in exposure-effect relationships to estimate the percentages of people annoyed by a noise source[1]. However, this index does not cover some acoustical features known to be particularly annoying, *i.e.* spectral content, irregular amplitude fluctuation or modulation-related sensations (*e.g.*[4–7]). In addition, combined noise exposure is a more frequent situation[8]. However as noise sources may interact[9], their characterization and therefore their prediction become more difficult.

It is also well known that non-acoustical factors, such as noise sensitivity, influence noise annoyance (*e.g.*[10]). Noise annoyance models, considering both psychoacoustic indices characterizing different annoying acoustical features and noise sensitivity, have already been proposed on laboratory data in the literature (*e.g.* for tramway noise[5,6], for aircraft noise[11,12]). However, these models may not be applicable on field data as some of the corresponding psychoacoustic indices cannot be predicted or measured on a wide area.

Below, annoyance due to combined urban road traffic noise and aircraft noise is studied. A combined noise experiment was carried out in a simulated environment. Partial and total annoyance were assessed in laboratory conditions simulating the combined noise exposure

studied during a field survey (presented in Gille *et al.*[2]). Data from the different single and combined laboratory experiments were aggregated to construct partial annoyance multilevel models for each noise source. These partial annoyance models accounted for different acoustical features of each noise source by considering psychoacoustic indices. Total annoyance models for combined urban road traffic noise and aircraft noise were constructed using data from the combined noise experiment. The data from the field survey were used to test the proposed models for partial and total annoyance. To test the models, a methodology was proposed to estimate noise index values from field L_{den}.

II. METHOD OF COLLECTION OF DATA

Data from three experiments were aggregated to study noise annoyance due to urban road traffic noise (denoted as $URTN$ in the following), annoyance due to aircraft noise (denoted as AN) and total noise annoyance due to combined noises (denoted as CN). The single noise experiments and the combined noise experiment are presented below (*cf.* Sections II A and II B, respectively). The models built using laboratory data were tested using data from a French survey, presented in Section II C.

A. Single noise experiments

A single noise experiment (*cf.* Gille *et al.* [13]) was carried out to study specific annoyance due to $URTN$ (*i.e.* annoyance due to $URTN$ heard in isolation). A similar experiment in terms of apparatus and procedure was carried out to study specific annoyance due to AN

(*cf.* Gille *et al.* [12]). They are briefly described below, for more details see Gille *et al.*[13] and Gille *et al.*[12], respectively.

1. Stimuli

a. URTN experiment. Twenty seven $URTN$ sequences of 3 minutes were constructed using different urban road vehicle pass-by noises and an urban background noise which were recorded in Lyon and its suburbs (*cf.* Gille *et al.* [13]). The A-weighted equivalent sound pressure level of the single pass-by noises ranged from 53.5 dB(A) to 67.1 dB(A), depending on the type of vehicle and of the driving condition. The background noise was equalized at 40 dB(A), in order to be masked by the pass-by noise with the lowest L_{Aeq}. Noise sequences comprised between 16 and 80 vehicles in 3 minutes, with a maximum of 44% of powered-two-wheelers or a maximum of 10% of buses and heavy vehicles. Resulting noise level of the $URTN$ stimuli ranged from 57.3 to 63.7 dB(A).

b. AN experiment. Twelve aircraft flyover noises were recorded in the French neighborhood of an international airport, approximately 5 km away in line with the runway (*cf.* Gille *et al.* [12]). The differences in A-weighted equivalent sound pressure level observed in field were kept, resulting in a range from 43.5 dB(A) to 54.6 dB(A), with stimulus duration ranging from 22.1 to 61.5 seconds.

c. For both experiments. No filter simulating facade transmission was applied to the stimuli as wall material and window types have an effect on auditory judgments [14] and the choice of one specific kind of facade might have been too limiting. Thus, the worst noise

exposure is considered at home such as being near the window open in private indoor spaces or in private outdoor spaces.

2. *Apparatus for URTN and AN experiments*

The experiments took place in a quiet room with a background noise of 19 dB(A). The stimuli were reproduced employing a 2.1 audio reproduction system consisting of two active loudspeakers and one active subwoofer. The center of the interaural axis of the participant and the loudspeakers formed an equilateral triangle (as recommended by Bech and Zacharov [15]). The loudspeakers were placed at a height of 1.20 m from the floor, and the subwoofer was placed on the floor between the loudspeakers.

3. *Procedure for URTN and AN experiments*

Participants were asked to imagine themselves at home while relaxing. For the $URTN$ experiment with 3-min stimuli, they could bring along their own reading stuff. After each stimulus, a reminder of the imaginary situation was presented to the participants and they were asked to give annoyance rating due to the stimulus on a continuous scale ranging from "0" to "10".

At the end of both experiments, the participants performed a verbalization task and they had to evaluate their noise sensitivity on a continuous scales ranging from "0" to "10".

4. Participants

The *URTN* experiment was performed by 34 participants (16 male, 18 female) aged between 20 and 55 years old (mean age: 32.3; standard deviation: 12.7). The AN experiment was performed by 33 participants (19 male, 14 female) aged between 20 and 56 years old (mean age: 32; standard deviation: 12.5). Eleven participants performed both single noise experiments. All participants declared normal hearing abilities and were paid for their participation.

B. The combined noise experiment

1. Stimuli

As recommended by Berglund and Nilsson[16], experiment stimuli were constructed on the basis of a complete matrix composed of 4 *URTN* sequences and 4 *AN* sequences. Sixteen 6 minute noise sequences were constructed to study different *URTN* and *AN* scenarios.

a. URTN sequences. To account for the urban road traffic observed in the surveyed cities[2], the mean daily traffic and the noise exposure expressed in terms of L_{den} were considered. These data allowed the traffic sequences 1T5 (16 light vehicles (LV), 1 bus or heavy vehicle (BHV) and 3 powered-two-wheelers (PTW) in 3 mins) and 1T8 (24 LV, 1 BHV, 3 PTW) built in Gille *et al.*[13] to be selected in order to simulate the road traffic in small streets. The traffic sequences 2T8 (48 LV, 2 BHV, 6 PTW) and 2T11 (62 LV, 2 BHV, PTW) were selected as they simulated road traffic in bigger streets (*cf.* Gille *et al.*[13]). In

order to construct 6 min CN sequences, the 3-min $URTN$ sequences taken from Gille *et al.*[13] were played twice.

b. AN sequences. In order to reproduce the aircraft traffic observed in the surveyed cities[2], the AN sequences were composed of 1 to 4 aircraft flyover noises, with an interval between 2 aircraft flyovers ranging from 90 secs to 6 mins. The aircraft flyover noises composing the sequences came from the experiment presented in Gille *et al.*[12]. Four different aircraft noises were selected, according to their mean annoyance rating[12]. Stimuli a2 (22.1 secs) and a6 (40.1 secs) were selected as they are respectively one of the least and one of the most annoying flyover noises. Stimuli a7 (44.3 secs) and a11 (61.5 secs) were selected among stimuli with mean annoyance ratings significantly different from a2 and a6 (*cf.*[12]). Table II gives the succession of aircraft flyover noises and the interval between 2 aircraft flyover events.

c. CN sequences. The 6-min CN sequences were composed of an $URTN$ sequence, an AN sequence and an urban background noise, recorded by Trollé *et al.*[5,6] early in the morning without distinguishable noise events and equalized at 32.4 dB(A). Noise level of aircraft flyovers and of urban road traffic noises was 20 dB(A) lower than field observations. This was done in order to simulate a window open and a distance between the street and the living room. Table I gives the A-weighted equivalent sound pressure level for the urban road traffic noise sequences combined with the urban background noise and for the aircraft flyover noises composing the aircraft noise sequences. Table II gives the resulting A-weighted equivalent sound pressure level, L_{Aeq} of the CN sequences.

TABLE I. The A-weighted equivalent sound pressure level $L_{Aeq,6min}$ of $URTN$ (Urban Road Traffic Noise) sequences and $L_{Aeq,flyover}$ of aircraft flyover noises composing the AN (Aircraft Noise) sequences.

$L_{Aeq,6min}$ of $URTN$ sequence		$L_{Aeq,flyover}$ of aircraft flyover noise	
	dB(A)		dB(A)
1T5	45.7	a2	35.0
1T8	46.6	a6	55.6
2T8	49.5	a7	51.3
2T11	51.9	a11	46.5

TABLE II. A-weighted equivalent sound pressure level L_{Aeq} of the combined noise sequences composed of $URTN$ (Urban Road Traffic Noise) and AN (Aircraft Noise). Interval between 2 aircraft flyovers is given between brackets.

	$URTN$ sequence			
AN sequence	1T5	1T8	2T8	2T11
a2	45.9	46.7	49.5	52.0
a2 a6 (3 mins)	49.1	49.4	51.1	53.0
a2 a7 a6 (2 mins)	49.9	50.4	51.6	53.3
a2 a7 a11 a6 (90 secs)	50.0	50.3	51.8	53.4

2. Apparatus

The experiment was run in a quiet simulated living room, with a background noise level below 22 dB(A). The noise sequences were reproduced employing a 2.1 audio reproduction system consisting of two active loudspeakers and one active subwoofer. Participants sat on a 3-place sofa. The center of the interaural axis of the participant seated in the middle of the sofa and the loudspeakers formed an equilateral triangle (as recommended by Bech and Zacharov[15]). The loudspeakers were placed at a height of 1 m, and the subwoofer was placed on the floor between the loudspeakers.

3. Procedure

Three participants performed the test simultaneously. They were asked to not speak together and to imagine themselves having a read at home. They could bring along their own reading material for the experiment. This procedure has been used in previous works (*e.g.*[17]). The stimuli were presented one by one in random order.

After each combined noise sequence, participants were asked about: i) the urban road traffic partial annoyance (*i.e.* annoyance due to a $URTN$ in combined noise exposure situation), ii) the aircraft partial annoyance and iii) the total annoyance due to combined noises: *"While you imagined yourself relaxing at home with this soundscape sequence, did (the road traffic noise) / (the aircraft noise) / (the global noise due to the road traffic noise and the aircraft noise) annoy you?"*. To answer these 3 questions, participants gave ratings

on continuous scales ranging from "0" to "10", with 11 evenly spaced numerical labels and two verbal labels at both ends (*"not at all annoyed"* and *"extremely annoyed"*).

At the end of the experiment, participants answered three following questions: *"Did (the aircraft noise) / (the road traffic noise) seem familiar to you?"*, *"How would you describe (the aircraft noise) / (the road traffic noise)? "* and *"In a general way, how do you judge (the aircraft noise) / (the road traffic noise)? "*. As in single noise experiments [2,12] (*cf.* Section II A), participants filled in a questionnaire with personal items such as non-acoustical factors at the end of the experiment. For noise sensitivity, participants were asked: *"Would you say you are sensitive to noise generally speaking ?"* They were then asked to make a judgment on a continuous scale ranging from "0" to "10" with two verbal labels at both ends (*"not at all sensitive"* and *"extremely sensitive"*). The experiment lasted two hours.

4. Participants

The experiment was performed by 32 participants (17 male, 15 female) aged between 20 and 67 years (mean age = 37.5; standard deviation = 14.5). Six participants had also performed the 2 single noise experiments and 2 had performed the $URTN$ experiment. All participants declared normal hearing abilities and were paid for their participation.

C. French survey

Data from a French survey focusing on annoyance due to combined transportation noises and conducted in 2012 (*cf.* Gille *et al.*[2]) were used to test the models built from laboratory data. In the following, only cities exposed to road traffic noise combined with aircraft noise

will be considered, that is to say data for 212 respondents (*cf.* Gille *et al.*[18]). Only people aged between 18 and 80 years old and living permanently in the dwelling since at least one year were face-to-face interviewed. The questions about the noise annoyance complied with the recommendations provided by the ISO 15666 standard[19]. Respondents were asked to give an annoyance rating on a continuous scale from "0" to "10", with 11 evenly spaced numerical labels and two verbal labels at both ends ("not at all" and "extremely"). Noise sensitivity was evaluated on the same format. The exposure of each respondent was determined using L_{den} from the strategic noise maps. Table III gives a summary of the survey. For more details, see Gille *et al.*[2].

III. RESULTS

Data from single noise and CN experiments were used to construct annoyance models for $URTN$ and for AN. These models were based on noise sensitivity and noise indices. These models used as variables of total annoyance models from the literature were considered to propose annoyance models for combined noises. Then a methodology was proposed to evaluate the different noise indices from L_{den} given by the strategic noise maps. Data from the French survey were used to test the annoyance models established under laboratory conditions.

TABLE III. Main information from the French survey[2,18], considering only the sample of respondents exposed to road traffic noise combined with aircraft noise.

French survey	
Date	2012
Cities	2
Respondents	212
Road traffic noise	L_{den}: [49.9;77.9] dB(A)
Aircraft noise	L_{den}: [42;54] dB(A)
	using strategic noise maps
Annoyance question	in agreement with ISO 15666[19]
Noise sensitivity	continuous scale: [0;10]

A. Verbalization description

Verbalizations from the CN experiment were compared to those collected during single noise experiments (*cf.*[12,13]) in Table IV. This step helps in the work of identifying and proposing relevant noise indices to characterize noise annoyance.

The highlighted influential acoustical features were similar in the different experiments. A deep analysis showed that occurrences might be different from the single noise experiment to the CN experiment. Actually, for $URTN$, only the global temporal variation was as

TABLE IV. Participant verbalizations on $URTN$ (Urban Road Traffic Noise) and AN (Aircraft Noise) within the 3 different experiments (separated by a year) and their occurrences. CN: Combined Noise; PTW: Powered-two-wheelers. %: percentage of respondents who cited this item.

Verbalizations	Single noise experiment		CN experiment	
	$URTN$	AN	$URTN$	AN
Presence of PTWs	100%		66%	
Perceived noise intensity	56%	39%	19%	44%
Global temporal variation	56%	55%	56%	59%
Spectral content	24%	58%		
Modulation-related	6%			
Timbre			13%	25%

frequently cited in both experiments. All the other acoustical features were less frequently cited in the CN experiment than in the $URTN$ experiment. For AN, both global temporal variation and perceived noise intensity were equally mentioned in both experiments. Only timbre was less cited in the CN experiment than in the AN experiment. These differences in occurrence for timbre might be due to masking effects between the CN or due to participant difficulties in describing acoustic content of CN sequences. The difficulty might be due to

the fact that they were having a read during the experiment. Nevertheless, this showed that the same acoustical features were noticed by participants and related to annoyance.

B. Analysis of variance for annoyance ratings in the *CN* experiment

Stimuli of the combined noise experiment were constructed on the basis of two factors: "URTN SEQ" for $URTN$ sequence and "AN SEQ" for AN sequence. The effects of these factors on partial annoyance due to URTN, on partial annoyance due to AN and on total noise annoyance were studied using two-factorial repeated-measures analyses of variance (RM ANOVA), with four levels per factor.

1. Partial annoyance due to URTN

The two main factors "URTN SEQ" and "AN SEQ" had a significant effect on $URTN$ annoyance (respectively, [F(3, 93)=75.59; p<0.05, ϵ=1] and [F(3, 93)=7.66; p<0.05, ϵ=0.92]) The proportion of variance (η^2) explained by the factor "URTN SEQ" was moderate, i.e. 26%, and the one explained by the factor "AN SEQ" was very small, i.e. 3%. The interaction between the factors "URTN SEQ" × "AN SEQ" had no effect on $URTN$ partial annoyance ([F(9, 279)=1.72; p>0.05, ϵ=0.85]).

2. Partial annoyance due to AN

Only the main factor "AN SEQ" had a significant effect on AN annoyance ([F(3, 93)=40.55; p<0.05, ϵ=0.84]). The proportion of variance explained (η^2) by the factor "AN SEQ" was moderate, i.e. 28%. Both the main factor "URTN SEQ" and the interaction between the factors "URTN SEQ" × "AN SEQ" had no effect on the AN partial annoyance (respectively, [F(3, 93)=1.73; p>0.05, ϵ=1] and [F(9, 279)=0.79; p>0.05, ϵ=0.78]).

3. Total annoyance

The two main factors "URTN SEQ" and "AN SEQ" had a significant effect on total noise annoyance (respectively, [F(3, 93)=20.34; p<0.05, ϵ=1] and [F(3, 93)=15.38; p<0.05, ϵ=1]). The proportion of variance explained (η^2) by the factor "URTN SEQ" was moderate, i.e. 15%, but higher than the proportion of variance explained by the factor "AN SEQ", i.e. 8%, indicating that $URTN$ sequences influenced total annoyance more than AN sequences did. The interaction between the factors "URTN SEQ" × "AN SEQ" had no effect on total annoyance ([F(9, 279)=1.72; p>0.05, ϵ=0.85]).

C. Single noise annoyance models

Multilevel regression analysis was performed to consider acoustical and individual data in annoyance models using various experiments. To consider both partial and specific an-

noyance ratings in the construction of models, as already seen in the literature (*e.g.*[1]), new models were computed by aggregating the aforesaid data. Furthermore, aggregating the results of different experiments increases noise exposure variability. Indeed, the noise sequences of the different experiments had neither the same duration ($URTN$ sequences lasted for 3 mins in $URTN$ experiment, aircraft flyover lasted between 22 and 64 secs, CN sequences lasted for 6 mins) nor the same noise level ($URTN$ sequences L_{Aeq} ranged from 57.3 to 63.7 dB(A) in $URTN$ experiment and from 45.7 to 51.9 dB(A) in CN experiment; aircraft flyover noise L_{Aeq} ranged from 43.9 to 54.6 dB(A) in AN experiment and from 35.0 to 55.6 dB(A) in CN experiment). The built models are therefore relevant for a wider noise exposure range (single and combined noise exposures, different durations and different noise levels).

The linear multilevel regression will be briefly presented below (for more details, see Hox[20] and Trollé *et al.*[5,6]).

Model specification: As data were obtained from several repeated measure experiments, the first level of the regression model refers to the stimulus ($URTN$ sequence or AN sequence, denoted as i in subscript). The second level refers to the individual (denoted as j in subscript) and the third level refers to the experiment (denoted as k in subscript - no variables and no error terms are introduced in the model at this level)[20]. An individual could therefore assess specific annoyance due to a noise sequence during one of the single noise experiments and partial annoyance due to the same noise sequence during the CN experiment. Considering a model with one variable at the individual level (the noise sensitivity, denoted as $Sens$) and M variables at the stimulus level (denoted as $Index_m$), the

formulae are as follows, with A_{ijk} the annoyance rating of the individual j for the stimulus i in the experiment k:

$$A_{ijk} = \pi_{0jk} + \sum_{m=1}^{M} \pi_{mjk} Index_{mi} + e_{ijk}$$

$$\pi_{0jk} = \beta_{000} + \beta_{011} \times Sens_{jk} + u_{0jk}$$

$$\pi_{mjk} = \beta_{m00} + \beta_{m11} \times Sens_{jk} + u_{mjk}$$

$$\begin{bmatrix} u_{0jk} \\ \vdots \\ u_{mjk} \\ \vdots \\ u_{Mjk} \end{bmatrix} \backsim \mathcal{N} \left(0, \begin{bmatrix} \sigma^2_{u_0} & \cdots & \cdots & \cdots & \sigma_{u_{0M}} \\ \vdots & \ddots & \vdots & & \vdots \\ \sigma_{u_{m0}} & \cdots & \sigma^2_{u_m} & \cdots & \sigma_{u_{mM}} \\ \vdots & & \vdots & \ddots & \vdots \\ \sigma_{u_{M0}} & \cdots & \cdots & \cdots & \sigma^2_{u_M} \end{bmatrix} \right)$$

$$\text{for } m = 1, ..., M, \text{ for } j = 1, ..., J \text{ and } k = 1, ..., K$$

$$e_{ijk} \backsim \mathcal{N}(0, \sigma^2_e) \text{ for } i = 1, ..., I, j = 1, ..., J$$

$$\text{and } k = 1, ..., K$$

Computation of multilevel regression is Bayesian and the studied parameter influence is made using Bayesian posterior distribution[20]. The posterior distributions of the model parameters are approximated via Markov Chain Monte Carlo (MCMC) simulation, with 350,000 iterations.

To select the model with the best goodness-of-fit, three criteria are used:

- R^2_1: the proportion of variance explained at the stimulus level. R^2_1 varies from 0 to 1. The closer R^2_1 is to 1, the better is the model goodness-of-fit.

- $R^2_{2,m}$ (m=0, ..., M): the proportion of variance explained at the individual level. This criterion is computed for each random coefficient at the individual level: the intercept π_{0j} and the slope(s) π_{mj}. The slope π_{mj} can be fixed, random or can account for a moderating effect due to noise sensitivity. If noise sensitivity explains the variation of each random coefficient (π_{0j} or π_{mj}), $R^2_{2,m}$ varies from 0 to 1. The closer $R^2_{2,0}$ (calculated for π_{0j}) is to 1, the more noise sensitivity has an effect on individual mean rating. The closer $R^2_{2,m}$ (calculated for π_{mj}) is to 1, the more noise sensitivity has a moderating effect on the relationship between the m^{th} index and annoyance ratings.

- Deviation Information Criterion (DIC): This criterion provides a measure of out-of-sample predictive error[20]. The lower the DIC is, the better is the predictive power of the model.

1. URTN annoyance models

As a first step, a classical model, with L_{Aeq}, without noise sensitivity or individual error term in the intercept (latter referred as CM_{road}) was calculated:

$$A_{ijk} = 3.58 + 0.02(n.s.) \times L_{Aeq} + e_{ijk}$$

$$e_{ijk} \backsim \mathcal{N}(0, 4.66) \text{ for } i = 1, ..., I, j = 1, ..., J$$

$$\text{and } k = 1, ..., K$$

The proportion of variance explained at the stimulus level (R^2_1) was equal to 3%, with a non-significant (denoted by $n.s.$) slope coefficient (0.02) of annoyance relative to L_{Aeq},

classically used as a single variable in an annoyance model. This highlights that such model, solely based on L_{Aeq}, is not relevant.

Then, null models, *i.e.* without explanatory variables at the stimulus level, without (M0a) and with (M0b) noise sensitivity in the intercept, were tested. For M0a, R^2_1=0.57 and DIC=5107. For M0b, R^2_1=0.57, $R^2_{2,0}$=0.20 and DIC=5107. Due to the $R^2_{2,0}$ value, noise sensitivity was considered for intercept modeling and M0b was further used as a baseline.

Finally, several combinations of indices were selected to account for the mentioned annoying acoustical features of the noise: L_{Aeq} or loudness N for perceived sound intensity, the indicator URA (*cf.* Klein *et al.* [7]) for perceived sound intensity and timbre and $\sigma'(N)$ the temporal derivative of loudness (*cf.* Gille *et al.* [13]) for temporal features. The best combinations were kept, according to DIC, R^2_1 and $R^2_{2,m}$ criteria, and including the combinations of noise indices already highlighted in Gille *et al.*[13] for $URTN$ annoyance:

- L_{Aeq} & random slope (a model denoted by $L_{\text{Aeq,road}}rand$),
- Loudness N & moderating effect ($N_{\text{road}}mod$),
- The indicator URA & moderating effect ($URA_{\text{road}}mod$),
- N, the derivative of loudness $\sigma'(N)$ & random slope ($LD_{\text{road}}rand$ for "Loudness and its Derivative")
- and URA, $\sigma'(N)$ & fixed slope ($URAD_{\text{road}}fix$ for "URA and loudness Derivative").

Using standardized coefficients, the contribution of each variable[21] to the model was determined. Noise sensitivity significantly contributed to the five models (37% for $L_{\text{Aeq,road}}rand$; 37% for $N_{\text{road}}mod$; 35% for $URA_{\text{road}}mod$; 31% for $LD_{\text{road}}rand$; 40% for $URAD_{\text{road}}fix$). In

TABLE V. Annoyance models for $URTN$. The values of L_{Aeq}, N, URA and $\sigma'(N)$ were grand mean centered with the grand-mean 56.3 dB(A), 5.67 sones, 4.74 and 33.49 sone/s respectively. Co.: Coefficient; [*St. Co.*]: Standardized Coefficient; 1^{st} L: Stimulus Level; 2^{nd} L: Individual Level. *: not significantly different from 0; σ_e^2, σ_{u0}^2, σ_{u1}^2, σ_{u2}^2: variances of residual errors; σ_{u01}, σ_{u02}; σ_{u12}: covariances of residual errors.

Model:	$L_{\text{Aeq,road}}rand$	$N_{\text{road}}mod$	$URA_{\text{road}}mod$	$LD_{\text{road}}rand$	$URAD_{\text{road}}fix$
Index:	L_{Aeq}	N	URA	N and $\sigma'(N)$	URA and $\sigma'(N)$
	Co. [*St. Co.*]	Co. [*St. Co.*]	Co. [*St. Co.*]	Co. [*St. Co.*]	Co. [*St. Co.*]
Fixed part					
β_{000}	1.31*	2.25	2.32	3.04	2.11
β_{011} ($Sens$)	0.54 [*0.56*]	0.47 [*0.49*]	0.50 [*0.52*]	0.34 [*0.34*]	0.45 [*0.46*]
β_{100} (L_{Aeq}, N or URA)	0.34 [*0.97*]	0.35 [*0.70*]	0.67 [*0.83*]	0.39 [*0.43*]	0.37 [*0.26*]
β_{111} ($Sens$ x (L_{Aeq} or N))	–	0.05 [*0.12*]	0.09 [*0.14*]	–	–
β_{200} ($\sigma'(N)$)	–	–	–	0.05 [*0.30*]	0.08 [*0.43*]
Random part					
σ_e^2 (1^{st} L)	1.29	1.27	1.28	1.23	1.39
σ_{u0}^2 (2^{nd} L)	6.93	4.94	6.52	5.35	4.03
(σ_{u01}; σ_{u1}^2) (2^{nd} L)	(-0.17; 0.03)	(0.30; 0.08)	(0.87; 0.32)	(0.47; 0.24)	–
(σ_{u02}; σ_{u12}; σ_{u2}^2) (2^{nd} L)	–	–	–	(-0.028;-0.031;0.007)	–
Explained variance					
R_1^2 (1^{st} L)	0.72	0.73	0.73	0.74	0.70
$R_{2,0}^2$ (2^{nd} L)	0.12	0.16	0.14	0.13	0.18
$R_{2,1}^2$ (2^{nd} L)	–	0.10	0.09	–	–
DIC	4529	4503	4506	4474	4594

particular, the interaction term between noise sensitivity and N or URA contributed to the models (9% for $N_{\text{road}}mod$; 9% for $URA_{\text{road}}mod$). The interaction also increased the contribution of noise sensitivity to the models. This highlights relevance of considering explanatory variables at the individual level in order to improve $URTN$ annoyance models. In fact, the model CM_{road}, based solely on L_{Aeq}, was not relevant to explain noise annoyance.

2. AN annoyance models

As in Section III C 1, a classical model, with L_{Aeq} but without noise sensitivity or individual error term in the intercept (CM_{air}) was calculated for AN annoyance:

$$A_{ijk} = -6.76 + 0.21 \times L_{Aeq} + e_{ijk}$$

$$e_{ijk} \backsim \mathcal{N}(0, 6.65) \text{ for } i = 1, ..., I, j = 1, ..., J$$

$$\text{and } k = 1, ..., K$$

The proportion of variance explained at the stimulus level (R_1^2) is equal to 17%.

Then, null models, without (M0a) and with (M0b) noise sensitivity in the intercept, were tested. For M0a, R_1^2=0.48 and DIC=4003. For M0b, R_1^2=0.48, $R_{2,0}^2$=0.15 and DIC=4003. Due to the $R_{2,0}^2$ value, noise sensitivity was therefore kept for intercept modeling and M0b was further used as a baseline.

Finally, several combinations of indices were selected to characterize the aforementioned annoying sound characteristics: L_{Aeq}, N or loudness exceeded 10% of the time N_{10} for perceived sound intensity, $\sigma'(N)$ for global temporal variation, the mean specific loudness integrated between Barks 1 and 12 N_{1-12} and the total energy of tonal components from

303 Barks 13 to 18 $TETC_{13-18}$ for spectral content (*cf.* Gille *et al.* [12]). The best combinations
304 were kept, according to DIC, R_1^2 and $R_{2,m}^2$ criteria:

305 - L_{Aeq} & random slope (a model denoted by $L_{\text{Aeq,air}}rand$),

306 - N & random slope ($N_{\text{air}}rand$),

307 - N_{10} & random slope ($N_{10,\text{air}}rand$),

308 - $\sigma'(N)$ & moderating effect ($\sigma'(N)_{\text{air}}mod$)

309 - and N_{1-12}, $TETC_{13-18}$ & fixed slopes ($LMLHT_{\text{air}}fix$, for "Low and Medium frequency
310 Loudness and High frequency Tonal component").

311 Noise sensitivity significantly contributed to the models (38% for $L_{\text{Aeq,air}}rand$; 24% for
312 $N_{\text{air}}rand$; 16% for $N_{10,\text{air}}rand$; 37% for $\sigma'(N)_{\text{air}}mod$; 51% for $LMLHT_{\text{air}}fix$). In particular,
313 the interaction term between noise sensitivity and $\sigma'(N)$ contributed to the models (9%
314 for $\sigma'(N)_{\text{air}}mod$), which also increased the contribution of noise sensitivity to the model.
315 Considering explanatory variables at the individual level with the aim to contributing to
316 AN annoyance model enhancement was shown to be relevant. In fact, the model CM_{air}
317 explained a smaller proportion of the variance in annoyance ratings than models considering
318 an individual characteristic.

319 **D. Noise indices for each surveyed respondent**

320 The models constructed from laboratory data were shown to be relevant compared to
321 models based solely on L_{Aeq}. But to test these noise annoyance models using the field mea-
322 sured partial annoyance ratings, the different noise indices involved in the models must be

TABLE VI. Annoyance models for AN. The values of L_{Aeq}, N, N_{10}, $\sigma'(N)$, N_{1-12} and $TETC_{13-18}$ were grand mean centered with the grand-mean 49.3 dB(A), 7.05 sones, 5.18 sones, 22.64 sone/s, 2.67 sones and 44.5 dB, respectively. All coefficients were significantly different from 0. σ_e^2, σ_{u0}^2, σ_{u1}^2: variances of residual errors; σ_{u01}: covariances of residual errors.

Model:	$L_{\mathrm{Aeq,air}}rand$	$N_{\mathrm{air}}rand$	$N_{10,\mathrm{air}}rand$	$\sigma'(N)_{\mathrm{air}}mod$	$LMLHT_{\mathrm{air}}fix$
Index:	L_{Aeq}	N	N_{10}	$\sigma'(N)$	N_{1-12} & $TETC_{13-18}$
	Co. [*St. Co.*]	Co. [*St. Co.*]	Co. [*St. Co.*]	Co. [*St. Co.*]	Co. [*St. Co.*]
Fixed part					
β_{000}	2.03	2.89	3.56	1.53	1.43
β_{011} ($Sens$)	0.30 [*0.24*]	0.24 [*0.18*]	0.28 [*0.22*]	0.39 [*0.31*]	0.40 [*0.32*]
β_{100} ($Index$)	0.21 [*0.39*]	0.39 [*0.58*]	1.14 [*1.19*]	0.08 [*0.44*]	0.87 [*0.30*]
β_{111} ($Sens$ x $Index$)	–	–	–	0.01 [*0.08*]	–
β_{200} ($TETC_{13-18}$)	–	–	–	–	0.06 [*0.17*]
Random part					
σ_e^2 (1^{st} L)	3.09	3.09	3.05	3.07	3.30
σ_{u0}^2 (2^{nd} L)	3.03	6.09	11.63	2.65	2.70
(σ_{u01}; σ_{u1}^2) (2^{nd} L)	(0.08; 0.01)	(0.45; 0.06)	(3.59; 1.34)	(0.06; 0.005)	–
Explained variance					
R_1^2 (1^{st} L)	0.62	0.62	0.62	0.62	0.59
$R_{2,0}^2$ (2^{nd} L)	0.16	0.05	0.11	0.22	0.23
$R_{2,1}^2$ (2^{nd} L)	–	–	–	0.10	–
DIC	3692	3689	3689	3684	3722

estimated as it is not possible to measure them on a wide area. Indeed, the survey database classically offered ratings for each respondent annoyance and noise exposure expressed in terms of L_{den} from strategic noise maps for each noise source. This section will investigate if the different noise indices may be estimated from L_{den} values and will assess limitations of such estimation.

Therefore, a 3-step methodology was proposed to estimate the indices using L_{den}.

1 - Noise index calculation in the survey area. Ninety urban road vehicle pass-by noises (30 powered-two-wheelers, 30 heavy vehicles and 30 light vehicles) and 30 urban road traffic noises were randomly selected from recordings[22] carried out at a receiver point of the surveyed area. Moreover, twelve aircraft flyover noises were recorded in the surveyed area. The mean values of the noise indices were calculated from these recordings (*cf.* Table VII).

2 - Noise index variation as a function of noise level variation. The variation of noise indices with equivalent noise level was assessed using noise sequences stemming from Gille *et al.*[13] for $URTN$ and from Gille *et al.*[12] for AN. Six $URTN$ sequences (1T3, 1T9, 1T13, 1T15, 2T5 and 2T7, *cf.*[13]) and three AN sequences (a2, a6 and a9, *cf.*[12]) (*i.e.* a quarter of the noise sequences studied in Gille *et al.*[13] and in Gille *et al.*[12]) were equalized in L_{Aeq} at 7 noise levels separated by a step of 5 dB(A), with a reference level $L_{Aeq\ ref}$ corresponding to a sound pressure level within the experiments. Then, relationships between noise indices and the variation $\Delta L_{Aeq} (= L_{Aeq} - L_{Aeq\ ref})$ were computed. $TETC_{13-18}$ evolved linearly with ΔL_{Aeq}, whereas the other noise indices (N, N_{10}, $\sigma'(N)$, N_{1-12} and URA) evolved exponentially with ΔL_{Aeq} (*cf.* Table VII).

TABLE VII. Field mean values, function depending on L_{Aeq} variation (denoted by ΔL_{Aeq}) and its α coefficient value of the different noise indices for urban road traffic noise and aircraft noise.

	Urban road traffic noise Field $L_{Aeq\ mean}$ = 71.1 dB(A)			**Aircraft noise** Field $L_{Aeq\ mean}$ = 68.1 dB(A)				
Index	N	$\sigma'(N)$	URA	N	N_{10}	$\sigma'(N)$	N_{1-12}	$TETC_{13-18}$
Field mean	23.99	113.87	13.29	14.83	25.78	75.52	11.58	53.7
Function	$e^{\alpha\times\Delta L_{Aeq}}$	$e^{\alpha\times\Delta L_{Aeq}}$	$e^{\alpha\times\Delta L_{Aeq}}$	$e^{\alpha\times\Delta L_{Aeq}}$	$e^{\alpha\times\Delta L_{Aeq}}$	$e^{\alpha\times\Delta L_{Aeq}}$	$e^{\alpha\times\Delta L_{Aeq}}$	$\alpha\times\Delta L_{Aeq}$
α	0.0747	0.0568	0.0528	0.0676	0.0647	0.0591	0.0656	0.5980

N.B.: In the literature, loudness is often given as a function of sound pressure, according to the equation given by Stevens [23]: $N = k \times p^{0.6}$. In the present study, loudness was calculated as $N = N_{mean} \times e^{\alpha\times(L_{Aeq}-L_{Aeq\ ref})}$, with α=0.0747 for $URTN$ and α=0.0676 for AN. This equation can also be written as:

$$N = \frac{N_{mean}}{e^{\left(\alpha\times L_{Aeq\ ref}+20\alpha\times\log_{10}(p_0)\right)}} \times p_A^{\frac{20\alpha}{\ln(10)}} \quad (1)$$

with p_A the A-weighted sound pressure and p_0 the reference sound pressure, equal to 2×10^{-5} Pa.

Using the α values given in Table VII, for $URTN$, $\frac{20\alpha}{\ln(10)} = 0.649$ and for AN, $\frac{20\alpha}{\ln(10)} = 0.587$. The equation in the present study is therefore equivalent to the one of Stevens, with k replaced by $\frac{N_{mean}}{e^{\left(\alpha\times L_{Aeq\ ref}+20\alpha\times\log_{10}(p_0)\right)}}$.

3 - Noise index estimation for each respondent. The noise level to which the respondent was exposed was estimated from L_{den}. In the functions given in Table VII,

ΔL_{Aeq} was therefore replaced by L_{den} - $L_{\text{Aeq mean}}$ with L_{den} from strategic noise maps and field $L_{\text{Aeq mean}}$. Such replacement of L_{Aeq} by L_{den} has already been shown in the literature (*e.g.*[24]). Using field mean values and α values (*cf.* Tables VII), indices for each surveyed respondent could be estimated (N, N_{10}, $\sigma'(N)$, N_{1-12} and URA were denoted as $ExpIndex$):

$$\begin{aligned} ExpIndex =& ExpIndex_{mean} \\ & \times e^{(\alpha \times (L_{den} - L_{Aeq\ mean}))} \end{aligned} \tag{2}$$

$$\begin{aligned} TETC_{13-18} =& TETC_{13-18mean} \\ & + \alpha \times (L_{den} - L_{Aeq\ mean}) \end{aligned} \tag{3}$$

E. Noise annoyance models tested using field measured partial annoyance ratings

Using equations 2 and 3, noise indices and therefore the proposed annoyance models could be estimated for each respondent (denoted as r in subscript) of the survey. Models were tested, as done by Miedema[25] and Klein *et al.*[26], using only fixed parameters of multilevel models, *i.e.* in a simpler form, with grand-mean value (denoted as $Index_{m\ GM}$) for each index m, as follows:

$$A_r = \pi_{0r} + \sum_{m=1}^{M} \pi_{mr}(Index_{mr} - Index_{m\ GM})$$

$$\pi_{0r} = \beta_{000} + \beta_{011} \times Sens_r$$

$$\pi_{mr} = \beta_{m00} + \beta_{m11} \times Sens_r$$

Annoyance models testing consisted in comparing predicted annoyance with individual partial annoyance ratings measured in field for road traffic noise and for aircraft noise respectively (*cf.* Table VIII). Three parameters (r, intercept and slope) were used to assess

the partial annoyance models: they resulted from the correlation and regression analysis between measured and predicted partial annoyance. These 3 parameters allowed the quality of the partial annoyance models (underestimation or overestimation) to be assessed. A perfect prediction by a model would lead to (r, intercept, slope) = (1, 0, 1).

For road traffic noise, the models could be divided into four significantly different groups: i) $L_{\mathrm{Aeq,road}}rand$ and $URA_{\mathrm{road}}mod$ ($0.42 \leq r \leq 0.45$), ii) $N_{\mathrm{road}}mod$ and $URAD_{\mathrm{road}}fix$ ($0.38 \leq r \leq 0.39$) and iii) the other models ($0.22 \leq r \leq 0.33$). For aircraft flyover noise, the different models simultaneously considering noise indices and noise sensitivity provided similar prediction of the partial annoyance ratings ($0.47 \leq r \leq 0.52$). All the models simultaneously considering noise indices and noise sensitivity were better correlated with the measured partial annoyance ratings than did the models CM_{road} and CM_{air} solely using the variable L_{den}. This result showed that the annoyance models and the methodology proposed to estimate field noise index values enabled to better predict partial noise annoyance ratings than L_{den} alone.

In the following, for road traffic noise, $L_{\mathrm{Aeq,road}}rand$ and $URA_{\mathrm{road}}mod$ were kept, as these models better correlated with field annoyance ratings. $L_{\mathrm{Aeq,road}}rand$ was very simple as there were no need to estimate an index and $URA_{\mathrm{road}}mod$ enabled to characterize different annoying acoustical features of road traffic noise *i.e.* perceived noise intensity, spectral content and modulation-related sensations. For aircraft noise, models with the highest correlation with field measured partial annoyance were kept: $L_{\mathrm{Aeq,air}}rand$ which was very simple to use and $\sigma'(N)_{\mathrm{air}}mod$ as this model enabled to characterize global temporal variation. As two

TABLE VIII. Correlation coefficients r between predicted partial annoyance and measured partial annoyance ratings for road traffic noise and for aircraft noise of the field survey. All correlation coefficients are significantly different from 0.

Road traffic noise

Model	intercept	slope	r
CM_{road} [a]	4.48	0.01	0.22
$L_{\mathrm{Aeq,road}}rand$ [a]	3.05	0.41	0.45
$N_{\mathrm{road}}mod$	4.83	0.54	0.39
$URA_{\mathrm{road}}mod$	5.13	0.54	0.42
$LD_{\mathrm{road}}rand$	5.97	0.40	0.32[b]
$URAD_{\mathrm{road}}fix$	5.37	0.41	0.38

Aircraft noise

Model	intercept	slope	r
CM_{air} [a]	2.16	0.16	0.40
$L_{\mathrm{Aeq,air}}rand$ [a]	1.53	0.28	0.51
$N_{\mathrm{air}}rand$	1.96	0.17	0.50
$N_{10,air}rand$	3.97	0.49	0.47
$\sigma'(N)_{\mathrm{air}}mod$	1.95	0.30	0.52
$LMLHT_{\mathrm{air}}fix$	1.98	0.31	0.51

[a] As done before, L_{Aeq} was replaced by L_{den}.

[b] A similar model using L_{den} (instead of N), $\sigma'(N)$ and noise sensitivity leads to a smaller correlation of 0.25.

noise indices had to be estimated, $LMLHT_{\text{air}}fix$ was not kept, whereas it characterized spectral content.

F. Total noise annoyance models built from laboratory data

The CN experiment data were used to construct total noise annoyance models. Different total annoyance models, described and discussed in the literature (*cf.*[27]), were adjusted in the current study using the mean measured total annoyance ratings stemming from the CN experiment.

Psychophysical models (*i.e.* total annoyance is described as a function of acoustical variables, *cf.* Table IX) used variables L_{Aeq}, and also N for each noise source in combination as proposed in Morel *et al.*[9]. Perceptual models (*i.e.* total annoyance is described as a function of perceptual variables, *cf.* Table IX) used mean calculated partial annoyance ratings as variables. The adjusted determination coefficient R^2_{adj} was used to compare total annoyance models: the higher R^2_{adj}, the better the goodness-of-fit. The models for which one coefficient was not significantly different from 0, or for which R^2_{adj} was inferior to 0.10, were not displayed in Table IX.

This model comparison highlighted that several psychophysical models had to be considered in the following to be tested using field survey data: i) considering L_{Aeq}, energy summation model, independent effect model and weighted summation model and ii) considering N, independent effect model. Linear regression model had to be kept and tested in the following using the field survey data considering the partial annoyance ratings pre-

TABLE IX. Psychophysical and perceptual total annoyance models from the literature (*cf.*[27]) and constructed in the current work using noise indices or the calculated partial annoyance of the combined noise experiment. *: not significantly different from 0. R^2_{adj}: The adjusted determination coefficient.

Psychophysical model	Equation	R^2_{adj}
Energy summation[a]	$A_T = -10.75 + 0.31 \times L_T$	0.86
Independent effects[b]	$A_T = -8.02 + 0.22 \times L_{\text{Aeq,road}} + 0.05 \times L_{\text{Aeq,air}}$	0.74
	$A_T = 2.41 + 0.66 \times N_{\text{road}} + 0.07 \times N_{\text{air}}$	0.76
Weighted summation[c]	$A_{\text{road}} = -8.40 + 0.27 \times L_{\text{Aeq,road}}$	0.97
(also named	$A_{\text{air}} = -6.03 + 0.19 \times L_{\text{Aeq,air}} \Rightarrow P_{\text{air}} = (2.37 - 0.08 \times L_{\text{Aeq,air}})/0.27$	0.74
quantitative model)	k=10 $A_T = -10.08 + 0.30 \times L_t$	0.74
	k=15 $A_T = -10.59 + 0.30 \times L_t$	0.76
Perceptual model	**Equation**	R^2_{adj}
Linear regression[d]	$A_T = 2.66 + 0.25 \times \sigma'(N)_{\text{air}} mod + 0.65 \times L_{\text{Aeq,road}} rand$	0.75
Strongest component[e]	$A_T = max(\sigma'(N)_{\text{air}} mod; L_{\text{Aeq,road}} rand)$	0.29
Vector summation[f]	$A_T = (\sigma'(N)_{\text{air}} mod^2 + L_{\text{Aeq,road}} rand^2 + 2 \times \sigma'(N)_{\text{air}} mod \times L_{\text{Aeq,road}} rand \times cos(1.26 rad))^{1/2}$	0.56
Linear regression	$A_T = 0.32^* + 0.25 \times \sigma'(N)_{\text{air}} mod + 1.13 \times URA_{\text{road}} mod$	0.70
Strongest component	$A_T = max(\sigma'(N)_{\text{air}} mod; URA_{\text{road}} mod)$	0.29
Vector summation	$A_T = (\sigma'(N)_{\text{air}} mod^2 + URA_{\text{road}} mod^2 + 2 \times \sigma'(N)_{\text{air}} mod \times URA_{\text{road}} mod \times cos(1.41 rad))^{1/2}$	0.43
Linear regression	$A_T = 1.52 + 0.23 \times L_{\text{Aeq,air}} rand + 0.65 \times L_{\text{Aeq,road}} rand$	0.74
Strongest component	$A_T = max(L_{\text{Aeq,air}} rand; L_{\text{Aeq,road}} rand)$	0.12*
Linear regression	$A_T = -0.82^* + 0.23 \times L_{\text{Aeq,air}} rand + 1.13 \times URA_{\text{road}} mod$	0.70
Strongest component	$A_T = max(L_{\text{Aeq,air}} rand; URA_{\text{road}} mod)$	0.12*

[a] Total annoyance is a function of the sound pressure level of the combined noises, calculated as an energy summation of the sound pressure level of each combined noise.

[b] Total annoyance is a function of each combined noise sound pressure level.

[c] Total annoyance is expressed as a function of the total sound rating, determined calculating level-dependent penalties of the noises in the combination with respect to a road traffic noise.

[d] Total annoyance is a function of each specific/partial annoyance due to each noise in the combination.

[e] Total annoyance is equal to the maximum of the specific/partial annoyances.

[f] Total annoyance is the result of a vector addition of the specific/partial annoyances of the combined noises.

dicted by the models $L_{\text{Aeq,road}}rand$ or $URA_{\text{road}}mod$ for road traffic noise and $\sigma'(N)_{\text{air}}mod$ or $L_{\text{Aeq,air}}rand$, for aircraft noise (*cf.* Section III C).

G. Total noise annoyance models tested using survey data

Total annoyance models (*cf.* Table IX) were tested using individual field total annoyance ratings (*cf.* Table X). Noise indices used in total annoyance models to consider field exposure to road traffic and aircraft noises (*i.e.* N, $\sigma'(N)$ and URA) were estimated following the methodology presented in Section III D. For each noise source, L_{Aeq} in models was replaced by L_{den} value of the survey.

Total annoyance predicted from independent effect model with N was not significantly correlated with field total annoyance. Perceptual total annoyance models provided a better prediction of individual total annoyance ratings than the psychophysical ones. In particular, linear regression model with $L_{\text{Aeq,road}}rand$ and $L_{\text{Aeq,air}}rand$, *i.e.* perceptual model using partial annoyance predicted from both L_{den} and noise sensitivity, provided a better prediction of individual total annoyance ratings (r=0.47) than the independent effect model using only L_{den} (r=0.20). These results show that total annoyance models based on partial annoyance models considering individual noise sensitivity associated to noise index estimation enabled a better prediction of field total annoyance ratings than total annoyance models based only on L_{den}.

TABLE X. Total annoyance models tested using individual total annoyance ratings of the field survey. Intercept, slope and r resulted from the correlation and regression analysis between field total annoyance ratings and predicted ones. *: not significantly different from 0.

Model with Index	intercept	slope	r
Energy summation with L_{den}	6.23	0.10	0.18
Independent effects with L_{den}	6.60	0.09	0.20
Weighted summation k=10 with L_{den}	7.01	0.09	0.16
Weighted summation k=15 with L_{den}	7.11	0.09	0.17
Independent effects with N	8.41	0.15*	0.12*
Linear regression with $\sigma'(N)_{\text{air}} mod$ & $L_{\text{Aeq,road}} rand$	5.07	0.33	0.48
Linear regression with $\sigma'(N)_{\text{air}} mod$ & $URA_{\text{road}} mod$	6.68	0.62	0.41
Linear regression with $L_{\text{Aeq,air}} rand$ & $L_{\text{Aeq,road}} rand$	3.80	0.32	0.47
Linear regression with $L_{\text{Aeq,air}} rand$ & $URA_{\text{road}} mod$	5.41	0.60	0.41

IV. DISCUSSION

Noise annoyance due to $URTN$, noise annoyance due to AN and total annoyance due to these combined noises were studied using laboratory and field data. A verbalization task enabled to highlight three main annoying acoustical features for $URTN$ and for AN: i) global temporal variation, ii) perceived noise intensity and iii) timbre (*cf.* Table IV). These

acoustical features were similar to the ones observed in the previous single noise experiments (*cf.*[12,13]) and are well-known to influence annoyance due to transportation noises [4,28,29].

Multilevel regressions were performed considering different annoyance ratings (*e.g.* specific and partial annoyance ratings) and different noise situations (*e.g.* single and combined noise exposures). Indeed, multilevel regression enables to consider experiments as a level of the equation. This mathematical model is therefore appropriate to consider the structure of the data. Moreover, such aggregation of data from different experiments has already been done in the literature: for the construction of exposure-effect relationships, Miedema and Oudshoorn[1] used data from several surveys, carried out in different countries, for different combined noise sources and single noise sources. Noise sensitivity highly contributed in all built models, confirming the findings of the literature (*e.g.*[5,6,11–13,30]). Some models highlighted a significant moderating effect of noise sensitivity on the relationship between noise indices and noise annoyance (*e.g.* $N_{\text{road}}mod$, $URA_{\text{road}}mod$ and $\sigma'(N)_{\text{air}}mod$), as found by some authors (*e.g.*[11,31]) whereas other authors did not observe such an effect (*e.g.*[5,6,30]). This may be explained by the fact that the moderating effect was weak when significant. The great influence of noise sensitivity on annoyance shows therefore the necessity to build new exposure-response relationships considering this factor, as proposed in Gille *et al.*[2], and to perform more surveys measuring noise sensitivity in order to be able to predict noise sensitivity for unmeasured samples of the population.

A methodology was proposed to calculate field psychoacoustic indices using L_{den} from noise maps. This methodology permits the calculation of loudness N and the corresponding equation is similar to Stevens' well-known relationship[23]. This similarity between equations

endorses the methodology. It seems therefore that the good correlations are not due to a particular situation (noise recordings in a specific area) but reflect the relationship between noise level and sensations evoked by noise (*i.e.* perceived sound intensity). Moreover, this methodology and the partial annoyance models enable a good prediction of field noise annoyance, both for $URTN$ and AN in comparison with annoyance models based solely on L_{den} (*cf.* Table VIII). However, it should be noted that the obtained index values are only an estimation of the corresponding indices. Indeed, they are predicted using L_{den}, which is an estimation of the field noise level, and using different equations involving mean values, measured in the survey area. These mean values also contribute to describe the soundscape of this area. More field noise recordings need to be performed in order to validate this methodology on wider areas, with different noise scenarios and to define mean values valid for these situations. Mapping of psychoacoustic indices for large areas are mentioned as potential alternatives to improve annoyance models (*cf.* Lercher *et al.*[32]). The methodology could be also used to estimate field psychoacoustic indices in other situations, *e.g.* soundscape studies. Further testings of this methodology are therefore of interest for different acoustic fields.

Total noise annoyance was studied using the CN experiment data. According to an ANOVA, both types of noise exposure had an influence on total noise annoyance, but the proportion of variance explained by $URTN$ is higher than the one explained by AN. Such result was already observed by Taylor[33], who hypothesized that the influence of each source is governed by its duration of apparition. Total annoyance models from the literature were built using L_{Aeq} and N, and using calculated partial noise annoyance. Total noise annoyance

calculated from several models was well correlated with measured total annoyance. The corresponding models were considered for further prediction of total annoyance. Psychophysical models under consideration allowed calculated total annoyance to be well correlated with measured total annoyance as some perceptual models. This result was not expected as several studies had shown more differences between perceptual and psychophysical models in terms of quality of adjustment (*e.g.* [26]). Actually, perceptual models generally calculated total annoyance more adequately than psychophysical models under laboratory conditions (*e.g.*[17,26]).

Considering individual total annoyance ratings (*cf.* Table X), predicted total annoyance from independent effect model based on N was not significantly correlated with field measured total annoyance, whereas the other psychophysical models were significantly correlated, but with small correlation coefficients ($r \leq 0.20$). On the other hand, perceptual total annoyance models using predicted individual partial annoyance as variables (*cf.* Table X) led to better correlation with field measured total annoyance than psychophysical models. This is in agreement with literature findings (*e.g.* [17]) when predicted and field measured annoyance responses are compared. Furthermore, results show that the methodology proposed to estimate psychoacoustic indices as a function of L_{den}, the partial and total annoyance models developed under laboratory conditions and considering noise sensitivity enabled a better field prediction of individual total noise annoyance than psychophysical models based only on L_{den} (*cf.* Table X). This result shows that it is still necessary to improve the prediction of individual annoyance ratings.

V. CONCLUSION

In this study, models for specific, partial and total annoyance due to *URTN* and/or to *AN* were constructed using laboratory acoustical and individual data, thanks to multilevel regression, an appropriate statistical analysis to such data structure. Then, a methodology is proposed to estimate the psychoacoustic indices involved in annoyance models, using L_{den} from noise maps. This methodology enables therefore to test the proposed annoyance models, using survey data. This methodology and the proposed annoyance models, for specific and total annoyance, enable a better prediction than L_{den} only.

In order to predict annoyance on wide areas, new surveys measuring noise sensitivity and noise exposure need to be performed. These new surveys will be useful to estimate noise sensitivity for unknown samples of the population and for testing the methodology to estimate noise indices.

Acknowledgements

This work was performed within the framework of the Labex CeLyA of Université de Lyon, operated by the French National Research Agency (ANR-10-LABX-0060/ANR-11-IDEX-0007).

This work used data from a survey funded by the French Ministry of Ecology (convention n^o 2100966391). The authors thank BruitParif for providing the field noise recordings in the survey area.

REFERENCES

[1]H. M. E. Miedema and C. G. M. Oudshoorn, "Annoyance from transportation noise: relationships with exposure metrics dnl and denl and their confidence intervals," Environmental Health Perspectives **109**(4), 409–416 (2001).

[2]L.-A. Gille, C. Marquis-Favre, and J. Morel, "Testing of the european union exposure-response relationships and annoyance equivalents model for annoyance due to transportation noises: The need of revised exposure-response relationships and annoyance equivalents model," Environment International **94**, 83–94 (2016).

[3]E. Commission *Commission Directive 2015/996 of 19 May 2015 establishing common noise assessment methods according to Directive 2002/ 49/ EC of the European Parliament and of the Council* (2015), p. 823.

[4]T. Kaczmarek and A. Preis, "Annoyance of time-varying road-traffic noise," Archives of Acoustics **35**(3), 383–393 (2010).

[5]A. Trollé, C. Marquis-Favre, and A. Klein, "Short-term annoyance due to tramway noise: determination of an acoustical indicator of annoyance via multilevel regression analysis," Acta Acustica united with Acustica **100**, 34–45 (2014).

[6]A. Trollé, C. Marquis-Favre, and A. Klein, "Corrigendum to: Short-term annoyance due to tramway noise: determination of an acoustical indicator of annoyance via multilevel regression analysis," Acta Acustica united with Acustica **101**, 205 (2015).

[7]A. Klein, C. Marquis-Favre, R. Weber, and A. Trollé, "Spectral and modulation indices for annoyance-relevant features of single urban road vehicle pass-by noises," The Journal

of the Acoustical Society of America **137**(3), 1238–1250 (2015).

[8]P. Meunier and C. Bouillon *Rapport d'information n° 3592 sur les nuisances sonores (Informative report n° 3592 on noise pollution)* (Assemblée Nationale, France, 2011).

[9]J. Morel, C. Marquis-Favre, and L.-A. Gille, "Noise annoyance assessment of various urban road vehicle pass-by noises in isolation and combined with industrial noise: A laboratory study," Applied Acoustics **101**, 47–57 (2016).

[10]T. L. Nguyen, T. Yano, S. Yokoshima, and T. Morihara, "Structural equation model of road traffic noise annoyance in vietnam," in *Inter-noise* (2014).

[11]H. M. E. Miedema and H. Vos, "Noise sensitivity and reactions to noise and other environmental conditions," The Journal of the Acoustical Society of America **113**(3), 1492–1504 (2003).

[12]L.-A. Gille, C. Marquis-Favre, and R. Weber, "Aircraft noise annoyance modeling: Consideration of noise sensitivity and of different annoying acoustical characteristics," Applied Acoustics **115**, 139–149 (2017).

[13]L.-A. Gille, C. Marquis-Favre, and R. Weber, "Noise sensitivity and loudness derivative index for urban road traffic noise annoyance computation," The Journal of the Acoustical Society of America **140**(6), 4307–4317 (2016).

[14]A. Trollé, C. Marquis-Favre, and J. Faure, "An analysis of the effects of structural parameter variations on the auditory perception of environmental noises transmitted through a simulated window," Applied Acoustics **69**, 1212–1223 (2008).

[15]S. Bech and N. Zacharov, *Perceptual audio evaluation - Theory, method and application* (Wiley edition, 2006), 462 p.

[16]B. Berglund and M. E. Nilsson, "Empirical issues concerning annoyance models for combined community noise," in *Inter-noise* (1997), pp. 1053–1058.

[17]C. Marquis-Favre and J. Morel, "A simulated environment experiment on annoyance due to combined road traffic and industrial noises," International Journal of Environmental Research and Public Health **12**(7), 8413–8433 (2015).

[18]L.-A. Gille, C. Marquis-Favre, and K.-C. Lam, "Partial and total annoyance due to road traffic noise combined with aircraft or railway noise: structural equation analysis," International Journal of Environmental Research and Public Health **14**, 12 p (2017).

[19]ISO/TS 15666:2003 *Acoustics – Assessment of noise annoyance by means of social and socio-acoustic survey* (2003).

[20]J. J. Hox, *Multilevel analysis. Techniques and applications (second edition)* (Routledge edition, 2010).

[21]The contribution of each variable to the model is calculated by dividing the corresponding standardized coefficient by the sum of all standardized coefficients.

[22]The recordings were provided by Bruitparif, under the supervision of C. Ribeiro, within the framework of the project related to the survey, funded by the French Ministry of Ecology (MEDDE, convention n°2100966391), and in which ENTPE was a partner (*cf.*[2]). The noises were recorded using an omnidirectional microphone.

[23]S. Stevens, "The measurement of loudness," The Journal of the Acoustical Society of America **27**(5), 815–829 (1955).

[24]T. L. Nguyen, H. Q. Nguyen, T. Yano, T. S. T. Nishimura, T. Morihara, and Y. Hashimoto, "Comparison of models to predict annoyance from combined noise in ho chi minh city and hanoi," Applied Acoustics **73**(9), 952–959 (2012).

[25]H. M. E. Miedema, "Relationship between exposure to multiple noise sources and noise annoyance," The Journal of the Acoustical Society of America **116**(2), 949–957 (2004).

[26]A. Klein, C. Marquis-Favre, and P. Champelovier, "Assessment of annoyance due to urban road traffic noise combined with tramway noise," The Journal of the Acoustical Society of America **141**(1), 231–242 (2017).

[27]C. Marquis-Favre, E. Premat, and D. Aubrée, "Noise and its effect - a review on qualitative aspect of sound. part ii: Noise and annoyance," Acta Acustica United With Acustica **91**, 626–642 (2005).

[28]B. Barbot, C. Lavandier, and P. Cheminée, "Perceptual representation of aircraft sounds," Applied acoustics **69**, 1003–1016 (2008).

[29]M. Paviotti and K. Vogiatzis, "On the outdoor annoyance from scooter and motorbike noise in the urban environment," Science of the Total Environment **430**, 223–230 (2012).

[30]I. van Kamp, R. F. S. Job, J. Hatfield, M. Haines, R. K. Stellato., and S. A. Stansfeld, "The role of noise sensitivity in the noise-response relation: A comparison of three international airport studies," The Journal of the Acoustical Society of America **116**(6), 3471–3479 (2004).

[31]E. Ohrström, A. Skanberg, H. Svensson, and A. Gidlof-Gunnarsson, "Effects of road traffic noise and the benefit of access to quietness," Journal of Sound and Vibration **295**, 40–59 (2006).

[32]P. Lercher, B. D. Coensel, L. Dekonink, and D. Botteldooren, "Community response to multiple sound sources: Integrating acoustic and contextual approaches in the analysis," International Journal of Environmental Research and Public Health **14**(6), 663 (2017).

[33]S. M. Taylor, "A comparison of models to predict annoyance reations to noise from mixed sources," Journal of Sound and Vibration **81**(2), 123–138 (1982).