Virjonen, P., Hongisto, V., Mäkelä, M. M. & Pahikkala, T. (2020). Optimized reference spectrum for rating the façade sound insulation. The Journal of the Acoustical Society of America, 148 (5), 3107-3116. doi: 10.1121/10.0002452

JASA ARTICLE

Optimized reference spectrum for rating the façade sound insulation

Petra Virjonen,[1,a)] Valtteri Hongisto,[2] Marko M. Mäkelä,[3] and Tapio Pahikkala[4,b)]
[1]*Finnish Institute of Occupational Health, Turku, FI-20520, Finland*
[2]*Turku University of Applied Sciences, Turku, FI-20520, Finland*
[3]*Department of Mathematics and Statistics, University of Turku, Turku, FI-20014, Finland*
[4]*Department of Future Technologies, University of Turku, Turku, FI-20014, Finland*

ABSTRACT:
Objectively determined single-number-quantities (SNQs) describing the airborne sound insulation of a façade should correspond to the subjective perception of annoyance to road traffic sounds transmitted through a façade. The reference spectra for spectrum adaptation terms C and C_{tr} in standard ISO 717-7 (International Organization for Standardization, 2013) are not based on psycho-acoustic evidence. The aim of this study is to develop reference spectra which result in SNQs that explain the subjective annoyance of road traffic sounds transmitted through a façade well. Data from a psycho-acoustic experiment by Hongisto, Oliva, and Rekola [J. Acoust. Soc. Am. **144**(2), 1100–1112 (2018)] were used. The data included annoyance ratings for road traffic sounds (five different spectrum alternatives) attenuated by the façade (twelve different sound insulation spectrum alternatives), rated by 43 participants. The reference spectrum for each road traffic spectrum was found using mathematical optimization. The performance of the acquired SNQs was estimated with nested cross-validation. The SNQs determined with the optimized reference spectra performed better than the existing SNQs for two road traffic spectra out of five and for an aggregate of the five road traffic sound types. The results can be exploited in the development of standardized SNQs.
https://doi.org/10.1121/10.0002452

(Received 2 June 2020; revised 11 September 2020; accepted 13 October 2020; published online 25 November 2020)
[Editor: S. K. Tang] Pages: 3107–3116

I. INTRODUCTION

Every day, road traffic noise affects many people: in Europe, more than 100 million people are exposed to adverse road traffic noise levels which are associated with health effects (European Environment Agency, 2017). Exposure to environmental noise levels exceeding certain limit values has been found to cause annoyance, sleep disturbance, tinnitus, cognitive impairment, and an increased risk of cardiovascular diseases (World Health Organization, 2011). The incidence of the effects depends on the noise exposure levels. Brink *et al.* (2019) found that the percentage of highly annoyed persons due to road traffic noise indoors increased from 3% to 46% as L_{den} (day-evening-night level) outdoors increased from 30–35 to 75–80 dB.

To ensure healthy living and working environments, the maximum indoor and outdoor sound levels are guided in many countries with legislation. For example, WHO recommends that the A-weighted equivalent sound pressure level L_{Aeq} should be below 30 dB for bedrooms during nighttime (World Health Organization, 1999). However, each country follows their own regulations. Buildings should be designed in such a way that low indoor sound levels can be attained. This requires adequate sound insulation of the façade.

The sound reduction index (SRI) of a façade can be measured using standardized measurement procedures in existing buildings using the ISO 16283-3 (2016) standard. In laboratory conditions, the SRI of a single façade element can be determined using the ISO 10140-2 (2010) standard. The measurements are carried out in one-third octave bands. Single-number-quantities (SNQs) reduce the one-third octave band data from the SRI measurements to a single number. They enable easier comparison between different constructions and facilitate the imposition of building regulations. Standard ISO 717-1 (2013) determines the calculation of SNQs for airborne sound insulation in buildings and building elements, such as the weighted sound reduction index R_{w}. It is based on comparing the measured SRIs to a standardized reference curve and by determining the sum of so-called unfavorable deviations (the measured value is lower than the value of the curve).

ISO 717-1 enables different frequency ranges for the calculation: the normal frequency range 100–3150 Hz and three enlarged frequency ranges 50–3150 Hz, 50–5000 Hz, and 100–5000 Hz. A reliable determination of the SRIs at low frequencies requires a special measurement procedure as SRI depends strongly on the measurement position being lower at the corners than in the middle of the room (Keränen *et al.*, 2019).

[a)] Also at: Department of Future Technologies, University of Turku, Turku, FI-20014, Finland. Electronic mail: pekavir@utu.fi, ORCID: 0000-0002-9731-7367.
[b)] ORCID: 0000-0003-4183-2455.

ISO 717-1 also includes two spectrum adaptation terms, C and C_{tr}, to take into account different spectra of environmental and living noises. The spectrum adaptation term is added to R_{w}. The reference spectrum for C is A-weighted pink noise and it is meant for living noise (living activities, children playing) as well as for certain kinds of traffic noises (railway traffic at medium and high speed, highway road traffic at higher speeds than 80 km/h, jet aircraft at short distances) and factory emission noise (medium- and high-frequency noise emissions). The reference spectrum for C_{tr} is A-weighted urban traffic noise, and it is meant for noise sources such as traffic noise (urban road traffic, railway traffic at low speeds, propeller driven aircraft), disco music, and factory emission noise (low and medium frequency noise emissions). The calculation of the SNQs with a spectrum adaptation term is based on determining the A-weighted level difference of the source sound pressure levels and the receiver sound pressure levels (source sound pressure level subtracted by the SRI of the façade).

It should be noted that the spectra applied in ISO 717-1 are political choices at a certain stage of their development in 1996. According to Rindel (2017), the method presented in the revised ISO 717-1 in 1996 was a combination of two methods used in Germany and France. In the harmonization process, the adaptation spectrum C_{tr} was adopted from the Nordtest Method NT ACOU 061 (Nordtest, 1987). The spectrum had been composed of two physical measurement sets and was not especially aimed for their current use. No psycho-acoustic experimental evidence was used in the derivation process either. Scientific work is needed to further develop SNQs which explain the annoyance of road traffic noise transmitting to dwellings. Objectively determined SNQs should explain the subjective perception of annoyance and rank different façades according to their subjective order of acoustic performance. In other words, if road traffic noise is experienced more annoying through façade A than façade B, then the SNQ value should be lower for façade A than for façade B. The performance of different SNQs has been studied with psycho-acoustic experiments only in a few studies, despite the fact that sound insulation of façades is globally dimensioned using those SNQs (Bailhache *et al.*, 2014; Hongisto *et al.*, 2018).

Hongisto *et al.* (2018) studied how 25 different SNQs explained the subjective annoyance (43 participants) of road traffic sound (five spectrally different alternatives) transmitted through a façade (twelve spectrally different alternatives). The composition of the road traffic sounds was S1: Light vehicles, urban street, 50 km/h, S2: Light vehicles, motorway, 80 km/h, S3: Light vehicles, motorway, 100 km/h, S4: Only heavy vehicles, urban street, 60 km/h, and S5: Both heavy and light vehicles, urban street, 60 km/h. The sound spectra of the road traffic sounds on the outer surface of the façade are presented in Fig. 1. The spectrum of S5 corresponded to ISO 717-1 (2013) spectrum for calculation of C_{tr}. Also, the scaled ISO 717-1 spectrum for C is shown in Fig. 1. Hongisto *et al.* (2018) concluded that a well performing SNQ depends on the spectrum of the road traffic sound. $R_{\mathrm{w}} + C_{50\text{–}3150}$ was found sufficient for most road traffic types. Bailhache *et al.* (2014) studied how twelve different SNQs explained the subjective ratings (24 participants) of exterior noise (seven alternatives) transmitted through a façade (ten alternatives). The exterior noise types were: pass-by of a plane, traffic in a busy street, construction works, church bell ringing, loud voices, pass-by of a scooter, and pass-by of an ambulance. In the second part of their study, they found that for road traffic sound type ("traffic in a busy street"), $R_{\mathrm{w}} + C_{100\text{–}3150}$ performed the best among those SNQs studied. Torija *et al.* (2011) studied the relationship between traffic noise annoyance and indoor sound levels. The participants (100) rated the annoyance to road (highway and local road) and railway noises transmitted through a façade. They found a reduced number (16 out of 27) of one-third octave bands to be relevant for annoyance of traffic noise. However, they studied only one façade type. Myllyntausta *et al.* (2020) studied how the road traffic noise transmitted through two façades having different SRI spectrum affected sleep. They found suggestive evidence that $R_{\mathrm{w}} + C_{50\text{–}3150}$ would better explain sleep disturbance than $R_{\mathrm{w}} + C_{\mathrm{tr}}$. However, they studied only two façades and one road traffic spectrum. There is a need for studying the best reference spectrum for road traffic noise based on psycho-acoustic evidence as well as to study the suitable frequency range. The analysis should be conducted with different types of noise spectra.

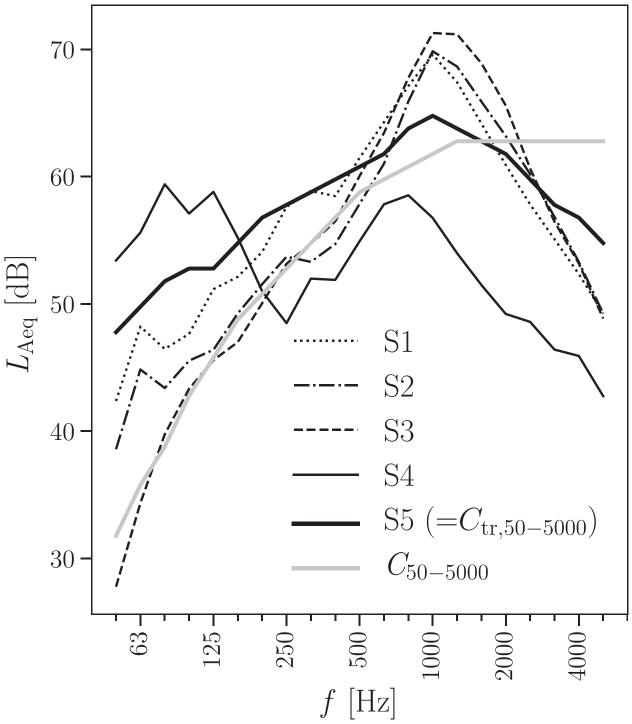

FIG. 1. A-weighted equivalent sound pressure level, L_{Aeq}, for one-third octave band frequencies f from 50 to 5000 Hz for all sound types S1,…,S5 on the outer surface of the façade. Sound type S5 spectrum conformed with $C_{tr,50\text{–}5000}$. Also scaled $C_{50\text{–}5000}$ is shown.

Mathematical optimization has been used twice to derive adequate reference spectra, first for airborne sound

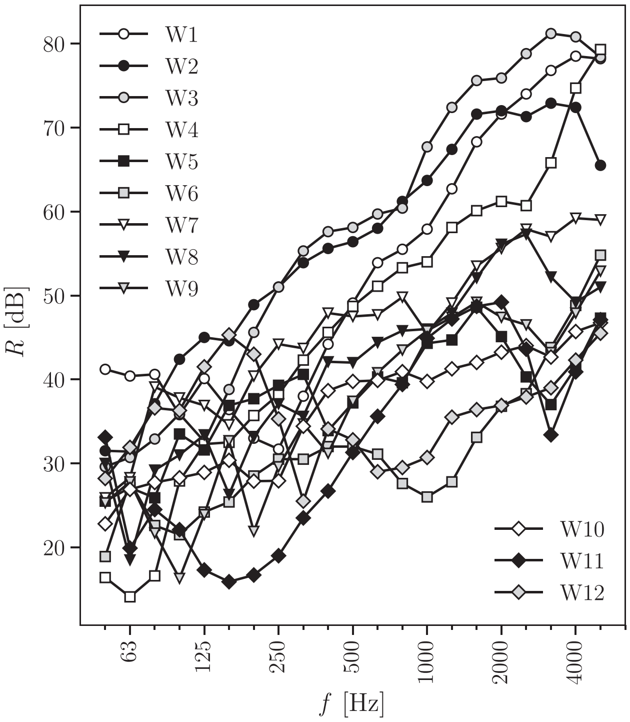

FIG. 2. Sound reduction indices R for one-third octave band frequencies f from 50 to 5000 Hz for the façade elements W1,…,W12.

insulation of partitions by Virjonen *et al.* (2016), and then for impact sound insulation of floors by Kylliäinen *et al.* (2019). The reference spectra were derived by constructing a nonlinear optimization problem with constraints and solving it with the sequential least squares programming method. The optimization method has not been previously used to determine the adequate reference spectra for the airborne sound insulation of façades.

The purpose of the study was to develop reference spectra which lead to SNQs that explain the subjective annoyance towards road traffic sounds transmitted through a façade well. Another purpose was to find the relevant frequency range employed to reach the best conformance between the subjective annoyance and the resulting SNQ. It also attempts to answer the question: is there a need for several spectrum adaptation terms or could some general solution be found, which would perform well for all kinds of road traffic spectra from heavy-weight vehicles driving in urban speeds to lightweight vehicles driving in highways?

II. MATERIALS AND METHODS

A. Experimental data

Experimental data from a previously published psycho-acoustic experiment were used (Hongisto *et al.*, 2018). Forty-three volunteers (28 women, 15 men, age between 21 and 50 years) participated in the experiment. They rated different road traffic sounds, one participant at a time, in a furnished experimental room, built for psycho-acoustic experiments. The experimental sounds were played through two active loudspeakers at 1.5 m height, and one subwoofer on the floor. The background noise level of the room was 20 dB L_{Aeq} between 50 and 5000 Hz.

The experiment contained 60 sounds. They were prepared from outdoor recordings including periods of both steady-state and intermittent road traffic. Five different sound types having different traffic content and traffic speeds were used (shown in Fig. 1). The outdoor sound samples were filtered according to the SRI of the façade elements. Twelve spectrally different alternatives were used, and their SRIs based on laboratory tests according to ISO 140-3 (1995) are presented in Fig. 2. Various SNQ values, determined from the SRIs, are presented in Table I for the façade elements. The levels outside the façade were adjusted between 68 and 77 dB L_{Aeq}. The resulting listening levels of the experimental sounds were thus audible, as well as realistic for residential dwellings (12–46 dB L_{Aeq}).

The participants rated the sounds with respect of loudness and annoyance. Annoyance ratings were used in the present study to determine the optimal reference spectra for the tested road traffic sound types, because annoyance is the most usual health impact of noise. The participants rated the annoyance by answering the question “How annoying is the sound?” using an 11-step response scale from 0 to 10. The extremes were verbally labeled by 0: “Not at all annoying” and 10: “Extremely annoying.” The participants were also given the option via a checkbox to indicate if they could not

TABLE I. The values of the existing SNQs [dB] for the façade elements W1–W12. R_{w} and its spectrum adaptation variations were determined according to ISO 717-1 (2013). *OITC* was determined according to ASTM E1332-10a (2010). $EA_{50\text{–}5000}$ was determined according to Park *et al.* (2008).

SNQ	W1	W2	W3	W4	W5	W6	W7	W8	W9	W10	W11	W12
R_{w}	49.8	61.5	59.3	50.5	41.4	32.5	50.5	45.7	40.3	40.7	32.7	34.2
$R_{\text{w}}+C_{50\text{–}3150}$	47.0	58.5	54.4	44.2	40.1	30.5	48.5	43.4	37.5	39.2	30.9	33.2
$R_{\text{w}}+C_{50\text{–}5000}$	48.0	59.3	55.4	45.2	40.6	31.5	49.4	44.1	38.4	40.0	31.8	34.1
$R_{\text{w}}+C_{100\text{–}3150}$	47.0	59.7	55.2	48.3	40.2	30.6	48.7	43.8	37.6	39.3	31.0	33.2
$R_{\text{w}}+C_{100\text{–}5000}$	48.0	60.5	56.1	49.2	40.7	31.5	49.6	44.5	38.5	40.0	31.9	34.1
$R_{\text{w}}+C_{tr,50\text{–}3150}$	42.8	50.0	46.4	33.1	38.5	28.7	43.8	37.4	32.0	35.9	26.9	31.9
$R_{\text{w}}+C_{tr,50\text{–}5000}$	43.0	50.1	46.5	33.3	38.6	28.8	44.0	37.6	32.2	36.0	27.1	32.1
$R_{\text{w}}+C_{tr,100\text{–}3150}$	42.9	55.5	49.3	43.0	39.8	28.9	45.9	40.0	32.4	36.6	27.0	31.9
$R_{\text{w}}+C_{tr,100\text{–}5000}$	43.1	55.7	49.5	43.2	39.9	29.1	46.0	40.2	32.6	36.7	27.2	32.1
$EA_{50\text{–}5000}$	39.7	40.7	37.9	23.8	33.1	26.6	36.1	30.0	26.4	30.9	22.9	32.8
OITC	41.2	51.7	44.9	34.9	37.8	29.1	44.4	37.7	29.7	34.9	24.8	32.4

hear the sound at all. Only 0.7% of the ratings were marked as inaudible. The distribution of the annoyance ratings for all façade elements and road traffic sound types are presented in Fig. 3.

B. Formulation of the optimization problem

The reference spectrum optimization procedure was introduced by Virjonen *et al.* (2016). They optimized the reference spectrum for SNQ rating airborne sound insulation

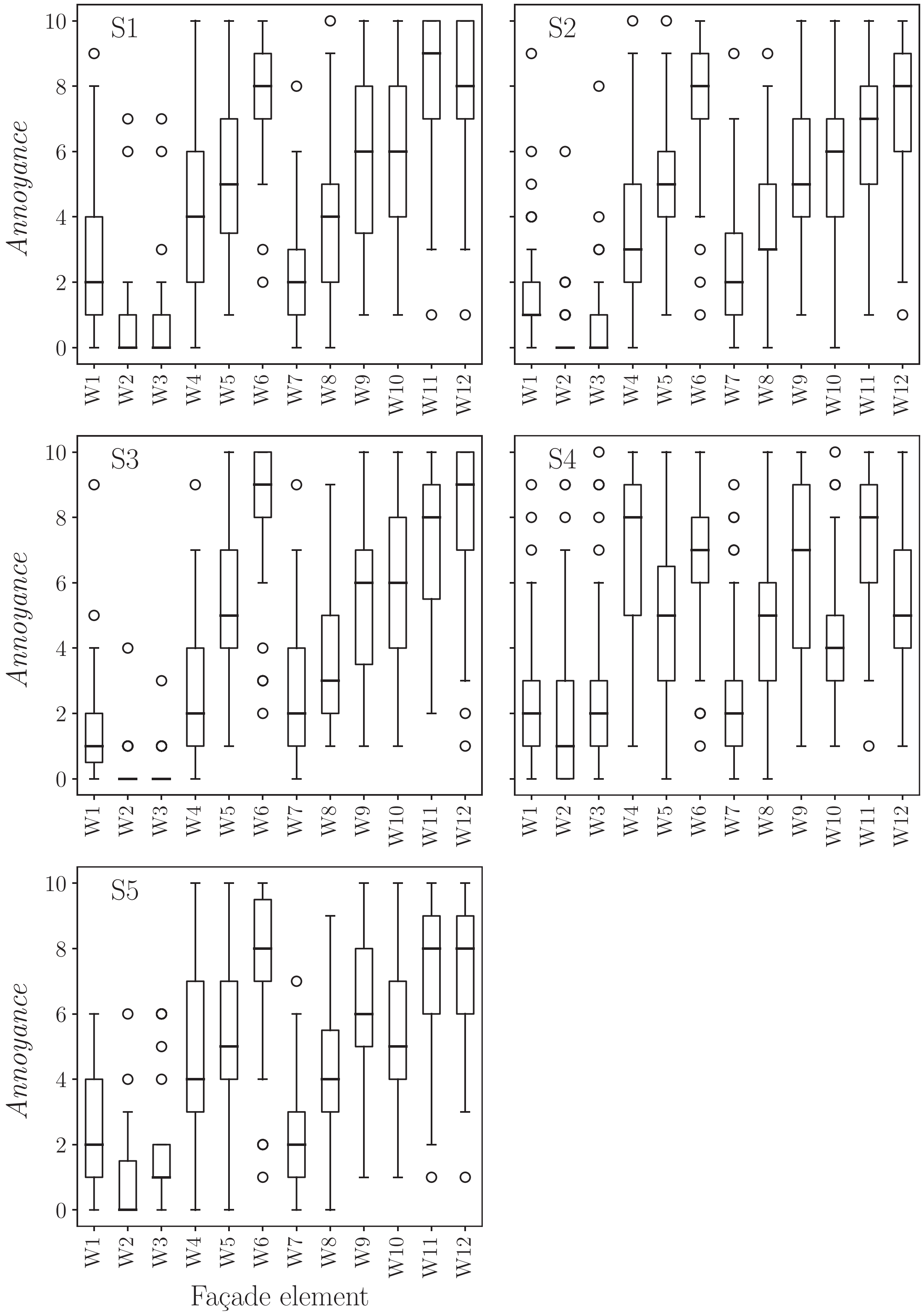

FIG. 3. Distribution of the annoyance ratings for each façade element with each road traffic sound type S1,…,S5. The horizontal line within the box presents the median Q2. The box extends from the lower quartile Q1 to upper quartile Q3 values of the annoyance ratings. The lower bound of the whiskers is the first datum greater than $Q1 - 1.5\cdot(Q3 - Q1)$. The upper bound of the whiskers is the last datum smaller than $Q3 + 1.5\cdot(Q3 - Q1)$. Outliers (outside the whiskers) are marked with circles.

for living sounds. The same procedure was also exploited in Kylliäinen *et al.* (2019). They optimized the reference spectrum for SNQ rating impact sound insulation for several natural impact sounds. The same optimization procedure deployed in the abovementioned studies was also used in the present study.

The optimal reference spectrum was calculated for each sound type S1,…,S5, separately. The resulting optimal reference spectra were named L_{S1},…,L_{S5}. To find a reference spectrum performing well for road traffic noise in general, the ratings from all sound types were averaged, and a reference spectrum was sought. The resulting optimal aggregate reference spectrum was named L_{S1-5}. For each sound type, the goal was to find such a reference spectrum L that a linear fit between the mean subjective ratings and the resulting SNQ values was optimal. A SNQ can be calculated from (ISO 717-1, 2013)

$$x_i = 10\lg \frac{\sum_{j=K_1}^{K_2} 10^{L_j/10}}{\sum_{j=K_1}^{K_2} 10^{(L_j\ \ R_{ij})/10}}. \quad (1)$$

Here, K_1 and K_2 determine the included one-third octave frequency bands, L_j is the level of the reference spectrum at frequency band j, and R_{ij} is the SRI for the façade element i at frequency band j. The reference spectrum was normalized to 0 dB, i.e.,

$$10\lg \sum_{j=K_1}^{K_2} 10^{L_j/10} = 0\,\mathrm{dB}. \quad (2)$$

To obtain a smoother solution, the maximum level difference between adjacent frequency bands δ was limited to 3 dB. The best frequency band range was selected within four options: 50–3150, 50–5000, 100–3150, and 100–5000 Hz. The frequency band range was selected using leave-one-out cross-validation (LOOCV) (Varma and Simon, 2006). The formulation of the optimization problem as well as the cross-validation scheme are explained in detail in the supplementary materials.[1]

C. Solution to the optimization problem

The optimal reference spectra were solved using an algorithm for finding the minimum of a constrained nonlinear multivariable function. The solution process of the optimization problem is explained in detail in supplementary materials. The reference spectrum for urban traffic noise to calculate C_{tr} from ISO 717-1 (2013) was chosen as the initial guess from which the algorithm started to proceed. The calculations were also made with two other initial guesses (Fig. 4) to test the convergence of the algorithm. Practically the same reference spectra were attained with the three initial guesses.

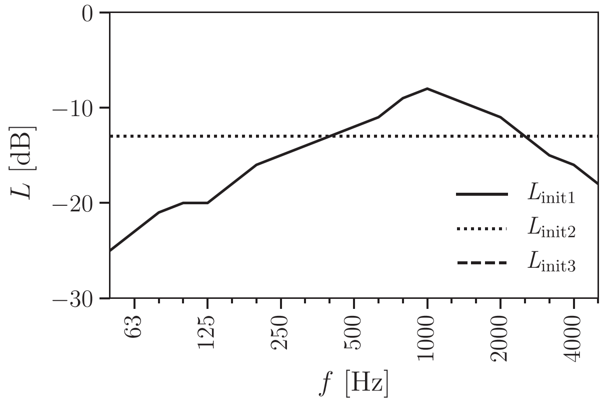

FIG. 4. Three tested initial spectra, from which the optimization algorithm started to proceed. The levels L are shown at one-third octave band frequencies f from 50 to 5000 Hz. L_{init1} is the reference spectrum for C_{tr} in ISO 717-1 (2013).

D. Uncertainty of the reference spectrum

To estimate the uncertainty of the optimized reference spectrum, bootstrap sampling (Chernik, 2008) was exploited. In bootstrapping, sampling is made with replacement, thus each datum can appear in the sample more than once. A sample from the participants ($n_{participants} = 43$) was drawn, and the optimal reference spectrum was determined using this bootstrap sample. For each frequency band, the difference between the optimized reference spectrum level acquired with the bootstrap sample, and the original sample, was calculated. The procedure was repeated 1500 times. For each frequency band, from the 1500 differences, the 2.5% and 97.5% quantiles were determined. This gave an estimation of the empirical 95% confidence intervals.

E. Estimation of the model performance

As the data set is rather small, it is beneficial to deploy it as a whole when finding the best model. However, this leaves no data for testing the performance of the selected model. How well would the ratings given by people outside the group of the participants (of similar distribution, e.g., of ages and genres) fit with the SNQs acquired with the optimized reference spectrum? To answer this, nested leave-one-out cross-validation (nested LOOCV) was used to estimate the model performance for all optimized reference spectra. Nested cross-validation (nested CV) (Varma and Simon, 2006) gives an estimation of how a model performs with data, that has not been a part of the model selection process ("model" in this case means the total process: optimizing the reference spectrum with frequency range selection). If the same data were used for the training of a model, as well as to estimate the performance of the model, this would result in an over-optimistic estimation. To overcome this, in nested cross-validation, the parameters (here: the frequency range) of the model are selected within the inner CV loop. The selected model is then tested in the outer CV loop. Different optimal frequency range may be found in different rounds of the CV. The variation of the optimal parameters in the nested CV also gives information on the stability of the selected model. The squared Pearson's correlation coefficient was used as the estimation parameter r^2. The Wilcoxon signed ranks test was used to test whether the estimation parameters for the SNQs acquired with the optimized reference spectra differed from the values obtained for the standardized SNQs.

III. RESULTS

For each road traffic sound type, the reference spectrum L in Eq. (1) was optimized, and the most relevant frequency range was selected using leave-one-out cross-validation.

The mean annoyance over all participants versus the resulting optimized SNQ values are shown in Fig. 5. The mean annoyance versus the best performing existing SNQs are also shown for each sound type.

A valid solution was found for each sound type (the stopping criterion was met before the maximum number of iterations, see supplementary materials for details), and the algorithm ended up in the same minimum with three different initial spectra.

The optimized reference spectra $L_{\mathrm{S1}},\ldots,L_{\mathrm{S5}}$ for road traffic spectra S1,…,S5, and the optimized aggregate reference spectrum $L_{\mathrm{S1\text{–}5}}$ together with their empirical confidence intervals are presented in Fig. 6.

The estimation of the predicting performance of the optimized spectra, and standardized SNQs for each sound type are presented in Table II. Also, one non-standardized SNQ is included, as it performed well in the study of Hongisto *et al.* (2018), namely, the energy average $EA_{50\text{–}5000}$ by Park *et al.* (2008). The best performing existing SNQs are indicated by footnote a for each sound type. If the difference between the SNQ acquired with the optimized reference spectrum and the best performing existing SNQ was statistically significant ($p < 0.05$), the SNQ acquired with the optimized reference spectrum is marked with footnote b. Table III shows the frequency ranges, which were most often selected as the best in nested LOOCV for each road traffic sound type.

IV. DISCUSSION

A. Main results

Figure 5 shows the squared Pearson's correlation coefficients for the best performing existing SNQs for each sound type. The correlations were already very high, when considering the average annoyance over all participants. The optimized reference spectrum resulted in slightly higher correlations for each sound type. Figure 5 shows the result of optimization of the reference spectra with all the available data. To estimate the model performance with data, which has not been a part of the model selection process, nested LOOCV was used. The ratings of one participant were left as test data, one participant in turn, and the rest were used to derive the model. This resulted in 43 model performance estimations. Table II shows how the SNQs acquired with the optimized reference spectra performed on average when each model was compared with the ratings given by the test participant left outside the model. Again, with the SNQs acquired with the optimized reference spectra, the squared Pearson's correlation coefficients were slightly improved for each road traffic sound type when compared with the existing SNQs. The differences for the estimation parameters between the optimized and existing SNQs were statistically significant ($p < 0.05$) for the sound types S2, S3 and the aggregate sound type S1–5.

Sound type S1 included only light vehicles on an urban street with 50 km/h speed. ISO 717-1 suggests C_{tr} for such purpose, however, the optimized reference spectrum for

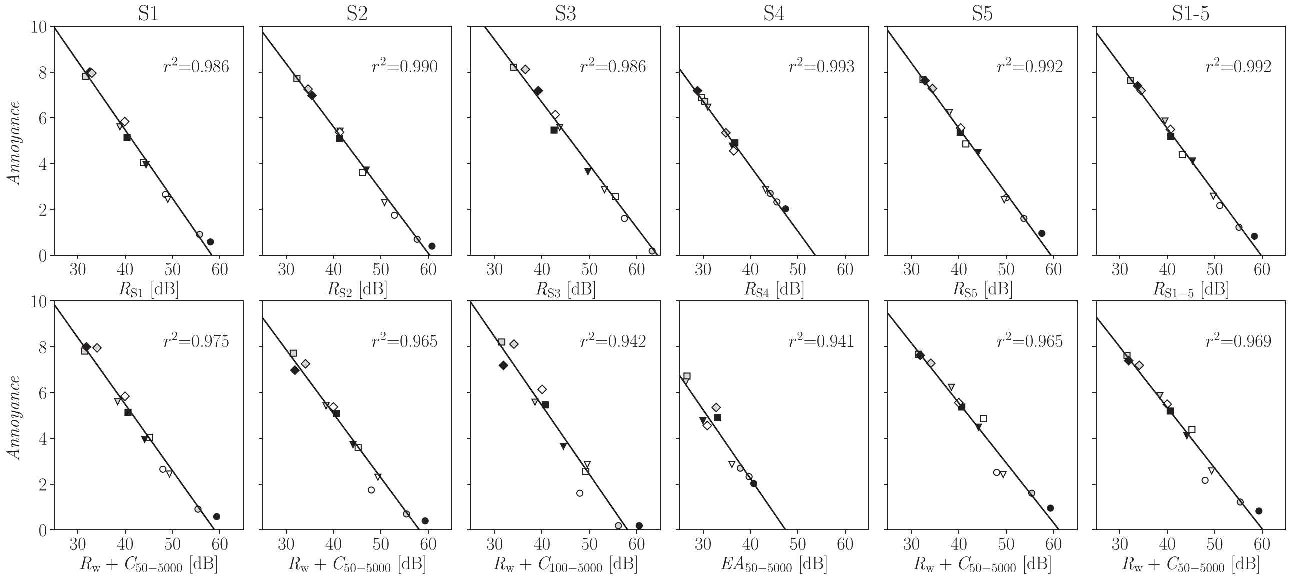

FIG. 5. Above: Mean annoyance versus the SNQs acquired with the optimized SNQs (optimized with the whole data). Below: Mean annoyance versus best performing existing SNQs. Squared Pearson's correlation coefficient for each linear fit is also shown. The different markers identify the façades W1–W12, see legend in Fig. 2.

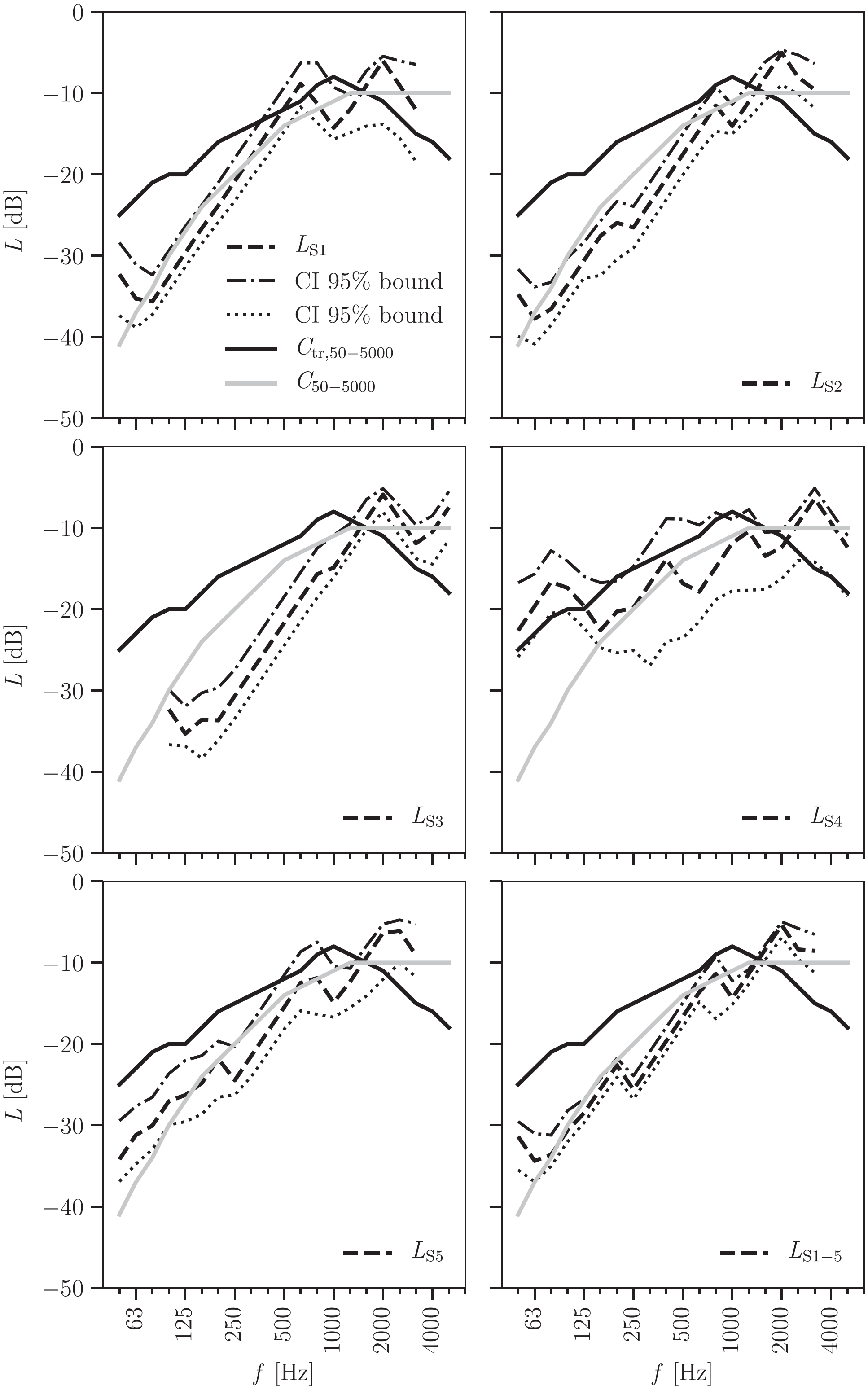

FIG. 6. Optimized reference spectrum L for each road traffic sound type S1,…,S5. L_{S1-5} is the aggregate reference spectrum acquired with the average annoyance rating from all road traffic sound types S1,…,S5. Reference spectra for C_{tr} and C by ISO 717-1 (2013), and empirical 95% confidence intervals are also shown. S1: Light vehicles, urban street, 50 km/h, S2: Light vehicles, motorway, 80 km/h, S3: Light vehicles, motorway, 100 km/h, S4: Heavy vehicles, urban street, 60 km/h, S5: Both heavy and light vehicles, urban street, 60 km/h.

sound type S1 conforms well with the reference spectrum for C (Fig. 6). The squared Pearson's correlation coefficients were the same up to two decimal places between R_{S1} and $R_w + C_{50–3150}$, and $R_w + C_{50–5000}$.

Sound types S2 and S3 included also only light vehicles on a motorway with 80 and 100 km/h speeds, respectively. ISO 717-1 suggests C for highway road traffic noise with speeds higher than 80 km/h. The spectra for S2 and S3 did conform better with the reference spectrum for C than C_{tr}. They still had lower values than the C spectrum in the middle frequencies, roughly from 125 to 500 Hz, especially for the sound type S3.

Sound type S4 included only heavy vehicles on an urban street with 60 km/h speed. Such roads hardly exist but

TABLE II. Estimation of the performance of the SNQs acquired with the optimized reference spectra and existing SNQs. The value describes the average squared Pearson's correlation coefficient between a participant's ratings and SNQs, for each road traffic sound type S1, S2, S3, S4, S5, and their aggregate S1–5.

Sound type	S1	S2	S3	S4	S5	S1–5
Existing SNQs						
$R_{\mathrm{w}}+C_{50\text{–}3150}$	0.745[a]	0.779	0.787	0.479	0.730	0.870
$R_{\mathrm{w}}+C_{50\text{–}5000}$	0.745[a]	0.780[a]	0.790	0.478	0.731[a]	0.871[a]
$R_{\mathrm{w}}+C_{100\text{–}3150}$	0.728	0.766	0.795[a]	0.429	0.701	0.839
$R_{\mathrm{w}}+C_{100\text{–}5000}$	0.727	0.767	0.797[a]	0.428	0.701	0.840
$R_{\mathrm{w}}+C_{tr,50\text{–}3150}$	0.684	0.699	0.648	0.587	0.714	0.828
$R_{\mathrm{w}}+C_{tr,50\text{–}5000}$	0.685	0.699	0.649	0.587	0.715	0.829
$R_{\mathrm{w}}+C_{tr,100\text{–}3150}$	0.707	0.732	0.736	0.465	0.697	0.820
$R_{\mathrm{w}}+C_{tr,100\text{–}5000}$	0.707	0.733	0.736	0.465	0.698	0.821
$EA_{50\text{–}5000}$	0.444	0.454	0.370	0.601[a]	0.522	0.582
$OITC$	0.663	0.675	0.628	0.552	0.692	0.793
Optimized SNQs						
R_{S1}	0.750					
R_{S2}		0.794[b]				
R_{S3}			0.828[b]			
R_{S4}				0.624		
R_{S5}					0.747	
$R_{S1\text{–}5}$						0.892[b]

[a] The best performing existing SNQs.

[b] The difference between the SNQ acquired with the optimized reference spectrum and the best existing SNQ is statistically significant ($p < 0.05$).

Hongisto *et al.* (2018) found it important to cover all possible spectra that road traffic noise could contain, even during short moments such as the pass-by of a single heavy vehicle. The values of L_{S4} were rather close to the reference spectrum for C_{tr} at low frequencies 50–125 Hz. The confidence intervals were clearly wider for sound type S4 than for the other sound types. The reference spectrum for urban road traffic noise suggested by ISO 717-1, C_{tr}, was well within the confidence intervals for sound type S4. The performance of R_{S4} remained rather low compared with other sound types (Table II).

Sound type S5 included both light and heavy vehicles on an urban street with 60 km/h speed, and its sound level spectrum on the façade surface was adjusted to meet with C_{tr} spectrum. That is, the sound represents the standardized urban road traffic noise of ISO 717-1 standard and deserves special attention. Compared to the reference spectrum for C_{tr}, L_{S5} had clearly lower values for frequency bands lower than 500 Hz. Again, L_{S5} conformed better with the reference spectrum for C.

TABLE III. The best frequency ranges selected for each road traffic sound type. The percentage of rounds for which it was selected as the best is also shown.

Sound type	Frequency range [Hz]	%
S1	50–3150	79
S2	50–3150	72
S3	100–5000	100
S4	50–5000	100
S5	50–3150	100
S1–5	50–3150	79

All in all, the optimized reference spectra L_{S1}, L_{S2}, L_{S3}, and L_{S5} for sound types S1, S2, S3, and S5 including light vehicles were rather similar, and closer to the reference spectrum C than $C_{\mathrm{tr.}}$ The optimized reference spectrum for sound type S4 including only heavy vehicles was closer to C_{tr} at low frequencies, however, this finding has very little meaning. The reason is that the relative share of heavy vehicles is usually under 20%. Although the pass-by sound level of a heavy vehicle is 5 to 10 dB higher than the pass-by sound level of light vehicles, the overall sound level and spectrum shape is dominated by light vehicles. The sound type S4 would be relevant only in roads having low traffic rates where the traffic consists mainly of single pass-bys, and the proportion of heavy vehicles is high. Such situation takes place in some main roads during night-time. In such rare cases the single pass-bys of heavy vehicles mainly explain the annoyance reactions. In most cases, when road traffic noise is an issue, the traffic density is so high that single pass-bys are not distinguishable and the spectrum resembles sound type S5 which is a mixture of light and heavy vehicles. In such cases, the reference spectrum of C was very close to the optimized reference spectra. It seems that spectrum C covers most of the sound types in real environments, and the actual need for C_{tr} may be negligible.

According to Table III, different optimal frequency ranges were found for different road traffic spectra. Sound types S1, S2, and S3 had only light vehicles but different speeds (50, 80, 100 km/h, respectively). For S1 and S2, the optimal frequency range was 50–3150 Hz, and for S3, 100–5000 Hz. For sound type S3, the optimal frequency range started from 100 Hz, which was expected, as the sound levels were very low for the lowest frequency bands. Sound types S4 and S5 were composed of vehicles driving at 60 km/h speed on an urban street but S4 had only heavy vehicles and S5 both light and heavy vehicles. The optimal frequency range for S4 was 50–5000 Hz, and for S5, 50–3150 Hz. The selection of optimal frequency range was rather stable: the same optimal frequency range was selected as the best in clear majority of the rounds of the nested LOOCV.

B. Method

The same optimization scheme as used in the present study, has been deployed for airborne sound insulation (Virjonen *et al.*, 2016) and impact sound insulation (Kylliäinen *et al.*, 2019). In the present study, the method was further developed to select the suitable frequency range for each optimized reference spectrum. Also, the interpretation of the results was improved: nested cross-validation was utilized to evaluate the performance of the selected model. In the previous studies, the optimized reference spectra ended up in a larger improvement of the squared correlation coefficient when compared with existing SNQs than in the present study. This was expected, as the correlations were already rather good with the standardized SNQs. Yet, a statistically better solution was found for two sound types.

The present study confirmed that the spectrum adaptation term C is an adequate descriptor for most road traffic sounds, and there is not much room for improvement, unlike in the cases with airborne and impact sound. The optimization method deployed here is well-suited and recommended for this kind of purposes, where physical parameters are tuned to correspond the subjective experience.

C. Strengths and limitations

The generalization of the reference spectrum depends on the representativeness of the subjective data: the chosen road traffic sound spectra, the façade structures, the playback levels of the test sounds, and the background noise levels. Different choices in producing the subjective data might have led into different reference spectra. However, Hongisto *et al.* (2018) and Bailhache *et al.* (2014) as well as Myllyntausta *et al.* (2020) performed independent studies and obtained similar results regardless of different selection of the abovementioned factors, which suggests that the data used in our study would be sufficiently representative taking into account the fact that current standardized reference spectra are not based on any psycho-acoustic evidence.

All the experimental sounds were played at relevant levels, i.e., the sound level on the outer surface of the façade was set to a realistic level. Also, the full range of experimental sounds, which affects the subjective rating scale, was the same for all sound types. Because of this, it was possible to acquire the optimized aggregate reference spectrum $L_{S1–5}$ by averaging the ratings for all sound types S1,…,S5. However, the relative importance of each road traffic sound was thus equal. As the spectrum varies according to the road traffic sound type, and the proportion of different sound types varies according to the place, there might not be a descriptive general composition of sounds which would fit each situation. The optimal reference spectrum would probably look different if the sound types were weighted in some other way.

Also, the intensity of the short-term variations of noise level over time affect the rated annoyance (Brink *et al.*, 2019). According to Hongisto *et al.* (2018), "the inherent temporal variation of the A-weighted SPL due to pass-by sounds was small but non-existent" for the experimental sounds. Thus, the results apply for the experimental samples used, and with different temporal variation, the results may have been different.

V. CONCLUSIONS

In this study, reference spectra which result in SNQs that explain well the subjective annoyance of road traffic sounds transmitted through façade were developed. The reference spectra were determined using psycho-acoustic experimental data and mathematical optimization. The optimization scheme, previously utilized with airborne sound insulation of partitions (Virjonen *et al.*, 2016) and impact sound insulation of floors (Kylliäinen *et al.*, 2019), was further developed to select the most suitable frequency range and to evaluate the performance of the selected models. The resulting optimized SNQs performed better than the existing standardized SNQs of ISO 717-1 (2013) and ASTM E1332-10a (2010) for two road traffic sound types out of five and for an aggregate of the five road traffic sound types, even though the performance of the existing SNQs was already rather good.

The frequency range of 50–3150 Hz was selected most often as the best frequency range. The selection of the most relevant frequency range was rather stable, the same frequency range was selected in clear majority of the cross-validation rounds for each road traffic sound type.

The results can be exploited in the development of standardized SNQs.

ACKNOWLEDGMENTS

The research concerning the psycho-acoustical experiment was a part of ÄKK project (2012–2014) which was mainly funded by Business Finland (formerly Tekes) (Grant No. 2296/31/2011). The other funders were the Ministry of the Environment, Betoniteollisuus Assoc., Kestävä Kivitalo, Karelia-Upofloor Ltd., Wärtsilä Finland Ltd., Saint-Gobain Rakennustuotteet Ltd., STX Finland Cabins Ltd., Skaala Ltd., University of Turku, Tampere University of Technology, and Finnish Institute of Occupational Health. The analyses and the manuscript writing for the optimization of the reference spectra were funded by MATTI, Doctoral Programme in Mathematics and Computer Sciences by University of Turku.

[1] See supplementary material at https://www.scitation.org/doi/suppl/10.1121/10.0002452 for the formulation and solving the optimization problem as well as the details of the cross-validation scheme.

ASTM E1332-10a (**2010**). *Standard Classification for Rating Outdoor-Indoor Sound Attenuation* (ASTM International, West Conshohocken, PA).

Bailhache, S., Jagla, J., and Guigou-Carter, C. (**2014**). "CSTB - Projet Environnement et ambiances: Effet des basses frequences sur le confort acoustique–tests psychoacoustiques" ("CSTB - Environment and atmosphere project: effect of low frequencies on acoustic comfort - psychoacoustic tests"), technical report, https://hal-cstb.archives-ouvertes.fr/hal-01045056 (Last viewed June 1, 2020).

Brink, M., Schäffer, B., Vienneau, D., Foraster, M., Pieren, R., Eze, I. C., Cajochen, C., Probst-Hensch, N., Röösli, M., and Wunderli, J.-M. (**2019**). "A survey on exposure-response relationships for road, rail, and aircraft noise annoyance: Differences between continuous and intermittent noise," Environ. Int. **125**, 277–290.

Chernik, M. R. (**2008**). *Bootstrap Methods: A Guide for Practitioners and Researchers* (Wiley-Interscience, Hoboken, NJ).

European Environment Agency (**2017**). "Managing exposure to noise in Europe," Publications Office of the European Union.

Hongisto, V., Oliva, D., and Rekola, L. (**2018**). "Subjective and objective rating of the sound insulation of residential building façades against road traffic noise," J. Acoust. Soc. Am. **144**(2), 1100–1112.

ISO 140-3 (**1995**). *Acoustics—Measurement of Sound Insulation in Buildings and of Building Elements—Part 3: Laboratory Measurements of Airborne Sound Insulation of Building Elements* (International Organization for Standardization, Geneva, Switzerland).

ISO 717-1 (**2013**). *Acoustics—Rating of Sound Insulation in Buildings and of Building Elements* (International Organization for Standardization, Geneva, Switzerland).

ISO 10140-2 (**2010**). *Acoustics—Laboratory Measurement of Sound Insulation of Building Elements—Part 2: Measurement of Airborne*

Sound Insulation (International Organization for Standardization, Geneva, Switzerland).

ISO 16283-3 (**2016**). *Acoustics—Field Measurement of Sound Insulation in Buildings and of Building Elements—Part 3: Façade Sound Insulation* (International Organization for Standardization, Geneva, Switzerland).

Keränen, J., Hakala, J., and Hongisto, V. (**2019**). "The sound insulation of façades at frequencies 50–5000 Hz," Build. Environ. **156**, 12–20.

Kylliäinen, M., Virjonen, P., and Hongisto, V. (**2019**). "Optimized reference spectrum for rating the impact sound insulation of concrete floors," J. Acoust. Soc. Am. **145**(1), 407–416.

Myllyntausta, S., Virkkala, J., Salo, P., Varjo, J., Rekola, L., and Hongisto, V. (**2020**). "Effect of the frequency spectrum of road traffic noise on sleep: A polysomnographic study," J. Acoust. Soc. Am. **147**(4), 2139–2149.

Nordtest (**1987**). *Windows: Traffic Noise Reduction Indices* (Nordtest, Espoo, Finland).

Park, H. K., Bradley, J. S., and Gover, B. N. (**2008**). "Evaluating airborne sound in terms of speech intelligibility," J. Acoust. Soc. Am. **123**(3), 1458–1471.

Rindel, J. H. (**2017**). *Sound Insulation in Buildings* (Taylor & Francis Group, Boca Raton), pp. 146–148.

Torija, A. J., Ruiz, D. P., Coensel, B. D., Botteldooren, D., Berglund, B., and Ramos-Ridao, A. (**2011**). "Relationship between road and railway noise annoyance and overall indoor sound exposure," Transport. Res. Part D Trans. Environ. **16**(1), 15–22.

Varma, S., and Simon, R. (**2006**). "Bias in error estimation when using cross-validation for model selection," BMC Bioinformatics **7**, 91.

Virjonen, P., Hongisto, V., and Oliva, D. (**2016**). "Optimized single-number quantity for rating the airborne sound insulation of constructions: Living sounds," J. Acoust. Soc. Am. **140**(6), 4428–4436.

World Health Organization (**1999**). "Guidelines for community noise," edited by B. Berglund, T. Lindvall, and D. H. Schwela, https://apps.who.int/iris/handle/10665/66217 (Last viewed June 1, 2020).

World Health Organization (**2011**). "Burden of disease from environmental noise: Quantification of healthy life years lost in Europe" https://www.who.int/quantifying_ehimpacts/publications/e94888/en/ (Last viewed June 1, 2020).