Acta Acustica united with Acustica
Vol. 100 (2014) 1186 – 1201

DOI 10.3813/AAA.918797

Repeatability and Reproducibility of *In Situ* Measurements of Sound Reflection and Airborne Sound Insulation Index of Noise Barriers

Massimo Garai1), Eric Schoen2), Gottfried Behler3), Beatriz Bragado4), Michael Chudalla5), Marco Conter6), Jérôme Defrance7), Patrick Demizieux8), Christ Glorieux9), Paolo Guidorzi10)

1) DIN, University of Bologna, Viale Risorgimento 2, 40136 Bologna, Italy. massimo.garai@unibo.it
2) TNO, 3700 AJ Zeist, The Netherlands. eric.schoen@tno.nl
3) RWTH Aachen University, Neustraße 50, 52066 Aachen, Germany. gkb@akustik.rwth-aachen.de
4) CIDAUT, Parque Tecnológico de Boecillo, P209,47151 Boecillo (Valladolid), Spain. beabra@cidaut.es
5) BASt, Brüderstrasse 53, D-51427 Bergisch Gladbach, Germany. chudalla@bast.de
6) AIT, Giefinggasse 2,1210 Vienna, Austria. marco.conter@ait.ac.at
7) CSTB, 24 rue Joseph Fourier, 38400 Saint-Martin-d'Hères, France. jerome.defrance@cstb.fr
8) LRPC Strasbourg, 11, rue Jean Mentelin - BP 9, 67035 Strasbourg cedex 2, France.
patrick.demizieux@developpement-durable.gouv.fr
9) K.U. Leuven, 24 Celestijnenlaan 200D, bus:2416, 3001 Leuven, Belgium. christ.glorieux@fys.kuleuven.be
10) DIN, University of Bologna, Viale Risorgimento 2, 40136 Bologna, Italy. paolo.guidorzi@unibo.it

Summary

In Europe, *in situ* measurements of sound reflection and airborne sound insulation of noise barriers are usually done according to CEN/TS 1793-5. This method has been improved substantially during the EU funded QUIESST collaborative project. Within the same framework, an inter-laboratory test has been carried out to assess the repeatability and reproducibility of the newly developed method when applied to real-life samples, including the effect of outdoor weather variability and sample ageing. This article presents the statistical analysis of the inter-laboratory test results, and the values of the repeatability and the reproducibility, both in one-third octave bands and for the single-number ratings. The estimated reproducibility values can be used as the extended measure of uncertainty at the 95% credibility level in compliance with the ISO GUM. The repeatability and reproducibility values associated with airborne sound insulation are also compared with the corresponding values for laboratory measurements in building acoustics and an acceptable agreement is found.

PACS no. 43.58.Gn, 43.58.Vb, 43.50.Gf, 43.55.Rg

1. Introduction

Noise barriers are one of the most frequently used means to reduce the impact of noise emanating from roads, railways and factories. Amongst their most important characteristics are their ability to not reflect part of the incident sound energy, and to reduce sound transmission from the noise source on one side to the neighborhood on the other side. These characteristics can be assessed with laboratory tests under a diffuse sound field [1, 2]. The European project ADRIENNE made it possible to measure these characteristics *in situ* under a direct sound field [3]. In Europe, *in situ* measurements of sound reflection and airborne sound insulation of noise barriers are currently done according to CEN/TS 1793-5 [3, 4]; a separate new normative standard has recently become available for airborne sound insulation [5]. This method has been validated by several authors, but the repeatability and reproducibility of the method has been only occasionally estimated and reported in the literature [6, 7, 8, 9]. The *in situ* method has been substantially improved during the EU funded QUIESST project [10, 11, 12, 13]. Moreover, in the same framework, an inter-laboratory test (ILT) has been carried out in order to assess the repeatability and reproducibility of the newly developed method when applied to real-life samples [13, 14, 15].

In this article, the most relevant aspects of the QUIESST inter-laboratory test are discussed. The values for repeatability and reproducibility are presented, both in one-third octave bands and for the single-number ratings. The values are also compared, whenever possible, with the corresponding values for laboratory measurement methods. Finally, some recommendations for future standards are presented.

Received 27 September 2013,
accepted 15 July 2014.

2. The QUIESST measurement method for sound reflection and airborne sound insulation

CEN/TS 1793-5 [4] describes a methodology for the *in situ* measurement of sound reflection and airborne sound insulation of noise barriers applicable to flat and non-flat products; a separate new normative standard has recently become available for airborne sound insulation [5]. Such a measurement is required for the characterization of the noise barrier and the verification of its intrinsic acoustic performance, once constructed aside of a road or railroad. In the framework of the EU funded QUIESST project, the measurement method was thoroughly revised and upgraded, increasing its robustness and reliability. Here only a brief outline of the QUIESST method is given, as the main focus of this article is on the repeatability and reproducibility assessment. Interested readers are referred to the QUIESST documentation [10, 11, 12] or reference [13].

2.1. Sound reflection

For measuring sound reflection a loudspeaker and a 3×3 square microphone array (0.80×0.80 m) are positioned in front of the noise barrier at a distance of 1.5 m and 0.25 m respectively from the vertical plane touching the most protruding part of the barrier on the traffic side, as shown in Figure 1; a multichannel impulse response measurement is taken. Then the sound source and the microphone grid, keeping the same relative distance, are moved away from any reflective object, allowing to acquire a multichannel "free-field" measurement. From the measured data a quantity called reflection index, ***RI***, is computed, defined as

$$\boldsymbol{RI}_j = \frac{1}{n_j}\sum_{k=1}^{n_j}\frac{\int_{\Delta f_j}\left|F[h_{r,k}(t)\,w_{r,k}(t)]\right|^2 \mathrm{d}f}{\int_{\Delta f_j}\left|F[h_{i,k}(t)\,w_{i,k}(t)]\right|^2 \mathrm{d}f} \cdot C_{\mathrm{geo},k}\,C_{\mathrm{dir},k}(\Delta f_j)\,C_{\mathrm{gain},k}(\Delta f_g), \qquad (1)$$

where $h_{i,k}(t)$ is the incident reference component of the free-field impulse response at the k-th measurement point (microphone); $h_{r,k}(t)$ is the reflected component of the impulse response taken in front of the sample under test at the k-th measurement point (microphone); $w_{i,k}(t)$ is the time window (Adrienne shape [4]) for the incident reference component of the free-field impulse response at the k-th measurement point (microphone); $w_{r,k}(t)$ is the time window (Adrienne shape) for the reflected component at the k-th measurement point (microphone); F denotes a Fourier Transform operation; j is the index of the one-third octave frequency bands (between 100 Hz and 5 kHz); Δf_j is the width of the j-th one-third octave frequency band; k is the microphone number ($k = 1$ to 9); n_j is the number of microphone positions on which to average; $C_{\mathrm{geo},k}$ is a correction factor used to compensate the geometrical divergence at the k-th measurement point and takes into account the path difference from the direct and reflected waves; $C_{\mathrm{dir},k}(\Delta f_j)$ is a correction factor used to compensate the difference of sound source directivity, at the k-th measurement point, due to different incidence angles of direct and reflected waves on the microphones; $C_{\mathrm{gain},k}(\Delta f_g)$ is a correction factor used to compensate a gain mismatch (if any) of the amplification settings between the "free-field" and "in front of barrier" measurement configurations [11, 13, 16].

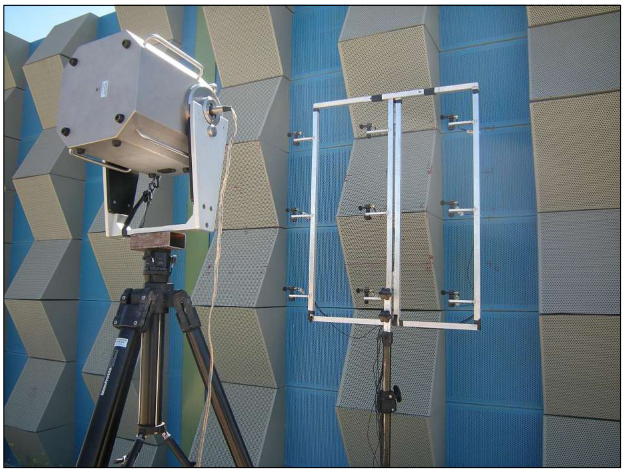

Figure 1. Example of a sound reflection index measurement *in situ*.

A careful choice of the Adrienne analysis window length (7,9 ms or 6,0 ms), on each impulse response gathered from the different microphones, allows to improve the low frequency resolution of the measurement and obtain valid data down to 200 Hz, for a 4 m high barrier [11, 13]. The measurement made in front of the barrier is "cleared" from the direct wave by subtracting in the time domain the anechoic part of the free-field impulse response. The latter is aligned to the barrier measurement by means of a time-shifting operation, performed in the frequency domain by taking the Fourier Transform of the free-field signal, changing its phase by a small amount and inverse transforming the data in order to obtain a fractional translation on the time axis of the original impulse response [11, 13]. Following the procedure specified in [11] it is sure that an overlap between the sound reflected by the sample under test and the sound reflected on the ground never occurred in the one-third octave bands from 200 Hz to 5 kHz. Measured values in the three one-third octave bands 100 Hz, 125 Hz and 160 Hz are kept for information. The analysis window has no influence on the repeatability and reproducibility of the measurement method for sound reflection, because it must always be placed by each laboratory according to strict rules [10, 11, 13, 5] in exactly the same position, with the same shape and the same width.

CEN/TS 1793-5 also specifies the computation of a single number rating, called DL_{RI} and expressed in dB, from the reflection index values in one-third octave bands and the noise spectrum given in EN 1793-3 [17], in order to categorize and compare different noise barriers.

2.2. Airborne sound insulation

The procedure described in CEN/TS 1793-5 for measuring airborne sound insulation is robust and easily appli-

Figure 2. Example of airborne sound insulation measurement *in situ*.

cable, and the new normative revision [5] did not involve notable changes. For insulation measurements, the same microphone grid and loudspeaker used for the reflection index measurement are employed, but the positioning and computation are different. As before two measurements are performed: in the first the microphone grid is placed at 0.25 m from the noise barrier on the receiver side and the sound source on the opposite side of the barrier (traffic side), at a distance of 1 m from the reference plane (Figure 2); the second measurement is taken after placing the microphone grid and the loudspeaker in free-field conditions (away from any obstacle), keeping the same distance between them.

The so called sound insulation index, *SI*, is then computed as

$$SI_j = -10\log_{10}\left\{\frac{1}{n}\sum_{k=1}^{n}\frac{\int_{\Delta f_j}\left|F[h_{t,k}(t)\,w_{t,k}(t)]\right|^2\mathrm{d}f}{\int_{\Delta f_j}\left|F[h_{i,k}(t)\,w_{i,k}(t)]\right|^2\mathrm{d}f}\right\}, \quad (2)$$

where $h_{i,k}(t)$ is the incident reference component of the free-field impulse response, measured at the k-th microphone of the grid; $h_{t,k}(t)$ is the transmitted component of the impulse response, measured at the k-th microphone of the grid; F is the symbol of Fourier Transform; k is the microphone identifier in the grid ($k = 1$ to 9); j is the index of the one-third octave frequency bands from 100 Hz to 5 kHz; Δf_j is the width of the j-th one-third octave frequency band; $w_{i,k}(t)$ is the time window (Adrienne temporal window [5]) for the incident reference component of the free-field impulse response at the k-th microphone position; $w_{t,k}(t)$ is the time window (Adrienne temporal window) for the transmitted component of the free-field impulse response at the k-th microphone position; $n = 9$ is the total number of microphones in the grid.

The analysis window has no influence on the repeatability and reproducibility of the method for airborne sound insulation, because it must always be placed by each laboratory according to strict rules [10, 12, 13, 5] in exactly the same position, with the same shape and the same width.

The European standard [5] also specifies the computation of a single number rating, called DL_{SI} and expressed in dB, from the sound insulation index values in one-third octave bands and the noise spectrum given in EN 1793-3 [17], in order to categorize and compare different noise barriers.

3. The inter-laboratory test

One of the goals of the QUIESST project was to evaluate the uncertainty of the new measurement method for sound reflection *in situ*. Some partial indications did exist, but always in terms of repeatability and reproducibility on a limited number of measurements and mostly for the former ADRIENNE method and not for the new QUIESST one [6, 7, 8, 9]. Actually, the new measurement method is too new and too complex to attempt a mathematical model or even to specify the different components of an uncertainty budget according to the GUM [18]. Therefore, it was decided to assess the uncertainty through an inter-laboratory test [14, 15], as well as:

- to involve eight European laboratories, the minimum number usually recommended (see for example [19, 20]), each one responsible for providing a complete measuring equipment and two skilled operators to apply the new method on the selected test sites;
- to set up two test sites in different European countries where to build the test samples;
- to keep the Grenoble test site (France), which was already used in the former ADRIENNE project [3]; this permits to test samples which for sure were relevant and to save some budget;
- to prepare some new samples, flat and non flat, sound absorbing and sound reflecting; some of them representative of the European market, some others designed to check critical aspect of the method; the new samples had to be built on a purposely prepared test site in Valladolid (Spain);
- to have a supervising panel, composed by an expert of this kind of *in situ* measurement from UNIBO and an expert of statistical analysis from TNO.

At a later stage, it was deemed necessary to widen the aim of the ILT including also the evaluation of the repeatability and reproducibility of the sound insulation index measurement method, now standardized in EN 1793-6 [5]. This additional task required a careful balance between the need for accuracy and the need to stay within the allocated budget (to prepare the test sites, construct the new samples,

Table I. Participating laboratories.

Name	Country
Austrian Institute of Technology (AIT)	Austria
Bundesanstalt für Straßenwesen (BASt)	Germany
Centre Scientifique et Technique Bâtiment (CSTB)	France
Fundación para la Investigación y Desarrollo en Transporte y Energía (CIDAUT)	Spain
Katholieke Universiteit Leuven	Belgium
Laboratoire Régional des Ponts et Chaussées de Strasbourg (LRPC)	France
Rheinisch-Westfälische Technische Hochschule Aachen – Institute of Technical Acoustics (RWTH–ITA)	Germany
University of Bologna	Italy

cover travel expenses to the test sites of two operators for each laboratory with their full equipment, etc.).

Overall, it was possible for 8 European laboratories to measure in a blind inter-laboratory test 13 samples placed on 2 test sites: Grenoble (France) and Valladolid (Spain). On average, each laboratory team spent one week on each test site.

The participating laboratories are shown in Table I in alphabetical order. In the rest of the paper, a randomly assigned code letter will anonymously refer to the laboratories.

Summarizing, it was possible to have the minimum number of laboratories and a relevant number of samples. On the other hand, the complexity of the method and budget limitations did not allow for more than one measurement cycle per site and per laboratory. This fact, together with the influence of outdoor weather variability and ageing of most samples resulted in an uncertainty estimate which is not as low as possible. It rather represents a realistic estimation of what can be obtained in real-life situation in the field.

3.1. Selection of the test samples

As mentioned above, the test samples had to cover the widest possible range about:

- geometry: some samples had to be flat, other moderately non flat and other strongly non flat;
- sound absorption: some of them had to be strongly sound absorbing, other moderately absorbing and other sound reflecting;
- materials commonly in use in Europe: at least concrete, timber, perforated metallic cassettes, wood chips and concrete, acrylic sheets, glass wool.

The choice of samples also had to respect the minimum dimension requirements specified in [4] and [5], with a span between adjacent posts of 4 m. The combination of all these requirements in few test samples on two test sites resulted in the combination of samples and sites specified below.

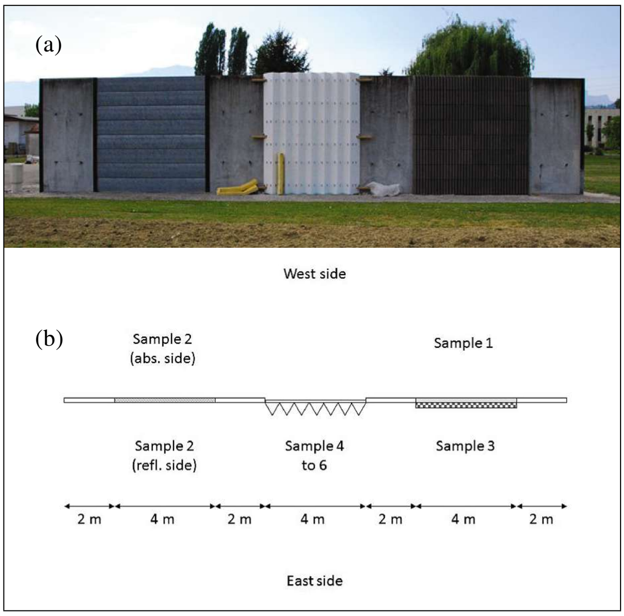

Figure 3. (a) Overview of the Grenoble test site from the East side. Left: Metallic cassettes filled with rock wool. Centre: Reflecting zigzag. Right: Porous wood chips and cement. The samples are separated by 2 m wide smooth concrete walls similar to sample 1. (b) Sketch of the plan view of the complete test wall. For the sample description, see Table II.

Table II. Samples in the Grenoble test site.

N	Name	Dimensions
1	Smooth concrete wall (on the rear of sample 3)	4×4 m
2	Metallic cassettes filled with rock wool	4×4 m
3	Porous wood chips and cement	4×4 m
4	Reflecting zigzag	4×4 m
5	Absorbing zigzag	4×4 m
6	Half-absorbing zigzag	4×4 m

3.1.1. The Grenoble test site

The first test site is located in Grenoble, France, and was already used in the former ADRIENNE project [3]. A general view of the site is shown in Figure 3a. A sketch of the layout of the samples is shown in Figure 3b. On the site there were six samples (i.e. 3 samples on each test facility side), listed in Table II, built one aside the other and separated by 2 m wide smooth concrete slabs.

The smooth concrete wall (sample 1) and the porous wood chips and cement sample (sample 3) were built back to back. In this configuration, sound reflection measurements were surely possible on the exposed sides, while the practicability of sound insulation measurements was questionable; in fact some laboratories did not measure it for samples 1 and 3 (see Table V). Samples 1 to 3, built more than 15 years ago, were well aged and supposed not to change too much from the time of the test by one laboratory to the other.

The absorbing zigzag (sample 5) and the half-absorbing zigzag (sample 6) were set up applying mineral wool

Table III. Samples in the Valladolid test site.

N	Name	Dimensions
7	Strongly non flat metallic wall (sound absorbing)	8×4 m
8	Flat absorbing wall (concrete + rock wool + perforated metallic plate)	8×4 m
9	Flat absorbing timber barrier	8×4 m
10	Non flat absorbing porous concrete A3	8×4 m
11	Non flat absorbing green wall	8×4.5 m
12	Non flat extra absorbing green wall	8×4.5 m
13	Non flat absorbing porous concrete A2	8×4 m

panels to the structure of the reflecting zigzag (sample 4), which was made of smooth painted plywood boards, 0.40 m wide, joined at right angles. In order to set up the half-absorbing zigzag (sample 6), only half of the plywood surface had to be covered with mineral wool, in alternating vertical strips. It is worth noting that during the tests the mineral wool panels have been put up and down manually each time: this may have left different layers of air between the mineral wool panels and the plywood, influencing the overall sound absorption. Thus, the absorbing and half-absorbing samples were actually slightly different for each laboratory. This may explain some unexpected differences found among laboratories for the sound reflection index of these samples. Samples 4 to 6 were intended for sound reflection measurements only, in order to check the measurement method over strongly non flat samples; in fact, they have a challenging shape. On the other hand, the airborne sound insulation of plywood panels is low and sound leaks were not sealed; therefore, airborne sound insulation measurements were not taken for these samples.

EN 1793-6 prescribes that airborne sound insulation should be measured across the acoustic elements as well as across posts for each sample. On the other hand, the compact arrangement of the test samples used in Grenoble implies that posts were generally shared between different samples, so that posts cannot generally be said to be part of a specific sample. Therefore, measurements across posts were not taken into account for the Grenoble test site.

3.1.2. The Valladolid test site

The second test site is located in Valladolid, Spain. A general view of the site is shown in Figure 4. On the site there are seven samples, listed in Table III, built one aside the other and free standing. It is worth noting that all samples were newly built, so that ageing outdoors in the initial part of their life could have had an undesirable effect, changing to same extent the characteristics of the samples from the time of the test by one laboratory to the other (measurements were done from spring 2011 to fall 2011). On the other hand, the time schedule allowed by the European project did not allow waiting several months for ageing all samples.

The strongly non-flat metallic wall (sample 7) was sound absorbing because the perforated non flat parts cover a thick mineral wool layer. The samples 10 and 13 were both made of porous concrete panels, moderately non flat, having different mixtures. The samples 11 and 12 were intended to be green walls, i.e. concrete structures to be filled with hearth and vegetation, but actually during the tests they were partly filled only with dry earth. As these two samples were built back to back and are very massive, sound reflection measurements were surely possible on the exposed sides, while the practicability of sound insulation measurements was questionable; in fact no laboratory did measure it for samples 11 and 12 (see Table V).

Figure 4. Overview of the Valladolid test site. From left to right: non flat absorbing green wall, non flat absorbing porous concrete A3, non flat absorbing porous concrete A2, flat absorbing timber barrier, Flat absorbing wall, strongly non flat metallic wall (the latter is the same sample in Figure 1).

It is important to keep in mind that all these new samples were originally intended for testing sound reflection only. Due to the average workmanship used, the samples were not perfectly homogeneous and had several sound leaks. At the beginning, this was not considered critical for sound reflection, but later on, during the statistical analysis of airborne sound insulation data, this fact turned out to be very important as a limiting factor in obtaining an optimal (low) standard deviation of repeatability and reproducibility.

4. Measurement results

The individual results are too abundant to be reported here in detail. As an example, Figures 5 and 6 show the sound reflection index values obtained for sample 5 on the Grenoble test site and sample 7 on the Valladolid test site, respectively. Figures 7 and 8 show the sound insulation index values obtained for sample 13 and sample 7 on the Valladolid test site, respectively.

Table IV reports the single number ratings for sound reflection. It is worth noting that some laboratories did not measure on sample 6, due to some misunderstanding of the instructions given to all participants. Table V reports the single number ratings obtained for the sound insulation across the acoustic elements. It is worth noting that some laboratories did not provide data for some samples,

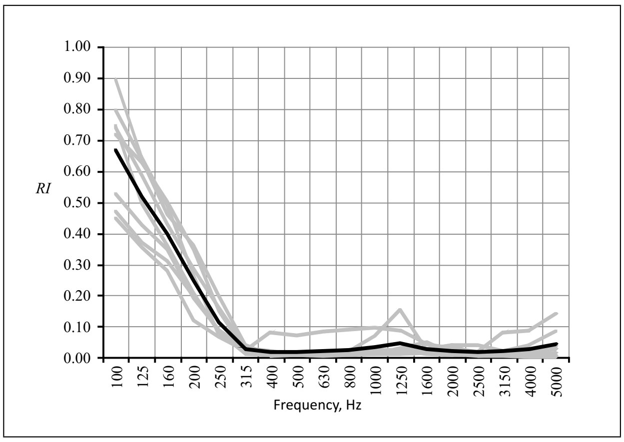

Figure 5. Sound reflection index values for sample 5 (absorbing zigzag) on the Grenoble test site. Gray lines: results for each laboratory. Black line: general mean.

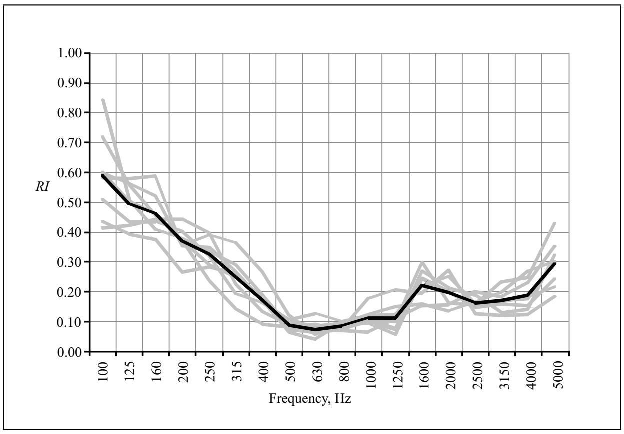

Figure 6. Sound reflection index values for sample 7 (strongly non flat metallic wall) on the Valladolid test site. Gray lines: results for each laboratory. Black line: general mean.

because specific samples were in a “back to back” configuration or because they failed to get valid data. Table VI reports the single number ratings of sound insulation across the posts on the Valladolid test site. It is worth noting that one laboratory forgot to measure the airborne sound insulation of sample 13 across posts.

5. Statistical analysis

5.1. General issues

The statistical analysis of the single number ratings has been carried out using values that have not been rounded, even though the guidelines [11] and [5] require that the single number rating values are rounded to the nearest non-negative integer.

The reason to work with more decimal places than required by the guidelines is that the rounded off part can be an appreciable percentage of the original measurements. As a result, standard deviations and other measures of variability will be inaccurate if based on the rounded values.

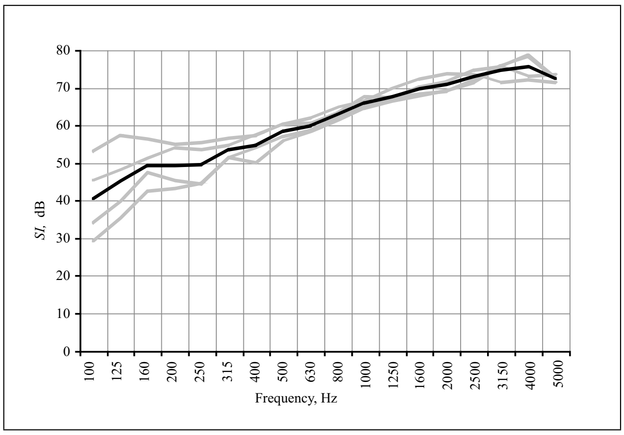

Figure 7. Sound insulation index values across acoustic elements for sample 13 (non flat absorbing porous concrete A2) on the Valladolid test site. Gray lines: results for each laboratory. Black line: general mean.

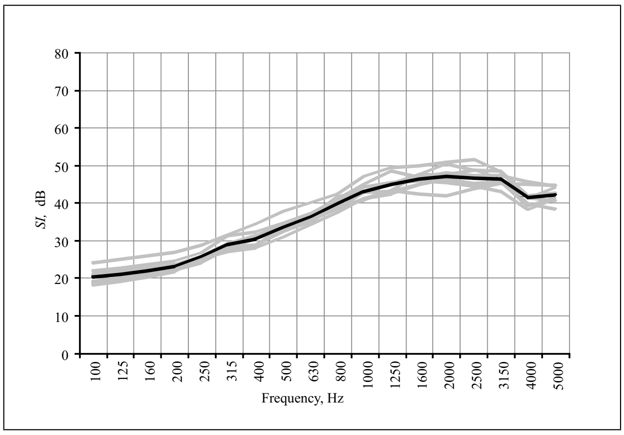

Figure 8. Sound insulation index values across acoustic elements for sample 7 (strongly non-flat metallic wall) on the Valladolid test site. Gray lines: results for each laboratory. Black line: general mean.

According to ISO GUM [18], the accuracy of measurement is the closeness of the agreement between the result of a measurement and a true value of the measurand, where “measurand” stands for a well-defined physical quantity, i.e. the sound reflection index, the airborne sound insulation index in the selected one-third octave bands and their respective single number ratings. In the context of the present inter-laboratory test, there are:

- 247 measurands for sound reflection, namely the 18 $RI(f)$ plus the DL_{RI} for each of the 13 samples;
- 152 measurands for airborne sound insulation of the acoustic elements, namely the 18 $SI(f)$ plus the $DL_{SI,E}$ for each of the 8 measurable samples;
- 95 measurands for airborne sound insulation across posts, namely the 18 $SI(f)$ plus the $DL_{SI,P}$ for each of the 5 measurable samples.

It is worth noting that the single number ratings have been considered as independent measurands as far as the uncertainty calculation is concerned (see also [21]), in order to derive their repeatability and reproducibility from the val-

Table IV. Single number ratings DL_{RI}, in dB, for sound reflection. Samples 1 to 6: Grenoble. Samples 7 to 13: Valladolid.

	Laboratory							
Sample	A	B	C	D	E	F	G	H
1	0	0	0	0	0	0	0	0
2	7	7	7	7	7	7	6	7
3	8	8	8	8	8	8	8	7
4	0	0	0	0	0	0	0	0
5	16	15	12	16	15	12	16	15
6	11	-	-	11	-	-	9	8
7	8	9	8	8	8	8	8	9
8	3	3	4	4	3	4	3	6
9	5	6	5	5	4	6	5	7
10	3	3	3	4	3	3	3	5
11	3	3	4	4	2	3	3	4
12	7	8	7	9	6	6	6	7
13	2	3	3	2	2	2	2	4

Table V. Single number ratings $DL_{SI,E}$, in dB, for airborne sound insulation across acoustic elements. Samples 1 to 3: Grenoble. Samples 7 to 13: Valladolid.

	Laboratory							
Sample	A	B	C	D	E	F	G	H
1	-	52	51	51	-	54	52	54
2	38	39	39	40	40	40	39	38
3	-	51	-	51	-	53	51	54
7	33	33	34	33	36	35	33	38
8	51	51	-	53	53	56	-	55
9	25	24	24	25	-	28	24	27
10	-	63	-	61	61	65	-	64
13	-	55	-	55	62	62	-	-

Table VI. Single number ratings $DL_{SI,P}$, in dB, for airborne sound insulation across posts, obtained in Valladolid.

	Laboratory							
Sample	A	B	C	D	E	F	G	H
7	24	23	24	24	27	26	24	27
8	23	23	23	22	24	24	23	23
9	24	23	24	24	26	27	24	24
10	19	19	19	17	22	23	20	19
13	22	22	22	20	-	21	22	21

ues given by each laboratory on each tested sample as for the one-third octave band values.

The true values of the measurands are unknown, however. Therefore, what can be determined are the repeatability and the reproducibility of the measurement process, including the measurement method itself plus the laboratory, weather, sample ageing and random variances. The repeatability r is the random variation under constant measurement conditions (same laboratory, same operator(s), same equipment, short interval of time). It is the best approach to quantify the variability under homogeneous conditions. It is worth noting that separate repeatability values can be calculated for each of the one-third octave bands as well as for the single number rating. As the single number rating is a kind of weighted average of the respective one-third octave band values, its repeatability will be smaller than any of the repeatability values of the constituent one-third octave bands.

The reproducibility R is the random variation under changed conditions of measurement (different laboratories, different operator(s), different equipment, any interval of time). Again, it is possible to calculate this value for the separate one-third octave band values as well as for the single number rating, where the reproducibility for the single number rating is probably smallest.

In this paper r and R are expressed as $2 \times s_r$ and $2 \times s_R$, respectively, where s_r is the standard deviation of measurements on one and the same object taken under similar conditions briefly after each other, while s_R is the standard deviation of measurements on one and the same object under different conditions. This is called an expanded uncertainty measure. It is worth noting that repeatability and reproducibility become better when r and R get smaller.

With this choice, the interval $[y - R; y + R]$, where y is the value of a single measurement, gives a 95% lower and upper bound for the true value of a single measurement taken by a randomly chosen laboratory. It is shown later in the paper that there is an appreciable inter-laboratory variation. Therefore, reproducibility and not repeatability should be chosen to declare the 95% credibility interval of a measurement.

Here and in the following the term "credibility interval" is used in place of the more widespread term "confidence interval" to point to the Bayesian approach to data analysis used in this work. Consider the reproducibility R for DL_{RI}. The frequentist approach to statistics considers R as an unknown but fixed quantity, which is estimated, based on data. A 95% confidence interval is an interval such that it includes the true value in 95% of the cases were data are collected. Therefore, a stated 95% confidence interval might not include the true value. The Bayesian approach to statistics (adopted in the GUM [18]) models R as a random variable, which has a probability distribution. Based on the data collected in the study, the so- called posterior distribution of R models our present state of knowledge about this quantity. A 95% credibility interval is a summary of this distribution and tells where R is positioned with a probability of 95%.

5.2. Sound reflection

In the following, we first propose a statistical model to evaluate the data. For ease of presentation, the discussion is focused on the single number rating DL_{RI} only. However, the entire discussion applies equally well to each constituent one-third octave band value $RI(f)$. In principle, using ANOVA the variation of the measurements around the overall mean can be decomposed in a contribution of the laboratory, a contribution of the sample, and a so-called residual. The main interest in an inter-laboratory test is in the size of the residual variation and in the variation among the laboratories. This is because these two sources of variation determine the repeatability and reproducibility of the measurements. Both repeatability and reproducibility include the residual variation. However, this makes sense only if the residual variation does not depend on the laboratory or on the sample. Otherwise, it is not appropriate to present a single number summary of the repeatability and reproducibility.

5.2.1. Modeling approach for the single number rating of sound reflection index

Figure 9 shows that the residual variation may depend on the sample that was measured. In particular the measurements on the absorbing zigzag (sample 5) seem to have a larger random variation than the measurements taken on the rest of the barriers. This dependency invalidates the usual statistical analysis method using ANOVA. Therefore, an alternative statistical modelling approach has been adopted, which can handle random variations that depend on experimental factors.

The model used in the statistical analysis is

$$DL_{RI,ijk} = L_i + e_{1,ij} + S_k + e_{0,ijk} \quad \text{dB}, \tag{3}$$

where DL_{RI},ijk is the measurement of laboratory i ($i = 1, \ldots, 8$) on the k-th sample ($k = 1, \ldots, 13$) at location j ($j = 1, 2$); S_k is the true value of $DL_{RI,ijk}$ for the k-th sample ($k = 1, \ldots, 13$); L_i is the effect of laboratory ($i = 1, \ldots, 8$); $e_{1,ij}$ is the random variation between the 16 measurement sessions (8 labs×2 sites); $e_{0,ijk}$ is the residual (random) variation.

$e_{1,ij}$ can be calculated by studying the eight differences in location mean values, one for each laboratory. It is assumed that the random contribution of the measurement sessions $e_{1,ij}$ is normally distributed with mean 0 and variance σ_1^2. It is assumed that the $e_{0,ijk}$ are identically and independently distributed normal variables with means 0 and variance σ_{0k}^2. So, the random error within the sessions depends on the sample.

The model was fitted to the data using Bayesian methodology [22]. Key elements of this methodology are (1) the model, (2) the prior distribution of the model parameters, (3) the data and (4) the posterior distribution of the model parameters given the data. The prior and posterior distributions are briefly discussed in the following.

A prior distribution of the model parameters models our present state of knowledge of these parameters in terms of probability distributions. In this paper, we use the so-called "uninformative priors" to express our initial ignorance at the beginning of the study. For example, the prior distribution of the 13 S_k values for DL_{RI} were modeled with a normal distribution with mean 10 and standard deviation 10. Given the usual values of this single number rating, this is a very broad distribution with 95% of the realizations between 10 ± 1.96 · 10. This can adequately express the state of affairs at the start of the study.

The posterior distribution is in fact an updated prior distribution based on the data. We summarize this distribution by giving the median and 95% or 90% credibility intervals. Note that we also give such intervals for standard deviations and *r* and *R* values.

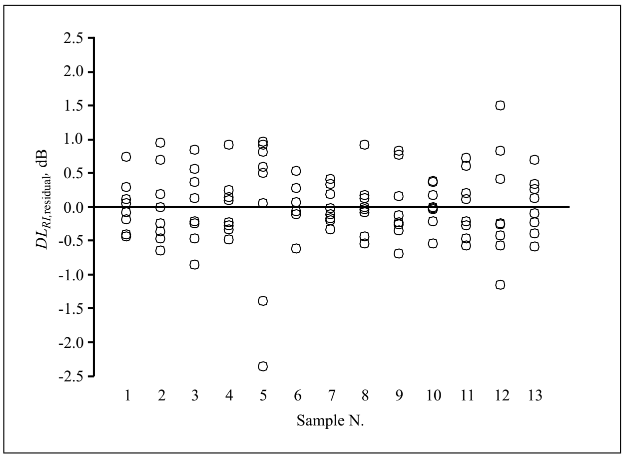

Figure 9. Residual variation for the 13 samples; single number rating of sound reflection index.

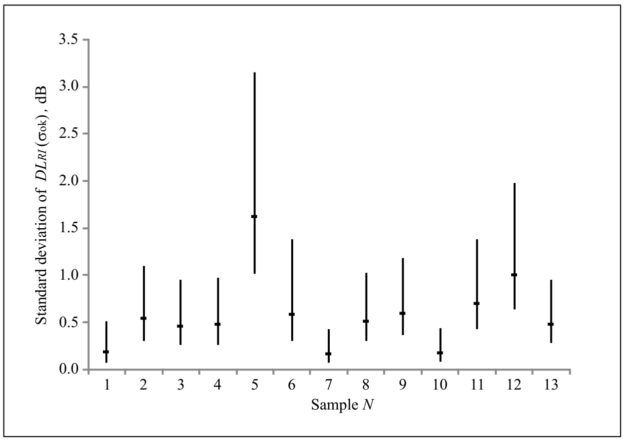

Figure 10. Random variation for the 13 samples; single number rating of sound reflection index.

Based on this model, the 95% credibility intervals for the random variation in each of the samples have been calculated; they are shown in Figure 10.

Also 95% and 99% credibility intervals for the 78 differences between the random variations in pairs of samples have been calculated. The results (not shown) point to a difference between the absorbing zigzag (sample 5) and the rest, although there are occasionally other sample pairs that differ in their random errors.

In the next modelling step, the data have been modelled with a separate random error for the absorbing zigzag (sample 5). Figure 11 depicts the laboratory effects. As these effects are deviations from the mean value of the sample, they are centred on zero. The plot shows medians and 90% credibility intervals of the laboratory effects. Here 90% intervals are used rather than 95% intervals because they are smaller, and may therefore be more discriminating. A study of pair wise differences did not reveal substantial between-laboratory differences. For this reason, an appropriate next modelling step is to assume that laboratories contribute to the random variation between measure-

Table VII. 95% credibility intervals for standard deviations, reproducibility and repeatability of DL_{RI}, in dB.

Parameter	min	max
Std. deviation session/laboratory	0.22	0.62
Std. deviation all samples excl. sample 5	0.44	0.62
Std. deviation sample 5	0.97	3.01
Repeatability (r)	0.88	1.23
Reproducibility (R)	1.08	1.62

ments together with the random effects between the 16 measurement sessions.

A plot for the sample means values and uncertainties follows (Figure 12). It is worth noting that the uncertainties for the samples differ. In particular, the uncertainty in the absorbing zigzag (sample 5) is larger than for the others.

Turning now to the random variation in the DL_{RI} values, Figure 13 shows the standard deviations of the three random variables that model this variation. The standard deviation for the absorbing zigzag (sample 5) is substantially larger than the remaining two components of random error.

For the calculation of repeatability (r) and reproducibility (R), it is therefore better to exclude the random variation of the absorbing zigzag (sample 5). The intervals for the various standard deviations, as well as for repeatability and reproducibility are shown in Table VII.

It is worth noting that r and R give the expanded uncertainty of DL_{RI} from measurements before rounding off. Denoting a single measurement before rounding with y, the two values round($y - 1.62$) and round($y + 1.62$) can be taken as a rounded 95% credibility interval for the true value of a measurement. The value of 1.62 is the upper bound for the reproducibility R in Table VII. It is therefore likely that the interval is too wide. For this reason, the constructed interval is called conservative.

For example, a measured value of $DL_{RI} = 14.12$ dB before rounding off the single number rating of the reflection index DL_{RI} has 95% credibility limits of $(14.12-1.62)$ dB and $(14.12 + 1.62)$ dB. After rounding, the single number rating is 14 dB with lower and upper margins of 13 dB and 16 dB, respectively.

5.2.2. r and R of sound reflection index in one-third octave bands

For the individual measurements in one-third octave bands, the same form of model as for the DL_{RI} can be assumed. That is, the random error for the absorbing zigzag (sample 5) is different from the random error of other samples, and there are random contributions of laboratory and measurement session. In the 315 Hz one-third octave band this is not true (see Figure 5), but for the sake of simplicity the same model has been kept for all bands. The reproducibility and repeatability based on the random error of the barriers excluding the absorbing zigzag sample have been calculated. The results are visualized in Figures 14 and 15. The exact values of repeatability and reproducibility are given in Table VIII.

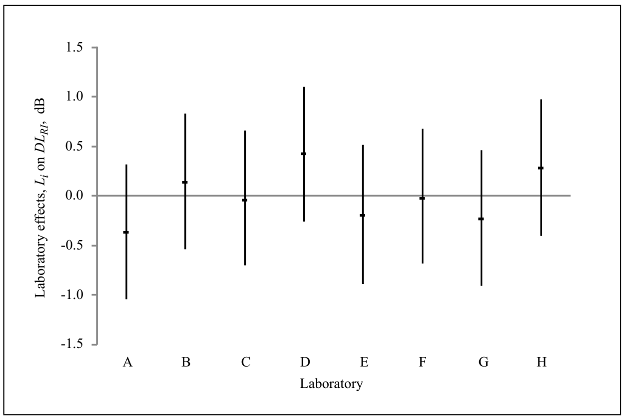

Figure 11. Laboratory effects; single number rating of sound reflection index.

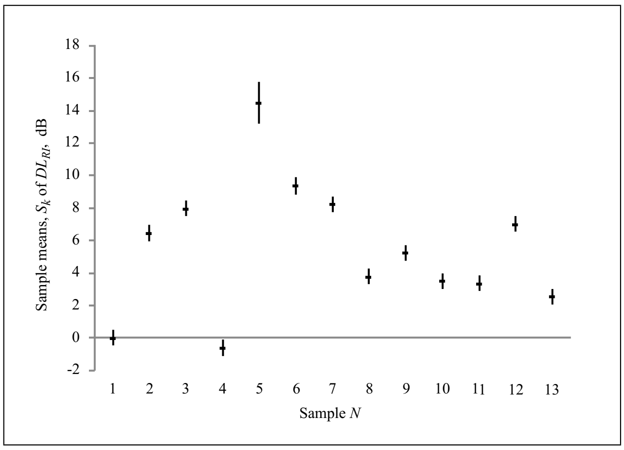

Figure 12. Sample means; single number rating of sound reflection index.

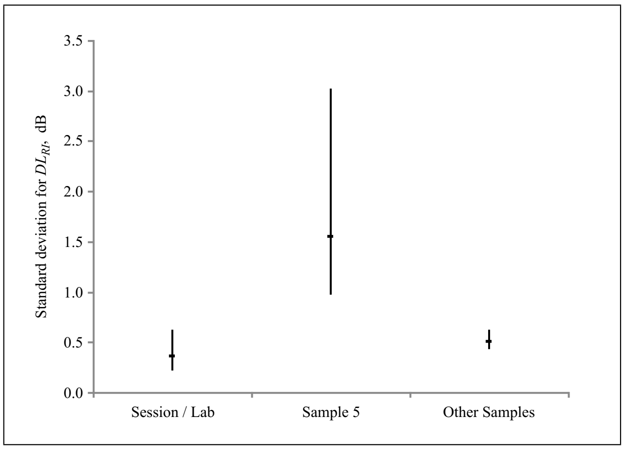

Figure 13. Random variation of the DL_{RI} values.

The plot of r and R values shows medians and 95% credibility intervals. Reproducibility values are never smaller than repeatability values because they include both inter-session and intra-session random variations. The most striking feature of the plot is the sharp decline in r and R values for the initial one-third octave bands.

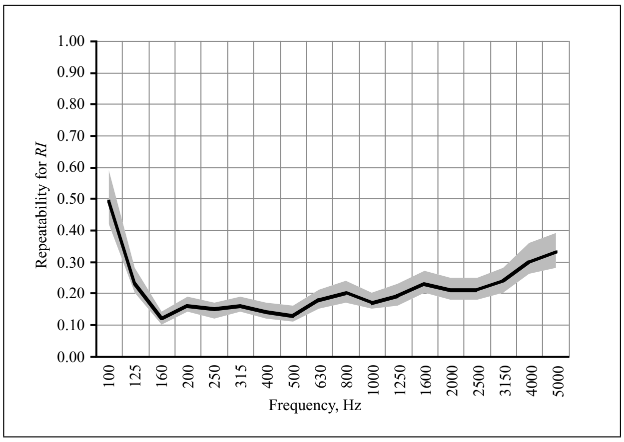

Figure 14. Repeatability ($r = 2s_r$) of the sound reflection index, ***RI***, in one-third octave bands. Black line: median; shaded area: 95% credibility interval.

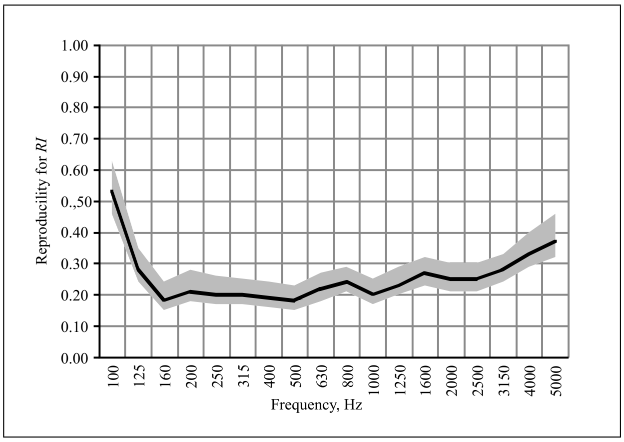

Figure 15. Reproducibility ($R = 2s_R$) of the sound reflection index, ***RI***, in one-third octave bands. Black line: median; shaded area: 95% credibility interval.

Table VIII. Repeatability and reproducibility intervals at 95% credibility level for ***RI***.

Band [Hz]	Repeatability Median	Low	High	Reproducibility Median	Low	High
100	0.49	0.42	0.59	0.53	0.46	0.63
125	0.23	0.20	0.28	0.28	0.24	0.35
160	0.12	0.10	0.14	0.18	0.15	0.24
200	0.16	0.14	0.19	0.21	0.18	0.28
250	0.15	0.12	0.17	0.20	0.17	0.26
315	0.16	0.14	0.19	0.20	0.17	0.25
400	0.14	0.12	0.17	0.19	0.16	0.24
500	0.13	0.11	0.16	0.18	0.15	0.23
630	0.18	0.15	0.21	0.22	0.18	0.27
800	0.20	0.17	0.24	0.24	0.21	0.29
1000	0.17	0.15	0.20	0.20	0.17	0.25
1250	0.19	0.16	0.23	0.23	0.20	0.29
1600	0.23	0.20	0.27	0.27	0.23	0.32
2000	0.21	0.18	0.25	0.25	0.21	0.30
2500	0.21	0.18	0.25	0.25	0.21	0.30
3150	0.24	0.20	0.28	0.28	0.24	0.33
4000	0.30	0.26	0.36	0.33	0.29	0.40
5000	0.33	0.28	0.39	0.37	0.32	0.46

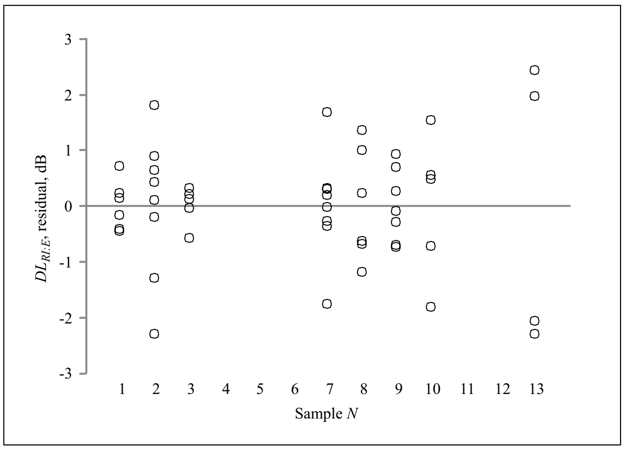

Figure 16. Residual variation for the 8 measurable samples; single number rating of sound insulation index of acoustic elements.

5.3. Airborne sound insulation

5.3.1. Modelling approach for the single number rating of sound insulation index across acoustic elements

There are eight measurable samples for airborne sound insulation. Figure 16 shows that the residual variation of $DL_{SI,E}$ may depend on the sample that was measured. In particular the measurements on sample 13 (non flat absorbing porous concrete A2) seem to have a larger random variation than the measurements taken on the rest of the samples. Therefore, the random error within the sessions depends on the sample. For this reason, also for airborne sound insulation a statistical model similar to that used for sound reflection – see equation (3) – has been adopted. Based on this model, the 95% credibility intervals for the random variation in each of the barriers have been calculated, and shown in Figure 17.

Also 95% and 99% credibility intervals for the 28 differences between the random variations in sample pairs have been calculated. The results (not shown) point to a difference between sample 13 and the rest, although there are occasionally other sample pairs that differ in their random errors. Thus, in the next modelling step, the data have been modelled with a separate random error for sample 13. Figure 18 is a plot of the laboratory effects. The plot shows medians and 95% credibility intervals of the laboratory effects.

A study of pair wise differences revealed between-laboratory differences between laboratories F and H on the one hand and A, B, C, D and G on the other hand, while E has an intermediary position. For this reason, the next modelling step has the following features:

1. as before, each sample has its own true value;
2. there are two groups of laboratories, viz. F and H versus the rest;
3. there are random differences between the laboratories within a group;

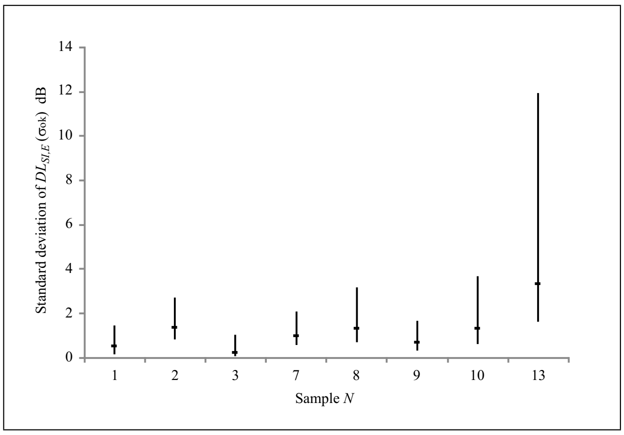

Figure 17. Random variation among the 8 measurable samples; single number rating of sound insulation index of acoustic elements.

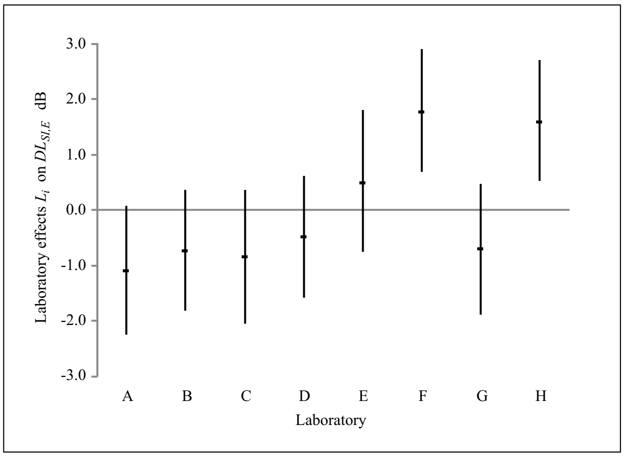

Figure 18. Laboratory effects; single number rating of sound insulation index of acoustic elements.

4. the measurement variability for sample 13 differs from the variability of other samples;
5. there is a random variation between the measurement sessions.

A plot for the sample means is shown in Figure 19. Note that the uncertainties for the samples differ. In particular, the uncertainty for sample 13 is larger than for the others. The measurements of laboratories F and H are 1.69–3.04 dB higher than those of the remaining laboratories; the stated interval is a 95% credibility interval. We do not choose to resolve the difference between the laboratories F and H versus the rest in the present paper. We leave this difference out of our estimation of repeatability and reproducibility and assume that two different values of the $DL_{SI,E}$ can be reported which have the same uncertainty.

Turning now to the random variation in the $DL_{SI,E}$ values, Figure 20 shows the standard deviations of the three random variables that model this variation. The standard deviation for sample 13 is substantially larger than the remaining two components of random error.

For the calculation of repeatability r and reproducibility R, it is therefore better to exclude the random variation of sample 13. The intervals for the various standard deviations, as well as for repeatability and reproducibility are shown in Table IX.

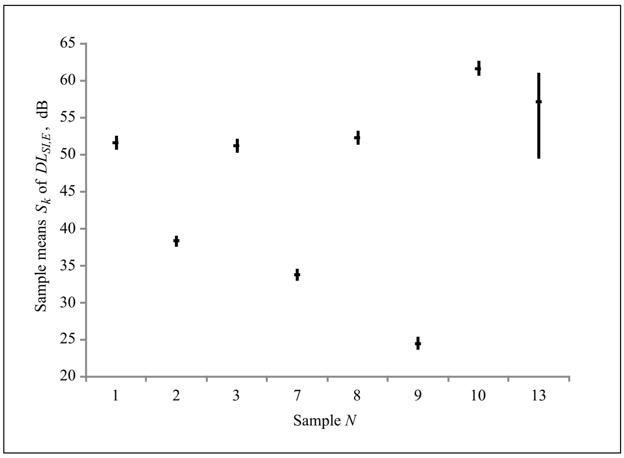

Figure 19. Sample means; single number rating of sound insulation index of acoustic elements.

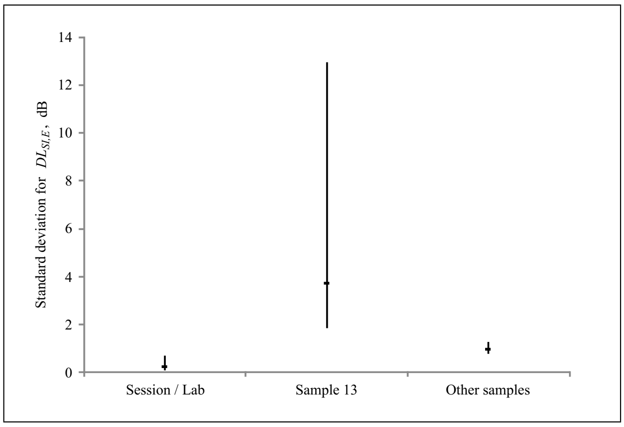

Figure 20. Random variation of the $DL_{SI,E}$ values.

Table IX. 95% credibility intervals for standard deviations, reproducibility and repeatability of $DL_{SI,E}$, in dB.

Parameter	min	max
Std. deviation session/laboratory	0.07	0.67
Std. deviation all samples excl. sample 13	0.77	1.24
Std. deviation sample 13	1.85	12.93
Repeatability r	1.54	2.48
Reproducibility R	1.62	2.61

Denoting a single measurement before rounding with y, the two values round$(y-2.61)$ and round$(y+2.61)$ can be taken as defining a rounded conservative 95% credibility interval for the true value of a measurement.

5.3.2. r and R of the sound insulation index of acoustic elements in one-third octave bands

For the individual measurements in one-third octave bands the same form of model as for the $DL_{SI,E}$ can be assumed, i.e., the random error for sample 13 is different from the random error of other barriers. The results from laboratories F and H differ from the rest, and there are random

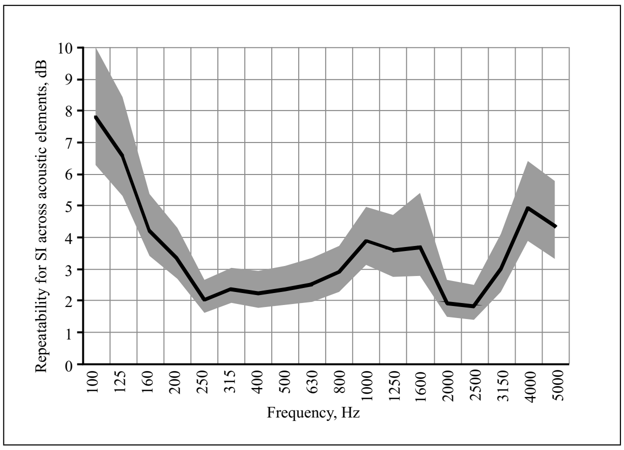

Figure 21. Repeatability ($r = 2s_r$) of the airborne sound insulation index, *SI*, across acoustic elements, in one-third octave bands. Black line: median; shaded area: 95% credibility interval.

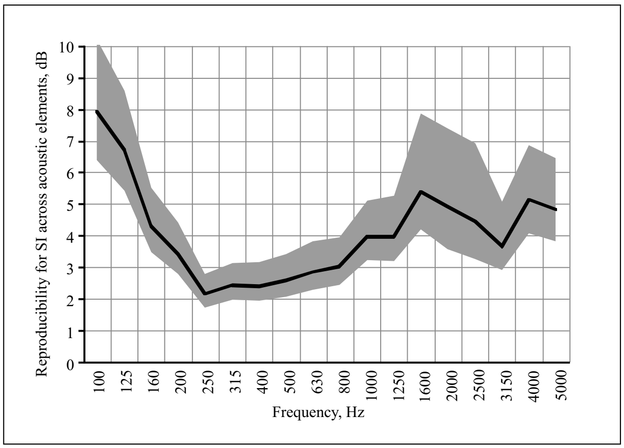

Figure 22. Reproducibility ($R = 2s_R$) of the airborne sound insulation index, *SI*, across acoustic elements, in one-third octave bands. Black line: median; shaded area: 95% credibility interval.

contributions of laboratory and measurement session. The reproducibility and repeatability based on the random error of the samples excluding sample 13 have been calculated. The results are visualized in figures 21 and 22. The exact values of repeatability and reproducibility are given in Table X. The plot of *r* and *R* values shows medians and 95% credibility intervals. The most striking differences in the plot are those in the 2000 Hz and 2500 Hz one-third frequency bands, for which there is a difference of about 3 dB between reproducibility and repeatability.

5.3.3. Modelling approach for the single number rating of airborne sound insulation across posts

Figure 23 shows the residual variation plotted versus the sample single number rating. There are three outlying values in the plot. These correspond to the measurements by laboratory F on samples 13 and 10 and the measurement by laboratory H on sample 7, all in the Valladolid test site. In view of this, we proceed with an analysis where these values are omitted.

Table X. Repeatability and reproducibility intervals at 95% credibility level for *SI* across acoustic elements.

Band	Repeatability			Reproducibility		
[Hz]	Median	Low	High	Median	Low	High
100	7.79	6.29	10.01	7.93	6.40	10.22
125	6.56	5.28	8.43	6.69	5.42	8.59
160	4.20	3.39	5.37	4.28	3.47	5.50
200	3.33	2.69	4.29	3.41	2.77	4.40
250	2.02	1.60	2.63	2.15	1.73	2.79
315	2.37	1.91	3.03	2.45	1.97	3.14
400	2.24	1.76	2.94	2.42	1.94	3.15
500	2.36	1.84	3.09	2.58	2.06	3.41
630	2.50	1.94	3.35	2.83	2.27	3.81
800	2.88	2.27	3.73	3.04	2.44	3.95
1000	3.86	3.11	4.96	3.98	3.21	5.11
1250	3.58	2.75	4.70	3.97	3.19	5.25
1600	3.69	2.77	5.38	5.38	4.19	7.86
2000	1.91	1.46	2.63	4.90	3.58	7.41
2500	1.83	1.39	2.50	4.45	3.26	6.91
3150	2.99	2.27	4.11	3.67	2.90	5.08
4000	4.90	3.88	6.41	5.15	4.08	6.85
5000	4.35	3.30	5.78	4.82	3.83	6.45

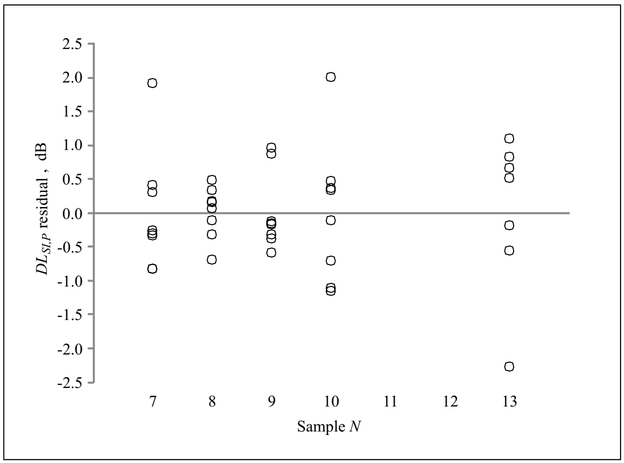

Figure 23. Residual variation for the 5 samples with measurable posts; single number rating for the sound insulation index across posts.

In the next modelling step, the laboratory effects have been studied; see Figure 24. The plot shows medians and 95% credibility intervals of the laboratory effects.

A study of pair wise differences revealed between-laboratory differences between laboratories E and F on one hand and the remaining laboratories on the other hand. For this reason, the proposed next modelling step has the following features:

1. as was the case for the element measurements, each sample has its own true value;
2. there are two groups of laboratories, viz. the pair of E and F, versus the rest;
3. there are random differences between the laboratories within a group.

A plot for the sample mean values follows (Figure 25).

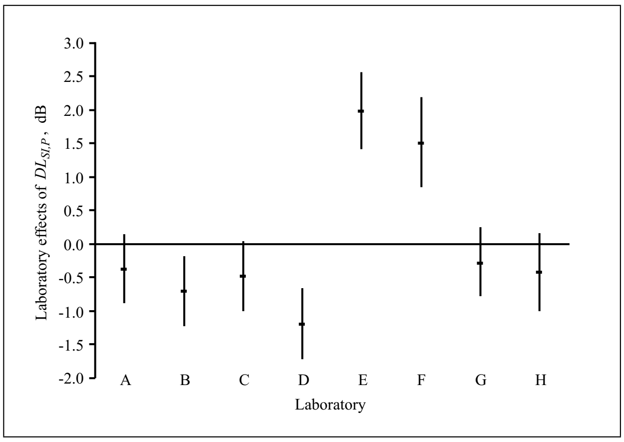

Figure 24. Laboratory effects; single number rating for sound insulation index across posts.

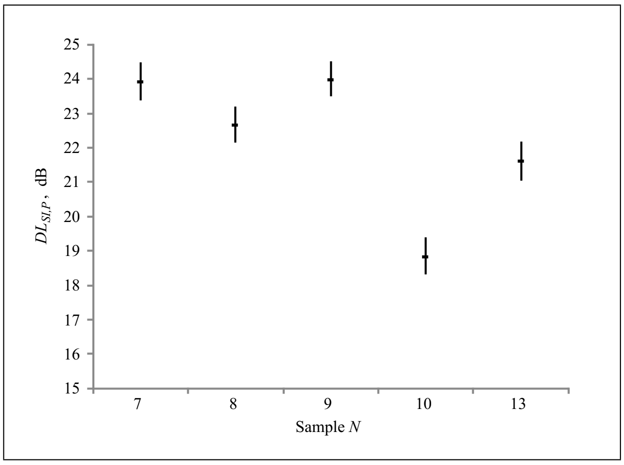

Figure 25. Sample means; single number rating for airborne sound insulation across posts.

Table XI. 95% credibility intervals for standard deviations, reproducibility and repeatability of $DL_{SI,P}$, in dB.

Parameter	min	max
Std. deviation session/laboratory	0.08	0.61
Std. deviation all samples	0.46	0.80
Repeatability (r)	0.92	1.60
Reproducibility (R)	1.03	1.83

The measurements of laboratories E and F are 1.71–3.03 dB higher than those of the remaining laboratories; the stated interval is a 95% credibility interval.

Turning now to the random variation in the $DL_{SI,P}$ values, the intervals for the various standard deviations, as well as for repeatability and reproducibility are shown in Table XI.

Denoting a single measurement before rounding with y, the two values round$(y - 1.83)$ and round$(y + 1.83)$ can be taken as defining a rounded conservative 95% credibility interval for the true value of a measurement.

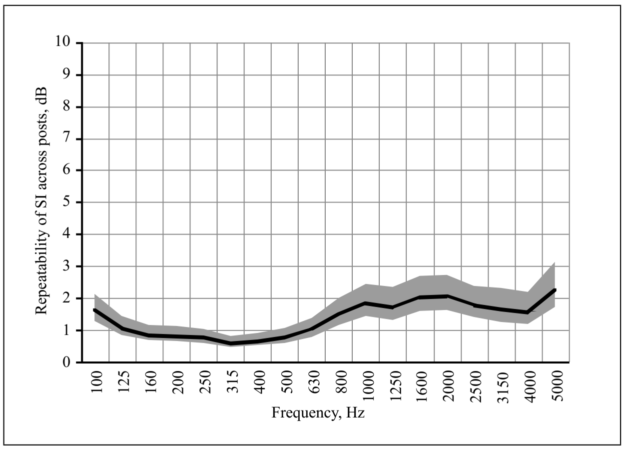

Figure 26. Repeatability ($r = 2s_r$) of the airborne sound insulation index, SI, across posts, in one- third octave bands.

5.3.4. r and *R* of the sound insulation index across posts in one-third octave bands

For the individual post measurements in one-third octave bands, the same form of model as for the $DL_{SI,P}$ can be assumed. That is, the results of laboratories E and F differ from the rest, and there is a random contribution of the laboratory on top of the measurement error. The reproducibility and repeatability based on the random error of the samples have been calculated. The results are visualized in Figures 26 and 27. The exact values of repeatability and reproducibility are given in Table XII. The plot of r and R values shows medians and 95% credibility intervals. The most striking differences in the plot are those in the 3150 Hz–5000 Hz one-third frequency bands, for which there is a difference of 0.5 dB to 1 dB between reproducibility and repeatability. It is worth noting that the dispersion of the data at high frequencies is larger than that at low frequencies. For a possible explanation of this fact, it should be recalled that airborne sound insulation across posts is usually influenced by sound leaks at panel-post junctions. These leaks are very directive at high frequencies, so that minimal differences in microphone positions among the different laboratories result in different values of the sound insulation index.

6. Comparison with the laboratory method – Airborne sound insulation

It would be interesting to compare the repeatability and reproducibility data obtained in the frame of the QUIESST project with similar data for laboratory measurements. Until now, ILT on laboratory measurements on noise barriers have not been performed. A few ILT have been done in the building acoustics sector, particularly on airborne sound insulation. The comparison should be done keeping in mind that field measurements generally have worst repeatability and reproducibility than laboratory ones and that QUIESST measurements are done outdoors under a directional sound field while building acoustics measure-

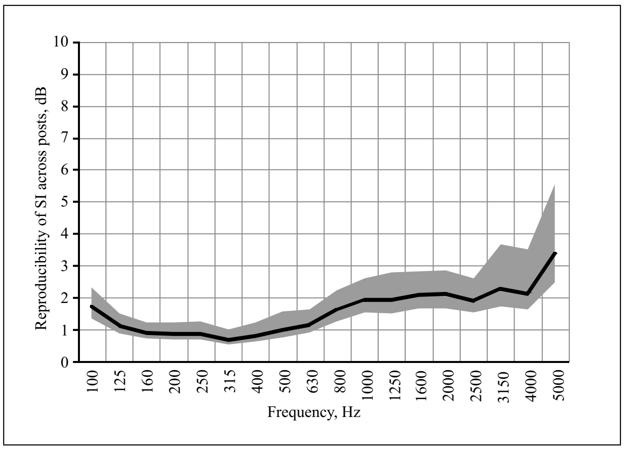

Figure 27. Reproducibility ($R = 2s_R$) of the airborne sound insulation index, *SI*, across posts, in one-third octave bands.

Table XII. Repeatability and reproducibility intervals at 95% credibility level of *SI* across posts.

Band [Hz]	Repeatability Median	Repeatability Low	Repeatability High	Reproducibility Median	Reproducibility Low	Reproducibility High
100	1.63	1.28	2.14	7.93	6.40	10.22
125	1.07	0.84	1.42	6.69	5.42	8.59
160	0.85	0.67	1.14	4.28	3.47	5.50
200	0.82	0.64	1.11	3.41	2.77	4.40
250	0.77	0.60	1.03	1.71	1.34	2.31
315	0.60	0.46	0.81	1.12	0.88	1.51
400	0.66	0.51	0.90	0.90	0.71	1.22
500	0.77	0.58	1.06	0.88	0.69	1.21
630	1.03	0.78	1.38	0.87	0.68	1.25
800	1.49	1.15	1.99	0.69	0.54	1.01
1000	1.83	1.42	2.43	0.79	0.61	1.21
1250	1.72	1.31	2.35	0.98	0.75	1.57
1600	2.03	1.60	2.68	1.15	0.90	1.62
2000	2.06	1.62	2.73	1.61	1.25	2.23
2500	1.79	1.39	2.39	1.92	1.52	2.60
3150	1.64	1.26	2.31	1.92	1.49	2.79
4000	1.55	1.18	2.19	2.09	1.64	2.80
5000	2.24	1.71	3.13	2.13	1.66	2.84

ments are done indoors under an almost diffuse sound field.

There is little literature work on the repeatability and reproducibility of *in situ* measurements of airborne sound insulation [23, 24, 21], and most previous work concerns laboratory tests [25, 26, 27, 28]. The most authoritative references still remain ISO 140-2:1991 and the most recent ISO 12999-1 [20, 29].

Figure 28 compares QUIESST data on repeatability with those reported in ISO 140-2 and ISO 12999-1, while Figure 29 compares QUIESST data on reproducibility with those reported in ISO 140-2 and ISO 12999-1.

It is worth noting that in ISO 140-2 the repeatability and reproducibility are given as

$$r = 2.8\sqrt{s_r^2}, \qquad R = 2.8\sqrt{s_R^2}. \tag{4}$$

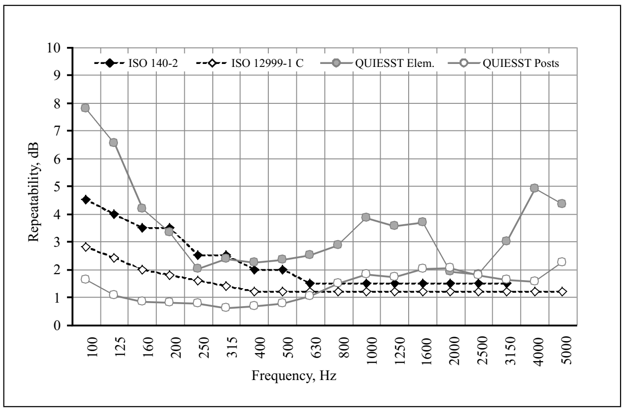

Figure 28. Repeatability *r* of airborne sound insulation measurements.

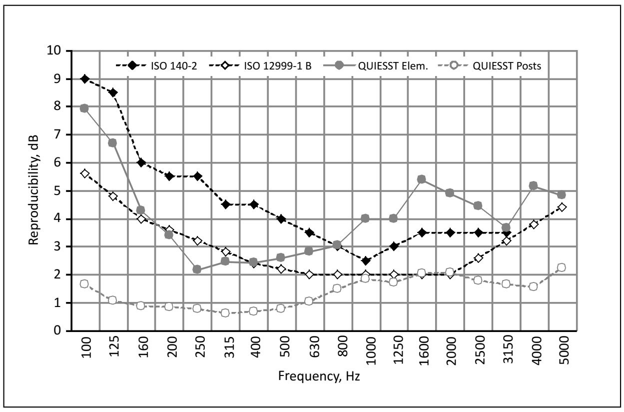

Figure 29. Reproducibility *R* of airborne sound insulation measurements.

In view of this, the values shown in Figures 28 and 29 have been taken from ISO 140-2 dividing by 2.8 and multiplying by 2. Also, in ISO 12999-1 three different measurement situations are distinguished:

- Situation A: A building element is to be characterized by laboratory measurements. All measurement results that may be obtained in another test facility or building also comply with this definition. The standard uncertainty is the standard deviation of reproducibility as determined by inter-laboratory measurements.
- Situation B: different measurement teams come to the same location to carry out measurements. The location may be a usual building or a test facility. The measurand thus is a property of one particular element in one particular test facility or the property of a building. The main difference to situation A is that many aspects of the airborne and structure-borne sound fields involved remain constant since the physical construction is unchanged. The standard uncertainty obtained for this situation is called *in situ* standard deviation.
- Situation C: the measurement is simply repeated in the same location by the same operator using the same equipment. The location may be a usual building or a test facility. The standard uncertainty is the standard deviation of repeatability as determined by interlaboratory measurements.

Table XIII. Repeatability and reproducibility at 95% credibility level of single number ratings of airborne sound insulation, in dB.

	Repeatability r	Reproducibility R
ISO 140-2	0.7	1.4
ISO 12999-1	1.0	2.0
QUIESST Elements	1.54–2.48	1.62–2.61
QUIESST Posts	0.92–1.60	1.03–1.83

Therefore, the ISO 12999-1 values for situation C are used for comparison with the other repeatability values, while values for situation B are used for comparison with the other reproducibility values.

It appears that ISO 140-2 and ISO 12999-1 give similar values for repeatability, especially above 500 Hz. The QUIESST data for repeatability across posts are comparable to the laboratory data for building acoustics: slightly better from 100 Hz to 630 Hz, slightly worse from 800 Hz to 5 kHz. The QUIESST data for repeatability across acoustic elements are worse than laboratory data for building acoustics. The QUIESST data for reproducibility across posts are better than laboratory data for building acoustics. The QUIESST data for reproducibility across acoustic elements are comparable to the laboratory data for building acoustics up to 630 Hz and from 630 Hz onward are generally worse. In sections 3.1.2 and 5.3, it was shown that the influence of sound leaks and ageing on the sound insulation performance of acoustic elements, especially on the Valladolid test site, does contribute to the overall variance. In fact, the figures across posts are definitely better, even if for posts less measurements were possible than for acoustic elements. An indirect support to this conclusion comes also from previous investigations, where sound insulation index measurements have been found quite reliable [6, 7, 8]; therefore a considerable part of the variance in the ILT for acoustic elements may be due to the sample variance.

Table XIII reports the repeatability and reproducibility for the single number ratings from ISO 140-2 and ISO 12999-1. Regarding the latter reference, the values for the single number ratings including the traffic correction term C_{tr} have been kept, because they have to be compared with the single number ratings for noise barriers which are obtained using a traffic noise spectrum. It seems that ISO 140-2 estimates are quite optimistic, while ISO 12999-1 estimates are comparable to QUIESST ones across posts; finally QUIESST values across acoustic elements are the worst.

7. Conclusions

In the framework of the QUIESST project, new measurement procedures for determining *in situ* the sound reflection and the airborne sound insulation of noise barriers have been defined. The repeatability and reproducibility of these new methods have been assessed through an inter-laboratory test involving 8 participating laboratories, 2 test sites and a total of 13 samples. The estimated reproducibility values can be used as an extended uncertainty at 95% credibility level to declare the measured values together with their uncertainty according to ISO GUM [18].

It is the first time that repeatability and reproducibility values for *in situ* sound reflection measurements are evaluated. It is worth noting that the samples were chosen in order to have different shapes, including strongly non flat products, and to be composed with different materials, a fact that made the evaluation more challenging.

The samples were real-size ones, built as in practice using an average workmanship, i.e. without the enhanced quality control measures that are typically performed in inter-laboratory tests done inside laboratory facilities. Nevertheless, the reproducibility values found for the sound insulation index are in general comparable to those coming from inter-laboratory tests of airborne sound insulation of building elements in the laboratory. In particular, the repeatability and reproducibility values of the sound insulation index across posts are better than the values across acoustic elements. As airborne sound insulation is usually better across the acoustic elements than across posts, which often present a weak point, this finding may suggest that different ranges of airborne sound insulation should be associated with different sets of repeatability and reproducibility values. A similar conclusion was advanced also by Scrosati *et al.* [21], inferring the need for more investigations on this point.

Acknowledgement

The inter-laboratory test was done in the frame of WP3 of the QUIESST collaborative project (EU 7th Framework Program: FP7-SST-2008-RTD-1–N. 233730).

References

[1] EN 1793-1:2013: Road traffic noise reducing devices – Test method for determining the acoustic performance – Part 1: Intrinsic characteristics of sound absorption.

[2] EN 1793-2:2013: Road traffic noise reducing devices – Test method for determining the acoustic performance – Part 2: Intrinsic characteristics of airborne sound insulation under diffuse sound field conditions.

[3] http://acustica.ing.unibo.it/Researches/barriers/adrienne.

[4] CEN/TS 1793-5:2003: Road traffic noise reducing devices – Test method for determining the acoustic performance – Part 5: Intrinsic characteristics - In situ values of airborne sound reflection and airborne sound insulation.

[5] EN 1793-6:2012: Road traffic noise reducing devices – Test method for determining the acoustic performance – Part 6: Intrinsic characteristics - In situ values of airborne sound insulation under direct sound field conditions.

[6] M. Garai, P. Guidorzi: European methodology for testing the airborne sound insulation characteristics of noise barriers in situ: experimental verification and comparison with laboratory data. J. Acoust. Soc. Am. **108** (2000) 1054–1067.

[7] G. Watts, P. Morgan: Measurement of airborne sound insulation of timber noise barriers: Comparison of in situ method CEN/TS 1793-5 with laboratory method EN 1793-2. Appl. Acoust. **68** (2007) 421–436.

[8] M. Garai, P. Guidorzi: In situ measurements of the intrinsic characteristics of the acoustic barriers installed along a new

high speed railway line. Noise Control Eng. J. **56** (2008) 342–355.

[9] P. Guidorzi, J. Klepáček, M. Garai: On the repeatability of reflection index measurements on noise barriers. Proceeding of Euronoise 2012, Prague, 2012, 1314–1319.

[10] http://www.quiesst.eu.

[11] QUIESST D3.3:2012: Noise reducing devices acting on airborne sound propagation – Test method for determining the acoustic performance – Intrinsic characteristics - In situ values of sound reflection under direct sound field conditions. http://www.quiesst.eu.

[12] QUIESST D3.4:2012: Noise reducing devices acting on airborne sound propagation – Test method for determining the acoustic performance – Intrinsic characteristics - In situ values of airborne sound insulation under direct sound field conditions. http://www.quiesst.eu.

[13] P. Guidorzi, M. Garai: Advancements in sound reflection and airborne sound insulation measurement on noise barriers. Open Journal of Acoustics **3** (2013) 25–38.

[14] QUIESST D3.5:2012: Inter-laboratory test to assess the uncertainty of the new measurement methods for determining the in situ values of sound reflection and airborne sound insulation of noise reducing devices under direct sound field conditions. http://www.quiesst.eu.

[15] M. Garai, P. Guidorzi, E. Schoen: Assessing the repeatability and reproducibility of in situ measurements of sound reflection and airborne sound insulation index of noise barriers. Proceedings AIA-DAGA 2013, Merano, 2013, Paper ID 106, 1–4.

[16] C. Glorieux, M. Rychtarikova: Self-calibrating method for sound reflection index measurements. Proceedings Acoustics 2012, Nantes, 2012, 2051–2055.

[17] EN 1793-3:1999: Road traffic noise reducing devices – Test method for determining the acoustic performance – Part 3: Normalized traffic noise spectrum.

[18] ISO/IEC Guide 98-3:2008: Uncertainty of measurement – Part 3: Guide to the expression of uncertainty in measurement. GUM: 1995.

[19] ISO 5725-2:1994 (Corrigendum 1:2002): Accuracy (trueness and precision) of measurement methods and results – Part 2: Basic method for the determination of repeatability and reproducibility of a standard measurement method.

[20] ISO 140-2:1991: Acoustics – Measurement of sound insulation in buildings and of building elements – Part 2: determination, verification and application of precision data.

[21] C. Scrosati, F. Scamoni, M. Bassanino, M. Mussin, G. Zambon: Uncertainty analysis by a round robin test of field measurements of sound insulation in buildings: single numbers and low frequency bands evaluation - Airborne sound insulation. Noise Control Eng. J. **61** (2013) 291–306.

[22] A. Gelman, J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, D. B. Rubin: Bayesian data analysis, 3rd edition. Chapman & Hall/CRC, New York, 2013.

[23] J. Lang: A round robin on sound insulation in buildings. Appl. Acoust. **52** (1997) 225–238.

[24] C. Simmons: Uncertainty of measured and calculated sound insulation in buildings - Results of a round robin test. Noise Control Eng. J. **55** (2007) 67–75.

[25] P. Fausti, R. Pompoli, R. S. Smith: An inter-comparison of laboratory measurements of airborne sound insulation of lightweight plasterboard walls. J. Building Acoust. **6** (1999) 127–140.

[26] R. S. Smith, R. Pompoli, P. Fausti: An investigation into the reproducibility values of the European inter-laboratory test for lightweight walls. J. Building Acoust. **6** (1999) 187–210.

[27] A. Schmitz, A. Meier, G. Raabe: Inter-laboratory test of sound insulation measurements on heavy walls: part I - preliminary test. J. Building Acoust. **6** (1999) 159–169.

[28] A. Meier, A. Schmitz, G. Raabe: Inter-laboratory test of sound insulation measurements on heavy walls: part II - results of main test. J. Building Acoust. **6** (1999) 171–186.

[29] ISO 12999-1:2014: Acoustics – Determination and application of measurement uncertainties in building acoustics – Part 1: Sound insulation.