MEASUREMENT UNCERTAINTY IN SOUND INSULATION

Bill Whitfield noise.co.uk Ltd
John Fenlon University of Warwick
Laura Church Vinci

1 INTRODUCTION

Eight acoustic laboratories took part in a Round Robin test in a Vinci Construction student accommodation block in East London in 2022. Both airborne and impact sound insulation testing was carried out on the separating floors following the test method requirements of Approved Document E of the Building Regulations. The paper describes the experiment procedure and the round robin repeatability and reproducibility results for the measurement systems.

2 METHOD

2.1 Field Testing - Sound Insulation

The proposed method was that each lab would test the sound transmission through floors using their own standard equipment, with two engineers each taking two samples. In BS ISO 5725[1,2] testing is generally assumed to be a chemical sample, which is then sent to several laboratories for analysis, and the result determines two derivatives: repeatability and reproducibility, the first being the variance (variability) due to the taking of multiple samples by a single lab, and the second being reproducibility, the overall variance comprising both the within lab variance (repeatability) and the between lab variance, essentially the variance between the results of the different labs. The repeatability gives the ‘average’ variability taken by a specific lab, while the reproducibility is the variability to be expected if the result is sent to a random lab, i.e., any lab in the system. With sound measurement, instead of a control sample being sent to each lab, each lab visits a ‘control’ room and measures the sound profile with its own equipment. This is analogous to chemical tests in that different labs often have different equipment (i.e., purchased from different manufacturers) for conducting a specific chemical analysis. In the case of acoustic research, we might also expect some difference between engineers, so we have built that into the procedure as well.

3 ANALYSIS

3.1 Statistical Analysis

The standard method of analysis for this type of data is called analysis of variance. For a particular set of observations, say $\mathbf{D_{nT,w}}$ + C_{tr} for airborne measurements there will be an overall measure of the variability of the data called the variance – for the record, it is a statistical measurement of the difference between all the individual measurements and the average of them, and is given by

$$\sum_{i=1}^{n} \frac{(x_i - \bar{x})^2}{n-1}$$

where x_i stands for the individual measurements (running from 1 to n), $\bar{x}$ for the mean (average) of them, and the Greek letter Σ indicates summation. Analysis of variance enables the variance to be divided into components that indicate where the variation is concentrated. In these tests, we can draw out the variance due to repetition, i.e., the variance due to each engineer taking two (independent) measurements, essentially the variability between the pairs of measurements combined across all engineers; what remains is the difference between engineers. In the following, the breakdown is into these two components. Had all organisations complied with the 'formal' rules of two readings by two engineers it would be possible to break down the 'between engineer' variance (which is combined across all organisations) into two further components: between engineers within organisations, and between organisations. It is possible to make some estimate of these given the data we have but it is not entirely reliable because of the imbalance in the data.

4 RESULTS

4.1 $D_{nT,w}$ + C_{tr} (airborne)

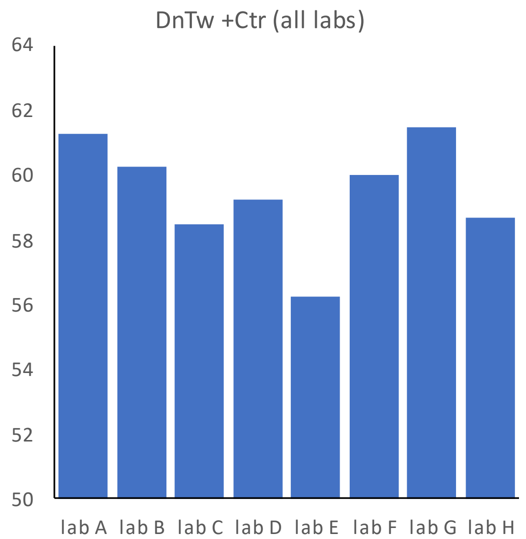
The mean results for $D_{nT,w}$ + C_{tr} are shown as a histogram and a table in Figure 1. The overall mean is 59.2 dB, although the mean of the lab means is 59.5 due to the different numbers of sample in each lab, and the lower mean of lab E. The statistical analysis relies on a method that takes the overall variance and splits it up into three components:

a) the difference between the lab means;

b) the difference between engineers in the same lab (combined across labs); and

c) the difference between duplicate samples by an engineer (combined across all engineers).

Item c is the repeatability, whereas item b includes the repeatability and the difference between engineers, while item a includes all three components.



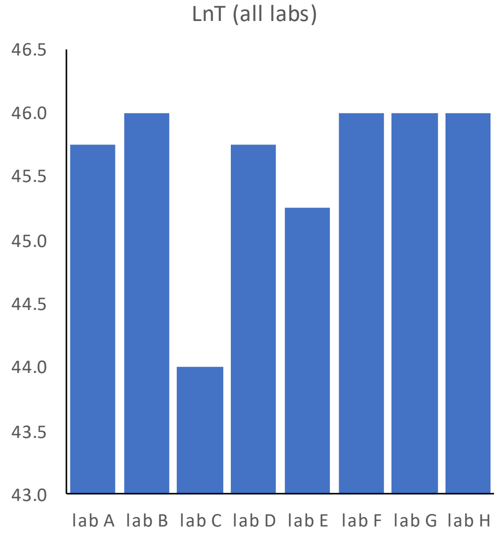
lab	n	mean
A	4	61.25
B	4	60.25
C	2	58.50
D	4	59.25
E	4	56.25
F	2	60.00
G	2	61.50
H	10	58.70

Figure 1: mean $D_{nT,w}$ + C_{tr} for all labs: labs means and replication (sample size) on the right

To calculate the reproducibility the three components need to be extracted so that we can add the single components together – this is not the same as variance (a) because of the different levels of combination for the laboratories. In actual fact, as stated in the previous section, (a) and (b) have been combined, so we have a value for (c) of 2.781 (repeatability), while the combined value of (a) and (b) is 0.903. According to ISO 5725, the reproducibility is the sum of these two: 3.684 (reproducibility): essentially, the variance of taking an engineer and a reading at random. The variance is a square measure of variability; if we take the square root, we get a measure in the same units as the mean (dB) called the standard deviation, in this case 1.919 dB. The implication of this is that if we were to choose an engineer at random from one of the participating companies, then any measurement he/she takes would lie within ±2 x s.d. of the 'true' measurement 95% of the time. We don't know the true measurement so we take the 'consensus' value of 59.2 dB, and we therefore expect a 'random' measurement (as defined above) to lie between 55.4 and 63.0 dB. If we try and break down the variation further, we are thwarted by the fact that different labs used different numbers of engineers, but the data suggest that the difference between engineers within the same lab was effectively zero, and the variation was due to lab differences. In practical terms this probably makes sense, accounting for differences between instruments and measurement practices between labs.

4.2 $L_{nT,w}$ (impact)

The mean results for $L_{nT,w}$ are shown as a histogram and a table in Figure 2. The overall mean is 45.7dB, and the mean of the lab means is 45.6 dB due to the different numbers of sample in each lab, so not moved much by the lower mean of lab C. Using the same methods as for $D_{nT,w} + C_{tr}$ above, we find that the repeatability variance (between pairs of measurements) is 0.281, while the between engineer variance is 0.902, which gives a reproducibility value of 1.183. The square root of this is 1.088 dB, so we expect our 'random' measurement to be between 43.5 and 47.9 dB, i.e., 45.7 ±2 x 1.1. If we were to try and take this further we would find that the variance between engineers within labs accounts for nearly all the variance between engineers, i.e., there is no real difference between labs.



lab	*n*	mean
A	4	45.75
B	4	46.00
C	2	44.00
D	4	45.75
E	4	45.25
F	2	46.00
G	2	46.00
H	10	46.00

Figure 2: mean $L_{nT,w}$ for all labs: labs means and replication (sample size) on the right.

4.3 Third octave bands

The airborne and impact results are given in Figure 3. The Figure contains two pairs of graphs: the top two for airborne and the lower two for impact. The plots on the left are the profiles of the measurements across the 16 ⅓-octave frequencies averaged for each lab – recall that the replication is different for the various labs (see Figs 1 & 2) – those on the right show the within and between engineer variances, the first being the repeatability, the reproducibility is the sum of the between and within variances. In a sense, the right-hand plots summarise the story on the left-hand side:

- airborne: the between engineer variance is very high at low frequency, and then falls to zero for much of the frequency band (from about 250 Hz with a couple of later blips) which is reflected in the raggedness between lab values at the two lower frequencies.

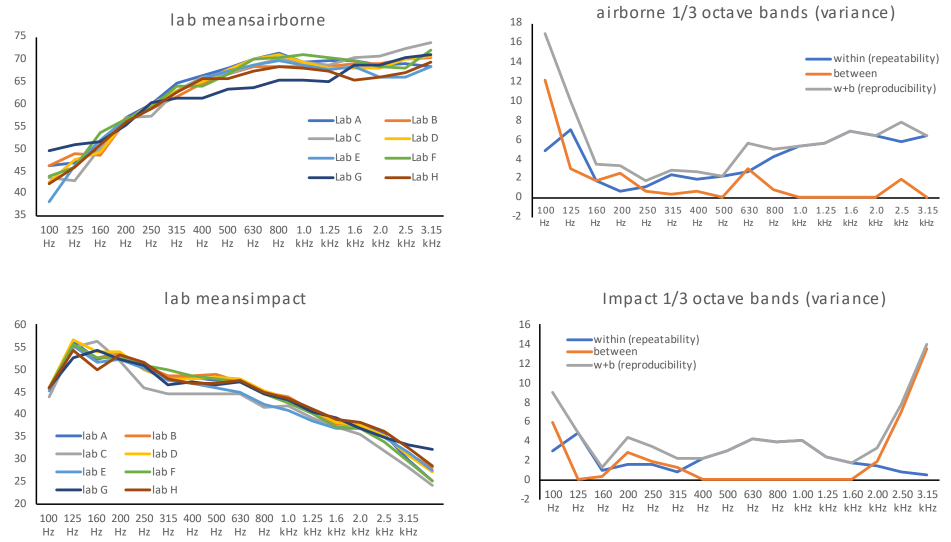


Figure 3: sound profiles across ⅓-octave bands for eight labs (A to H) – airborne and impact; corresponding repeatability and reproducibility across the same spectrum.

The repeatability is also fairly high at low frequencies, but stabilises through the middle frequencies before gradually rising. Note that with the between engineer variance virtually disappearing at high frequency, the repeatability and reproducibility are much the same from about 500 Hz. The sound profiles for most of the labs are very similar – the lab that trails a bit during the middle frequencies does not have a major effect ion the result.

- Impact: the frequency profiles for impact are fairly tight, only getting ragged at high frequency. This is reflected in the between engineer variance in the right-hand plot, which also inflates the reproducibility as the repeatability is quite low across the spectrum. The between engineer variance is quite low (zero in many places) except at the high frequency end.

Overall, this seems like a useful and consistent exercise and is helpful in submission to UKAS for the assessment of measurement uncertainties in sound insulation measurement in the field. If it is repeated it would be good to ensure that all labs did the same thing, i.e., two replicate samples from each of two engineers – though that could lead to competition, as differences between labs would be more apparent!!

5 REFERENCES

1. BS ISO 5725-1:1994 Accuracy (trueness and precision) of measurement methods and results – Part 1: General principles and definitions
2. BS ISO 5725-2:1994 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